cosyne.org

COSYNE 2024

MAIN MEETING

Lisbon, Portugal

29 Feb – 3 March

Program Summary

All times are GMT (Greenwich Mean Time)

Thursday, 29 February 2024

08:00	**NWB Tutorial**
09:45	**IBL Tutorial**
11:00	**Mind Matching (Neuromatch)**
12:30	**Cosyne 2024 Tutorial sponsored by the Simons Foundation**
14:00	**Registration opens**
16:30	**Welcome reception**
19:00	**Opening remarks**
19:30	**Session 1:** Invited speaker: Lars Chittka
20:30	**Poster Session 1**

Friday, 01 March 2024

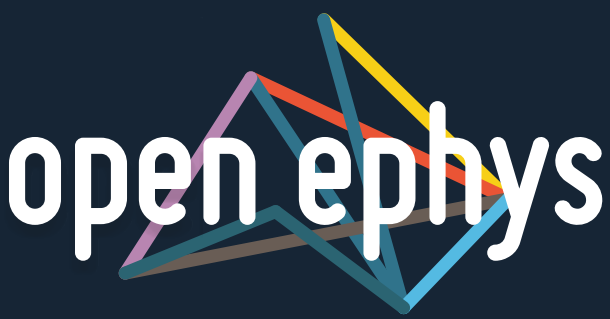
09:00	**Session 2: Reward and reinforcement learning** Invited speaker: Alex Pouget; 3 accepted talks
10:45	**Session 3: Learning** Invited speaker: Blaise Aguera y Arcas; 3 accepted talks
12:30	**Poster Session 2**
13:00	**Spike-Sorting: The Game!**
16:00	**Session 4: Sensorimotor dynamics** Invited speaker: Claire Wyart; 4 accepted talks
17:00	**Session 5: Structure of neural activity** Invited speaker: Silvia Arber; 3 accepted talks
19:30	**Panel discussion**
20:30	**Simons Foundation Social**

Saturday, 02 March 2024

09:00	**Session 6: Decision-making** Invited speaker: Rajesh PN Rao; 3 accepted talks
10:45	**Session 7: Movement and motor control** Invited speaker: David Sussillo; 3 accepted talks
12:30	**Poster Session 3**
13:00	**Mini Chess Tournament**
16:00	**Session 8: Sensing and learning** Invited speaker: Elizabeth Hong; 4 accepted talks
18:00	**Session 9: Neural circuits** Invited speaker: Kafui Dzirasa; 3 accepted talks

Sunday, 03 March 2024

09:00	**Session 10: Neural coding and representations** Invited speaker: Mac Shine; 3 accepted talks
11:00	**Session 11:** Invited speakers: Laura Busse, Kay Tye; 2 accepted talks
12:30	**DEI Lunch (Opera Gale)**



Open-source Tools, Support,
Training, and Custom Design

For electrophysiology, microscopy, and behavior

open-ephys.org
info@oeps.tech

Made with ♥ in Lisbon & Atlanta

DataJoint

The SciOps Company

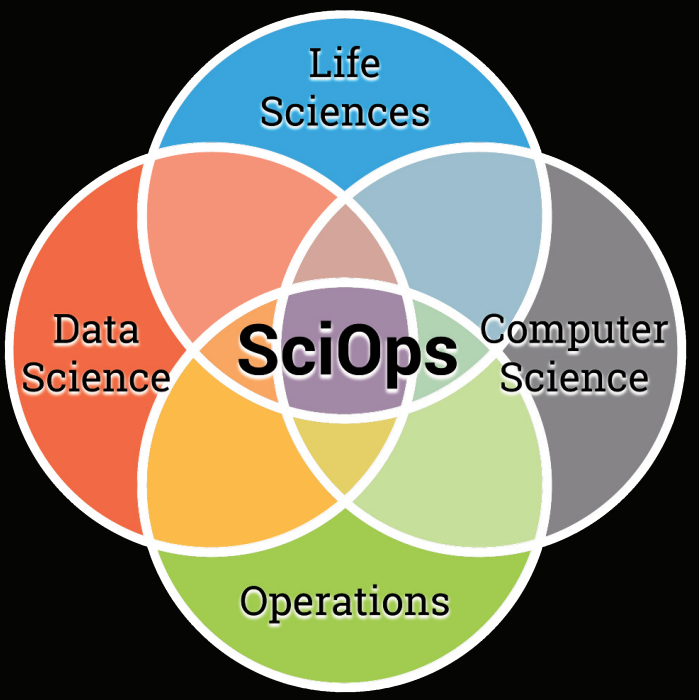
Research tech is accelerating...

petabytes of data • multiple modalities • new ML tools

...but data ops haven't kept up.

DIY • no standards • no systems • constant change

DataJoint can help:



See:
A Maturity Model for Operations in Neuroscience Research at
https://arxiv.org/abs/2401.00077

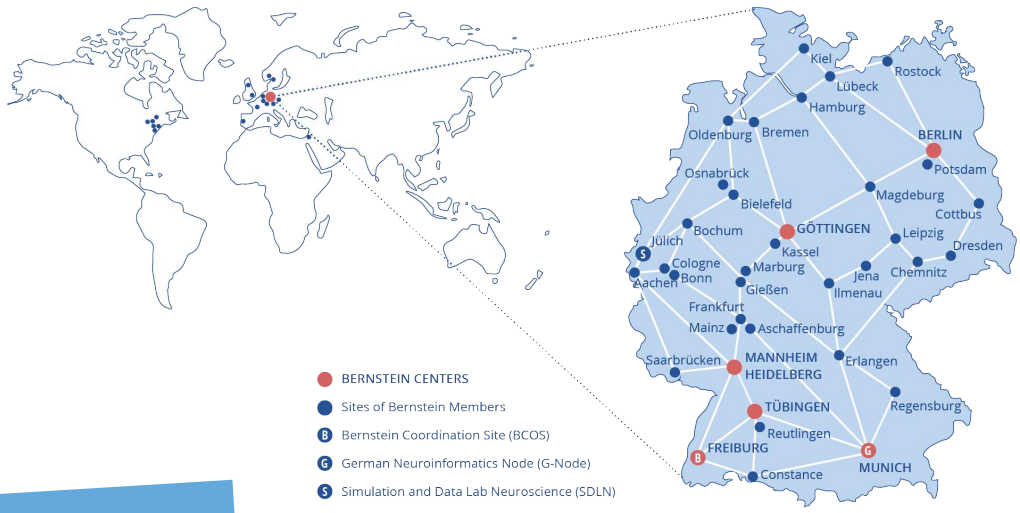
Elevate your data ops.
Prepare for AI in science.

info@datajoint.com

The Bernstein Network

Computational Neuroscience connects experimental and theoretical scientists. It comprises more than 200 research groups and 450 individual scientists from all over the world who combine experimental neuroscientific approaches with theoretical models and computer simulations. The scientists can rely on central infrastructural facilities of the network, which support the scientific dialogue.

bernstein-network.de



The Bernstein Conference

Sep 29 - Oct 2, 2024
Frankfurt am Main

Dmitriy Aronov (USA)
Alison Barker (Germany)
Helen Barron (UK)
Elizabeth Buffalo (USA)
Alex Cayco Gajic (France)
Jan Drugowitsch (USA)
Jakob Macke (Germany)
Tirin Moore (USA)
Mala Murthy (USA)
Memming Park (Portugal)
Xiao-Jing Wang (USA)

bernstein-conference.de

CATALYST NEURO

EMPOWERING **NEUROSCIENCE** RESEARCH THROUGH **ADVANCED SOFTWARE SOLUTIONS**

Core Software

NWB GUIDE

A desktop application that provides a code-free solution for conversion to NWB.

github.com/NeurodataWithoutBorders/nwb-guide

NeuroConv

Automate conversion from common neurophysiology data formats to NWB.

github.com/catalystneuro/neuroconv

NWB Inspector

Scan NWB files for potential errors and areas of improvement and generate a comprehensive report.

github.com/NeurodataWithoutBorders/nwbinspector

NWB Widgets

Generate interactive visualizations for NWB data

github.com/NeurodataWithoutBorders/nwbwidgets

Services

Data **Sharing**

CatalystNeuro can aid in converting your data to the Neurodata Without Borders (NWB) format, publishing it on the DANDI Archive, and providing access to a comprehensive ecosystem of analysis and visualization tools.

Schedule a data publication consultation

Spike Sorting **Pipelines**

CatalystNeuro offers bespoke spike sorting pipeline development using SpikeInterface, tailored to your task, recording technology, and research needs.

Schedule a spike sorting consultation

Data Infrastructure **on Grant Applications**

CatalystNeuro can enhance large research projects as a sub-award in grant applications, providing expertise in data and software engineering.

Schedule a consultation

catalystneuro.com inquiry@catalystneuro.com @CatalystNeuro github.com/CatalystNeuro

PlexStim™ Electrical Stimulator System

PlexStim™ is a 16-channel, constant current electrical stimulator that is electrically isolated, individually programmable, and is compatible with Plexon's OmniPlex® data acquisition systems and/or the CinePlex® Behavioral Research System.

Scan for more info

- With Robust API Support.
- Stimulation Patterns and Arbitrary Waveforms.

Looking For Electrodes?

Scan to view customizable specialty electrodes, probes and micro electrode arrays

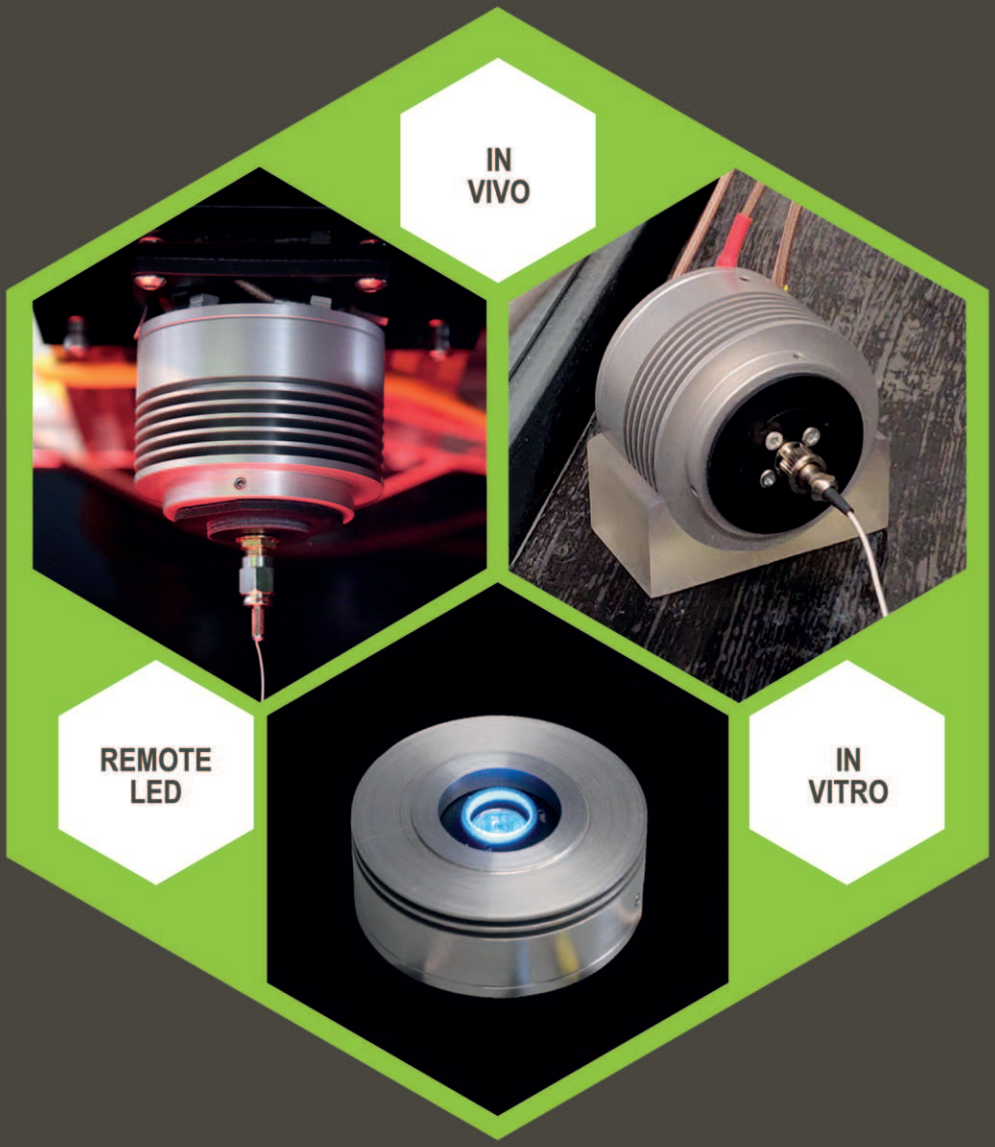
PLEXON

Contact a Plexon Sales Engineer for more information visit us at plexon.com or email us at info@plexon.com.

CyNexo

INNOVATIVE CUSTOM SOLUTIONS

NEW OPTOGENETICS PRODUCTS



The aoLED controller is a compact and efficient optogenetic stimulus system. It integrates a robust LED driver with a fiber-coupled LED, capable of producing high-bandwidth (up to 0.5 MHz) and high-intensity light signals. These signals can be delivered through a 200 um core high NA fiber or concentrated for in vitro applications. The electronic design enables precise analog signal modulation and fast digital modulations. Additionally, the built-in light power sensor ensures accurate monitoring of performance and stability over time.

More info at www.cynexo.com

Follow us on:

Contact us:

www.cynexo.com - email: info@cynexo.com - Phone: + 39.0432.1843913

CyNexo srl | Via Roma n. 6 | 33050 Trivignano Udinese (UD) | Italy

COSYNE 2025

Cosyne returns to Canada, save the date!

MAIN MEETING
Montreal, Mar 27–Mar 30

WORKSHOPS
Mont-Tremblant, Mar 31–Apr 01

About Cosyne

The annual Cosyne meeting provides an inclusive forum for the exchange of empirical and theoretical approaches to problems in systems neuroscience, in order to understand how neural systems function.

To encourage interdisciplinary interactions, the main meeting is arranged in a single track. A set of invited talks are selected by the Executive Committee and Organizing Committee, and additional talks and posters are selected by the Program Committee, based on submitted abstracts.

Cosyne topics include (but are not limited to): neural basis of behavior, sensory and motor systems, circuitry, learning, neural coding, natural scene statistics, dendritic computation, neural basis of persistent activity, nonlinear receptive field mapping, representations of time and sequence, reward systems, decision-making, synaptic plasticity, map formation and plasticity, population coding, attention, machine learning for neuroscience, and computation with spiking networks. Participants include pure experimentalists, pure theorists, and everything in between.

Cosyne 2024 Leadership

Organizing Committee

General Chairs

Jessica Cardin (Yale University) and Blake Richards (McGill University)

Program Chairs

Bing Brunton (University of Washington) and Chandramouli Chandrasekaran (Boston University)

Workshops Chairs

Andrew Saxe (University of Oxford) and SueYeon Chung (New York University, Flatiron)

Social Media Chair

Sabera Talukder (California Institute of Technology)

DEIA Chairs

Hysell Oviedo (The City University of New York) and Luke Sjulson (Albert Einstein)

Tutorial Chair

Il Memming Park (Champalimaud)

Fundraising Chair

Michael Long (New York University)

Audio-Video Media Chair

Carlos Stein Brito (Champalimaud)

Undergraduate Travel Chairs

Kimberly Stachenfeld (DeepMind) and Marcelo Mattar (New York University)

Poster Design

Maja Bialon

Executive Committee

Stephanie Palmer (University of Chicago)
Anne-Marie Oswald (University of Chicago)
Zachary Mainen (Champalimaud)
Alexandre Pouget (University of Geneva)
Anthony Zador (Cold Spring Harbor Laboratory)

Program Committee

Bing Brunton (U Washington) Co-chair
Chandramouli Chandrasekaran (Boston U) Co-chair
Yashar Ahmadian (U Cambridge)
Athena Akrami (UCL)
Adrian Bondy (Princeton)
Timothy Buschman (Princeton)
Alex Cayco Gajic (Ecole Normale Superieure)
Hannah Choi (Gatech)
Brian DePasquale (Boston U)
Sridhar Devarajan (Indian Inst Sci)
Laura Driscoll (Stanford)
Ann Duan (UCL)
Lea Duncker (Stanford)
Annegret Falkner (Princeton)
Rainer Friedrich (Friedrich Miescher Institute)
Juan Gallego (Imperial)
Matthew Golub (U Washington)
Bilal Haider (Georgia Tech)
Kiah Hardcastle (Harvard)
Kameron Harris (W Washington U)
Monika Jadi (Yale)
Santiago Jaramillo (U Oregon)
Jonathan Kao (UCLA)
Kohitij Kar (York U)
Ann Kennedy (Northwestern)
Guillaume Lajoie (MILA)
Laura Lewis (MIT)
Camilo Libedinsky (National U Singapore)
Scott Linderman (Stanford)
Ashok Litwin-Kumar (Columbia)
Emily Mackevicius (Basis)
Jorge Mejias (U Amsterdam)
Leenoy Meshulam (U Washington)
Jonathan Michaels (York U)
James Murray (U Oregon)
Hendrikje Nienborg (NIH)
Gouki Okazawa (Chinese Acad Sci)
Marino Pagan (UCL)
Angelique Paulk (Harvard)
Chethan Pandarinath (Emory)
Joseph Paton (Champalimaud)
Il Memming Park (Champalimaud)
Hannah Payne (Columbia)
Talmo Pereira (Salk)
Supratim Ray (Indian Inst Sci)
Erin Rich (Mount Sinai)
Ben Scott (Boston U)
Alireza Soltani (Dartmouth)
Nicholas Steinmetz (U Washington)
Carsen Stringer (HHMI)
Marie Suver (Vanderbilt)
Aparna Suvrathan (McGill)
John Tuthill (U Washington)
Ali Weber (Bryn Mawr)
Brady Weissbourd (MIT)
Alex Williams (NYU)
Klaus Wimmer (CRM)
Brad Wyble (Penn State)
Dan Yamins (Stanford)

Cosyne 2024 reviewers

Elliott Abe, Nixon Abraham, Emily Aery Jones, Everton J Agnes, Sweta Agrawal, Tosif Ahamed, Osama Ahmed, Hessam Akhlaghpour, Matteo Alleman, Kelsey Allen, Theoklitos Amvrosiadis, Stefano Anzellotti, Ahmet Arac, Salva Ardid, Abdoreza Asadpour, Hiroki Asari, Vivek Athalye, Lilach Avitan, Anthony Azevedo, Raymundo Baez Mendoza, Andrew Bahle, Shahab Bakhtiari, Abhishek Banerjee, Arkarup Banerjee, Pinglei Bao, João Barbosa, Alessandro Barri, Pouya Bashivan, Manuel Beiran, Arash Bellafard, Riccardo Beltramo, Marcus Benna, Brandon Bhasin, Vezha Boboeva, Scott Bolkan, Adrian Bondy, Tyler Bonnen, Conrado Bosman, Flora Bouchacourt, Anna Bowen, Jeff Bowers, Tyler Boyd-Meredith, Colin Bredenberg, Braden Brinkman, Lindsey Brown, Roberto Budzinski, Mathew Bull, Diana Burk, Nicholas Bush, Joana Cabral, Malcolm Campbell, Fanny Cazettes, Gloria Cecchini, Angus Chadwick, Sucheta Chakravarty, Julie Charlton, Thomas Chartrand, Xiaomo Chen, Selmaan Chettih, Feng-Kuei Chiang, Edmund Chong, Raeed Chowdhury, Amy Christensen, Radek Cichy, Justine Clery, Olivier Codol, Pip Coen, Uri Cohen, Christine Constantinople, Jonathan Cornford, Antonio Carlos Costa, Julia Cox, Bruno Cruz, Timothy Currier, Liset M de la Prida, Cristina Delgado Sallent, Sachira Denagamage, Xinyi Deng, Daniel Denman, Emily Dennis, Eric Denovellis, Darrel Deo, David Deutsch, Anita Devineni, Eric DeWitt, Ashesh Dhawale, Serena Di Santo, Efthymia (Mika) Diamanti, Bradley Dickerson, George Dimitriadis, Mario Dipoppa, Clementine Domine, Shirin Dora, Nicholas Dotson, Anne Draelos, Timothy Dunn, Eleonore Duvelle, Samuel Eckmann, Michael Economo, Seth Egger, Alan Emanuel, Rainer Engelken, Linlin Fan, Matthew Farrell, Jenelle Feather, Lisa Fenk, Andrew Fink, Steve Flavell, Ming-fai Fong, Lorenzo Fontolan, Jessica Fox, Rainer Friedrich, Sean Froudist-Walsh, Aniruddh Galgali, Razvan Gamanut, Richard Gao, Richard Gast, Jeff Gauthier, Matthew Getz, Arna Ghosh, Marcus Ghosh, Chad Giusti, Soledad Gonzalo Cogno, Iris Groen, Olivia Guest, Diksha Gupta, Harsha Gurnani, Eartha Guthman, Caroline Haimerl, Gregory Handy, Balazs Hangya, Richard Hardstone, Roy Harpaz, Julia Harris, Sarah Harvey, James Heald, Guillaume Hennequin, Roger Herikstad, Mark Histed, Helen Hou, Marc Howard, Chengcheng Huang, Alex Hyafil, Kyo Iigaya, Michele Insanally, Leyla Isik, Len Jacob, Caroline Jahn, Mohsen Jamali, Jorge Jaramillo, Jamie Jeanne, Kristopher Jensen, Danique Jeurissen, Aditi Jha, Xiaoxuan Jia, Jeff Johnston, Rebecca Jordan, Adrien Jouary, Naama Kadmon Harpaz, Louis Kang, Yul Kang, Tomomi Karigo, David Kastner, Stephen Keeley, Sepiedeh Keshavarzi, Adil Khan, Preeya Khanna, Mikail Khona, Soon Ho Kim, Sung Soo Kim, Iku Kimura, Shin Kira, Michael Kleinman, Ugne Klibaite, Leo Kozachkov, Nina Kudryashova, Anna Lakunina, Joonyeol Lee, Eric Leonardis, Jennifer Li, Anthony Lien, Caitlin Lienkaemper, Sukbin Lim, Grace Lindsay, Yuhan Helena Liu, Laureline Logiaco, Michael Lohse, Silvia Lopez-Guzman, Isabel Low, Thomas Luo, Tzuhsuan Ma, Xuan Ma, Amadeus Maes, Maxime Maheu, Joseph Makin, Tomoyuki Mano, Jeff Markowitz, Owen Marschall, Stefano Martiniani, Paul Masset, Francesca Mastrogiuseppe, Daniel McNamee, Christoph Miehl, Ana P Millan, Kevin Miller, Bin Min, Andrew Miri, Abhishek Mishra, Yalda Mohsenzadeh, Manuel Molano-Mazón, Mitchell Morton, Alice Mosberger, Asma Motiwala, Lyle Muller, Alex Murphy, Peter Murphy, Apurva Ratan Murty, Samuel Muscinelli, Devika Narain, Abdulraheem Nashef, Torbjørn Ness, Cristopher Niell, Ramon Nogueira Manas, Marcella Noorman, Timothy O'Leary, Daniel O'Shea, Emily Oby, Gabriel Ocker, Umberto Olcese, Amy Orsborn, Ivana Orsolic, Marius Pachitariu, Srikanth Padmala, Marino Pagan, Agostina Palmigiano, Rich Pang, Lia Papadopoulos, Philip Parker, John Patterson, John Pearson, Ulises Pereira, Matthew G Perich, Andrew Peters, Warren Pettine, Noah Pettit, Eugenio Piasini, Alex Piet, Viktor Plattner, Mark Plitt, William Podlaski, Roman Pogodin, Cindy Poo, Lorenzo Posani, Dean Pospisil, Dominique Pritchett, Raghav Rajan, Mohsen Rakhshan, Arjun Ramakrishnan, Hamidreza Ramezanpour, Jorge Ramirez, Chen Ran, Jonas Ranft, Ryan Raut, Supratim Ray, Stefano Recanatesi, Jasmine Reggiani, Sandra Reinert, Arbora Resulaj, Dylan Rich, Fabio Rizzoglio, Drew Robson, Chris Rodgers, Erica Rodriguez, Alex Roxin, Christopher Rozell, Mikail Rubinov, Caroline Runyan, Peter Rupprecht, Simone Russo,

Hannes Saal, Sadra Sadeh, Shervin Safavi, Alessandro Sanzeni, Anupama Sathyamurthy, Shreya Saxena, Evan Schaffer, Christian Schmid, Marieke Scholvinck, Manuel Schottdorf, Sylvia Schroeder, Friedrich Schuessler, Audrey Sederberg, Nishal Shah, Yuxiu Shao, Weikang Shi, Yanliang Shi, Daisuke Shimaoka, Mac Shine, Vinay Shirhatti, Josh Siegle, Luke Sjulson, Joana Soldado Magraner, Saray Soldado Magraner, Devarajan Sridharan, Ramanujan Srinath, Frederic Stoll, Iris Stone, Katherine Storrs, Divya Subramanian, Xulu Sun, Brian Sweis, Sina Tafazoli, Bharath Chandra Talluri, Hui Min Tan, Craig Taswell, Tatjana Tchumatchenko, Hazem Toutounji, Max Turner, Timo van Kerkoerle, Rahul Venkatesh, Catalina Vich Llompart, Amit Vinograd, Anna Vlasits, Saurabh Vyas, Mark Wagner, Siwei Wang, Siyu Wang, Adrian Wanner, Moritz Weglage, Xuexin Wei, Ziqiang Wei, Caleb Weinreb, Sam Whitehead, Ross Williamson, Katharina Wilmes, Steffen Wolff, Anqi Wu, Yue Kris Wu, Brad Wyble, Ji Xia, Ruobing Xia, Xize Xu, Qianli Yang, Zhiwen Ye, Murat Yildirim, Roxana Zeraati, Xingjian Zhang, Yuan Zhao, Lin Zhong, Chengxu Zhuang, Corey Ziemba, and Christopher Zimmerman.

Special thanks to Titipat Achakulvisut, Daniel Acuna, and Konrad Kording for writing and managing the automated software for reviewer abstract assignment.

Conference Support

Administrative Support, Registration, Hotels

Leslie Weekes, Cosyne

Travel Grants

The Cosyne community is committed to bringing talented scientists together at our annual meeting, regardless of their ability to afford travel. Thus, a number of travel grants are awarded to students, postdocs, and PIs for travel to the Cosyne meeting. Five award granting programs were available for Cosyne 2024. Each award covers at least $500 towards travel and meeting attendance costs (less for Childcare grants).

The generosity of our sponsors helps make these travel grant programs possible. Cosyne Travel Grant Programs are supported entirely by the following corporations and foundations:

- Simons Foundation
- Reality Labs
- Burroughs Wellcome Fund

Cosyne Presenters Travel Grant Program

These grants support early career scientists with highly scored abstracts to enable them to present their work at the meeting.

The 2024 recipients are: Leonardo Agueci, Mojtaba Abbaszadeh, Yoel Sanchez Araujo, Ikhwan Bin Khalid, Caitriona Costello, Martha Garcia Garcia, Cristina Giossi, Mrugsen Gopnarayan, Shailee Jain, Mitra Javadzadeh, Aditi Jha, Yeowon Kim, Tianqing Li, Qi Lin, Hung Lo, Nischal Mainali, Asma Motiwala, Somang Paeng, Danilo Perez, Eryn Sale, Yuria Shimizu, Alexander Silva, Selena Singh, Tahereh Toosi, Shirin Vafaei, Ying Yu, Roxana Zeraati, and Zhanqi Zhang.

Cosyne New Attendees Travel Grant Program

These grants help bring scientists that have not previously attended Cosyne to the meeting for exchange of ideas with the community.

The 2024 recipients are: Hamza Abdelhedi, Anamika Agrawal, Destinee Aponte, Shayan Ashfar, Eleanor Brown, Elisa Chinigo, Shoshana Chipman, Oumaima Essamadi, Zulfar Ghulam-Jelani, Kelvin Q Laracuente, Jin Hwa Lee, Hildelith Leyser, Emily Lopez, Antara Majumdar, Emiliano Jimenez Marquez, Paulina Paiz, Mohamed Athif Mohammadu Rizan, Magdalena Sabat, and Katya Tsimring.

Cosyne Undergraduate Travel Grant Program

These grants help bring promising undergraduate students with strong interest in neuroscience to the meeting.

The 2024 recipients are: Gideon Alex, Rodrigo Anjos, Zeineb Ben Fredj, Tomás Caldeira, Hossam El-Basiouny, Pranjal Garg, Shaurya Goyal, Chaitanya Kapoor, Alisa Leshchenko, Camila Maura, Vatsala Nema, Margaret Pozo, and Cintia Ricaele Ferreira da Silva.

Cosyne Childcare Travel Grant Program

Cosyne Childcare Grants help cover childcare expenses incurred by participation in Cosyne 2024.

The 2024 recipients are: Maximilian Eggl, Avital Hahamy, Anat Leibovici, Shujing Li, Srikanth Ramaswamy, Lee Susman, and Anqi Wu.

Social media policy

Cosyne encourages the use of social media before, during, and after the conference, so long as it falls within the following rules:

- Do not capture or share details of any unpublished data presented at the meeting.
- If you are unsure whether data is unpublished, check with the presenter before sharing the information.
- Respect presenters' wishes if they indicate the information presented is not to be shared.

Stay up to date with Cosyne 2024 #Cosyne24

Cosyne Code of Conduct

Purpose

At Cosyne, we strive for open and honest intellectual debate as part of a welcoming and inclusive atmosphere. This requires a community and an environment that recognizes and respects the inherent worth of every person.

Sources

This code of conduct is based on standards and language set at other meetings, whose organizing boards convened special working groups of scientific and legal experts to set their policies. We follow, in particular, those guidelines established for the Gordon Research Conferences, the Society for Neuroscience Annual Meeting, and NeurIPS.

The following code of conduct has been adapted from:

https://www.grc.org/about/grc-policies-and-legal-disclaimers

https://www.sfn.org/about/professional-conduct/code-of-conduct-at-sfn-events

https://nips.cc/public/CodeOfConduct

Other online resources:

http://changingourcampus.org

https://www.sfn.org/about/professional-conduct/sfn-ethics-policy

Responsibilities

All participants, volunteers, organizers, reviewers, speakers, sponsors, and volunteers (referred to as "Participants" collectively throughout this document) at our Conference, workshops, and Conference-sponsored social events—are required to agree with this Code of Conduct both during an event and on official communication channels, including social media.

Sponsors are equally subject to this Code of Conduct. In particular, sponsors should not use images, activities, or other materials that are of a sexual, racial, or otherwise offensive nature. This code applies both to official sponsors as well as any organization that uses the Conference name as branding as part of its activities at or around the Conference.

Organizers will enforce this Code, and it is expected that all Participants will cooperate to help ensure a safe and inclusive environment for everyone.

Policy

The conference commits itself to providing an experience for all Participants that is free from the following:

Harassment, bullying, discrimination which includes but is not limited to:

- Offensive comments related to age, race, religion, creed, color, gender (including transgender/gender identity), sexual orientation, medical condition, physical or intellectual disability, pregnancy, or medical conditions, national origin or ancestry.
- Intimidation, personal attacks, harassment, unnecessary disruption of talks or other conference events.

Inappropriate or unprofessional behavior that interferes with another's full participation including:

- Sexual harassment, stalking, following, harassing photography or recording, inappropriate physical contact, unwelcome attention, public vulgar exchanges, derogatory name-calling, and diminutive characterizations.
- Use of images, activities, or other materials that are of a sexual, racial, or otherwise offensive nature that may create an inappropriate or toxic environment.
- Disorderly, boisterous, or disruptive conduct including fighting, coercion, theft, damage to property, or any mistreatment or non-businesslike behavior towards participants.
- "Zoom bombing" or any virtual activity that is not related to the topic of discussion which detracts from the topic or the purpose of the program. This includes inappropriate remarks in chat areas as deemed inappropriate by presenters/monitors/event leaders.

Scientific misconduct: including fabrication, falsification, or plagiarism of paper submissions or research presentations, including demos, exhibits or posters. Cosyne asks each session chair and organizing and reviewing committee member to promote rigorous analysis of all science presented for or at the meeting in a manner respectful to all attendees.

This Code of Conduct applies to the actual meeting sites and Conference venues where Cosyne business is being conducted, including physical venues, online venues, and official virtual engagement platforms, including video, virtual streaming, and chat-based interactions. Cosyne is not responsible for non-sponsored activity or behavior that may occur at non-sponsored locations such as hotels, restaurants, or physical or virtual locations not otherwise a sanctioned space for sponsored events. *Nonetheless, any issues brought to the Hotline Relations Counselors will be explored.* Moreover, Cosyne will not actively monitor social media platforms but will follow up on issues of harassment and violations of the code of conduct that occur on those platforms that are specifically related to the Cosyne program, during the course of Cosyne, if and when they are brought to our attention.

Complaint reporting

The Conference encourages all Participants to immediately report any incidents of discrimination,harassment, unprofessional conduct, and/or retaliation so that complaints can be quickly and fairly resolved. There will be no retaliation against any Participant who brings a complaint or submits an incident report in good faith or who honestly assists in investigating such a complaint. If you have concerns related to your participation/interaction at the Conference or Conference sanctioned events, or observe someone else's difficulties, or have any other concerns you wish to share, please write to *CosyneHotline@gmail.com* or by calling the *Cosyne Hotline phone number at +1-858-208-3810* where you can speak with an HR Consultant.

Action

If a Participant engages in any inappropriate behavior as defined herein, the Conference organizers may take action as deemed appropriate, including: a formal or informal warning to the offender, expulsion from the conference with no refund, barring from participation in future conferences or their organization, reporting the incident to the offender's local institution or funding agencies, or reporting the incident to local authorities or law enforcement. A response of "just joking" is not acceptable. If action is taken, an appeals process will be made available. There will be no retaliation against any Participant who brings a complaint or submits an incident report in good faith or who honestly assists in investigating such a complaint. All issues brought forth to the onsite HR Consultant during the course of a Conference will be immediately investigated.

Program

Note: Institutions listed in the program are the primary affiliation of the first author. For the complete list, please consult the abstracts.

All times are GMT (Greenwich Mean Time)

Thursday, 29 February 2024

08:00	**Neurodata Without Borders Data Standard and DANDI Archive for Neurophysiology Data Tutorial**
09:45	**Learn to Use the International Brain Laboratory (IBL) Brainwide Map Dataset**
11:00	**Mind Matching (Neuromatch)**
12:30	**Cosyne 2024 Tutorial sponsored by the Simons Foundation**
	Get some SLEAP: Automating behavior quantification using deep learning
	Talmo Pereira
14:00	**Registration opens**
16:30	**Welcome reception**
19:00	**Opening remarks**

Session 1:

(Chair: Bing Brunton, Chandramouli Chandrasekaran)

19:30 The Mind of a Bee
Lars Chittka, Queen Mary University of London **(invited)** 35

20:30 **Poster Session 1**

Friday, 01 March 2024

Session 2: Reward and reinforcement learning

(Chair: Hannah Choi)

09:00 Neural compositionality at last!
Alex Pouget, University of Geneva **(invited)** 35

09:40 Dopamine neurons encode a multidimensional probabilistic map of future reward
M. Sousa, P. Bujalski, B. Cruz, K. Louie, D. McNamee, J. Paton, Champalimaud Research 36

09:55 Two types of locus coeruleus norepinephrine neurons drive reinforcement learning
Z. Su, K. Jung, K. Hagihara, J. Cohen, Allen Institute for Neural Dynamics 37

10:10 A cerebellar granule cell-climbing fiber computation to learn to predict reward
M. Garcia Garcia, M. Wagner, A. Kapoor, O. Akinwale, L. Takemaru, T. H. Kim, C. Paton, A. Litwin-Kumar, M. Schnitzer, L. Luo, National Institutes of Health 38

10:30 **Coffee break**

Session 3: Learning

(Chair: Alex Cayco Gajic)

11:00 What is intelligence?
Blaise Aguera y Arcas, Google Research **(invited)** 35

11:40 Estrogenic control of reward prediction errors and reinforcement learning
C. Golden, D. Kaur, A. Mah, A. Martin, D. Levy, T. Yamaguchi, D. Lin, C. Aoki, C. Constantinople, New York University 38

Saturday, 02 March 2024

Sunday, 03 March 2024

Session 10: Neural coding and representations

(Chair: Erin Rich)

09:00 The Role of the Thalamus in Complex Adaptive Dynamic Behaviour
Mac Shine, The University of Sydney **(invited)** 36

09:40 Comparing the geometry of static and dynamic coding across the auditory pathway.
O. Lahrach, S. Bagure, B. Bathellier, S. Ostojic, Ecole Normale Superieure 50

09:55 Representational geometry of hierarchical category structures in the monkey inferotemporal cortex
N. Y. Lo, R. Kiani, S. Chung, Flatiron Institute 50

10:10 Rapid, concerted switching of the neural code in inferotemporal cortex
Y. Shi, D. Bi, J. Hesse, F. Lanfranchi, S. Chen, D. Tsao, California Institute of Technology . 51

10:30 **Coffee break**

Session 11:

(Chair: Luke Sjulson)

11:00 Neuronal circuit dynamics in the visual thalamocortical system
Laura Busse, LMU Munich **(invited)** 36

11:40 Emergence of state modulation in a developing cortical circuit
J. Majnik, S. Zangila, R. Cossart, J. Platel, University of Aix-Marseille 51

11:55 Evolving a Toxic Hierarchy into a Sustainable Ecosystem *(DEI talk)*
Kay Tye, Salk Institute for Biological Studies **(invited)** 36

12:25 **Closing remarks**

12:30 **DEI Lunch (Opera Gale)**

15:00 **Transit to Cascais**

Poster Session 1 Thursday 29 February 2024

Poster Session 2 — Friday 01 March 2024

Poster Session 3 Saturday 02 March 2024

Abstracts

Abstracts for talks appear first, in order of presentation; those for posters next, in order of poster session and board number. An index of all authors appears at the back.

T-1. The Mind of a Bee

Lars Chittka
Queen Mary University of London

T-2. Neural compositionality at last!

Alex Pouget
University of Geneva

T-3. What is intelligence?

Blaise Aguera y Arcas
Google Research

T-4. Uncovering structure and individuality in motor patterns deployed by larval zebrafish during navigation

Claire Wyart
ICM (Brain & Spine Institute)

T-5. Brainstem circuits regulating body movements

Silvia Arber
University of Basel

T-6. A Sensory-Motor Theory of the Neocortex based on Active Predictive Coding

Rajesh PN Rao
University of Washington Seattle

T-7. Flexible Multitask Computation in Recurrent Networks Utilizes Shared Dynamical Motifs

David Sussillo[1,2]

[1]Meta Reality Labs
[2]Stanford University

T-8. Odor relationships in the natural world as an organizational axis of the fly olfactory code

Elizabeth Hong

California Institute of Technology

T-9. Decoding emotional brain states

Kafui Dzirasa

Duke University

T-10. The Role of the Thalamus in Complex Adaptive Dynamic Behaviour

Mac Shine

The University of Sydney

T-11. Neuronal circuit dynamics in the visual thalamocortical system

Laura Busse

LMU Munich

T-12. Evolving a Toxic Hierarchy into a Sustainable Ecosystem

Kay Tye

Salk Institute for Biological Studies

T-13. Dopamine neurons encode a multidimensional probabilistic map of future reward

Margarida Sousa[1] MARGARIDA.SOUSA@RESEARCH.FCHAMPALIMAUD.ORG
Pawel Bujalski[1] PAWEL.BUJALSKI@RESEARCH.FCHAMPALIMAUD.ORG
Bruno Cruz[2,3] BRUNO.CRUZ@ALLENINSTITUTE.ORG
Kenway Louie[4,5] KL837@NYU.EDU
Daniel McNamee[1] DANIEL.MCNAMEE@RESEARCH.FCHAMPALIMAUD.ORG
Joe Paton[6] JOE.PATON@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Research
[2]Allen Institute
[3]Neural Dynamics

[4]New York University
[5]Center for Neural Science
[6]Champalimaud Foundation

Learning to predict rewards is a fundamental driver of adaptive behavior. Midbrain dopamine neurons (DANs) play a key role in such learning by signaling reward prediction errors (RPEs) that teach recipient circuits about expected rewards given current circumstances and actions. However, the algorithm that DANs are thought to provide a substrate for, temporal difference (TD) reinforcement learning (RL), learns the mean of temporally discounted expected future rewards, discarding useful information concerning experienced distributions of reward amounts and delays. Here we present time-magnitude RL (TMRL), a multidimensional variant of distributional reinforcement learning that learns the joint distribution of future rewards over time and magnitude using an efficient code that adapts to environmental statistics. In addition, we discovered signatures of TMRL-like computations in the activity of optogenetically identified DANs in mice during a classical conditioning task. Specifically, we found significant diversity in both temporal discounting and tuning for the magnitude of rewards across DANs, features that allow the computation of a two dimensional, probabilistic map of future rewards from just 450ms of neural activity recorded from a population of DANs in response to a reward-predictive cue. In addition, reward time predictions derived from this population code correlated with the timing of anticipatory behavior, suggesting the same temporal information is used to guide decisions regarding when to act. Finally, by simulating behavior in a foraging environment, we highlight benefits of access to a joint probability distribution of reward over time and magnitude in the face of dynamic reward landscapes and internal physiological need states. These findings demonstrate surprisingly rich probabilistic reward information that is learned and communicated to DANs, and suggest a simple, local-in-time extension of TD learning algorithms that explains how such information may be acquired and computed.

T-14. Two types of locus coeruleus norepinephrine neurons drive reinforcement learning

Zhixiao Su ZHIXIAOSU@GMAIL.COM
Kanghoon Jung KANGHOON.JUNG@ALLENINSTITUTE.ORG
Kenta Hagihara KENTA.HAGIHARA@ALLENINSTITUTE.ORG
Jeremiah Cohen JEREMIAH.COHEN@ALLENINSTITUTE.ORG

Allen Institute for Neural Dynamics

The cerebral cortex generates flexible behavior by learning. Reinforcement learning is thought to be driven by error signals in dopamine neurons. However, they project more densely to basal ganglia than cortex, leaving open the possibility of another source of learning signals for cortex. The locus coeruleus (LC) contains most of the brain's norepinephrine (NE) neurons and project broadly to the cortex. We measured activity from identified mouse LC-NE neurons during a behavioral task requiring ongoing learning from reward prediction errors (RPEs). We found two types of LC-NE neurons: neurons with wide action potentials (type I) were excited by positive RPE and showed an increasing relationship with change of choice likelihood. Neurons with thin action potentials (type II) were excited by lack of reward and showed a decreasing relationship with change of choice likelihood. NE dynamics in prelimbic cortex (PL), which has been previously shown to encode task-critical decision variables, is excited by RPE, indicating its contribution to learning from errors. Indeed, silencing LC-NE neurons changed future choices, as predicted from the electrophysiological recordings and a model of how RPEs are used to guide learning. We reveal functional heterogeneity of a neuromodulatory system in the brain and show that NE inputs to cortex as a quantitative learning signal for flexible behavior.

T-15. A cerebellar granule cell-climbing fiber computation to learn to predict reward

Martha Garcia Garcia[1,2] MARTHA.GARCIAGARCIA@NIH.GOV
Mark Wagner[1,3] MARK.WAGNER@NIH.GOV
Akash Kapoor[1] AKASHKAPOOR2017@GMAIL.COM
Oluwatobi Akinwale[1] OLUWATOBI.AKINWALE@NIH.GOV
Lina Takemaru[1] LINATAKEMARU@OUTLOOK.COM
Tony Hyun Kim[4] KIMTH@STANFORD.EDU
Casey Paton[1] CMP99@CORNELL.EDU
Ashok Litwin-Kumar[5] AK3625@COLUMBIA.EDU
Mark Schnitzer[4] MSCHNITZ@STANFORD.EDU
Liqun Luo[4] LLUO@STANFORD.EDU

[1]National Institutes of Health
[2]National Institute of Neurological Disorders and Stroke
[3]NINDS
[4]Stanford University
[5]Columbia University

The cerebellum helps animals learn to predict events. More recently, cerebellar activity was found to encode reward. Yet it remains unclear how the cerebellum uses reward signals. Cerebellar Purkinje cell (PkC) computations depend on relative activity of granule cells (GrCs) and climbing fibers (CFs)—but these have never been recorded together. Here, we imaged simultaneous activity of GrCs and CFs while animals learned to make forelimb movements for delayed water reward. As mice learned reward timing, GrCs and CFs jointly developed reward expectation signals: many GrCs sustained anticipatory activity throughout the delay until reward, which triggered widespread time-locked CF spiking. We computed CF-dependent plasticity rules governing GrC→PkC synaptic changes, demonstrating that reward-evoked CF spikes sufficed to grade synaptic strengths of many GrCs by their anticipatory timing. We predicted and confirmed that PkCs could thereby continuously estimate seconds of time passage prior to reward, broadening possible roles for cerebellar learning.

T-16. Estrogenic control of reward prediction errors and reinforcement learning

Carla Golden[1,2] CG163@NYU.EDU
Daljit Kaur[3] DALJIT.KAUR@EINSTEINMED.EDU
Andrew Mah[1,2] AM9056@NYU.EDU
Audrey Martin[1] ACM868@NYU.EDU
Diana Levy[1] EDL389@NYU.EDU
Takashi Yamaguchi[4] TAKASHI.YAMAGUCHI@NYULANGONE.ORG
Dayu Lin[4] DAYU.LIN@NYULANGONE.ORG
Chiye Aoki[1] CA3@NYU.EDU
Christine Constantinople[1,2] CMC472@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Albert Einstein
[4]New York University Langone Medical Center

Despite the broad influence of gonadal hormones throughout the brain, little is known about how these hormones influence cognitive behaviors and their underlying neural substrates. Exogenous 17β-estradiol modulates dopamine signaling in the nucleus accumbens core (NAcc)1, which receives dopaminergic input from the midbrain that is thought to support reinforcement learning. According to this framework, animals or agents learn the value of different actions from experience and use those value estimates to guide behavior2. Action values are iteratively updated based on reward prediction errors (RPEs), or the difference between received and expected rewards. A wealth of evidence suggests that dopamine released in the NAcc instantiates RPEs3, raising the intriguing possibility that 17β-estradiol can influence reinforcement learning. Here we show that estrogenic hormones that fluctuate over female rats' reproductive cycles enhance reinforcement learning by increasing the dynamic range of dopamine signaling in the NAcc, producing a multiplicative gain on RPEs. We trained rats to perform a temporal wagering task with different reward states. Rats adjusted how quickly they initiated trials across states, balancing effort against expected rewards. In fertile stages, females showed greater sensitivity to reward states, which we show is driven by enhanced encoding of dopamine RPEs in the NAcc that increase or decrease the perceived value of the environment. During fertile stages, dopamine transporters were reduced in expression, and computational modeling showed that reduced reuptake could increase the gain of RPEs. Genetic suppression of estrogen

receptors in midbrain eliminated hormonal modulation of behavior. Estrogenic hormones therefore control the rate of reinforcement learning by regulating RPEs via dopamine reuptake, providing a mechanism by which hormones influence neural dynamics for motivation and learning.

T-17. Predictive auxiliary objectives in deep RL mimic learning in the brain

Ching Fang[1] CHINGFANG17@GMAIL.COM
Kimberly Stachenfeld[2] STACHENFELD@GOOGLE.COM

[1]Columbia University
[2]Google DeepMind; Columbia University

The ability to predict upcoming events has been hypothesized to comprise a key aspect of natural and machine cognition. This is supported by trends in deep reinforcement learning (RL), where self-supervised auxiliary objectives such as prediction are widely used to support representation learning and improve task performance. Here, we study the effects predictive auxiliary objectives have on representation learning across different modules of an RL system and how these mimic representational changes observed in the brain. We find that predictive objectives improve and stabilize learning particularly in resource-limited architectures. We identify settings where longer predictive horizons better support representational transfer. Furthermore, we find that representational changes in this RL system bear a striking resemblance to changes in neural activity observed in the brain across various experiments. Specifically, we draw a connection between the auxiliary predictive model of the RL system and hippocampus, an area thought to learn a predictive model to support memory-guided behavior. We also connect the encoder network and the value learning network of the RL system to visual cortex and striatum in the brain, respectively. This work demonstrates how representation learning in deep RL systems can provide an interpretable framework for modeling multi-region interactions in the brain. The deep RL perspective taken here also suggests an additional role of the hippocampus in the brain– that of an auxiliary learning system that benefits representation learning in other regions.

T-18. A neural mechanism for learning from delayed postingestive feedback

Christopher Zimmerman[1,2] CZIMMERMAN@PRINCETON.EDU
Alejandro Pan-Vazquez[1] APV2@PRINCETON.EDU
Bichan Wu[1] BICHANW@PRINCETON.EDU
Emma Keppler[1] EK1856@PRINCETON.EDU
Eartha Guthman[1,2] EGUTHMAN@PRINCETON.EDU
Robert Fetcho[1] RF6456@PRINCETON.EDU
Scott Bolkan[1] SBOLKAN@PRINCETON.EDU
Brenna McMannon[1] BM1327@PRINCETON.EDU
Junuk Lee[1] JUNUKL@PRINCETON.EDU
Austin Hoag[1] AHOAG@PRINCETON.EDU
Laura Lynch[1] LL3@PRINCETON.EDU
Sanjeev Janarthanan[1] SJ0470@PRINCETON.EDU
Juan Lopez Luna[1] JL4459@PRINCETON.EDU
Adrian Bondy[1,2] ABONDY@PRINCETON.EDU
Annegret Falkner[1] AFALKNER@PRINCETON.EDU
Samuel Wang[1] SSWANG@PRINCETON.EDU
Ilana Witten[1] IWITTEN@PRINCETON.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute

Summary. Animals learn the value of foods based on their postingestive effects and thereby develop aversions to foods that are toxic and preferences to those that are nutritious. However, it remains unclear how the brain is able to assign credit to flavors experienced during a meal with postingestive feedback signals that can arise after a substantial delay. Here, we reveal an unexpected role for postingestive reactivation of neural flavor representations in this temporal credit assignment process. To begin, we leverage the fact that mice learn to associate novel, but not familiar, flavors with delayed gastric malaise signals to investigate how the brain represents flavors that support aversive postingestive learning. Surveying cellular resolution brainwide activation patterns reveals that a network of amygdala regions is unique in being preferentially activated by novel flavors across every stage of the learning process: the initial meal, delayed malaise, and memory retrieval. By combining high-density recordings in the amygdala with optogenetic stimulation of genetically defined hindbrain malaise cells, we find that postingestive malaise signals potently and specifically reactivate amygdalar novel flavor representations from a recent meal. The degree of malaise-driven reactivation of individual neurons predicts strengthening of flavor

responses upon memory retrieval, leading to stabilization of the population-level representation of the recently consumed flavor. In contrast, meals without postingestive consequences degrade neural flavor representations as flavors become familiar and safe. Thus, our findings demonstrate that interoceptive reactivation of amygdalar flavor representations provides a neural mechanism to resolve the temporal credit assignment problem inherent to postingestive learning.

T-19. A complete sensorimotor pathway underlying altitude control in flight

Anne Erickson[1] AERICKSON@CALTECH.EDU
Jaison Omoto[2] JOMOTO@CALTECH.EDU
S. David Stupski[3] SSTUPSKI@UNR.EDU
Floris van Breugel[3] FVANBREUGEL@UNR.EDU
Michael Dickinson[2] FLYMAN@CALTECH.EDU

[1]Caltech
[2]California Institute of Technology
[3]University of Nevada, Reno

All flying animals must possess the ability to regulate altitude. Yet, due to the challenge of measuring neural activity in flying organisms, the circuits underlying altitude control remain poorly understood. We studied the control of flight altitude in fruit flies using an integrative approach that linked sensory, motor, biomechanical, and aerodynamic phenomena across scales. We conducted an optogenetic screen to activate specific bilateral pairs of descending neurons—interneurons that receive inputs in the brain and project to downstream motor centers in the ventral nerve cord—while recording the kinematics of wings flapping at ~220 Hz using high-speed videography and markerless 3D tracking. Our screen identified a single pair of descending neurons, DNa10, that upon activation, elicit wing kinematics consistent with a decrease in altitude. To understand the response properties of these cells, we used 2-photon calcium imaging to record DNa10 neurons during tethered flight and presentation of different patterns of optic flow. DNa10 neurons were most responsive to full-field downward translational motion—a visual stimulus that induces a bilateral decrease in wing stroke amplitude. Next, to identify the aerodynamic consequences of DNa10 activation, we replayed the wing motion observed before and during optogenetic activation on a dynamically scaled robot equipped with force and torque sensors. Robotic replay of wingbeats during DNa10 activation decreased upward force production relative to pre-activation wingbeats, a change characteristic of a reduction in flight altitude. To test this aerodynamic prediction, we conducted a free-flight optogenetic activation experiment using a specialized wind tunnel, and confirmed that DNa10 activation reliably induced a decrease in altitude. To our knowledge, these results provide the first characterization of a neural circuit for the control of flight altitude.

T-20. Predictive representations for rapid learning of complex tasks via composition of behavioral primitives

Pablo Tano PABLOTANORETAMALES@GMAIL.COM
Alex Pouget ALEXANDRE.POUGET@UNIGE.CH

University of Geneva

Animals exhibit an astonishing capacity to rapidly adapt their behavior to novel challenges. We show that such adaptability is made possible by combining two cornerstones of biological intelligence: (1) the ability of the brain to predict how the world changes as a consequence of behavior and (2) the fact that behavior can be generated by composing together multiple smooth behavioral primitives with high predictability. We simulated an agent that learns complex tasks according to these two principles. In the first phase, the agent gradually builds a predictive representation (PR) of the consequences of following sequences of behavioral primitives. In the second phase, the desired task is solved via gradient descent by composing differentiable behavioral primitives. Crucially, most of this goal directed learning can be performed offline (i.e. without having to execute actions in the real world), using the predictive representation to simulate the real world while also providing an analytical expression for the gradient of the loss function. We show that this division of labor allows extremely sample-efficient learning that is several orders of magnitude faster than model-free RL. Finally, we show that in order for the behavioral primitives to be useful, their spatiotemporal complexity must strike a balance: primitives must be complex enough to move the agent through the environment efficiently, yet sufficiently simple to ensure their predictability.

T-21. Mechanisms for working memory and evidence accumulation during olfactory navigation

Aaron Lanz[1,2] AJL787@NYU.EDU
Nicholas Kathman[1] NICHOLAS.KATHMAN@NYULANGONE.ORG
Katherine Nagel[1] KATHERINE.NAGEL@NYULANGONE.ORG

[1]NYU Grossman School of Medicine
[2]Neuroscience

Olfactory plume navigation is an ethological task in which information about the location of a goal (an odor source) is available only intermittently. Here we use Drosophila melanogaster to show how a recurrent circuit in the fly navigation center generates working memory and evidence accumulation to support navigation towards an intermittent olfactory goal. Using in vivo calcium imaging in behaving flies, we show that hΔCK neurons of the fan-shaped body display persistent bump activity that can build up in response to successive odor encounters and outlast the odor by several seconds. Persistent activity is associated with maintenance of a straight trajectory in the direction adopted during odor. Connectomic analysis suggests that these dynamics are generated locally by a recurrent network resembling a ring attractor. hΔCK neurons show both local and global recurrent connectivity with populations of putative excitatory and inhibitory neurons. Using patch clamp electrophysiology, we find that recurrent excitation onto hΔCK neurons uses slow metabotropic signaling, and that inhibition onto hΔCK undergoes robust short-term synaptic depression. Based on both connectome data and our synaptic physiology, we construct a dynamical model of this network and show that it generates persistent activity in response to simulated odor inputs. Slow metabotropic excitation at recurrent synapses endows our model with more robust and graded persistence than fast synaptic transmission, while short-term depression at inhibitory synapses allows increasing numbers of odor encounters to produce longer memory durations. Together, our data and model provide insight into how synaptic dynamics within an attractor-like network can produce a history-dependent working memory signal for navigation.

T-22. Representational similarity modulates neural and behavioral signatures of novelty

Sophia Becker[1,2] SOPHIA.BECKER@EPFL.CH
Alireza Modirshanechi[1] ALIREZA.MODIRSHANECHI@EPFL.CH
Wulfram Gerstner[3] WULFRAM.GERSTNER@EPFL.CH

[1]EPFL
[2]School Life Sciences, School of Computer and Communication Sciences
[3]ecole Polytechnique Federale de Lausanne

Novelty signals in the brain gate learning and drive exploratory behaviors, yet it remains elusive how stimulus representations influence novelty computation. In particular, existing novelty models in computational neuroscience fail to account for the experimentally observed impact of stimulus similarities on novelty computation in the brain. Here, we present a unifying, biologically plausible model that captures how stimulus similarities modulate novelty signals in the brain and, as a result, influence novelty-driven learning and exploration. By applying our model to two open data sets, we quantify and explain (i) how generalization across similar visual stimuli affects novelty responses in the mouse visual cortex and (ii) how generalization across nearby locations in a labyrinth impacts mouse exploration. Our model unifies and explains distinct neural and behavioral signatures of novelty-related processing and enables theory-driven experiment design to understand the neural mechanisms of novelty computation in naturalistic tasks and environments.

T-23. Beyond Geometry: Comparing the Temporal Structure of Neural Computation with Dynamical Similarity Analysis

Mitchell Ostrow[1,2] OSTROW@MIT.EDU
Adam Eisen[1] EISENAJ@MIT.EDU
Leo Kozachkov[3,2] LEOKOZ8@MIT.EDU
Ila Fiete[1] FIETE@MIT.EDU

[1]MIT
[2]Brain and Cognitive Sciences
[3]Massachusetts Institute of Technology

How can we tell whether two neural networks utilize the same internal processes for a particular computation? This question is pertinent for multiple subfields of computational and systems neuroscience, including the analysis of latent neural dynamics, neuroAI, and brain-machine interfaces. Standard approaches for comparing neural networks focus on the spatial geometry of latent states. Yet in recurrent networks, computations are implemented at the level of dynamics, and two networks performing the same computation with equivalent dynamics need not exhibit the same geometry. To bridge this gap, we introduce a novel similarity metric that compares two systems at the level of their dynamics, called Dynamical Similarity Analysis (DSA). Our method incorporates two components: Using recent advances in data-driven dynamical systems theory, we learn a high-dimensional linear system that accurately captures core features of the original nonlinear dynamics. Next, we compare different systems passed through this embedding using a novel extension of Procrustes Analysis that accounts for how vector fields change under orthogonal transformation. In four case studies, we demonstrate that our method disentangles dynamically equivalent and distinct recurrent neural networks (RNNs), while geometric methods fall short: (1) We identify a set of RNN architectures that when trained on the same task, are different geometrically but are dynamically equivalent. (2) A set of neural networks that are similar geometrically but are dynamically very different. (3) A ring attractor network whose geometry we can smoothly and substantially deform while preserving the attractor topology. (4) A line attractor network whose topology is transformed into a ring attractor by adding periodic boundary conditions. We additionally show that our method can distinguish learning rules in an unsupervised manner. Our method opens the door to comparative analyses of the essential temporal structure of computation in neural circuits.

T-24. A Biologically Inspired Neural Attention Model for the Analysis of Sequential Spiking Patterns

Nicolas Skatchkovsky[1,2] NICOLAS.SKATCHKOVSKY@CRICK.AC.UK
Alexander Egea Weiss[1] ALEXANDER.EGEA-WEISS@CRICK.AC.UK
Sadra Sadeh[3,4] S.SADEH@IMPERIAL.AC.UK
Maria Florencia Iacaruso[1] FLORENCIA.IACARUSO@CRICK.AC.UK

[1]The Francis Crick Institute
[2]Neuroscience
[3]Imperial College London
[4]Brain Sciences

Advances in experimental neuroscience are enabling scientists to record neuronal activity with higher spatial and temporal resolution, with the most recent techniques allowing for simultaneous recordings of hundreds to thousands of neurons over extended periods of time. However, interpreting this neural data in relation to behaviour remains challenging. Previous research suggests that neural information can be represented in low-dimensional manifolds, which has prompted a shift from studying relationships between neural and behavioural parameters to finding meaningful representations of neural data using machine learning (ML). Despite recent progress, existing ML techniques still often overlook the sequential structure inherent to neural data. To address this limitation, we propose a novel self-attention mechanism that leverages single spikes as tokens by incorporating the biological principle of Hebbian learning. The approach involves training a variational autoencoder in which the encoder comprises the proposed neural attention block. To enforce temporal coherence, our model is trained using an information-theoretic criterion inspired by predictive coding, which consists in maximizing the mutual information between current latent samples and future targets. The resulting model is a Sequential Predictive Autoencoder for the Representation of spiKing Signals (SPARKS), and offers versatile training, supporting both unsupervised learning similarly to traditional autoencoders, but also supervised learning using e.g. behavioural data. We demonstrate the approach can produce directly interpretable latent embeddings and offers state-of-the-art prediction capabilities on publicly available datasets. Notably, we can decode video frames from the Allen visual dataset using less than a tenth of the number of neurons required by previous methods. We also perform the reconstruction of images from naturalistic movie clips using our own recordings from the mouse visual cortex, obtaining high fidelity within and across days. Overall, combining ML models with biologically inspired mechanisms, SPARKS provides a promising solution for extracting meaningful information from neural data.

T-25. Graded representation of economic value across regions and projections of the frontal cortex

Antara Majumdar[1,2] ANTARA.MAJUMDAR@KELLOGG.OX.AC.UK
Caitlin Ashcroft[1] CAITLIN.ASHCROFT@ORIEL.OX.AC.UK
Matthias Fritsche[1] MATTHIAS.FRITSCHE@DPAG.OX.AC.UK
Lauren Strickland[1] LAUREN.STRICKLAND@DPAG.OX.AC.UK
Lukas Bijoch[1] L.BIJOCH@NENCKI.EDU.PL
Adam Packer[1] ADAM.PACKER@DPAG.OX.AC.UK
Simon Butt[1] SIMON.BUTT@DPAG.OX.AC.UK
Armin Lak[1] ARMIN.LAK@DPAG.OX.AC.UK

[1]University of Oxford
[2]Physiology, Anatomy and Genetics

Economic decision-making under risk - the process of selecting between options with different values and uncertain outcomes - concerns many aspects of our lives. Past studies have demonstrated representations of economic decision variables across various regions of the frontal cortex, however, with coarse spatial and temporal resolution (e.g. using fMRI) or in small neural populations using electrophysiology. Here, we sought to bridge this gap by investigating how fine grained neural signals from various frontal regions contribute to economic decisions using high-density large-scale electrophysiology across frontal cortical regions and their defined projection pathways. To this end, we devised a visual economic decision-making task in head-fixed mice, akin to those previously used in studies of non-human primates. In each trial, mice chose between two simultaneously presented abstract visual stimuli that differed in their magnitude of associated water reward, and reward probability, resulting in choice options with different expected values and risks. We found that mice's choices were sensitive to the expected value of stimuli: mice consistently selected stimuli with higher expected values. Moreover, we observed diverse risk attitudes across mice. We used high-density large-scale electrophysiological recordings to measure neural signals across many frontal regions during the task. Furthermore, we combined these recordings with optotagging to track neural signals in frontal cortical neurons projecting to specific downstream targets. Using population decoding, we show that neural signals across various frontal regions and projection-defined neuronal populations encode economic value, albeit with graded strength. Our work reveals graded representations of economic value across the frontal cortical regions and in their projection pathways, and provides a platform for investigating the neural basis of economic decision-making at a large scale with high spatial and temporal resolution.

T-26. Transitions in dynamical regime and neural mode underlie perceptual decision-making

Thomas Luo[1,2] ZHIHAOL@PRINCETON.EDU
Timothy Kim[1] TDKIM@PRINCETON.EDU
Diksha Gupta[3] DIKSHA.GUPTA@UCL.AC.UK
Adrian Bondy[1,2] ABONDY@PRINCETON.EDU
Charles Kopec[1,2] CKOPEC@PRINCETON.EDU
Verity Elliott[1] VELLIOTT@PRINCETON.EDU
Brian DePasquale[4,5] BDDEPASQ@BU.EDU
Carlos Brody[1,2] BRODY@PRINCETON.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute
[3]Sainsbury Wellcome Centre
[4]Boston University
[5]BME

Perceptual decision-making is the process by which an animal uses sensory stimuli to choose an action or mental proposition. This process is thought to be mediated by neurons organized as attractor networks. However, whether attractor dynamics underlie decision behavior and the complex neuronal responses remains unclear. Here we use an unsupervised, deep learning-based method to discover decision-related dynamics from the simultaneous activity of neurons in frontal cortex and striatum of rats while they accumulate pulsatile auditory evidence. We show that contrary to prevailing hypotheses, attractors play a role only after a transition from a regime in the dynamics that is strongly driven by inputs to one dominated by the intrinsic dynamics. The initial regime mediates evidence accumulation, and the subsequent intrinsic-dominant regime subserves decision commitment. This regime transition is coupled to a rapid reorganization in the representation of the decision process in the neural population (a change in the "neural mode" along which the process develops). A simplified model approximating the coupled transition in the dynamics and neural mode allows inferring, from each trial's neural activity, the internal decision commitment time in that trial, and captures diverse and complex single-neuron

temporal profiles, such as ramping and stepping. It also captures trial-averaged curved trajectories, and reveals distinctions between brain regions. Our results show that the formation of a perceptual choice involves a rapid, coordinated transition in both the dynamical regime and the neural mode of the decision process, and suggest pairing deep learning and parsimonious models as a promising approach for understanding complex data.

T-27. Understanding atypical decision making behavior with recurrent neural networks

Jin Zida[1] RIE.ACAD@GMAIL.COM
Li Ji-An[2] JIAN.LI.ACAD@GMAIL.COM
Marcelo Mattar[3,4] MARCELO.MATTAR@NYU.EDU

[1]University of Science and Technology of China
[2]University of California, San Diego
[3]New York University
[4]Psychology

A major goal of cognitive science is to characterize the cognitive processes underlying healthy and pathological decision making. A traditional approach involves developing cognitive models grounded in normative principles such as reinforcement learning and Bayesian inference. These models are typically composed of few interpretable parameters that are fit to a subject's decisions and then used to characterize the differences between individuals or between groups. Recently, artificial neural networks have emerged as an alternative modeling framework, enabling better predictions of behavior and requiring less domain-specific knowledge than classical cognitive models. However, neural networks are notoriously difficult to interpret, and therefore seen as inadequate for characterizing distinct patterns of decision making. Here, we leverage tiny recurrent neural networks (RNNs) and dynamical-systems visualization shown to facilitate the identification of interpretable cognitive strategies in reward-learning tasks (Ji-An et al. 2023). We used this framework to analyze human decisions in a two-armed bandit task with rare rewards, contrasting healthy, depression, and bipolar subjects. Our research uncovered various strategies that drive diverse, atypical behavioral patterns missed by classical cognitive modeling, such as a tendency to shift actions after a reward. Utilizing features from tiny RNNs, we designed a diagnostic procedure that predicts an individual pathological status based solely on their decisions, with accuracy comparable to those from larger, black-box RNNs. Overall, our findings demonstrate the significant role of tiny RNNs and dynamical-systems interpretability in understanding individual differences in computational psychiatry.

T-28. Mechanisms of brain-wide inter-area communication

Ulises Pereira[1] PIMUNTUE@GMAIL.COM
Sean Froudist-Walsh[2,3] SEAN.FROUDIST-WALSH@BRISTOL.AC.UK
Xiao-Jing Wang[4] XJWANG@NYU.EDU

[1]Allen Institute for Neural Dynamics
[2]University of Bristol
[3]Bristol Computational Neuroscience Unit
[4]New York University

For organizing brain-wide activity, local dynamics within brain areas need to be flexibly routed. Understanding the neural mechanisms underlying brain-wide inter-area communication remains a major challenge. Here we study a class of multi-regional recurrent neural networks to analyze inter-area communication at a brain-wide scale. We show that local dynamics can be routed through low-dimensional communication sub-spaces depending on the alignment of the long-range projections with the local circuit's connectivity. Spatial patterns of long-range projections can embed dynamically distinct functional sub-networks in the brain-wide dynamics. We show that dynamic routing of functional networks can be performed by subcortical control of trans-thalamic projections, neuromodulation, and top-down control. We apply our modeling framework to large-scale models of the mammalian cortex constrained by multimodal data recapitulating brain-wide electrophysiological and imaging studies, and provide testable experimental predictions. Our modeling framework provides new insights into how local and long-range connectivity interact to coordinate brain-wide communication during cognition.

T-29. Probing Movement Preparation Dynamics Using BCI-based Causal Perturbations.

Asma Motiwala[1] AMOTIWAL@ANDREW.CMU.EDU
Emily Oby[2] EMO22@PITT.EDU
Erinn Grigsby[2] ERINN.GRIGSBY@GMAIL.COM
Juliana Couras[2] JULIANA.COURAS@PITT.EDU
Alan Degenhart[2] ALANDEGENHART@GMAIL.COM
Aaron Batista[2] AARON.BATISTA@PITT.EDU
Byron Yu[1] BYRONYU@CMU.EDU

[1]Carnegie Mellon University
[2]University of Pittsburgh

Nearly all behaviors and cognitive processes involve dynamical processing, such as integrating and transforming information across time. We seek to interrogate the dynamical mechanisms that underlie movement preparation activity in the motor cortex using a novel perturbation paradigm that leverages animals' own volitional control of population activity in the motor cortex. We use a memory guided arm reaching task and use a brain-computer interface (BCI) to causally perturb population activity during the memory period. By instructing animals to volitionally modulate population activity during the memory period using BCIs, we perturb activity either towards or away from the preparatory state for a given arm movement. By leveraging volitional modulation of population activity, we ensure that our perturbation is within animals' neural repertoire where task-relevant dynamics unfold. We asked–to what extent and in what manner does volitional control of motor cortical activity during the delay affect arm movement preparation? Prior to the perturbation, population activity in the motor cortex evolves based on the instructed reach target. We find that our BCI based perturbation modulates population activity along dimensions related to arm movement preparation and interferes with animals' behavior in a spatially specific manner. When the perturbation requires animals to modulate their neural activity away from the preparatory state, they are more likely to reach to the wrong target after the perturbation. Surprisingly, we find that information about reach target direction is not preserved in the activity of this population during the BCI perturbation. After the perturbation population activity does not evolve deterministically as a function of population activity state at the end of the perturbation, but evolves as a function of the subsequent reach animals perform. This suggests that population dynamics for movement preparation are driven by inputs external to the motor cortex that signal the direction of animals' upcoming movement.

T-30. A combinatorial code of action and context for motor memory

Jae-Hyun Kim[1,2] RUMINSIGHT@GMAIL.COM
Kayvon Daie[3] KAYVON.DAIE@ALLENINSTITUTE.ORG
Nuo Li[1] NUO.LI@BCM.EDU

[1]Baylor College of Medicine
[2]Neuroscience
[3]Allen Institute for Neural Dynamics

Motor skill repertoire can be stably retained long-term, but the neural mechanism underlying stable memory storage remains poorly understood. Moreover, it is unknown how existing motor memories are maintained while new motor skills are continuously acquired. Here we tracked neural representation of learned actions throughout a significant portion of a mouse's lifespan. We found that learned actions are retained in motor memory in combination with context, which protects existing memories from erasure during new motor learning. We used automated home-cage training to establish a continual learning paradigm. Mice learned to perform directional licking in different tasks. In-cage optogenetics shows that learned directional licking is driven by preparatory activity in anterior lateral motor cortex (ALM). We used two-photon calcium imaging to chronically track ALM activities for multiple months. Within the same task context, the preparatory activity for directional licking exhibited little representational drift over time. As mice learned new task contexts, new preparatory activity emerged to drive the same licking actions, while activity related to sensory stimulus and movement execution remained stable. Interestingly, the previously acquired motor memories were retained: re-learning to make the same licking actions under the previous context re-activated the previous preparatory activity, even months later. Across multiple task learnings, distinct preparatory states were created to drive the same licking action in a context-dependent manner. Our results show that preparatory activity reflects motor memories that stably encode learned actions in combination with their context. A feedforward network model that stored sensorimotor combinations in high-dimensional hidden layers could explain multiple aspects of neural data. Context-specific memory, as we observed in the motor system, may provide a general solution for stable memory storage in the face of continual learning. Learning in new contexts produces parallel new representations instead of modifying existing representations, thus protecting existing motor memories from erasure.

T-31. Self-supervised behavior modeling with dense keypoint tracking

Ying Yu[1] YY334@UW.EDU
Jingyuan Li[1,2] JINGYLI6@UW.EDU
Kun Su[3] SUK4@UW.EDU
Anna Bowen[1,4] ABOWEN5@UW.EDU
Carlos Campos[3] CAMPOSCA@UW.EDU

[1]University of Washington, Seattle
[2]ECE
[3]University of Washington
[4]Biological Structure

Behavior is the final substrate of neural computation. Accordingly, rich descriptions of behavior are essential for brain decoding efforts, allowing us to understand action generation and distinguish between motor and cognitive processes. We hypothesized that leveraging a rich postural representation of the body would allow us to learn more precise, transferable representations of behavior from video. Here, we utilize a near-continuous 3D map of the mouse body consisting of over 30,000 points to extrapolate posture from 2D video data. Our goals were to optimally process this high-dimensional spatial data in both space and time to 1) identify discrete postures and 2) parse sequences of postures to segment behavior. To address these challenges we embed frame-wise spatial data in latent space and compress latent representations into posture tokens. Subsequently, posture tokens are used to train a transformer-based sequence prediction model. Initial efforts demonstrate this multistep approach is adept at learning temporal representations up to 4 s in duration, which enables one-shot behavioral sequence retrieval that bypasses the annotation and training steps previously required for supervised behavioral analysis. Our ongoing studies focus on segmenting behavior and making these resources readily available to support research in neuroscience and related fields.

T-32. Weight transfer in the reinforcement learning model of songbird learning

Khue Tran KTRAN@CSHL.EDU
Alexei Koulakov KOULAKOV@CSHL.EDU

Cold Spring Harbor Laboratory

Song learning behavior observed in the songbird system provides a notable example of learning through trial-and-error. We present a computational model of song learning that integrates reinforcement learning (RL) and unsupervised learning (UL) and agrees with known songbird circuitry. The song circuit outputs activity from nucleus RA, which receives two primary inputs: timing information from area HVC and stochastic activity from nucleus LMAN. Additionally, song learning relies on Area X, a basal ganglia area that receives dopaminergic inputs from VTA. In our model, song learning begins with the HVC-to-Area X connectivity, employing an RL mechanism that involves node perturbation. This acquired information is then consolidated in the HVC-to-RA weight matrix through a UL mechanism. The transfer of weights from Area X to RA takes place via the thalamus, utilizing a specific form of spike-timing-dependent plasticity. Thus, we present a computational model grounded in songbird circuitry in which the optimal policy is initially guided by reinforcement and subsequently transferred to another circuit through Hebbian plasticity.

T-33. Decoding Decision-Making in Fruit Flies: Examining Perseverance with a High-Throughput 2AFC Assay and a Computational Model of Mushroom Body Circuitry.

Rishika Mohanta[1] RMOHANTA@ROCKEFELLER.EDU
Aparna Dev[2] DEVA@JANELIA.HHMI.ORG
Glenn Turner[2] TURNERG@JANELIA.HHMI.ORG

[1]The Rockefeller University
[2]HHMI Janelia Research Campus

Animals navigating the world must frequently choose between different options, a task complicated by uncertainty and dynamic conditions over varying timescales. Adapting behavior based on past experiences is crucial, yet understanding the computational nature of these adaptations is an ongoing challenge. Fruit flies, with their well-documented neuronal connectivity and extensive genetic tools for circuit manipulation, offer a unique platform for investigating the circuit basis of decision-making. Recent findings suggest that fruit flies show operant matching

by predicting rewards through Dopaminergic neurons (DANs) in the Mushroom Body (MB), their primary learning center. However, to better understand how the MB's complex recurrent connectivity influences decision-making, examining a large number of choice trajectories against diverse reward schedules is essential. For this, we developed a high-throughput 2-alternative forced choice assay, allowing simultaneous monitoring of choice behaviors in 16 flies. This approach produced a substantial dataset from 360 flies, typically involving ~443 [203-891; 95% CI] sequential decision trials. Our analysis reveals a strong reliance on past choice-reward associations and a habit-like preference for recent choices over historically more rewarding ones, termed 'perseverance.' To explain this within the MB's neural circuitry, we formulated a computational model focusing on DAN-mediated synaptic depression between Kenyon Cells (KCs) and Mushroom Body Output Neurons (MBONs). This model includes aversive and appetitive MBON inputs integrated by previously identified downstream neurons. Crucially, we incorporated a feedback loop to appetitive DANs via feedback pathways observed in the connectome. Our model successfully replicates the perseverance behavior, and disrupting the feedback pathway results in the loss of this feature. This suggests that the feedback loop provides an effective reward-like signal in the appetitive DANs, reinforcing preference even without an external reward. In summary, our high-throughput assay reveals a habit-like element in Drosophila's choice behavior, potentially linked to MB's circuitry through a simple feedback loop.

T-34. Distance scaling and coherent motion allow NNs to match behavioral performance in electrolocation

Denis Turcu[1,2] DT2626@CUMC.COLUMBIA.EDU
Abigail Zadina[1] AZ2361@COLUMBIA.EDU
Nathaniel B. Sawtell[1] NS2635@COLUMBIA.EDU
Larry F Abbott[1] LFA2103@COLUMBIA.EDU

[1]Columbia University
[2]Neurobiology and Behavior

Mormyrid electric fish generate pulsed electric fields and use specialized mormyromast electroreceptors to identify nearby objects through the distortions they make in these fields. Object distortions are typically in the microvolt and sub-microsecond range (von der Emde 1992), yet the fish process these small perturbations to identify objects by determining their resistive and capacitive properties, allowing them to forage for preferred food. The amplitudes of the perturbations being used for this at distance d scale as 1/d^4 (Chen 2005). Nevertheless, freely swimming fish can detect objects up to 7cm away. Paradoxically, responses from single electrosensors can only be detected at distances less that 1cm (Pedraja 2020). We combined NN modeling, experimental data and an electric field model to account for this discrepancy. Our primary result is that an ANN readout can match the performance of the fish if 1) it is provided information about the scaling of electric signals, and 2) it makes use of coherent spatial cues provided by the motion of the fish. We used LFP data collected from projection layers of mormyromast afferents to extract temporal filters that characterize electroreceptor responses and developed a novel adaptive model of mormyromast sensory processing. We also constructed an electric field model to generate large simulated data sets for developing and analyzing NNs. We trained a variety of NN architectures on multiple electrolocation tasks and evaluated their performance. The best models could extract the spatial properties of a nearby object but not its electrical properties, but this problem was fixed by providing the NNs with information about how features scale with distance. Furthermore, we found that noise suppression due to multi-receptor averaging was not sufficient to explain the observed behavior. Instead, the use of coherence due to self motion could explain the discrepancy between the single-receptor and behavioral detection ranges.

T-35. Long timescales needed for memory tasks arise from distinct mechanisms shaped by learning curricula

Roxana Zeraati[1] ROXANA.ZERAATI@UNI-TUEBINGEN.DE
Sina Khajehabdollahi[2] SINA.ABDOLLAHI@GMAIL.COM
Emmanouil Giannakakis[3,4] GIANNAKAKISMANOS@GMAIL.COM
Tim Schafer[1] TIM.SCHAEFER@UNI-TUEBINGEN.DE
Georg Martius[5] GEORG.MARTIUS@TUEBINGEN.MPG.DE
Anna Levina[2] ANNA.LEVINA@UNI-TUEBINGEN.DE

[1]University of Tubingen, MPI for Biological Cybernetics
[2]University of Tubingen
[3]University of Tuebingen
[4]Computer Science
[5]University of Tubingen, MPI for Intelligent Systems

The brain solves complex tasks with intricate temporal dependencies by maintaining the memory of previous inputs over long periods. Long timescales required for solving such tasks may arise from the biophysical properties of individual neurons (single-neuron timescale, e.g., membrane time constant) or recurrent interactions among them. While both mechanisms operate in brain networks, their interplay and individual contributions to optimally solving memory-dependent tasks remain poorly understood.

We investigate the role of different mechanisms by training recurrent neural networks (RNNs) to solve N-parity and N-delayed match-to-sample (N-DMS) tasks with increasing memory requirements controlled by N. Networks are trained using two distinct curricula with gradually increasing N: (i) in single-N curriculum, networks learn a new N at each curriculum step; (ii) in multi-N curriculum, they learn a new N while maintaining the solutions for previous Ns, similar to biological learning. Each neuron has a leak parameter indicating the single-neuron timescale, optimized alongside recurrent weights. We estimate the network-mediated timescales from the autocorrelation decay of each neuron's activity. We find that in both curricula, RNNs develop longer timescales with increasing N, but via distinct mechanisms. Single-N RNNs operate in a strong inhibitory state and mainly rely on increasing their single-neuron timescales with N. However, multi-N RNNs operate closer to a balanced state and use only recurrent connectivity to develop long timescales, while keeping their single-neuron timescales constant. The latter is compatible with findings in primate cortex. We show that using network-mediated mechanisms to develop long timescales, as in multi-N RNNs, increases training speed and stability to perturbations, and allows generalization to tasks beyond the training set. Our results suggest that adapting timescales to task requirements via recurrent connectivity enables learning more complex objectives (holding multiple concurrent memories) and improves computational robustness, which can be a beneficial strategy for implementing brain computations.

T-36. Deep neural networks reveal context-sensitive speech encoding in single neurons of human cortex

Shailee Jain[1,2] SHAILEE.JAIN@UCSF.EDU
Matthew K. Leonard[1] MATTHEW.LEONARD@UCSF.EDU
Edward F. Chang[1] EDWARD.CHANG@UCSF.EDU

[1]University of California, San Francisco
[2]Neurological surgery

Speech is a dynamic acoustic signal that requires listeners to continuously extract and integrate information at multiple timescales. Local neural populations and single neurons in human superior temporal gyrus (STG) encode many different types of information in speech, including acoustic-phonetic features, prosodic cues like pitch, and word-level surprisal. Although these features explain significant variance in neuronal firing, they do not reflect continuous, context-sensitive computations. Here, we used deep-neural-networks (DNNs) to investigate contextual processing in hundreds of STG neurons recorded from all layers of human cortex with Neuropixels probes. Specifically, we built encoding models to predict neuronal firing from hidden states of speech DNNs (HuBERT, WavLM) that can capture contextual, multi-timescale information. To investigate the degree of context sensitivity in different neurons, we varied the amount of prior context available to the DNN (20-1000ms) and DNN layer from which states were extracted. Overall, DNNs better predicted neuronal firing in 73% neurons compared to previously used context-agnostic features (3.5$\pm$5% more variance explained). We used linear dimensionality reduction methods to identify encoding patterns across the neural population, finding that the primary source of variation was sensitivity to the amount of prior context incorporated in DNN states. In some neurons, encoding performance improved with upto ~500ms of prior speech, while others were best predicted by shorter context windows. Different recording sites (putative cortical columns) could be characterized by their context sensitivity, independent from spectrotemporal tuning, DNN layer selectivity, spiking patterns or cortical depth. Finally, we used population-level decoding to investigate the role of context in spectrotemporal representations, and found that neurons and columns with long context sensitivity faithfully represented speech over timescales consistent with higher-order word and phrase-level information (~1sec). Together, these results demonstrate that a defining property of neurons across the depth of high-level speech cortex is the ability to perform context-sensitive computations.

T-37. Neocortical long-range inhibitory neurons coordinate state-dependent network synchronization and promote sleep

Jacob Ratliff[1,2] JACOB.RATLIFF@EINSTEINMED.EDU
Renata Batista-Brito[1,3] RENATA.BRITO@EINSTEINMED.EDU
Geoffrey Terral[1] GEOFFREY.TERRAL@EINSTEINMED.EDU
Julie Mota[1] JULIE.MOTA@EINSTEINMED.EDU
Charu Ramakrishnan[4] CHARUR@STANFORD.EDU
Leif E Fenno[5] LIEF.FENNO@AUSTIN.UTEXAS.EDU
Karl Deisseroth[4] DEISSERO@STANFORD.EDU
Thomas Kilduff[6] THOMAS.KILDUFF@SRI.COM

[1]Albert Einstein College of Medicine
[2]Dominic P Purpura Department of Neuroscience
[3]Neuroscience
[4]Stanford University
[5]The University of Texas Austin
[6]SRI International

Cortical brain states are typically associated with changes in network synchronization and are key determinants for behavioral responses, such as sleep and wake. Though the mechanisms enabling these distinct modes of cortical operation remain largely unknown, inhibitory neurons (INs) have been suggested repeatedly as a regulator of behavioral-state dependent neocortical activity. Here we investigate how a unique subpopulation of neocortical long-range projecting INs impact cortical states. These cells are defined by the co-expression of somatostatin (SST) and neuronal nitric oxide synthase (nNOS), namely SST/nNOS cells. There deep evolutionary conservation, from amphibians to humans, despite being a small minority of INs, in addition to their unique morphology with dense arborization and cross area projections suggest an important role for SST/nNOS cells in the regulation of neocortical activity. Using 2-photon calcium imaging with high precision state monitoring, we find that SST/nNOS cells are active during low-arousal states, specifically NREM sleep and quiet wake, characterized by synchronized local field potentials. By optogenetically manipulating SST/nNOS cells, we show that the activity of SST/nNOS cells is sufficient to induce a synchronized local network state, with both increases in low-frequency LFP power and increases in synchronous spiking, regardless of the global behavioral state of the animal. We find that this synchronous spiking propagates across cortical space mirroring the arborization patterns of SST/nNOS cells. Lastly, we find that by pan-cortically activating SST/nNOS cells using chemogenetics, sparing subcortical cells, we can promote sleep both increasing sleep time and decreasing latency to fall asleep. Taken together, our data suggests that SST/nNOS cells read out the more broadly generated behavioral state of the animal and act as effectors of cortical state change. Further this neocortical cell type is also able to influence the more broadly generated behavioral state, promoting sleep, and challenging the dogma of solely subcortical sleep generation.

T-38. Back to the present: self-supervised learning in neocortical microcircuits

Kevin Kermani Nejad[1] KEVIN.KERMANINEJAD@DPAG.OX.AC.UK
Paul Anastasiades[2] PAUL.ANASTASIADES@BRISTOL.AC.UK
Loreen Hertag[3] LOREEN.HERTAEG@TU-BERLIN.DE
Rui Ponte Costa[1] RUI.COSTA@DPAG.OX.AC.UK

[1]University of Oxford
[2]University of Bristol
[3]Berlin Institute of Technology

The brain is believed to predict incoming sensory input, yet the underlying mechanisms remain unclear. Leveraging recent experimental findings and self-supervised algorithms, we propose a model where Layer-2/3 (L2/3) learns to predict sensory inputs by incorporating previous inputs from Layer-4 (L4) and comparing its prediction with current inputs received by Layer-5 (L5). First, we tested the model on temporal sensory prediction tasks thus demonstrating L2/3's ability to predict future input. Moreover, we show that L2/3 predictions result in sensory representations that are robustness to sensory noise and occlusions. Next, using a sensorimotor task, we show that prediction errors in our model can explain mismatch responses observed in both L2/3 and L5 in behaving animals, while making new experimentally testable predictions. In summary, our findings suggest that the classical L4 -> L2/3 -> L5 motif implements temporal self-supervised learning in the brain.

T-39. Comparing the geometry of static and dynamic coding across the auditory pathway.

Othman Lahrach[1,2] OTHMAN.LAHRACH@ENS.PSL.EU
Sophie Bagure[3] SOPHIE.BAGUR@PASTEUR.FR
Brice Bathellier[4] BRICE.BATHELLIER@CNRS.FR
Srdjan Ostojic[5] SRDJAN.OSTOJIC@ENS.PSL.EU

[1]Ecole Normale Superieure
[2]Cognitive Science
[3]Institut Pasteur
[4]Institut de l'Audition
[5]ecole Normale Superieure

Neural processing across sensory areas is often conceptualised as a series of feed-forward steps analogous to successive layers in a deep network. Under that view, neural coding is approached as a static phenomenon, where the code is studied with disregard to its temporal component. In contrast, the role of temporal dynamics generated by recurrent connectivity is only starting to be explored. How do network dynamics shape the geometry of sensory representations? We examined large-scale calcium-imaging and electrophysiological recordings performed across the mouse inferior colliculus (IC), auditory thalamus (Th) and auditory cortex (AC) in response to 140 auditory stimuli. We analyzed the geometry of responses to tonic sounds, either from the static point of view or by focusing on dynamics. The static code approach examines the manifold of time-averaged responses as captured by a non-linear feedforward model. The dynamic approach considers manifolds at different time periods, and compares them to feed-forward and recurrent network models fitted to neural trajectories. We then examine how well these models predict the responses to temporally modulated sounds. We found that simple feed-forward models described well the manifolds of time-averaged responses across all areas. Examining the dynamics however showed that the response manifold changed significantly over time in the auditory cortex, but not in upstream areas. Manifolds at the onset and offset of sounds are orthogonal in AC but not in IC, and mildly in Th. Recurrent Neural Networks (RNN) suggest that this orthogonality can emerge from recurrent connectivity. We show that this cortical effect is crucial for describing responses to temporally modulated sounds, as it provides a mechanism that breaks temporal symmetry in the responses to time-reversed sounds. This paves the way for a transformation from a simple sequence code to a population code that represents the direction of change of sound properties over time.

T-40. Representational geometry of hierarchical category structures in the monkey inferotemporal cortex

Nga Yu Lo[1] NYULO@FLATIRONINSTITUTE.ORG
Roozbeh Kiani[2] RK97@NYU.EDU
SueYeon Chung[3,4] SCHUNG@FLATIRONINSTITUTE.ORG

[1]Flatiron Institute
[2]New York University
[3]New York University; Flatiron Institute
[4]Center for Computational Neuroscience

Psychological theories suggest that humans categorize objects hierarchically based on abstraction of common and shared attributes. Prior studies have also found evidence for hierarchical representation of visual categories through clustering and representational similarity analyses of neural activity in the monkey and human ventral visual pathway. However, the format and organization of these hierarchical representations remain unknown. We address this question by extending the manifold capacity theory, a recent advance that builds on machine learning and statistical physics to connect the geometry of neural representations with their ability to support classification. We apply our analytical approach to the responses of monkey inferotemporal neurons to a large stimulus set selected from a hierarchy of object categories. A neural manifold is defined as the aggregated neural responses to each category. Quantifying changes in manifold geometry across different levels of the hierarchy, we discover a gradual reduction of distances between manifold centers from the lower levels of the hierarchy to the higher levels, accompanied by minimal changes in the width of manifolds. The combination results in reduced separability of superordinate categories (reduced capacity) compared to subordinate and entry level categories, matching behavioral studies. We further probe the representational geometry of manifolds by manipulating the manifold shape and orientation, and exploring the impact of different manipulations on the manifold capacity. Our results suggest manifold orientation and overlap attribute to the trends we observe, and hierarchical representation of categories stems largely from judicious spacing of their manifolds.

T-41. Rapid, concerted switching of the neural code in inferotemporal cortex

Yuelin Shi[1] YSSHI@CALTECH.EDU
Dasheng Bi[2] DBI@BERKELEY.EDU
Janis Hesse[2] JANISHESSE@GOOGLEMAIL.COM
Frank Lanfranchi[3] FLANFRAN@CALTECH.EDU
Shi Chen[2] SHICHEN@BERKELEY.EDU
Doris Tsao[4] DORTSAO@BERKELEY.EDU

[1]California Institute of Technology
[2]University of California, Berkeley
[3]California Institute of Technology; University of California, Berkeley
[4]University of California, Berkeley; Howard Hughes Medical Institute

A fundamental paradigm in neuroscience is the concept of neural coding through tuning functions. According to this idea, neurons encode stimuli through fixed mappings of stimulus features to firing rates. Here, we report that the tuning of visual neurons can rapidly and coherently change across a population to attend to a whole and its parts. We measured responses of cells in macaque inferotemporal face patches to faces and objects and used a deep neural network to generate two feature spaces: a general object space capturing face and object variations and a face space capturing face variations. We found that initially, cells used a common encoding axis for faces and objects; this axis pointed towards the direction of faces in the object space, enabling robust face detection. However, at ~100 ms, a sudden and drastic switch occurred in the neural code for faces. The encoding axis for faces in the object space completely reversed direction, and robust new tuning developed along multiple dimensions of the face space, supporting fine face discrimination. These dynamics were face specific and did not occur in response to objects. Overall, these results show that face cells achieve domain specificity over time through a rapid, concerted change in the neural code they use to represent faces. Our findings conclusively resolve a debate that has raged for decades concerning whether face processing is domain specific. They also strongly challenge the notion that object recognition can be explained by largely feedforward processes, revealing an essential role for recurrent dynamics. More broadly, our results suggest a novel mechanism for neural representation: concerted, stimulus-driven switching of the neural code used by a cortical area.

T-42. Emergence of state modulation in a developing cortical circuit

Jure Majnik[1,2] MAJNIK.JURE@GMAIL.COM
Sofia Zangila[3] SZAGGILA@HOTMAIL.COM
Rosa Cossart[3] ROSA.COSSART@INSERM.FR
Jean-Claude Platel[3] JEAN-CLAUDE.PLATEL@INSERM.FR

[1]University of Aix-Marseille
[2]INSERM, INMED (Mediterranean Institute of Neurobiology)
[3]Mediterranean Institute of Neurobiology (INMED)

Behavioral state-dependence is a hallmark of neural dynamics in adult neocortical circuits. Conversely, during development, most activity is thought to be internally generated and thus independent of behavior. To bridge this dichotomy, we performed longitudinal recordings tracking neural activity from the same cells throughout postnatal development in mice. Firstly, characterizing the changes to single cell and population dynamics showed a gradual maturation of the circuit towards adult-like activity patterns. Secondly, applying decoding analysis to predict global movements, we observed a striking rise in state-dependence of neural responses during the second postnatal week. Once developed, the state representation is stable, with the same neurons exhibiting state-coupling, allowing for cross-day decoding. Finally, we show that the optimal movement decoding axis gradually aligns with the principal axis of neural variability during the course of development, providing a direct link to observations in adult mice. Our work thus provides an important insight into the development and maturation of state-dependent neocortical computation.

1-001. A consistent map in the medial entorhinal cortex supports spatial memory

Yi Gu YI.GU@NIH.GOV

National Institute of Neurological Disorders and Stroke

The medial entorhinal cortex (MEC) is hypothesized to function as a “cognitive map” for memory-guided navigation. However, how this map develops during learning and influences memory remains unclear. To answer this

question, we imaged MEC calcium dynamics while mice learned a novel virtual linear track with a fixed reward location over ten days. The mice exhibited different levels of spatial learning based on their reward-predictive behaviors. In the “good performer” mice with successful learning, MEC calcium dynamics gradually became spatially consistent and then stabilized. The MEC also showed an increased c-Fos expression in novel environments, suggesting the induction of synaptic plasticity that shaped the consistent dynamics. In contrast, in the “poor performer” mice with unsuccessful learning, the MEC neural dynamics showed lower spatial consistency that did not improve during learning. c-Fos expression in the MEC was high in familiar environments and did not increase in novel environments, indicating a hyperactive network with impaired synaptic plasticity. Furthermore, the difference in MEC spatial activity consistency between the good and poor performers was larger in track areas around cues and immediately before the reward, suggesting that consistent activity in these areas supports effective spatial memory. Indeed, optogenetic stimulation of the MEC in the good performers in a spatially random pattern in cue areas disrupted their reward-predictive behaviors in a learned environment, supporting the necessity of consistent activity at cues for spatial memory. Moreover, optogenetic stimulation of the MEC in the poor performers in a spatially consistent pattern immediately before the reward improved their behaviors in the environment, indicating the sufficiency of consistent activity before the reward for spatial memory. Thus, we demonstrated that the establishment of a spatially consistent MEC map during learning both correlates to, and is causally associated with, successful spatial memory.

1-002. Duality of Bures and Shape Distances with Implications for Comparing Neural Representations

Sarah Harvey[1,2] SHARVEY@FLATIRONINSTITUTE.ORG
Brett Larsen[3] BRETTLARSEN@FLATIRONINSTITUTE.ORG
Alex Williams[4] ALEX.H.WILLIAMS@NYU.EDU

[1]Flatiron Institute
[2]Center for Computational Neuroscience
[3]New York University; Flatiron Institute
[4]New York University

How should neuroscientists mathematically evaluate whether two individuals or networks have similar neural representations? This question of quantifying representational similarity between high-dimensional neural datasets—e.g. multi-neuron spike train recordings, hidden layer activation patterns in recurrent and feedforward artificial network models, or BOLD signals across fMRI voxels—is now an active research direction in computational neuroscience, with implications for understanding variability in neural computations across individuals and species, as well as for comparisons between biological systems and model circuits. Excitement around these applications has sparked a proliferation of approaches including Representational Similarity Analysis (RSA), Canonical Correlation Analysis (CCA), linear regression, and Centered Kernel Alignment (CKA), and a body of empirical work comparing different methods. The result is a rich but confusing landscape of (dis)similarity measures guided by intuition, and an insufficient theoretical understanding of their relationships.

This work provides a unifying theoretical account of several of these methods. We prove that two existing and independently proposed (dis)similarity measures—Riemannian shape distance and Normalized Bures Similarity (NBS)—are equivalent. This is significant because the mathematical definitions of each method appear quite different: shape distance involves optimizing an explicit alignment mapping between two neural firing rate spaces (similar to CCA and linear regression), while NBS is a measure of consistency between stimulus-by-stimulus similarity matrices (similar to RSA and CKA). Thus, we uncover a bridge between two major categories of methods. To further relate existing approaches, we derive upper and lower bounds for CKA—a popular approach in the machine learning literature, which is closely related to RSA—in terms of NBS. By revealing these connections between competing similarity measures, we consolidate the menu of options that users face, and develop a foundation to theoretically characterize how these methods behave in experiments with finite durations, noisy responses, and limited neural measurements.

1-003. Network of biologically plausible neuron models can solve motor tasks through heterogeneity

Yigit Demirag[1] YIGIT@INI.ETHZ.CH
Giacomo Indiveri[2] GIACOMO@INI.UZH.CH

[1]Institute of Neuroinformatics, University of Zurich and ETH Zurich
[2]Institute of Neuroinformatics, University of Zurich, ETH Zurich

Unlike neuron models in artificial neural networks (ANNs), biologically plausible spiking neuron models exhibit rich

temporal dynamics. However, the computational implications of these dynamics and their operational advantages in recurrent neural networks are still underexplored. Previous research on weight agnostic neural networks (Gaier and Ha, 2019) has shown that using varied activation functions (e.g., relu, tanh, gaussian) in feedforward ANNs, optimized via neural architecture search, can address reinforcement learning (RL) tasks without training network weights. Inspired by this, we explored the computational expressivity of common spiking neuron models, i.e., AdExp (Gerstner and Brette, 2009), Izhikevich (Izhikevich, 2013) and Leaky-Integrate and Fire (LIF) by optimizing only the neuron parameters in a recurrent neural network while keeping the weights randomly initialized and fixed throughout the training. We found that networks of biologically-plausible neuron models train suboptimally with backpropagation through time (BPTT) due to the vanishing gradient problem of long sequences. To mitigate, we applied a gradientfree optimization, Evolutionary Strategies (ES) (Salimans et al., 2017), to exclusively train the neuron parameters. The networks interface with the environment through a dynamic and trainable encoding layer that transforms real-valued multimodal temporal sensory signals from a physics simulator into spike trains. Through evolution of generations, we confirmed that biologically-plausible heterogeneous neuron models can be potentially as effective as network weights, especially in simple task environments. We observed the best-performing networks comprise clusters of neurons exhibiting rich characteristics, such as tonic firing or bursting. We showed that evolution can find neuron parameters that are robust to perturbations of the synaptic weights. The effectiveness of biologically plausible heterogeneous neuron models motivates the search for plasticity rules in neuron parameters as well as synapses, and enables more scalable neuromorphic hardware with orders of magnitude less trainable parameters while maintaining expressivity.

1-004. Whole-brain fMRI reveals the notes, chords, and conductors of the cortical orchestra

Brandon Munn[1,2] BRANDON.MUNN@SYDNEY.EDU.AU
Eli Muller[1] ELI.MULLER@SYDNEY.EDU.AU
Mac Shine[1,3] MAC.SHINE@SYDNEY.EDU.AU

[1]The University of Sydney
[2]Faculty of Medicine and Health
[3]Brain and Mind Center

The brain produces rich and dynamic patterns of neural activity spanning multiple spatial and temporal scales over seconds to minutes and in circuits and systems to enable a broad spectrum of complex behaviours. How does the brain instantiate this wide variety of spatiotemporal dynamics? A prevailing assumption in the field, which is currently undergoing revision, is that these features arise due primarily to the organisation of and the interactions within the cerebral cortex1. In this study, we demonstrate that many features of coordinated neural activity measured by standard whole-brain functional magnetic resonance imaging (fMRI) approaches are inherently linked to the organisation of the subcortex (e.g., thalamus, neuromodulatory brainstem structures, cerebellum – Fig. 1a). Leveraging a geometrically-based eigenmode decomposition of (i.e., the “notes” of the brain2; Fig. 1b) and a temporal dimensionality reduction technique, identifying time-lagged independent components (i.e., the “chords”, which are different combinations of notes – Fig. 1c), we establish a connection between slow subcortical fMRI temporal dynamics with spatially coarse cortical patterns and vice versa for fine spatial patterns. Our methodology reveals that activation in subcortical structures significantly precedes resting-state component time series in the cerebral cortex (explained variance > 0.6). These patterns further explain the well-known alternating oscillation in dorsal attention and default mode networks (Fig. 1d). In particular, we focus on the anatomical role of these systems, as we find that diffuse projecting thalamic and neuromodulatory systems preferentially explain slow and coarse spatiotemporal patterns. This extension can be seen as a whole-brain extension to previous theoretical3–5 and empirical6–8 analyses focusing on one subcortical structure at a time. In particular, we confirm that the independent spatiotemporal whole-brain modes typically involve a coordinated orchestra of subcortical structures.

1-005. Influenceability and predictability of C. elegans action selection through closed-loop interrogation

Raymond Dunn[1,2] RAYMOND.DUNN@UCSF.EDU
Jackson Borchardt[1,2] JACKSON.BORCHARDT2@UCSF.EDU
Grace Chiu[1] GRACE.CHIU@UCSF.EDU
Julia Miller[1] JULIA.MILLER@UCSF.EDU
Noelle L’Etoile[1] NOELLE.LETOILE@UCSF.EDU
Saul Kato[1] SAUL.KATO@UCSF.EDU

[1]University of California, San Francisco
[2]Neurology

Brains are active nonlinear control systems, which integrate time-varying sensory input with evolving internal state to produce effective continuous behavioral output. Here, we probe the neural basis of motor command sequence coordination in C. elegans towards understanding how it implements contingent (i.e. influenceable) action selection. While crawling on agar in search of food, worms execute a stereotyped behavioral sequence, crawling forwards and backwards with interspersed dorsal or ventral (D/V) turns. Whole-brain calcium imaging and linear dimensionality reduction has shown that high-level locomotory states correspond to fixed sub-volumes of neural state space, which segment a global attractor [1]. We examined influenceability of brain state trajectory by state-triggered optogenetic activation of single neurons during whole-brain imaging to determine neural mechanisms of (a) duration of reversal and (b) branch selection of post-reversal D/V turn. We find brief stimulations of specific neurons halt ongoing reversal in a seemingly probabilistic manner, evoking D/V turns. We fit predictive models on pre-stimulus neural activity to predict animal response to stimulation, and vary what neurons, embeddings, or time-history the model is fit on. We find that the stochasticity of reversal halting response disappears and the low-dimensional brain-wide neural state at the time of stimulation determines the response of the animal. Moreover, we find oscillatory-like subnetwork dynamics, not adequately captured by global embeddings like PCA, which are highly predictive of post-reversal turn branch selection. Our models correctly predict about 80% of individual trials for both reversal halting and D/V turn selection. Fascinatingly, turn selection is most predictable ~5s prior to perturbation. We further evaluate predictive properties of neural dynamics with nonlinear and dynamical models. More broadly, this work suggests contingent actions are the consequence of temporal sequencing of multiscale dynamics, which serve as computational scaffold for sensory input, in biological neural networks.

1-006. Human Action-Outcome Inference through Weighted Evidence Accumulation with Subjective Uncertainty

Naoyuki Okamoto[1,2] OKAMOTO.NAOYUKI.35R@ST.KYOTO-U.AC.JP
Shin Ishii[1] ISHII@I.KYOTO-U.AC.JP
Benedetto De Martino[3] BENEDETTODEMARTINO@GMAIL.COM
Aurelio Cortese[4] CORTESE_A@ATR.JP

[1]Kyoto University
[2]Graduate School of Informatics
[3]University College London, UK
[4]Advanced Telecommunications Research Institute International, Japan

A fundamental problem in learning and decision-making is for an agent to determine whether outcomes are worth considering (i.e., resulting from their actions vs randomness). However, it remains unclear how people can do this efficiently when multiple sources of errors/uncertainty coexist. We designed a task where uncertainty arises from environmental causes and motor variability. Participants had to hit a target appearing and disappearing quickly on a touchscreen, obtaining a binary score ('good'/'bad') on every trial. The relationship between their hitting location and the displayed score depended on two alternating hidden task states ('skill': the score depended on the hit location, or 'random': the score was random). In each trial, participants inferred the state of the game and rated their confidence about the inference. We first confirmed that participants adapted their choices to task state switches while also showing low confidence in those trials. Next, we hypothesised that participants would infer the current state by accumulating task state evidence while also considering the evidence's reliability. Our model-based analyses showed that, while participants used previous outcomes to estimate the current state, they accumulated evidence with different strengths based on the scores received. This effect mapped to the participants' overall difference in performance and confidence for the two outcomes. These results suggest that participants assigned reliability to their actions' consequences depending on whether the trial was subjectively easy or difficult to estimate. Further, participants' trial-by-trial confidence was strongly related to the probability the model outputted, meaning confidence increased as the model's internal evidence accumulated towards either state. The reliability difference and confidence data suggest that participants evaluate their motion error when making the inference (i.e., how accurate they assume their hitting ability is). Our findings shed light on the computational mechanisms of adaptive decisions and inference by considering action-based interactions in noisy environments.

1-007. Learned navigation strategies rely on distributed neural computation in C. elegans

Kevin Chen KSCHEN@PRINCETON.EDU
Jonathan Pillow PILLOW@PRINCETON.EDU
Andrew Leifer LEIFER@PRINCETON.EDU

Princeton University

Learning plays a fundamental role in naturalistic behaviors such as sensory-guided navigation. However, the field currently lacks rigorous measurements and models of this complex behavior and its relationship to neural mechanisms. To fill this gap, we develop the first quantitative model of learned navigation behavior in the nematode C. elegans. This worm has a compact nervous system, tractable behavior, and can associate odor with food (appetitive) or starvation (aversive). We study the behavioral strategies and neural computation underlying learned navigation by measuring, modeling, and perturbing the brain's transformation from olfactory signal into navigational motor output. We design a novel apparatus to precisely record animal's behavior while monitoring the odor environment it experiences. We develop a new statistical model that describes how worms dynamically alternate between two known strategies: (1) biased random walk (BRW), in which the probability of a sharp turn is modulated by a change in odor concentration and (2) weathervaning (WV), in which the animal steers based on the odor gradient. By applying this model to the measured time-varying odor navigation trajectories, we find that appetitive learning up-regulates BRW and aversive learning down-regulates WV. The model can be used to decode the animal's learned experience at > 90% accuracy, outperforming traditional metrics. To probe neural mechanisms, we optogenetically perturb the odor-sensing neuron AWC during navigation. Both measurement and model show that the worm's response to AWC perturbation reflects the learned experience. Furthermore, we investigate the contribution to learned odor navigation of interneurons downstream from AWC by genetically disrupting individual neurons. In contrast to previous work suggesting that learning alters single neuron, we find that navigation strategies are distributed in the network and different neuron's contributions are altered depending on learning. Together, we present a flexible navigation algorithm that is supported by distributed neural computation in a compact brain.

1-008. Local lateral connectivity is sufficient for replicating cortex-like topographical organization in deep neural networks

Xinyu Qian[1] XINYU.QIAN@MAIL.MCGILL.CA
Amirozhan Dehghani[1] AMIROZHAN.DEHGHANI@MAIL.MCGILL.CA
Asa Borzabadi Farahani[1] ASA.BORZABADIFARAHANI@MAIL.MCGILL.CA
Pouya Bashivan[2] POUYA.BASHIVAN@MCGILL.CA

[1]McGill University
[2]Mila, McGill University

Across the primate cortex, neurons that perform similar functions tend to be spatially grouped together. In the high-level visual cortex, this principle manifests as distinct cortical patches of category-selective neurons. Recent advances have shown that unit activity in Artificial Neural Networks (ANNs) exhibits a remarkable resemblance to brain activity observed in human and non-human primates. However, despite these similarities, ANNs fail to replicate the systematic arrangement of neurons across the cortex. Recent advancements have aimed to incorporate topographical organization into ANNs, but these models face several critical shortcomings, including performance gaps, limited scope, segregated representation and topography learning, and increased biological unrealism. Our research presents a novel topographical neural network model that addresses these flaws. Building on the well-established concept of local lateral connectivity (LLC), prevalent in the cortex of species with pronounced topographical organization, we integrated LLC into the layers of a convolutional neural network. We found that this method not only facilitates the learning of robust representations but also engenders a naturally emerging smooth topographic organization as an inherent byproduct of computation with this inductive bias. More specifically, our model: 1) replicates the arrangement of neurons in the early visual cortex, aligning with orientation, spatial frequency, and color; 2) forms object-selective clusters in deep layers analogous to those in the human inferotemporal cortex; 3) predicts unit selectivity in inferotemporal regions with previously uncharted object-selectivity; 4) improves the trade-off between object recognition performance and cortex-like topography compared to previous models and; 5) demonstrates notable improvement in robustness against adversarial noise, suggesting a potential functional role for LLC in learning robust representations. Our results demonstrate that incorporating local lateral connectivity in ANN models is sufficient for reproducing cortex-like topographical organization in these models, importantly, without the need for any topography-inducing learning objectives or learning rules.

1-009. Dendritic properties shape the responses of collicular wide-field neurons to behaviorally relevant stimuli

Norma Kuhn[1] NORMA.KUEHN@GMAIL.COM
Bram Nuttin[2,3] BRAM.NUTTIN@NERF.BE
Karl Farrow[2] KARL.FARROW@NERF.BE
Chen Li[4] CHEN.LI.CL2538@YALE.EDU
Natalia Baimacheva[5] NBAYMACHEVA@GMAIL.COM
Katja Reinhard[6] KATJA.REINHARD@SISSA.IT
Vincent Bonin[1] VINCENT.BONIN@NERF.BE

[1]NERF
[2]Neuro-Electronics Research Flanders (Imec, KU Leuven, VIB)
[3]Biology
[4]Yale School of Medicine
[5]ITMO University
[6]Sissa

Delineating brain circuitry is the first step to understanding brain function. The next is to decipher the rules by which neurons combine their several thousand inputs to generate the right output. To tackle this, we focused on wide-field neurons of the mouse superior colliculus. These neurons, associated with innate defensive behaviors to predator-like visual stimuli, can be genetically targeted, and receive direct input from the retina. To understand how these neurons filter their inputs to guide behavior, we used a three-step approach. First, we measured their responses to relevant visual cues in dendrites and cell bodies. Second, we examined inputs from the retina and collicular inhibitory interneurons using a combination of transsynaptic viral tracing and two-photon calcium imaging. Third, to replicate wide-field neuron responses and identify important physiological parameters, we tested different computational models. We made three important findings: First, twelve retinal ganglion cell types provide layered input to wide-field neuron dendrites. Second, wide-field neuron cell bodies, known to respond to expanding discs, turn out to also respond strongly to shrinking discs and cluster into three types based on contrast preference: On, Off, and On-Off. While an expanding disc is commonly associated with looming overhead predators, a shrinking disc might emulate receding prey, in line with the neurons' role in prey detection (Hoy et al. 2019). Third, a linear-nonlinear point neuron model is not sufficient to capture wide-field neuron responses to the "receding prey" stimulus. Instead, the location of inputs along wide-field neuron dendrites and dendritic filtering properties need to be explicitly modeled to accurately capture wide-field cell body responses. This demonstrates that if we neglect neural dendrites, responses to important stimuli will be missed and a neuron's function within a circuit might be misinterpreted. Hence, accurately describing a neuron's input-output function is essential to understanding brain function.

1-010. The role of oscillations in generating toroidal topology in grid cell ensemble

Giovanni di Sarra[1] G.DISARRA@GMAIL.COM
Siddharth Jha[2] SIDDHARTHJHA98@G.UCLA.EDU
Mayank Mehta[2] MAYANKMEHTA@UCLA.EDU
Yasser Roudi[1] YASSERROUDI@GMAIL.COM

[1]Kavli Institute for Systems Neuroscience
[2]UCLA

Topological data analysis shows that the activity of grid cell populations lies in a toroidal manifold. However, the underlying mechanism of this striking topology and its precise relationship to the spatio-temporal properties of neurons is unclear. Analyzing both real data together with a simplified model, here we show that neural oscillations can play a key role in generating the observed toroidal topology. We first define the degree of toroidal topology of the data as the normalized bottleneck distance between persistent barcodes from the data and a perfect torus. We find that slight perturbations of the spike times that have little impact on the grid score destroy the toroidal topology as quantified in this way. This leads to the hypothesis that the temporal structures present in the spiking of neurons play an important role and that the hexagonal spatial selectivity is not sufficient on its own. We confirmed this hypothesis using a simplified model for the activity of grid cells as follows. We modeled the rate maps of the ensemble of grid cells using independent, rate-modulated poisson processes. Although a population of such neurons with hexagonal-arranged firing does not express a toroidal topology, when the rates are also modulated by temporal oscillations, the degree of toroidal topology substantially increases. Specifically, adding theta modulation significantly increases the chance of observing the toroidal topology, and adding more oscillations increases this to a hundred percent. Finally, we show how adding spatial noise to the positioning of the fields even with oscillations destroys the torus. We thus conclude that a combination of oscillatory modulations of

the single-neuron firing rate, together with the repetitive position of the fields, is a robust and biologically realistic mechanism underlying the toroidal topology.

1-011. Elucidating neuronal circuit function by comparing representations with deep RL agents

Peter Buttaroni[1] PETER.BUTTARONI@FMI.CH
Yael Bitterman[2] YAEL.BITTERMAN@FMI.CH
Julien Courtin[3] JULIEN.COURTIN@FMI.CH
Andreas Luthi[1] ANDREAS.LUTHI@FMI.CH
Friedemann Zenke[1] FRIEDEMANN.ZENKE@FMI.CH

[1]Friedrich Miescher Institute for Biomedical Research
[2]Hebrew University of Jerusalem
[3]Neurocenter Magendie

Goal-directed behavior requires effective internal representations of the external and choosing actions toward desired outcomes. While several higher-order brain circuits participate in this processing, we lack a deep understanding of what they represent and how their internal representations are shaped through learning. To address these questions, we develop a framework in which we compare the representational geometry of deep reinforcement learning (RL) agents and mice trained on the same instrumental goal-directed task. The agents have to perform a variable number of button pushes before collecting a food reward from a separate reward location. As a proof-of-principle, we systematically compare the internal representations in different network locations of trained actor-critic (AC) deep RL agents with in-vivo calcium recordings from the basolateral amygdala (BLA) in behaving mice using representational similarity analysis (RSA). We find a remarkable similarity between the representational geometry of the virtual agents and the animals. By replaying the sensory experience of a trained agent to an untrained/naive agent, we see that the task explains some but not all of this similarity, demonstrating that learning-induced changes render the representations of virtual and real agents more similar. Differentially comparing the BLA's geometry to the actor and the critic branches in our virtual agents, we found that BLA's geometry is more consistent with the critic, underscoring its well-documented role in value coding. However, our analysis also reveals essential differences in the geometry consistent across animals but not captured by the AC agent. This gap suggests that the BLA learns to encode more than plain value. We explore and discuss several hypotheses concerning the role of this encoding. In summary, we show that abstract RL agents display shared representational geometry with higher-order brain areas in behaving animals and how analyzing differences in similarity helps elucidating neuronal circuit function in goal-directed behavior.

1-012. A biologically constrained model of motion-processing neuron circuitry in the optomotor response

Whit Jacobs[1,2] WHITNEY.JACOBS@DUKE.EDU
Kaitlyn Fouke[1] KAITLYN.FOUKE@DUKE.EDU
Matthew Loring[1] MATTHEW.LORING@DUKE.EDU
Joe Choo-Choy[1] JPCCHOY@GMAIL.COM
Rishab Pulugurta[1] RISHAB.PULUGURTA@DUKE.EDU
Maxim Nikitchenko[1] NIKITCHM@GMAIL.COM
Timothy Dunn[1,3] TIMOTHY.DUNN@DUKE.EDU
Eva Naumann[1] EVA.NAUMANN@DUKE.EDU

[1]Duke University
[2]Neurobiology
[3]Biomedical Engineering

The transformation of visual input into behavior requires circuits of diverse neurons, yet there exists a gap in understanding how these neurons are connected and influence neural response dynamics. Recent studies into the visually guided optomotor response (OMR) in larval zebrafish have attempted to bridge this gap by proposing circuit models that predict connectivity among common types of motion-processing neurons. However, these models ignore the diversity of these neurons and fail to capture inherent circuit dynamics. Here, we generate hypotheses about functional connectivity of motion-processing neurons by training recurrent neural networks (RNNs) with calcium imaging data to model connections within the pretectum, a visual processing region involved in the OMR. Our model considers common and underrepresented classes of motion-processing neurons and is constrained by known anatomical connections, with model neurons receiving input from the contralateral retina and connecting to the ipsilateral and contralateral pretectum. We find our model excels at reproducing activity from neurons with

diverse motion preferences and that model neuron weights tend to cluster according to the directional selectivity of the target neuron. Additionally, model neuron currents rise during presentation of motion in the tuned direction of their target neuron, providing support for the biological plausibility of our model's solutions. By removing connections between groups of neurons in the model, we determine connectivity among similarly tuned neurons is necessary for successful training. Indeed, with holographic photostimulation, we confirm that neurons with shared directional tuning tend to be strongly connected, a feature of our model's solutions. To enhance our model's biological fidelity, we aim to permit connections among units only if their target neurons belong to motion-responsive classes found connected across fish. In summary, we propose a realistic, dynamic circuit model that is supported and constrained by experimental results to provide accurate predictions of neural circuitry underlying the OMR.

1-013. Decoupled neuronal activities during sleep give rise to synaptic downscaling

Aziza Yusupova[1] AZIZA.YUSUPOVA@UNIBAS.CH
Everton J Agnes[2,3] EVERTON.AGNES@UNIBAS.CH

[1]University of Basel, Biozentrum
[2]University of Basel
[3]Biozentrum

Sleep may induce specific alterations in synapses and neural activity patterns, however the underlying mechanism governing synapse modifications during sleep is not well understood. In some brain regions, e.g., rodent's somatosensory, prefrontal, and motor cortical areas, synapses have been shown to weaken during sleep. In others, e.g., rodent's visual cortex, synapses have been shown to remain unaltered. In this study, we investigated how excitatory synapses change in response to alterations in pre- and postsynaptic spike patterns, designed to mimic sleep/awake patterns of neuronal activity. We simulated changes in excitatory synapses via a weight-dependent spike-timing-dependent plasticity (STDP) rule with additive long-term potentiation (LTP) and multiplicative long-term depression. The combination of correlated spike patterns and STDP resulted in a steady state Gaussian-like distribution of synaptic weights, with mean and standard deviation well described by the synaptic plasticity model's parameters as well as the correlation between pre- and postsynaptic activity. Based on these results, we fitted experimentally reported distribution of excitatory synapse strengths from mice motor and somatosensory cortices as a composite of multiple Gaussian-like functions. From this fit, we then hypothesised that each Gaussian would correspond to a group of connections undergoing plasticity with unique parameter sets. To test this hypothesis, we incorporated metaplasticity dynamics to the amplitude of LTP, set to integrate the weight dynamics, which separated synapses into these distinct groups. Our findings show that excitatory synapses exhibit a natural strengthening (weakening) during periods of correlated (uncorrelated, i.e., decoupled) pre- and postsynaptic activity, reminiscent of awake (sleep) states, mirroring experimental observations. Importantly, metaplasticity, which can be linked to consolidation of synaptic weights, gives rise to long-tail distribution of synaptic weights as reported experimentally. Our results suggest that the weakening of excitatory synapses may arise from decoupled synaptic inputs during sleep rather than an active scaling-down process.

1-014. Distinct excitatory and inhibitory connectivity structures control the dynamics and computational capabilities of recurrent networks

Emmanouil Giannakakis[1,2] GIANNAKAKISMANOS@GMAIL.COM
Victor Buendia[3] VBUENDIAR@ONSAGER.UGR.ES
Oleg Vinogradov[4] OLEG.VINOGRADOV@UNI-TUEBINGEN.DE
Sina Khajehabdollahi[4] SINA.ABDOLLAHI@GMAIL.COM
Anna Levina[4] ANNA.LEVINA@UNI-TUEBINGEN.DE

[1]University of Tuebingen
[2]Computer Science
[3]Bocconi University
[4]University of Tubingen

Connectivity among cortical pyramidal neurons is far from random. Notably, it is known that there is a higher connection probability between similarly tuned neurons across cortical layers. In theoretical studies, this type of connectivity has been shown to generate rich dynamics in spiking networks. While the connectivity of inhibitory neurons is usually less specific than that of excitatory neurons, several studies have found different degrees of inhibitory specificity and diverse connectivity patterns, whose impact on population dynamics and computational implications are not fully understood. Here we link the presence of inhomogeneous localised, neuron-type specific clusters with complex dynamics that are associated with optimal performance in computational tasks.

Using a reservoir computing framework, we evaluate the computational capabilities of balanced, recurrent networks of rate neurons with neuron-type specific connectivity patterns. The reservoirs are composed of several interconnected clusters of E/I populations and are trained to simultaneously predict the trajectories of multiple chaotic time series. We study the impact of varying E and I clustering levels on network dynamics and identify the optimal topology for a complex time series reconstruction task.

We find that the presence of different levels of excitatory and inhibitory clustering enables the precise control of the network's dynamical state. In particular, we show that E and I clustering levels distinct and non-trivial effects on network dynamics and identify structures that control the network's distance from the chaotic state. Finally, we demonstrate that a commonly observed cortical connectivity pattern of highly specific excitation and less specific (but not uniform) inhibition among similarly tuned neurons can maintain network dynamics close to the edge of chaos and may significantly contribute to the computational efficiency of brain networks.

1-015. Multi-subject neural decoding via relative representations

Valentino Maiorca[1,2] MAIORCA@DI.UNIROMA1.IT
Simone Azeglio[3,4] SIMONE.AZEGLIO@GMAIL.COM
Marco Fumero[1] FUMERO@DI.UNIROMA1.IT
Clementine Domine[5] CLEMENTINE.DOMINE.20@UCL.AC.UK
Emanuele Rodola[1] RODOLA@DI.UNIROMA1.IT
Francesco Locatello[6] FRANCESCO.LOCATELLO@ISTA.AC.AT

[1]Sapienza University of Rome
[2]Computer Science
[3]Sorbonne University & Ecole Normale Superieure
[4]Vision Institute & Laboratoire des Systemes Perceptifs
[5]UCL Gatsby
[6]Institute of Science and Technology Austria

Neural decoding, i.e., the task of deciphering environmental signals and cognitive states from neural activity, is a fundamental direction in neuroscience research. The decoding process is generally specific to each subject and faces limitations in generalizability across different individuals, primarily due to the significant variability observed among subjects. We propose to adopt Relative Representations, a flexible computational framework, to efficiently map neural representations of distinct subjects into a common representational space agnostic of subject-specific attributes. The proposed framework builds on the assumptions of the existence of a second-order isomorphism between the neural activity of different individuals subject to the same stimuli, similarly to established neural analysis techniques like RSA. Relative representations introduces fundamental differences by: i) leveraging the expressiveness of artificial neural encoders, capable of extracting meaningful nonlinear patterns from the sparse and redundant high dimensional structure in fMRI signals; ii) being a better fit to capture the geometrical structure of the isomorphism, through explicit control on the invariances enforced on the representation. We demonstrate the applicability of the proposed framework on retrieval tasks of fMRI representations across subjects on the Natural Scenes Dataset benchmark, showing the usefulness and robustness of the proposed tool for processing neuronal data.

1-016. Towards the neuroethology of vocal communication in the Mongolian gerbil

Ralph Peterson[1,2] RALPH.EMILIO.PETERSON@GMAIL.COM
Aramis Tanelus[1] ARAMIS.TANELUS@NYU.EDU
Aman Choudhri[3] AMAN.CHOUDHRI@COLUMBIA.EDU
Violet Ivan[1] VJI2004@NYU.EDU
David Schneider[1] DS5577@NYU.EDU
Dan Sanes[1] DHS1@NYU.EDU
Alex Williams[1] ALEX.H.WILLIAMS@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Flatiron Institute

Social animals congregate in groups and communicate with vocalizations. To study the dynamics of natural vocal communication and their neural basis, one must characterize signals used for communication and determine the sender and receiver of the signal [1]. To this end, we established two complementary approaches for (1) quantifying vocal repertoire using a variational autoencoder (VAE) with longitudinal audio recordings in a naturalistic

social environment, and (2) vocal call attribution using a deep neural network. We pursued this research by establishing a unique and favorable model organism — the Mongolian gerbil — which has a sophisticated vocal repertoire and complex social hierarchy, including pair bond formation [2]. Here, we made continuous acoustic recordings of three separate gerbil families for 20 days each, and used a VAE for unsupervised representation learning of acoustic features to show that gerbil families have family specific vocal repertoires. Although this result positions the gerbil as an intriguing model of social vocal interactions, the inability to attribute vocalizations to individuals in a group limits interpretability of these family vocal differences, and remains a persistent problem for others in the field. We have therefore developed (1) a supervised deep learning framework with calibrated uncertainty estimates that achieves state-of-the-art sound source localization performance, (2) novel hardware solutions to generate benchmark datasets for training/evaluating sound source localization models across labs, and (3) curated and released the first large-scale benchmark datasets for vocal call localization in social rodents.

1-017. Emulating Human Visual Representational Structure in DNN via Relational Knowledge Distillation

Yuria Shimizu[1,2] YURIA19971120@GMAIL.COM
Masaru Sasaki[1] M-SASAKI@G.ECC.U-TOKYO.AC.JP
Takato Horii[3] TAKATO@SYS.ES.OSAKA-U.AC.JP
Masafumi Oizumi[1] C-OIZUMI@G.ECC.U-TOKYO.AC.JP

[1]The University of Tokyo
[2]Graduate School of Arts and Sciences
[3]The University of Osaka

Deep Neural Networks (DNNs) have become increasingly successful at solving visual classification tasks, often reaching or exceeding human-level accuracy. However, comparable classification performance does not necessarily imply that their internal representations are similar. In fact, closer examination reveals that the internal representations of DNNs differ significantly from human representations [1]. To address this gap, our study explores the potential of methods that directly make the internal representations of DNNs more human-like via Relational Knowledge Distillation (RKD) [2], which directly teaches DNNs the similarity relations of human representations. Specifically, we used RKD to teach DNNs the structure of the human psychological embedding of natural objects, as estimated by massive data of human similarity judgments collected in two open datasets (THINGS-odd-one-out [3] and Imagenet-HSJ [4]). The degree of similarity between the internal representations of DNNs and human psychological representations is evaluated by the conventional representational similarity analysis (RSA) and a recently developed unsupervised alignment method based on Gromov-Wasserstein optimal transport (GWOT) [5]. We show that after RKD, DNNs exhibit representational structures that are more similar to human visual representations, to the extent that two structures can be aligned even in the purely unsupervised alignment method. This research underscores the importance of representational structure in DNNs and posits that directly emulating this aspect may be a potent strategy for promoting human-like representational structures.

1-018. Towards Neural-Fidelity in Cognitive Task Modeling via Procrustes Distance Optimization

Yihao Li[1] LIYIHAO020302@GMAIL.COM
Wenxin Che[2] CHEWX@MIT.EDU
Zhaoze Wang[3] WANGZHZ@BU.EDU
Nathan Cloos[4] NACLOOS@MIT.EDU
Guangyu Robert Yang[4] YANGGR@MIT.EDU
Christopher Cueva[4,5] CCUEVA@GMAIL.COM

[1]Tsinghua University
[2]Southern University of Science and Technology
[3]University of Pennsylvania
[4]Massachusetts Institute of Technology
[5]Department of Brain and Cognitive Sciences

Neural networks are widely used for modeling neural activity in the brain. However, there is a growing recognition that the similarity between models and brains relies on a number of ad hoc design choices, including the specification of appropriate model inputs, outputs, and regularization hyperparameters applied during training. To address these challenges, we present an efficient and highly-generalizable approach that directly optimizes the Procrustes distance between recurrent neural network (RNN) activity and neural data with gradient descent. Directly tuning RNN connectivity via Procrustes distance minimization results in significantly better alignment be-

tween network activity and neural data, even when comparing models to neural data from experimental conditions that were never seen during training. Additionally, we have demonstrated that neural recordings collected during different tasks, albeit from the same region (M1), can enhance zero-shot neural alignment in neural network models trained on distinct tasks. Thus, we view this method as a general regularizer that leverages a small amount of neural data to select, from the huge space of potential models, those that are aligned with the brain. We subsequently leverage our method to facilitate an efficient search of the optimal inputs and outputs that guarantee higher neural-fidelity. Notably, the refined inputs show potential for mirroring neural perceptions of task variables, and the outputs appear to align more closely with downstream neural signals.

1-019. Representations of the intrinsic value of information in mouse orbitofrontal cortex

Jennifer Bussell[1,2] JBUSSELL@GMAIL.COM
Ryan Badman[3,4] RYAN.BADMAN113@GMAIL.COM
Christian Marton[5] CDMARTON@GMAIL.COM
Ethan Bromberg-Marton[6] BROMBEE@WUSTL.EDU
Larry Abbott[7] LFABBOTT@COLUMBIA.EDU
Kanaka Rajan[8] RKANAKA@GMAIL.COM
Richard Axel[9] RA27@COLUMBIA.EDU

[1]Columbia University, Zuckerman Institute
[2]Neuroscience
[3]1. Harvard Medical School 2. Kempner Institute, Harvard University
[4]Neurobiology
[5]Icahn School of Medicine at Mt. Sinai
[6]Washington University School of Medicine, St. Louis
[7]Columbia University
[8]Harvard University Kempner Institute
[9]Columbia University / HHMI

Animals often seek information of no apparent extrinsic value, suggesting that information is a source of intrinsic value. We developed an odor-based task in mice to study the neural mechanisms that compute the value of informative stimuli that drive information seeking. Mice are provided with odor cues in a center port that dictate entry into one of two side ports. The side ports offer water reward with identical probability and differ only in whether they provide an information odor cue that reveals the reward outcome. We observed that mice choose to acquire knowledge about uncertain reward, with their preference for information scaling with the delay length between information cues and the reward outcome. We also observed that mice were willing to pay for information by sacrificing water reward, suggesting that knowledge is of intrinsic value to a mouse. We fit cognitive models to choices in the information seeking task, and the animals' behavior was best fit by a model with separate reinforcement learning terms for extrinsic water value and intrinsic information value. We imaged neural activity in orbitofrontal cortex during the information seeking task. We observed different but overlapping populations of neurons responsive to odors predictive of information and odors predictive of water reward. We modeled OFC population neural activity using CEBRA, a nonlinear latent variable model. We were able to decode information and water reward separately within the model embedding. We also found that distances between state-space trajectories of information and no information scaled with the delay following informative odor cues. These data suggest that mice have evolved distinct pathways in OFC that represent the intrinsic value of knowledge and the extrinsic value of water reward. Thus, the desire to acquire knowledge is observed in mice and the value of knowledge is represented in the OFC.

1-020. Inferring Neuronal Identity from Functional Correlation Patterns in Mouse Hippocampal Circuits

Margaret Conde-Paredes[1] MC5155@COLUMBIA.EDU
Stephanie Herrlinger[1] SAH2245@COLUMBIA.EDU
Bovey Rao[1] BYR2104@CUMC.COLUMBIA.EDU
Attila Losonczy[1] AL2856@COLUMBIA.EDU
Erdem Varol[2,3] EV2240@NYU.EDU

[1]Columbia University
[2]New York University
[3]Computer Science

Neuronal type is hypothesized as a critical factor in a neuron's functional role within a neural circuit. This study explores the extent to which neuron type can be inferred from interaction patterns with other neuron types within local circuits. Such an understanding would facilitate correlating functional interactions with molecular, proteomic, or transcriptomic profiles, providing insight into whether the correlation structure within neural circuits reflects molecular identity. To this end, we developed a probabilistic graphical model to represent cell-specific functional activities as random variables from a hierarchical distribution capturing both single-cell to cell-type as well as cell-type-to-cell-type interactions. Using this model, we infer the molecular signatures of cells based on their correlations to other cells in circuits with known cell types. Validation was performed on data from calcium imaging of mouse hippocampal interneurons in the CA1 region, followed by immunohistochemistry imaging to identify six interneuron subtypes. Results indicate that functional correlations can help predict cell types, revealing a quantitatively stereotypical pattern of cell- type interactions in the mouse hippocampus.

1-021. Connectome-based models of feature selectivity in a cortical circuit

Jacopo Biggiogera[1,2] JACOPO.BIGGIOGERA@GMAIL.COM
Camilla Bianco[1] CAMILLA.BIANCO@STUDBOCCONI.IT
Victor Buendia[1] VBUENDIAR@ONSAGER.UGR.ES
Alessandro Sanzeni[1,2] ALESSANDRO.SANZENI@UNIBOCCONI.IT

[1]Bocconi University
[2]Computing Sciences

Feature selectivity, i.e. neurons' heightened responses to specific configurations of stimuli, constitutes a fundamental building block of cortical functions. Various mechanisms have been proposed to explain its origins, differing primarily in their assumptions about the connectivity between neurons. Some models ascribe selectivity to structured, tuning-dependent feedforward or recurrent connections, whereas others have demonstrated that selectivity can emerge within randomly connected networks when interactions are sufficiently strong. This diversity of plausible explanations poses a challenge in identifying the genuine mechanism for feature selectivity in cortex. We developed a novel approach that seeks to minimize preconceived assumptions about the underlying connectivity by utilizing connectomic data at synaptic resolution to construct network models. With this approach, we investigate the mechanisms governing selectivity to oriented visual stimuli in mouse visual cortex. We show that a connectome-constrained network model mimics neural responses seen in experiments and points at randomness in connectivity as the dominant factor shaping selectivity in the circuit. These findings provide novel insights on the mechanisms underlying feature selectivity in cortex and highlight the potential of connectome-based models for exploring the mechanistic underpinnings of brain functions.

1-022. Visuomotor control in virtually swimming Danionella larvae

Leonardo Demarchi[1,2] LEONARDO.DEMARCHI@SORBONNE-UNIVERSITE.FR
Monica Coraggioso[1] MONICA.CORAGGIOSO@SORBONNE-UNIVERSITE.FR
Antoine Hubert[1] ANTOINE.HUBERT@SORBONNE-UNIVERSITE.FR
Thomas Panier[1] THOMAS.PANIER@SORBONNE-UNIVERSITE.FR
Ghislaine Morvan-Dubois[1] GHISLAINE.MORVAN-DUBOIS@SORBONNE-UNIVERSITE.FR
Volker Bormuth[1] VOLKER.BORMUTH@SORBONNE-UNIVERSITE.FR
Georges Debregeas[1] GEORGES.DEBREGEAS@SORBONNE-UNIVERSITE.FR

[1]Sorbonne University
[2]Laboratoire Jean Perrin

The ability to maintain a stable state in the presence of perturbations, known as robustness, is a defining feature of living systems. Most animals demonstrate adaptability by adjusting their position when confronted with external stimuli, a crucial skill for survival. How the brain dynamically regulates behavior in response to a changing environment remains a fundamental question in neuroscience. We decided to investigate it using Danionella cerebrum, the smallest known vertebrate amenable to brain-wide functional imaging at cellular resolution across all developmental stages. We developed a 2D virtual reality system in which partially tethered larvae can navigate their visual environment. The system makes use of fluid dynamics estimates of the fish's intended movements to restore their expected visual feedback. We observed that they can continuously stabilize their position when subjected to external visual flows of varying speed and direction. We mathematically modeled this regulation process with a system of delay differential equations that can exhibit limit cycle oscillations, consistent with the observed speed fluctuations. Moreover, we were able to perform calcium imaging during these virtual reality experiments and identify neural populations spanning the entire brain with activities that correlate with specific features of both behavior and visual stimulation. Notably, we found assemblies of neurons that activate differentially during spontaneous or visually-evoked swimming. In conclusion, our study not only significantly advances our understanding of

how animals integrate sensory input in real-time to drive motor actions, but also introduces analysis and modeling tools with broader applicability, which may prove useful to other researchers in the field.

1-023. Learning, selectivity, and robustness: an AI-inspired model of the cholinergic system

Maija Filipovica[1] MAIJA.FILIPOVICA@BRISTOL.AC.UK
Kevin Kermani Nejad[1] KEVIN.KERMANINEJAD@BRISTOL.AC.UK
Will Greedy[1,2] WILL.GREEDY@BRISTOL.AC.UK
Heng Wei Zhu[3] HZ3791@NYU.EDU
Jack Mellor[1] JACK.MELLOR@BRISTOL.AC.UK
Rui Ponte Costa[4] RUI.COSTA@DPAG.OX.AC.UK

[1]University of Bristol
[2]Computer Science
[3]New York University
[4]University of Oxford

The cholinergic system has been associated with learning, but also with cognitive decline in dementia, aging and injury. Yet, to date, no computational models have been put forward to explain how the cholinergic system contributes to both learning and cognitive decline. Here we introduce a model that combines a recently proposed model of cortex-wide credit assignment with a cholinergic adaptive module based on adaptive deep learning rules. According to this model, the cholinergic system opens a cortex-wide gate for learning, but its end effect is controlled by local adaptive processes. Using both reinforcemet learning and perceptual discrimination tasks we show that cholinergic adaptive learning leads to rapid cortex-wide learning when compared with non-adaptive models. Such cholinergic adaptive modulation results in a constant redistribution of learning across the cortex making task-encoding sparser. Consequently, we show that the network becomes more robust to perturbations such as simulated cell death. Moreover, we demonstrate that to obtain such rapid and robust learning, global mechanisms are not sufficient, suggesting the need for a tight cholinergic interaction with local cortical circuits. Overall, our work provides a novel theoretical framework for cellular-systems neuroscience with which to link cholinergic cortical modulation to health and disease.

1-024. Different state-dependence of population codes across cortex

Akhil Bandi[1,2] ACBANDI@ANDREW.CMU.EDU
Caroline Runyan[3,4] RUNYAN@PITT.EDU

[1]Carnegie Mellon University
[2]Neuroscience Institute
[3]University of Pittsburgh
[4]Neuroscience, Center for the Neural Basis of Cognition

During perceptual decision-making, behavioral performance varies with changes in internal states of arousal, motivation, engagement, and strategy. It is unknown how internal states related to behavioral performance affect information coding in the sensory and association cortices involved in sensory perception and decision-making. To answer this question, we recorded neural activity from the primary auditory cortex (A1) and posterior parietal cortex (PPC) in mice performing a navigation-based sound localization task. We then modeled transitions in the behavioral strategies mice used during task performance with a latent-state model. Mice transitioned between three latent behavioral states with differing decision-making strategies: an ‘engaged’ state with near-perfect task performance and optimal choice formation, and two ‘disengaged’ states characterized by choice bias and frequent errors. Using a linear classifier, we could successfully classify the identity of the inferred behavioral states using PPC but not A1 population activity. This suggests that behavioral state strongly influences population activity patterns in association cortex but not sensory cortex. We next compared neural encoding of task variables across behavioral states by training encoding models with information from exclusively engaged state trials or trials spanning all three states. Task variable encoding in A1 was unaffected by state, while task variable encoding in PPC was surprisingly stronger during the disengaged, biased states. However, in the engaged state, functional coupling between neurons better explained PPC neuronal activity than task variables alone. Functional coupling did not improve model performance when including disengaged states. Together, these findings indicate that transitions in latent behavioral state differently impact sensory and association cortex and show how behavioral states of high task engagement drive PPC neural populations into highly correlated activity states that benefit optimal choice formation and task performance.

1-025. Interpretable representations of neural dynamics using geometric deep learning

Adam Gosztolai[1] GOSZTOLA@MIT.EDU
Alexis Arnaudon[2,3] ALEXIS.ARNAUDON@EPFL.CH
Robert Peach[4] R.PEACH13@IMPERIAL.AC.UK
Mauricio Barahona[5] M.BARAHONA@IMPERIAL.AC.UK
Pierre Vandergheynst[1] PIERRE.VANDERGHEYNST@EPFL.CH

[1]Ecole Polytechnique Federale de Lausanne
[2]EPFL
[3]BBP
[4]University of Wurzburg
[5]Imperial College London

It is increasingly recognised that computations in the brain and artificial neural networks can be understood as outputs of a high-dimensional dynamical system conformed by the activity of large neural populations. Yet revealing the structure of the underpinning latent dynamical processes from data and interpreting their relevance in computational tasks remains a fundamental challenge. A prominent line of research has observed that task-relevant neural activity often takes place on low-dimensional smooth subspaces of the state space called neural manifolds. However, there is a lack of theoretical frameworks for the unsupervised representation of neural dynamics that are interpretable based on behavioural variables, comparable across systems, and decodable to behaviour with high accuracy.

To address these challenges, we introduce Manifold Representation Basis Learning (MARBLE), a fully unsupervised representation-learning framework for non-linear dynamical systems. Our approach combines empirical dynamical modelling and geometric deep learning to transform neural activations during a set of trials into statistical distributions of local flow fields (LFFs). Our central insight is that LFFs vary continuously over the neural manifold, allowing for unsupervised learning, and are preserved under different manifold embeddings, allowing the comparison of neural computations across networks and animals.

We show that MARBLE offers a well-defined similarity metric between different neural systems that is expressive enough to compare computations and detect fine-grained changes in dynamics due to task variables, e.g., decision thresholds and gain modulation. Being unsupervised, MARBLE is uniquely suited to biological discovery. Indeed, we show that it discovers more interpretable neural representations in several motor, navigation and cognitive tasks than generative models such as LFADS or (semi-)supervised models such as CEBRA. Intriguingly, this interpretability implies significantly higher decoding performance than state-of-the-art. Our results suggest that using the manifold structure yields a new class of algorithms with higher performance and the ability to assimilate data across experiments.

1-026. Retrosplenial cortex dynamics underlying psilocybin-enhanced fear extinction

Sophie Rogers[1,2] RSOPHIE@PENNMEDICINE.UPENN.EDU
Stephen Wisser[1] STEPHEN.WISSER@PENNEMEDICINE.UPENN.EDU
Kyle Czarnecki[1] KYLE.CZARNECKI@PENNMEDICINE.UPENN.EDU
Elizabeth Heller[1] EHELLER@PENNMEDICINE.UPENN.EDU
Gregory Corder[1] GREGORY.CORDER@PENNMEDICINE.UPENN.EDU

[1]University of Pennsylvania
[2]Neuroscience Graduate Group

The psychedelic serotonin 2 receptor (5HT2R) agonist psilocybin has demonstrated rapid and long-lasting efficacy across neuropsychiatric disorders characterized by cognitive inflexibility. Psilocybin induces rapid and stable dendritic growth, suggesting that psilocybin enhances cognitive flexibility by literally creating new pathways for information flow and transformation in the brain. However, the direct impact of psilocybin on patterns of neural activity underlying sustained changes in cognitive flexibility has not been characterized. Behavioral flexibility is implemented in a variety of cortical structures, including the 5HT2R-enriched retrosplenial cortex (RSC). The RSC is necessary for both appetitive and aversive reversal learning behaviors, including extinction of trace fear conditioning (TFC). To test the hypothesis that psilocybin enhances cognitive flexibility by rapidly and persistently altering activity in RSC neural ensembles underlying behavior, we applied tensor component analysis (TCA) to single cell calcium activity recorded in the RSC over a five-day TFC extinction assay in saline- and psilocybin-administered mice. TCA revealed components of neural activity in each animal that drove RSC dynamics in trials within particular sessions of TFC training. The evolution of RSC dynamics through these components predicted extinction rate in mice, suggesting we had successfully isolated task-relevant dynamics. By identifying

heavily weighted neurons in components dominating particular sessions, we found that psilocybin induced a larger turnover in the ensembles driving RSC activity between acquisition and early extinction sessions, with the recruitment of a novel ensemble during early extinction that persisted more throughout subsequent days. Psilocybin also persistently suppressed activity in fear acquisition-related neurons and enhanced late extinction-related neurons in responders. This work reveals novel mechanisms of psilocybin-enhanced cognitive flexibility and demonstrates how the application of computationally sophisticated techniques to neurophysiology data can generate novel contributions to neuropharmacology.

1-027. Networks of descending neurons transform command signals into population-based behavioral control

Femke Hurtak[1] FEMKE.HURTAK@EPFL.CH
Jonas Braun[1] JONAS.BRAUN@EPFL.CH
Sibo Wang-Chen[2] SIBO.WANG@EPFL.CH
Pavan Ramdya[1] PAVAN.RAMDYA@EPFL.CH

[1]Ecole Polytechnique Federale de Lausanne
[2]EPFL

Behavioral control is a complex problem for animals and robots, involving interactions from the brain's descending modulation to low-level mechanical feedback. Resolving how nervous systems solve this multiscale challenge is a major goal in neuroscience. In particular, the nature of descending signals that drive activity in the spinal (or ventral) cord remains poorly understood. In this work we resolve a long-standing contradiction between two models —population versus command-like— of descending control. The population control model suggests that the coactivation of multiple descending neurons (DNs) is required to coordinate complex behaviors ; the command-like control model suggests that the drive from small sets of DNs is sufficient. Both models are supported by experimental evidence. We uncover a unifying DN control mechanism in Drosophila whereby small sets of command-like DNs can recruit larger networks of interconnected DNs. We first observed this recruitment through functional recordings of DN populations while simultaneously optogenetically activating known command-like DNs. Through behavioral experiments, we show that this DN recruitment is necessary to generate complex behaviors: without it flies generate simpler movements akin at motor primitives. Remarkably, the extent of a command-like DN's functional recruitment of other DNs can be explained by direct chemical synaptic connections —as measured in the fly brain connectome (FAFB). Using graph analysis on the DN-DN network, we find that DNs are segregated into clusters coarsely corresponding to behaviors such as walking, anterior grooming, or takeoff. We hypothesize that these predominantly excitatory clusters enable the coordination of flexible behaviors. Additionally, we observe that inhibitory DNs belonging to clusters inhibit behaviorally distinct clusters. Our work —combining functional imaging, behavioral experiments and graph analysis — introduces a novel paradigm for the descending control of behavior which intrinsically enables flexibility and the capacity to evolve novel behaviors through the combination of simpler motor primitives.

1-028. Learning efficient backprojections across cortical hierarchies in real time

Kevin Max[1,2] KEVIN.MAX@UNIBE.CH
Laura Kriener[1,3] LAURA.KRIENER@UNIBE.CH
Walter Senn[1] WALTER.SENN@UNIBE.CH
Garibaldi Pineda Garcia[4] G.PINEDA-GARCIA@SUSSEX.AC.UK
Thomas Nowotny[4] T.NOWOTNY@SUSSEX.AC.UK
Mihai Petrovici[1] MIHAI.PETROVICI@UNIBE.CH

[1]University of Bern
[2]Dept of Physiology
[3]Institute for Physiology
[4]University of Sussex

Models of sensory processing in the cortex need to efficiently assign credit to synapses in early areas. In deep learning, a known solution is backpropagation, which however requires biologically implausible weight transport from feed-forward to feedback paths. We introduce Phaseless Alignment Learning (PAL), a bio-plausible method to learn useful top-down weights in layered cortical hierarchies. This is achieved by exploiting the noise naturally found in bio-physical systems as an additional carrier of information. In our fully dynamical system, all weights are learnt simultaneously with always-on plasticity and using only information locally available to the synapses. Hence, our method is completely phase-free and allows for efficient error propagation across multi-layer cortical

hierarchies, while maintaining biologically plausible signal transport.

1-029. Structure of subjective representations predicts the efficiency of transfer learning

Anna Szekely[1,2] SZEKELY.ANNA.95@GMAIL.COM
Gergo Orban[3,4] ORGERGO@GMAIL.COM
Balazs Torok[5] BALAZSTOEROEK@GMAIL.COM
Mariann M. Kiss[6] KISSMARYANNE@GMAIL.COM
Karolina Janacsek[7] JANACSEKKAROLINA@GMAIL.COM
Dezso Nemeth[8] NEMETHD@GMAIL.COM

[1]Wigner Research Centre for Physics // Budapest University of Technology and Economics
[2]Department of Computational Sciences // Department of Cognitive Science
[3]Wigner RCP
[4]Dep Computational Sciences
[5]Mozalearn Ltd
[6]Institute of Psychology, Faculty of Education and Psychology, Eotvos Lorand University
[7]School of Human Sciences, Faculty of Education, Health and Human Sciences, University of Greenwich
[8]Centre de Recherche en Neurosciences de Lyon, Universite Claude Bernard Lyon 1

Surviving in a natural environment demands agents to be able to cope with previously unseen situations. This challenge requires agents to utilise their previous knowledge effectively and generalise to novel situations, known as transfer learning. Previous attempts to charactherise the computational basis of transfer learning mainly focused on the characterisation of behaviour via ideal observer models. Here, we propose an alternative approach, emphasising the importance of characterising individual learning trajectories to gain more insights into individual internal models (IMs) and their use during transfer. We argue that successful transfer learning hinges upon the acquisition of the tasks' underlying structure, but this learning often shows high variance across individuals. Thus, it is essential to assess the diversity of individual learning strategies. We introduce an infinite Hidden Markov Model-based (iHMM) method to characterize the evolution of individualized internal models and its use in transfer. To study transfer learning, we conducted an implicit sequence learning experiment with 25 human participants, in which a sequence was trained for eight days, after which a new sequence was introduced which retained the high-level structure of the first sequence. We used the Cognitive Tomography (CT) method to infer the iHMM from reaction time data alone. CT captures the gradual development of IMs and enables us to compare the reuse of previously acquired IMs across different tasks. We show that i) participants are able to develop an inventory of IMs that apply to different tasks, ii) they are able to alternate between IMs according to the environmental statistics, and that iii) the acquired imperfect task representation guides learning in a new task. Taken together, our study provides novel insights into the computational underpinnings of transfer learning, and demonstrates the power of inferring complex individualized mental representations from easy-to-obtain behavioral data.

1-030. Vector Field Learning on Latent Manifolds

Robert Peach[1] R.PEACH13@IMPERIAL.AC.UK
Matteo Vinao-Carl[2] MATTEO.VINAO-CARL16@IMPERIAL.AC.UK
Nir Grossman[2] NIRG@IMPERIAL.AC.UK
Michael David[2] MICHAEL.DAVID12@IMPERIAL.AC.UK
Emma Mallas[2] E.MALLAS@IMPERIAL.AC.UK
David Sharp[2] DAVID.SHARP@IMPERIAL.AC.UK
Paresh Malhotra[2] P.MALHOTRA@IMPERIAL.AC.UK
Pierre Vandergheynst[3] PIERRE.VANDERGHEYNST@EPFL.CH
Adam Gosztolai[4] ADAM.GOSZTOLAI@EPFL.CH

[1]University of Wurzburg
[2]Imperial College London
[3]Ecole Polytechnique Federale de Lausanne
[4]EPFL

There is increasing recognition that the activity of neural populations can be captured by a much smaller number of key dimensions. This has led to the development of tools and algorithms aimed at generating neural embeddings that capture low-dimensional manifold structures. Yet, many datasets are not merely points in space but represent dynamic, smoothly varying vector fields over these manifolds. Examples include the firing rate trajectories of neural populations, changes in gene expression in neurodevelopment, or voltage gradients between electrical

leads in EEG recordings. Despite these numerous applications, there is a lack of theoretical frameworks can explicitly model vector fields on latent manifolds.

To address these challenges, we introduce Riemannian manifold Vector field Gaussian Process (RVGP), a generalisation of Gaussian Processes (GPs) for learning the vector fields of non-linear dynamical systems. Our approach combines ideas from probability theory and differential geometry to learn smooth continuous representations of vector signal data over latent Riemannian manifolds. RVGP is data-driven, approximating the manifold using only local similarities between sparse data points, yet can reconstruct intricate singular vector field structures by leveraging the global continuity of dynamics as an inductive bias. This property makes RVGP exceptionally robust to sampling density and noise, making it practical in real-world laboratory and clinical scenarios.

Using several synthetic and real-world examples, we demonstrate that RVGP can make accurate out-of-sample predictions on complex neural manifolds from very sparse training data. This makes RVGP uniquely adept at facilitating predicting neural dynamics in unrecorded state-space regions for testing biological hypotheses or enhanced clinical diagnostics. To exemplify this, we use RVGP to reconstruct high-density, laboratory-grade neural dynamics from low-density clinical EEG recordings in healthy individuals and Alzheimer's patients. We show that RVGP accurately reconstructs vector field singularities, leading to comparable classification accuracy of disease states to high-density recordings, overcoming significant practical limitations.

1-031. Unmasking cortical interactions: Understanding the influence of basal and apical dendrites in pyramidal cell computations

Sebastian Onasch[1] SEBASTIAN.ONASCH@TUM.DE
Julijana Gjorgjieva[1,2] GJORGJIEVA@TUM.DE

[1]Technical University of Munich
[2]School of Life Sciences

Pyramidal (PY) cells in layer 2/3 (L2/3) receive their inputs at two distinct loci. The basal dendrites receive bottom-up sensory information from the first-order thalamus and layer L4. The apical dendrites extend into L1 and receive top-down predictions. Recently, multiple experiments revealed cortical computations in V1, where top-down input from higher visual areas plays an important role. One example is the generation of inverse receptive fields, where neurons in L2/3 but not L4 respond strongly to input that stimulates everything but their 'classical' receptive field (Keller et al., Nature, 2020). Another example is surround facilitation, where the response of cells in V1 increases when an orthogonal surround accompanies the stimulus in the receptive field (Kirchberger et al., Science Advances 2021). Here, we established a new PY cell model with basal and apical dendrites, where inputs are filtered by a nonlinearity, reflecting active dendritic processes like N-methyl-D-aspartate (NMDA) receptor-induced plateaus. Using this model, we separately examined the influence of the two distinct pathways on the output of PY cells and the computations performed by the circuit. Furthermore, we investigated the effect of the subcellular targeting bias of inhibitory interneurons – parvalbumin-expressing (PV) neurons target the somatic area. In contrast, somatostatin-expressing (SOM) neurons target the dendrites and can modulate their integration properties. We found that top-down and bottom-up inputs differently shape the circuit responses and dissected the distinct roles of the basal and apical dendrites in phenomena like the paradoxical effect, where inhibitory input to a population can paradoxically lead to an increase in its firing rate. Moreover, we showed that the top-down inputs to PY cells alone are sufficient to explain complex network responses like the generation of an inverse receptive field. This framework allows us to explore multiple cortical computations involving top-down feedback.

1-032. Smoothly switching linear dynamical systems via Gaussian process models

Amber Hu[1] AMBERHU@STANFORD.EDU
David Zoltowski[1] DZOLTOWSKI@STANFORD.EDU
Aditya Nair ADI.NAIR@CALTECH.EDU
David Anderson[2] WUWEI@CALTECH.EDU
Lea Duncker[1] LDUNCKER@STANFORD.EDU
Scott Linderman[1,3] SCOTT.LINDERMAN@STANFORD.EDU

[1]Stanford University
[2]California Institute of Technology
[3]Statistics Department and Wu Tsai Neurosciences Institute

A key goal in neuroscience is to characterize how the collective activity of neurons in recurrent circuits relates to computation and ultimately behavior. In constrained experimental settings, it has been repeatedly reported that

high-dimensional neural population activity may be adequately described in terms of a set of low-dimensional, time-varying latent variables. Thus, methods to infer flexible descriptions of potentially nonlinear dynamics of these latent variables from neural measurements have become important tools for analyzing neural systems. Previous work has modeled such nonlinear dynamics using either (1) recurrent neural networks (RNNs) or Gaussian processes (GPs), which represent flexible function classes but are often challenging to interpret, or (2) linear or piecewise linear models, such as the recurrent switching linear dynamical system (rSLDS), which can be easily analyzed but are also more restrictive. Such simpler models may also fail to accurately describe nonlinearities, and can produce discontinuities at the boundaries between different linear regimes that hinder interpretability. Here, we propose to overcome these challenges using a novel formulation of a GP dynamical system – the smoothly switching linear dynamical system (sSLDS). Our approach relies on the design of a novel covariance function defining a GP that can smoothly interpolate between linear regimes, with arbitrarily complex partitions across regimes. Conceptually, our model maintains the switching linear structure and interpretability of the rSLDS, while extending it through the flexibility warranted by nonlinear GP models. The sSLDS incorporates the fixed point of each linear regime as a learnable parameter, which makes it directly accessible for further analysis, without post-hoc fixed-point-finding commonly required in RNN models. We leverage a variational inference framework to infer latent trajectories and dynamics along with estimates of their posterior uncertainty. We demonstrate the effectiveness of the sSLDS and its advantages over previous methods on simulated and real data.

1-033. Curriculum learning inspired by behavioral shaping effectively trains RNNs to mimic rat behaviors

David Hocker[1,2] DH148@NYU.EDU
Christine Constantinople[1,2] CMC472@NYU.EDU
Cristina Savin[1,3] CS5360@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Center for Neural Science, Center for Data Science

Recurrent neural network (RNN) models are a ubiquitous tool in neuroscience that aims to capture both neural activity and behaviors from living systems. However, as the complexity of cognitive tasks increases, conventional approaches for training RNNs on such tasks can fail to capture key aspects of animal behavior. In order to meet these increasing demands, we leverage a commonly used (though rarely appreciated) approach from the experimental neuroscientist's toolkit: behavioral shaping. Specifically, we decomposed a target temporal wagering task previously studied in rats, and designed a pretraining curriculum of simpler cognitive tasks that are prerequisites for performing this task well. These pretraining tasks are not simplified variants of the temporal wagering task, but rather out-of-context tasks that involve relevant sub-computations. We show that this approach is required to train RNNs on the temporal wagering task in order to learn long-timescale inference of latent states, and that conventional pretraining approaches fail to capture rat behavior. The resulting RNN dynamics are shaped by pretraining, and require several key dynamical systems features for implementing both inference and value-based decision making. Our approach addresses a missing gap in neural network models of biological systems by using pretraining to incorporate prior animal experiences, which is important when modeling complex behaviors that rely on multiple computations

1-034. Mesolimbic dopamine encodes reward prediction errors independent of learning rates

Andrew Mah[1,2] AM9056@NYU.EDU
Carla Golden[1,2] CG163@NYU.EDU
Christine Constantinople[1] CONSTANTINOPLE@NYU.EDU

[1]New York University
[2]Center for Neural Science

Subjective uncertainty about the state of our environment plays a crucial role in shaping how we learn from previous outcomes. In reinforcement learning tasks, uncertainty modulates the degree to which we learn from outcomes via the learning rate. When state uncertainty is high, for example in a volatile environment, more recent outcomes should be given more weight with a faster learning rate, because those outcomes are more likely to be informative about future states compared to more distant outcomes. However, neural mechanisms that modulate learning rates are unclear. While mesolimbic dopamine is classically thought to instantiate error signals for learning, previous accounts suggest it reflects the product of those error signals with the learning rate, or more recently, perhaps encodes learning rates directly. To study how dopamine interacts with dynamic

learning rates, we measured dopamine activity in rats performing a temporal wagering task with hidden reward states (blocks of trials with differing reward distributions). Previous work has suggested that rats used a state-uncertainty-based dynamic learning rate to update how quickly they initiated trials in this task1, which allows us to relate dopamine to learning rates within subjects. Using a combination of behavioral analyses, computational modeling, and statistically powerful behavioral datasets from hundreds of rats, we present compelling evidence that rats modulate their learning rates based on state uncertainty. Next, we measured dopamine release using GRABDA, a fluorescent dopamine sensor, in the nucleus accumbens core of rats performing the task. Despite the correlation between dopamine release and trial-by-trial update of trial initiation times, we found that dopamine encoding of reward prediction errors was not affected by the dynamic learning rate. These results confirm the classical view that dopamine release conveys a learning-rate-independent error signal, and suggests that other neuromodulators, for example serotonin2, might adapt the rate of learning from outcomes.

1-035. Neural Dynamics of Cognitive Flexibility: Unraveling mPFC Activity in Learning and Reversal

Madelyn Hjort[1,2] MADELYN.HJORT@GMAIL.COM
Garret Stuber[1] GSTUBER@UW.EDU

[1]University of Washington
[2]Graduate Program in Neuroscience

It is important that the brain can flexibly update learning in response to changing environments, and maintain stable responding when conditions are stable. Although the medial prefrontal cortex(mPFC) has long been implicated in learning and cognitive flexibility (PMID34408280), how large scale neural dynamics adapt and contribute to behavioral flexibility remain unclear. We assessed learning in stable and changing conditions using a Pavlovian contingency reversal flanked by days of stable Pavlovian behavior while longitudinally recording neurons with two-photon calcium imaging to profile mPFC activity. When we modeled learning as the rate of sustained change in value-guided behavior via a generalized-linear model(GLM), we identified a unique population of 'reversal' neurons that activates during the contingency reversal. This population minimally overlaps with cue, value, or value update encoding before, during, or after the reversal. Kernel regression suggests that reversal cell signal is strongest at cue onset and following reward delivery. Previously a population of neurons in mPFC were identified where neural dynamics were stable across contingency-matched cue-sets (PMID37382590). We found that a stable cue population persists across reversal for high value cues, even when these sessions are separated by multiple weeks. This stable cue population shares little overlap with the population of 'reversal' neurons. Complimentarily, when we performed k-nearest neighbors (KNN) spectral clustering on principle components of trial averaged neural activity post-reversal, we isolated four active clusters, each preferentially enriched in GLM-identified reversal, value, value update, or cue cells. Experiments underway will assess whether (1) 'reversal' cell activity causally impacts behavior or cue cell remodeling via patterned holographic optogenetics and (2) dopaminergic release in the mPFC during Pavlovian reversal learning (PMID11917003) impacts activity of GLM-defined cell populations including reversal cells. The isolation of a cognitive flexibility signal will be interesting to experimentalists, while the models that allowed that isolation may interest theorists.

1-036. Zero-code containerization and data annotation for sharing computational neuroscience methods

Shay Neufeld[1,2] SHAY.NEUFELD@BRUKER.COM
Srinivas Gorur-Shandilya[1] S.GORUR_SHANDILYA@BRUKER.COM
Alina Quereilhac[1] AL.QUEREILHAC@BRUKER.COM
Bruno Boivin[1] BRUNO.BOIVIN@BRUKER.COM
Braden Neufeld[1] BRADEN.NEUFELD@BRUKER.COM

[1]Inscopix
[2]Data Products and Analytics

Computational methods developed by neuroscientists are often not widely used, despite the desire to openly share them. Larger datasets and the growing complexity of computational methods have led to a key gap in our ability to fully benefit from the effort spent by scientific programmers in academic labs developing innovative analytical tools. Here, we propose a novel framework to share computational methods that makes it possible to widely adopt and execute scientific workflows on diversely organized neuroscience data with minimal extra effort required. Our approach uses code introspection to ingest Python modules and generate wrappers that allow them to run in a Docker container. Crucially, this process requires zero extra lines of code to be written by hand, and the entire procedure from import of code to its execution on user-supplied data is handled either automatically or

via annotations in an interactive web app. Annotations of input and output files also identify file types and data structures within a flexible ontology, formalizing implicit contracts between code and arbitrarily formatted data inputs and results. While not required to adhere to any specific data formatting, this approach is compatible with standards like Neurodata Without Borders (NWB) to further facilitate the dissemination and reuse of neuroscience research data. The combination of automated containerization and streamlined data annotation distinguishes our work from other efforts, making it possible to easily interconnect analytical methods in a computational graph that embodies a data pipeline. This process further allows analytical methods and corresponding data to integrate seamlessly with cloud computing and data storage systems. Our work has been focused on rodent calcium imaging and behavior data but is generalizable to most neuroscience workflows and data.

1-037. Local field potentials, excitation-inhibition balance, and network communication in rodent mPFC

Geoffrey Diehl[1,2] GDIEHL@UMN.EDU
David Redish[1] REDISH@UMN.EDU

[1]University of Minnesota
[2]Neuroscience

The dynamic balance between excitatory (E) and inhibitory (I) drive is thought to be critical for proper cognitive functioning and effective decision-making. Yet, gaining access to measures of network communication, and E:I balance, has remained challenging. Recent work has posited that the aperiodic spectral component of local field potentials (LFPs) provides a direct window into the dynamic balance between excitation and inhibition (Gao et al., 2017; Donoghue et al., 2020). However, to date this hypothesis has not been directly tested in complex behavioral tasks.

To probe these network features, we recorded single unit activity and LFPs from the medial prefrontal cortex (mPFC) of rats performing a decision-making task. Computing cross-correlation histograms (CCHs) between pairs of simultaneously recorded cells revealed clear instances of monosynaptic connections, both excitatory and inhibitory. Notably, examination of these coupled pairs can reveal features of mPFC network architecture. For example, we found an inverse relationship between coupling probability and pairwise distance between cells, highly consistent with computational models of network architecture. Interestingly, this relationship held for inhibitory synapses as well, directly challenging models that posit excitation is local but inhibition global.

To test the proposed relationship between E:I balance and LFP aperiodic, we computed the exponential slope of LFP spectra in 512ms bins. We then divided our recording into chunks of aggregate time, grouping bins according to slope magnitude. Computing CCHs in each chunk of time allowed us to directly quantify excitatory and inhibitory synaptic drive, and the E:I balance between them, for different degrees of exponent slope. Interestingly, our data did not support the hypothesis that exponent slope is inversely related to E:I balance. Instead, we found that overall broadband power of the LFP spectrum was significantly related to E:I ratio, presenting a possible alternative model for the relationship between cellular spiking and emergent LFP.

1-038. Dissecting modular recurrent neural networks with different cell types trained to perform un-cued task switching

Yue Liu ALLENLIU3@HOTMAIL.COM
Xiao-Jing Wang XJWANG@NYU.EDU

New York University

Behavioral flexibility relies on the brain's ability to switch rapidly between multiple tasks, even when the task rule is not explicitly cued but must be inferred through trial and error. The underlying neural circuit mechanism remains poorly understood. We investigated recurrent neural networks (RNNs) trained to perform an analog of the classic Wisconsin Card Sorting Test. The networks consist of two modules responsible for rule representation and sensorimotor mapping, respectively, where each module is comprised of a circuit with excitatory neurons and three major types of inhibitory neurons. We found that rule representation by self-sustained persistent activity across trials, error monitoring and gated sensorimotor mapping emerged from training. Systematic dissection of trained RNNs revealed a detailed circuit mechanism that is consistent across networks trained with different hyperparameters. The networks' dynamical trajectories for different rules reside in separate subspaces of population activity; they became virtually identical and performance was reduced to chance level when dendrite-targeting somatostatin-expressing interneurons were silenced, demonstrating that rule-based gating critically depends on the disinhibitory motif.

1-039. Scaling neural network training enables neuromotor interface decoders to generalize to new users

Jorge A Menendez[1] JAMENENDEZ11@META.COM
Diogo Peixoto[1] DIOGOPEIXOTO@META.COM
Ali Farshchian[2] AFAR@META.COM
Jean-Christophe Gagnon-Audet[2] JCAUDET@META.COM
Ning Guo[2] NINGGUO@META.COM
Nirag Kadakia[2] NIRAGKADAKIA@META.COM
Patrick Kaifosh[1] KAIFOSH@META.COM
Calvin Kao[2] TCK@META.COM
Michael Mandel[2] MMANDEL@META.COM
Jesse D Marshall[1] JESSEDMARSHALL@META.COM
Josh Merel[1] JSMEREL@META.COM
Ricardo Monti[2] RPMONTI@META.COM
Tejaswy Pailla[2] TEJASWY@META.COM
Eftychios Pnevmatikakis[2] EFTYCHIOS@META.COM
Thomas R Reardon[1] TRR@META.COM
David Schwab[2] DSCHWAB@META.COM
David Sussillo[1] SUSSILLO@META.COM
Jimmy Wang[2] JIMMYWANG@META.COM
Daniel Wetmore[1] WETMORE@META.COM
Ctrl Labs[1] TRR@FB.COM

[1]Meta Reality Labs
[2]Meta Reality Labs (CTRL Labs group)

A long-standing goal in myoelectric and brain computer interface research is to develop decoding models that work for anyone out-of-the-box, without the need for participant- or session-specific calibration1,2. This has been a challenge because recorded signals can vary substantially across participants and sessions, due to idiosyncrasies in individuals' anatomy, physiology, and behavior3,4. Here, we illustrate this point by showing that surface electromyography (sEMG) signals at the wrist evoked by the same gesture vary substantially across participants (Fig. 1a). As a result, a decoding model trained on an individual participant's data performs well only for that participant, and does not generalize to others (Fig. 1b). To test whether a decoding model trained on multiple participants' data can generalize, we developed a novel non-invasive dry electrode sEMG wristband that is easily adaptable to behaviorally and physiologically diverse participants. Using this wristband we unblock large-scale data collection (greater than 6,000 consenting participants) and use this data to train deep neural network decoding models on three tasks (discrete gesture classification, 1D wrist-based continuous navigation, and handwriting). We find that the decoders' ability to generalize to new participants scales with the number of participants' data in the training set (Fig. 1c-e), following previously reported deep neural network scaling laws for large language models5. Upon analyzing the decoders' internal representations, we find that they learn abstractions of the sEMG signal that are invariant to differences between participants and sessions (Fig. 1f-h). We then deploy these decoders to unseen participants in closed-loop tasks and achieve a median performance of 0.9 gesture detections per second in a discrete gesture task, 0.5 target acquisitions per second in a continuous control task, and handwriting at 18 words-per-minute. To our knowledge this is the first demonstration of performant out-of-the-box generalization across users for a neuromotor interface.

1-040. Highly accelerated high spatiotemporal resolution whole-brain fMRI with deep learning reconstruction

Mehmet Akcakaya[1,2] AKCAKAYA@UMN.EDU
Burak Demirel[1] DEMIR035@UMN.EDU
Luca Vizioli[1] LVIZIOLI@UMN.EDU
Steen Moeller[1] MOELL018@UMN.EDU
Essa Yacoub[1] YACO0006@UMN.EDU
Kamil Ugurbil[1] UGURB001@UMN.EDU

[1]University of Minnesota
[2]Center for Magnetic Resonance Research

Functional MRI (fMRI) is a widely-employed non-invasive imaging tool for studying brain function in many neuroscientific studies targeting the human brain. In such studies, improving temporal resolution of fMRI time series while maximizing coverage and spatial resolution is desirable, which in turn necessitates fast MRI techniques. Current methods rely on linear parallel imaging methods, which exhibit noise amplification and residual artifacts

at high accelerations. This hinders acquisition of high-resolution fMRI data required for studying mesoscopic scale neuronal organizations. In the broader MRI community, physics-driven deep learning (PD-DL) reconstruction has been proposed to improve acceleration rates. However, its application to fMRI has been limited due to lack of ground-truth training data at high acceleration rates, and lack of studies investigating its impact on subsequent fMRI analyses. In this study, we develop a PD-DL reconstruction for high spatiotemporal resolution whole-brain fMRI acquired at twice the acceleration of the state-of-the-art 7T fMRI scans (20-fold versus 10-fold). Our PD-DL algorithm is trained in a self-supervised manner without reference data. The quality of our reconstruction is first verified using retrospective 20-fold acceleration on the conventional 10-fold accelerated acquisitions. Subsequently a high spatiotemporal resolution whole-brain fMRI dataset (1.6mm isotropic resolution, TR=643ms. 3 runs of 12s on/off face localizer and 6 runs of 12s on/off block design face detection task paradigms) is acquired with prospective 20-fold acceleration, and reconstructed with the same PD-DL approach. The results show visibly superior reconstructions compared to parallel imaging, while subsequent fMRI analyses clearly identify four of the most commonly studied face areas (bilateral occipital and fusiform face areas), which are not visible with standard parallel imaging reconstructions due to high noise amplification in these regions. These findings show the promise of our PD-DL reconstructions for achieving very high accelerations in high-resolution fMRI, and potential for enabling functional mapping at unprecedented details.

1-041. Identifying the impact of local connectivity features on network dynamics

Yuxiu Shao[1,2] IVYEROSION@GMAIL.COM
David Dahmen[3] D.DAHMEN@FZ-JUELICH.DE
Stefano Recanatesi[4] STEFANO.RECANATESI@GMAIL.COM
Eric Shea-Brown[5] ETSB@UW.EDU
Srdjan Ostojic[6] SRDJAN.OSTOJIC@ENS.FR

[1]Beijing Normal University
[2]School of Systems Science
[3]Inst. for Neuroscience and Medicine (INM-6), Julich Research Centre, Julich, Germany
[4]Technion
[5]University of Washington
[6]Ecole Normale Superieure

Understanding how connectivity structure shapes network dynamics is paramount in the field of neuroscience. Theoretical investigations of multi-population neuronal networks often consider statistically homogeneous populations and incorporate either only the population-averaged mean or i.i.d. fluctuations in synaptic couplings. A newly released synaptic physiology dataset highlighted the strong presence of motifs – specific connectivity patterns between pairs and triplets of neurons–beyond the scope of mean connectivity[1]. However, it is a priori not clear which of the experimentally identified connectivity motifs exert a strong influence on neural dynamics. While most previous works focused on reciprocal motifs, here we show that another feature of connectivity, chain motifs, has a much stronger impact on the dynamics of neural activity. We compared the effects of chain and reciprocal motifs within two-population excitatory-inhibitory networks using an analytical framework that approximates the connectivity in terms of low-rank structures that incorporate motifs. We mathematically derived the dominant eigenvalues and exploited matrix perturbation theory to determine the statistics of corresponding eigenvectors. We then used these results to perform a low-rank approximation that predicts the effects of connectivity motifs on linear network dynamics. Our results show that chain motifs have a much stronger impact on dominant eigenmodes than reciprocal motifs[2,3]. Moreover, an overrepresentation of chain motifs induces an additional eigenmode with an eigenvalue of sign opposite to the dominant one, thus modifying the network's effective rank. This additional eigenmode substantially influences network dynamics, offering a new perspective on how local EI motifs shape the network's excitability. Our exploration of the physiological connectivity dataset for the first time revealed the significant impact of EI chain motifs on altering the network's effective rank, permitting the discovery of richer dynamics associated with these specific connectivity motifs.

1-042. LFP transient events in macaque subcortical areas reveal network coordination across scales and structures: a simultaneous fMRI-electrophysiology study

Michel Besserve[1] MICHEL.BESSERVE@TUEBINGEN.MPG.DE
Shervin Safavi[2,3] NEUROPRINCIPLIST@GMAIL.COM
Bernhard Scholkopf[1] BERNHARD.SCHOELKOPF@TUEBINGEN.MPG.DE
Nikos Logothetis[4] NIKOS.LOGOTHETIS@TUEBINGEN.MPG.DE

[1]Max Planck Institute for Intelligent Systems
[2]TU Dresden
[3]Faculty of Medicine
[4]International Center for Primate Brain Research, Institute of Neuroscience, Chinese Academy of Scien

Local field potentials (LFP) are rich mesoscopic signals for studying multi-scale coordinated dynamics in the brain. Indeed, activities identifiable by their LFP signature, such as sharp-wave ripples (SWR), are often a marker of key brain phenomena occurring on a broad range of scales. For instance, SWRs have been shown to co-occur with other markers of coordination across scales: macroscopic (whole-brain) up/down-regulation in thalamocortical system, as well as microscopic coordination at the level of individual spiking activity and synaptic dynamics. In spite of the importance of identification of such characteristic neural activity (neural event) in LFPs, unsupervised methodologies beyond SWR remain elusive.

Thus, we developed an unsupervised method for single-channel LFPs, and, along with SWRs themselves, we found five radically new event types, which are as prevalent, and potentially as important as SWRs. We found these various events to have characteristic signatures across almost all scales. They have distinct *meso-scale* spectral profiles in LFP signals (that we use to identify them), and also, unique brain-wide macro-scale profiles. Using concurrent electrophysiology-fMRI recording, we showed that our neural events are associated with characteristic up/down-regulation of the thalamic and cortical BOLD signals. Lastly, through realistic biophysical simulations, we demonstrated micro-scale signature of neural events. We showed our newly identified neural events co-occur with distinct cellular activities. Overall, our results suggest neural events can be a window for gaining a multi-scale understanding of brain dynamics.

1-043. Bayesian bi-clustering of neural spikes by latent structures

Ganchao Wei[1,2] WEIGANCHAO@GMAIL.COM

[1]Duke University
[2]Department of Statistical Science

Modern neural recording techniques allow us to simultaneously obtain spiking activity of neurons from multiple regions over long time periods. This requires new statistical methods to be developed for understanding structure of the large-scale data, in terms of both neuron number and recording duration. Here, we develop a bi-clustering method to cluster the neural spiking activity spatially and temporally, according to their low-dimensional latent structures. The spatial (neuron) clusters are defined by the latent trajectories within each neural population, while the temporal (state) clusters are defined by synchronous local linear dynamics shared across different periods. To flexibly extract the bi-clustering structure, we build the model in a Bayesian nonparametric way, and develop an efficient Markov chain Monte Carlo (MCMC) algorithm to sample the posterior distributions of model parameters. Validating our proposed MCMC algorithm through simulations, we find the method can recover unknown parameters and true bi-clustering structures successfully. We then apply the proposed bi-clustering method to multi-regional neural recordings under different experiment settings (Neuropixels data), where we find that simultaneously considering latent trajectories and spatial-temporal clustering structures can provide us with a more accurate and interpretable result. Overall, the proposed method provides scientific insights for large-scale neural spiking data with elongated recording periods, and it can have application beyond neuroscience.

1-044. Probing neural representations of language in the human right hemisphere

Felix Waitzmann[1,2] FELIX.WAITZMANN@TUM.DE
Laura Schiffl[3] LAURA.SCHIFFL@TUM.DE
Lisa M. Held[3] LISA.HELD@TUM.DE
Bernhard Meyer[3] BERNHARD.MEYER@MRI.TUM.DE
Jens Gempt[4] J.GEMPT@UKE.DE
Simon N. Jacob[3] SIMON.JACOB@TUM.DE
Julijana Gjorgjieva[1,5] GJORGJIEVA@TUM.DE

[1]Technical University of Munich
[2]Computation in Neural Circuits Group
[3]Klinikum rechts der Isar, TUM
[4]Universitatsklinikum Hamburg-Eppendorf
[5]School of Life Sciences

The human language system comprises multiple processing stages across multiple brain regions typically lateralized to the left cerebral hemisphere. In the event of stroke or other injury to these systems, neural circuits in the brain reorganize in an attempt to restore lost function. However, the neural processes and learned representations underlying language network reorganization remain unclear. Here, we present data from a first-of-its-kind human case study involving chronic intracortical recordings from frontal and parietal cortices in the right cerebral hemisphere following loss of left hemisphere function due to stoke. Multi-electrode arrays were implanted in regions of the language network known from the left hemisphere to be pertinent to language comprehension, concept retrieval, and speech preparation to study to which extent networks in the right hemisphere assume new language-relevant tasks. Neuronal activity was recorded while the subject performed linguistic tasks designed to probe language comprehension and generation mechanisms. We found that neural responses in frontal regions correlate with speech preparation and production, consistent with a role in planning future actions. On the other hand, population activity in parietal areas exhibits a transient increase post-stimulus presentation, suggesting an assumed role in language comprehension. Analysis of response variability between correct and incorrect trials exposed intriguing differences. Words with a high degree of response accuracy exhibited strong decorrelation on correct trials. In animal studies, decorrelation has been linked to attentional processes and optimal arousal states, hinting that our data captures better network control for certain words. An intriguing possibility is that response confidence, or ease of word production, may be linked to a decorrelated network state. Our results thus expose a remarkably rich and interesting dataset exploring language function at both a single-neuron and neural population level while pointing to distinguishable roles of different brain regions in controlling rehabilitated speech in an adult human.

1-045. Diverse signals mediate feedforward and feedback communication in early visual cortex

Francesca Mastrogiuseppe[1] FRAN.MASTROGIUSEPPE@GMAIL.COM
Joana Carmona[2,3] JOANA.CARMONA@RESEARCH.FCHAMPALIMAUD.ORG
Byron Yu[4] BYRONYU@CMU.EDU
Adam Kohn[5] ADAM.KOHN@EINSTEINMED.EDU
Christian Machens[1] CHRISTIAN.MACHENS@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Research
[2]Champalimaud Foundation
[3]Champalimaud Research, Champalimaud Neuroscience Programme
[4]Carnegie Mellon University
[5]Albert Einstein College of Medicine

Brain computations rely on the interactions among large populations of neurons across distinct cortical areas. Recent work indicates that those interactions are low-dimensional, and therefore can be well described in terms of latent signals. However, the computational nature of the signals that travel along the cortical hierarchy in the feedforward and feedback directions has largely remained elusive so far. In this work, we used simultaneously-recorded electrophysiological data from non-human primates to characterize the dynamics, directionality, and computational role of the signals mediating communication between neurons in visual cortical areas V1 and V2. To this end, we employed a novel statistical approach based on dimensionality reduction that characterizes communication between pairs of areas by extracting, via a linear dynamical model, the latent signals expressed in one area that predict trial-to-trial activity in the other one. In agreement with previous work, we found that latents mediating inter-area interactions between V1 and V2 are low-dimensional - much lower-dimensional than those mediating within-area interactions. Among inter-area latents, we identified two major types of signals. The first

type encoded stimulus-related trial-to-trial variability, and mirrored the dynamics of the visual stimulus (drifting gratings). The second type was related instead to changes in cortical state: it evolved on slower timescales (100-400 ms), and expressed global trial-to-trial variability that is positively correlated across a large majority of V1 and V2 neurons. We investigated whether those signals were simultaneously expressed in V1 and V2 or preferentially traveled from one area to the other one. We found that the two different types of signals traversed the cortex in opposite directions, with stimulus-related signals mostly moving in the feedforward direction (V1→V2), and global signals in the feedback one (V2→V1). Our findings indicate that signals with different dynamics, directionality and behavioural relevance contribute to shared variability in multi-area cortical circuits.

1-046. Enumerating and discovering discriminative tasks for probing diverse foraging strategies

Tzuhsuan Ma[1,2] MAT@JANELIA.HHMI.ORG
Rishika Mohanta[3] RMOHANTA@ROCKEFELLER.EDU
Aparna Dev[1] DEVA@JANELIA.HHMI.ORG
Glenn Turner[1] TURNERG@JANELIA.HHMI.ORG
Ann Hermundstad[4] HERMUNDSTADA@JANELIA.HHMI.ORG

[1]HHMI Janelia Research Campus
[2]Theory & Computation
[3]The Rockefeller University
[4]HHMI Janelia

Foraging, an indispensable behavior for survival, consists of long sequences of searches, encounters, and decisions. To forage successfully, animals are thought to leverage the statistical regularities and dynamical rules of their habitats to maximize long-term utility. Since animals encounter different habitats that demand different decision rules, it is important to infer the behavioral strategies that are most relevant to a particular species. One approach is to observe an animal in its natural habitat, but this comes with technical and conceptual challenges of recording and manipulating behavior in naturalistic settings. To circumvent this without compromising the richness of environmental features that evoke an animal's foraging strategy, we sought to manipulate a vast set of such features in a controlled lab setting. To this end, we designed a two-choice foraging task with complex contingencies in reward delivery (controlled by up to 13 past decisions). By enumerating different reward-delivery rules, we simulated half a million different task conditions—each resembling a slightly different environment and varying in their putative relevance to the animal. As a proof of principle, we tested these tasks on fruit flies. We selected tasks that best discriminate between two classes of strategies: one that requires "one-shot memory" (like Boolean logic), and one that does not. Even though flies' decisions are highly stochastic from trial to trial, we identify flies that use one-shot memory to perform well, a finding that cannot be explained by rare sampling events from any "memoryless" strategies. This finding suggests that flies can exploit short-timescale decision rules in a manner that differs from longer-timescale adaptive behaviors mediated by synaptic plasticity. Our framework is agnostic to specific model systems and can flexibly perform inferences in different hypothesized strategy spaces. This will allow us to compare foraging strategies across species and studying their dependence on underlying environmental structure.

1-047. Highly-connected subnetworks of neurons in mouse visual cortex dominate visual processing

Bradley Akitake[1,2] BRADLEY.AKITAKE@NIH.GOV
Remy Yovanno[3] REMY.YOVANNO@NIH.GOV
Connor Phillips[1] CONNOR.PHILLIPS@NIH.GOV
Nina Friedman[4] NINA.FRIEDMAN@NIH.GOV
Paul LaFosse[4] PAUL.LAFOSSE@NIH.GOV
Jonathan O'Rawe[1] JONATHAN.O'RAWE@NIH.GOV
Yanting Deng[1] YANTING.DENG@NIH.GOV
Zhishang Zhou[1] ZHISHANG.ZHOU@NIH.GOV
Victoria Scott[1] VICTORIA.SCOTT@NIH.GOV
Mark Histed[5,6] MARK.HISTED@NIH.GOV

[1]National Institute of Mental Health
[2]Unit on Neural Computation and Behavior
[3]National Institute of Mental Health; Harvard Medical School
[4]National Institute of Mental Health; University of Maryland College Park
[5]NIH

[6]NIMH

Neural responses in sensory cortex support representations of the world that are used to guide behavior. How do local recurrent networks create and shape these representations? Within each cortical area, excitatory neurons are highly interconnected, with a majority of inputs to any excitatory cell arising from other local excitatory neurons (Braitenberg and Schuz, 2013). Here, we examine recurrent network computation using two-photon optogenetics. We hypothesize that certain V1 input patterns are amplified and others are suppressed. Such local recurrent amplification may structure the latent space ('intrinsic manifold') of cortical responses. Using two-photon stimulation and imaging in layer 2/3 of mouse V1, we employed a compressed sensing approach (3 cells/second), to rapidly determine how single cells responded to input. We identify a subset of "high-connectivity" neurons whose optogenetically evoked activity is accompanied by greater than average off-target activation ($>$2-5% ΔF/F). Grouped stimulation of these cells generated amplified network responses. We constructed a "most excitatory" group of cells with the highest field-evoked activity (off-target $\sum \Delta$F/F). This pattern produced 2.5-times larger network responses versus selecting cells with the highest optogenetic responses alone; and 15-times larger than patterns designed to be "least excitatory". While visual input was not used to select neurons (see Marshel et al., 2019), highly-connected neurons overlapped strongly with neurons most responsive to natural visual stimuli. In contrast, neurons that were most responsive to optogenetic input were not visually responsive, and produced little network response when stimulated. Thus, we find that some neurons in the visual cortex are part of a highly-connected subnetwork that is amplified by input to many neurons but weakly responsive to single-cell stimulation. The highly-connected subnetwork is used in visual processing. The pool of less-connected neurons may provide room for learning, or may be used for other, non-visual purposes.

1-048. Spatiotemporal modeling of cortical cholinergic and calcium signaling dynamics across learning

Josue Ortega Caro[1,2] JOSUEORTC@GMAIL.COM
Sweyta Lohani[1] SWEYTA.LOHANI@YALE.EDU
David van Dijk[1] DAVID.VANDIJK@YALE.EDU
Jessica Cardin[1] JESS.CARDIN@YALE.EDU

[1]Yale University
[2]Neuroscience

Neuromodulatory systems, such as acetylcholine and norepinephrine, profoundly influence cortical functions including attention, perception, and cognition. Recent investigations revealed the state-dependent and spatiotemporally complex cholinergic signaling that exhibits variable coupling with cortical neural activity [1], suggesting the potential for plasticity. However, due to the high-dimensional nature of both cholinergic and neural signaling, the precise patterns of cholinergic influence on cortical activity remain unclear. To identify the role of cholinergic signaling in learning and performance of sensory-guided tasks, we used mesoscopic imaging to record cholinergic release (ACh3.0 sensor [2]) and neural calcium (Ca) signaling (jRCaMP1b [3]) longitudinally in animals as they learned to perform a visual contrast detection task (Fig. 1). We find distinct changes in cortical ACh and Ca dynamics during learning, suggesting coordinated plasticity (Fig. 1D,E). Understanding this multimodal system requires robust computational methods capable of handling diverse dynamics. We therefore developed the Multimodal Transformer (MMT), a machine learning model trained through self-supervised learning [4]. MMT adeptly reconstructs ACh, Ca, and behavioral dynamics from partially observed information (Fig. 2A). Furthermore, MMT surpasses single-modality transformers and other machine-learning models, exhibiting superior predictive accuracy for calcium dynamics in new mice (Fig. 2B) and trial-by-trial performance (Fig. 2C,D). Employing complete trial dynamics is essential for achieving high trial-by-trial predictability (Fig. 2E). We further developed a novel computational approach designed to extract spatiotemporal motifs from nonlinear embedding machine learning model (Fig. 3A). To validate this method, we use simple recurrent neural network system dynamics (Fig. 3B,C) and then apply the model to real data. Notably, our model discerns differences between recordings of naive and expert mice, highlighting longitudinal changes in within-trial dual-color dynamics (Fig. 4). Specifically, the model identifies increased activity in visual cortex pre and post-trial and increase in somatosensory and motor cortex post-trial, furthermore, this pattern decreases in higher contrast trials.

1-049. Fast behavioral learning with an imprecise hippocampal code on a dynamic, multi-step linear maze

John Widloski[1] JOHNWIDLOSKI@BERKELEY.EDU
Heike Stein[2] HEIKE.C.STEIN@GMAIL.COM
Jared Collina[3] JARED.COLLINA@PENNMEDICINE.UPENN.EDU
David Foster[1] DAVIDFOSTER@BERKELEY.EDU

[1]Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley
[2]Ecole Normale Superieure
[3]Neuroscience Graduate Group, University of Pennsylvania, Philadelphia, PA

Goal-directed navigation in complex environments relies on the inference of self-localization together with learning the environment's behavioral affordances as a prerequisite for memory-dependent spatial decision making. It is classically assumed that hippocampal place cell coding underlies this process, yet experimental support for this hypothesis is scattered and often conflicting. To address this, we developed a novel task that requires rats to learn a precise mapping between location and action, by navigating a sequence of occluded left-right decision points arranged on a linear track. Decision points were reconfigured unpredictably across sessions. Rats were trained to run the maze repeatedly while neural activity was recorded from the dorsal hippocampus. The task had two phases: recall of the previously learned sequence and learning of a new one. Recall was poorest in the middle of the maze, exhibiting an asymmetric inverted U-shape characteristic of primacy-recency effects in serial learning. Intriguingly, place fields exhibited ambiguous tuning across boxes that matched this U-shaped performance curve. During the learning phase, animals were forced to update the mapping between spatial contexts and the decision pattern. In principle, the task allows one-shot re-learning, given correct location inference. Yet, rats needed 10-20 repetitions to learn new sequences and learning was coupled across boxes. To understand why one-shot learning was not observed and how it might depend on ambiguous spatial coding, we modeled performance with three different Rescorla-Wagner models: (1) Actions are selected and rewarded independently across boxes; (2) Action selection is "chunked" across boxes, with a scalar reward inversely related to the number of errors but agnostic to where they occurred; (3) Same as (2) but with reward propagation to unchosen but similar sequences. Only the last model captured qualitative aspects of animal behavior. Our findings demonstrate a novel and unexpected link between spatial learning and hippocampal spatial codes and indicate that rats exploit structural knowledge about the environment to guide learning.

1-050. Inferring feedback signaling from the temporal evolution of representational geometries.

Abhimanyu Pavuluri[1,2] ABHIMANYU.PAVULURI@EINSTEINMED.EDU
Adam Kohn[1] ADAM.KOHN@EINSTEINMED.EDU

[1]Albert Einstein College of Medicine
[2]Neuroscience

Cortical feedback is poorly understood, especially with respect to its role in shaping sensory representations. Feedback interactions between macaque visual areas V1 and V2 have been shown to use distinct communication channels from feedforward interactions. However, the purpose of this organization is unclear. Here we assess the role of these connections in defining the geometry of neuronal population representations for textures. We find that low dimensional representations of textures emerge in V2 population activity before V1, although spiking activity in V1 precedes that in V2. To understand the role of recurrent (within area) and between area circuits in building this representation, we fit regression models and recurrent neural network models to the measured population responses. Regression models show that recurrent and feedback inputs to V1 reside in distinct neuronal subspaces, whereas feedforward and recurrent inputs in V2 occur in more aligned subspaces. Analysis of recurrent neural networks suggests that the feedforward input from V1 to V2 is important for compressing the dimensionality of population responses, whereas feedback signals from V2 to V1 are important for shaping the low dimensional representation in V1. Our work suggests that studying the dynamics of neuronal population representation geometries can reveal distinct contributions of recurrent, feedforward and feedback signals in building visual representations.

1-051. When is a non-canonical olfactory system optimal?

Caitlin Lienkaemper[1,2] CLIENKAEMPER.PSU@GMAIL.COM
Meg A Younger[1] MYOUNGER@BU.EDU
Gabriel Ocker[1,2] GKOCKER@BU.EDU

[1]Boston University
[2]Mathematics and Statistics

The early olfactory system is classically described by a canonical "one receptor, one neuron, one glomerulus" organization: each olfactory receptor neuron (ORN) expresses one type of olfactory receptor (OR) and all neurons expressing the same receptor project to a shared glomerulus (Chess et al., 1994). This model roughly describes the early olfactory systems of Drosophila melanogaster, Mus musculus, and Homo sapiens, as well as other species. It is, however, not universal. The mosquito Aedes aegypti expresses at least twice as many ORs as it has glomeruli (Bohbot et al, 2007), so each glomerulus must pool inputs driven by multiple ORs. Herre et al. (2022) found that individual Aedes aegypti ORNs coexpress multiple ORs.

Why does A. aegypti violate the canonical model of early olfaction? We take a normative approach, assuming that the glomerular inputs encode olfactory stimuli in a noisy environment. We analytically compute the patterns of receptor expression which maximize the mutual information between olfactory stimuli and glomerular inputs under a resource constraint for OR expression. We find that pooling ORs for correlated odorants is optimal when the olfactory signal is weak or the olfactory environment is noisy, relative to neural responses.This echoes the classic result that the top principal components maximize mutual information in Gaussian environments (Linsker, 1988). On the other hand, selecting one high-variance receptor in each glomerulus is optimal when the olfactory signal is strong or neural responses are unreliable. By training a feedforward network to perform an odor classification task, we show that this pattern persists in the presence of response nonlinearities. Confirming our predictions, D. melanogaster ORNs coexpress two receptors, Or33c and Or85e, whose ligands are correlated across natural stimuli (Dweck et al, 2016). Our work suggests that the organization of early olfactory systems reflects the statistics of the ethologically relevant olfactory environments experienced by different species.

1-052. Reverse-engineering biological limb motor control using a neuromechanical model of Drosophila

Sibo Wang-Chen[1] SIBO.WANG@EPFL.CH
Victor Alfred Stimpfling[1] VICTOR.STIMPFLING@EPFL.CH
Pavan Ramdya[2] PAVAN.RAMDYA@EPFL.CH

[1]EPFL
[2]Ecole Polytechnique Federale de Lausanne

Animals are capable of highly efficient, adaptive, and robust behaviors unmatched by even the most advanced robots. These unfold through the dynamics of lower motor centers — the spinal cord in vertebrates and the ventral nerve cord (VNC) in insects — which carry out important neural computations to orchestrate movements of the body and appendages. However, how the organization of lower motor circuits give rise to these computations and adaptive limb movements remains largely unknown in any species. Here, we are discovering critical circuit motifs within the VNC of the fly, Drosophila melanogaster, by building connectome- and fly behavior-constrained artificial neural networks (ANNs). Initial connectivity of ANNs are informed by Drosophila VNC connectomes — synapse-resolution wiring diagrams reconstructed using electron microscopy. The outputs of ANNs then drive movements of NeuroMechFly, a morphologically realistic biomechanical model of the adult fly, which was reconstructed from micro-CT scans of a real animal and embedded in a physics simulator. The parameters of the neural network model are then optimized through imitation learning, a process in which we match the behavior of the simulated fly with experimental recordings of real, freely walking flies. This novel modeling approach allows us to study the behavioral outcome of ANN dynamics in closed loop, taking into account the biomechanics of the body and the physics of the surrounding environment. This ongoing work will ultimately allow us to reverse-engineer Drosophila limb motor control by leveraging a digital twin at whole-organism scale, constrained by experimental data including neural network topology, biomechanics, behavioral output, and neural activity.

1-053. Neuron-Astrocyte Associative Memory

Leo Kozachkov[1,2] LEOKOZ8@MIT.EDU
Jean-Jacques Slotine[1] JJS@MIT.EDU
Dmitry Krotov[3] KROTOV@IBM.COM

[1]Massachusetts Institute of Technology
[2]Brain and Cognitive Sciences
[3]MIT-IBM Watson AI Lab & IBM Research

Astrocytes, a unique type of glial cell, are thought to play a significant role in memory due to their involvement in modulating synaptic plasticity. However, no existing theories explain how neurons, synapses, and astrocytes could collectively contribute to memory function. To address this, we propose a biophysical model of neuron-astrocyte interactions that unifies various viewpoints on astrocyte function in a principled, biologically-grounded framework. A key aspect of the model is that astrocytes mediate long-range interactions between distant tripartite synapses (synapses ensheathed by an astrocyte process). This effectively creates “multi-neuron synapses" where more than two neurons interact at the same synapse. Such multi-neuron synapses are ubiquitous in models of Dense Associative Memory (also known as Modern Hopfield Networks) and are known to lead to superlinear memory storage capacity, which is a desirable computational feature [**?**]. We establish a theoretical relationship between neuron-astrocyte networks and Dense Associative Memories and demonstrate that neuron-astrocyte networks have a larger memory storage capacity per compute unit compared to previously proposed biological implementations of Dense Associative Memories. This theoretical correspondence suggests the exciting hypothesis that memories could be stored, at least partially, within astrocytes instead of in the synaptic weights between neurons.

1-054. Persistent activity bump on a ring without a continuous ring attractor

Memming Park[1,2] MEMMING.PARK@RESEARCH.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Centre for the Unknown

We propose a robust recurrent memory mechanism for integrating and maintaining head direction without a ring attractor. The ring attractor is a popular theoretical mechanism for maintaining a continuous internal representation of a periodic variable. For instance, they are used to describe the neural dynamics underlying head direction cells, grid cells, and visual working memory. At the same time, it is well-known that mathematically, continuous attractors are prone to fine tuning, where small changes in the connectivity or biophysical parameters lead to the destruction of the mechanism. We describe an alternative, robust mechanism for working memory of a periodic variable using the theory of stable limit cycles. In particular, we detail a biophysical implementation that can explain the bump of activity observed over the ellipsoid body of insects. The content of working memory is stored in the relative phase difference between two independent oscillators, whose periods are constrained by conduction delay. A phase independent readout is achieved through a delay line and nonlinear coincidence detectors that can be mapped onto the readout system, i.e., the ellipsoid body. Our theory makes several strong predictions: the bump of activity on the ellipsoid body is tightly periodic on a short time scale, is robust to partial lesions, and exhibits small, consistent systematic drifts due to misadjustments within each animal. Furthermore, we investigate the impact of diffusion in information encoding and memory due to various noise sources.

1-055. Endocannabinoid modulation of stimulus-driven projection-specific neuronal activity in the prefrontal cortex

Saptarnab Naskar[1,2] SAPTARNABN@GMAIL.COM
Luis E Rosas-Vidal[1] LUIS.ROSASVIDAL@NORTHWESTERN.EDU
Farhana Yasmin[1] FARHANA.YASMIN@NORTHWESTERN.EDU
Sachin Patel[1] SACHIN.PATEL@NORTHWESTERN.EDU

[1]Northwestern University
[2]Psychiatry and Behavioral sciences

Endocannabinoids including 2-Arachidonoylglycerol (2-AG) are lipid neurotransmitters that regulate presynaptic strength and mediate forms of long-term synaptic plasticity. Endocannabinoid dependent neuromodulation is expressed in many limbic nodes including the prefrontal cortex (PFC). Neurons in the PFC, a brain area involved in working memory, decision making, goal-directed behaviors, send projections to brain areas that are implicated in both appetitive, consummatory behaviors (such as Ventral Tegmental Area (VTA)) and aversive, escape behaviors (such as Periaqueductal Grey (PAG)). This study investigates how modulating the 2-AG signaling can in

turn lead to changes in ethologically relevant, innate behavioral outputs and the neural representation of environmental stimuli that trigger divergent behavioral responses. We show that increased endocannabinoid signaling can elevate innate fear expression while also increasing exploratory drive for food and social interaction. These behavioral changes are associated with altered activity in PFC-PAG and PFC-VTA neuronal sub-populations. Our study provides an important framework in understanding 1) the relative specificity/overlap of behavioral representations in PFC neuronal subpopulations and 2) differential modulation of circuit-specific neuronal activity through the endocannabinoid signaling pathway.

1-056. Dynamical changes of attractor landscapes during reinforcement learning in macaque prefrontal cortex

Siyu Wang[1,2] WANGXSIYU@GMAIL.COM
Yuan Zhao[2] YUAN.ZHAO@NIH.GOV
Ramon Bartolo[1] RAMON.BARTOLOOROZCO@NIH.GOV
Francisco Pereira[1] FRANCISCO.PEREIRA@NIH.GOV
Bruno Averbeck[1] AVERBECKBB@MAIL.NIH.GOV

[1]National Institutes of Health
[2]National Institute of Mental Health

In reinforcement learning (RL), animals can flexibly change their choice preference as they learn through trial and error. Theoretical work suggests that attractor dynamics in networks can account for the choice process. Attractor models operationalize the idea that decision-related neural activity evolves in a way that can be modeled as an object frictionally sliding on a landscape with two basins separated by a hill. The two attractor basins represent the two alternative choices in a binary decision-making task. Artificial network simulations predict that as animals learn, the attractor basins get deeper and the energy barriers between the two choices become higher, which reflect a learned preference for one option over another. Despite the theory, there is little empirical evidence directly examining the development of attractor energy landscapes during learning. In this work, we fill this gap and explicitly examine this hypothesis in neural data. We simultaneously recorded hundreds of neurons using high-channel-count multielectrode arrays in the dorsal lateral prefrontal cortex (dlPFC) of rhesus monkeys while they performed a variant of the two-armed bandit RL task. We analyzed population dynamics in the state space spanned by the firing rates of individual neurons, by estimating low-dimensional linear and non-linear dynamics using two different methods. With converging results from both methods, we provide neural evidence that attractor basins get deeper as animals learn.

1-057. A unifying normative model of decision confidence

Amelia Johnson[1] AMELIAMJ@UW.EDU
Michael Buice[1] MICHAELBU@ALLENINSTITUTE.ORG
Koosha Khalvati[2] KOOSHA.KHALVATI@ALLENINSTITUTE.ORG

[1]Allen Institute + University of Washington
[2]Allen Institute

Self-assessment of one's choices, i.e., confidence, is the topic of many decision neuroscience studies. Mostly focused on in perceptual decision-making studies, confidence is only mathematically defined for scenarios where different choices have the same potential reward. In these situations, the goal is to pick the "correct" choice as opposed to all other equally "incorrect" ones based on some perceptual cues. Consequently, confidence is defined as the probability of choosing the correct option [Pouget et al., 2016, Khalvati et al., 2021]. In these setups, confidence in perception and confidence in decision are equivalent measures. In the real world, though, different actions lead to different potential rewards, and this variation in reward could influence confidence in decisions, differentiating it from perception confidence. Notably, most experiments where different choices vary in value, known as value-based decision-making, do not involve uncertainty in perception. These works are mostly about memory retrieval, and do not study the interaction between value and perceptual uncertainty [Shadlen and Shohamy, 2016, Brus et al., 2021]. Here we present a normative framework to formally define and assess decision confidence in a general decision-making scenario involving uncertainty about the perception, and variation in choice prior and reward. Particularly, this framework defines confidence as the probability of making the best decision. We further show that our model, which is a planning as inference framework, represents the confidence of a Bayesian optimal observer with maximum entropy Reinforcement Learning objective function [Botvinick and Toussaint, 2012, Levine, 2018]. We tested our framework and other alternative hypotheses for confidence in two experiments. The superiority of our method in explaining the confidence report in both experiments validates our approach and provides evidence in favor of the implementation of planning as inference in the brain. The latter

is especially interesting as there has been substantial evidence supporting the capability of the brain to perform Bayesian inference. Therefore, the same mechanism could be used for decision-making in the brain.

1-058. Deep-layer projection neurons develop representations of perceptual categories and behavioral choice

Nathan Schneider[1,2] NAS290@PITT.EDU
Michael Malina[3] MMALINA@ANDREW.CMU.EDU
Rebecca Krall[1] RFK12@PITT.EDU
Ross Williamson[1] ROSS.S.WILLIAMSON@PITT.EDU

[1]University of Pittsburgh
[2]Otolaryngology
[3]Carnegie Mellon University

Auditory-guided behavior is a fundamental aspect of our daily lives, whenever auditory information guides our decisions and actions. Nestled amongst several populations, extratelencephalic (ET) neurons reside in the deep layers of auditory cortex (ACtx) and provide a primary means of routing auditory information to sub-cortical targets associated with decision-making and action. To investigate the behavioral role of L5 ET neurons, we trained head-fixed mice to categorize the rate of sinusoidal amplitude-modulated (sAM) noise bursts as either fast or slow to receive a water reward. We then used two-photon calcium imaging alongside selective GCaMP8s expression to monitor the activity of L5 ET, as well as layer (L) 2/3 and 5 intratelencephalic (IT) populations. Clustering analyses of these distinct populations revealed heterogeneous responses that correlated with various stimulus and task variables. One motif, present in L5 ET neurons, corresponded to “categorical” firing patterns (i.e., preference for slow or fast sAM rates). This categorical selectivity was not present in untrained mice, and longitudinal recording revealed that L5 ET neurons dynamically shifted their responses across training to reflect these learned categories. This categorical selectivity was stronger during active behavioral engagement as compared to selectivity derived from passive listening on the same day, suggesting top-down modulatory input. L5 ET population activity also reflected behavioral choice, regardless of stimulus identity or reward outcome. Consistent with this observation, behavioral choice could be robustly predicted from L5 ET activity. Choice activity preceded motion onset, emphasizing that these signals were separate from motor activity. The L2/3 population contained notably less information about choice and categorical selectivity while L5 IT neurons contained some categorical and choice information early in training which then degraded across learning. These results suggest that ACtx projection neuron sub-types differentially encode behaviorally relevant stimuli, emphasizing the divergent pathways from ACtx and their contributions to auditory-guided behavior.

1-059. The functional role of diverse cortical responses for network dynamics and task performance

Jade Toth[1,2] JACKMTOTH@PITT.EDU
Badr Albanna[1] BADR.ALBANNA@PITT.EDU
Michele Insanally[1] MNI@PITT.EDU

[1]University of Pittsburgh School of Medicine
[2]Otolaryngology, Neurobiology

Spike trains recorded from the cortex of behaving animals can be complex, highly variable from trial to trial, and therefore challenging to interpret. A fraction of cells exhibit obvious trial-averaged, task-related responses such as pure tone frequency tuning in auditory cortex. However, a substantial number of cells do not appear to fire in a task-related manner and are often neglected from analysis. Previous work used a novel single-trial, spike-timing-based analysis to show that both classically and non-classically responsive cortical neurons contain significant information about sensory stimuli and behavioral decisions suggesting that non-classically responsive cells may play an underappreciated role in perception and behavior. A recent study presented at Cosyne 2022 introduced a novel, task-performing spiking recurrent neural network (RNN) model incorporating excitatory and inhibitory spike-timing-dependent plasticity (STDP) that successfully recapitulates the distribution of classically and non-classically responsive neurons measured from the cortex of behaving animals. Here, we leverage this model to explore the synaptic origin and functional contribution of heterogeneous response profiles to network performance and dynamics. To validate our model, we demonstrated that local patterns of synaptic connectivity which predict the response properties of network units in silico also predicted the spiking response properties of auditory cortical neurons from in vivo whole-cell recordings during behavior. Detailed inactivation experiments revealed that classically responsive and non-classically responsive model units contributed to task performance via output and recurrent connections, respectively. To evaluate how classically and non-classically responsive

units influence network dynamics, we examined how these selective inactivation experiments altered the global dimensionality of network activity. Strikingly, only non-classically responsive units alter network dimensionality suggesting they play a unique role in restricting network dynamics to a task relevant subspace. Our approach provides a powerful framework for understanding how diverse neural responses constrained by plasticity rules shape large-scale neural dynamics and behavior.

1-060. The internal head-direction sense is subjectively anchored to the environment during head fixation

Simon Carrillo Segura[1,2] SC8633@NYU.EDU
Alex Pak[1,2] AP7454@NYU.EDU
Janna Aarse[3] JANNA.AARSE@RUHR-UNI-BOCHUM.DE
Dora Angelaki[1] DA93@NYU.EDU
Andre Fenton[1] AF95@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Ruhr University Bochum

An accurate sense of allocentric direction is crucial for navigation and for reliably registering cognitive maps to environmental features, despite crucially relying on the vestibular system, which is independent of other environmental sensations. Azimuth-tuned head direction cells (HDCs) persist even during sleep, suggesting that strong attractor dynamics organize head-direction neural networks independently of allocentric stimuli. Here, we investigate the vestibular contribution to the direction sense by recording HDCs from the anterior-dorsal nucleus (ADN) of the thalamus. We compare the same ADN cells in mice while they are freely-behaving in a fixed environment or able to locomote during head fixation in an environment that counter moves on an air cushion. This head fixation eliminates changes in vestibular stimulation, while maintaining access to the same types of other stimuli as during free behavior; locomotion creates a vestibular conflict with the other senses. The allocentric azimuthal tuning of HDCs was lost during head-fixed vestibular conflict, whereas the cofiring of HDC pairs was unchanged, indicating HDC discharge remained neurocentrically organized. Indeed, the cofiring maintained a ring manifold in the neuronal population dynamics that did not reliably register to allocentric azimuth because ensemble activity decoded to “neurocentric” azimuths on the manifold, independent of allocentric direction. A continuous-attractor model explains how a dynamically erroneous estimation of angular velocity can produce a neurocentric direction sense that erroneously and episodically registers to allocentric azimuth. Although the ring manifold and allocentric azimuthal tuning maintained during immobility when the mice were free to move, a new neuronal state appeared as off-manifold activity during head fixed immobility. These findings challenge the standard model of the reliable allocentric direction sense resulting from a persistent ring attractor network, and instead emphasize its subjectivity, internal organization, and independence from external stimuli, in addition to demonstrating a novel and distinct state during head-fixed immobility.

1-061. Neural choice-selective sequences across regions align with sequential evidence accumulation models

Lindsey Brown[1] LINDSEYSBROWN@PRINCETON.EDU
Jounhong Ryan Cho[1] JOUNHONG@PRINCETON.EDU
Scott Bolkan[1] SBOLKAN@PRINCETON.EDU
Edward Nieh[2] ENIEH@VIRGINIA.EDU
Manuel Schottdorf[1] MSCHOTTDORF@PRINCETON.EDU
Sue Ann Koay[3] KOAYS@JANELIA.HHMI.ORG
David Tank[1] DWTANK@PRINCETON.EDU
Carlos Brody[1,4] BRODY@PRINCETON.EDU
Ilana Witten[1] IWITTEN@PRINCETON.EDU
Mark Goldman[5] MSGOLDMAN@UCDAVIS.EDU

[1]Princeton University
[2]University of Virginia
[3]Janelia Research Campus
[4]Princeton Neuroscience Institute
[5]University of California, Davis

Accumulating evidence for decision-making is a fundamental cognitive operation. Canonical models have posited that this operation is performed neurally through low-dimensional attractors with persistent ramping activity. In

contrast to predicted ramping, recent studies have observed sequential neural activity across the brain during decision-making, necessitating new models. We develop two new models that accumulate evidence through sequences with different evidence encoding mechanisms. In both models, information is gated to a set of neurons active at a particular position and that information is transferred to the next position. These models differ in their computations within the active position. While the competing chains model accumulates evidence through competition between opposing populations, with the amplitude of different neurons monotonically representing evidence for each choice, the position-gated bump attractor accumulates evidence by shifting a bump of activity in which the neurons active in the bump indicate the evidence level non-monotonically. We tested the diverging predictions of these models in a dataset of over 14,000 neurons across four cortical and subcortical regions, which have choice-selective sequences during an evidence accumulation task. Despite this seemingly similar sequential activity, different regions have different evidence coding. We find that cortex and dorsomedial striatum show largely monotonic evidence tuning, as predicted by the competing chains model, while hippocampus displays non-monotonic evidence tuning, as predicted by the position-gated bump attractor. Furthermore, the monotonic tuning curves in cortex shift from graded evidence representations, required for evidence accumulation, earlier in the task to more step-like choice representations later in the task. Interestingly, while retrosplenial cortex is dominated by these monotonic tuning curves, there is a population of non-monotonically tuned cells in this region, possibly representing input from hippocampus. Together, this analysis across brain regions suggests different mechanistic underpinnings to the observed choice-selective sequences in different regions, aligned with predictions of our models.

1-062. What are large language models mapping to in the brain.

Ebrahim Feghhi[1,2] EFEGHHI@GMAIL.COM
Jonathan Kao[3] KAO@SEAS.UCLA.EDU
Nima Hadidi[4] NHADIDI@G.UCLA.EDU
Idan Blank[4] IBLANK@PSYCH.UCLA.EDU

[1]University of california, Los Angeles
[2]Neuroscience
[3]UCLA
[4]University of California, Los Angeles

Language encoding models predict neural activity during language comprehension tasks from stimulus representations. State-of-the-art performance in language encoding uses learned representations from large language models (LLMs), suggesting similarities between how LLMs and humans process language. However, little progress has been made in the interpretation of these predictive features. Here, we sought to understand how much of the linearly mappable information between LLMs and brains is accounted for by explicitly interpretable feature spaces. Towards this end, we collected a set of interpretable feature spaces that we hypothesized would be useful for language encoding. These consisted of 1) positional signals tracking the comprehender's location in a passage of text, 2) static semantic embeddings derived from GloVe, and 3) surprisal values computed using an LLM. We then examined what fraction of LLM neural predictivity can be explained by these interpretable feature spaces through a variance partitioning approach. We applied our method to an fMRI dataset in which participants read short passages, where both untrained and trained GTP2-large (GPT2) models were reported to display high neural predictivity in an influential prior study. We found that 1) positional signals and random word embeddings (included to account for repeated words in train and test splits) fully account for untrained GPT2 neural predictivity, and that 2) a combination of positional, semantic, and surprisal feature spaces accounted for 80.1% of the neural variance explained by GPT2. Additionally, GPT2 did not explain significantly more neural variance in total than the interpretable feature spaces. Together, these results suggest that the majority of neural variance explained by GPT2 on a passage-level dataset can be accounted for by a concise set of interpretable feature spaces.

1-063. Quantifying Behavioral State Instability and Belief-updating Deficits in a Novel Mouse Model for Schizophrenia

Yi-Yun Ho[1,2] YI.YUN.A.HO@GMAIL.COM
Tingting Zhou[1] TTZHOU@MIT.EDU
Amanda Fath[1] AFATH@MIT.EDU
Kathleen He[1] KYHE@MIT.EDU
Nolan Hartley[3] NHARTLEY@BROADINSTITUTE.ORG
Jonathan Wilde[1] WILDEJ@MIT.EDU
Xian Gao[1] XIANGAO@MIT.EDU
Cui Li[1] LICUI@MIT.EDU
Zhanyan Fu[3] ZFU@MIT.EDU
Matt Nassar[4] MATTHEW_NASSAR@BROWN.EDU
Michael Halassa[5] MICHAEL.HALASSA@TUFTS.EDU
Guoping Feng[6] FENGG@MIT.EDU

[1]Massachusetts Institute of Technology
[2]McGovern Institute for Brain Research
[3]Broad Institute
[4]Brown University
[5]Helsinki Institute of Life Science; Tufts University School of Medicine
[6]Massachusetts Institute of Technology; Broad Institute

Schizophrenia is a severe neuropsychiatric disorder affecting 1% of the global population, however, it has been difficult to mechanistically study its cognitive symptoms due to a lack of appropriate animal models and behavior assays. In this study, we developed an easy-training and computational trackable assay, lever pressing task, capable of capturing subtle cognitive deficits in belief updating in a dynamic environment. We used mice with point mutation (Grin2a Y700X+/-) identified in human schizophrenic patients by large-scale exome sequencing studies as a mouse model for schizophrenia. We found these mutants showed more sampling trials when the values of options were close and had difficulty committing to the better option, which was also reported in human patients. Using the Hidden Markov Model, we further quantified the state transition between commitment and exploration state and found that mutants showed more frequent state transition, thus more unstable states. The performance of the behavior can also be captured by a belief updating model based on Bayesian framework, where mutants showed a slower update rate and higher uncertainty of the option values. Using the resting-state functional ultrasound, we identified the mediodorsal thalamus (MD) as one of the prominent regions showing greater differences between wild-types and mutants. In vivo single-unit electrophysiological recording showed that MD encodes dynamic values and behavioral states in the lever pressing task, elucidating the role of MD in task performance. Interestingly, mice with optogenetic suppression of MD also showed unstable behavioral state, phenocopying behavioral deficits of Grin2a Y700X+/- mice in the lever-pressing task. Activating MD neuronal activity with SSFO in mutants rescued the cognitive deficits in Grin2a Y700X+/- mice. Together, we identified a defective belief updating process similar to human patients in a newly developed genetic schizophrenia mouse model using a lever pressing assay, and identified a potential treatment target for cognitive deficit in schizophrenia.

1-064. Stereotyped Propagating Activity in Developing Neonatal Visual Cortex

Luna Kettlewell[1,2] CKETTLEW@UMN.EDU
Audrey Sederberg[3] SEDE0018@UMN.EDU
Gordon Smith[1] GBSMITH@UMN.EDU

[1]University of Minnesota
[2]Neuroscience
[3]Georgia Institute of Technology

Visual perception relies on precise network organization in the visual cortex. Evidence of this organization is found before eye-opening in early development, where activity patterns exhibit modular structure and long-range spatial organization. These early long-range networks appear to serve as precursors for mature network organization (Smith 2018, Tragenap 2023), yet they have not been fully characterized. While previous research has focused on the spatial features of early long-range networks, their temporal structure remains largely unexplored, yet likely contributes significantly to the plasticity and maturation of cortical networks. Direct long-range horizontal projections are immature at this age, suggesting that large-scale network structure arises via propagating network activity through short range connectivity. To explore this, we acquired a novel dataset utilizing a fast calcium indicator (GCaMP8m) and widefield imaging at high temporal resolution (50Hz) to capture hours of spontaneous activity in neonatal ferret visual cortex. Consistent with prior work, spontaneous activity was modular and spatially

segregated. Temporal dynamics of activity were complex, with approximately one thousand activity sequences per animal (mean=1352 sequences, n=6). We find that the majority of sequences have a dynamic component where modules activate sequentially across the field of view (mean=91.4%, CI=81.5%-95.7%, n=6), whereas only a small fraction of sequences arise as a global response. These sequences share many features of traveling waves and we find that a large fraction of the dynamic sequences can be fit to an underlying linear wave propagating in a modular manner across the cortex (mean=35.6%, CI=24.0%-51.5%, n=6). We find that spatiotemporal sequences occur in repeated and stereotyped motifs across hours of spontaneous activity. In conclusion, we demonstrate a rich spatiotemporal structure to spontaneous activity in the early developing cortex, providing important constraints for models of cortical development and plasticity.

1-065. Multi-assay learning to discover the anxious internal state

Michael Klein[1,2] MICHAEL.KLEIN413@DUKE.EDU
Dalton Hughes[1] DALTON.HUGHES@DUKE.EDU
Diana Waters[1] DIANA.WATERS@DUKE.EDU
Stephen Mague[1] STEPHEN.MAGUE@DUKE.EDU
Jake Benton[1] JACOB.BENTON@DUKE.EDU
Kathryn Walder-Christensen[1] KATHRYN.WALDER@DUKE.EDU
Rainbo Hultman[3] RHULTMAN@HEALTHCARE.UIOWA.EDU
Gwenaelle Thomas[4] GWENAELLE.THOMAS@LIBD.ORG
Elise Adamson[1] ELISE.ADAMSON@DUKE.EDU
Noah Lanier[1] NOAH.LANIER@DUKE.EDU
Neil Gallagher[5] NGA4004@MED.CORNELL.EDU
Austin Talbot[6] TALBOTA@PILLARBIOSCI.COM
Jack Goffinet[1,7] JACK.GOFFINET@DUKE.EDU
Alexandra Fink[8] ALEXANDRA.FINK@ICAHN.MSSM.EDU
William Carson[1] WILLIAM.CARSON@DUKE.EDU
David Carlson[1] DAVID.CARLSON@DUKE.EDU
Kafui Dzirasa[1] KAFUI.DZIRASA@DUKE.EDU

[1]Duke University
[2]Electrical and Computer Engineering
[3]University of Iowa
[4]Leiber Institute for Brain Development
[5]Weill Cornell Medicine
[6]Pillar Biosciences
[7]Computer Science
[8]Icahn School of Medicine at Mount Sinai

Pathological anxiety represents a significant comorbidity and precursor of many mental health disorders. A better understanding of neural mechanisms underlying anxiety may lead to better diagnostic tools and treatments. While several individual brain regions and projections have been shown to be associated with anxiety in specific contexts, the interaction of these mechanisms remains unexplored on a brain-wide multi-region scale and across many anxiogenic contexts. To elucidate these multi-region interactions, we record electrical activity from eight brain regions as mice engage in three anxiogenic assays consisting of environmental and pharmacological stimuli. We then use machine learning to discover three distinct multi-region networks of local field potential activity that jointly encode anxious behavior in all three assays. We find that using single-assay approaches fail to generalize to multiple contexts, highlighting the necessity of multiple-assay learning to capture complex emotional states. We validate that these networks encode five new anxiogenic contexts using holdout mice but do not encode new control contexts that relate to novelty and arousal. The networks reliably distinguish mice that underwent a chronic unpredictable stress paradigm from controls in a homecage environment, implying that the networks can detect increased baseline anxiety in a clinically relevant paradigm. Additionally, these networks modulate in concert with optogenetic stimulation of an established anxiety-related pathway. More importantly, recent works have highlighted the desirability of representational consistency in brain-wide association studies. Here, we also present novel methodology for evaluating representational consistency of predictive networks. We use this approach for model selection, and show that the methodology learns robust, reproducible networks within the cohort. In conclusion, we present three networks that jointly encode anxious behavior in multiple assays and across many subjects while providing new model selection methodology to improve representational stability.

1-066. Characterizing noise performance in Drosophila vision

Hyosun Kim[1,2] HOYSUN0407@HANMAIL.NET
Anmo Kim[3] ANMOKIM@HANYANG.AC.KR

[1]Hanyang University
[2]Department of Artificial Intelligence
[3]Hanging University

Animals can recognize visual objects in challenging conditions, such as viewing through a window covered with water droplets. However, the neural mechanisms underlying the robustness of visual object recognition against noise remain unknown. To understand how visual systems suppress noise to extract critical visual features, we characterized the performance of visual object detection in flying Drosophila, by measuring both behavioral and neuronal responses. We designed simple visual patterns, such as translating bars, gratings, spots, and looming objects overlaid with varying levels of salt-and-pepper noise that is either temporally correlated or not. The analysis of the wing stroke responses to these stimuli revealed a decrease in wing response amplitude and an increase in the latency for increasing noise levels. The bar and grating responses exhibited relatively stronger noise tolerance (more gradual reduction in the wing response for increasing noise levels) than the spot and looming patterns (Figure 1). We repeated the same experiments for humans and 3D ResNet and found that humans exhibited overall higher performance than flies, albeit with a narrow margin and longer latencies, whereas 3D ResNet performed the worst when trained only with noiseless patterns. Subsequently, we performed two-photon calcium imaging of visual neurons (T4/T5, LC15, LC11, LC4, LPLC2) previously reported to be sensitive to these visual stimuli. The bar-sensitive LC15 neurons exhibited the highest level of noise performance, responding significantly up to 60% of noise. Finally, when we genetically silenced one of these neuron types, the behavioral response reduced for a specific range of noise levels, suggesting the role of these visual neurons in noise suppression. Altogether, our study established the role of feedforward visual neurons in extracting specific features from sensory stimuli in noisy environments.

1-067. GPe arkypallidal neurons can mediate inhibitory control by disrupting competition in the striatum

Cristina Giossi[1,2] CRISTINAGIOSSI@GMAIL.COM
Jyotika Bahuguna[3] JYOTIKA.BAHUGUNA@GMAIL.COM
Jonathan Rubin[4] JONRUBIN@PITT.EDU
Timothy Verstynen[3,5] TIMOTHYV@ANDREW.CMU.EDU
Catalina Vich Llompart[1,6] CATALINA.VICH@UIB.ES

[1]Universitat de les Illes Balears
[2]Department of Mathematics and Informatics
[3]Carnegie Mellon University
[4]University of Pittsburgh
[5]Department of Psychology, Neuroscience Institute
[6]Departament de Ciencies Matematiques i Informatica

Inhibitory control involves either restraining (proactive control) or stopping (reactive control) actions and relies on the cortico-basal ganglia-thalamic (CBGT) network. Traditional models posited that reactive control was solely managed by the hyperdirect pathway and information in the CBGT network flowed in one direction: downward from cortex to the basal ganglia output regions. Recent research, however, has highlighted an underappreciated complexity of neuron types and connections in the external globus pallidus (GPe) region of the basal ganglia, highlighting the limitations of the traditional view. Our study provides the first use of a biologically-constrained spiking neural network that incorporates these recent findings and demonstrates how they lead to a radically revised understanding of the roles of the GPe in inhibitory control and information regulation. This artificial CBGT network includes recently identified connections from arkypallidal neurons in the GPe (GPe-A) to the striatum. These connections carry ascending information, counterbalancing the traditional descending signals controlled by prototypical GPe neurons (GPe-P). We explore various scenarios in which a stop-signal is sent to different CBGT neuron populations: subthalamic nucleus (STN; hyperdirect pathway input), indirect spiny projection neurons (iSPNs) in the striatum, and GPe-A cells. These experiments demonstrate STN's involvement in inhibition via signals relayed through the GPe. Additionally, disrupting the GPe-A pathway, in combination with activating STN and iSPN, diminishes the network's stopping capability, underscoring GPe's pivotal role in regulating information flow. Notably, GPe-A neurons affect the balance between striatal iSPNs and dSPNs (direct pathway SPNs), thereby influencing the transmission of a stop-signal through pallidostriatal projections. This study elucidates how GPe-A pathways facilitate reactive control by managing the bidirectional information flow within the CBGT network, offering new insights into the complex mechanisms of inhibitory control.

1-068. The Interplay of Behavioral Rules, Internal State, and Feedback Cues in Shaping Social Interactions

Sarath Ravindran Nair[1,2] SARATH.NAYAR@GMAIL.COM
Adrian Palacios-Munoz[1] A.PALACIOS@ENI-G.DE
Jan Clemens[3] J.CLEMENS@ENI-G.DE

[1]European Neuroscience Institute
[2]Neural Computation and Behavior Group
[3]European Neuroscience Institute; Carl von Ossietzky University Oldenburg; BCCN

Social interactions are shaped by the behavioral decisions and feedback of interacting organisms. Interactions vary towards different partners, but the relative contributions of an individual's actions and partner's feedback to this variation is often unclear. Here, we use Drosophila melanogaster courtship as a model to dissect the role of behavioral rules and feedback in shaping social interactions. During courtship, Drosophila males dynamically process sensory feedback from females to pattern their courtship song. Interestingly, males also sing to other males, but these interactions result in interaction contexts distinct from female-directed courtship. To understand whether the males alter their sensorimotor strategies to pattern courtship songs in target-specific contexts, we used computational modeling and optogenetics to characterize these transformations during female- and male-directed singing. We find that the males use the same set of rules towards both sexes with target-agnostic and target-specific models performing similarly, indicating a target-agnostic sensory subspace relevant for song decision. We further dissected how sensory cues induce behavioral feedback through changes in internal state. The unique head-to-head context during male-directed interactions is not initiated by the singer but arises from the male target's turning response and this response is caused by courtship song perception. However, the song does not act as a trigger but rather through a slower change in the target's internal arousal state. Optogenetic manipulations show that this arousal state is encoded in the central brain neurons P1a. Thus, while the experimental findings identify the causes of target-specific interaction contexts, a computational approach reveals target-invariant sensory representations crucial for behavioral decisions. In summary, our research sheds light on the intricate interplay of behavioral rules, internal state, and feedback cues in shaping social interactions, providing valuable insights into the dynamics of social behavior.

1-069. Riemannian geometry of neural object representations

Jacob Zavatone-Veth[1,2] JZAVATONEVETH@G.HARVARD.EDU
Sheng Yang[1] SHENGYANG@G.HARVARD.EDU
Julian Rubinfien[3] JULIAN.RUBINFIEN@YALE.EDU
Cengiz Pehlevan[1] CPEHLEVAN@SEAS.HARVARD.EDU

[1]Harvard University
[2]Physics
[3]Yale University

Neural codes are distributed: the identity of a particular stimulus is usually signaled not by the activity of a single specialized neuron, but by a collective change in the activity of a population. To understand distributed neural codes, we need to identify the right set of collective observables one should measure. Here, we propose that the Riemannian geometry of neural representations provides a compelling candidate framework. This perspective gives a precise way to measure how distributed codes shape the similarity of stimuli. It is particularly powerful because it does not require a priori knowledge of the geometric structure for which one is looking. As a proof of principle, we investigate the geometry of deep convolutional neural networks trained to perform object classification, which are a popular model for the computations of the ventral visual stream. We show that these networks learn to magnify areas of the input space around object class boundaries. Networks trained with self-supervised learning also learn to magnify areas near decision boundaries. This provides an important prerequisite for applying these analyses to biological data: even when training is performed without label information, object information can be extracted from the geometry of the neural representation. This leads us to hypothesize that the ventral visual stream may share this object-centric geometric structure. In total, this Riemannian framework provide a new lens through which to compare natural and artificial distributed neural representations.

1-070. Individual differences in prefrontal coding of visual features

Qi Lin CBLINQI@GMAIL.COM
Hakwan Lau HAKWAN.LAU@RIKEN.JP
RIKEN Center for Brain Science

The lateral prefrontal cortex (LPFC) is commonly associated with high-level cognition, such as attention, language and cognitive control. However, recent work has demonstrated that it is also critical for basic perceptual functions including object recognition. Here we characterize the role of LPFC in visual processing with computational models. Using a dataset of human fMRI data at 7T, we built encoding models relating visual features extracted from a deep neural network (the image encoder of a CLIP [Contrastive Language–Image Pre-training] network) to brain responses to thousands of natural images. Within each of the eight subjects, we were able to robustly predict responses in patches of LPFC, most notably in FEF (frontal eye field) and vlPFC (ventrolateral PFC) regions. Leveraging these robust encoding models, we then screened for images with high predicted responses. We found striking individual differences in which images tend to drive the overall activation in LPFC of a specific subject. In contrast, images with high predicted responses in the ventral visual stream remain highly consistent across individuals. Overall, our study demonstrates the under-appreciated role of LPFC in visual processing and suggests that LPFC may underlie the idiosyncrasies in how different individuals experience the visual world. Methodologically, these findings may also explain why previous group studies have often failed to observe robust visual functions in LPFC, as subjects' responses may need to be calibrated individually.

1-071. Robust and Efficient Grid Code Transformation for Rapid Task Transfer

Heejun Kim[1,2] HJKDAVID@KAIST.AC.KR
Sang Wan Lee[3] SANGWAN@KAIST.AC.KR
[1]Korea Advanced Institute of Science and Technology
[2]Bio and Brain Engineering
[3]KAIST

Humans use entorhinal grid and hippocampal place codes to navigate complex mazes. Although the grid codes represent maze structures independent of behavioral policy, little is known about how the task structure-specific grid codes generalize in new environments. Most reinforcement learning algorithms transfer task representations based on their decision policies. However, this approach lacks biological plausibility and needs multiple representations biased toward policies to solve the tasks. To circumvent these issues, we propose a novel computational principle called robust grid code transformation (RGCT). Our model learns a single robust grid code for each task structure and transforms it into diverse tasks with the same structure. Specifically, RGCT calculates policy-independent state prediction errors (SPE) from grid codes to rapidly learn robust grid codes reflecting task structures and transforms grid codes pretrained on similar sub-structures for faster adaptation. RGCT has a generic and scalable architecture since it can be viewed as a multi-head self-attention, in which two heads calculate robust place codes for decision-making and decode task structures to calculate SPE. First, we show the robustness of grid codes in various structure-preserving task transformation scenarios. We also demonstrate our model's biological plausibility by replicating neural findings showing the robustness of place codes across diverse scenarios. RGCT learns novel task structures faster than other training algorithms. Notably, RGCT outperforms other transfer learning models while using only a single grid code per task structure and fewer place codes. This study suggests an efficient computational principle of hippocampal transfer learning and advances our understanding of hippocampal information processing.

1-072. Feedback controllability constrains learning timescales during motor adaptation

Harsha Gurnani[1] HARSHA84@UW.EDU
Bing Brunton[2] BBRUNTON@UW.EDU
[1]University of Washington
[2]University of Washington, Seattle

The ability to produce new neural dynamics is a key feature of motor learning, and likely involves plasticity within distributed circuits; this learning is also relevant in the context of brain-computer interfaces (BCI) that rely on real time decoding of neural activity. Previous work exploring the structure of M1 activity has largely assumed autonomous dynamics (i.e. activity unfolding from initial states dominated by local recurrent interactions), and related work on BCI learning has focused on local mechanisms (such as M1 synaptic plasticity). However, recent

experimental evidence suggests that M1 activity during BCI use is continuously modified by sensory feedback [1,2] and produces corrections for noise and external perturbations [1,3], suggesting a critical need to model this interaction between feedback and intrinsic M1 dynamics. Here we propose that for learning on short timescales, such as adaptation to visuomotor rotations or BCI decoder changes, M1 dynamics can be effectively modified by changing inputs, including by flexible remapping of sensory feedback. Using recurrent network models of BCI under feedback control, we show how the rate of such adaptation is constrained by pre-existing recurrent and input connectivity. We investigate how intrinsic dynamical structure shapes the flow of neural activity in time and poses additional constraints on which activity trajectories are easier to produce. This insight extends previous analyses of neural covariance structure that describes the statistical distribution of activity patterns, the so-called "intrinsic manifold". Using this framework, we further show that designing BCI decoders using only statistical features of neural data leads to less robust performance, which may be critical in BCI applications. By incorporating adaptive controllers upstream of M1, our work highlights the need to model input-dependent latent dynamics, and clarifies how constraints on learning arise from both the statistical characteristics and the underlying dynamical structure of neural activity.

1-073. A view from latent space: mapping spikes to rates in low-rank, excitatory-inhibitory networks

William Podlaski[1,2] WILLIAM.PODLASKI@RESEARCH.FCHAMPALIMAUD.ORG
Christian Machens[3] CHRISTIAN.MACHENS@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Neuroscience Programme
[3]Champalimaud Research

A ubiquitous characteristic of both neural data and network models is the use of continuous-valued firing-rates (or linear combinations of rates at the population level) as the definitive language of neural coding in the brain. The validity of this choice relies upon a correspondence between such rates and the underlying spike-based nature of neural communication, which has long been the subject of theoretical study. However, most previous approaches attempt to transform the discontinuous spiking threshold into a smooth firing-rate at the single-neuron level, which may involve averaging over neurons, time, or trials, and which may not be applicable in the general case. Here, we present a new approach to the spike-rate correspondence using a neural population geometry perspective. The key insight that we propose is that, rather than smoothing the spiking threshold at the level of each neuron, we first map spiking dynamics to a low-dimensional latent space, and then smooth the dynamics in this latent space at the population level. Specifically, we consider balanced spiking network models with low-rank, Daleian (excitatory-inhibitory) recurrent connectivity, and show that population activity trajectories are confined to manifolds in low-dimensional latent space. Moreover, these manifolds are situated at the boundary between sub-threshold and supra-threshold (i.e., balanced) inputs. This boundary is hard and infinitely-steep in spiking networks with fast (delta) synapses, but becomes softer and more sigmoidal with finite (e.g., exponential) synaptic dynamics and noise. We demonstrate the corre- spondence between the latent trajectories of such spiking networks and their equivalent rate networks on a single-trial basis, despite trial-to-trial spiking variability. Overall, this work provides a much-needed mecha- nistic basis for a large class of abstract rate models and links classic work on the spike-rate correspondence to more recent ideas on low-dimensional neural population geometry.

1-074. Hebbian and heterosynaptic plasticity regulate orientation matching in binocular cortical circuits during the critical period

Katya Tsimring[1,2] KTSIMRIN@MIT.EDU
Claudia Cusseddu[3] CLAUDIA.CUSSEDDU@TUM.DE
Kyle Jenks[1] KRJENKS@MIT.EDU
Greggory Heller[1] GREGGH@MIT.EDU
Julijana Gjorgjieva[3,4] GJORGJIEVA@TUM.DE
Jacque Ip[5] JACQUEIP@CUHK.EDU.HK
Mriganka Sur[1] MSUR@MIT.EDU

[1]Massachusetts Institute of Technology
[2]Department of Brain and Cognitive Science
[3]Technical University of Munich
[4]School of Life Sciences
[5]The Chinese University of Hong Kong,

Experience-dependent plasticity refines immature sensory circuits during critical periods in development. In the

binocular visual cortex (bV1), visual experience aligns information from the ipsilateral and contralateral eye onto bV1 neurons. It remains unclear how this experience-dependent alignment between the two eyes takes place at the synaptic level. We propose that Hebbian and heterosynaptic mechanisms regulate the alignment of inputs onto bV1 neurons during the critical period by modifying synapses based on their correlation to the post-synaptic neuron and their synaptic neighbors, respectively. To investigate the synaptic mechanisms that underlie somatic orientation matching, we used in vivo two photon calcium imaging to chronically track eye-specific and binocular visual responses of neurons and dendritic spines. We observed significant turnover of dendritic spines from ~postnatal (p)22 (D1) to ~p32 (D10), as well as shifts in the functional properties of retained spines. To examine Hebbian-like mechanisms of regulating dendritic spine turnover, we compared tuning curves of spines with the soma's. Early in development, we found that newly added spines were less aligned to the soma's tuning properties than retained spines. By D10, spines that were misaligned to the soma were lost, suggesting Hebbian refinement of visual circuits. To examine heterosynaptic interactions, we measured trial-to-trial correlations between spine pairs and found a distance-dependent relationship between co-active spines. Furthermore, this correlation between spine neighbors increased from D1 to D10 with the addition of co-tuned spines. Our results suggest that heterosynaptic mechanisms stabilized newly added spines. To quantify contributions of Hebbian and heterosynaptic plasticity in somatic orientation matching, we built a computational model to simulate dendritic spine turnover. Simulations removing heterosynaptic plasticity prevented orientation matching at the soma, indicating that Hebbian mechanisms alone cannot drive this alignment. Overall, our research provides critical insights into how synaptic mechanisms shape sensory circuits and enhance information encoding across development.

1-075. Selective filtering of sequences of neural activity by recurrent circuits of sensory cortex

Ciana Deveau[1] CIANA.DEVEAU@NIH.GOV
Zhishang Zhou[2] ZHISHANG.ZHOU@NIH.GOV
Paul LaFosse[3] PAUL.LAFOSSE@NIH.GOV
Yanting Deng[2] YANTING.DENG@NIH.GOV
Saghar Mirbagheri[4] SAGHARM@UW.EDU
Nicholas Steinmetz[4] NSTEINME@UW.EDU
Mark Histed[5,6] MARK.HISTED@NIH.GOV

[1]National Institutes of Health
[2]National Institute of Mental Health
[3]National Institute of Mental Health; University of Maryland College Park
[4]University of Washington
[5]NIH
[6]NIMH

As we interact with the world, we receive streams of sensory input that change over time, yet how cortical networks process temporally dynamic input is not well-understood. All cortical areas have extensive local recurrent excitatory-excitatory connections. In some cortical regions, these recurrent circuits are proposed to generate dynamic activity, as with delay responses in prefrontal cortex. However, generative dynamics are not seen in sensory cortices like visual cortex, where responses largely turn on and off with stimulus presentations. Here we directly study the computational role of recurrent circuits in mouse V1 using two-photon holographic stimulation (Fig. 1). We find a new role for recurrent networks in sensory cortex: they selectively filter dynamic sequences of input, amplifying sequences corresponding to natural sensation. First, if V1 circuits are sensitive to input dynamics, cortical responses should differ depending on sequential ordering of input patterns. We test this using two-photon induced patterns, and find sequential order changes network responses. We next examine potential network mechanisms by splitting up component sequences in time. We find that a single pattern of input modulates the firing rates of non-stimulated neurons, and these off-target suppressed responses interact with later inputs, changing responses to inputs based on pattern history (Fig. 2A). We next examine how this sequential selectivity relates to natural visual responses. We extract one frame from a natural movie, measure V1 responses, then play back those responses using two-photon stimulation in the proper dynamic context and in a mismatched context. We find the matched context yields an amplified response (Fig. 2F), showing the V1 network amplifies inputs when they are embedded in a sequence corresponding to natural vision. This work suggests a novel function for recurrent connections in the cerebral cortex: selective filtering of input sequences. This sequence filtering boosts natural input sequences while attenuating irrelevant input.

1-076. Analytical study of learning dynamics in compositional tasks in the teacher-student setup

Jin Hwa Lee[1] HAYA.JINLEE@GMAIL.COM
Stefano Sarao Mannelli[2] S.SARAOMANNELLI@UCL.AC.UK
Andrew Saxe[3] A.SAXE@UCL.AC.UK

[1]Sainsbury Wellcome Centre, UCL
[2]Gatsby Computational Neuroscience Unit, UCL
[3]Gatsby Computational Neuroscience Unit & Sainsbury Wellcome Centre, UCL

Complex tasks in the real world are often composed of multiple simpler primitive tasks. This compositionality allows animals to flexibly combine, reuse, and generalise previously acquired knowledge (Schwartenbeck et al., 2023). Nevertheless, we lack theoretical understanding of the learning dynamics of compositional tasks, especially when tasks unfold over time (Luettgau et al., 2023) as in Reinforcement Learning (RL) when reward is provided after successfully performing a sequence of actions. Although there is evidence that different curricula influence compositional generalisation in humans (Dekker et al., 2022), there is no analytical framework to explain and quantitatively characterise these findings. This study aims to fill this gap by proposing a parsimonious model of compositional generalisation based on the teacher-student framework (Biehl and Schwarze, 1995; Saad and Solla, 1995). Our model allows for an exact solution of the learning dynamics and a characterisation of different learning strategies including vanilla training, and what we call primitives pre-training. In vanilla training, a naive agent is trained directly on a compositional task. In primitives pre-training, an agent follows a curriculum strategy where primitives are learnt to expert level before experiencing the full compositional task. While the agent is able to learn the primitives and the full compositional task in both protocols, our results show that primitives pre-training provides a significant speed-up in learning. The theory also reveals that compositionality helps most in temporally extended RL settings. Overall, our theory provides an analytical understanding of the benefits of compositionality, and a quantitative account of how training protocols can disclose useful task primitives, ultimately speeding learning.

1-077. Serotonin predictively encodes value

Emerson Harkin[1,2] EMERSON@EFHARKIN.COM
Cooper Grossman[3] COOPERGROSSMAN@CALTECH.EDU
Jeremiah Cohen[4] JEREMIAH.COHEN@ALLENINSTITUTE.ORG
Jean-Claude Beique[1] JBEIQUE@UOTTAWA.CA
Richard Naud[1] RNAUD@UOTTAWA.CA

[1]University of Ottawa
[2]Centre for Neural Dynamics and Artificial Intelligence
[3]California Institute of Technology
[4]Allen Institute for Neural Dynamics

The in vivo responses of dorsal raphe nucleus (DRN) serotonin neurons to emotionally-salient stimuli are a puzzle. Existing theories centred on reward, surprise, salience, and uncertainty individually account for some aspects of serotonergic activity but not others. Here we find a unifying perspective in a predictive code for value, a quantity that combines biological constraints with the representation of future reward used in reinforcement learning. Through simulations of trace conditioning experiments common in the serotonin literature, we show that our theory, called value prediction, explains phasic activation of serotonin neurons by both rewards and punishments, preference for surprising rewards but absence of a corresponding preference for punishments, and contextual modulation of tonic firing—observations that currently form the basis of many and varied serotonergic theories. To empirically test our theory, we analyzed tetrode recordings of identified serotonin neurons in mice undergoing trace conditioning, finding single-neuron and population-level activity patterns well within 0.1 Hz / neuron of our predictions; a surprisingly close match. Finally, we directly compared value prediction against quantitative formulations of existing ideas and found that our theory best explains both within-trial activity dynamics and trial-to-trial modulations, almost always by a large margin. Overall, our results show that previous models are not wrong, but incomplete, and that reward, surprise, salience, and uncertainty are simply different faces of a predictively encoded value signal. By unifying previous theories, our work represents an important step towards understanding the potentially heterogeneous computational roles of serotonin in learning, behaviour, and beyond.

1-078. Spiking Neural Networks on Multiple Instruction Multiple Data Processors for Interdisciplinary HPC applications

Jan Finkbeiner[1,2] JAN.FINKBEINER@T-ONLINE.DE
Catherine Schofmann[1] C.SCHOEFMANN@FZ-JUELICH.DE
Jan Vogelsang[1] J.VOGELSANG@FZ-JUELICH.DE
Susanne Kunkel[1] SU.KUNKEL@FZ-JUELICH.DE
Emre Neftci[1] E.NEFTCI@FZ-JUELICH.DE

[1]Juelich Research Center
[2]PGI-15

This work explores the potential of leveraging MIMD (Multiple Instruction, Multiple Data) processors and local computation for both efficient large-scale training of spiking neural networks (SNNs) and simulating SNNs with biologically plausible synaptic density and connectivity. It addresses the limitations of current High-Performance Computing (HPC) hardware in handling the demands of SNN training and simulations due to their architectures optimized for traditional machine learning tasks. By approaching these challenges from both a machine learning inspired and neuroscientific simulation perspective, we identify common elements supporting the implementation of SNNs on an accelerator and document significant performance gains over conventional hardware counterparts.

The use of SNNs in machine learning has gained traction as an alternative approach to process time series and event-based sensor data owing to their spatiotemporal sparsity. However, GPUs, while adept at accelerating deep learning computations, are suboptimal for the distributed local computation and sparse communication required by SNNs. Similarly, established simulators for bio-realistic networks still primarily rely on CPUs. Neuromorphic chips tailored for inference lack algorithmic flexibility required to adapt to large SNNs. Here, we explore the use of Graphcore's IPU (Intelligence Processing Unit), featuring MIMD design with near-memory computing. It consists of parallel tiles, each hosting a processing unit and local SRAM memory accessible within few computing cycles. This design enables handling irregular memory access patterns inherent in dynamical activation sparsity and sparse connectivity of SNNs. This kind of distributed near-memory computing aligns with the brain's principle of distributed local computation. It allows neuron placement on dedicated tiles, reduces on-chip data transfer, and increases efficiency, marking a step towards non-von Neumann neuromorphic computing architectures while retaining algorithmic flexibility.

1-079. Order from chaos: Interplay of development and learning in recurrent networks of structured neurons

Laura Kriener[1,2] LAURA.KRIENER@UNIBE.CH
Kristin Volk[3] VOELK@CAT-LAB.COM
Ben von Hunerbein[1] BEN.VONHUENERBEIN@UNIBE.CH
Federico Benitez[1] FEDERICO.BENITEZ@UNIBE.CH
Walter Senn[1] WALTER.SENN@UNIBE.CH
Mihai Petrovici[1] MIHAI.PETROVICI@UNIBE.CH

[1]University of Bern
[2]Institute for Physiology
[3]Catlab Engineering GmbH, University of Bern

Behavior can be described as a temporal a sequence of actions, which are grounded in neural activity. In order for a neuronal network to learn a complex sequential pattern, memories of past activities need to persist on behavioral time scales – significantly longer than characteristic relaxation times of single-neuron activity. Typically, network recurrence is assumed to produce such long transients, but learning in such networks remains a challenging problem. Commonly used approaches can be divided into two groups. First, one can simply forego recurrent plasticity and only learn the weights of a "readout" layer (reservoir computing). Second, the recurrent weights can be learned as well, with the most successful algorithms using propagated errors to update synaptic weights. However, their biological plausibility suffers from issues with locality (BPTT), resource allocation and scaling (RTRL and its approximations), or parameter scales and tuning (FORCE, E-PROP). We suggest that many of these issues can be alleviated by considering dendritic information storage and computation. In particular, we consider the interplay between two populations of cortex (for example, in motor and premotor areas) that undergo two phases of network evolution. During early development, neurons start out by forming a sparse, random scaffold of somato-somatic connections. During later learning, this scaffold is used to transmit errors that serve the growth and tuning of new synapses along dendritic trees. This fully local, always-on plasticity ultimately carves out strong, dynamic attractors that generate robust activation sequences. Importantly, our model makes efficient use of its neuronal resources, allowing the learning of complex sequences with only a small number of neurons. We demonstrate these features in a mock-up of birdsong learning, in which our networks learn a long, non-Markovian sequence (a sample of Beethoven's "Fur Elise") that they can reproduce robustly despite severe

external disturbances.

1-080. SiBBlINGS: Similarity-driven Building-Block Inference using Neural-Graphs across States

Noga Mudrik[1,2] NOGAMUDRIK@GMAIL.COM
Gal Mishne[3] GMISHNE@UCSD.EDU
Adam Charles[1] ADAMSC@JHU.EDU

[1]Johns Hopkins University
[2]Biomedical Engineering
[3]University of California San Diego

Neural data, collected through multiple trials within distinct states (e.g., varied tasks), are often subject to complex variability due to underlying latent factors (e.g., unmeasured environmental variables), affecting dynamics both within and across known states. A promising approach to studying neural variability involves identifying functional ensembles (neural Building Blocks) of co-active neurons that can adjust their temporal activity and neural membership across trials. Existing methods for finding such ensembles underlying multi-way data (e.g., trial-by-condition-by-time-by-neuron) often rely on tensor factorization (TF) approaches, which require aligned trials of the same duration, constant sampling rates, no missing data, or identical trial structure—assumptions that do not align with the properties of real-world neural data. Here, we present a method for Similarity-driven Building-Block Inference using Neural-Graphs across States (SiBBlINGS), which employs a graph-guided dictionary learning approach for neural ensemble discovery. SiBBlINGS considers shared temporal activity, inter- and intra-state relationships, non-orthogonal ensembles, and variations in trial duration and counts across states, while remaining resilient to noise, initializations, and missing samples. Additionally, it enables the identification of state-specific vs. state-invariant ensembles and allows for cross-state controlled variations in ensemble structure and per-trial temporal variability—qualities often unattainable with alternative methods. We demonstrate SiBBlINGS's capability to 1) recover ground truth components in noisy synthetic data with random initialization—outperforming existing approaches, 2) identify neural patterns in the cortex, and 3) capture emerging neural activity that aligns with clinically identified seizure locations.

1-081. Estimating flexible across-area communication with neurally-constrained RNN

Joao Barbosa[1,2] PALERMA@GMAIL.COM
Adrian Valente[1] ADRIAN.VALENTE@ENS.FR
Scott Brincat[3] SBRINCAT@MIT.EDU
Earl Miller[3] EKMILLER@MIT.EDU
Srdjan Ostojic[1] SRDJAN.OSTOJIC@ENS.FR

[1]Ecole Normale Superieure
[2]Group for Neural Theory
[3]Massachusetts Institute of Technology

Neural computations supporting complex behaviors involve multiple brain regions, and large-scale recordings from animals engaged in complex tasks are increasingly common. A current challenge in analysing these data is to identify which part of the information contained within a brain region is shared with others. For instance, using linear decoding, one might find that a given area encodes all task-related variables but using decoding alone it is hard to identify which variables are actually communicated to a specific downstream area. This is particularly challenging when considering more than two interconnected areas (Semedo et al., 2019).

Here, to address this limitation, we trained multi-region recurrent neural networks (RNN) models to reproduce the dynamics of large-scale single-unit recordings (more than 6000 neurons across 7 cortical areas) from monkeys engaged in a two-dimensional (color and motion direction) context-dependent decision-making task (Siegel et al., 2015). After fitting, we partitioned the activity of each area, separating recurrent inputs from those originating in other areas (Perich et al., 2021). Decoding analyses show that all areas encode both stimuli (color and direction), but selectively project different dimensions of their activity. Sensory areas (V4, MT and IT) project only one variable (color or direction) while compressing others, irrespective of the context or downstream area. In contrast, we observed that the prefrontal cortex (PFC) and frontal eye fields (FEF) projected different aspects of the stimulus, depending on the downstream area or context. In the model, PFC/FEF strongly compress the irrelevant stimulus dimension in their projections to fronto-parietal areas but not as much towards sensory areas. These preliminary results motivate a novel approach to study how different regions coordinate their activity to solve context-dependent tasks.

1-082. A power law of cortical adaptation in neural populations

Dario Ringach[1,2] DARIO@UCLA.EDU
Elaine Tring[3] ETRING@GMAIL.COM
Mario Dipoppa[4] MDIPOPPA@G.UCLA.EDU

[1]University of california, Los Angeles
[2]Neurobiology and Psychology
[3]University of California Los Angeles
[4]UCLA

Sensory adaptation is a ubiquitous computation across modalities and brain regions. Here we provide a mathematical description of adaptation in cortical populations. The experimental design begins with the selection of finite stimulus set, S=s_i . We define a visual environment, A, as one where the probability of observing s_i is given by p_A (s_i). To examine how neurons adapt in this environment, we present a stimulus sequence by independently drawing stimuli from p_A (s_i) while recording their activity. We define the vector r_A (s_i) as the mean response of the population over repeated presentations of s_i in the environment A. Similarly, the responses of the population to the same stimulus set can be measured in a different environment, B, where the probability of observing a stimulus is dictated by p_B (s_i). This measurement yields another mean population vector r_B (s_i). Given two environments, A and B, can we describe how r_A (s_i) relates to r_B (s_i)? If so, can such a model predict how the population will behave when it adapts to a new environment C? Here we offer affirmative answers to these questions.

We report that two properties of adaptation capture how the population response to a given stimulus, viewed as a vector, changes across environments. First, the ratio between the response magnitudes is a power law of the ratio between the stimulus probabilities. Second, the response direction to a stimulus is largely invariant. These rules can be used to predict how cortical populations adapt to novel, sensory environments. We show how the power law enables the cortex to preferentially signal unexpected stimuli and to adjust the metabolic cost of its sensory representation to the entropy of the environment. Finally, we offer a model for the emergence of the power law based on auto-encoder models that balance the quality of stimulus representation with its metabolic cost.

1-083. Premotor theta oscillation coordinates articulatory movements during continuous speech production

Yitzhak Norman[1,2] NORMANIK@GMAIL.COM
Loren M. Frank[1] LOREN.FRANK@UCSF.EDU
Edward F. Chang[1] EDWARD.CHANG@UCSF.EDU

[1]University of California, San Francisco
[2]Neurological surgery

Speaking fluently requires rapid and precisely timed motor commands that orchestrate the movement of nearly 100 muscles governing our speech organs. Prior research has demonstrated the significant role of distributed neuronal representations within the sensorimotor cortex (SMC) in governing this motor process. Nevertheless, the mechanism through which the underlying neuronal populations achieve precise temporal coordination and synchronization during continuous speech remains elusive. Using high-density electrocorticography (ECoG) recordings in patients, we identified a prominent ~8 Hz theta oscillation within the local field potential of the sensorimotor cortex (SMC), persisting during both speech and silent intervals. Following speech initiation, a significant increase in theta coherence was observed among speech-responsive sites, primarily clustered within the SMC and the superior temporal gyrus (STG). The distributed theta-coherent network displayed strong phase-amplitude coupling, with increased high-gamma amplitude around theta troughs. Notably, this theta oscillation maintained a stable frequency despite marked variations in speech rate, attesting to its intrinsic origin. To explore potential correlations between this cortical oscillation and concurrent articulatory movements, we employed a task where participants spoke hundreds of sentences in English. Using a deep learning acoustic-to-articulatory inversion technique, we derived articulator movements from the produced speech acoustics and obtained the Articulatory Kinematic Trajectories (AKTs) of the jaw, lips, and tongue in each participant. Analysis of the AKTs revealed semi-rhythmic motor events characterized by abrupt, coordinated changes in AKT velocity, indicative of vocal tract reconfiguration. Strikingly, these coordinated movements occurred at an average rate of 7.9 Hz and exhibited strong coupling to the ongoing theta oscillation in SMC. Our findings reveal an uncharted function of theta oscillation in motor control, hinting at its potential role as an internal pacemaker. This rhythmic maintenance fosters synchronized phases of excitation among distant neuronal populations, facilitating precise timing for the coordinated activation of distributed motor representations and articulatory commands.

1-084. Differential Contributions of Anterior Cingulate and Orbito-Frontal Cortex to action timing and its self-monitoring in rats

Lea Barillier[1] LEA.LEBARILLIER@GMAIL.COM
Valerie Doyere[2] VALERIE.DOYERE@UNIVERSITE-PARIS-SACLAY.FR
Tadeusz Kononowicz[2] T.W.KONONOWICZ@ICLOUD.COM

[1]Institute of Psychology, Polish Academy of Science
[2]CNRS/Universite Paris-Saclay

Performance monitoring is a hallmark of cognition, typically studied in paradigms involving processing of external stimuli. However, whether and how animals monitor performance that originate from internal processing, such as temporally precise self-generated actions, remains unknown. To address that gap, we developed an unprecedented task allowing to study inference of their own timing errors in rats. We then sought to understand which frontal cortical areas carry computations relevant to performance monitoring of internally-generated timing behavior. In this paradigm, rats self-reported temporal errors in their time production (TP) without any external feedback. They produced precise time intervals (>2.4s) by holding a lever and reporter their errors by betting on reward access in Short Error (SE) or Long Error (LE) ports, accurately demonstrating error monitoring of self-generated actions. To investigate the neural bases underlying this ability, we inhibited either the orbitofrontal cortex (OFC), which carries confidence signals, or the anterior cingulate cortex (ACC), which carries error-dependent signals. Rats underwent training until they reached consistent error monitoring behavior. On alternating days, muscimol (GABAergic agonist) or saline was infused in OFC (n=12) and ACC group (n=12). Inhibiting the OFC impaired their ability to produce accurate time intervals. Conversely, for rats in the ACC group, timing performance was not impaired. Fitting the behavioral model showed an overestimation bias to judge their TP as longer under muscimol, suggesting a disruption of temporal error monitoring in ACC. In addition, we fitted several models investigating utilization of history of previous outcomes and learning from negative feedback, thereby uncovering a multitude of performance monitoring aspects inherent in the ACC. The finding of hierarchical and independent from timekeeping inference of temporal errors in the ACC offers valuable insight into timing and performance monitoring architectures.

1-085. Noise resilience of memory stored in low-dimensional manifolds through multiple synaptic timescales

Georg Chechelnizki[1,2] GEORG.CHECHELNIZKI@GMAIL.COM
Nimrod Shaham[1] NIMROD.SHAHAM@MAIL.HUJI.AC.IL
Alon Salhov[1,3] ALON.SALHOV@MAIL.HUJI.AC.IL
Yoram Burak[1,4] YORAM.BURAK@ELSC.HUJI.AC.IL

[1]Hebrew University of Jerusalem
[2]Edmond and Lily Safra Center for Brain Sciences
[3]Racah Institute of Physics
[4]Edmond and Lily Safra Center for Brain Sciences, and Racah Institute of Physics

The reliable storage of information in short term memory is a crucial function of the brain. This function is challenged by the inherently noisy nature of neurons. Representations of continuous variables are particularly sensitive to noise, due to random drift that accumulates over time and can severely degrade memory accuracy. Here we identify a neural mechanism which effectively counteracts such deterioration by employing slow excitatory and fast inhibitory synapses. The mechanism is inspired by the derivative feedback mechanism (Lim and Goldman, 2013), which was previously shown to improve stability to certain forms of parameter mistuning in noise-free networks. We start by examining the simple case of a linear continuous attractor network (CAN), where we show that diffusivity can be made arbitrarily small if the timescale of excitation is larger than that of inhibition. We then show that similar principles generalize to a far more general class of nonlinear CANs. We successfully apply these principles to ring attractor networks, inspired by the insect head direction system. We find that our theory correctly predicts the improvement of memory stability as a function of synaptic timescale differences in such models, when endowed with the derivative feedback mechanism. This offers a plausible explanation for how the brain can stably store memories of continuous parameters, despite the ubiquity of noise. Furthermore, we identify how to engineer connectivity such that modes of activity other than motion along the attractor are not slowed down by the stabilization mechanism, allowing them to relax rapidly and tightly confining neural activity patterns to a one dimensional manifold. Insights from our theory allow us to conclude that neurons in head direction cell networks that are commonly thought to be utilized for velocity integration can also aid in stabilization against noise-driven motion.

1-086. Inferring the control signal driving zebrafish locomotion

Thomas Nathaniel Soares Mullen[1,2] THOMAS.MULLEN@RESEARCH.FCHAMPALIMAUD.ORG
Marine Schimel[3] MMCS3@CAM.AC.UK
Christian Machens[1] CHRISTIAN.MACHENS@NEURO.FCHAMPALIMAUD.ORG
Guillaume Hennequin[3] G.HENNEQUIN@ENG.CAM.AC.UK
Michael Orger[1] MICHAEL.ORGER@NEURO.FCHAMPALIMAUD.ORG
Adrien Jouary[1,4] ADRIEN.JOUARY@RESEARCH.FCHAMPALIMAUD.ORG

[1]Champalimaud Research
[2]Neuroscience Unit
[3]University of Cambridge
[4]Neuroscience

A central objective in neuroscience is to understand how the brain orchestrates movement. Recent advances in automated tracking technologies have made it possible to generate rich behavioral datasets that can be exploited to gain insights into the neural control of movement. One approach to analyzing such data is to identify stereotypical motor primitives using cluster analysis. However, this categorical description can limit our ability to model the effect of more continuous control schemes. Here we take a control theoretic approach to behavioral modeling and argue that movements can be understood as the output of a controlled dynamical system. Previously, models of movement dynamics, trained solely on behavioral data, have been effective in reproducing observed features of neural activity. These models addressed specific scenarios where animals were trained to execute particular movements upon receiving a prompt. In this study, we extend this approach to analyze the full natural locomotor repertoire of an animal: the zebrafish larva. Our findings demonstrate that this repertoire can be effectively generated through a sparse control signal driving a latent Recurrent Neural Network (RNN). Our model's learned latent space preserves features relevant to the fish's navigation while disentangling different categories of movements. Collectively the control signal and dynamics we identified offer a novel framework for understanding neural activity in relation to movement.

1-087. Quantitative modeling of the emergence of macroscopic grid-like representations

Ikhwan Bin Khalid[1,2] IKHWANKHALID92@GMAIL.COM
Eric T. Reifenstein[1] E.REIFENSTEIN@GOOGLEMAIL.COM
Naomi Auer[1] NAOMI.AUER@POSTEO.DE
Lukas Kunz[3] DRLUKASKUNZ@GMAIL.COM
Richard Kempter[4,5] R.KEMPTER@BIOLOGIE.HU-BERLIN.DE

[1]Humboldt Universitat zu Berlin
[2]Institute for Theoretical Biology
[3]University of Bonn Medical Center
[4]Humbold-Universitat zu Berlin
[5]Department of Biology, Institute for Theoretical Biology

Grid cells are neurons in the entorhinal cortex that are thought to perform neural computations in support of spatial navigation. As direct recordings of grid cells from the human brain are only rarely possible, functional magnetic resonance imaging (fMRI) studies proposed and described an indirect measure of entorhinal grid-cell activity, which is quantified as a hexadirectional modulation of fMRI activity as a function of the subject's movement direction through a virtual environment. However, the contributing role of the aggregated activity of grid cells to this modulation remains unclear. Our research addresses the unresolved question concerning the origin of hexadirectional modulation of activity in the entorhinal cortex, as observed in fMRI, iEEG, and MEG studies (e.g. Doeller et al., Nature, 2010; Staudigl et al., Curr Biol, 2018; Convertino et al., Brain, 2023). Here, we explored three hypotheses through both numerical simulations and analytical calculations: head-direction tuning (conjunctive grid by head-direction cell hypothesis); firing-rate adaptation (repetition suppression hypothesis); or a bias towards a certain grid phase offset (structure-function mapping hypothesis). Our findings indicate that, in principle, all three hypotheses can account for hexadirectional modulation of sum grid-cell activity in ideal conditions. However, when including grid-cell properties found in the literature, our simulations most strongly support the conjunctive grid by head-direction cell hypothesis. In contrast, our simulations do not support the structure-function mapping hypothesis. With respect to the repetition-suppression hypothesis, our simulations are insufficient to substantiate or refute it, and further experiments on the adaptation properties of single grid cells are required. Our study is grounded in computational simulations and offers useful insights for systems neuroscience since hexadirectional modulation has been used as a measure of grid cell activity in studies of the brain's navigation and orientation systems.

1-088. Neuromodulatory recurrent neural networks on a timing task

Julia Costacurta[1,2] JCOSTAC@STANFORD.EDU
Shaunak Bhandarkar[1] SHAUNAKB@STANFORD.EDU
David Zoltowski[1] DZOLTOW@STANFORD.EDU
Scott Linderman[1,3] SCOTT.LINDERMAN@STANFORD.EDU

[1]Stanford University
[2]Electrical Engineering
[3]Statistics Department and Wu Tsai Neurosciences Institute

Neuromodulatory signals are powerful and prevalent influences on behavior. Dopamine, a well-known example, is implicated in timing deficits resulting from Parkinson's disease. However, the specific role of neuromodulation within neural circuits is poorly understood. A common way to probe circuit dynamics is by analyzing task-optimized recurrent neural networks (RNNs), yet currently most models fail to incorporate neuromodulatory effects. We propose the neuromodulated RNN (NM-RNN), which consists of two linked subnetworks corresponding to neuromodulation and output generation. We model the output generation subnetwork as a low-rank RNN, a popular model choice due to the constrained dimension of its neural trajectories. In our setting, the low-rank RNN is also appealing because it implies a natural way to implement neuromodulation–we allow the output of the neuromodulatory subnetwork to scale the low-rank factors of the output generation weight matrix. We train NM-RNN models with two scaling schemes (global and factor-specific) on the measure-wait-go (MWG) interval reproduction task, where the network must measure interval durations and produce a ramping output with the same duration. We compare to a low-rank RNN with no neuromodulation. When generalizing to unseen intervals, the factor-specific NM-RNN produces the smallest generalization loss. We next investigated how the neuromodulatory subnetwork contributes to computation. We found that the factor-specific NM-RNN's neuromodulatory signal has channels which are clearly associated with distinct aspects of the task, and ablating these signals destroys performance in a predictable way. For example, one channel of the neuromodulatory signal is associated with terminating the output ramp, and when ablated the network is no longer able to produce a meaningful output. Our results indicate how the NM-RNN distributes different pieces of the task computation across different low-rank factors. In general, the NM-RNN framework offers a novel method to investigate the effects of neuromodulation on neural dynamics in a variety of settings.

1-089. Inter-hemispheric prefrontal mechanisms of within- and across-trial working memory

Melanie Tschiersch[1] MEL.TSCHIERSCH@GMAIL.COM
Joao Barbosa[2,3] PALERMA@GMAIL.COM
Akash Umakantha[4] AKASH.UMAKANTHA@GMAIL.COM
Ryan C. Williamson[4] RWILLIAMSON88@GMAIL.COM
Matthew A. Smith[4] MATTSMITH@CMU.EDU
Albert Compte[1] ACOMPTE@RECERCA.CLINIC.CAT

[1]IDIBAPS
[2]Ecole Normale Superieure
[3]Group for Neural Theory
[4]Carnegie Mellon University

Working memory (WM) refers to the short-term information maintenance and its processing. Neurons in the prefrontal cortex (PFC) exhibit persistent activity to items in WM and display a contralateral preference [1]. Recent evidence points towards an interplay between persistent activity and activity-silent mechanisms in the PFC leading to the attraction of current memories towards previously remembered items (serial dependence) [2]. While memory representations have been shown to transfer between hemispheres [3], it is unclear how and if neural representations of serial dependence relate across hemispheres. Here, we analyzed behavioral and bilateral prefrontal cortex neural data recordings (5708 multi-units) from three monkeys performing a visuo-spatial delayed response task. Monkey behavior showed diffusing memories and serial dependence, and prefrontal neuronal activity was selective to the cue location. Moreover, we found that the correlation between neural predictions and the monkeys' responses increased during the delay, suggesting slowly drifting memory traces as in bump attractor networks [4]. Intriguingly, this memory drift was significantly, but weakly, correlated across hemispheres suggesting inter-hemispheric interactions during memory maintenance. As shown in previous research on the neural correlates of serial dependence, we found that previous-trial delay code reactivated prior to the upcoming stimulus. Additionally, we found that current-trial neural predictions drifted towards the previously memorized location. Surprisingly, while memory diffusion within a trial was correlated across-hemispheres, memories drifted exclusively towards the reactivations occuring within the same hemisphere. Overall, our analyses suggest interacting hemispheres during the delay, but private serial dependence mechanisms. To comprehend the essential

requirements for the observed interactions during the memory and inter-trial interval periods, we propose a two-area bump-attractor network with weak connectivity between the two hemispheric networks and local short-term synaptic plasticity mechanisms in excitatory synapses. References: [1] Funahashi et al., J. Neurophysiol. (1989), [2] Barbosa, Stein et al., Nat. Neurosci. (2020), [3] Brincat, Donoghue et al., Neuron (2021), [4] Wimmer et al., Nat. Neurosci. (2014)

1-090. A multi-area network model of adaptive motor control

Rui Xia[1,2] RX220@CAM.AC.UK
Marine Schimel[1] MMCS3@CAM.AC.UK
Guillaume Hennequin[1] G.HENNEQUIN@ENG.CAM.AC.UK

[1]University of Cambridge
[2]Engineering

Non-stationary environments pose a formidable challenge for biological motor control. To maintain proficiency following a change in context, state estimation dynamics and feedback control loops must be rapidly and jointly reconfigured in a consistent manner. What network principles might underlie coordinated adaptation from limited experience? Here, we propose a multi-area neural network model, in which a prefrontal cortical (PFC) network continuously integrates input from cerebellum (CR) and primary motor cortex (M1). Both CR and M1 are in turn modulated by PFC through simultaneous gating of thalamo-cortical and thalamo-cerebellar loops, enabling coordinated low-rank modifications of their effective connectivity.

We meta-train the network on a synthetic non-stationary motor task, whereby each context is characterized by a 2D flow representing passive hand dynamics. CR is trained to perform adaptive probabilistic state estimation: it must infer (and broadcast) the current hand position from noisy delayed observations and efferent copies of motor commands. M1 is trained to perform control-as-inference: it must infer (and issue) motor commands consistent with optimality of predicted future states w.r.t. a reaching target. PFC is tasked with learning a representation of the current (unobserved) context, used to modulate CR and M1 via thalamus.

We find that this neural architecture supports few-shot learning of the environmental context: state estimation and control both rapidly improve with experience in every novel context. Following a change in passive dynamics, cerebellar predictions approach the fundamental limit of Kalman filtering within a few reaching trials. Concurrently, M1 rapidly learns to speed up target acquisition relative to passive dynamics. Adaptation arises from PFC having learned to orchestrate CR and M1 in a context-dependent manner. This working model can be further refined to incorporate more detailed aspects of the mammalian motor circuitry; it will help integrate neural and behavioral data, and guide future experiments in adaptive motor control.

1-091. Atypical corticocollicular feedback underlies suboptimal Bayesian inference in autism model mice

Leiron Ferrarese[1] LEIRON.FERRARESE@EMBL.IT
Hiroki Asari[1,2] ASARI@EMBL.IT

[1]European Molecular Biology Laboratory
[2]Epigenetics and Neurobiology Unit

Bayesian inference provides a theoretical framework for understanding how the brain integrates current sensory information with past experiences to predict future events and guide an animal's behaviour. Anomaly in this process is thought to underlie various symptoms in autism spectrum disorder (ASD), such as sensory abnormalities. It remains unclear, though, which neuronal circuits are involved and how. Here we focused on the mouse visual system to address how neural circuit computation supports implicit learning about visual environmental contexts and how it is altered in a mouse model of ASD. Specifically, we developed a Bayesian implicit visual learning task, following the ones used in human studies; and combined in vivo two-photon calcium imaging, chemogenetic perturbation, pupillometry, and computational modelling to characterize both neurophysiological and behavioural traits of mice engaged in this task. Our major findings are as follows: 1) wild-type mice showed contextual modulation of the neuronal dynamics in the superior colliculus as well as pupil dynamics in a way consistent with the Bayesian inference principles; 2) Scn2a-haploinsufficient (Scn2a+/-) ASD-model mice exhibited only suboptimal contextual modulations; and 3) the corticotectal feedback input from the primary visual cortex mediated the observed contextual modulations, both at neurophysiological and behavioural levels. These results highlight that top-down sensory modulation plays a key role in the integration of contextual sensory information, and that anomaly in this process is associated with atypical sensory processing in Scn2a+/- ASD-model mice. While we focused on a specific brain region involved in a specific computation, our findings may extend as a more generic

computational pattern of ASD within complex, recursive feedforward-feedback interactions of sensory and associative systems in the framework of the Bayesian brain theory.

1-092. The structure of population activity in mouse visual cortex is stable for weeks

Celian Bimbard C.BIMBARD@UCL.AC.UK
Enny van Beest E.BEEST@UCL.AC.UK
Kenneth Harris KENNETH.HARRIS@UCL.AC.UK
Matteo Carandini M.CARANDINI@UCL.AC.UK

University College London

The activity of neural populations in the cortex is remarkably structured, and this structure is thought to reflect the underlying circuits and guide the resulting computations. For example, some neurons in visual cortex have high correlations with the population ("choristers") while others are more independent ("soloists"). Furthermore, neurons in auditory cortex exhibit stereotyped sequences of activation both in responses to stimuli and in ongoing activity. Is this structure invariant across days and months? Or does it change over time, similarly to the 'representational drift' reported for some sensory representations?

We recorded the activity of large populations of neurons in the visual cortex of awake mice with chronic Neuropixels probes, and tracked cells for over 100 days. For each daily recording and each pair of tracked neurons, we calculated the spontaneous correlation (correlation in spontaneous activity), the typical delay in spontaneous firing (time of peak of the cross-correlogram), and the signal correlation (correlation in the mean responses to two sets of repeats of the same visual stimuli). We also measured stimulus response sequences from the average response times of each neuron to natural images.

Both spontaneous and signal correlations were remarkably stable across days and months (correlation ~0.9 across matrices measured >100 days apart), especially when correlations were calculated at fast timescales (<30 ms bins). The neurons fired in sequences that were also stable across days (~0.8 correlation >100 days apart). These sequences had the same order during spontaneous activity and in response to visual stimuli. Putative excitatory and putative inhibitory neurons were similarly stable.

We conclude that the structure of neuronal population activity in visual cortex is highly stable, opposite to the described representational drift of sensory responses. This suggests that the underlying circuits are hard-wired, limiting not only the patterns that a population may produce but also their plasticity.

1-093. Visual experience instructs the organization of cortical feedback input to primary visual cortex

Rodrigo Dias[1,2] RODRIGO.DIAS@NEURO.FCHAMPALIMAUD.ORG
Radhika Rajan[1] RADHIKA.RAJAN@NEURO.FCHAMPALIMAUD.ORG
Margarida Baeta[1] MARGARIDA.BAETA@RESEARCH.FCHAMPALIMAUD.ORG
Beatriz Belbut[1,3] BEATRIZ.BELBUT@RESEARCH.FCHAMPALIMAUD.ORG
Tiago Marques[1] TIAGO.MARQUES@NEURO.FCHAMPALIMAUD.ORG
Leopoldo Petreanu[1] LEOPOLDO.PETREANU@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Research
[3]Champalimaud Neuroscience Programme

Cortical feedback (FB) projections are thought to modulate lower-order activity depending on learned expectations. Thus, predictions might be constituted of a learned association between activity patterns in higher- and lower-order neurons mediated by FB connections. However, whether FB inputs become bound to specific lower-order neurons depending on experience is unknown. We measured the effects of manipulating visual experience on the retinotopic specificity of supragranular and infragranular projections from the lateromedial (LM) visual area to layer(L) 1 of the mouse primary visual cortex (V1). Using simultaneous dual-color 2-photon calcium imaging we compared the spatial receptive fields (RF) of boutons of LM axons in V1 with those of the neurons below them. While LM inputs were, on average, retinotopically matched with V1 neurons even in dark-reared animals, visual experience degraded their retinotopic precision. Surrounding information was differentially conveyed in the two FB streams, with L5 inputs having a larger proportion of retinotopically unmatched boutons than those originating in L2/3. L2/3 inputs were enriched in distal information from locations lying colinearly to their preferred orientations in dark-reared mice, but this organization was lost with visual experience. Conversely, LM inputs from L5 were depleted of distal information from locations colinear to their preferred orientations in mice with visual

experience but not in dark-reared ones. Thus, the functionally dependent fine scale organization of LM inputs in V1 reflects experienced visual statistics and visual experience exerts different influences on the two FB streams. Restricting visual experience to a narrow range of orientations revealed visual experience exerts an instructive role on the organization of FB inputs. Together, our results show that the organization of LM inputs in V1 reflects experienced visual statistics, supporting theories advocating for a role of descending inputs in shaping the activity of lower-order neurons according to learned associations.

1-094. Stimulus-tuned interneurons accelerate Bayesian sampling in recurrent circuits

Eryn Sale[1,2] ERYN.SALE@UTSOUTHWESTERN.EDU
Wen-Hao Zhang[3] WENHAO.ZHANG@UTSOUTHWESTERN.EDU

[1]University of Texas Southwestern Medical Center
[2]Lyda Hill Department of Bioinformatics
[3]UT Southwestern Medical Center

There is accumulating evidence that stochastic cortical circuits perform sampling-based Bayesian inference to compute the latent stimulus posterior. The canonical cortical circuits consist of excitatory (E) pyramidal neurons and diverse types of inhibitory (I) interneurons. Nevertheless, how sampling is implemented by nonlinear recurrent circuits with types of interneurons remains poorly understood. To provide theoretical insight, we built a nonlinear recurrent circuit model consisting of E neurons and two types of I neurons including Parvalbumin (PV) and Somatostatin (SOM) interneurons. The E neurons are modeled as a canonical ring (attractor) circuit model. To emphasize the role of interneurons, our model simplifies that PV neurons do not have stimulus tunings and only provide global inhibition to E neurons, while SOM neurons are tuned and send tuning-dependent feedback to E neurons. Our theoretical analysis analytically derives the network's sampling dynamics on the stimulus manifold in neuronal population responses, hence theoretically identifying the sampling algorithm of the circuit. We found SOM neurons with tuning-dependent inhibition speed up the sampling in the network via running a Hamiltonian-like sampling algorithm. The Hamiltonian framework also leads to no direct feedforward connections targeting SOM neurons, consistent with anatomy that SOM receives few feedforward synapses. Inactivating SOM neurons in the model slows down the sampling speed via reducing the Hamiltonian sampling into Langevin sampling. Our work provides overarching connections between nonlinear circuits with various types of interneurons and sampling algorithms, deepening our understanding of circuit mechanism of Bayesian inference.

1-095. Competitive learning through fast inhibitory regulation of neural plasticity

Patricia Rubisch[1,2] PATRICIA@RUBISCH.NET
Matthias H. Hennig[1] M.HENNIG@ED.AC.UK

[1]University of Edinburgh
[2]Institute for Adaptive and Neural Computation

While inhibition is a building block of biological neural networks, its role in synaptic plasticity is yet to be fully understood. So far its role has only been addressed in rate and spike-timing based models, yet experimental results show that the hyperpolarising effect of inhibition on the sub-threshold membrane potential can alter the expression of plasticity at excitatory synapses. To address this gap, we explore the effects of inhibition in two voltage-dependent models: the voltage-dependent plasticity model by Clopath et al. (2010), and our new Voltage-Dependent Pathway model (VDP) which captures the experimentally characterised dynamics of Long Term Potentiation (LTP) and Long Term Depression (LTD) induction more closely. We find that our model displays sensitivity to fast sub-threshold fluctuations, and that this reproduces experimental results showing that plasticity is sensitive to the timing of GABAergic inputs. Inhibitory spiking within temporal proximity of a pre-post spike pair in a Spike-Timing-Dependent Plasticity protocol causes the potentiation-to-depression ratio to decrease and thereby strengthens LTD. In recurrent networks, this mechanism further increases circuit-level competition. Laterally connected neurons or networks reacting with short latency to a given input actively increases LTD activity in afferent synapses from the same pathway. We find that this network-driven competition of synaptic weight growth is sufficient for learning factorised neural representations. A fully plastic EI network whose synapses evolve according to the VDP develops receptive fields matching the input statistics of the natural images used for training. This work predicts that competition facilitated by inhibitory plasticity regulation is sufficient for the development of functional networks, and suggests that plasticity is sensitive to fast sub-threshold fluctuations.

1-096. Stable contrast computation of natural visual scenes under dynamic luminance conditions

Luisa Ramirez LUFRAMIREZOC@GMAIL.COM
Burak Gur BGUER@UNI-MAINZ.DE
Marion Silies MSILIES@UNI-MAINZ.DE

Johannes Gutenberg-Universitat Mainz

Visually guided behaviors depend on the reliable extraction of contrast information from natural environments. Under dynamic luminance conditions, e.g. a forest with alternating dark and bright patches, animals can encounter fast luminance changes that can lead to inaccurate contrast estimations. To overcome this problem the visual system implements rapid luminance gain control postsynaptically of photoreceptors, to suppress luminance dependence and drive stable behavior, as recently discovered in flies. However, the visual circuits and computational principles underlying the implementation of this gain control remain unclear. Using two-photon imaging at different levels of the direction-selective (DS) circuits - the main drivers of optomotor behavior - of D. melanogaster, we identify that luminance-invariant contrast responses emerge in the dendrites of the pre-synaptic medulla neurons, Tm1 and Tm9, and propagate to the DS T4/T5 neurons. We propose shunting normalization via spatial pooling as the computational principle underlying this fast gain control. Specifically, we develop a data-driven model to quantify contrast coding reliability in dynamic natural conditions. Our analysis of natural scenes as seen from the perspective of the behaving animal, suggests the existence of an optimal spatial pooling range at which contrast information remains reliable and invariant to fast luminance variations. Using Flywire connectomics analysis, we identify the glutamatergic wide-field neuron Dm12 as a major input to Tm9, and show that GluCla -a driver of shunting inhibition- is a necessary synaptic mechanism to evoke fast luminance gain control in Tm9 medulla neurons. Our results also provide evidence of local spatial pooling, altogether supporting our finding that shunting normalization via spatial pooling is the main mechanism underlying reliable contrast coding in dynamically changing conditions.

1-097. How trees see the forest: A global address space for locally coordinating brain development

Stan Kerstjens[1] KERSTJE@CSHL.EDU
Florian Engert[2] FLORIAN@MCB.HARVARD.EDU
Anthony M Zador[1] ZADOR@CSHL.EDU
Rodney J Douglas[3] RJD@INI.UZH.CH

[1]Cold Spring Harbor Laboratory
[2]Harvard University
[3]Institute of Neuroinformatics, UZH & ETH Zurich

The developing brain faces the daunting task of generating billions of neurons, each with their proper type, morphology, position, and connectivity. This task is challenging from an algorithmic perspective, as each cell only has access to its own internal molecular state, and its immediate extracellular environment. However, targeted long range migration of cells and axons suggest cells have implicit access to spatial knowledge at the whole-brain scale. This raises the question of how global spatial information is distributed to a cell's local environment, and interpreted by the cell. The question is complicated by the limited information capacity of the genome, which disqualifies naive strategies that encode exhaustive instructions for each cell. Here, we propose a simple developmental model whereby cell division induces a global address space. The model incorporates two simple constraints on cell division: Mitotic daughters exhibit gene expression similar to their parent; and daughters remain close to one another in physical space. These constraints implicitly embed the hierarchical lineage relations among cells in spatial patterns of gene expression. This hierarchy acts as an address space by allowing cells, in principle, to estimate their global position from their local environment by comparing local expression to an ancestral expression profile. We tested the predictions of this model with respect to the spatiotemporal organization of gene expression by analyzing spatial transcriptomic data acquired at various developmental stages in mouse and zebrafish. We find that, despite the highly dynamic nature of gene expression during brain development, the spatial covariance of gene expression is organized as a global static hierarchy, as predicted by the model. This hierarchy persists throughout mouse development, and is conserved across mouse and zebrafish brains, suggesting that the address space may play a fundamental role in the developmental algorithms that coordinate tissue organization throughout both ontogeny and phylogeny.

1-098. Representational drift across cortical areas and task demands during spatial navigation

Sofia Soares[1,2] SOFIA_SOARES@HMS.HARVARD.EDU
Shih-Yi Tseng[3] SHIHYI.TSENG@UCSF.EDU
Hanwen Zhang[1] HANWENZHANG@G.HARVARD.EDU
Elena A. Westeinde[1] E.WESTEINDE@G.HARVARD.EDU
Charlotte Arlt[1] CHARLOTTE.ARLT@NATURE.COM
Timothy O'Leary[4] TSO24@ENG.CAM.AC.UK
Christopher Harvey[1] CHRISTOPHER_HARVEY@HMS.HARVARD.EDU

[1]Harvard Medical School
[2]Neurobiology
[3]Harvard Medical School, University of California, San Francisco
[4]University of Cambridge

Representational drift, the change in neural activity patterns linked to sensation, cognition and action over days and weeks, occurs across brain regions and behaviors. However, whether it is a nuisance for brain circuits or serves specific functions remains unclear. Comparing representational drift across brain areas and task demands, while technically challenging, is crucial for uncovering possible functions and mechanisms of drift. Here we provide the first comprehensive comparison of drift across cortical regions and across tasks of varying complexity. We used a random access 2-photon mesoscope to simultaneously image neurons in primary visual (V1), retrosplenial (RSC), posterior parietal (PPC) and secondary motor (M2) cortices, crucial areas for navigational decision-making. We tracked the same neuronal populations over weeks while mice first navigated a maze using only visual landmarks and then learned to navigate using an abstract visual discrimination rule. Across all areas, we found higher drift rates during the landmark task than during the association task. At a population level, task encoding geometry drifted during the landmark task, but stabilized during the association task, indicating an invariant task encoding geometry developed with increased task demands. Additionally, primary motor (M1) and sensorimotor (S1) cortex, together with V1, had greater stability compared to PPC, RSC and M2, indicating that drift is more prominent in relatively more associative areas. These differences persisted after matching encoding properties across areas, indicating that drift is not solely determined by the type of information neurons encode. Our findings are consistent with theoretical models of drift where noise in neural circuits drives them to explore the subspace of possible task representations. Instead of a ubiquitous process inherent to any biological system, we propose that this exploration rate is constrained by task dimensionality and brain-region specific plasticity factors, guiding future experimental and theoretical work.

1-099. Estimating Noise Correlations Across Continuous Conditions With Wishart Processes

Amin Nejatbakhsh[1,2] ANEJATBAKHSH@FLATIRONINSTITUTE.ORG
Isabel Garon[1] IGARON@FLATIRONINSTITUTE.ORG
Alex Williams[1] AWILLIAMS@FLATIRONINSTITUTE.ORG

[1]Flatiron Institute
[2]CCN

The signaling capacity of a neural population depends on the scale and orientation of its noise covariance across trials. Estimating this covariance is challenging and is thought to require a large number of stereotyped trials of a repeated action or stimulus presentation. New approaches are therefore needed to interrogate the structure of neural noise across rich, naturalistic behaviors and sensory experiences, with few trials per condition. Here, we exploit the fact that conditions are smoothly parameterized in many experiments and leverage Wishart process models to pool statistical power from trials in neighboring conditions. We demonstrate favorable performance on experimental data from the monkey motor cortex relative to standard covariance estimators. Moreover, Wishart Processes produce smooth estimates of covariance as a function of stimulus parameters, enabling estimates of noise correlations in unseen conditions as well as continuous estimates of Fisher information—a commonly used measure of signal fidelity. Together, our results suggest that Wishart processes are broadly applicable tools for quantification and uncertainty estimation of noise correlations in trial-limited regimes, paving the way toward understanding the role of noise in complex neural computations and behavior.

1-100. Systems Consolidation of Sequential Dynamics in Model-Based Planning

Oliver Vikbladh[1] O.VIKBLADH@UCL.AC.UK
Neil Burgess[1] N.BURGESS@UCL.AC.UK
Evan Russek[2] EVRUSSEK@GMAIL.COM

[1]Inst of Cognitive Neuroscience, UCL
[2]PNI, Princeton University

It is commonly held that Model-Based (MB) planning is achieved by rollouts through a transition model, implemented through sequential dynamics in the hippocampus (HC). Little direct evidence supports this notion, in part due to the lack of highly powered behavioural paradigms that can measure MB planning, independently from other strategies for action evaluation like the successor representation (SR) (Momennejad et al., 2017). Furthermore, while MB planning needs the HC (Vikbladh et al 2019), rodent work shows that its role may be transient (Bradfield et al 2020), and that transition-models may be consolidated to the cortex over time, consistent with systems consolidation theory (Alvarez & Squire, 1994). Here, across two sessions, seven days apart, we used MEG to measure sequential representational dynamics while people performed a task capable of measuring MB planning using rollouts. Participants that made use of MB planning showed MEG signatures of sequential representational dynamics. Consistent with systems consolidation theory we also show that across sessions, the representations underlying these sequential dynamics shifted from right HC to PFC.

1-101. Neuromodulated mixture of experts: A prefrontal cortex inspired architecture for lifelong learning

Kanha Batra[1,2] KBATRA@UCSD.EDU
Anousheh Bakhti-Suroosh[1] ABAKHTIS@HEALTH.UCSD.EDU
Clara Yi[1] CNYI@UCSD.EDU
May Chan[1] MCHAN@SALK.EDU
Ruby Tseng[1] RUTSENG@UCSD.EDU
Kelbi Banducci[1] KBANDUCCI3@CSUB.EDU
Jiaqi Zhang[1] JIZ100@UCSD.EDU
Christopher Lee[1] CRLEE@HEALTH.UCSD.EDU
Tanisha Roy[1] TAROY@UCSD.EDU
Kelly Chang[1] K7CHANG@UCSD.EDU
Romy Wichmann[1] RWICHMANN@SALK.EDU
Caroline Jia[1] CLJIA@HEALTH.UCSD.EDU
Ben Tsuda[1] BTSUDA@HEALTH.UCSD.EDU
Bryan Nielsen[1] BRYAN@SALK.EDU
Laurel Keyes[1] LKEYES@SALK.EDU
Tristan Tuazon[1] TTUAZON@SALK.EDU
Talmo Pereira[1] TALMO@SALK.EDU
Terrence Sejnowski[1] TERRY@SNL.SALK.EDU
Kay Tye[1] TYE@SALK.EDU

[1]Salk Institute for Biological Studies
[2]Systems Neuroscience

Lifelong learning is the ability of a system to learn and retain knowledge of multiple tasks without catastrophic forgetting, switch between them seamlessly and use old knowledge to facilitate more efficient learning of new tasks. Despite recent advances in artificial intelligence, this problem still hasn't been solved efficiently, with most solutions focused on network expansion (Rusu et al. 2016, Vecoven et al. 2020) – a costly mechanism with limited connect to biology. In this study, we have devised a novel modular deep learning network architecture called Neuromodulated Mixture of Experts (NeMoE) inspired by the prefrontal cortex (PFC) that utilizes the distributed learning framework of the classical Mixture of Experts model and the context-dependent signal-to-noise ratio mediating capabilities of neuromodulators like Dopamine (Vander Weele et al. 2018, Tsuda et al. 2020). To test the model, we developed a novel multi-context "seasonal" foraging task where mice are presented with different environmental contexts indicated by ambient lighting (green/UV) – each context is paired with a high/low cost (shock) and a high/low reward (ensure). We found that mice learn context-specific behavioral strategies and that context is predictable from behavioral features. Further, we used deep reinforcement learning simulations to train NeMoE on the task and found that the model converges to the same policy as real mice for each context – displaying bio-mimicking capabilities. Lastly, we recorded neural activity from the medial prefrontal cortex (mPFC) while the mice were performing the task. We were able to identify ensembles of neurons that have context-specific signal profiles and found that context is decodable from neural activity. Together these findings suggest

that neuromodulation driven flexibility can enable models to perform lifelong learning and such “experts” can be found in mPFC ensembles.

1-102. Spiking neurons discover features from few examples

Timo Wunderlich TIMO.WUNDERLICH@BIH-CHARITE.DE
Robert Gutig ROBERT.GUETIG@BIH-CHARITE.DE

Charite – Universitatsmedizin Berlin

Humans and other animals can learn and generalize remarkably well from little training data. For instance, a few exposures to an unfamiliar face or an unknown word are sufficient for us to reliably recognize the novel sensory feature in different contexts and noise. In machine learning, good generalization is associated with realizing large distances (i.e., large margins) between training examples and a classifier's decision surface. However, the implementation of large-margin techniques in spiking neurons has been prevented by the lack of suitable distance measures for a neuron's spiking responses. Here we use the recently introduced notion of critical thresholds to implement margin learning in spiking neurons. This learning enables spiking neurons to discover, without receiving timing information, recurring input spike patterns embedded in random activity from orders of magnitude fewer training examples than previous work and achieve higher robustness to noise. We derive an upper bound for the noise robustness of spiking neuron models. To obtain a biologically plausible approximation of gradient-based learning, we average gradients conditioned on spike timing and postsynaptic voltage. The resulting voltage dependent STDP kernels match the performance of the exact gradients and provide testable predictions for electrophysiological experiments. We validate our model by applying it to in vivo recordings of mice watching a movie as well as to a spoken digit recognition task. In the first case, neurons discover repeating patterns in experimentally recorded spike trains that correspond to repeating visual input. When applied to human speech, neurons learn to discover recurring digits contained in unsegmented sequences of spoken digits. Our work translates the pivotal concept of margin learning to spiking neurons. By validating the enhanced learning capabilities of single neurons, we highlight the theory's relevance for electrophysiology as well as for learning in larger networks of spiking neurons.

1-103. Neural dynamics in prefrontal regions as a candidate mechanism for instantiating belief states

Sandra Romero Pinto[1,2] SROMEROPINTO@G.HARVARD.EDU
Jay Hennig[3] JAY.A.HENNIG@GMAIL.COM
Daigo Okada[4] OKD-D5226@G.ECC.U-TOKYO.AC.JP
Celia Benquet[5] CELIA.BENQUET@EPFL.CH
Mark Burrell[3] MHBURRELL@FAS.HARVARD.EDU
Scott Linderman[6,7] SCOTT.LINDERMAN@STANFORD.EDU
Sam Gershman[3] SAM.GERSHMAN@GMAIL.COM
Naoshige Uchida[3,8] UCHIDA@MCB.HARVARD.EDU

[1]Harvard university
[2]Center for Brain Science , Department of Molecular and Cellular biology
[3]Harvard University
[4]Tokyo University
[5]Ecole Polytechnique Federale de Lausanne
[6]Stanford University
[7]Statistics Department and Wu Tsai Neurosciences Institute
[8]Department of Molecular and Cellular Biology

It is thought that animals learn to predict future outcomes through reinforcement learning (RL). RL models typically assume a set of fully observable ‘states’, yet in natural environments, states are often hidden and must be inferred. Under such circumstances, a sensible approach is to perform RL over ‘belief states’ (the posterior probability over possible states). Previous studies1,2 showed that dopaminergic reward prediction errors (RPEs) in probabilistic rewards – a task condition involving state uncertainty – can be better explained by RL models harnessing belief states. Furthermore, recent work has shown that recurrent neural networks (value-RNNs) trained on the same tasks exhibit neural dynamics that can be interpreted as belief states3. To elucidate neural mechanisms, here we performed electrophysiological recordings across prefrontal cortical regions in mice. We observed dynamic activity patterns transitioning between two states corresponding to the inter-trial and inter-stimulus interval (ITI and ISI) states, which resemble the dynamics observed in the value-RNNs. When a cue only partially predicted reward (P=90%; Task 2), the ISI state was leaky, spontaneously drifting back to the ITI state, while the ISI state

was substantially stable when a cue predicted reward deterministically (P=100%; Task 1). To infer the underlying mechanisms, we fit switching linear dynamical system models. In addition to a fixed stable point corresponding to the ITI state, another stable fixed point for the ISI state emerged in the model fit to the activity in Task 1, but not in Task 2. Furthermore, the stable ISI state was absent early during learning, indicating that the observed dynamics were not just a result of transient responses to sensory stimuli. Altogether, we propose that the task-dependent neural dynamics observed in the prefrontal regions are a likely candidate mechanism for instantiating belief states in the brain.

1-104. Time-limited integration windows constrain and organize hierarchical computation in ferret auditory cortex.

Magdalena Sabat[1,2] MAGDALENA.SABAT@ENS.PSL.EU
Hortense Gouyette[1] HORTENSE.GOUYETTE@CLUB-INTERNET.FR
Flavien Feral[1] FERAL@BIOLOGIE.ENS.FR
Quentin Gaucher[1] GAUCHER.QUENTIN@GMAIL.COM
Samuel Norman-Haignere[3,4] SAMUEL_NORMAN-HAIGNERE@URMC.ROCHESTER.EDU
Yves Boubenec[1] YVES.BOUBENEC@ENS.FR

[1]Ecole Normale Superieure
[2]Department of Cognitive Studies
[3]University of Rochester Medical Center
[4]Biostatistics & Computational Biology, Neuroscience

The auditory cortex must integrate across a wide range of timescales to derive meaning from natural sounds such as speech, music, and animal vocalizations. But despite longstanding interest, much remains unknown about the cellular mechanisms of multiscale computation in the auditory cortex. One central question is whether neurons in the auditory cortex analyze sounds within a time-limited integration window, and if so, what range of timescales these windows span, how they are organized, and whether they reflect a property of the auditory system or a property of the sounds analyzed by the auditory system (e.g., the duration of a vocalization). To answer these questions, we leveraged a new experimental and computational method that tests whether there exists a time-limited window outside of which stimuli no longer alter the response (the temporal context invariance or TCI paradigm). Using this paradigm, we show that virtually all neurons in ferret auditory cortex exhibit a clear integration window (ranging from ~15 to 125 ms), outside of which surrounding stimuli have very little effect on the response. These windows show clear anatomical organization, being similar for spatially proximal neurons and increasing substantially from primary- to non-primary regions. Moreover, we find that integration windows are largely unaffected by the category of sound or the duration of sound structures (e.g., vocalization duration), even in higher-order, non-primary regions. These results indicate demonstrate that multiscale neural computation in the auditory cortex is strongly constrained and organized by a set of short-duration (<125 ms) integration windows whose temporal extent is predominantly determined by the structure of the auditory system and not the structure of natural sounds.

1-105. Neural landscape diffusion as an organizing framework for brain state dynamics

Ethan Richman[1] RICHMAN@STANFORD.EDU
Karl Deisseroth[1] DEISSERO@STANFORD.EDU
Liqun Luo[1] LLUO@STANFORD.EDU
Nicole Ticea[1] NTICEA@STANFORD.EDU
William Allen[2] WE.ALLEN@GMAIL.COM

[1]Stanford University
[2]Harvard University

Animals perform flexible goal-directed behaviors to satisfy their basic physiological needs. However, little is known about how unitary behaviors are chosen under conflicting needs. Here we reveal principles by which the brain resolves such conflicts between needs across time. We developed an experimental paradigm in which a hungry and thirsty mouse is given free choices between equidistant food and water. We found that mice collect need-appropriate rewards by structuring their choices into persistent bouts with stochastic transitions. High-density electrophysiological recordings during this behavior revealed distributed single neuron and neuronal population correlates of a persistent internal goal state guiding future choices of the animal. We captured these phenomena with a mathematical model describing a global need state that noisily diffuses across a shifting energy landscape. This model is based on a novel conceptual resemblance between our findings of intermixed neural networks

and Langevin diffusion dynamics. Model simulations successfully predicted behavioral and neural data, including population neural dynamics before choice transitions and in response to optogenetic thirst stimulation. These results provide a general framework for resolving conflicts between needs across time, rooted in the emergent properties of need-dependent state persistence and noise-driven shifts between behavioral goals. Moreover, the results suggest an organizing framework for internal states of the brain, in which spontaneous shifts between distributed internal states structure regional and sensory-evoked neural dynamics.

1-106. Adaptation shapes the representational geometry in mouse V1 to encode the statistics of the environment

Mario Dipoppa[1] MDIPOPPA@G.UCLA.EDU
Ramon Nogueira Manas[2,3] RN2446@COLUMBIA.EDU
Stephane Bugeon[4] BUGEON.STEPHANE@GMAIL.COM
Yoni Friedman[5] YONIIFRIEDMAN@GMAIL.COM
Charu Reddy[4] C.REDDY@UCL.AC.UK
Dario Ringach[1] DARIORINGACH@ICLOUD.COM
Kenneth Miller[2] KENDMILLER@GMAIL.COM
Matteo Carandini[4] M.CARANDINI@UCL.AC.UK
Stefano Fusi[6] SF2237@COLUMBIA.EDU

[1]UCLA
[2]Columbia University
[3]Center for Theoretical Neuroscience
[4]University College London
[5]Massachusetts Institute of Technology
[6]Columbia University, Zuckerman Institute

Neural responses dynamically change as a function of previous stimuli, a phenomenon known as sensory adaptation, which can profoundly impact our perception. While adaptation has been mostly studied at the single neuron level, to understand its impact on the neural code, it is key to study it at the neural population level. Therefore, we asked how the geometry of the neural representations adapts to environments with different sensory statistics and what the computational benefits of these adaptation effects are. We addressed these questions experimentally and theoretically. We recorded the responses of thousands of neurons in the mouse's primary visual cortex to gratings whose orientation was drawn from uniform or biased distributions. We fully characterized the dependence of the representational geometry on the stimuli distribution. When considering the discrimination performance between any pair of stimulus orientations, we observed that in the biased sequence, the distance between the representation of the adaptor and stimuli with similar orientation increased, leading to more accurate decoding of the adaptor stimulus. This is surprising, given that the responses of neurons tuned to the adaptor decreased. We adopted an efficient coding approach to determine the optimal geometry under different environment statistics. We trained autoencoder models with metabolic constraints to reconstruct the stimuli under uniform or biased statistics. Under biased statistics, the autoencoder was more penalized when misclassifying the adaptor's orientation. Within a broad hyperparameter region, the changes in tuning curves and population geometry in the model's hidden layer were consistent with our experimental observations. In conclusion, adaptation-induced changes in the neuronal populations are orchestrated so that the decoding accuracy of overrepresented stimuli increases even though responses to those stimuli decrease in magnitude. Our model suggests that these changes allow the brain to improve the representation of frequently presented stimuli while limiting the metabolic cost.

1-107. How brains "jazz": neural mechanisms for producing flexible vocal sequence in parrots

Zhilei Zhao[1,2] ZZ367@CORNELL.EDU
Caleb Jones[1] CJ397@CORNELL.EDU
Han Kheng Teoh[1] HT452@CORNELL.EDU
Julie Carpenter[1] JC2824@CORNELL.EDU
Brian Kardon[1] BMK27@CORNELL.EDU
Jesse Goldberg[1] JESSEHGOLDBERG@GMAIL.COM

[1]Cornell University
[2]Neurobiology and Behavior

Novel action sequences can be generated by varying and combining basic building blocks. Important examples include human speech and music, where an infinite number of vocal utterances can be constructed from a finite

set of acoustic units. Circuit mechanisms underlying this process remain elusive. We study this question using a new bird model, budgerigar parrots. Like humans, budgerigars learn vocalizations throughout life and produce highly flexible songs called warble, where the birds combine simple syllables into complex ones and vary syllable order to produce novel sequences. These features of warble present a unique opportunity to study how brains generate flexible action sequences. We took a comparative approach between parrots and well-studied songbirds, e.g. zebra finches, that produce stereotyped songs. We first developed deep learning methods to uncover warble acoustic and syntactic structures. Using reversible brain inactivation, we found that an anterior forebrain pathway (AFP) is critical: bilateral AFP inactivation abolishes warble production while unilateral inactivation reduces the sequence length and complexity, indicating inter-hemisphere coordination. Blocking NMDA receptors in the primary motor region that receives AFP inputs also disrupts warble production. These results form stark contrasts to those in songbirds, where neither AFP nor NMDA receptors in the song system output nuclei are necessary for producing stereotyped songs, demonstrating the importance of synaptic plasticity for flexible sequence generation. We recorded single unit activity in AFP and observed many neurons that show premotor bursts before the onset of certain syllables or at a specific moment within the syllable. So far, some of these representations, including ultra-sparse bursts time-locked to acoustic elements, resemble HVC neurons that famously participate in neural synfire chains. However, we also found neuronal types in AFP that are unique to parrots. Together, our data revealed novel circuit mechanisms underlying the production of complex and naturalistic sequences of actions.

1-108. Behavioral adaptation to changing energy constraints via altered frequency of movement selection

Geoffrey Goodhill[1] G.GOODHILL@WUSTL.EDU
Robert Wong[1] ROBERT.WONG@WUSTL.EDU
Thomas Darveniza[1] DARVENIZA.T@WUSTL.EDU
Shuyu Zhu[1] ZHU.S@WUSTL.EDU
Zac Pujic[2] ZAC.PUJIC@GMAIL.COM
Biao Sun[2] B.SUN1@UQ.EDU.AU
Matthew Levendosky[1] MATTHEWL@WUSTL.EDU
Ramesh Agarwal[1] RKA@WUSTL.EDU
Michael McCullough[3] MICHAEL.MCCULLOUGH@ANU.EDU.AU

[1]Washington University in St Louis
[2]University of Queensland
[3]Australian National University

Animal behavior is strongly constrained by energy consumption. A natural manipulation which provides insight into this constraint is development, where an animal must adapt its movement to a changing energy landscape as its body grows. Unlike many other animals, for fish it is relatively easy to estimate the energy consumed by their movements via fluid mechanics. Here we simulated the fluid mechanics of >100,000 experimentally-recorded movement bouts from larval zebrafish across different ages and fluid conditions as they hunted Paramecia. We find that these fish adapt to their changing relationship with the fluid environment as they grow by adjusting the frequency with which they select different types of movements, so that more expensive movements are chosen less often. This strategy was preserved when fish were raised in an unnaturally viscous environment. This work suggests a general principle by which animals could minimize energy consumption in the face of changing energy costs over development.

1-109. Dissociation of efference copy and afferent feedback signals in somatosensory cortex

Xinyue An[1,2] XINYUE.AN@NORTHWESTERN.EDU
Raeed Chowdhury[3,4] RAEED.CHOWDHURY@PITT.EDU
Josh Glaser[1] J-GLASER@NORTHWESTERN.EDU

[1]Northwestern University
[2]Department of Neurology
[3]University of Pittsburgh
[4]Bioengineering

Motor brain regions send efferent signals not just to descending spinal pathways, but also to many other brain areas. This enables discriminating self-generated from externally-generated movements, and allows for more timely estimation of body state (proprioception). One area that likely participates in the integration between in-

coming sensory (afferent) feedback and efference copy signals is area 2 of somatosensory cortex, a brain area that processes proprioceptive information. While we know area 2 receives information from motor cortex, we don't understand how these signals interact. Here, we investigated the interaction of efferent and afferent signals in area 2, using a neural decoding-based approach. Different relationships between efferent and afferent signals would support opposing hypotheses of their interaction within area 2. If the two population signals are orthogonal, it would suggest separate representations of efference and afference. However, if the two signals are aligned, it would suggest a direct update on the efference copy by afferent feedback. Our decoding analyses discovered two separable signals within the area 2 population, occurring at distinct time shifts relative to movement, and which resided on orthogonal neural axes– this supports a hypothesis enabling simultaneous separate representations. To further explore the nature of these representations, we compared these signals during active movements and passive perturbations. We found the pre-movement signal to be absent during passive perturbations, confirming our interpretation of it being related to efference copy. Unique from previous results, we discovered feedback signals for voluntary movements arrived slower than those from external perturbations. A simple model with suppression of afferent feedback during active movements recapitulates this finding. Overall, our findings differentiate between hypotheses of efferent/afferent interactions within area 2, find a novel time-delay between afferent feedback in active and passive movements, and demonstrate a broadly applicable approach to dissect efferent and afferent signals within neural population data.

1-110. Mixed coding in the mouse cerebellar nuclei during a locomotion obstacle avoidance task

Ramin Khajeh[1,2] RK2899@COLUMBIA.EDU
Qianyun Zhang[1] QZ2311@COLUMBIA.EDU
Richard Warren[3] RICHARDWARREN2163@GMAIL.COM
Larry F Abbott[1] LFA2103@COLUMBIA.EDU
Nathaniel B. Sawtell[1] NS2635@COLUMBIA.EDU

[1]Columbia University
[2]Neurobiology & Behavior
[3]CTRL-Labs

The cerebellum is involved in a wide range of sensory, motor, cognitive, and affective processes, but how the output of the cerebellum, most of which flows through the deep cerebellar nuclei (DCN), supports these diverse functions remains poorly understood. Towards this long-term goal, we combined multi-site extracellular recordings from the mouse DCN with high-resolution kinematic tracking in a previously developed head-fixed obstacle avoidance task (Authors et al., 2021). Analysis of neuronal activity in relation to fine locomotor kinematics and speed, obstacle anticipation, detection and traversal, jaw and whisker kinematics, and reward (water) delivery, revealed robust firing rate modulations in neurons recorded across all three cerebellar nuclei. Strikingly, individual neurons across all three nuclei often encoded multiple task events and kinematic variables concurrently. Furthermore, we commonly observed individual neurons that encoded details of limb kinematics, interlimb changes or orofacial movement on timescales of tens of milliseconds while also exhibiting firing rate modulations that ramp up or down to obstacle contact over hundreds of milliseconds. We consider these results in the context of both traditional and emerging views of the anatomical and functional organization of cerebellar output pathways.

1-111. Subspace transformations underlie decision-driven working memory prioritization

Huidi Li[1,2] HUIDILI@MIT.EDU
Jonas Rose[3] JONAS.ROSE@RUHR-UNI-BOCHUM.DE
Scott Brincat[1] SBRINCAT@MIT.EDU
Earl Miller[1] EKMILLER@MIT.EDU

[1]Massachusetts Institute of Technology
[2]brain and cognitive sciences
[3]Institute for Cognitive Neuroscience Ruhr-University Bochum

Decision making and working memory are two core cognitive functions that have been thoroughly examined in separate behavioral paradigms. But in real-life problem solving, they interact—decisions can guide what information is prioritized in working memory. To understand this interaction, we recorded from lateral prefrontal cortex (LPFC) of non-human primates performing a task in which reward-based decisions guided selection of items in working memory. Subjects were presented with a series of two spatial targets, each paired with a cue indicating its reward value, and they had to select the location of the higher-value target. We found that decisions rotated

orthogonal working memory subspaces to alignment based on their priority states. Before decisions, LPFC represented the locations of both targets shown in a trial with approximately equal strength with orthogonal subspaces, consistent with maintaining separate high-quality representations of both potential choices. After decisions, the subspace for the selected target was rotated so that location representations were aligned whether they arose from the first- or second-presented target, simplifying read-out for behavior. Moreover, we observed that decision led to a greater decoding accuracy and drove a neural code expansion for selected targets, consistent with it being prioritized in working memory. These findings illustrate that decision guides the prioritization of working memory by inducing subspace transformations which rotate working memory subspaces based on their priority states and expand neural codes for the prioritized working memory.

1-112. Allothetic and Idiothetic Control of Theta Phase Coding

Yotaro Sueoka[1,2] YSUEOKA1@JHMI.EDU
Ravikrishnan Jayakumar[1] RPERURJ1@JHU.EDU
Manu Madhav[3,4] MANU.MADHAV@UBC.CA
Francesco Savelli[1] FSAVELLI.RESEARCH@GMAIL.COM
Noah Cowan[1] NCOWAN@JHU.EDU
James Knierim[1] JKNIERIM@JHU.EDU

[1]Johns Hopkins University
[2]Department of Neuroscience
[3]University of British Columbia
[4]School of Biomedical Engineering

Phase precession by hippocampal place cells is one of the most well-studied forms of phase coding (O'Keefe & Recce, 1993). As an animal traverses through a place field, the place cell fires at progressively earlier phases of the LFP theta (~8 Hz) oscillation. Place cells also exhibit phase procession towards the end of the field (Wang et al., 2020). Phase precession and procession occur at different phase ranges of the theta rhythm, suggesting the co-existence of distinct phase coding mechanisms segregated by theta phase. While the activity of hippocampal place cells is controlled by external landmarks (allothetic cues) and self-motion-based idiothetic cues, with allothetic cues showing dominant control over the hippocampal map when the two cues are in conflict (Jayakumar, Madhav, et al., 2019), how these spatial cues interact to form the intricate phase coding remains unknown. Recently, Chu, Ji, et al. (2023) reported that a continuous attractor network with a feedback inhibition term can produce units that exhibit both phase precession and procession. Using this model, we show that coherence in allothetic and idiothetic inputs is selectively required for phase procession, but not precession. To test the model's predictions, we analyzed the activity of hippocampal place cells as rats circumnavigated a planetarium-style, VR environment that creates a persistent conflict between allothetic and idiothetic inputs (Jayakumar, Madhav, et al., 2019). In 40/51 sessions across five rats, place fields were maintained in the landmark frame. While the overall phase coding structure was maintained relative to the allothetic cues, increased conflict between allothetic and idiothetic cues disrupted phase procession with minimal effect on phase precession. The results i) provide experimental support to the attractor network-based model of phase coding and ii) uncover the differential control mechanism of the two modes of phase coding by allothetic and idiothetic inputs.

1-113. Simultaneous, cortex-wide and cellular-resolution neuronal population dynamics reveal an unbounded scaling of dimensionality with neuron number

Alipasha Vaziri[1] VAZIRI@ROCKEFELLER.EDU
Jason Manley[2] JMANLEY@ROCKEFELLER.EDU

[1]Rockefeller University
[2]jmanley@rockefeller.edu

The brain's remarkable properties arise from collective activity of millions of neurons. Widespread application of dimensionality reduction to multi-neuron recordings implies that neural dynamics can be approximated by low-dimensional "latent" signals reflecting neural computations. However, what would be the biological utility of such a redundant and metabolically costly encoding scheme and what is the appropriate resolution and scale of neural recording to understand brain function? Imaging neural activity at cellular resolution and near-simultaneously across mouse cortex, we demonstrate an unbounded scaling of dimensionality with neuron number up to one million neurons. While half of the neural variance lies within sixteen behavior-related dimensions, we find this unbounded scaling of dimensionality corresponds to an ever-increasing number of internal signals without immediate behavioral or sensory correlates. The activity patterns underlying these higher dimensions are fine-grained

and cortex-wide, highlighting that large-scale recording is required to uncover the full neural substrates of internal neuronal computations.

1-114. The hippocampus as an unpredicted map: Hippocampal traces of uncertainties induced by changes in reward distributions.

Charline Tessereau[1] CHARLINE.TESSEREAU@INTERNATIONALBRAINLAB.ORG
Jack Mellor[2] JACK.MELLOR@BRISTOL.AC.UK
Peter Dayan[1] PETER.DAYAN@TUEBINGEN.MPG.DE
Feng Xuan[3] FENG.XUAN@NORTHWESTERN.EDU
Dan Dombeck[3] D-DOMBECK@NORTHWESTERN.EDU

[1]Max Planck Institute for Biological Cybernetics
[2]University of Bristol
[3]Northwestern University

In volatile environments, humans and animals face different forms of uncertainty to which they must adapt to thrive. However, our understanding of the neural basis of this adaptation is incomplete, despite, for instance, long-standing arguments about its possible dependence on neuromodulation. Here, we take advantage of the well-known spatial remapping of hippocampal place cells in the face of environmental change to interrogate these processes. We performed calcium imaging in CA1 place cells in the Uncertain Reward virtual reality Task (UR-Task) [1]. In this, mice run along a linear track and lick for a water reward whose precise location on any run may be more or less certain in a block (a form of expected uncertainty), and which might also translate without warning (unexpected uncertainty). From limpid changes in reward locations, the place map undergoes remapping that is zone-dependent, affecting most the cells between reward zones, and visible as early as within the first trial after the switch. Given inherently variable reward locations, the place map within the reward zone fractionates, with some cells coding distance from reward, other cells firing reliably in relationship to space, and a third group being co-modulated, with reward and space-related fields. Single cell analysis highlights that, while some cells are uniquely modulated by reward or position, there is a rich interplay between position and reward contributions to the place code. We design generalized linear mixed models that can quantify the unique contribution of position and reward (both predictive and responsive) on each individual cell activity and help cluster the place map. We find more co-modulation between reward and position of the cells under expected than unexpected uncertainty, in which cells seem to cluster as only reward- or position-based.

1-115. The representational geometry of hierarchical decision-making processes

Isabella Rischall[1,2] IMR8238@NYU.EDU
Braden Purcell[1] PURCELBA@GMAIL.COM
SueYeon Chung[3,4] SCHUNG@FLATIRONINSTITUTE.ORG
Roozbeh Kiani[1] ROOZBEH@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]New York University; Flatiron Institute
[4]Center for Computational Neuroscience

Goal-directed behavior in dynamic, naturalistic environments depends on a hierarchy of decision-making processes that operate at different timescales to infer context, adjust decision policy, and map stimuli to actions (Purcell & Kiani, 2016; Sarafyazd & Jazayeri 2019). Still, the neural mechanisms underlying hierarchical decisions are poorly understood. Here, we study such decision processes by recording from large neural populations in the lateral intraparietal area (LIP), supplementary eye field (SEF), and dorsolateral prefrontal cortex (dlPFC) of macaque monkeys performing a variant of the direction discrimination task with random dots. Monkeys classify the motion direction of a random dots stimulus, while also tracking a hidden, spontaneously changing context variable. When the context changes, monkeys receive negative feedback until a 'switch' target is chosen. Characteristic of hierarchical decisions, negative feedback ambiguously marks either a failed direction classification or failure to recognize an environment change. Monkeys perform the task well, integrating motion information within trials to make direction choices, as well as integrating confidence about motion direction and feedback across trials to make switch decisions. Both integration processes are reflected in neural responses, with many neurons in each area representing the motion and switch decision variables (DVM and DVS). In each region, neural population activity organizes into curved manifolds representing the DVM. The DVM manifolds shift along axes in state space that represent the environments and DVS, with unique geometric motifs in each brain region.

Critically, the code for DVM and DVS are factorized in the neural state space, enabling the population to concurrently integrate sensory information within a trial and feedback and confidence across trials in order to meet the demands of hierarchical decisions.

1-116. Integration of behavioral related correlation from top-down and bottom-up pathways in mouse V1

Peijia Yu[1] PEIJIAY@UCHICAGO.EDU
Ha Yun Anna Yoon[2] ANNA_YOON@BERKELEY.EDU
Yuhan Yang[2] YUHANY1024@GMAIL.COM
Olivia Gozel[1] GOZEL@UCHICAGO.EDU
Na Ji[2] JINA@BERKELEY.EDU
Brent Doiron[1] BDOIRON@UCHICAGO.EDU

[1]University of Chicago
[2]University of California, Berkeley

Brain-wide neuromodulation by behavioral variables, such as locomotion, pupil area, and face motion, have been observed in mice (Stringer et al., 2019; Musall et al., 2019). To study this mechanism at higher spatial resolution, we used two-photon imaging that allows recording of individual neuronal and synaptic bouton activity in mouse primary visual cortex (V1), while the animal's face was simultaneously videotaped. We aim to understand how facial motion is related to the population activity of cortical neurons and their lateral geniculate nucleus (LGN) afferents, both during spontaneous and stimuli evoked periods. To avoid spurious correlations due to overfitting from naive linear regression analysis, we applied a variant of ridge regression analysis, where its regularization hyperparameter is optimized to minimize the correlation between facial motion and neuron/bouton activity after random trial permutations. We observed a robust correlation between facial motion and neuronal population activity, which is higher for visually evoked response compared to spontaneous activity. In contrast, LGN bouton activity does not correlate with facial motion during spontaneous periods, but surprisingly becomes significantly correlated for visually evoked responses. To explain this last observation we show that LGN boutons are almost silent during the spontaneous period, in contrast to their high activity during evoked states, implying that the LGN is subthreshold during the spontaneous period and cannot transfer any received information about facial motion. This prompts the hypothesis that the improved encoding of facial motion variables in V1 cortical neurons when visually evoked is mainly due to the onset of facial motion correlated bottom-up visually evoked LGN inputs, rather than stronger top-down movement-related cortical inputs. In total, our work gives an unprecedented analysis of the circuit pathways that underlie the recent observations that mouse V1 activity is related, in part, to non-visual inputs.

1-117. Thalamocortical interactions shape hierarchical neural variability during stimulus perception

Raul Adell[1,2] RAULADELLSEGARRA@GMAIL.COM
Adria Tauste[1] ADRIA.TAUSTE@GMAIL.COM
Antonio Zainos[3] AZAINOS@IFC.UNAM.MX
Yuriria Vazquez[4] YUVAZUSTER@GMAIL.COM
Manuel Alvarez[3] ALVAREZ@ASTROSEN.UNAM.MX
Gustavo Deco[5] GUSTAVO.DECO@UPF.EDU
Sergio Parra[3] SPARRA@IFC.UNAM.MX
Ranulfo Romo[3] RROMO@IFC.UNAM.MX
Roman Rossi-Pool[3] ROMANR@IFC.UNAM.MX

[1]Universistat Politecnica de Catalunya
[2]Computational Biology and Complex Systems Research group (BIOCOM-SC)
[3]Universidad Nacional Autonoma de Mexico
[4]Rockefeller University
[5]Universitat Pompeu Fabra

Brain circuits require functional diversification to efficiently process sensory signals. The hierarchy in its neural code permits increased complexity and representation of abstract concepts. But, to what extent do functional connections within and across areas shape this hierarchical order? We addressed this problem in the thalamocortical network, while monkeys judged the presence or absence of a vibrotactile stimulus. We quantified the variability and functional connectivity in simultaneously recorded neurons sharing the same cutaneous receptive field from the somatosensory thalamus (VPL) and areas 3b and 1 from the somatosensory cortex. We examined two

sources of spiking variability: Within-trial temporal variability (intrinsic timescale), and inter-trial variability (Fano Factor). Neuronal variability increased along the VPL-3b-1 network during the stimulus in both metrics, showing slight differences between VPL and 3b, and a notable distinction between VPL and area 1. Upon unraveling information flow dynamics across and intra-areas, we report that VPL establishes transient parallel feedforward connections multiplexing to 3b and 1 areas. VPL and 3b display fast dynamics with rapid feedforward interactions. In contrast, area 1 shows slower timescales with persistent intra-area interactions. Our findings indicate that the lower variability of VPL and 3b facilitates feedforward thalamocortical communication, while the higher variability of area 1 supports intra-cortical interactions during sensory processing. Moreover, contrary to the hypothesis suggesting 3b acts as a relay to area 1 and both cortical areas sharing similar characteristics, our results suggest 3b $\rightarrow$ area 1 can be regarded a secondary feedforward station in the touch processing route and these cortical areas exhibit clear distinct functionalities. The coordination of variability measures, intrinsic timescales, and directed information measures provides compelling evidence, for the first time to our knowledge, for the existence of a variability hierarchy within the somatosensory network. Furthermore, complementary and interlinked aspects of the network infrastructure are obtained.

1-118. The spatial structure of surround modulation in mouse visual cortex depends on layer and experience

Beatriz Belbut[1,2] BEATRIZ.BELBUT@RESEARCH.FCHAMPALIMAUD.ORG
Serena Di Santo[3,4] SERENADISANTO@UGR.ES
Agostina Palmigiano[5] AP3676@COLUMBIA.EDU
Margarida Baeta[1] MARGARIDA.BAETA@RESEARCH.FCHAMPALIMAUD.ORG
Tiago Marques[1] TIAGO.MARQUES@RESEARCH.FCHAMPALIMAUD.ORG
Kenneth D. Miller[5] KDM2103@COLUMBIA.EDU
Leopoldo Petreanu[1] LEOPOLDO.PETREANU@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Neuroscience Programme
[3]Universidad de Granada
[4]Electromagnetismo y Fisica de la Materia
[5]Columbia University

Contextual information is key for visual perception. Neurons in the primary visual cortex (V1) of mice respond differently to stimuli in their receptive field (RF) depending on the context given by surrounding locations. Surround modulation (SM) depends on stimuli features and may reflect expected natural visual statistics. Consistent with this, the local and feedback connections thought to underlie SM follow anisotropic, tuning-, and experience-dependent wiring rules. However, whether SM shows systematic asymmetries that depend on experienced visual statistics is unknown. Here, we used large scale two-photon recordings in mouse V1 to investigate the spatial structure of SM in L2/3 and L5 neurons. Small circular disks, with moving gratings, aligned to the recorded neurons' RF, were presented, either alone, or in the presence of surrounding gratings patches. The magnitude of SM varied depending on the position, orientation, and direction of the stimuli surrounding their RF. Notably, higher suppression was elicited when surround stimuli were above the RF position, when surround gratings were collinearly aligned to the center position, or when gratings were nasally-moving. Moreover, SM anisotropies depended on the soma's layer, on anesthesia and running state. Further, multiple SM asymmetries, but not all, depended on visual experience, as they were absent or altered in dark reared mice. The observed SM anisotropies were recapitulated by a circuit model with spatial and orientation specific asymmetric connectivity. The model successfully links competitive interactions across space with connectivity anisotropies to give rise to anisotropic SM. Together, these results show systematic asymmetries in the SM of V1 neurons and suggest that contextual modulations in V1 may reflect adaptations of the cortical circuitry to the visual statistics experienced by mice.

1-119. Neural population latent dynamics predict cognitive state errors related to schizophrenia

Samantha Brunson[1,2] BRUNS319@UMN.EDU
Matthew Chafee[1] CHAFE001@UMN.EDU
Audrey Sederberg[3,4] AUDREY.SEDERBERG@GATECH.EDU

[1]University of Minnesota
[2]Graduate Program in Neuroscience
[3]Georgia Institute of Technology
[4]School of Psychology and School of Physics

Schizophrenia is a devastating mental illness affecting 1% of the global population. Cognitive symptoms in schizophrenia are highly predictive of patient outcomes, but are untreated by current medications. It is currently unclear how these symptoms, such as deficits in cognitive control, are generated in the brain. We hypothesize that cognitive control requires the creation and use of cognitive states, such as holding a certain stimulus in working memory and using that to determine a course of action to take, and that these cognitive states are related to neural states - the coordinated patterns of activity across populations of cells in relevant brain regions. To examine the link between neural states and cognitive states, we analyzed single-trial neural population activity recorded from nonhuman primates (NHPs) performing a cognitive control task. In the framework of a hidden Markov model, we inferred latent states from neural population recordings (N = 20 to 40 neurons) to identify shared patterns of cognitive states and transitions in this task across disparate brain regions, including the prefrontal and parietal cortex. Under a pharmacological model of schizophrenia in NHPs that has been previously shown to replicate some of the cognitive deficits of the disease, we found that neural states trajectories were disrupted. Errors can be predicted from the inferred neural activity states during stimulus presentation, before a response has been made, potentially implicating changes to the stability of cognitive states in the cognitive symptoms of schizophrenia. The results of this project will help to guide the design of future experiments continuing to investigate cognitive states and cognitive symptoms in schizophrenia. This research into improving the understanding of cognitive changes in schizophrenia can lead to newer and more efficacious treatments targeting an important aspect of the disease.

1-120. Normalized Cuts Characterize Visual Recognition Difficulty of Amorphous Image Sub-parts

Shuchen Wu[1] SHUCHEN.WU@TUEBINGEN.MPG.DE
Mehmet Yoerueten[2] MEHMETYORUTENN@GMAIL.COM
Felix Wichmann[2] FELIX.WICHMANN@UNI-TUEBINGEN.DE
Eric Schulz[1] ERIC.SCHULZ@TUEBINGEN.MPG.DE

[1]Max Planck Institute for Biological Cybernetics
[2]University of Tuebingen

Upon glimpsing at an image, we effortlessly perceive structures from within. What charac- terizes this process? Historically, gestalt psychologists have suggested that people tend to group nearby similar image parts together as a whole. Can an algorithm that partitions images into sub-parts based on similitude characterize visual perception behavior? We look at the normalized min-cut algorithm and its correlation to the recognition difficulty of image parts. The algorithm transfers an image seg- mentation problem into a graph-cutting problem, approximating an energy optimization problem that preserves within-graph similarities. We study whether the number of computational steps needed for the algorithm to isolate an image part correlates with participants' difficulty in recognizing that part, and whether higher exposure time correlates with further computational steps. We propose a psy- chophysics paradigm to study subjects' recognition behavior upon seeing images tiled by amorphous subparts. We found that an increasing cut no. of image subpart (more computation steps) is harder for subjects to recognize after a brief exposure time, and longer exposure time increases the recognition ease, consistent with the model's prediction that higher cut no. demands more computation steps to isolate. Our study relates the recognition difficulty of image parts with the computational resources needed to solve an optimization problem of grouping by similarity.

1-121. Excitatory-inhibitory assemblies in olfactory memory networks

Claire Meissner-Bernard[1] CLAIRE.MEISSNER-BERNARD@FMI.CH
Friedemann Zenke[1] FRIEDEMANN.ZENKE@FMI.CH
Thomas Frank[2] THOMAS.FRANK@BI.MPG.DE
Rainer Friedrich[1] RAINER.FRIEDRICH@FMI.CH

[1]Friedrich Miescher Institute for Biomedical Research
[2]European Neuroscience Institute Gottingen

Classically, memories are thought to be encoded in assemblies of highly connected excitatory (E) neurons. However, recent studies indicate that inhibitory (I) neurons also exhibit specificity in their connectivity, suggesting that information is represented by assemblies containing both E and I neurons. To understand how E-I assemblies affect memory function, we use the telencephalic area Dp, the zebrafish homolog of olfactory cortex, as a model system. Indeed, this brain area shows structural and functional features of a memory network as well as signatures of tuned inhibition. We built a biologically plausible spiking neural network model of Dp in which learned odors are subsequently stored by increasing the connectivity between defined subsets of E and I neurons. This establishes co-tuning of excitatory and inhibitory currents in individual neurons, as observed experimentally. We

found that within neural population state space, E-I assemblies confine odor-evoked activity to lower-dimensional subspaces than random connectivity, but to a lesser extent than E assemblies which form basins of attraction. Learned odors and degraded versions thereof drive activity onto specific, yet only partially isolated regions of neural state space, resulting in a tradeoff between stimulus discrimination and robustness to noise. We then refined the model to account for our observation that I neurons in Dp are divided into at least two subtypes that differ in their connectivity profiles, and uncovered another interesting phenomenon: in the presence of E-I assemblies, for some odor pairs, correlations between odor-evoked activity patterns increase strongly upon partial silencing of one of the two inhibitory subtypes. Importantly, we observed the same phenomenon during optogenetic manipulation of the same inhibitory subtype in ex-vivo Dp, suggesting that E-I assemblies are likely to exist in Dp. Taken together, our results provide valuable insights into how specific E-I circuit motifs shape network activity and function.

1-122. A neural circuit mechanism in medial entorhinal cortex for integrating event duration

John Bowler[1] JACK.BOWLER@UTAH.EDU
Hyun-Wo Lee*[2] HYUNWOO.LEE@NEURO.UTAH.EDU
Erin Bigus[2] ERIN.BIGUS@UTAH.EDU
Carlos Martinez-Navarro[2] CARLOS.MARTINEZ-NAVARRO@NEURO.UTAH.EDU
Jim Heys[2,3] JIM.HEYS@NEURO.UTAH.EDU

[1]University of utah
[2]University of Utah
[3]Neurobiology

Animals utilize interval timing—measuring time in the range of seconds to minutes—to plan and execute a wide variety of behaviors, including foraging, capturing prey and evading threats. In this project we have focused on how medial temporal lobe structures might play a previously unrecognized role in interval timing by developing novel timing paradigms that require more flexible, context-dependent learning. We have discovered that medial entorhinal cortex (MEC), traditionally known for spatial memory, may also function as a timing system. To test this hypothesis, we developed a temporal delayed nonmatch-to-sample task (tDNMS) in which mice must learn to report whether pairs of odor stimuli are match or nonmatch in duration (Bigus et al., 2023, bioRxiv). We combined this behavioral approach with cellular resolution 2-photon Ca2+ imaging in MEC and chemogenetic manipulation. Our results show that from the start of each trial, elapsed time is encoded by the regular, sequential activity of populations of MEC "time cells." This work underscores the critical role of the MEC in learning more complex and flexible timing behaviors. In this abstract, we use a combination of experimental and computational approaches to answer three key questions to pinpoint the neural mechanisms in MEC that are the basis for this learning: Neural Dynamics: How do the neural dynamics of MEC time cells evolve to support learning, shifting from context-independent to context-dependent activity during the tDNMS task? Circuit Mechanism: What circuit mechanism within the MEC drives the time-locked sequential firing of time cells, and does it resemble the continuous attractor network dynamics used by MEC grid cells for spatial navigation? Cognitive Strategy: Which cognitive strategy do mice employ to solve the tDNMS time task, how does this strategy impact their ability to perform the task and how is the strategy reflected in the neural dynamics?

1-123. Reactivation in the human brain connects the past with the present

Avital Hahamy[1] A.HAHAMY@UCL.AC.UK
Haim Dubossarsky[2] HAIM.DUB@GMAIL.COM
Timothy Behrens[3] BEHRENS@FMRIB.OX.AC.UK

[1]University College London
[2]Queen marry University of London
[3]University of Oxford & University College London

To understand why you are reading this abstract, you may recall you volunteered to review Cosyne abstracts. This exemplifies that our present experiences are linked to relevant past events, yet the neural mechanisms underlying this process are currently unknown. We hypothesized that a process termed "replay" underlies this linking. Originally observed in rodents during navigation tasks, replay involves the rapid reactivation of cell firing-patterns related to previous locations, seemingly binding these locations into an internal model of the environment. We hypothesize that replay could similarly bind episodic events into an internal model of evolving experience. This hypothesis cannot be studied in experiments based on repeated trials, typical of rodent studies, since the neural representations of repeated events are virtually indistinguishable. However, in evolving experiences, like

narratives, each event has a unique representation that can be studied in humans. We therefore developed a method to measure the reactivation of past events in fMRI data of participants engaged in narratives. We indeed found that the human brain reactivates neural representations of past narrative-events. Similar to replay in rodents, these reactivations occur in hippocampus and regions of the default mode network. However, in humans, these reactivations do not occur during prolonged offline periods, but at the boundaries between ongoing narrative-events. Furthermore, in the precuneus/retrosplenial cortex, these reactivations consist of past information that is specifically relevant for the comprehension of the current narrative stage. These results, replicated across two independent datasets, demonstrate selective reactivations of relevant individual experiences from the past. Such a mechanism has been postulated, for example, as a substrate for solving the challenging temporal credit assignment problem in reinforcement learning, but has never been demonstrated. Furthermore, our findings dramatically expand the scope of replay, suggesting it acts as an online sense-making mechanism that interprets incoming information in light of past experiences.

1-124. Persistent representation of socioemotional information in parallel with encoding of context

Nick Frost[1,2] NICKFROSTNEURO@GMAIL.COM
Kevin Donohue[3] KEVIN.DONOHUE@UCSF.EDU
Vikaas Sohal[3] VIKAAS.SOHAL@UCSF.EDU

[1]University of utah
[2]Department of Neurology
[3]University of California, San Francisco

The prefrontal cortex is critical for the regulation of normal social interactions and anxiety-related behaviors which take place in varied contexts over the life of an individual. Critically, contextual information is also encoded by the prefrontal cortex, necessitating that information relating to location or context be encoded in parallel with information relevant to social and emotional behaviors. We sought to understand how dynamic context-driven changes in prefrontal activity affect the representation of information relevant to different behaviors. To answer this question, we employed microendoscopic imaging of prefrontal neurons in freely moving mice engaged in social interaction or anxiety-related behaviors across different contexts. We show that ensemble activity within the prefrontal cortex undergoes dynamic, context-dependent re-organization such that distinct patterns of activity underlie social interactions which occur in different contexts. Despite these widespread changes, individual neurons which are active during social interactions in one context remain active during social interactions occurring within other contexts, and that these persistent representations of social information are possible at the ensemble level because social ensemble representations are orthogonal from those which encode context. Similarly, while we observe context dependent remapping of prefrontal activity as mice explore different anxiety-provoking mazes, invariant representations of anxiety-related regions emerged within each maze. Like social encoding, information related to anxiety and context are encoded by population vectors oriented orthogonally from each other. These anxiety-related ensembles were distinct from activity observed during exploration of other mazes and were driven by the persistent modulation of anxiety-encoding neurons within anxiety-provoking regions within each maze. Taken together these data suggest that orthogonal representations of different types of information are encoded by prefrontal microcircuits working in parallel to permit simultaneous encoding of both context and behavior-specific information.

1-125. Learning dynamics in the PFC can be explained by an external controller

Joe Pemberton[1,2] JOEPEMBERTON1995@GOOGLEMAIL.COM
Michal Wojcik[3] MICHAL.WOJCIK@PSY.OX.AC.UK
Rui Ponte Costa[3] RUI.COSTA@DPAG.OX.AC.UK

[1]University of Bristol
[2]Computer Science
[3]University of Oxford

The prefrontal cortex (PFC) exhibits a remarkable capacity to employ two distinct strategies when engaging in cognitive tasks. Upon encountering a novel task, it leverages high-dimensional representations, well positioned for rapid linear decoding. However, with growing task familiarity, the PFC transitions to employing generalisable low-dimensional neural codes. Through a system-level modelling approach, we propose that these two strategies can operate harmoniously at different timescales, ensuring a shift from high-to-low dimensional representations. Specifically, we introduce a controller-based framework in which: (i) on a faster timescale control signals drive

recurrent networks to generate task-encoding but relatively unstructured, high-dimensional representations, which is then followed by (ii) a slower optimisation of recurrent connections and consequently more structured, low-dimensional representations. We validate these predictions by analysing PFC neural dynamics at fast and slow learning scales in non-human primates that were trained to learn a complex cognitive task from scratch. In summary, our results suggest a learning-dependent control of prefrontal dynamics via a separate brain-region for high-to-low representational switching.

1-126. A biophysically detailed cortical circuit model to map spike sorting biases in dense recordings.

Steeve Laquitaine[1,2] STEEVE.LAQUITAINE@EPFL.CH
Milo Imbeni[3] MILO.IMBENI@GMAIL.COM
Joseph Tharayil[4] JOSEPTH.THARAYIL@EPFL.CH
Michael W. Reimann[5] MICHAEL.REIMANN@EPFL.CH

[1]EPFL - The Swiss Federal Institute of Technology
[2]Blue Brain Project
[3]Universiteit Maastricht
[4]Blue Brain Project, The Swiss Federal Institute of Technology
[5]ecole Polytechnique Federale de Lausanne (EPFL)

Why do modern dense extracellular recordings and spike sorting still miss most isolatable neocortical neurons? This “dark matter” problem in neuroscience may stem from spike sorters’ frequent failure to assign spikes to neurons, based on extracellular features that vary due to cortical tissue heterogeneity. Left uncorrected for, poor sorting undersamples the neuron population activity, produces spurious spike trains and neurons, and hides the population’s true firing rate distribution, which can lead to erroneous claims on neural coding and computations. If one knows the ground truth, it becomes possible to measure the true sorting accuracy, identify the nature and sources of errors, and improve existing algorithms to maximize the neuron yield. Hybrid synthetic datasets, generated by adding a small homogeneous set of spike waveforms sorted from real recordings to a simulated background noise, have been crucial to developing spike sorting algorithms. Yet they underestimate the diversity of spike spatiotemporal shapes, and strong assumptions about firing rate statistics are made during their generation. Here, we pursue a complementary approach, evaluating Kilosort 3.0, the state-of-the-art spike sorter, against simulated Neuropixels 1.0 recordings generated from a recently published, real-scale, comprehensive biophysically detailed in silico model of the rat’s neocortical circuitry. The resulting traces, spike shapes, and sorted firing rate distributions were analogous to those from real recordings. We found that Kilosort 3.0 severely undersampled the cortical population activity, correctly sorting 10% of the isolatable neurons, in agreement with reported in vivo yields. We also found that poor sorting was caused by overmerging and oversplitting biases (for 25% of the units). Sparse firing rates and significant unit distance to the nearest recording site also significantly predicted poor sorting accuracy.

1-127. Bridging reinforcement learning and causal inference: Emergence of grid cells in abstract tasks

Adithya Gungi[1,2] AG4472@COLUMBIA.EDU
Pradyumna Sepulveda[1] PS3345@CUMC.COLUMBIA.EDU
Ines Aitsahalia[1] IFA2108@CUMC.COLUMBIA.EDU
Matthew Schafer[1] MS6883@CUMC.COLUMBIA.EDU
Kyo Iigaya[1] KI2151@COLUMBIA.EDU

[1]Columbia University
[2]Physics

The ability to discover causal, rather than correlational, structures in the world is a hallmark of biolog- ical intelligence. Recent behavioral and neural evidence (Jeong et al., Science 2022) has challenged the sufficiency of reinforcement learning (RL) models in capturing causal learning. This raises a question — Do we need to abandon RL models to explain causal learning? Here, we propose a novel computational model that captures causal structure learning, while at the same time preserving the RL paradigm. Existing model-based computational frameworks, such as model-based RL, typically assume pre- defined world structures and avoid structure learning. Bayesian inference models, although capable of learning full probabilistic structures, lack biological plausibility due to their high computational demands. Most existing models also overlook the human inclination to infer deterministic structures even when none exist (Redelmeier and Tversky, 1996). To address these gaps, we propose a computational model that builds a deterministic causal graph. Our model leverages a nonlinear

transformation of the successor representation (Dayan, 1993), learned through an RL algorithm. Our model explains various empirical findings, including spatial navigation, classical conditioning, a mixture of model-based vs. model-free learning, and information seeking behavior. Our model also offers novel interpretations of grid cells as a neural basis for building causal graphs. Importantly, we predict the emergence of grid-cell structures in almost all experimental tasks with trial regularities, including two-step tasks and classical conditioning. Taken together, our study offers a novel computational framework to bridge RL and causal learning. Our findings suggest the deterministic causal graph as a unifying mechanism for a range of phenomena that have been studied separately. Our model predicts that grid cells support abstract causal planning across various decision-making tasks.

1-128. Reversed depth illusion in central vision revealed by backward masking as theoretically predicted

Li Zhaoping[1,2] LI.ZHAOPING@TUEBINGEN.MPG.DE

[1]Max Planck Institute for Biological Cybernetics and University of Tubingen
[2]Sensory and Sensorimotor Systems

Human observers can discriminate depths of surfaces in 3-dimensional (3D) random-dot stereogram (RDS) scenes. In anti-correlated RDSs, a black dot in one eye corresponds to a white dot in the other eye, making V1 neurons respond as if the depth order between surfaces is reversed from that defined by binocular disparities (Cummings & Parker 1997). Zhaoping & Ackermann (2018) found that peripheral but not central vision can perceive this reversed depth according to V1 responses, when observers reported whether a disk was in front of a surrounding ring. They explained this finding in terms of a Central-peripheral Dichotomy (CPD) theory, proposing that feedback from higher to lower visual areas (e.g., V1) mainly targets central, rather than peripheral, visual fields and aids recognition by verifying potential perceptual outcomes using analysis-by-synthesis. Accordingly, the reversed depth illusion from misleading V1 responses is vetoed in central vision by the verification process. Feedback verification is weaker in peripheral vision, and so the illusion is perceived.

The verification process in central vision can be compromised by presenting difficult visual inputs briefly, followed by a mask — backward masking. This replaces the original details by the mask when feedback verification occurs, making verification difficult (DiLollo et al 2000). If the reversed depth illusion in central vision is indeed normally vetoed, then backward masking should make the illusion visible. We tested this prediction by making the RDS dynamic: the random set of dots in the RDS was replaced every 10 milliseconds by another random set while keeping the 3D scene unchanged. Each replacement random set acts as a backward mask for the previous random set. Indeed, in central vision, observers see the reversed depth illusion better in the dynamic than static RDSs. This finding supports the analysis-by-synthesis computation in the feedback process, and the CPD theory in particular.

1-129. Reinforcement of valence through action

Jasmine Stone[1,2] JASMINETSTONE@GMAIL.COM
Fatima Amin[3] FATIMA.AMIN@LIN-MAGDEBURG.DE
Benjamin Bargeron[4] BENJAMIN.BARGERON@MPFI.ORG
Oliver Barnstedt[3] OLIVER.BARNSTEDT@LIN-MAGDEBURG.DE
Salil Bidaye[4] SALIL.BIDAYE@MPFI.ORG
Bertram Gerber[3] BERTRAM.GERBER@LIN-MAGDEBURG.DE
Ilona C. Grunwald Kadow[5] ILONA.GRUNWALD@UNI-BONN.DE
Marcel Heim[6] MICHAEL-MARCEL.HEIM@CHARITE.DE
Christian Konig[3] CHRISTIAN.KOENIG-BETHKE@LIN-MAGDEBURG.DE
Utsab Majumder[5] MAJUMDER@UNI-BONN.DE
Nino Mancini[4] NINO.MANCINI@MPFI.ORG
David Owald[6] DAVID.OWALD@CHARITE.DE
Anna Pierzchlinska[7] APIERZCH@UNI-KOELN.DE
Ashok Litwin-Kumar[1,8] A.LITWIN-KUMAR@COLUMBIA.EDU

[1]Columbia University
[2]Neurobiology & Behavior; Center for Theoretical Neuroscience
[3]Leibniz Institute for Neurobiology
[4]Max Planck Florida Institute for Neuroscience
[5]University of Bonn
[6]Charite – Universitatsmedizin Berlin
[7]University of Cologne
[8]Neuroscience

Associating a stimulus with rewarding or punishing reinforcement leads to the production of approach or avoidance behaviors in response to subsequent stimulus presentations. We examined whether such behaviors themselves, absent external reinforcement, may serve as reinforcers of stimuli, using the Drosophila fruit fly model system. In this animal, the activation of descending neurons called "moonwalker neurons" promotes backward walking, an avoidance behavior. We found that odors paired with optogenetic activation of moonwalker neurons acquire a negative valence, leading to subsequent avoidance of those odors. Dopamine neurons in the fly mushroom body, an associative learning center, are activated during moonwalker neuron activation. A model of the mushroom body as a reinforcement learning system demonstrates that this influence of avoidance behavior on dopamine neuron activity prolongs the lifetime of avoidance memories. Such an effect may provide a resolution to the so-called "avoidance paradox"—the observation that animals continue to avoid a conditioned stimulus even after they have learned to successfully avoid the associated punishment. Our study establishes for the first time that, in flies, the production of an avoidance behavior can support associative learning. Such observations may also apply to other species, including mammals, in which dopamine activity has also been observed to correlate with both reinforcement prediction error and movement.

1-130. Distinct and asymmetric neuronal responses to pitch-tilt axis and roll-tilt axis vestibular stimulation in larval zebrafish

Sharbatanu Chatterjee[1,2] SHARBATANU444@GMAIL.COM
Geoffrey Migault[3] MIGAULT@HOTMAIL.FR
Natalia Beiza-Canelo[3] NBBEIZA@GMAIL.COM
Georges Debregeas[4] GEORGES.DEBREGEAS@SORBONNE-UNIVERSITE.FR
Volker Bormuth[4] VOLKER.BORMUTH@SORBONNE-UNIVERSITE.FR

[1]Sorbonne Universite
[2]Laboratoire Jean Perrin
[3]Sorbonne Universite, CNRS, Institut de Biologie Paris-Seine (IBPS), Laboratoire Jean Perrin (LJP)
[4]Sorbonne University

Posture is crucial to executing motor functions, and perturbations of posture trigger various behavioural strategies to regain the preferred posture. The vestibular system plays a vital role in responding to movement-related cues to maintain the preferred direction, and its dysfunction is associated with neurological disorders. Previous work has introduced a miniaturised, rotating, light-sheet microscope for brain-wide neuronal recordings during controlled vestibular stimulation in head-restrained zebrafish larvae to study whole-brain circuits. In this study, we present significant enhancements to such a miniaturised rotating light-sheet microscope setup to enable precise rotational movements along arbitrary axes between roll-tilt and pitch-tilt (nose moving up and down) directions. The improved design is 3D printable and incorporates a double galvanometer mirror configuration, facilitating replication and enhancing scanning capabilities. Using this design, we conducted the first mapping of larval zebrafish brain-wide responses to dynamic vestibular stimulation along the pitch-tilt axis, a dimension of vestibular processing that is important for movement initiation but has not been explored with whole-brain imaging. Our findings revealed an asymmetry in the number of responsive neurons between nose-up and nose-down tilt responses, displaying distinct spatial patterns. Furthermore, we conducted recordings during roll-tilt stimuli in the same fish, revealing a greater number of responsive neurons across the brain, especially in the cerebellum, during roll stimuli, compared to pitch. In addition, we identified a transgenic line that exhibited significant overlap with these functional maps, offering exciting opportunities for further investigation using fluorescence-mediated tracing, optogenetics, or targeted ablation studies. These results reveal how whole-brain circuits respond to vestibular stimuli by recruiting distinct and asymmetric neuronal populations based on the axis of the stimulus.

1-131. Inhibitory cell-type-specific-connectivity underlie computations in mean-field model of V1

Soon Ho Kim[1,2] SOONHOKIM@GATECH.EDU
Hannah Choi[1] HANNAHCH@GATECH.EDU

[1]Georgia Institute of Technology
[2]School of Mathematics

Inhibitory neurons in the cortex have distinct cell types with heterogeneous connectivity and physiological properties. Parvalbumin- (PV) and somatostatin-expressing (SST) neurons, two major inhibitory neuron subtypes, perform distinct arithmetic operations on their target excitatory (Exc) neurons. However, debate remains about the precise effects that they have and their computational roles. Here we develop a spatially extended neu-

ral network model of the primary visual cortex (V1) with inhibitory cell types with distinct structural and dynamical properties. In particular, we incorporate recent evidence of long-range spatial projections in SST neurons. A mean field theory is developed in order to compute the stationary firing rates in response to visual stimuli combined with optogenetic stimulation. Simulations show that previous models are unable to simultaneously reproduce experimental results from local and lateral inhibition of PV and SST neurons. By incorporating realistic connectivity motifs, physiological parameters, and cell type-specific projection lengths, our model is able to reproduce both effects. Three-population position-dependent mean field equations are derived and validated against simulations, and balanced solutions are derived from scaling arguments. It is shown that PV and SST stimulation have different effects on excitatory-inhibitory balance of the network, which critically effects their modulations. Our approach offers insight into the relationship between cell type-specific structure and their computational function.

1-132. Connectome-constrained deep mechanistic networks enable hypothesis generation and refinement

Janne Lappalainen[1] JANNE.LAPPALAINEN@UNI-TUEBINGEN.DE
Fabian D. Tschopp[2] NAIBAF710@GMAIL.COM
Sridhama Prakhya[2] PRAKHYAS@JANELIA.HHMI.ORG
Mason McGill[2] MASON.B.MCGILL@GMAIL.COM
Aljoscha Nern[2] NERNA@JANELIA.HHMI.ORG
Kazunori Shinomiya[2] SHINOMIYAK@JANELIA.HHMI.ORG
Shin-ya Takemura[2] TAKEMURAS@JANELIA.HHMI.ORG
Eyal Gruntman[2] GRUNTMANE@JANELIA.HHMI.ORG
Jakob Macke[1] JAKOB.MACKE@UNI-TUEBINGEN.DE
Srinivas C. Turaga[2] TURAGAS@JANELIA.HHMI.ORG

[1]University of Tubingen
[2]HHMI Janelia Research Campus

We can now measure the connectivity of every neuron in a neural circuit, but we are still blind to other biological details, including the dynamical characteristics of each neuron. The degree to which connectivity measurements alone can inform understanding of neural computation is an open question. Empirically, many neural circuit architectures can perform the same large diversity of computational tasks. Further, circuits with the same connectivity can function differently. For these reasons, there has been considerable debate about the utility of connectomic measurements for understanding brain function. Is it possible to use measurements of neural connectivity to generate accurate predictions about how the neural circuit functions? Even in the complete absence of direct measurements of neural activity from a living brain? Here we show that with measurements of only the connectivity of a biological neural network, we can predict the neural activity underlying a specified neural computation. We constructed a model neural network with the experimentally determined connectivity for 64 cell types in the motion pathways of the fruit fly optic lobe but with unknown parameters for the single neuron and single synapse properties of each cell type. We then optimized the values of these unknown parameters using techniques from deep learning, to allow the model network to detect visual motion. Our mechanistic model generates detailed experimentally testable predictions for each neuron in the connectome. We found that model predictions agreed with experimental measurements of neural activity across 26 studies. Our work demonstrates a strategy for generating and testing detailed hypotheses about the mechanisms of neural circuit function from connectivity measurements.

1-133. Ultrafast connectivity optimization of large-scale biophysical network models with deep learning

Nicholas Tolley[1,2] NICHOLAS_TOLLEY@BROWN.EDU
Stephanie Jones[1] STEPHANIE_JONES@BROWN.EDU

[1]Brown University
[2]Neuroscience

Understanding the relationship between network connectivity and emergent neural dynamics is a fundamental and unsolved problem in neuroscience. Detailed biophysical models can simulate highly realistic neural circuits with biologically interpretable parameters, providing a powerful way to study neural dynamics. However, their use is challenged by an overwhelming number of model parameters, computationally expensive simulations, and complex mappings from model parameters to simulation outputs. Previous work has demonstrated that deep neural networks (surrogate models) can be trained to approximate compartmental neuron models, offering simulation speeds that are orders of magnitude faster. Here, we extend this work by implementing surrogate models of individual cells and using a spiking neural network (SNN) architecture to connect them into a biophysical

network model. This construction permits efficient gradient-based optimization of cell-cell connectivity parameters, and the ability to optimize to complex neural activity patterns. We demonstrate the effectiveness of this approach by using surrogate models to approximate a detailed model of the neocortex, the Human Neocortical Neurosolver (HNN). As a proof-of-concept, we infer the strength of connectivity among neurons that gives rise to 15-60 Hz oscillations from noisy background drive, with corresponding predictions on cell spiking activity.

1-134. Distinct cortical mechanisms for egocentric vs. allocentric planning.

Jingjie Li[1] JINGJIE.LI.21@UCL.AC.UK
Chaofei Bao[2] CHAOFEI.BAO@UCL.AC.UK
Liujunli Li[3] LL3462@NYU.EDU
Ziqian Ariel Xu[2] ZIQIAN.XU.21@UCL.AC.UK
Qianbo Grayson Yin[2] YGRAYSON@UMICH.EDU
Jeffrey Erlich[2,4] J.ERLICH@UCL.AC.UK

[1]Sainsbury Wellcome Centre, UCL
[2]University College London
[3]New York University Shanghai
[4]Sainsbury Wellcome Centre

In the real world, movement planning is complex. We can receive instructions in world-centred (allocentric) or self-centred (egocentric) reference frames, which can have different meaning depending on our current heading and position. For example, when you ask for directions to the nearest pub, you can be told to turn to the right or to turn north. Data from typical 2-alternative forced choice tasks indicates that the frontal orienting field (FOF) in rodent M2 causally contributes to planning of orienting movement. Can FOF plan in both egocentric and allocentric reference frames? To answer this, we developed two novel auditory memory-guided orienting tasks using an 8-port wall. In the world-centered 'allo' task, sound A tells the animal to move to the bottom-left port and sound B to the bottom-right port, regardless of their start position. In the self-centered 'ego' task, sound A tells the animal to move to the left and sound B to the right of their start position. In both tasks, FOF activity represents the current position of the animal throughout the fixation period. However, in the ego task, the FOF uses a common neural code for planning across all start positions (and across correct and error trials). In the allo task, the code did not generalize across start positions, and planning activity is later and weaker than in the ego task. Our results suggest that the FOF does not play a universal role in planning for orienting movements, but instead its role seems specific to situations where the planning can be easily done in an egocentric reference frame. In the allo task, an egocentric strategy would require remembering one of 12 movement vectors, whereas the egocentric task involves remembering only 1 of 2. This suggests that animals employ a mnemonic strategy that minimises cognitive load.

1-135. Odor motions and gradients inform Drosophila navigation differently in diffuse and sparse plumes

Samuel Brudner[1,2] SAMUEL.BRUDNER@YALE.EDU
Baohua Zhou[1] BAOHUA.ZHOU@YALE.EDU
Damon A Clark[1] DAMON.CLARK@YALE.EDU
Thierry Emonet[1] THIERRY.EMONET@YALE.EDU

[1]Yale University
[2]Quantitative Biology Institute

Adult Drosophila have been shown to navigate odor plumes using three directional cues: the wind direction, the direction of local odor concentration gradients, and the motion direction of drifting odor packets. The wind points coarsely away from the source within an odor plume cone, but it cannot distinguish the two crosswind directions. Instead, gradient and motion cues may inform flies which crosswind direction contains the plume centerline. However, the usefulness of odor motion and gradient may differ from plume to plume, since odor dispersal patterns vary greatly across environments.

We studied crosswind navigation strategies in two plumes: a diffusive plume with a smooth odor concentration field, and a sparse plume where the odors appear in discrete packets generated through random air motion. We calculated the local gradient and motion cues that stationary upwind-facing flies would experience in these plumes. Local odor gradients were more informative of the correct turning direction in the diffusive plume, while odor motion was more informative in the sparse plume. Next, we used data-driven methods to identify the odor features that upwind-facing flies should calculate to guide their turns. Models trained in the diffusive plume learned to extract bilateral concentration gradients, while models trained in the sparse plume learned to extract odor motion. Finally,

we presented our diffuse and sparse plumes to navigating flies and measured their crosswind turning decisions. Odor gradients predicted Drosophila turning in the diffuse plume, but not in the sparse plume. By contrast, odor motion predicted fly turning decisions in both plume types.

Bilateral odor signals inform olfactory navigation in many animals. Our results indicate how this sensory architecture promotes gradient and motion sensing for navigating diverse natural plumes. Furthermore, our analysis predicts the behavior of flies: when navigating two different plume environments, they use distinct informative cues in each environment.

1-136. Representational drift without synaptic plasticity

Caroline Haimerl CAROLINE.HAIMERL@RESEARCH.FCHAMPALIMAUD.ORG
Christian Machens CHRISTIAN.MACHENS@RESEARCH.FCHAMPALIMAUD.ORG

Champalimaud Research

Neural computations support stable behavior despite relying on many dynamically changing biological processes. One such process is representational drift (RD), which describes changes in neurons' responses over the timescale of minutes to weeks. Across many brain areas, neurons change their tuning profile or even stop/start being active, while population encoding stays intact. Generally, RD is believed to be caused by changes in synaptic weights, which alter individual neurons' tuning properties. These changes impact the population readout and consequently require adaptation of downstream areas to maintain stable function, a costly and non-local problem. Here we propose that much of the observed drift phenomenon can be explained by a simpler mechanism: changes in the excitability of cells without changes in synaptic weights. Fluctuations in excitability due to intrinsic homeostatic properties or neuromodulation can occur at different timescales and change individual neuron's response gain. Here we show that given recurrent connections, such excitability changes can also change the apparent tuning of neurons without requiring any adaptation of population readouts in downstream areas. We use spike coding networks (SCN) to show that the extent of these tuning shifts matches experimentally observed changes and that a specific decoder trained only on one fixed excitability setting performs poorly on others, while a general decoder can perform optimally across excitability changes. This suggests that experimentally observed decline in decoder accuracy across sessions may be due to overfitting of the decoder to one particular population configuration (i.e. the experimental session it was trained on), but that a biological downstream decoder does not need adaptation if it is general. Overall our work proposes a simple mechanism that explains experimentally observed RD and leaves downstream decoding and, by extension, behavior intact.

1-137. A Novel Pain Measurement Tool by Modelling Free-operant Foraging Behaviour in Immersive Virtual Reality

Shuangyi Tong[1,2] SHUANGYI.TONG@ENG.OX.AC.UK
Timothy Denison[1] TIMOTHY.DENISON@ENG.OX.AC.UK
Sang Wan Lee[3] SANGWAN@KAIST.AC.KR
Ben Seymour[1] BEN.SEYMOUR@NDCN.OX.AC.UK

[1]University of Oxford
[2]Institute of Biomedical Engineering
[3]KAIST

Pain measurement through ratings or quantitative sensory testing may not objectively reflect the impact of pain in daily activities. Additionally, pain naturally involves cognitive-motor interactions and is topographic and embodied (Seymour, 2019). Motivated by this observation, we designed a realistic free-operant task and modelled the behaviour by taking pain as one of the inputs fed to decision models. We present a novel framework powered by virtual reality to study foraging behaviour in a realistic environment, where participants are required to collect pineapples in a jungle, some of which are paired with pain stimuli. We tested different pain stimulation sites and types of stimulation. We then proposed mathematical models to address two main questions: which choice (pineapple) participants will take and when they will perform the action. The first choice problem can be solved with classical reinforcement learning algorithms if the state and action spaces are finite (Sutton & Barto, 2020), but they can be more difficult to solve efficiently in our case where the state and action spaces may be continuous and partially observable. Eye-tracking technology that works seamlessly with the virtual environment provides a solution to track the current observable space and possible action spaces. Given a time point, our model assigns a value to each pineapple by linearly adding functions representing different effects. The second vigour-opportunity cost model predicts the delay between two choices by formulating the problem as a trade-off between vigour cost and opportunity cost with additional effect terms derived from the first choice model. We ran model-fitting against the data we collected from human participants, and we found that the fitted model parameters

reflect pain intensities and stimulation configurations. With real-time quantification of pain and movement actions, these models lay the foundation for future pain research and potential treatment in a realistic environment.

1-138. Semi-supervised animal action segmentation with switching nonlinear dynamical systems

Ari Blau[1,2] ARI.BLAU@COLUMBIA.EDU
Neeli Mishra[1] NEELIMISHRA@GMAIL.COM
Evan Schaffer[1] ESS2129@COLUMBIA.EDU
Liam Paninski[1] LMP2107@COLUMBIA.EDU
Matthew R. Whiteway[1] M.WHITEWAY@COLUMBIA.EDU

[1]Columbia University
[2]Statistics

Action segmentation from behavioral video data is a crucial component of many studies that investigate the effects of genetic, neural, and social manipulations on animal behavior. A wide range of supervised and unsupervised methods exist to parse discrete animal behaviors. Supervised methods (i.e., classification) excel at identifying human-defined behaviors of interest, but require a time-consuming annotation step while ignoring the vast amount of data that remains unlabeled. Unsupervised methods (i.e., clustering) excel at discovering novel behaviors via statistical regularities in the data, but may fail to accurately cluster particular behaviors of interest to the experimenter. We address some of these challenges by introducing a semi-supervised approach that learns from both labeled and unlabeled data. We construct a switching nonlinear dynamical system – a general unsupervised model of time series – and propose a novel amortized variational inference strategy that allows for the inclusion of sparse labels for the discrete states. We show in two challenging real world datasets – a head-fixed fly and a freely moving mouse – that our approach outperforms supervised, unsupervised, and existing semi-supervised approaches.

1-139. Large-scale pretraining on neural data allows for transfer across subjects, tasks and species

Mehdi Azabou[1] MAZABOU@GATECH.EDU
Vinam Arora[1] VINAM@GATECH.EDU
Patrick Mineault[2] PATRICK.MINEAULT@MILA.QUEBEC
Venkataramana Ganesh[1] VENKY911@GATECH.EDU
Ximeng Mao[2] XIMENG.MAO@UMONTREAL.CA
Santosh Nachimuthu[1,3] SNACHIMUTHU7@GATECH.EDU
Michael Mendelson[1] MICHAELMENDELSON@GATECH.EDU
Blake Richards[2] BLAKE.RICHARDS@MCGILL.CA
Matthew G Perich[4] MATTHEW.PERICH@UMONTREAL.CA
Guillaume Lajoie[5,6] G.LAJOIE@UMONTREAL.CA
Eva Dyer[1] EVADYER@GATECH.EDU

[1]Georgia Institute of Technology
[2]MILA
[3]Neural Data Science Lab
[4]University of Montreal
[5]Universite de Montreal
[6]Mila

Current population-level models of neural data are built to operate on a single dataset, condition, or task at a time. This both limits the amount of training data we can build models on, and also makes it difficult to build unified models where it may be possible to compare across individuals and over distinct tasks. To address this gap, we introduce a novel approach for large-scale training over many diverse datasets spanning many sessions and tasks. Much like words are passed into transformers in language models, our approach centers around a novel way of breakdowning neural population activity into parts or "tokens" that can be processed by a transformer model. Rather than binning data, we represent each individual spike as a token and associate each with a learned unit-level embedding which captures information about the identity of the unit. We combine this tokenization method with an encoder-decoder architecture that learns interactions across spikes in time and across neurons. To demonstrate the flexibility and scalability of our multi-dataset training framework, we apply it to a wide range of motor, and visual tasks from electrophysiology datasets collected in monkeys, humans, and mice. In particular, we train three large models, the first on 7 nonhuman primates (27,373 units) engaging in four different motor tasks

across three laboratories, the second combining human and monkey motor cortical datasets (9,789 monkey units + 1,920 human units) in diverse tasks, and the third from 58 mice (99,180 units) engaged in visual tasks during recordings in the cortex, deep, and midbrain structures. In these cases, we demonstrate that our pretrained models can be rapidly adapted to new, unseen sessions, enabling few-shot performance across animals and species. Our results highlight the power of scale and demonstrate how pretraining on large datasets can benefit neural data analysis.

1-140. The Helmholtz Hippocampus: A biologically plausible generative model of the Hippocampal formation

Tom George[1] TOMGEORGE1@BTINTERNET.COM
Caswell Barry[2] CASWELL.BARRY@UCL.AC.UK
Kimberly Stachenfeld[3] STACHENFELD@DEEPMIND.COM
Claudia Clopath[4] C.CLOPATH@IMPERIAL.AC.UK
Tomoki Fukai[5] TOMOKI.FUKAI@OIST.JP

[1]Sainsbury Wellcome Centre, UCL
[2]UCL
[3]Google DeepMind
[4]Imperial College London
[5]Okinawa Institute of Science and Technology

Generative models have recently revolutionised machine learning and long been thought fundamental to biological intelligence. In animals, data suggests the hippocampal formation learns and uses a generative model to support its role in spatial and non-spatial memory. Here we introduce a biologically plausible model of the hippocampal formation tantamount to a Helmholtz machine that we apply to a continuous stream of inputs. Fast theta-band oscillations (5-10 Hz) gate the direction of information flow through the network, training it akin to a high-frequency wake-sleep algorithm. Our model can accurately infer the latent state from sensory stimuli and generate realistic sensory predictions offline. Trained on a navigation task it learns to path integrate by developing a ring attractor and can flexibly transfer this structure between environments matching previous theoretical but biologically implausible proposals. Whereas many models trade-off biological plausibility with generality, our model captures a variety of hippocampal cognitive functions under one simple and local learning rule.

1-141. Mapping changes in oscillatory activity across the brain in response to optogenetic stimulation

Jack Goffinet[1,2] JACK.GOFFINET@DUKE.EDU
Kathryn Walder-Christensen[1] KATHRYN.WALDER@DUKE.EDU
Kafui Dzirasa[1] KAFUI.DZIRASA@DUKE.EDU
David Carlson[1] DAVID.CARLSON@DUKE.EDU

[1]Duke University
[2]Computer Science

It has recently been demonstrated that local field potential (LFP) signatures that distinguish between behavioral conditions can be used to heuristically design optogenetic neurostimulation protocols that successfully manipulate behavior. These manipulations impact neurons, which then coordinate in networks to causally impact the behavior of interest. These network-level responses can be measured through LFP signatures to capture neural coordination across many regions. However, the effect of optogenetic stimulation on neural networks as measured in LFPs across the brain is not well understood, hindering the design of effective stimulation protocols. Thus, we advance a novel experimental and statistical framework where LFPs are both used to design neurostimulation protocols and measure their neural response. Here, we use multisite electrophysiological recordings across the mouse brain along with site specific optogenetic stimulation to systematically map how frequency- and region-specific optogenetic stimulation results in perturbations to LFP oscillations. Specifically, we record from six cortical and subcortical regions important for emotion and cognitive processing, and we stimulate each region independently across frequencies up to 30Hz using soma-targeted ChR2 based optogenetic stimulation. Our findings reveal complex functional connectivity changes in response to optogenetic stimulation. We observe both local and long-range effects, indicating the presence of direct and indirect connections between stimulated regions and their downstream targets. Additionally, we observe that optogenetic stimulation commonly affects LFP phases without associated changes in power. Lastly, we use a novel variant of a recently proposed probabilistic model of phase variables, the torus graph, to determine the functional connectivity between frequencies and brain regions of both stimulated and non-stimulated LFPs. Thus, this work provides a step towards understanding network-level

responses of neurostimulation.

1-142. Neuromorphic dreaming: A pathway to efficient learning in artificial agents

Ingo Blakowski[1] INGOBLAKOWSKI@GMAIL.COM
Dmitrii Zendrikov[2] DMITRII@INI.UZH.CH
Cristiano Capone[3,4] CRISTIANO0CAPONE@GMAIL.COM
Giacomo Indiveri[5] GIACOMO@INI.UZH.CH

[1]Institute of Neuroinformatics, UZH and ETH Zurich
[2]Institute of Neuroinformatics
[3]Istituto Superiore di Sanita
[4]PRORA
[5]Institute of Neuroinformatics, University of Zurich, ETH Zurich

In contrast to humans and animals, current reinforcement learning algorithms require extensive data to master new skills. Recent model-based approaches are reducing the number of necessary interactions with the environment. However, they often use biologically implausible elements, such as detailed storage of past experiences and offline learning. Inspired by biological processes, we introduce a model-based reinforcement learning approach that uses spiking neural networks (SNNs) and leverages on a "dreaming" phase as a strategy to efficiently use internal models for reducing interactions with the environment and for addressing the challenge of optimal learning and exploitation of world models. In addition, the SNN nature of the model makes it naturally well-suited for implementation on neuromorphic hardware. Our architecture is tested with the Atari game Pong and features two on-chip feedforward SNNs with readout layers on the computer: (i) an agent network that determines action probabilities, and (ii) a model network that predicts the next state and reward. A key innovation in our approach is the alternating "awake" and "dreaming" learning phases. The "awake" phase involves direct environmental interaction, enhancing both networks, while the "dreaming" phase refines only the agent network through simulated interactions with the model network, avoiding potentially costly real-world interactions. Training exclusively focuses on the readout weights. It starts with a 100-frame "awake" episode in the real environment where the model network is adjusted frame-by-frame in a supervised way and the reward-based policy gradient is accumulated. The agent network is then updated using the accumulated policy gradient. This is followed by a 50-frame "dreaming" episode, updating the agent network based on the accumulated policy gradient obtained from imaginary interactions. We emphasize that our system avoids storing detailed experiences, learns policy and model online, uses low-power optimized neuromorphic hardware, and successfully learns Pong with reduced environmental interactions through dreaming.

1-143. Reward and perceptual difficulty drive distinct changes in behavior and motor cortical activity

Adithya Narayan Chandrasekaran[1,2] ADC136@PITT.EDU
Megan McDonnell[3] MMCDONNELL@CMU.EDU
Chris Ki[3] CSKI@ANDREW.CMU.EDU
Adam Smoulder[3] ASMOULDE@ANDREW.CMU.EDU
Byron Yu[3] BYRONYU@CMU.EDU
Aaron Batista[1] AARON.BATISTA@PITT.EDU
Matthew A. Smith[3] MATTSMITH@CMU.EDU
Steven Chase[3] SCHASE@ANDREW.CMU.EDU

[1]University of Pittsburgh
[2]Neuroscience
[3]Carnegie Mellon University

Our behavior depends on both external factors, such as sensory stimuli, and internal factors, like motivation. We may expend more effort when the potential payoff is greater. We may also expend more effort when the task is more demanding. Do reward and difficulty engage a common motivational drive? To investigate, we recorded from motor cortex (M1) while monkeys performed a delayed reach task where the go-cue was a difficult perceptual change. We varied the offered reward on each trial. We also manipulated perceptual difficulty using blocks of hard- or easy-to-detect changes. To improve performance, monkeys could either increase their perceptual sensitivity (ability to detect changes), shift their impulsivity (tendency to respond), and/or improve their motor accuracy. We found that perceptual sensitivity was increased for both larger rewards and greater difficulty. In contrast, impulsivity and motor accuracy were differently impacted by reward and difficulty: First, motor accuracy improved

for larger rewards, but was not affected by the perceptual difficulty context. Analyzing M1 activity revealed that reward impacted the fidelity of the motor plan. Larger rewards led to increased target direction separability in M1 during preparation, reflecting improved motor planning. In contrast, perceptual difficulty context did not impact target direction separability. Second, monkeys were more impulsive when the perceptual difficulty was greater, but less impulsive for larger rewards. We found that perceptual difficulty context altered the initial condition of M1 activity. During the hard context, M1 activity was closer to a movement initiation state before the go-cue, reflecting increased impulsivity. To summarize, our results suggest that reward and difficulty operate through distinct processes. Reward increased motor accuracy by improving the fidelity of the reach plan, while perceptual difficulty context increased impulsivity by altering initial conditions. Our work provides insight into how the brain flexibly guides behavior in response to a changing environment.

1-144. Optimization of retinal electrical stimulation for vision restoration using a bidirectional neural interface

Andrew Phillips[1,2] ANDREWJP@STANFORD.EDU
Nishal Shah[1,3] NISHALPS@STANFORD.EDU
Praful Vasireddy[1,4] PVASIRED@STANFORD.EDU
Amrith Lotlikar[1] LOTLIKAR@STANFORD.EDU
Alex Gogliettino[1] GOGLIETTINO@STANFORD.EDU
Jeff Brown[1] JDOTCAFE@STANFORD.EDU
Pawel Hottowy[5] HOTTOWY@AGH.EDU.PL
Alexander Sher[6] SASHAKE3@UCSC.EDU
Alan Litke[6] ALAN.LITKE@CERN.CH
Eduardo Chichilnisky[1] EJ@STANFORD.EDU

[1]Stanford University
[2]Department of Electrical Engineering
[3]Neursurgery
[4]Electrical Engineering
[5]AGH University of Science and Technology
[6]University of California, Santa Cruz

Neural interfaces such as retinal implants have the potential to restore and even augment our senses. However, all present and envisioned devices are limited by an imperfect ability to evoke neural signals that match the neural code. Thus, a central challenge is to computationally define and execute spatiotemporal patterns of neural stimulation that optimize the sensory experience within the measured constraints of the neural interface. Here, we develop a closed-loop approach to this problem and test its effectiveness empirically using large-scale multi-electrode stimulation and recording in the retina ex vivo — an experimental laboratory prototype for a future implanted device. The computational framework consists of two main components: visual decoding and electrical encoding. In the visual decoding stage, we use measured light responses of retinal neurons to infer a mapping between retinal activity and visual perception, allowing for direct optimization of the inferred sensory experience. Then, in the electrical encoding stage, we define and solve an optimization problem to determine the spatiotemporal pattern of electrical stimulation which would be expected to produce the most accurate evoked visual experience. In our closed-loop experiment, the electrical encoding uses a rapidly interleaved sequence of calibrated, single-electrode stimuli, selected greedily to minimize the error between the target visual stimulus and the putative visual perception inferred as a linear combination of the visual decoders. The results demonstrate experimentally, for the first time, that a closed-loop computational approach to electrical stimulation for vision restoration can deliver optimized visual sensations from the retina to the brain.

1-145. Locomotion shapes visually-evoked neural dynamics through modulation of collicular populations

Bram Nuttin[1,2] BRAM.NUTTIN@NERF.BE
Norma Kuhn[1] NORMA.KUHN@NERF.BE
Anna Chrzanowska[1] ANNA.CHRZANOWSKA@NERF.BE
Arnau Sans Dublanc[1] ARNAU.SANSDUBLANC@NERF.BE
Karl Farrow[1] KARL.FARROW@NERF.BE

[1]Neuro-Electronics Research Flanders (Imec, KU Leuven, VIB)
[2]Biology

In the face of danger, an animal's survival hinges on its ability to react appropriately. These instinctive reac-

tions are influenced by the animal's behavioral state. In the brain, signals correlated with behavioral state are widespread across early sensory regions, including the superior colliculus ('colliculus'), a sensorimotor area that drives defensive behaviors. However, the mechanism that enables behavioral state to modify the processing of visual information to affect action selection remains unknown. We propose a simple mechanism for action selection based on the selective modulation of visual inputs from the retina. We recorded the visual responses in the colliculus using Neuropixels, while mice were free to run. We have three key findings. First, fitting a generalized linear model to behavior reveals that current locomotion predicts the response of mice to visual threat. Second, clustering neurons based on their visual and locomotion responses demonstrates that the strength of modulation is cell type dependent. Finally, projecting the neural population activity into a low-dimensional space reveals different effects of modulation in the superficial versus deep layers of the colliculus. In superficial layers, locomotion decreases the magnitude of neural trajectories suggesting a general gain modulation of most neurons. In deeper layers, locomotion increases the angle between responses suggesting selective modulation of different neurons. We are now linking these effects of modulation to visually-evoked behaviors in freely moving mice using chronic Neuropixels recordings. Our results demonstrate that behavioral state differentially modulates the gain of distinct neural populations in the colliculus. This population-specific modulation allows for the independent control of the different output pathways of the colliculus. This work contributes to our understanding of how behavioral and contextual factors shape an animal's defensive responses to threats. More broadly, it establishes a framework for understanding the relationship between neural modulation across the brain and its impact on behavior.

1-146. Respiration coordinates the olfactory cortical code

Robin Blazing[1,2] ROBIN.BLAZING@DUKE.EDU
Kevin Franks[1] FRANKS@NEURO.DUKE.EDU

[1]Duke University School of Medicine
[2]Neurobiology

A primary goal of systems neuroscience is to determine how information is communicated across networks of neurons. One common framework for intra-areal communication is phase coding, in which the phase of responses relative to ongoing neural oscillations gates the transmission of stimulus information. While phase-locked spiking activity has been identified in many brain regions, including the hippocampus and olfactory system, there is little empirical evidence that the precise phase of neural responses is read out by downstream networks. Here, we investigate how the phase of inputs from the olfactory bulb (OB) to the piriform cortex (PCx) impacts cortical responses. In this system, odors activate select subsets of OB glomeruli, which respond at different phases of the respiration cycle. The glomeruli project via mitral and tufted cells (M/TCs) to PCx, where individual neurons receive input from multiple glomeruli. While glomerular response phase contains odor information, how this phase code is read out in PCx is not known. To address this question, we optogenetically activated small subsets of OB glomeruli at different phases of the respiratory cycle while recording spiking responses in PCx of awake, head-fixed mice. We revealed that PCx neurons are tuned to glomerular stimulation at specific respiration phases. Remarkably, the preferred phase of each PCx neuron was highly consistent across different glomeruli. Blocking airflow through the nostril considerably diminished phase-tuning, indicating that respiratory-induced oscillatory activity is required for phase-specific PCx responses. Finally, we show that OB M/TCs prefer stimuli arriving during early inhalation, while phase preferences in PCx tile the sniff cycle. Experimental manipulations indicate that recurrent circuitry in PCx is required for this shift in tuning, suggesting that cortical circuits filter incoming inputs in a phase-specific manner. Together, our results indicate that cortex implements a phase-to-rate transformation to facilitate efficient representation of inputs arriving throughout the sniff cycle.

1-147. Unveiling the origin of the word-specific area with the object space model Summary

Jia Yang[1,2] JIAYANG@PKU.EDU.CN
Pinglei Bao[1] PLBAO.USC@GMAIL.COM
Yipeng Li[1] 1800013731@PKU.EDU.CN
Haoxuan Yao[1] HAOXUANYAO@PKU.EDU.CN
Jingqiu Luo[1] 1900013728@PKU.EDU.CN
Hongyu Li[1] HONGYULI@PKU.EDU.CN
Xiaoya Chen[3,4] XIAOYACHEN2333@GMAIL.COM
Shiming Tang[1] TANGSHM@PKU.EDU.CN

[1]Peking University
[2]School of Psychological and Cognitive Sciences
[3]Vanderbilt University

[4]Psychology and Human Development

Words constitute a unique, experience-dependent category within the representational space of the human ventral pathway. The prevailing view maintains that learning to read repurposes a pre-existing region in the ventral occipitotemporal cortex for the recognition of written words. However, the initial function of this prototypical region (visual word form area, VWFA) remains elusive. In this study, by leveraging deep learning neural networks, we initially show that considerable word discrimination capacity can be derived from general non-word object recognition training. We find that objects similar to words in the network's 'object space' share more features that help in word recognition. This is mirrored in the primate brain: our fMRI studies show that the VWFA responds more to objects that are closer to the words in the object space. Such an effect is even true in the macaque inferotemporal (IT) cortex with the measurements of the responses to the words and objects using wide-field imaging. Through integrating findings from CNN, human fMRI, and wide-field imaging in macaques, our work highlights the possibility that VWFA may have initially evolved to represent features of non-word objects that are closely related to words in the object space and shed light on the development of category-specific areas.

1-148. 3D cross-modality mapping of tissue scale brain atlas to cellular scale spatial transcriptomics

Kaitlin Stouffer[1,2] KSTOUFF4@JHMI.EDU
Xiaoyin Chen[3] XIAOYIN.CHEN@ALLENINSTITUTE.ORG
Mara Rue[3] MARA.RUE@ALLENINSTITUTE.ORG
Alain Trouve[4] ALAIN.TROUVE@ENS-PARIS-SACLAY.FR
Benjamin Charlier[5] BENJAMIN.CHARLIER@UMONTPELLIER.FR
Michael Miller[6] MIM@JHU.EDU

[1]Johns Hopkins University
[2]Biomedical Engineering
[3]Allen Institute for Brain Science
[4]Centre Borelli, ENS Paris-Saclay, University Paris-Saclay
[5]Universite de Montpellier
[6]Dept of Biomedical Engineering and Kavli Neuroscience Discovery Institute, Johns Hopkins University

Developing spatial transcriptomics technologies can reveal patterns in cell types and transcriptional signatures across brain regions and disease models that has prompted interest in understanding their association to region-specific functions and behavior-evoked activity. Associating molecular information with functional measures on a brain-wide scale, however, requires integration of datasets across specimens and technologies and with existing anatomical atlases. Computational tools for achieving integration remain sparse with challenges posed by the datasets' sizes and necessary multi-modal alignment of measurements taken at different scales, across variable 2D/3D scopes of tissue, and of incomparable gene, protein, and cell-specific feature readouts. We address these challenges through a cross-modality mapping method of tissue scale atlases to molecular/cellular scale data that models both datatypes equivalently as "particle" measures termed image varifolds. Unlike classical image-based methods, which model data as regularized grids, image varifolds accommodate the potentially different scales, sparsity, and spatial irregularity of data in a computationally efficient manner without explicit modeling of background space. To align datasets of different feature readouts, we rely on the segmentation schemes defined in the tissue scale atlas as a scaffold, with each brain region prescribed a single distribution over target feature values that is assumed to hold throughout the entire region. Distributions per atlas region are estimated simultaneously with smooth, invertible geometric deformations (diffeomorphisms) using large deformation diffeomorphic metric mapping (LDDMM) extended from the imaging setting. We demonstrate our method in aligning the CCFv3 to mouse brain hemispheres of spatially resolved, cell-typed BARseq data. Results show global alignment of coronal hemi-sections within the CCFv3 3D coordinate framework, with local alignment at 5-10 micron exhibited by overlap in corresponding cortical layers to layer-specific cell types. Future work includes cross-specimen analysis facilitated by integration into the CCFv3 coordinates and integration of more diverse datasets from spatial transcriptomics to brain-wide functional imaging measures.

1-149. How context representations emerge during training: a linear network perspective

Alexandra Proca[1,2] A.PROCA22@IMPERIAL.AC.UK
Kai Sandbrink[3,4] KAI.SANDBRINK@LMH.OX.AC.UK
Jan Bauer[5] JAN.BAUER@RWTH-AACHEN.DE
Ali Hummos[6,7] AHUMMOS@MIT.EDU

[1]Imperial College London
[2]Department of Computing
[3]University of Oxford
[4]Department of Experimental Psychology
[5]Hebrew University of Jerusalem
[6]MIT
[7]Brain and Cognitive Science

Contextual responding is a common feature in neuroscience experiments where animals identify two contexts in the task and switch between them flexibly, often requiring only a few error trials to infer a context switch. Context representations are then thought to gate the neural population dynamics to perform the correct computation in each context. However, how these cognitive abstractions of context emerge in neural systems is unknown. Here, we consider the emergence of context in a linear network with gating variables, both analytically and in simulations. During learning we update both the network weights and the gating variables along the objective function gradient. Simulations show that the weight matrices specialize to solve the computation required in each task, while the gating variables represent the active context. Transitioning between contexts initially relies on updating weights and later on updating gating variables only, leading to adaptable behavior and minimizing interference and forgetting. This separation between task computations and representation of task context emerges from an interplay of scales in the respective parts of the network. Gating variables encoding of context is incentivized by the dynamics once the weight matrices learn the task computation accurately. In addition, analytical expressions of the gating variable dynamics and its interactions with weights dynamics show that gating not only influences behavioral output, but also gates the effective learning rate of the weights responsible for other behaviors. Overall, our work studies linear networks to propose a mechanism for how abstract cognitive representations of context emerge, identifying the pertinent components for behavioral flexibility and protecting knowledge.

1-150. A Spectral Theory of Neural Alignment and Prediction

Abdulkadir Canatar[1,2] CANATARA@GMAIL.COM
Jenelle Feather[1,2] JFEATHER@FLATIRONINSTITUTE.ORG
Albert Wakhloo[3,4] AJW2232@CUMC.COLUMBIA.EDU
SueYeon Chung[5,2] SCHUNG@FLATIRONINSTITUTE.ORG

[1]Flatiron Institute
[2]Center for Computational Neuroscience
[3]Columbia University; Flatiron Institute
[4]Center for Theoretical Neuroscience; Center for Computational Neuroscience
[5]New York University; Flatiron Institute

The representations of neural networks are often compared to those of biological systems by performing regression between the neural network responses and those measured from the brain. Many state-of-the-art deep neural networks perform better than past hand-designed models at predicting neural data. However, many different architectures and training procedures often lead to similar predictions of neural responses, even when model representations are notably different from each other using other metrics of comparison. It is an open question of why current neural prediction benchmarks are less sensitive to model modification, and how to design future experiments and stimulus sets to better test our models. To investigate the model properties leading to good neural predictions, we use a recent theoretical framework that relates the generalization error from regression to the spectral properties of the model and the neural data. We apply this theory to the case of regression between model activations and neural responses and decompose the neural prediction error in terms of the model eigenspectra, alignment of model eigenvectors and neural responses, and the training set size. Using this decomposition, we introduce geometrical measures to interpret the neural prediction error. We test a large number of deep neural networks that predict visual cortical activity and show that there are multiple types of geometries that result in low neural prediction error as measured via regression. The work demonstrates that carefully decomposing representational metrics can provide interpretability of how models are capturing neural activity and points the way towards improved models of neural activity.

1-151. Learning to infer transitively: ranking symbols on a mental line in premotor cortex

Sofia Raglio[1] SOFIA.RAGLIO@UNIROMA1.IT
Gabriele Di Antonio[2] GABRIELE.DIANTONIO@CREF.IT
Emiliano Brunamonti[1] EMILIANO.BRUNAMONTI@UNIROMA1.IT
Stefano Ferraina[1] STEFANO.FERRAINA@UNIROMA1.IT
Maurizio Mattia[3,4] MAURIZIO.MATTIA@ISS.IT

[1]Sapienza University of Rome
[2]Istituto Superiore di Sanita, Roma Tre University, Research Center "Enrico Fermi"
[3]Istituto Superiore di Sanita
[4]Natl. Center for Radioprotection and Computational Physics

Transitive inference (TI) is a form of deductive reasoning that allows to infer unknown relations among premises. Cognitive resolution of this task is thought to involve a mental linear workspace, usually referred to as the "mental line," organizing stimuli sequentially based on their ranks. Here, we introduce the Geometric Mental Line (GML) model for TI task solution, defining the mental line direction as the linear combination of the symbols representations weighted by their ranks. This model can be adapted to specific task designs and implemented in a linear dynamical system approximating more biologically-realistic recurrent neural networks. Tuning few key parameters, the GML model can successfully replicate the behavioral effects observed in monkeys and its unfolded dynamics offers insights into the neural organization of stimuli representations during the task. Therefore, a crucial question arises: does the GML model faithfully capture the geometry of neural activity? Specifically, does a neuronal "mental line" exist, and if so, where is it encoded and learned in the brain? This work explores the role of dorsal premotor cortex (PMd) in representing the GML, challenging the assumption that this region solely represents motor planning, and showing its involvement in encoding task-relevant information, such as symbols representations. Our results provide evidence that PMd plays a key role in manipulating these representations, efficiently transforming the ordinal knowledge into a proper motor decision. The GML implemented in PMd is predictive of animal behavior, forecasting the decision both in correct and error trials and explaining the reaction times distribution. Moreover, we found striking evidence that the representations of the stimuli are plastic: the learning process leads to a realignment of the GML to the motor decision axis, elucidating an optimization strategy pursued by the PMd which eventually shrinks the solution dimensionality, as predicted by our GML model.

1-152. A Dual-Input Firing Rate Model for CA1 Place Cell Phase Precession and Theta Sequences

Yiqing Lu[1,2] YL4698@NYU.EDU
Antonio Fernandez-Ruiz[3] AFR77@CORNELL.EDU
John Rinzel[1] RINZELJM@GMAIL.COM

[1]New York University
[2]Courant Institute of Mathematical Sciences
[3]Cornell University

Hippocampal place cells, crucial for spatial memory and navigation, fire when an animal enters a specific location in the environment. During locomotion, the hippocampus exhibits theta rhythm, representing an online state. A notable characteristic of place cell activity is phase precession, where the firing phase relative to the theta rhythm progressively shifts earlier. Despite various theories, the underlying mechanism of phase precession remains elusive. This study investigates the dual-input model (Fernandez-Ruiz et al., 2017), which proposes that the interplay of the strength and phases of location/time-dependent inputs from CA3 and Entorhinal Cortex layer III (EC3) accounts for phase precession of CA1 place cells.

We implement this idea with a compartmental firing rate model, representing activity of an ensemble of CA1 place cells with identical place fields. The model incorporates mean-membrane potentials in three compartments – soma, proximal dendrites, distal dendrites, along with the soma firing rate and the mean synaptic output across cells. With CA3 input innervating the proximal dendrite and EC3 input innervating the distal dendrite, the soma firing rate in our simulation shows a phase precession in a range consistent with experimental observations. Moreover, our model replicates several experimental findings where parts of the circuit are optogentically silenced. Notably, it demonstrates that rate coding predominantly relies on CA3 input, while phase coding depends mostly on EC3 input (Zutshi et al., 2022).

Furthermore, we extend the dual-input model to explain theta sequence generation, where the place coding of distinct ensembles of place cells is compressed within a single theta-cycle to represent spatial trajectories. In particular, we capture the bimodal cells (Wang et al., 2020) that display both phase precession (advance) and recession (delay) by increased EC3 input and study their role in generating both forward and reversed theta

sequences. Overall, our computational model provides supporting evidence for the dual-input hypothesis.

1-153. Hippocampal neuronal populations over weeks and months in a simple continual learning task

Gabriela Michel[1] MICHELG@HHMI.ORG
Michalis Michaelos[1] MICHAELOSM@JANELIA.HHMI.ORG
Johan Winnubst[1] JOHAN@E11.BIO
Weinan Sun[1] SUNW2@JANELIA.HHMI.ORG
Kevin Miller[2] KEVINJMILLER@DEEPMIND.COM
Boaz Mohar[1] MOHARB@JANELIA.HHMI.ORG
Matthew Botvinick[3] BOTVINICK@GOOGLE.COM
Kimberly Stachenfeld[4] STACHENFELD@DEEPMIND.COM
Nelson Spruston[1] SPRUSTONN@JANELIA.HHMI.ORG

[1]HHMI Janelia Research Campus
[2]Google DeepMind, University College London
[3]Deep Mind
[4]Google DeepMind

"Continual Learning" – the ability to combine memory and generalization to acquire, update, and integrate knowledge over long timescales – is a natural part of how animals interact with the world and a longstanding challenge for AI [1]. In order to investigate the neural mechanisms of continual learning, we developed a blocked, multitask experiment in a head-fixed virtual reality setup in which water restricted mice discriminate between pairs of visual cues. Mice learned this task in 2-3 days, and transferred this knowledge to quickly acquire new cue pairs. Furthermore, mice exhibited long memory retention, maintaining high accuracy on previously experienced cue pairs despite intervening distractor tasks or long delays (~50 days). Interestingly, some cue pairs were challenging for the mice and disrupted performance on previously learned cue pairs. To study the role of the hippocampus in this process [2], we used 2-photon calcium imaging to record the activity of thousands of hippocampal CA1 neurons across the entire months-long trajectory of learning. Preliminary analyses reveal aspects of the hippocampal representations that persist over long timescales and others that change with time. Our results provide an early insight into how the hippocampus contributes to the acquisition and consolidation of new information over long timescales.

1-154. Random walk length modulates structure acquisition in modular graphs

Tejas Savalia[1,2] TSAVALIA@UMASS.EDU
Jeffrey Starns[1] JSTARNS@UMASS.EDU
Andrew Cohen[1] ALC@UMASS.EDU

[1]University of Massachusetts Amherst
[2]Cognition and Cognitive Neuroscience, Psychological and Brain Sciences

We consistently segment our temporally continuous sensory experience into meaningful chunks of information. Perceived temporal segmentation has been shown to occur both when the information stream consists of explicit, perceptual breaks and through more implicit breaks determined by the underlying transition probabilities between the stimuli[1]. In modular graph environments, response time differences between cross module transitions and within module transitions have been used to indicate the acquisition of the statistics of the underlying graph structure[1]. It has further been shown that both the graph topology[2] and the parsing mechanism[3] modulate this response time difference and hence the statistical learning of graph structure. Particularly, structure acquisition is the strongest when participants experience a random walk through a modular graph. In our work, we first present an information theoretic account using the Successor Representation (SR)[4, 5] framework to show that walk type and graph topology lead to varying levels of uncertainty difference in the within vs cross-module transitions leading to the observed response time difference. The model further predicts the length of a random walk as a modulatory factor in structure acquisition. We next test the model prediction by presenting participants the same modular structure as prior studies, using the same motor stimuli used in Kahn et al. (2018) by modulating the lengths of random walk they are exposed to. Using a Bayesian hierarchical linear model fit to account for effects of stimulus lag (the number of trials before a stimulus is repeated), reset (a trial when the shorter random walk resets) and motor learning, participants (1) improve performance with a longer random walk and (2) respond slower to cross-module transitions with a longer random walk. Our results thus provide support for predictive representations of upcoming stimuli as a key mechanism behind statistical learning.

1-155. The low-dimensional evolution of neural connectivity over learning

Arthur Pellegrino[1,2] PELLEGRINO.ARTHUR@ED.AC.UK
N. Alex Cayco Gajic[3] NATASHA.CAYCO.GAJIC@ENS.FR
Angus Chadwick[1,4] ANGUS.CHADWICK@ED.AC.UK

[1]University of Edinburgh
[2]Institute for Adaptive and Neural Computation
[3]Ecole Normale Superieure Paris
[4]School of Informatics

Learning relies on coordinated synaptic changes in recurrently connected populations of neurons. Therefore, understanding the collective evolution of synaptic connectivity is a key challenge for understanding how neural circuits learn to perform complex tasks. To address this, we investigate the structure of the 3-tensor formed by stacking the connectivity matrices over trials. By fitting RNNs of varying rank to large-scale neural recordings during a motor learning task, we find that the inferred connectivity is low-tensor-rank. A consequence of this constraint is that neural connectivity evolves over a fixed low-dimensional subspace throughout the entire course of learning. We next validate the observation of low-tensor-rank learning in RNNs which are trained to solve common neuroscience tasks. Finally, we present a set of mathematical results bounding the matrix and tensor ranks of gradient descent learning dynamics which show that low-tensor-rank connectivity emerges naturally in RNNs trained to solve behavioural tasks. Taken together, our findings suggest that low-tensor-rank connectivity is a general property of learning in recurrently connected neural populations. Moreover, based on this result we introduce an interpretable method to reverse-engineer learning-induced changes in recurrent dynamics from large-scale neural recordings.

1-156. Decomposing thermodynamic dissipation of neural dynamics via spatio-temporal oscillatory modes

Daiki Sekizawa[1,2] SEKIZAWA-DAIKI963@G.ECC.U-TOKYO.AC.JP
Sosuke Ito[1] SOSUKE.ITO@UBI.S.U-TOKYO.AC.JP
Masafumi Oizumi[1] C-OIZUMI@G.ECC.U-TOKYO.AC.JP

[1]The University of Tokyo
[2]Department of General Systems Studies

Recent developments in stochastic thermodynamics have elucidated various relations between the entropy production rate (thermodynamics dissipation) and the physical limits of information processing in non-equilibrium dynamical systems, which have been actively used and opened new perspectives in the analysis of real biological systems. Even in neuroscience, the importance of quantifying the entropy production has attracted increasing attention to understand the properties of information processing in the brain [1, 2]. However, the relationships between entropy production rate and neural oscillations, such as delta, theta, and alpha waves, which are prevalent in the brain, are unclear. Here, we derive a novel decomposition of the entropy production rate. We show that one of the components of the entropy production rate, called the housekeeping entropy production rate, can be decomposed into independent positive contributions from spatio-temporal oscillatory modes. Our decomposition enables us to calculate the contribution to the housekeeping entropy production rate from oscillatory modes, as well as the spatial distribution of the contributions. To demonstrate the utility of our decomposition, we applied our decomposition to the electrocorticography (ECoG) dataset recorded during awake and anesthetized conditions, where the properties of oscillations change drastically. We showed that the contributions of oscillatory modes from the delta band were larger in the anesthetized state than in the awake state, while those from the theta and alpha bands were smaller. These results allow us to interpret the change in the neural oscillation in terms of stochastic thermodynamics and the physical limit of information processing.

1-157. Neural Heterogeneity Controls the Computational Properties of Spiking Neural Networks

Richard Gast[1,2] RICHARD.GAST@NORTHWESTERN.EDU
Ann Kennedy[1,3] ANN.KENNEDY@NORTHWESTERN.EDU
Sara A. Solla[1] SOLLA@NORTHWESTERN.EDU

[1]Northwestern University
[2]Department of Neuroscience
[3]Neuroscience

Neurons and the synapses connecting them are inherently diverse in their structure and electrophysiological response properties. To understand how brain function arises from networks of neurons, we must understand how the diversity of neural tissue relates to that function. Is this diversity, which we call neural heterogeneity, an unintended consequence of messy biology, or does it have some computational benefit? Recent results suggest that neural heterogeneity can situate the brain in a computationally beneficial critical regime and that the loss of neural heterogeneity might underlie pathological brain dynamics as seen in epilepsy.

In this work, we provide novel insight into the computational role of neural heterogeneity by directly relating heterogeneity-induced changes in neural dynamics to changes in neural population function. To this end, we leverage a set of mean-field equations which allows to determine a direct relationship between the heterogeneity of spike thresholds across neurons and the resulting low-dimensional network dynamics.

Building on this model, we show that the heterogeneity of inhibitory interneurons plays a crucial role in shaping the dynamic regimes of neural circuits: heterogeneous inhibitory interneuron populations preserve the dynamic repertoire of local excitatory populations, whereas homogeneous interneurons overwrite excitatory dynamic repertoires and facilitates synchronized dynamics. Furthermore, we find that neural heterogeneity directly controls the encoding capacity of populations of recurrently coupled excitatory neurons by affecting the multistable dynamic regime of the population. Together, our results suggest that neural heterogeneity should be considered as a crucial "knob" that can be adjusted in neural circuits to tune their computational function.

1-158. Working memory and recall with expander networks

Anandita De[1,2] ANANDITA@UOREGON.EDU
Rishidev Chaudhuri[3] RCHAUDHURI@UCDAVIS.EDU

[1]University of Oregon
[2]Institue of Neuroscience
[3]University of California, Davis

Working memory is the ability to hold things ‘in mind’. Persistent, content specific activity underlying this ability is observed in multiple regions in the brain such as prefrontal, parietal and sensory cortices. Recall is the ability to retrieve information stored in memory. It relies on interaction between sensory cortices and prefrontal cortices. Both processes involve maintaining or reconstructing signals corresponding to memories from incomplete information and activate multiple areas. These areas are connected by long-range reciprocal connections. For example, there are convergent feedforward connections starting from sensory cortices into association cortices all the way up to prefrontal cortices and reciprocal divergent connections from prefrontal cortices down to sensory cortices. What is the role of the long range reciprocal connections in supporting working memory and recall? We investigate this question by modeling two regions interacting with each other via sparse random projections. These networks have a special property known as the expander property. We show that working memory and recall correspond to different initial conditions for the dynamics on this network. They converge to the same final state which corresponds to a memory being maintained or reconstructed. Combining results from expander graphs and numerical simulations we demonstrate that any sparse binary pattern with a given number of active neurons which depend on the network parameters can be maintained in this network. We extend our model to three regions and show that recall is possible in the three-region network. Although local recurrent connectivity is not a necessity to maintain signals in this network as in previous work, we speculate that it will lead to richer dynamics in each coritcal area during working memory as observed experimentally.

1-159. Inferring neural communication dynamics from field potentials using graph diffusion autoregression

Felix Schwock[1,2] FSCHWOCK@UW.EDU
Julien Bloch[1] JULIENB@UW.EDU
Karam Khateeb[1] KKHATEEB@UW.EDU
Jasmine Zhou[1] JZHOU33@UW.EDU
Les Atlas[1] ATLAS@UW.EDU
Azadeh Yazdan-Shahmorad[1] AZADEHY@UW.EDU

[1]University of Washington
[2]Electrical and Computer Engineering

The goal to better understand cognitive processes, combined with rapid advancements in multi-electrode neural recording technologies, has spurred increased interest in estimating dynamic network communication. Yet, traditional methods, which infer communication from statistical dependencies between neural activity from different recording sites, such as coherence or measures based on vector autoregressive (VAR) models, face core limita-

tions: They lack an explicit communication model, neglect spatial information from the recording setup, and yield predominantly static estimates that cannot capture rapid changes in brain states. Alternatively, communication between brain regions can be described using more mechanistic models that give rise to highly dynamic communication signals on top of structural networks. For example, recent results have shown that communication between different brain regions can be well modeled via diffusive processes. However, it remains unclear how to effectively integrate these models with neural recordings. To address this, we introduce graph diffusion autoregressive (GDAR) models. Designed for distributed field potential recordings, the GDAR model combines vector autoregression with diffusion, thereby adding biophysical realism to traditional VAR model and naturally producing a communication signal with millisecond temporal resolution. Using simulated data from a network of couples Wilson-Cowan oscillator we demonstrate that the GDAR model can estimate fast communication dynamics more accurately than an enhanced VAR model. Next, using electrocorticography (ECoG) recordings from non-human primates, we demonstrate how the model can be used to study transient communication dynamics evoked by cortical stimulation, changes in resting state communication during a stroke experiments, and the neural trial-to-trial variability underlying an animal's reach behavior. For all datasets, we show that the GDAR model outperforms standard VAR models and other FC measures and provides insights that cannot be obtained by other models.

1-160. Alignment between stimuli and prediction can explain deviant detection in a parsimonious model

John Meng[1,2] HM97@NYU.EDU
Xiao-Jing Wang[1] XJWANG@NYU.EDU

[1]New York University
[2]Center of Neural Science

A brain is not only a passive sensory information receiver but also generates predictions about the external world. We use predictions to focus more on surprise signals than known events. One protocol to untangle these mechanisms is the oddball paradigm, where an animal receives an unexpected or "deviant" stimulus interspersed among a series of repetitive "standard" stimuli. Neuronal recordings from different experimental protocols (1, 2, 3) robustly detect a subpopulation (referred to as deviant detector, DD) response stronger if a stimulus is presented as an “oddball” than presented as a “normal ball”. In addition, for these DDs, the oddball response is stronger than a many-standard control (1), suggesting the role of top-down prediction and opposing the sole contribution of bottom-up adaptation. However, how these DDs are generated in the neocortex is poorly understood. Here, we explore potential mechanisms by using parsimonious biologically feasible models that include a decision-making PFC-like circuit generating internal expectations. To serve the function in the predictive coding, these DDs must subtract the prediction from the stimuli: R(rate)= X(stimulus) – Y(prediction). The observation that DDs respond least to the last repetitive “normal ball” can be due to increased Y over repetition. Similarly, the higher response to the “oddball” can be due to decreased Y, achieved through competition between predicting an oddball next or a normal ball. Recently, (4) suggests that V1 L23 pyramidal cells with different genetic markers respond to predictions and stimuli differently. (5) suggests that excitation and inhibition can align on single pyramidal cells through plasticity. Based on these, a rate model that aligns the stimuli and the prediction, directly on pyramidal cells or bi-synaptically through Somatostatin (SST) interneuron (IN) successfully reproduced the DDs observed in experiments. Further, the model includes Parvalbumin interneurons and nonlinear dendritic compartments and is tested with realistic biological constraints.

1-161. Inference of neural activity in connectome-constrained recurrent neural networks

Manuel Beiran[1] MB4878@COLUMBIA.EDU
Ashok Litwin-Kumar[1,2] A.LITWIN-KUMAR@COLUMBIA.EDU

[1]Columbia University
[2]Neuroscience

Recent experimental advances have enabled the measurement of the ‘connectome’ or synaptic wiring between neurons across large brain areas in different animal models. Theoretical work has long focused on linking connectivity features of large neural networks to their emerging dynamics. Nevertheless, the extent to which connectomics data can provide insights into the functionality of neuronal networks remains an open question, given that the physiological properties of individual neurons are unknown. To address this, we developed a theory that describes the space of solutions when inferring activity in connectivity-constrained neural networks with unknown single-neuron parameters. We studied recurrent neural networks (RNNs) in a teacher-student set-up, where the teacher RNN represents a biological neural network for which a connectome is available. The student has ac-

cess to this connectome, but not the teacher's single-neuron parameters (bias, gains, etc). We then optimized the student to match the activity of a subsample of teacher neurons. We found that the activity of unrecorded neurons can be inferred well above chance given a small number of recorded neurons. The minimum number of recorded neurons depends on the dimensionality of the network dynamics, but not the number of units in the network. In contrast, when the connectome is not known, inference of unrecorded activity is unsuccessful. We further demonstrated that the connectivity constraints modify the landscape of solutions, making it consistent with a high-dimensional convex optimization problem. Altogether, we characterized how measured connectivity of a network, beyond a coarse statistical description, can bridge micro- and mesoscale descriptions of neural dynamics. Our theory provides a guide for when connectome data can, or cannot, facilitate the inference of neural circuit function and draws attention to a qualitative difference in the solution spaces of connectome-constrained and unconstrained networks.

1-162. Revisiting efficient representations of space in hierarchical place field populations

Zach Cohen[1,2] ZCOHEN1@G.HARVARD.EDU
Jan Drugowitsch[3] JAN_DRUGOWITSCH@HMS.HARVARD.EDU

[1]Harvard University
[2]Neurobiology
[3]Harvard Medical School

Hippocampal place cell populations putatively encode allocentric spatial position. Recent experiments in bats and rodents found that these place cells exhibit heterogeneously sized and overlapping place fields. These characteristics suggest a hierarchical spatial tuning landscape, in which the (relatively fewer) larger place fields of some place cells encompass the (relatively more numerous) narrowly tuned place fields of other cells. In this work, we investigate the effectiveness of such hierarchical codes for encoding allocentric spatial position. We do so by deriving closed-form expressions for how well joint population activity discriminates between nearby spatial locations, a metric given by the population Fisher information. In contrast to periodically tuned grid cells, in which hierarchical modules have been shown to expand the encoding capacity of a population, we found that a hierarchical place field code is no more powerful in encoding position than a non-hierarchical one. Furthermore, we show that if we additionally require neurons to obey metabolic constraints, hierarchical place fields become strictly inferior to non-hierarchical codes with minimal tuning widths. Our results generalize to arbitrary stimulus dimensionality, beyond the 2-dimensional spatial stimulus case considered here. Further, they stand in conflict with recent simulation-based theoretical work which argues that hierarchical place fields are the optimal tuning strategy for encoding position - a conflict that we attribute to the use of strict modeling assumptions and ambiguous statistics. The influence of these factors becomes obvious once one applies an analytic approach, as we have. Overall, our results demonstrate that conceptualizing place cells as solely implementing an allocentric positional code fails to predict their hierarchical tuning structure. Instead, we speculate that this structure finds its purpose in supporting more complex computations underlying navigation rather than simply implementing a positional code.

1-163. Spatial integration properties in MT neurons affect spatiotemporal motion discrimination

Lucia Arancibia[1,2] LARANCIBIA@CRM.CAT
Klaus Wimmer[1] KWIMMER@CRM.CAT
Alexandre Hyafil[1] ALEXANDRE.HYAFIL@GMAIL.COM
Jacob L. Yates[3] YATES@BERKELEY.EDU
Alexander Huk[4] ALEXHUK@G.UCLA.EDU

[1]Centre de Recerca Matematica
[2]Computational Neuroscience Group
[3]UC Berkeley
[4]UCLA

Perception requires integrating noisy dynamic visual information across the visual field to identify stimuli and guide conscious decisions. While the temporal integration process has been studied extensively in experiments with highly controlled visual stimuli and reverse-correlation techniques, the nonlinear mechanisms underlying spatial integration are often neglected. Here, we show that the spatial structure of the stimulus modulates neural responses in visual cortex and impacts perceptual choices. In particular, we find that monkeys integrate spatial evidence sublinearly in a motion discrimination task due to (i) surround suppression effects causing an attenuation of the responses to motion in the center of the stimulus, and (ii) weaker im- pact of motion further away from the

fovea. To test whether these effects originate in the Middle Temporal area (MT), we estimate and validate the spatiotemporal direction sensitivity kernels of MT neurons using nonlinear regression models. Our findings can be synthesized in a two-stage model of perceptual decision making in which spatial context effects modulate spatial stimulus integration in neurons of visual cortex, and a decision area supports temporal integration to give rise to perception. Taken together, our work challenges the assumption that spatially distributed stimuli are integrated uniformly and calls for a new class of decision making models that take into account spatial processing.

1-164. Combining the connectome and neural imaging to infer causal whole-brain dynamics in C. elegans

Matthew Creamer[1,2] MATTHEW.S.CREAMER@GMAIL.COM
Jonathan Pillow[1] PILLOW@PRINCETON.EDU
Andrew Leifer[1] LEIFER@PRINCETON.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute

A fundamental goal in neuroscience is to characterize the underlying network that generates recorded neural activity. However, in most model organisms, the anatomical connectivity is unknown, or the causal functional connections between neurons is unknown. As a result, the relationship between a network's functional activity and its anatomical connectivity remains unclear. Here, we investigate the relationship between the connectome and causal functional connections using whole-brain recordings of neural activity in C. elegans. We fit a linear dynamical systems model to a dataset with optogenetic perturbations, where each neuron in the brain was stimulated independently, thus measuring the causal interactions between neurons. We show that our model captures these neuron-to-neuron causal interactions, generalizes to novel animals, and even predicts the activity of held-out neurons. Although this model could explain the causal connections of data, the model weights did not correlate to the connectome, suggesting that the model was distinct from the animal's nervous system. Therefore, we constrained the model to only have connections between neurons with known synapses. We show that this sparser model loses little performance, despite having only 10% of the connections in the full model. We then wanted to determine if synapse count alone predicted the model weights. We show that the learned model weights in the anatomically constrained model do not correlate with synapse count, suggesting that connectomics alone were not sufficient to generate the model. Taken together, we show that neural recordings alone were insufficient to infer the connectome, and synapse count alone could not explain our model. However, by combining both neural recordings and the connectome, we were able to construct an anatomy constrained whole-brain model of the worm that can explain how the causal interactions of the nematode brain arise from anatomy.

1-165. Local mechanisms lead to global anticipation in structured and random networks

Jared Salisbury JSALISBURY@UCHICAGO.EDU
Stephanie E Palmer SEPALMER@UCHICAGO.EDU

University of Chicago

Prediction is a crucial function of the nervous system due to the existence of substantial sensory and motor delays. While prediction is well-understood from an engineering perspective, its neural implementation, particularly in the high-dimensional and disordered setting of neocortex, remains largely unknown. To address this, we present a framework for understanding the basic mechanisms of prediction in the simplest possible setting–a linear network responding to linearly changing input–which we term 'anticipation.' Previous work on scalar inputs revealed simple mechanisms of anticipation, namely, feedback and feedforward inhibition, and pointed to the importance of the time constant associated with that inhibition. Here, we generalize to vector inputs in order to explore population-level anticipation in networks with structured or random lateral connections. Typically, lateral connections slow down dynamics, making anticipation all the more crucial. Insights from the scalar case are applied to yield anticipatory population-level dynamics. We show that fine-tuned, exact solutions can be constructed using the eigendecomposition of the connection matrix, but these lack biological plausibility. More intriguingly, local mechanisms embedded within a network give rise to approximate global anticipation without fine-tuning. In particular, using random matrix theory, we derive analytically the densities of eigenvalues (in the large N limit) which determine the dynamics, gain, and time shift of random networks with local anticipatory mechanisms. Random matrices serve as a useful minimal model of the complex, disordered connectivity observed in cortex, and the resulting eigenvalue densities are universal in the sense that they do not depend on the precise distribution of connections. The results suggest that local mechanisms, such as spike frequency adaptation inherent to individual neurons, may be used by the brain as a robust approximate solution to anticipation at the population

level.

1-166. How inhibition shapes spiking and bursting activity in stochastic recurrent networks

Audrey Teasley[1,2] AET67@BU.EDU
Gabriel Ocker[1,2] GKOCKER@BU.EDU

[1]Boston University
[2]Mathematics and Statistics

Pyramidal neurons in mammalian cortical layer 5 (L5) send long-range projections to other cortical areas and subcortical structures. Understanding the outputs of the cortex thus requires understanding the activity of these neurons. Thick-tufted L5 pyramidal neurons (TPNs) generate two types of action potentials: 1) the classic sodium-potassium action potentials (Na-K APs) generated at the axon hillock, which propagate down the axon to trigger neurotransmitter release and backpropagate across the soma and dendritic tree, and 2) slow calcium spikes in the apical dendrite, triggered by the coincidence of a back-propagating Na-K AP and dendritic depolarization [1]. These calcium spikes can then trigger a burst of Na-K APs. Parvalbumin positive (PV+) and Somatostatin-positive (SOM+) interneurons specialize in inhibiting TPNs' somata and apical dendrites, respectively, suggesting a role for regulating the bursting vs non-bursting regimes of pyramidal activity [2, 3, 4, 5]. Dendrite-dependent bursting has been proposed as a mechanism for coincidence detection [6], multiplexed spike-burst sensory coding [7], and a substrate for powerful and biologically plausible learning algorithms, e.g.,[8, 9]. Understanding these theoretically requires a theory for the joint spiking and bursting activity in cortical networks. Here, we study a simple stochastic model of dendrite-dependent bursting. Using tools from statistical field theory, we develop the joint probability density functional of spikes and bursts in recurrent networks of pyramidal and inhibitory cells. This allows the prediction of the TPNs' mean spike and burst rates, as well as pairwise and higher-order correlations between spikes and bursts. We show how somatic vs dendritic-targeting inhibition shapes the population activity. Our approach opens new avenues to investigating how dendritic nonlinearities interact with network structure to shape population activity.

1-167. Learning only a handful of latent variables produces neural-aligned CNN models of the ventral stream

Yudi Xie[1,2] YU_XIE@MIT.EDU
Esther Alter[1] ALTERS@MIT.EDU
Jeremy Schwartz[1] JEREMYES@MIT.EDU
James J. DiCarlo[1] DICARLO@MIT.EDU

[1]Massachusetts Institute of Technology
[2]Department of Brain and Cognitive Sciences

Image-computable modeling of primate ventral stream visual processing has made great strides via brain-mapped versions of convolutional neural networks (CNNs) that are optimized on thousands of object categories (ImageNet), the performance of which strongly predicts CNNs' neural alignment. However, human and primate visual intelligence extends far beyond object categorization, encompassing a diverse range of tasks, such as estimating the latent variables of object position or pose in the image. The influence of task choice on neural alignment in CNNs, compared to CNN architecture, remains underexplored, partly due to the scarcity of large-scale datasets with rich known labels beyond categories. 3D graphic engines, capable of creating training images with detailed information on various latent variables, offer a solution. Here, we asked how the choice of visual tasks that are used to train CNNs (i.e., the set of latent variables to be estimated) affects their ventral stream neural alignment. We focused on the estimation of variables such as object position and pose, and we tested CNNs' neural alignment via the Brain-Score open science platform. We found some of these CNNs had neural alignment scores that were very close to those trained on ImageNet, even though their entire training experience has been on synthetic images. Additionally, we found training models on just a handful of latent variables achieved the same level of neural alignment as models trained on a much larger number of categories, suggesting that latent variable training is more efficient than category training in driving model-neural alignment. Moreover, we found that these models' neural alignment scores scale with the amount of synthetic data used during training, suggesting the potential of obtaining more aligned models with larger synthetic datasets. This study highlights the effectiveness of using synthetic datasets and latent variables in advancing image-computable models of the ventral visual stream.

1-168. Grid-like representation in value-based decision-making

Mark Orloff[1,2] MAORLOFF@UCDAVIS.EDU
Seongmin Park[3] SEONGMIN.A.PARK@GMAIL.COM
Jake Blumwald[4] JBLUMWALD@UCDAVIS.EDU
Philippe Domenech[5] PHILIPPE.DOMENECH@INSERM.FR
Erie Boorman[6] EDBOORMAN@UCDAVIS.EDU

[1]UC Davis
[2]Center for Mind and Brain
[3]French National Centre for Scientific Research
[4]University of California Davis
[5]Institut National de la Sante et de la Recherche Medicale
[6]University of California, Davis

Theoretical models propose that value-based decisions utilize a common neural currency for utility to compare choice options, where rewarding options such as foods, money, and social stimuli are encoded using a common neural value code in the brain (sometimes called the "brain valuation system [BVS]"). This value coding reflects an individual's subjective valuation—how rewarding something is based on an individual's preference. However, how a given choice option is transformed to its subjective value remains poorly understood. One candidate for how this transformation could happen is via the brain's the grid coding system. To test this hypothesis, we designed a novel task where human participants are asked to make value-based risky choices between two shapes that span a 2D attribute space between reward probability and reward amount while undergoing fMRI scanning. First, we utilized well-established computational models to calculate the subjective value of each participant's choices according to the Cumulative Prospect Theory (CPT) model. We show that blood-oxygen-level-dependent (BOLD) activity is associated with individuals' subjective value difference between the two options in the BVS, replicating previous findings in our new task. Second, using well-established analyses, we also find a grid-like representation of individuals' decision vectors in a subjective value space that is distorted as predicted by CPT. Strength of activation in these two systems (BVS and grid-coding system) is correlated across participants, suggesting these neural codes operate in tandem to construct and compare values. This finding reveals a previously unknown mechanism in the transformation of a reward's attributes into subjective value and has fundamental implications for understanding how the brain constructs and represents subjective value.

1-169. How populations of neurons encode perspective: insights from mouse visual cortex

Judith Hoeller[1] HOELLERJ@JANELIA.HHMI.ORG
Lin Zhong[1] ZHONGL@JANELIA.HHMI.ORG
Marius Pachitariu[1] PACHITARIUM@JANELIA.HHMI.ORG
Sandro Romani[2] ROMANIS@JANELIA.HHMI.ORG

[1]HHMI Janelia Research Campus
[2]Janelia Research Campus

As we move around the world, we see the same visual scene from different perspectives. While we might be aware of such perspective deformations, our percept of the scene remains stable. This suggests that scene representations in the brain must be both perspective-selective and invariant. Here we propose and test the hypothesis that perspective-selective and -invariant representations are encoded in mutually orthogonal subspaces of the same population of visual cortical neurons. This hypothesis can be viewed as a consequence of a general theoretical framework we developed that describes how visual representations ought to transform under perspective image transformations. Our basic premise is that visual representations form maps from the space of images to the space of neuronal responses. Then, under mild conditions on these maps, we show that image transformations act linearly on the space of neuronal responses. In particular, if an image is rotated, neuronal responses simply rotate in a high-dimensional space. An eigendecomposition of the neuronal rotation matrix then decomposes the high-dimensional space into one- and two-dimensional, mutually orthogonal subspaces, some of which are rotation-invariant while others are rotation-selective. We tested if populations of visual cortical neurons in mice followed our predictions when images were rotated either in- or out-of-plane (these generate distinct subgroups of perspective transformations). We presented many such images to mice while recording more than 50,000 neurons simultaneously in primary visual cortex cortex (V1) and surrounding higher-order areas (LM and AL). For in-plane rotations, we identified orthogonal subspaces that were rotation-selective and -invariant across image identities in V1 and LM but not in AL, whereas for out-of-plane rotations these subspaces only emerged in the higher-order areas. Our findings provide insights into how perspective transformations constrain neuronal coding. We suggest that different higher-order visual cortical areas are specialized to different perspective transformations.

1-170. Prefrontal gamma oscillations and task feature representations in a mouse rule shifting task

Caitriona Costello[1,2] CAITRIONA.COSTELLO@UCSF.EDU
Vikaas Sohal[1] VIKAAS.SOHAL@UCSF.EDU

[1]University of California, San Francisco
[2]Neuroscience

Dynamic environments require organisms to constantly update their behavioral strategies. We can learn about how the brain implements such flexibility by studying neural activity during a rule shifting task, in which mice sample combinations of odor and texture cues and learn to associate one cue with a reward. After mice learn this initial cue-reward association, the task rule changes such that a cue from the previously uninformative cue dimension becomes rewarded. This rule shift learning is known to depend on the synchronized gamma-frequency (~30-80 Hz) activity of parvalbumin-expressing inhibitory neurons (PVINs) in the medial prefrontal cortex (mPFC)[1,3]. Fascinatingly, gamma-frequency stimulation of PVINs delivered out-of-phase between mPFC hemispheres induces perseveration on the outdated rule[3]. By contrast, the same stimulation delivered in-phase does not disrupt learning in otherwise normal mice, and it rescues learning in mutant mice that are perseverative[1,3]. Although the mPFC is known to represent values, actions, and outcomes to facilitate flexible, goal-driven behavior[2], little is known about how the mPFC represents behaviorally-relevant information during this rule shifting task, or about how gamma synchrony interacts with these representations to facilitate learning. Here, we use silicon probes to record from the mPFC during the rule shifting task. We find spiking that correlates with task-relevant actions, outcomes, and stimulus value. We see an increase in gamma bursts (transient increases in gamma power) during unexpected negative outcomes and as mice begin to make decisions that deviate from the outdated rule. Units with significant spike-LFP phase locking exhibit selectivity for a variety of task features. We optogenetically drive gamma oscillations either in-phase or out-of-phase between mPFC hemispheres and characterize the resulting effects on neural dynamics. These results begin to uncover how gamma oscillations may modulate neural representations of task-relevant information as behavioral strategies are updated in response to changing environmental contingencies.

1-171. Hippocampal spatial representations are modulated by cyclic endocrine factors

Nora Wolcott[1,2] NWOLCOTT@UCSB.EDU
William Redman[3,4] WREDMAN@UCSB.EDU
Michael Goard[3] MICHAEL.GOARD@LIFESCI.UCSB.EDU
Marie Karpinska[1] GOARD.LAB@GMAIL.COM
Lori Mandjikian[1] MANDJIKIAN@UMAIL.UCSB.EDU
Emily Jacobs[1] EMILY.JACOBS@PSYCH.UCSB.EDU

[1]University of California Santa Barbara
[2]Molecular, Cellular, and Developmental Biology
[3]University of California, Santa Barbara
[4]Dynamical Neuroscience

It has become increasingly clear that the neural circuits underlying spatial navigation are fundamentally shaped by cyclic endocrine factors. Despite this, the systems-level mechanisms underlying hormone modulation of these circuits remain poorly understood. Here, we used chronic two-photon (2P) imaging to demonstrate that dendritic spine turnover, intracellular activity, and place cell remapping are modulated by cyclic fluctuations in sex steroid hormones. These phenomena were observed using a supervised deep neural network for endocrine state classification and a custom surgically implanted microperiscope for imaging. Using these techniques, we first measured dendritic spine turnover over the course of several weeks, and found ~15.0% fluctuation in turnover, with spine density peaking during the high estradiol stage (proestrus). To understand how these structural changes affect higher-level dendritic dynamics, we tracked dendritic coupling across the cycle, and found that both somatodendritic and interdendritic coupling were highest during high estradiol stages, possibly due a higher ratio of synaptic input to intrinsic excitability. Next, to characterize population-level hippocampal responses, we used a remapping paradigm in which the mice first run through environment A, then environment B, followed by a reintroduction to environment A (A > B > A'). Both remapping between environments (A vs B) and stability within environment (A vs A') were greatest during stages with elevated estradiol and synaptic density, results that were mirrored in the decoding accuracy of spatial position. The described shifts in spine density, dendritic coupling, and population-level coding suggest that endocrine state plays a significant role in shaping hippocampal network dynamics.

1-173. A mesoscopic electrophysiology platform to measure the spectro-temporal dynamics of large-scale brain states and connectivity

Tobias Teichert[1,2] TEICHERT@PITT.EDU

[1]University of Pittsburgh School of Medicine
[2]Psychiatry

Current models of large-scale brain states and connectivity are largely built around measurements of slow BOLD fluctuations whose temporal dynamics are far removed from the millisecond time scale that governs neural activity. However, electrophysiological recordings are often not suitable for understanding and modeling brain-wide states and dynamics because they either have a small field of view (microscopic electrophysiology) or because they have relatively coarse spatial resolution (macroscopic electrophysiology). To address this methodological need, we developed the first electrophysiological mesoscope by tiling the volume of an entire monkey hemisphere with 992 electrode contacts that were distributed across 62 chronically implanted multi-electrode shafts. The rich mesoscopic electrophysiology data sets provide important new insights for the development and refinement of large-scale models of brain-wide states and connectivity. (1) We measured functional connectivity of the monkey brain with a field of view that cannot be achieved using microscopic electrophysiology and a spectro-temporal resolution that is not possible using fMRI. (2) We revealed the emergence and recurrence of stable brain-wide states that last on the order of tens of seconds. (3) We show that ketamine strongly attenuates functional connections between brain regions and abolishes the emergence and recurrence of stable brain-wide states. These findings suggest that long-range functional connections between brain regions are critical for the emergence of stable brain states.

1-174. Multi-Region Markovian Gaussian Process: an efficient method for multi-region analysis

Weihan Li[1,2] WEIHANLI@GATECH.EDU
Anqi Wu[3] ANQIWU@GATECH.EDU

[1]Georgia Institute of Technology
[2]CSE
[3]georgia institute of technology

Uncovering the intricate interactions among various brain regions stands as a pivotal pursuit in neuroscience research. Numerous statistical methodologies have been diligently applied to tackle this challenge by probing communication patterns across multiple brain regions. Broadly categorized into Gaussian Process (GP) and Linear Dynamical System (LDS) classes, these methods offer distinct advantages. The GP-based approach, with the powerful representation ability from multi-output kernels, excels in discovering latent variables, encompassing crucial information like frequency bands and directions in brain communications. Conversely, while computationally efficient with a linear cost in time points, the LDS-based approach easily enables multi-state analysis but lacks the nuanced expressiveness of the GP in latent representation. In this study, we amalgamate the strengths of both methodologies by constructing an LDS that mirrors a multi-output GP. Introducing the Multi-Region Markovian Gaussian Process (MRM-GP) as a probabilistic dimensional reduction framework, we exploit the discrete state space representation of Markovian GPs. Notably, to the best of our knowledge, our work is the first to establish a connection between an LDS and a multi-output GP explicitly modeling frequencies and phase delays in the latent space of neural data. The resultant latent variable model adeptly discerns both across- and within-region communications, elucidating phase delays across regions and dynamics within regions. Consequently, this model yields an interpretable low-dimensional representation of multi-region neural data, revealing the direction of communication flow among regions and untangling oscillatory interactions in diverse frequencies.

1-175. ATP demand and supply in neurons: Perfectly balanced, as all things should be.

Chaitanya Chintaluri[1] CCLURI@GMAIL.COM
Tim Vogels[2,3] TIM.VOGELS@IST.AC.AT
Anjali Amrapali Vishwanath[4] ANJALI.AMRAPALI@ICM-INSTITUTE.ORG
Jaime De Juan-Sanz[4] JAIME.DEJUANSANZ@ICM-INSTITUTE.ORG

[1]Institute of Science and Technology Austria
[2]IST, Austria
[3]_

[4]Sorbonne Universite, Institut du Cerveau

Neurons are energetically expensive. A substantial part of this energy is required to process action potentials and to release neurotransmitters from their presynaptic terminals. It is also known that the firing frequency of a neuron can vary abruptly in response to inputs and therefore the metabolic costs to process them. How these sudden changes in energy demands are conveyed to the ATP-manufacturing mitochondria is largely unknown. Here, based on previous studies and some new unpublished experimental observations, we built a computation model that links the spiking of neurons to the influx and efflux of Ca2+ into the mitochondria, which results in the acceleration of select mitochondrial enzymes and ATP production. We show that spike-dependent mitochondrial calcium upregulates ATP production rates to rapidly match ATP demand.

We began with a simplified tricarboxylic acid cycle (TCA) and ATP production model by Nazareth et al. To this, we added a Ca2+ dependency for relevant enzymatic reactions. We then tested the differences between three particular experimental perturbations involving mitochondrial Ca2+ levels. Firstly i) the control case when each neuronal spike included a small rapid influx of Ca2+ into the mitochondria which is also rapidly removed. Next ii) the case when Ca2+ is prohibited from entering the mitochondria (mitochondrial calcium uniporter knockdown, MCUKD), and lastly iii) the case when the Ca2+ efflux is restricted (Letm1KD). We also explored five different spontaneous firing rates of the neurons and showed that in all these cases, the ATP levels in the Letm1KD case increased more than in the control due to activity-dependent Ca2+. This is a first-of-its-kind model that links biochemical reactions of mitochondrial ATP production with activity-dependent Ca2+ signaling and provides a framework to study the balancing act of ATP demand and supply in neurons.

1-176. Few-shot temporal pattern learning via spike-triggered boosting of somato-dendritic coupling

Gaston Sivori[1,2] GAST.SIVORI@GMAIL.COM
Tomoki Fukai[1] TOMOKI.FUKAI@OIST.JP

[1]Okinawa Institute of Science and Technology
[2]Neural Coding and Brain Computing Unit

Neurons can detect temporal patterns amidst noise in continuous input streams, but their tuning rapidity tends to be overlooked. Consequently, previous spiking neuron models lack mechanistic features that explain how biological neurons rapidly converge to target responses. We propose a novel biologically-plausible learning rule that relies on known mechanisms implicated in Hebbian synaptic plasticity. In our models, an internal error signal propagates somatic spiking activity to dendritic compartment facilitating unsupervised learning of repeatedly activated presynaptic-neuron communities. Our work predicts that, under proposed calcium-based plasticity traces, spike-triggered transient boosting of compartmental coupling—akin to backpropagating action potential (bAP)—is essential for pattern learning, and that the inclusion of N-methyl-D-aspartate receptor (NMDAR) dynamics significantly increases the signal-to-noise ratio of the detection capacity. We show the utility of our proposed plasticity rules in three scenarios: we replicate STDP protocol data from CA3-CA1 Schaffer collateral synapses, we demonstrate rapid learning of place field-reward association in spatial navigation tasks, and we show how pre-existing cell assemblies in recurrent networks require few pattern presentations when their within-assembly connectivity is strong enough. Our results shed light on the self-supervising function of bAP for pattern learning, its tight relationship with calcium-based synaptic plasticity mechanisms, and the ability of pre-existing cell assemblies to provide a reservoir of activity patterns ready to rapidly self-organize and respond to novel temporal patterns of information.

1-177. Dynamic value codes in the medial prefrontal cortex to inform decision-making

Xulu Sun[1] XULU.SUN@UCSF.EDU
Alison Comrie[1,2] ALISON.COMRIE@UCSF.EDU
Emily Monroe[1] EMONROE@UCDAVIS.EDU
Ari Kahn[3] ARIK@PRINCETON.EDU
Abhilasha Joshi[4] ABHILASHA.JOSHI@UCSF.EDU
Jennifer Guidera[1] JENNIFER.GUIDERA@UCSF.EDU
Lulu Tong[5] LULU.TONG@UCSF.EDU
Eric Denovellis[5,6] ERIC.DENOVELLIS@UCSF.EDU
Timothy Krausz[1] TIM.KRAUSZ@UCSF.EDU
Donghoon Shin[7] DONGHOON.SHIN@BERKELEY.EDU
Joshua Berke[1] JOSHUA.BERKE@UCSF.EDU
Nathaniel Daw[3] NDAW@PRINCETON.EDU
Loren M. Frank[5] LOREN.FRANK@UCSF.EDU

[1]UCSF
[2]Physiology
[3]Princeton University
[4]UCSF, HHMI
[5]University of California, San Francisco
[6]Department of Physiology and Psychiatry, Howard Hughes Medical Institute
[7]University of California, San Francisco, University of California, Berkeley

Storing information about previous decision outcomes allows animals to estimate the values of different options. These values can guide future decisions, increasing the odds of making a favorable choice. The medial prefrontal cortex (mPFC) is critical for value-based decision-making and mPFC neurons can encode value, but how value-tuned neurons inform decisions is not fully understood. It also remains unclear how value representations interact with other task-structure representations in mPFC. We addressed these questions using rats performing a spatial bandit task with probabilistic rewards distributed across multiple foraging patches. On each trial, rats chose to stay within their current patch or switch to a distant patch. An HMM fit to the rats' behavior allowed us to infer subjective option values and capture value-based foraging decisions. Generalized linear models (GLMs) applied to 100-500 simultaneously recorded mPFC neurons (per day per rat) revealed that 20-80% of neurons significantly encoded the relative value of staying in the current patch as compared to switching to a different patch, a decision variable directly relevant for the task. These value-encoding neurons typically exhibited strong tuning to task progression (fractional distance along journeys between reward locations), implementing a multiplicative code for abstract task structure and option values. These findings indicate that the value-coding axis in the neural population activity changes as animals progress through each trial. We confirmed this dynamic value axis in the neural state space using cross-validated decoders: a time-varying value axis led to the optimal readout of values and optimal prediction of stay/switch decisions. Strikingly, utilizing the neurally-decoded values to predict single-trial decisions was often more accurate than predictions of the HMM that provided the behavioral value estimates. Our results reveal a systematic interaction between value and task-progression representations in the mPFC neural population, which defines a dynamic, state-space axis for translating values into decisions.

1-178. Hierarchical neural dynamics across motor cortex and striatum during naturalistic movement

David Xing[1,2] DAVID.XING@NORTHWESTERN.EDU
Josh Glaser[1] J-GLASER@NORTHWESTERN.EDU
Andrew Miri[1] ANDREWMIRI@NORTHWESTERN.EDU

[1]Northwestern University
[2]Neurobiology

Naturalistic behaviors involve many movement modalities, from precise, targeted reaches, to innate behavioral sequences such as grooming, forming a hierarchy of behavioral complexity. One challenge the nervous system faces is the need to accurately carry out each of these movements while also being able to flexibly switch between them. Motor cortex and striatum have previously been implicated in the execution and selection of motor actions, but how their activity dynamics are organized across this behavioral hierarchy is still poorly understood. Whether the nervous system engages specific subpopulations, or specific modes of coordinated activity across a wide range of behaviors is unknown. Here, we developed a novel paradigm that allows for the investigation of neural dynamics across behaviors that require agility and dexterity, such as climbing and walking across an irregular grid, as well as innate behaviors, such as eating and grooming. We utilized UMAP projections of muscle activity during

free behavior to parcellate the behavioral space into distinct states. In order to interrogate the organization of neural dynamics across these behavioral states, we chronically implanted Neuropixels probes to simultaneously record large numbers of neurons in the caudal forelimb area (forelimb M1) and dorsolateral striatum. We identified striatal as well as cortical neurons whose activity range from highly selective for specific behaviors to uniform across all behaviors. On a population level, we discovered both distinct neural subspaces corresponding to separate sets of behaviors as well as highly overlapping subspaces among more similar movement modalities. These results suggest that the neural population does transition between different dynamical regimes throughout naturalistic behaviors, and the neural dynamics corresponding to different behaviors are organized hierarchically.

2-001. Estimating shape distances on neural representations with limited samples

Brett Larsen[1] BRETTLARSEN@FLATIRONINSTITUTE.ORG
Dean Pospisil[2] DP4846@PRINCETON.EDU
Sarah Harvey[3,4] SHARVEY@FLATIRONINSTITUTE.ORG
Alex Williams[3] AWILLIAMS@FLATIRONINSTITUTE.ORG

[1]New York University; Flatiron Institute
[2]Princeton University
[3]Flatiron Institute
[4]Center for Computational Neuroscience

Quantitative comparisons of neural population dynamics across biological systems—e.g. different subjects, animal species, or brain areas—and to artificial network dynamics are of longstanding interest to systems neuroscience. Many metrics of functional, population-level similarity have been proposed including Representational Similarity Analysis (RSA), Centered Kernel Alignment (CKA), and shape distances. However, we still have a poor grasp on fundamental questions: How many neurons, trials, and behavioral conditions do we need to experimentally measure in order to accurately assess the similarity of two neural populations? Here, we mathematically derive concrete answers to these questions for the Procrustes shape distance—a measure of representational distance with desirable theoretical properties (symmetry and triangle inequality; Williams, 2021). We find that the problem is challenging for high-dimensional manifolds—for example, to compensate for a twofold increase in dimensionality, there must be a fourfold increase in the number of sampled conditions. To mitigate these challenges, we introduce a new method-of-moments estimator with a tunable bias-variance tradeoff. We show that this estimator achieves superior performance to standard estimators, particularly in high-dimensional settings. Furthermore, since our approach bounds the bias and variance of the estimate, it naturally produces a confidence interval, which we show to be an accurate reflection of uncertainty in simulation and on semi-synthetic experimental datasets with established ground truth. Finally, we leverage this new estimator to analyze mouse visual cortical responses to 2800 natural images. Thus, we lay the foundation for a rigorous statistical theory for high-dimensional shape analysis, and we contribute a new estimation method well-suited to practical scientific settings.

2-002. Dynamic dimensionality reduction of neural data by timescale separation using Predictable Mode Decomposition

Tosif Ahamed AHAMEDT@HHMI.ORG
Carsen Stringer STRINGERC@JANELIA.HHMI.ORG
Marius Pachitariu PACHITARIUM@JANELIA.HHMI.ORG
Qingqing Zhang ZHANGQ@JANELIA.HHMI.ORG

HHMI Janelia Research Campus

Recent methodological advances have enabled high dimensional recordings of neural activity. These recordings have revealed complex dynamics operating over a hierarchy of interacting timescales. Although several approaches exist to reduce the dimensionality of such datasets, few methods do so while disentangling the dynamics explicitly according to timescales. Inspired by ideas of timescale separation in model reduction of nonlinear dynamical systems, here we present "Predictable Mode Decomposition" (PrMD), a method for reducing dimensionality that captures the most predictable timescales of ongoing neural dynamics. Specifically, PrMD finds a linear projection that maximizes the "average predictability time" (APT), which is defined as the integral timescale of the fraction of variance explained by a time-series model. Regular oscillations and slowly changing dynamical variables are ranked as the top predictable modes (PrMs), while noisy or rapidly changing processes are ranked lowest. PrMD combines and generalizes two well-known methods: slow feature analysis, which extracts slowly changing variables, and jPCA, which extracts dominant oscillatory modes. Solving PrMD is efficient, even on

datasets containing ~10,000 neurons, as it can be framed as an eigenvalue problem. We use PrMD to study visual processing of time-varying stimuli by analyzing calcium recordings in mouse V1 driven by dynamical gratings, in which contrast, orientation or spatial frequency oscillate. PrMD extracts stimulus-related neural dynamics from these recordings whereas other methods like PCA fail. Specifically, PrMD finds additional oscillatory timescales other than the timescale of the stimuli, revealing nonlinear neural processing of the oscillating gratings. On other neural datasets, PrMD discovers interpretable components that provide insights into the hierarchy of behavioral timescales. It can further provide a starting point for exploring interactions across timescales, identifying dynamically coherent motifs as well as building simple dynamical systems models that capture the most predictable timescales.

2-003. The non-specific matrix thalamus facilitates the cortical information processing modes relevant for conscious awareness

Eli Muller[1] ELI.MULLER@SYDNEY.EDU.AU
Brandon Munn[1,2] BRANDON.MUNN@SYDNEY.EDU.AU
Yuri B. Saalmann[3] SAALMANN@WISC.EDU
Mac Shine[1,4] MAC.SHINE@SYDNEY.EDU.AU
Michelle Redinbaugh[3] MREDINBA@STANFORD.EDU
Joseph Lizier[1] JOSEPH.LIZIER@SYDNEY.EDU.AU
Michael Breakspear[5] MJBREAKS@GMAIL.COM

[1]The University of Sydney
[2]Faculty of Medicine and Health
[3]University of Wisconsin-Madison
[4]Brain and Mind Center
[5]The University of Newcastle

The neurobiological mechanisms of arousal and anaesthesia remain poorly understood. Recent evidence highlights the key role of dynamic interactions between the cerebral cortex and the diffusely-projecting matrix thalamic nuclei. Here, we interrogate these processes in a whole-brain corticothalamic neural mass model endowed with targeted and diffusely projecting thalamocortical nuclei inferred from empirical data. This model captures key features seen in propofol anaesthesia, including diminished network integration, lowered state diversity, impaired susceptibility to perturbation, and decreased corticocortical coherence. Collectively, these signatures reflect a suppression of information transfer across the cerebral cortex. We recover these signatures of conscious arousal by selectively stimulating the matrix thalamus, recapitulating empirical results in macaque monkeys, as well as wake-like information processing states that reflect the thalamic modulation of large-scale cortical attractor dynamics. Our results highlight the role of matrix thalamocortical projections in shaping many features of complex cortical dynamics to facilitate the unique communication states that support conscious awareness.

2-004. Identifying interpretable latent factors within and across brain regions

Andrew Ulmer[1,2] ANDREW.ULMER@NORTHWESTERN.EDU
Andrew Zimnik[3] AZ2412@COLUMBIA.EDU
Abigail Russo[3] RUSSOAA@GMAIL.COM
Vladislav Susoy[4] VLADISLAV_SUSOY@FAS.HARVARD.EDU
Laura Driscoll[5] LNDRISCO@STANFORD.EDU
Xinyue An[1,6] XINYUE.AN@NORTHWESTERN.EDU
Ann Kennedy[1,7] ANN.KENNEDY@NORTHWESTERN.EDU
Mark Churchland[3] MC3502@COLUMBIA.EDU
Josh Glaser[1] J-GLASER@NORTHWESTERN.EDU

[1]Northwestern University
[2]Interdepartmental Neuroscience
[3]Columbia University
[4]Harvard University
[5]Allen Institute for Neural Dynamics
[6]Department of Neurology
[7]Neuroscience

All behaviors result from the interplay of multiple neural computations spread across brain regions. Understanding how a behavior is generated requires delineating these computations from one another. Traditionally, identifying these critical 'computational building blocks' requires a researcher to impose structure on the data: task epochs

are analyzed separately or population-level structure is sought using supervised dimensionality reduction methods. While these approaches can be successful, they have obvious drawbacks – many behaviors cannot be easily parsed into distinct epochs, and it is often unclear, particularly in exploratory data analyses, what type of structure to look for in neural data. Unsupervised methods, like principal component analysis (PCA) identify large neural signals, yet often do not provide interpretable low-dimensional representations of the data. Moreover, existing approaches fail to dissect the role separate brain regions play within a given low-dimensional computation.

Here, we present multi-region sparse component analysis (mSCA), an unsupervised dimensionality reduction method that produces interpretable, low-dimensional representations within and across neural populations. Our method builds upon the finding that distinct computations are dissociable in time from one another, and thus aims to find sparsely occurring latent factors. We first demonstrated the power of encouraging sparse factors within a single neural population across diverse datasets, including monkey motor cortex during reaching and C. elegans during mating. Our approach found not only structure previously reported using supervised methods, but also novel structure, such as posture-related signals distinct from movement-related signals. We then examined how mSCA, when applied to recordings from primary motor cortex and supplementary motor area during a cycling task, found interpretable factors that were unique to, or shared across multiple neural populations. Finally, mSCA accurately learns time-delays at which individual factors affect different neural populations, allowing users to demix neural computations into interpretable factors, and understand their flow across populations.

2-005. The role of striatal dopamine in learning to perform a new task

Yoel Sanchez Araujo[1,2] SANCHEZ.ARAUJO@PRINCETON.EDU
Alejandro Pan-Vazquez[3] APV2@PRINCETON.EDU
Jonathan Pillow[3] PILLOW@PRINCETON.EDU
Nathaniel Daw[3] NDAW@PRINCETON.EDU
Ilana Witten[3] IWITTEN@PRINCETON.EDU

[1]PNI, Princeton University
[2]Neuroscience
[3]Princeton University

Dopamine in the striatum is well-recognized in supporting reward-based learning. However, dopamine is normally studied in animals that have already learned a task's structure, thus the role of dopamine in learning a task de-novo is largely unknown. We address this gap by comparing behavioral and neural data from 22 mice learning to associate visual gratings of varying contrasts with actions to obtain rewards, without shaping or experimenter intervention. Our goal is to understand if and how dopamine signals across the striatum predict individual differences in task acquisition. Fiber photometry was used to acquire calcium signals in dopaminergic axon terminals in 3 striatal subregions: the nucleus accumbens core (NACC), dorsomedial striatum (DMS), and dorsolateral striatum (DLS). To model the dopaminergic signals across learning, we built a Gaussian generalized linear model (GLM) that captures the contributions of task events such as the visual stimuli, actions, and reward delivery. To model the behavioral data, we built a hierarchical Bernoulli GLM that captures the relationship between the animals choices, a bias towards either choice, the stimulus presented to the animal, previous choice history, and lapse. We find learning trajectories are idiosyncratic across mice, with most mice exhibiting different learning trajectories for visual stimuli presented on the left versus right side. Within each animal, in all regions, we found a strong relationship between side-specific behavioral performance and side-specific dopaminergic responses. Surprisingly, before rewards were delivered, responses to visual stimuli were present in dopamine signals in DMS, but not the other regions. Across mice, the strength of pre-existing DMS responses predicted the extent of learning on the side contralateral to the recording. In summary, while dopaminergic responses track learning across striatal regions, signals in DMS respond to a visual stimulus even before being paired with reward, and these pre-existing sensory responses predict the extent of learning.

2-006. Continuous decoding of complex vocalizations from multi-region latent neural dynamics

Pablo Tostado[1,2] PATOSTAD@UCSD.EDU
Jingya Huang[3] JIH201@UCSD.EDU
Ezequiel Arneodo[1] ZEKE.ARNEODO@GMAIL.COM
Lauren Ostrowski[1] LOSTROWS@HEALTH.UCSD.EDU
Daril Brown[1] DEBROWN@UCSD.EDU
Lauren Stanwicks[1] LSTANWIC@HEALTH.UCSD.EDU
Abdullah Alothman[1] AALOTHMA@UCSD.EDU
Timothy Gentner[1] TGENTNER@UCSD.EDU
Vikash Gilja[1] VGILJA@UCSD.EDU

[1]University of California, San Diego
[2]Bioengineering
[3]University of California San Diego

Skilled vocal behavior, like speaking or singing, requires the coordinated and precise temporal control of vocal effector muscles. The learning and execution of complex vocalizations is dependent upon the engagement of populations of neurons distributed across interconnected brain networks, including motor and auditory pathways. In this study, we investigate the neural correlates of vocal production in the context of song behavior through multi-region, single-cell-resolution neural recordings from awake-singing zebra finches. We used Neuropixels probes to capture the simultaneous activity of hundreds of neurons in the HVC and RA regions of the posterior motor pathway of the avian brain along with the bird's own song. State-space analyses revealed structured trajectories in low-dimensional manifolds in both HVC and RA tied to vocal behavior similar to those reported for human and non-human primate motor cortex during motor actions [1, 2]. Analysis of the inferred dynamics unveiled key differences in their relation to vocal output, and suggested an HVC manifold of higher dimensionality than that of RA during vocal production. Next, we utilized a foundational model for audio compression to derive audio embeddings facilitating the reconstruction of birdsong. We coupled this model to a neural network trained to translate inferred neural latents to song embeddings. Our results demonstrate high-fidelity reconstruction of song spectrograms from neural trajectories.

2-007. Network models for distinguishing population-level learning mechanisms

Jacob Sacks[1,2] JSACKS6@CS.WASHINGTON.EDU
Emily Oby[3] EMO22@PITT.EDU
Jay A. Hennig[4] JHENNIG@FAS.HARVARD.EDU
Alan D. Degenhart[3] ALAN.DEGENHART@GMAIL.COM
Patrick T. Sadtler[3] PATRICK.T.SADTLER@GMAIL.COM
Kristin M. Quick[3] KRISTINMQUICK@GMAIL.COM
Stephen I. Ryu[5] SEOULMAN@STANFORD.EDU
Elizabeth C. Tyler-Kabara[6] ELIZABETH.TYLERKABARA@AUSTIN.UTEXAS.EDU
Steven M. Chase[7] SCHASE@CMU.EDU
Byron Yu[7] BYRONYU@CMU.EDU
Aaron Batista[3] AARON.BATISTA@PITT.EDU
Matthew D. Golub[1,8] MGOLUB@CS.WASHINGTON.EDU

[1]University of Washington
[2]Paul G. Allen of Computer Science & Engineering
[3]University of Pittsburgh
[4]Harvard University
[5]Stanford University & Palo Alto Medical Foundation
[6]University of Pittsburgh & University of Texas at Austin
[7]Carnegie Mellon University
[8]Computer Science & Engineering

How do populations of neurons modify their neural activity during learning to improve behavior? Learning can alter the dynamics of a neural population by modifying the inputs it receives or by rewiring local connectivity. To assess whether these population-level mechanisms are consistent with learning experiments, we developed task-optimized and data-constrained recurrent neural networks (RNNs), which explicitly consider a local population and its upstream inputs. We used task-optimized networks to characterize each mechanism's capabilities and limitations in improving behavior and modifying neural activity. We used data-constrained networks to determine whether each mechanism could reproduce empirical changes in single-trial neural population activity

during learning. We applied these models to investigate the mechanisms driving changes to primary motor cortex (M1) activity in monkeys learning brain-computer interface (BCI) tasks. In a BCI, the experimenter defines the mapping from neural population activity to behavior and can perturb that mapping to study learning. Prior experiments have shown that perturbations requiring M1 to generate activity patterns within the manifold (WM) of pre-existing activity are relatively fast to learn (1-2 hours), whereas perturbations requiring off-manifold (OM) activity are slower to learn (~10 days). Here, we show that altering inputs of task-optimized networks results in behavioral improvements and local activity changes consistent with fast WM learning. By contrast, altering local connectivity is required to achieve comparable behavioral proficiency in OM learning tasks. Next, we pre-trained data-constrained networks to reproduce before-learning single-trial monkey M1 population activity. In these models, altering inputs reproduced M1 activity following fast WM learning, whereas altering M1 connectivity was required to comparably reproduce M1 activity following multi-day OM learning. Altogether, these findings suggest multiple learning mechanisms operate on M1 at distinct timescales, with input changes on a timescale of hours and local connectivity changes on a timescale of days.

2-008. Modeling Visual Memorability Assessment with Autoencoders

Elham Bagheri[1,2] ELHAM.BAAGHERI@GMAIL.COM
Yalda Mohsenzadeh[3,2] YMOHSENZ@UWO.CA

[1]Western University, Canada
[2]Computer Science
[3]The University of Western Ontario

Image memorability refers to the phenomenon where certain images are more likely to be remembered by humans. It is a quantifiable and intrinsic attribute of an image. Understanding how visual perception and memory interact is important in both cognitive science and artificial intelligence. It reveals the complex processes that support human cognition and helps to improve machine learning algorithms by mimicking the brain's data processing mechanisms. To explore the computational underpinnings of image memorability, a pretrained autoencoder is used to replicate human-like memorability assessment, inspired by the visual memory game employed in memorability estimations. The relationship between an image's reconstruction error, distinctiveness in latent space, and its memorability score are then examined. This study leverages a VGG-based autoencoder pretrained on the ImageNet dataset. An empirical analysis is conducted using the MemCat dataset, including 10,000 images and their memorability scores. The memorability score assigned to each image represents the probability of that image being remembered by participants after a single exposure. The autoencoder is finetuned for one epoch, attempting to create a scenario similar to human memorability experiments where memorability is quantified by the likelihood of an image being remembered after being seen once. The reconstruction error of each image and its distinctiveness are calculated and correlated with the memorability score. The results indicate that there is a strong correlation between the reconstruction error and distinctiveness of images and their memorability scores. This suggests that images with more novel features that challenge the autoencoder's encoding capacities are inherently more memorable. These insights suggest a new pathway for evaluating image memorability, which could potentially impact industries reliant on visual content and mark a step forward in merging the fields of artificial intelligence and cognitive science. The current research opens avenues for utilizing neural representations as instruments for understanding and predicting visual memory.

2-009. Ensemble reactivations during brief rest periods drive fast sequence learning in non-human primates

Sandon Griffin[1,2] SANDON.GRIFFIN@UCSF.EDU
Preeya Khanna[3,4] PKHANNA@BERKELEY.EDU
Karunesh Ganguly[5] KARUNESH.GANGULY@UCSF.EDU
Hoseok Choi[1] HOSEOK.CHOI@UCSF.EDU
Lisa Novik[6] LMNOVIK@UCDAVIS.EDU
Katherina Thiesen[6] KLTHIESEN@UCDAVIS.EDU
Robert J. Morecraft[7] ROBERT.MORECRAFT@USD.EDU

[1]University of California, San Francisco
[2]Neurology
[3]University of California, Berkeley
[4]EECS and Neuroscience
[5]University of California at San Francisco
[6]University of California, Davis
[7]University of South Dakota

During motor learning, breaks in practice are known to facilitate behavioral optimizations. While this process has traditionally been studied over hours to days, recent studies in humans have demonstrated that rapid performance gains during early motor sequence learning are most pronounced following very brief rest periods. However, the precise causal neural mechanisms that facilitate performance gains after short breaks remain poorly understood. Here, we recorded neural ensemble activity in the motor cortex of non-human primates (NHPs) while they performed a visuomotor sequence learning (VMSL) task interspersed with brief rest periods. We found that task-related neural co-firing patterns were reactivated during brief breaks in practice. The rate and content of reactivations predicted the magnitude and pattern of subsequent performance gains. Interestingly, we found that performance gains and reactivations were positively correlated with cortical ripples (80-120 Hz oscillations) but anti-correlated with beta (β) bursts (13-30 Hz oscillations), which ultimately dominate rest periods after the fast-learning phase plateaus. We then applied 20 Hz epidural alternating current stimulation (ACS) to reduce reactivation rates in a phase-specific and dose dependent manner. Notably, this eliminated performance gains. Using a neural network model, we found that 20 Hz ACS may be more effective than 100 Hz ACS because of longer sustained inhibition during each cycle. Overall, our results suggest that the reactivations of task ensembles during brief rest periods are causal drivers of subsequent performance gains. In contrast, β bursts compete with this process to support stable performance.

2-010. Rhythmic Timing in Continuous-time Recurrent Neural Networks

Manav Shardha[1,2] MANAV.SHARDHA@GMAIL.COM
Matin Yousefabadi[1,3] YOUSEFAM@MCMASTER.CA
Jonathan Cannon[1] CANNOJ9@MCMASTER.CA

[1]McMaster University
[2]Psychology Neuroscience & Behaviour
[3]Psychology, Neuroscience, and Behavior

Humans' ability to anticipate rhythmic sequences across various tempos allows for synchronization and music making. The ability of various animal species to learn to discriminate between isochronous (steady) and irregular rhythms suggests that tempo-flexible rhythmic timing can be learned without the specialized brain mechanisms that humans possess. To study mechanisms underlying learnable rhythm perception, we investigated solutions produced by continuous-time recurrent neural networks (ctRNNs). The ctRNNs were trained using either full-FORCE or backpropagation to produce a train of isochronous pulses shifted earlier than the given input train over a range of tempos. The full-FORCE trained network makes precise predictions preceding each input pulse and predicts an additional pulse after the last input, thus demonstrating a generalized mechanism for rhythmic prediction. It reproduces several key elements of experimental data from human timing tasks such as increased accuracy and reduced bias towards the mean tempo of its training set as the number of input pulses increases. Further, the network exhibits human-like adherence to "Weber's Law" as the width of the generated pulse and the standard deviation of the difference in the inter-pulse intervals increase proportionally with the input period. The network produced circular oscillatory dynamics in neural phase space similar to the neuronal data from monkey medial frontal cortices when performing synchronization tasks. We were unable to replicate these results in the backpropagation-trained network. The consistency of the full-FORCE-trained network's behavior with human psychophysics results shows that human-like perception and anticipation of isochronous sequences can be attained through semi-supervised (predictive) learning in generic recurrent networks without special timing mechanisms. Future studies should explore the capacity of such networks to perform more complex rhythm perception tasks. Studying the solutions to these tasks can provide great insight into the mechanisms underlying rhythm perception in human and animal brains.

2-011. Brain-wide neural dynamics in dynamics foraging

YoungJu Jo[1,2] YJJO@STANFORD.EDU
Tony X. Liu[1] TXLIU@STANFORD.EDU
Daniel O'Shea[1,3] DJOSHEA@STANFORD.EDU
Jenny Shi[1] JSHI7@STANFORD.EDU
Thanh-Nga C. Shenoy[1] TSHENOY@STANFORD.EDU
Kishandra Anne Patron[1] KPATRON@STANFORD.EDU
Doo Kyung Kim[1] EDKKIM@STANFORD.EDU
Charu Ramakrishnan[1] CHARUR@STANFORD.EDU
Krishna V. Shenoy[1] SHENOY@STANFORD.EDU
David Sussillo[1] SUSSILLO@STANFORD.EDU
Karl Deisseroth[1] DEISSERO@STANFORD.EDU

[1]Stanford University
[2]Applied Physics & Bioengineering
[3]Bioengineering

Building, maintaining, and updating models of the world represent fundamental computations essential to the survival of animals. Past interaction with the environment, such as the history of choice and reward, may shape the subjective value of presently available actions and guide future decision-making. Here we investigated the brain-wide neural dynamics implementing such computations in the mouse brain. We modeled the dynamic environment using a head-fixed dynamic foraging task in which mice made sequential two-alternative choices, and experienced stochastic outcomes, across trials. The underlying reward probabilities, unobserved by the subject, remained constant for tens of consecutive trials before changing to new values. Upon training, the animals learned to dynamically match their choice preferences to their experience of rewarded choices over a time window of several minutes. This behavioral strategy could be described by a classic Q-learning model, which infers trial-by-trial estimates of each choice's reward probability ("action values"). Using multi-Neuropixels extracellular electrophysiology, we surveyed spiking activity across the brain to identify the neural substrates for the action value computation. We identified single-units for which activity correlated with action values in multiple brain regions. At the population level, data-driven dynamical systems modeling revealed brain-wide line attractor dynamics aligned to the relative value axis, thereby enabling bidirectional update of the value representation. Behavioral performance degradation upon optogenetic inhibition of the retrosplenial cortex was more potent during action value maintenance than during action value update. Interestingly, the opposite pattern was observed when inhibiting a subregion of the prefrontal cortex, suggesting potentially distinct roles in the action value computation. In addition, inhibition of the medial habenular Tac1 neurons reproduced and generalized previous findings to a brain-wide context. Collectively, these results provide a brain-wide picture of neural dynamics and causal influences in dynamic foraging, an important paradigm to study long-timescale computations in the brain.

2-012. Unsupervised quantification and classification of free-moving human behavior in euthymic bipolar disorder.

Zhanqi Zhang[1,2] ZHZ091@UCSD.EDU
Chi Chou[1] CKCHOU@UCSD.EDU
Holden Rosberg[1] HROSBERG@HEALTH.UCSD.EDU
William Perry[1] WPERRY@HEALTH.UCSD.EDU
Jared Young[1] J9YOUNG@HEALTH.UCSD.EDU
Arpi Minassian[1] AMINASSIAN@HEALTH.UCSD.EDU
Mikio Aoi[1,3] MAOI@UCSD.EDU
Gal Mishne[4] GMISHNE@UCSD.EDU

[1]University of California, San Diego
[2]Computer Science
[3]Neurobiology & Data Science
[4]University of California San Diego

Free-moving spontaneous behavior is the window to probe the brain and mind. Individuals with neuropsychiatric conditions such as bipolar disorder (BD) can exhibit distinctive patterns of behavior (McReynolds, 1962). Our objective was to quantify free-moving spontaneous human behavior in real-world contexts among euthymic BD individuals and differentiate them from a healthy control (HC) population based on these identified behavioral features.

We analyzed videos of 25 BD patients and 25 HC participants freely moving in an unexplored room for 15 minutes (Young et al, 2007). Utilizing a key-point estimation toolbox (Mathis et al., 2018), we extracted human poses and represented them through a latent variable model (Luxem et al., 2022). Clustering the latent representations

identified repeated behavioral motifs, revealing unique features of BD aligned with known clinical observations. For comparison, we employed action detection from computer vision (CV) models (e.g., MMAction and S3D) and from expert human annotation to extract motifs.

Notably, our unsupervised framework revealed differences between BD and HC populations. We detected differences in dwell time for identified motifs (two sample t-test, p-value: 0.04), in contrast to manually labeled or CV identified actions. We found a smaller behavioral repertoire in BD, as measured by transition probabilities, but with higher latent volume, indicating a higher-variance yet more stereotyped behavior in BD. Our approach outperformed CV models, expert human annotation, and even established clinical assessment scales (Hamilton Depression Scale and Young Mania Rating Scale) in distinguishing BD from HC (accuracy mean std; ours: 0.72 $\pm$ 0.11, CV actions: 0.62 $\pm$ 0.13: assessment scales: 0.60 $\pm$ 0.13, human annotation: 0.50 $\pm$ 0.14).

Our results underscore the potential of data-driven identified behavioral motifs to effectively differentiate BD from HC. This insight promises the development of novel diagnostics and treatment approaches for understanding neuropsychiatric conditions from behavior.

2-013. Serial tutoring reveals a composite song template

Kanghwi Lee[1] KANLEE@INI.ETHZ.CH
Richard H.R. Hahnloser[2] RICH@INI.ETHZ.CH
Hazem Toutounji[3,4] H.TOUTOUNJI@SHEFFIELD.AC.UK
Dina Lipkind[5] DINA.LIPKIND@GMAIL.COM

[1]Institute of Neuroinformatics, UZH and ETH Zurich
[2]ETH Zurich
[3]University of Sheffield
[4]Psychology
[5]City University of New York

A popular animal model for vocal learning is the zebra finch. According to the standard theory (Konishi, 1965), a young zebra finch first stores a memory of a heard song, a so-called template, and later uses the template to refine its own vocal output until it matches the song it heard. Despite more than 50 years of evidence for template-driven song learning, it is not yet clear how birds extract this template from their auditory experience. Is the template simply a single song instance or does it derive from several instances? To understand the template formation process, we conducted serial tutoring experiments on juvenile zebra finches. We serially exposed juveniles to two different tutor playbacks, first a source song and later a target song. When the second syllable in the target deviates by one semitone from the corresponding syllable in the source (AB -> AB+), juveniles ended up learning neither the source nor the target syllable, but a syllable in between, showing that a template (i.e., the song they end up learning) can be a composite generalizing over auditory memories. In another experiment where the target song contained two different variants of the corresponding source (one variant one semitone higher in pitch than the original, the other one semitone lower: ABAB -> AB+AB-), juveniles shifted their syllable to one variant and created a new syllable that matched the other variant, paradoxically learning a more complex auditory input better. To model this behavior, we hypothesized that birds have a built-in mechanism of generalizing over auditory memories. When forming a template, they first infer the causal structure underlying the heard songs and then they compute the optimal template, given the posterior of possible causal structures. This model is able to reproduce the templates extracted by juveniles in all experiment groups.

2-014. Possible Optimal Strategies for Orientation Coding in Macaque V1 Revealed with a Self-Attention Deep Neural Network (SA-DNN) Model

Xin Wang[1,2] 2301110701@STU.PKU.EDU.CN
Cai-Xia Chen[1] CHENCAIXIA@PKU.EDU.CN
Sheng-Hui Zhang[1] SHENGHUI.ZHANG@PKU.EDU.CN
Dan-Qing Jiang[1] JIANGDANQING8@PKU.EDU.CN
Shu-Chen Guan[3] ME@SHUCHEN.DE
Shiming Tang[1] TANGSHM@PKU.EDU.CN
Cong Yu[1] YUCONG@PKU.EDU.CN

[1]Peking University
[2]School of Psychological and Cognitive Sciences
[3]Justus-Liebig-Universitat Giessen

The orientation tuning bandwidths of individual V1 neurons are not sufficiently narrow to support fine psychophysical orientation discrimination thresholds. Here we explore the possibility that V1 neurons as a population may

apply optimal orientation coding strategies to achieve superb orientation tuning. We trained a self-attention deep neural network (SA-DNN) model to reconstruct a Gabor stimulus image from neuronal responses obtained through two-photon calcium imaging in five awake macaques. Each response field of view (FOV) contains 1400-1700 neurons, and their responses to a Gabor stimulus are used as the model inputs. The SA-DNN model consists of a self-attention mechanism followed by a feedforward layer. The self-attention mechanism can reveal cooperative coding by neurons activated by the Gabor stimulus, yielding attention maps that display two-way connections among neurons. Key findings include: (1) Neurons tuned to the stimulus orientation tend to have higher attention scores with all other neurons. The top 25% of orientation-tuned neurons with the highest mean attention scores can best reconstruct the stimulus images, while the bottom 50% neurons are unable to do so. (2) The responses of the top 25% neurons, after self-attention transformation, generate significantly sharpened population orientation tuning functions, with the amplitude increased by 3-5 times and bandwidth narrowed by approximately 30%. (3) After excluding the self-attention component, the forward propagation of the model would only reconstruct very coarse stimulus images. (4) The tuning sharpening displays an oblique effect: attention maps have higher variabilities at cardinal than at oblique orientations, producing more sharpened orientation tuning functions at cardinal orientations. These modeling results suggest that the self-attention mechanisms optimize orientation coding in macaque V1, reweighting responses to accentuate neurons based on attention scores. The results provide new insights into V1 neuronal connectivity, elaborating how self-attention refines neuronal interactions and reweights responses to process orientation information.

2-015. Efficient coding of a complex goal-directed behaviour in mouse medial-frontal cortex

Peter Doohan[1,2] PETER.DOOHAN@CCC.OX.AC.UK
Beatriz Godinho[1,3] BEATRIZ.GODINHO@NDCN.OX.AC.UK
Chongyu (Xiao) Qin[4] CHONGYU.QIN.20@UCL.AC.UK
Tim Behrens[5] TIMOTHY.BEHRENS@NDCN.OX.AC.UK
Thomas Akam[1] THOMAS.AKAM@PSY.OX.AC.UK

[1]University of Oxford
[2]Nuffield Department of Clinical Neuroscience
[3]Nuffield Department of Clinical Neurosciences
[4]University College London
[5]University of Oxford & University College London

Neuroscience has made substantial progress understanding the principles that determine how brains represent the external world, often appealing to notions of efficiency – representations that reduce statistical redundancy in sensory input, and utility – representations that support useful computations. However, the representations brains use to represent the rich structure of our own behaviour remain less well understood, and it is unclear whether, or how, these principles apply. Here we present data which suggests that medial frontal cortex (mFC) contains an efficient code for the sequential structure of goal directed behaviours, which may support hierarchically organised action selection. We trained mice to navigate to visually cued goal locations in complex elevated mazes, and recorded neurons in mFC using silicon probes. We observed two prominent patterns in mFC activity: First, a population code for path-distance to goal in which different neurons responded at different distances, such that the population tiled the navigation trajectory. Second, place-direction tuning in which neurons responded when subjects traversed particular sections of a maze in a given direction. This often took the form of tuning to extended route segments, or motion towards bottle-neck locations, reminiscent of 'options' in hierarchical reinforcement learning (RL). The low-dimensional structure of this place-direction tuning across neurons spanned the same low-dimensional space as that of subject's behavioural trajectories across trials - suggesting they may comprise an efficient code for behavioural sequences. These perspectives may be complementary, as work in machine learning suggests that sequence compression can discover useful options for hierarchical RL.

2-016. The computational geometry of flexible decision-making in prefrontal cortex

Kenji Lee[1,2] KENJILEE@BU.EDU
Tian Wang[1] TWANG@BU.EDU
Nicole Carr[1] NCARR@BU.EDU
Pierre Boucher[1] PBOUCHE1@BU.EDU
Chandramouli Chandrasekaran[1,3] CCHANDR1@BU.EDU

[1]Boston University
[2]Psychological and Brain Sciences

[3]Anatomy & Neurobiology

In many decision-making scenarios, sensory evidence is available before or after possible actions are known and need to be flexibly combined to achieve behavioral goals. We investigated the geometry of neural dynamics behind flexible decision-making by recording from the dorsolateral prefrontal cortex (DLPFC) of a monkey performing two decision-making tasks and compared the data to dynamics of a low-rank recurrent neural network (LR-RNN) model trained on the same tasks. In both tasks the goal was the same: discriminate a red-green checkerboard and reach to and touch the corresponding-colored target. In the Targets First task (TF), targets appeared before the checkerboard and in the Checkerboard First with Delay task (CFD), the checkerboard is presented first and then the targets appear. With Neuropixels probes, we recorded from the same neurons in DLPFC across both tasks and assessed if dynamics are similar or different between the two tasks to arbitrate between the following hypotheses. In the "universal computation" hypothesis [1], an overlapping population encodes both targets and sensory evidence with near-identical neural trajectories for both tasks. In the "non-overlapping" hypothesis, population trajectories are sensitive to order of inputs and evolve along distinct axes to a common choice-related output. Both DLPFC data and RNN were more consistent with the "non-overlapping" hypothesis. Single neurons rarely encoded more than one task variable within tasks and those that did encode targets or sensory evidence, rarely did so consistently between tasks. Neural population trajectories were task dependent and evolved within near-orthogonal axes after the first stimulus and finally converged to a common "action choice" representation just before movement initiation. Trajectories also depended on the order in which targets and sensory stimuli were presented. Together, our results suggest that dynamics in DLPFC adapt to task demands and timing of decision-related inputs to mediate flexible decision-making.

2-017. Groups of Neurons Form Synaptic Sequences on Short Stetches of Dendrite that Repeat Across Cortex

Saarthak Sarup[1,2] SSARUP@STANFORD.EDU
Nick Riedman[1] NRIEDMAN@STANFORD.EDU
Naijing Guo[1] NJGUO@STANFORD.EDU
Kwabena Boahen[1] BOAHEN@STANFORD.EDU

[1]Stanford University
[2]Department of Electrical Engineering

It has been hypothesized that the spatiotemporal pattern of spikes from an assembly of neurons encodes behaviorally-relevant information important for tasks like visual discrimination [1,2] and spatial navigation [3], and that this pattern is decoded by a 10-micrometer stretch of dendrite [4]. An assembly's axons synapse on a stretch of dendrite in a spatial order that matches the temporal order of its spikes [5,6], and these sequentially-activated synapses could generate a 'plateau potential' across the stretch and thus detect a particular spatiotemporal pattern [7]. If another dendrite stretch uses this mechanism to detect the same spatiotemporal pattern, then the assembly's axons would synapse on that stretch in a similar spatial order, and this synaptic sequence would be more common than expected. We searched for repeated synaptic sequences among 125,000 neurons connected by 523 million synapses, reconstructed from about 1 cubic millimeter of mouse primary visual cortex (Cortical MM^3 dataset) [8]. After skeletonizing the dendritic trees of 56,000 excitatory pyramidal neurons (PNs) into approximately 1.1 million branches, we uncovered over 170,000 synaptic sequences of 2 or more excitatory synapses that repeated in the dendritic branches of other pyramidal neurons with a median inter-synapse distance of 5 micrometers. Synaptic sequences repeated significantly more frequently than in an axon-constrained shuffle that preserved the inter-synapse distance distribution (5% and 39% increase with Z-scores of 31 and 12 for length-2 and -3 subsequences, respectively). Thus our findings reveal an intricate micro-scale synaptic organization of assemblies of axons and stretches of dendrite in the cortex: a stretch coordinates both the combination and permutation of axons that synapse along it. This organization could represent distinct packets of information stored in the wiring pattern of short stretches of dendrite, reflecting a fundamental aspect of the cortical neural code.

2-018. Integration of visually tuned inputs weighted by dendritic organization in the mouse visual cortex.

Kyle Jenks[1] KRJENKS@MIT.EDU
Greggory Heller[1] GREGGH@MIT.EDU
Katya Tsimring[1] KTSIMRING@GMAIL.COM
Emma Odom[2,3] EMMAODOM@MIT.EDU
Kendyll Brunell[1] KBURNELL@MIT.EDU
Mriganka Sur[1] MSUR@MIT.EDU

[1]Massachusetts Institute of Technology
[2]Massachusetts Institute for Technology
[3]Brain and Cognitive Sciences

Neurons in the mouse visual cortex (V1) receive broadly tuned inputs at thousands of dendritic spines. Despite this, only a fraction of neurons respond to any given stimulus. It is unclear how variability in neuronal responses is influenced by the structural and functional properties of individual spines. Here, we address this by performing in-vivo two-photon imaging of more than 4,500 dendritic spines on L2/3 pyramidal neurons in the mouse binocular visual cortex during presentation of drifting gratings. With this dataset we compare the orientation tuning of spines to their respective soma. Visually responsive neurons have a higher fraction of responsive spines compared to unresponsive neurons. Additionally, responsive spines on responsive somas have higher orientation selectivity and are more similarly tuned to the soma than responsive spines on unresponsive somas. The results demonstrate that the quantity and quality of visual selectivity of spine inputs correlates with the visual selectivity of the soma. The structure and location of dendritic spines within the dendritic arbor are hypothesized to weight the influence of a given spine on the soma's response. To test how dendritic organization impacts the spine-input to soma-output relationship we developed a linear integration model which linearly sums recorded inputs from spines and generates an in-silico orientation tuning curve. We compared the model to the in-vivo visual tuning of the soma and found that the model generates a similar tuning curve for responsive somas, but not for unresponsive somas as expected. Next, we differently weight the inputs by biophysical features such as spine size or distance to identify those which are relevant to the neuron's input-output computation. Results from our study will not only enhance our understanding on how presynaptic inputs are transformed into an informative and reliable single neuron output but will also provide insight into building biologically constrained neural networks.

2-019. Uncertainty encoded in a recurrent neural network trained to predict visual input during navigation

Yeowon Kim[1,2] YEOWON724@GMAIL.COM
Yul Kang[3,4] YUL.HR.KANG@GMAIL.COM

[1]Sungkyunkwan University
[2]Global Biomedical Engineering
[3]KAIST
[4]Bio and Brain Engineering

Navigation requires localizing oneself in an environment. However, one's location ("latent state") is not directly observable, and must be inferred from noisy and partial information such as egocentric visual input and self-motion signals. Recent findings suggest that the uncertainty about the latent state from such noisy information must be considered for optimal localization, and that it is indeed considered by people during navigation and represented by place/grid fields in the brain. However, it is unclear how such probabilistic representation (i.e., representation of the uncertainty about the latent state) is acquired. Here we show that such representation naturally arises when a neural network is trained to predict the upcoming sensory input. We develop a variant of an autoencoder that receives noisy egocentric visual stimuli/self-motion signals of an agent navigating an environment, and recurrently updates its hidden state to predict the upcoming visual stimulus. Then we show its hidden state represents a handcrafted ideal observer's belief about its location given noisy sensory inputs. The representation matched not only the optimal estimate of the location but also the estimate's uncertainty. Also, the decoded uncertainty about the distance from the nearest wall correlated with the distance, matching that of the ideal observer and paralleling the finding about how human homing behavior depends more on closer landmarks, and how a place field's size correlates with the distance from the nearest wall. A control network trained to reproduce the visual stimulus of the current time step failed to represent the ideal observer's belief as reliably. Thus, our results suggest that learning to predict the upcoming noisy sensory input may be a potential mechanism for the learning of probabilistic representation in the brain, even in a natural task like spatial navigation where the relationship between the sensory input and the latent state is complex.

2-020. What is the relationship between neural manifolds and field models of neuronal activity?

Louis Pezon[1,2] LOUIS.PEZON@EPFL.CH
Valentin Schmutz[3] V.SCHMUTZ@UCL.AC.UK
Wulfram Gerstner[4] WULFRAM.GERSTNER@EPFL.CH

[1]EPFL
[2]Brain Mind Institute
[3]University College London
[4]ecole Polytechnique Federale de Lausanne

Experimentally, the high-dimensional activity of large brain networks can often be parameterised as a point on a low-dimensional neural manifold with latent variables, whose dynamics characterise computations in the brain. On the modelling side, Wilson-Cowan-type field models postulate a low-dimensional embedding of the neurons in some space that can represent either their spatial location on the cortical sheet, or the topology in an abstract feature space, where e.g. the similarity of receptive fields implies “neighbourhood”. Here we ask whether, and if so how, this embedding space is related to neural manifolds.

We propose a framework that unifies models of recurrent neural networks (RNNs) associated with each perspective: on one side low-rank networks, that yield low-dimensional neural manifolds with latent variables obeying collective dynamics; and on the other side field models, defined over a space in which neurons are embedded. We find that the properties (dimension and topology) of this embedding space determine neither the neural manifold nor the collective dynamics, and vice versa. For example, a contextual decision-making task can be equivalently implemented in a rank-2 network by a field model over one- or two-dimensional space. However, we show how the topology of the neurons’ embedding can be extracted from the neuronal activity, so as to distinguish from one another models that yield similar neural manifolds.

In summary, we argue that there is no one-to-one relationship between the collective dynamics in neural manifolds characterising the computations performed by an RNN, and the structure of its neuronal population, characterised by the neurons’ embedding. Importantly, we suggest a simple method for probing this relationship using large-scale recordings.

2-021. Generalizability under sensor failure: tokenization + transformers enable more robust latent spaces

Geeling Chau[1] GCHAU@CALTECH.EDU
Yujin An[1,2] YAN2@CALTECH.EDU
Ahamed Raffey Iqbal[1] RAFFEY@CALTECH.EDU
Soon-Jo Chung[1] SJCHUNG@CALTECH.EDU
Yisong Yue[1] YYUE@CALTECH.EDU
Sabera Talukder[3] SABERA@CALTECH.EDU

[1]California Institute of Technology
[2]Bioengineering
[3]Caltech

One of the loftiest goals in neuroscience is to discover generalizable latent representations. This goal is challenged by differences between subjects (e.g. varying underlying neural structures), sessions (e.g. attention level), and sensors (e.g. sensor noise). Recent work addresses generalization across sessions and subjects [1, 2, 3, 4], but few study robustness to sensor failure [3] which is highly prevalent in neuroscience experiments. We study generalizability across three dimensions: subjects, sessions, and sensors by leveraging our own human electroencephalography (EEG) recordings across two subjects (A & B) and two sessions (1 & 2) per subject. To contextualize generalizability, we define: Within Session (wSes) - training and testing on the same session [e.g. A1 train, A1 test]; Cross Session (xSes) - training on one session, testing on another [e.g. A1 train, A2 test]; Cross Subject (xSub) - training on one subject, testing on another [e.g. A1 train, B2 test]; Sensor Failure - randomly failing X% of sensors, where X is contained in 0,10,20, ...,100. On the modeling front, we focus on TOTEM [4], the most recent prior work which leverages tokenization + transformers, and EEGNet [1], the most used prior work which leverages a convolutional neural network (CNN) upon which many other models [2] are built. Our contributions are as follows: we (1) collect four EEG datasets that enable studying generalizability (2) demonstrate that TOTEM outperforms or matches EEGNet across wSes, xSes, xSub & sensor failure, (3) explore TOTEM’s latent spaces to probe why tokenization enables more generalizable representations.

2-022. Temporal scaling in behavioral and dopaminergic learning

Annie Taylor[1] ANNIE.TAYLOR@UCSF.EDU
Huijeong Jeong[1,2] HUIJEONG.JEONG@UCSF.EDU
Vijay Mohan K Namboodiri[1] VIJAYMOHAN.KNAMBOODIRI@UCSF.EDU
Dennis Burke[1] DENNIS.BURKE@UCSF.EDU
Brenda Wu[1] BRENDA.WU@UCSF.EDU
Seul Ah Lee[3] SEULAHLEE@BERKELEY.EDU
Joey Floeder[1] JOEY.FLOEDER@UCSF.EDU

[1]University of California, San Francisco
[2]Neurology
[3]University of California, Berkeley

Animals can learn associations from experiences of paired events (e.g., cue-reward associations; each experience is commonly called a "trial"). In contemporary models of associative learning, such as temporal difference reinforcement learning (TDRL) models, learning progresses sequentially across trials. Thus, these models predict that the rate of learning is primarily driven by the number of experienced trials. In contrast, we show that reward frequency, and not number of trials, is the primary factor controlling the rate of associative learning in real animals. Specifically, when head-fixed mice received ten times fewer cue-reward experiences over the same total time as other mice, a single experience produced as much learning as ten experiences in the group with more frequent experiences. In other words, removing nine out of ten trials had no effect on overall learning, since the remaining trial fully compensated for the nine missing trials. A similar quantitative effect was found when removing five out of ten, or nineteen out of twenty trials, thereby showing a linear quantitative scaling of learning rate across a wide range of reward frequencies. This quantitative scaling also holds for mesolimbic dopaminergic learning, with the increase in learning rate being so high that the group with fewer experiences exhibits dopaminergic learning in as few as four cue-reward experiences and behavioral learning in nine. An algorithm implementing reward-triggered retrospective learning explains these findings. The temporal scaling and few-shot learning observed here fundamentally changes our understanding of the neural algorithms of associative learning.

2-023. Dynamics of orientation tuning during perceptual learning in the mouse visual hierarchy

Sarah Armstrong[1,2] SARMSTG1@GMAIL.COM
Rodrigo Carrasco Davis[3] RODRIGO.CD.20@UCL.AC.UK
Adam Packer[1] ADAM.PACKER@GMAIL.COM
Andrew Saxe[4] A.SAXE@UCL.AC.UK

[1]University of Oxford
[2]Department of Experimental Psychology
[3]University College London
[4]Gatsby Computational Neuroscience Unit & Sainsbury Wellcome Centre, UCL

Learning is fundamental to diverse behaviours. Ultimately learning-induced improvements in behaviour must trace back to changes in neural responses across distributed brain regions, yet the principles underlying learning remain unclear. A variety of theoretical ideas have been proposed, ranging from gradient descent–the workhorse algorithm of deep learning–to predictive coding and several Hebbian proposals. However, these theories have been challenging to test empirically, because their predictions often require measuring neural activity in many brain regions over long term learning. Here we present such a dataset on neural responses from four visual cortical areas as 7 mice learn a visual perceptual learning task over the course of several months. The areas recorded are thought to capture two functional hierarchies corresponding to dorsal and ventral 'streams' of visual processing (Tong et al., 2023, bioRXiv). Mice performed a multi-stage orientation discrimination task, and by the end of training detected differences of $\pm$20 degrees and under. To make clear theoretical predictions for this task, we extended a framework for considering a space of learning rules, including Gradient Descent (SGD), Hebbian (H) and Anti-Hebbian (AH) learning, Predictive Coding (QPC), and Contrastive Hebbian learning (CH) (Cao et al. 2020, NeurIPS), to operate within a hierarchy of orientation-tuned layers. We identify several key predictions of these models, including whether lower or higher layers change more, and whether the most informative neurons change tuning the most, characterised by their change in slope through learning, among other features. To test these predictions, we measured orientation tuning curves at high resolution, as well as spatial receptive fields, in areas V1, LM, LI and AL at up to six temporal points during learning, comprising recordings from ~100k excitatory neurons in total. We hope this dataset will provide an important benchmark for evaluating candidate learning algorithms.

2-024. A tensor decomposition uncovers centro-frontal EEG disturbances during reward learning in alcohol dependence.

Mica Komarnyckyj[1] MICA.KOMARNYCKYJ@MANCHESTER.AC.UK
Anna Murpphy[2] A.MURPHY2@HUD.AC.UK
Chris Retzler[2] C.RETZLER@HUD.AC.UK
Ioannis Delis[3] I.DELIS@LEEDS.AC.UK
Elsa Fouragnan[4] ELSA.FOURAGNAN@PLYMOUTH.AC.UK

[1]University of Manchester
[2]University of Huddersfield
[3]University of Leeds
[4]University of Plymouth

EEG has been widely used to study neuropsychiatric disorders, however robust biomarkers which link symptoms to underlying neurobiology are lacking. The common event-related potential (ERP) method relies on a priori selection of time windows and electrodes, meaning most data collected is overlooked, and by averaging data over space and time, important variability in the data is neglected. To overcome these limitations, we applied a tensor (non-negative canonical polyadic) decomposition (Cong et al., 2015), to a multidimensional EEG reward learning dataset (including 26 healthy-controls and 20 alcohol-dependents). This unsupervised machine learning approach, agnostic to a priori criteria, uncovered spatiotemporal components of reward-learning in alcohol dependence.

The first component, explaining most variance in the data, had centro-frontal distribution and was differentially recruited between the groups (alcohol vs control). Component recruitment was enhanced in alcohol-dependents compared to healthy-controls (p < 0.001) and positively correlated with alcohol units consumed (r = 0.39, p = 0.008). This component peaked within time windows previously highlighted as potential biomarkers for alcohol-dependency (250 ms = feedback-related negativity, 316 ms = P3)(Jurado-Barba et al., 2020). Classification demonstrated that component recruitment was predictive of group (area under ROC = 0.74 and classifier accuracy = 76%). Comparison ERP analyses found group differences at a priori feedback-related negativity (180 – 260 ms) (p = 0.020) but not P3 (280 - 380 ms).

Tensor decomposition demonstrated enhanced sensitivity to group differences within the P3 time-window compared with ERP analyses, and differences in component recruitment were related to clinical severity. Considering the groups had comparable performance on this complex reward-learning task, increased centro-frontal activity in alcohol-dependence may indicate a compensatory mechanism and greater attentional allocation (Xie et al., 2023). Tensor decomposition is recommended as a data-driven whole-brain approach for exploring differences in neural mechanisms between populations and uncovering biomarkers which may otherwise go unnoticed.

2-025. Abrupt transitions interrupt slow, ongoing representational drift in experiment and model

Jens-Bastian Eppler[1] EPPLER@FIAS.UNI-FRANKFURT.DE
Simon Rumpel[2] SIRUMPEL@UNI-MAINZ.DE
Matthias Kaschube[1] KASCHUBE@FIAS.UNI-FRANKFURT.DE

[1]Frankfurt Institute for Advanced Studies
[2]Johannes Gutenberg University Mainz

Both synaptic connections and neuronal representations within mouse cortex have been found to drift. However, the relationship between these two processes remains elusive. Experimentally, addressing this interplay between synaptic and representational drift proves challenging due to vast differences in temporal and spatial scales at which these processes occur. Here, we integrated representational dynamics and synaptic drift into a firing rate model to explore the influence of synaptic drift on representational drift. In mouse auditory cortex, neuronal populations have been shown to respond to stimuli in a clustered manner, with different stimuli evoking highly similar local response patterns (Aschauer et al., 2022). We constructed a random firing rate model comprising inhibitory and excitatory neurons and showed that this model effectively reproduced experimentally observed representational dynamics, including the clustering of stimulus responses, in an inhibition-dominated regime. Subsequently, we introduced synaptic drift into our networks within this specific regime, implementing it as a random process, gradually altering synaptic connectivity, according to experimentally observed synaptic drift (Loewenstein et al. 2011). This gradual, ongoing synaptic drift led to periods of nearly stable network responses, occasionally punctuated by abrupt transitions toward different activity patterns. The resulting bimodal distribution hints at two distinct processes of response changes, also present in experimental data. We explored these continuous and abrupt response transitions within our network model by examining the numerically identified stable and unstable fixed points of the system, using a method adapted from Sussilo and Barak (2013). We found two specific fixed point

changes responsible for these two types of response transitions. In conclusion, our study demonstrates that in a generic network model, random synaptic changes produce both gradual and abrupt response changes, similar to experimental findings. In our network model these qualitatively different response changes manifest through specific alterations of the network's fixed points.

2-026. Relating linear response and spontaneous input-spikes-output-spikes cross-correlations

Jakob Stubenrauch[1,2] JAKOB.STUBENRAUCH@RWTH-AACHEN.DE
Benjamin Lindner[1] BENJAMIN.LINDNER@PHYSIK.HU-BERLIN.DE

[1]Humboldt Universitat zu Berlin
[2]Physics Department

In the absence of an external stimulus, neurons in networks often remain active and fluctuate spontaneously. When a stimulus is presented, they respond to it, which enables information processing in neural networks. The spontaneous fluctuations and the response properties are not independent: subtle links, corresponding to the famous fluctuation-dissipation theorem in theoretical physics, connect them, as has recently been demonstrated [1]. Such relations constrain the signal-to-noise ratio, they can be used to fit models, and to advance theories. However, for the biologically relevant case of shot-noise driven neurons, such relations have not been reported yet. We consider an arbitrary spiking neuron model driven by a rate-modulated Poisson process with rate λ(t) = λ0 + ϵs(t). We derive an exact relation between the input-spikes-output-spikes cross-correlation in the spontaneous state (ϵ = 0) and the response function for a weak time-dependent modulation of the input firing rate ($\epsilon > 0$). This can be regarded as a variant of the famous Furutsu-Novikov theorem (FNT) [2, 3] for the case of shot noise. As we demonstrate, we can use the new FNT to obtain a fluctuation-response-relation (FRR) between the spontaneous fluctuations of a neuron's output and its systematic response to a time-dependent stimulus, extending the approach of [1] from Gaussian noise to shot noise. The relations are numerically tested and their limitation to Poissonian input exemplified for the important example of a leaky integrate-and-fire neuron with alpha synapses. However, the new FNT is valid for arbitrary neuron models, and including transient dynamics. [1]: B. Lindner, Phys. Rev. Lett. (2022) [2]: K. Furutsu, Res. Natl. Bur. Stand. (1963) [3]: E. A. Novikov, J. Exp. Theor. Phys. (1965)

2-027. Early life stress reduces cognitive flexibility through population-specific alterations to ventral hippocampal circuits.

Gabrielle Gregoriou[1,2] G.GREGORIOU@UCL.AC.UK
Karyna Mishchanchuk[1] KARYNA.MISHCHANCHUK.17@UCL.AC.UK
Andrew MacAskill[1,3] A.MACASKILL@UCL.AC.UK

[1]University College London
[2]Neuroscience, Physiology and Pharmacology (NPP)
[3]Neuroscience, Physiology and Pharmacology

In humans and animals, experiencing stress during early life has long-lasting effects on the neural systems that control goal-directed flexible behaviour. These changes are persistent and are thought to drive cognitive inflexibility in adults, a hallmark symptom of several mental illnesses, including addiction, depression, obsessive compulsive disorder and schizophrenia. Neural circuits in the ventral hippocampus (vH) are heavily impacted by early life stress, however, their role in flexible behaviour remains unclear, partially due to the fact that the vH is comprised of several heterogenous populations of projection neurons. Using a combination of computational modelling with circuit analysis, the current study investigates the role of individual vH projection populations in cognitive flexibility, and the mechanisms by which early life stress compromises this control. Using retrograde tracing and patch-clamp electrophysiology in mouse brain slices, we characterise three major populations of vH neurons and show that early life stress alters the vH in a population-specific manner, reducing the activity of vH neurons that project to the prefrontal cortex (PFC). Mimicking this effect in vivo by optogenetically silencing vH-PFC activity impairs flexible behaviour in a probabilistic reversal learning task. Specifically, our quantitative models reveal that vH-PFC inactivation impairs behaviour by reducing animals' ability to utilise memories of past outcomes to guide current choices, suggesting that the vH-PFC circuit contributes to cognitive flexibility by integrating past memories to infer task structure and select optimal actions. Intriguingly, mice that experience early life stress also exhibit deficits in flexible behaviour through similar strategy impairments, suggesting the vH-PFC may be a neural substrate for the negative impacts of early life stress on cognitive flexibility. Ongoing work aims to understand the mechanisms that underly population-specific changes at the molecular level, then use this understanding to rescue cognitive inflexibility in adults following early life stress.

2-028. Acetylcholine determines dopamine's role in learning versus moving

Heejae Jang[1,2] HJJ296@NYU.EDU
Carla Golden[1,3] CG163@NYU.EDU
Christine Constantinople[1,3] CMC472@NYU.EDU

[1]New York University
[2]Center for neural science
[3]Center for Neural Science

Midbrain dopamine neurons (DA) and their terminals in the striatum are implicated in reinforcement learning and motor control. A major outstanding question in the field is understanding how dopamine can satisfy these dual functions via the same circuit elements. Some attempts to reconcile dopamine's dual functions in learning versus moving have emphasized that different DA and recipient striatal regions are differentially engaged in learning and movement. A wealth of evidence suggests that DA release in the ventral striatum acts as a reward prediction error (RPE) to support cue-outcome association while that in the dorsolateral striatum produces gross effects on movement. The dorsomedial striatum (DMS) is located between the ventral and dorsolateral striatum and may represent an intermediate position along the continuum of reward to action coding in the striatum and its associated DA innervation. Here we reconcile these dual functions of DA using rats performing a value-based decision-making task that includes reward- and movement-related events at distinct time points. Fiber photometry measurement of DA release in the DMS revealed that only some event-aligned phasic DA signals predicted the vigor of upcoming contralateral movements. Phasic DA at other task events signaled RPEs that modulated rats' behavior and DMS firing rates on subsequent trials. These data demonstrate DA-dependent synaptic plasticity and behavioral change following RPE on a trial-by-trial basis. To explain the dual role of DA in supporting movement and learning at distinct time points within a single trial, we hypothesized that acetylcholine (ACh) may act as a gating mechanism1. Fiber photometry measurement of ACh release revealed dips at RPE events, but not movement events. These results show that ACh dynamically gates whether DA in the striatum is used for learning or moving on a moment-by-moment basis.

2-029. Adaptive algorithms for shaping behavior

William Tong[1,2] WTONG@G.HARVARD.EDU
Venkatesh Murthy[1] VNMURTHY@FAS.HARVARD.EDU
Gautam Reddy[3] GAUTAM.NALLAMALA@NTT-RESEARCH.COM

[1]Harvard University
[2]School of Engineering and Applied Sciences
[3]NTT Research

From Skinnerian pigeons trained as missile pilots, to canine companions trained to fetch, to present-day lab mice trained on a diversity of behavioral tasks, animal training remains a vital procedure for science, service – and at times – national security. By guiding an animal through a curriculum of simpler tasks that gradually approximate the desired outcome ("shaping"), an animal trainer can consistently teach highly nontrivial tasks to novice animals. Analogously, training artificial agents for many machine learning tasks is notoriously difficult, particularly in deep reinforcement learning (RL). Like animals, artificial agents frequently benefit from a graded curriculum that approximate the hardest task. However, behavior shaping remains a largely ad hoc and heuristic endeavor. To clarify the principles that underlie effective shaping, we present a curriculum learning framework that discovers approximately optimal algorithms for teaching complex behaviors. Specifically, we analyze behavior shaping as a teacher-student interaction. A teacher agent proposes calibrated tasks based only on extrinsic observations of the student's performance, optimizing the time required for success at the hardest task. Starting from a simple sequence learning task, we discover that curricular effects are driven by competing tendencies to reinforce and extinguish learned behaviors based on task difficulty. We demonstrate that a near-optimal curriculum modulates the difficulty to achieve a careful balance between reinforcement and extinction. Based on this intuition, we derive an adaptive shaping procedure that achieves near-optimal performance on the sequence learning task. We validate these results for deep reinforcement learning agents in naturalistic navigation tasks that involve sparse, delayed rewards. Our work offers a robust curricular strategy for training artificial agents, while suggesting a starting point towards a general computational framework for shaping both animal and artificial behavior.

2-030. Fast and tight control of motor vigor by accumulated decision evidence

Alexandre Garcia-Duran[1,2] ALEXGADUCA@GMAIL.COM
Manuel Molano-Mazon[3] MANUELMOLANOMAZON@GMAIL.COM
Jordi Pastor-Ciurana[4] PASTORC.JORDI@GMAIL.COM
Lluis Hernandez-Navarro[4] LLUISHN@GMAIL.COM
Lejla Bektic[4] BEKTIC@GMAIL.COM
Debora Lombardo[4] DEBORA.LOMBARDO@OUTLOOK.COM
Jaime de la Rocha[4] DELAROCHA.JAIME@GMAIL.COM
Alexandre Hyafil[1] ALEXANDRE.HYAFIL@GMAIL.COM

[1]Centre de Recerca Matematica
[2]Computational Neuroscience
[3]IDIBAPS
[4]August Pi i Sunyer Biomedical Research Institute

The link between the dynamics of decision formation and the kinematics of response movement remains poorly understood. Here we investigate how the accumulation of decision evidence shapes response movements in a task where freely-moving rats combine prior expectations and acoustic information to select between two possible options (Hermoso-Mendizabal et al. 2020). We show that the vigor with which rats execute their decision is greatly influenced by prior expectations and stimulus evidence. We extracted the rats' trajectories using an automatic video analysis (Mathis et al. 2018) and found that the vigor is initially set by the prior and incorporates the stimulus information after the response onset: rats accelerate if the stimulus supports their initial choice, and slow down otherwise. Detailed analysis of the trajectories shows that prior expectations influence the rats' trajectories already at the movement onset while the impact of stimulus evidence can be detected as early as 60 ms after stimulus onset. We further identified putative changes of mind, whereby rats reverse their initial, prior-guided choice when the stimulus accumulated evidence clearly contradicts their initial decision. Human subjects performing an equivalent task display a remarkably similar behavior to that of rats. We encapsulated this behavior in an extended drift-diffusion model that describes the mapping between the dynamics of evidence accumulation and the full response trajectory. The model, fitted with Mixed Neural Likelihood Estimate (Boelts et al. 2020), parsimoniously explains the subjects choices, the kinematics of response trajectories and the conditions yielding trajectory reversals. Together, our results show the tight and graded relationship between the evidence accumulated during perceptual decisions and the kinematics of the response trajectories described by rats and humans to execute their choices.

2-031. Why spikes? An axonal communication perspective

Micha Grutter[1] MICHA.GRUETTER@UNIBE.CH
Jonas Stapmanns[1,2] JONAS.STAPMANNS@UNIBE.CH
Jean-Pascal Pfister[1] JEANPASCAL.PFISTER@UNIBE.CH

[1]University of Bern
[2]Department of Physiology

The majority of neurons in the vertebrate nervous system communicate via spikes. These all-or-none events were discovered nearly 100 years ago by Lord Edgar Adrian. However, until today it remains unclear why - or rather when - neuronal information is preferably represented by spikes instead of analog signals. It has been proposed that event-based communication is preferable over analog-based communication over long distances, but a systematic and fair comparison is still lacking. Here, we address this question by systematically comparing event-based, i.e. digital, with analog communication. As a performance measurement, we use both the Shannon mutual information (in bits) as well as the information efficiency (bits/Joule) as a function of the axonal distance between the source signal position (soma) and the receiver position (at the synapse location). We derive the axonal noise properties from the cable equation and assume that the communication energy depends on the source energy as well as the channel energy - which scales linearly with distance in the digital channel case. Our results are threefolds. First, they confirm the intuition from the literature that for short distances the analog communication is preferable. This holds true up to some critical cable length, from where on the digital communication outperforms the analog communication. Second, we can compute analytically how the information and the information efficiency decay as a function of distance. Finally, we show that the critical cable length gets smaller in noisier environments while it increases for larger variance in the input signal.

2-032. Isolating single cycles of neural oscillations in spiking activity

Ehsan Sabri[1,2] EHSAN.SABRI@GMAIL.COM
Renata Batista-Brito[1,3] RENATA.BRITO@EINSTEINMED.EDU

[1]Albert Einstein College of Medicine
[2]Department of Neuroscience
[3]Neuroscience

Neural oscillations are prominent features of brain activity, observable as frequency-specific power changes in electroencephalograms (EEG) and local field potentials (LFP). Although identifying oscillations has primarily relied on EEG and LFP, the intrinsic relation between neural oscillations and neuronal spiking is noteworthy. We investigate the potential to detect individual cycles of neural rhythms solely through the spiking activity of neurons, using recent advances in densely recording large populations of neurons within a local network. Numerous spikes from neurons within a local network estimate the network's activity over time, where cycles from different time scales may combine in various ways to shape the network's spiking probability. We simulated the population spiking probability by synthesizing a signal using the known neural cycles as the basis functions. We also incorporated noise and variations in the width of each cycle instance to match the spectral profile of the recorded population spikings. We then used this synthesized signal to train a multi-layer Long Short Term Memory (LSTM) network for detecting the timing of the underlying cycles. We used this network to robustly isolate and align the neural cycles across different time scales in different brain regions of mice. Using isolated gamma cycles driven by sensory input, we achieve a more precise alignment of trials in sensory stimulation experiments in the primary visual cortex (V1) of mice. This alignment compensates for the biological variation in the transmission times of sensory signals from the retina to V1 across trials. As a result, we retrieve more accurate neural dynamics in response to sensory stimulation. Moreover, we applied this method to isolate cycles of ultra-slow rhythms spanning durations of up to hundreds of seconds in population spiking, pupil size, and facial motion. We observed that the delay in population spiking between brain regions varies with different time scales.

2-033. Choice and Deliberation in a Complex Planning Game in Monkeys

Jordan Lei[1] HL3976@NYU.EDU
Min-Yoon Park[2] MPARK61@JHU.EDU
Mariann Oemisch[2] M.OEMISCH@GMAIL.COM
Bas Van Opheusden[3] BASVANOPHEUSDEN@GMAIL.COM
Kristian Osborne[2] KRISTIAND.OSBORNE@JHU.EDU
Hexin Liang[2] HLIANG23@JHMI.EDU
Milan Ferguson[2] MFERGU35@JH.EDU
Daeyeol Lee[2] DAEYEOL@JHU.EDU
Wei Ji Ma[1] WEIJIMA@NYU.EDU

[1]New York University
[2]Johns Hopkins University
[3]Princeton University

Planning is a hallmark of complex cognition and decision making. Planning is difficult: in order to plan effectively, agents must develop abstract representations of their present and future states, consider the actions of others as well as their own, and do so while encountering states that they have never encountered before. Animal models are critical for understanding the cognitive and neural mechanisms of planning because they allow researchers to directly relate the activity of individual neurons to behavior. However, existing animal studies in planning and decision making operate on relatively simple tasks where subjects see the same states repeatedly and are not required to generalize to new states. To address this gap, we trained a monkey to play against a computer opponent on a board game called Four-in-a-row (4IAR). The game is a variant of tic-tac-toe with the goal of placing four pieces in a row. The game has over 10^16 board states and requires the player to think ahead in never-before-seen states; it has been used to study planning in humans. Behaviorally, we find that the monkey takes into account abstract task-relevant features and the relative weights of these features are well-ordered in the context of 4IAR. This is evidence that they are able to generalize the task demands. We also analyze the eye movements of the monkey during the task, demonstrating that our behavioral model, which is trained only to predict the next available move, is also able to predict the gaze fixations. Finally, we find that neurons in dorsomedial (DMPFC) and dorsolateral (DLPFC) prefrontal cortex as well as the caudate nucleus represent these features during game play. This work has broad implications for the role of PFC and striatum in complex planning and motivates the use of strategy games in understanding animal cognition.

2-034. Metabolic constraints on growth explain how developmental temperature scales synaptic connectivity relevant for behaviour.

Carlotta Martelli CMARTELL@UNI-MAINZ.DE
Pascal Zufle PZUEFLE@STUDENTS.UNI-MAINZ.DE
Leticia Batista BATISTA@UNI-MAINZ.DE
Sofia Brandao ADECASTR@UNI-MAINZ.DE
Giovanni D'Uva GDUVA@UNI-MAINZ.DE
Christian Daniel CDANIEL@STUDENTS.UNI-MAINZ.DE

Johannes Gutenberg University Mainz

The wiring of the nervous system follows a complex genetic plan during development. However, due to stochastic processes and environmental factors, genetically identical individuals seldomly show the same phenotypic outcome. Temperature is the environmental factor with the broadest effects in biology, as it determines the rates of all biophysical reactions of an organism. In poikilothermic animals, such as insects, worms, fish, amphibians, and reptiles, temperature determines the speed of development. However, whether a different developmental speed leads to the same wiring of the brain remains unclear. In Drosophila, temperature scales synaptic connectivity in the visual system, yet the underlying reasons for such scaling, the generality of this scaling across neural circuits, and the functional implications for behaviour are not understood. Here we combine anatomical, functional, and theoretical approaches to gain insights into the nature and consequences of temperature dependent synaptic scaling within the fly olfactory system. We show that synaptic scaling leads to heterogeneous functional effects in different olfactory subcircuits, with striking consequences for odor-driven behaviours. A first-principle model that imposes different metabolic constraints on the neural system and organism development explains these findings, and generalizes to predict brain wiring under ecologically relevant temperature cycles. Our data argue that metabolic constraints dictate the extent of synaptic scaling within neural subcircuit and that the resulting circuit architecture and function are contingent upon the availability of synaptic partners. This complex interplay between synaptic scaling and partner availability underscores the intricate impact of temperature-dependent developmental plasticity on the behaviour of poikilothermic animals.

2-035. A simple mathematical model unifies place field statistics across dimensionalities and species

Nischal Mainali[1,2] NISCHAL.MAINALI@MAIL.HUJI.AC.IL
Rava Azeredo da Silveira[3] RAVA@IOB.CH
Yoram Burak[1,4] YORAM.BURAK@ELSC.HUJI.AC.IL

[1]Hebrew University of Jerusalem
[2]ELSC
[3]ENS, Paris; CNRS; Institute of Molecular and Clinical Ophthalmology, Basel; University of Basel
[4]Edmond and Lily Safra Center for Brain Sciences, and Racah Institute of Physics

According to the classical view of spatial coding in the hippocampus, place cells express unimodal firing fields with a stereotyped, bell-shaped profile. Recent recordings from area CA1, however, reveal that this picture breaks down in large environments. Place cells in large environments typically fire in multiple locations. Furthermore, the multiple firing fields of individual cells, as well as those of the whole population, vary in shape and size, often deviating substantially from the classical bell-shaped form. Here, we report that a surprisingly simple mathematical model, in which firing fields are generated by thresholding a realization of a random Gaussian process, explains in quantitative detail a wide range of statistics of the observed place fields. The model simultaneously provides excellent fits to the distribution of field sizes, the distribution of inter-field distances, and the number of fields per cell, in several data sets that differ in the species and the dimensionality of the environment: from bats and rodents, in 1d, 2d, and 3d enclosures. In addition, the model makes testable quantitative predictions on the statistics of field shapes – the distribution of the number of local maxima within a field and the joint distribution of a field's width and its peak firing rate – as well as on topological properties of the fields. These predictions are all borne out when checked against experimental data, without any refitting of the parameters inferred from fitting field sizes. The description of place fields in terms of threshold crossings of a Gaussian process is phenomenological, yet it is suggestive of a mechanistic interpretation. Gaussian processes naturally arise when many statistically independent spatial fields are summed. Thus, the model's successful explanation of the place field statistics is consistent with a picture in which synaptic weights associated with projections into CA1 from its input regions are predominantly random.

2-036. Two opposing forces in inhibitory spike-timing-dependent plasticity differentially regulate network connectivity

Dylan Festa[1,2] DYLAN.FESTA@TUM.DE
Claudia Cusseddu[1] CLAUDIA.CUSSEDDU@TUM.DE
Julijana Gjorgjieva[1,2] GJORGJIEVA@TUM.DE

[1]Technical University of Munich
[2]School of Life Sciences

Inhibitory neurons (I), akin to their excitatory (E) counterparts, adjust synaptic efficacies in response to neural activity. Contrary to the traditional view of inhibition as a purely global activity modulator, recent studies highlighted the importance of specificity in (E/I) wiring and motif structures. These motifs, encompassing mutual and lateral inhibition, take essential functional roles in tuning of receptive fields, learning of representations, predictive coding, and in the regulation of the degrees of freedom of the network's dynamics (i.e. the network dimensionality).

Despite numerous characterizations of inhibitory plasticity rules, integrating them into a cohesive understanding of inhibition's various aspects — circuit stabilization, motif formation, and specific motif selection — remains challenging. Here we tackle this question, revealing how rules based solely on local interactions and spike-timing can effectively mediate both excitatory stabilization and structure formation. We distinguish the effects of spike-timing dependent plasticity (STDP) into two components: one dependent on firing rates, which tunes inhibitory synapses broadly and hinders motif formation, and another exclusively regulated by higher order statistics, which instead fosters sparser, more structured connectivity.

Selecting STDP rules that are either rate- or covariance- dominated, we compare the effect of the two regimes in simple circuits. Rate-dominated rules show low sensitivity to structure. By contrast, covariance-dominated regimes promote either mutual E/I connections, or later inhibition, contingent on the shape of the pairwise STDP interaction. Applying our theory to large recurrent circuit simulations, we further reveal how covariance-dominated rules can create sparse connectivity with specific E/I motifs, and stabilize overall circuit dynamics, thus bridging the gap between local synaptic plasticity and network-level organization.

2-037. Scalable gaussian process inference of neural responses to movies

Matthew Chalk[1,2] MATTHEWJCHALK@GMAIL.COM
Simone Azeglio[3,4] SIMONE.AZEGLIO@GMAIL.COM
Thomas Buffet[5] THOMAS.BUFFET@INSERM.FR
Matias Goldin[1] MATIGOLDIN@GMAIL.COM

[1]Sorbonne Universite
[2]Institut de la Vision
[3]Sorbonne University & Ecole Normale Superieure
[4]Vision Institute & Laboratoire des Systemes Perceptifs
[5]Sorbonne Universite, Institut de la Vision

Predicting the responses of sensory neurons to natural stimuli is a long-standing goal. While state-of-the art deep neural networks can now perform well at this in certain cases, their performance can be substantially degraded when trained on insufficient data, e.g. when recording time is limited. Moreover, constructing these models requires multiple structural choices (e.g. the network architecture, non-linearities, hyperparameters etc.), and it may be unclear a priori how these will affect their performance. On the other hand, gaussian processes (GPs) require few assumptions and perform well with limited data, but are typically poor at predicting responses to high-dimensional stimuli, such as natural images or movies. Recently however, it was shown how incorporating structured priors, e.g. for local and smooth receptive fields (RFs), can be used to scale up GPs to predict retinal neurons responses to static images. But it was unclear whether this approach could be used to predict neural responses to dynamic stimuli and movies, which have substantially higher dimensionality. Here we show that, by incorporating a recently proposed 'temporal relevance determination' (TRD) prior, which imposes a variable degree of smoothness as a function of time-lag, GPs can outperform a state-of-the art convolutional neural network (CNN) in predicting retinal responses to movies. Performance improvements were particularly striking when both models were trained on short recordings less than 30 minutes. This is important: it is often only possible to obtain short recordings with a given stimulus type during one experiment. The GP had the additional advantage of outputting the uncertainty in its predictions (which could be used e.g. in optimal stimulus design), and requiring relatively few prior assumptions about network architecture. Moreover, it could be trained quickly, based on the responses of single neurons, while the CNN required multiple neurons to train the middle layers.

2-038. Dendritic computation for context-dependent flexible decision-making

Yuan Zhang[1] YZH@ION.AC.CN
Yanhe Liu[1] YHLIU@ION.AC.CN
Lele Cui[1] LLCUI@ION.AC.CN
Yachuang Hu[1] YCHU@ION.AC.CN
Qi Liu[1] LIUQI@ION.AC.CN
Lin Zhong[2] ZHONGL@JANELIA.HHMI.ORG
Yu Xin[1] YXIN@ION.AC.CN
Ruiming Chai[1] RMCHAI@ION.AC.CN
Li Deng[1] DENGLI@ION.AC.CN
Jingwei Pan[1] JWPAN@ION.AC.CN
Ninglong Xu[3,4] XUNL@ION.AC.CN

[1]Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences
[2]HHMI Janelia Research Campus
[3]Chinese Academy of Sciences
[4]Institute of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology

Nonlinear dendritic integration in cortical pyramidal neurons implements powerful computations; however, its role in complex behavior and higher cognition is unknown. Here, we investigate dendritic computations in rule-switching flexible decision-making behavior in a projection-defined type of cortical layer 5 (L5) neurons. Using two-photon imaging and optogenetics we recorded and manipulated subcellular dendritic and somatic activity during task performance. We found that the L5 pyramidal tract-like (PTl) neurons in mouse auditory cortex exhibit compartmentalized representations for key behavioral variables, including sensory, rule, choice and prediction errors, in distinct dendritic and somatic subcellular domains. The distal dendritic tuft preferentially encodes task rule/context information in highly localized dendritic regions, whereas the somatic activity encodes comparatively greater amount of sensory information. Choice information is globally represented in both dendrites and soma, and is subject to task-context modulation. Simultaneous optogenetic inhibition of dendritic activity and somatic imaging show that dendrite-encoded rule information is required for contextual modulation of choice-related activity in the soma. Our results reveal an essential role of single-neuron dendritic computation in cognitive processes during context-dependent flexible decision-making.

2-039. Spontaneous activity generation and its role in refinement of developing neural circuits

Shreya Lakhera[1,2] SHREYA.LAKHERA@GMAIL.COM
Zhuoshi Liu[3] Z.S.L.LIU@ERASMUSMC.NL
Jan Kirchner[4] JAN.KIRCHNER@BRAIN.MPG.DE
Julijana Gjorgjieva[1,2] GJORGJIEVA@TUM.DE

[1]Technical University of Munich
[2]School of Life Sciences
[3]Dept. of Neuroscience, Erasmus University Medical Center, Rotterdam, Netherlands
[4]Max Planck Institute for Brain Research, Frankfurt

Developing neural circuits generate spontaneous activity in the absence of sensory input. This activity guides the refinement of synaptic connections; however, how the activity evolves during development and interacts with the connectivity it instructs remains unclear. During the first two postnatal weeks in mice, two patterns of spontaneous activity are observed in the visual cortex: retina-dependent low synchronicity (L) events and internally-generated high synchronicity (H) events (Siegel et al., 2012). Building a spatially organized model of the early visual cortex, we found that spatiotemporally correlated input can generate spatially local, inhibition-dominated L-events and intrinsic activity of the circuit can generate global, excitation-dominated H-events. We further showed that a combination of several factors drives the developmental sparsification of cortical spontaneous activity observed experimentally (Rochefort et al., 2009): increasing recurrent inhibition, increasing feedforward drive and decreasing intrinsic activity. This suggests that multiple processes underlie an important developmental change preparing the circuit for visual processing. Since neural activity and connectivity evolve simultaneously and interact in the developing circuit, with our model we could investigate how spontaneous activity reciprocally influences the refinement of feedforward connections through a Hebbian synaptic plasticity rule. We found that synaptic plasticity acting on L-events topographically refines and strengthens the feedforward synaptic connections to excitatory neurons in the cortex, hence providing a key drive for sparsification. Further, we found that the refinement and strengthening of feedforward excitatory connections and increasing inhibitory connections interact non-trivially in the model; increasing feedforward inhibitory connection can suppress the refinement of connections, but increasing recurrent inhibitory connections has no effect. Our work sheds light on the symbiotic relationship between

spontaneous neural activity patterns and connectivity in a recurrent cortical circuit highlighting the multiple axes of developmental changes.

2-040. Neural and behavioral signatures of online learning in probabilistic models

Jeroen Olieslagers[1,2] JO2229@NYU.EDU
Camille Rullan Buxo[1] CER454@NYU.EDU
Cristina Savin[1,3] CS5360@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Center for Neural Science, Center for Data Science

The statistics of the world are constantly in flux, so the brain must be able to adapt its internal model in real time to make sense of this evolving environment. This problem is further complicated by the fact that sensory information is itself dynamic and ambiguous, so neural circuits have to disambiguate meaningful changes from noise through a dynamic inference process. Despite a rich literature on probabilistic brain representations, we know very little about the signatures of learning internal models from moment-by-moment experience. Here we propose a framework for online learning in probabilistic models that can infer model parameters from the stochastic, continuous-time dynamical systems underlying inference over time. Our sampling-based approach allows for unified learning solutions for different internal models. The learning process is temporally local and does not require an evaluation of gradients, both computational prerequisites for biological learning. Furthermore, the phenomenological effects of such learning can be mapped into neural responses and behavior via a spiking recurrent network encoding process allowing us to directly compare its properties to experimental data. As a test case, we demonstrate the utility of this approach in the context of smooth pursuit eye movements, a task known to involve spatio-temporal integration of motion signals. Our model can account for the effects of adaptation to different speeds of motion at both the behavioral and neural levels, in particular observations that firing rates and eye speeds change during learning in a manner that depends on the magnitude of the sensory noise. More generally, this work brings probabilistic neural models into a more naturalistic setting, where both inference and learning happen concurrently.

2-041. Semantic extraction via systems consolidation

Albert Albesa Gonzalez A.ALBESA-GONZALEZ22@IMPERIAL.AC.UK
Claudia Clopath C.CLOPATH@IMPERIAL.AC.UK

Imperial College London

The theory of Complementary Learning Systems (CLS) [1] suggests the existence of two memory systems (a fast learner and a slow learner) as a solution to the plasticity-stability dilemma. The fast learner -usually associated with the hippocampus (HPC)- quickly forms new memories from sensory experience and then replays them to the slow learner. During this replay, the slow learner -attributed in biology to the Prefrontal Cortex (PFC)- integrates these memories via a process known as systems consolidation. Existing computational models of systems consolidation typically focus on the transfer of single, non-overlapping memories (episodes) (e.g. [2, 3, 4, 5, 6]), and don't investigate the extraction of reusable schemas. However, when they do so, these schemas are typically isolated [7] and don't include the relationships between them, which would lead to a semantic web. Similarly, they many times obviate [5, 7] the role of PFC recurrent synapses [8] in systems consolidation. In this work, we investigate a potential mechanism of semantic extraction via systems consolidation using standard modelling approaches in CLS: encoding and retrieval of episodes in HPC with Hebbian plasticity and attractor dynamics, providing PFC with a slower learning rate than HPC, and transfer of memories imprinted in HPC via replay to PFC during sleep. A key component of the model is the competition between outgoing weights, implemented with multiplicative renormalization. Under this paradigm, recurrent connections of PFC: (1) form maximally interconnected assemblies that represent latent elements in the generation of the input, and (2) establish graded and directional synapses between engrams that represent the posterior probabilities between the associated latents. While we reproduce phenomena like pair association [9, 10], memory extinction [9], or schema reusing [10], our model conceptually challenges how and why these occur, and makes testable experimental predictions based on the notion of semantic extraction.

2-042. To head fix or not to head fix: head-fixation alters neural circuit dynamics during immobility

Alex Pak[1,2] AP7454@NYU.EDU
Simon Carrillo Segura[1,2] SC8633@NYU.EDU
Janna Aarse[3] JANNA.AARSE@RUHR-UNI-BOCHUM.DE
Heng Wei Zhu[1] HZ3791@NYU.EDU
Jean-Paul Noel[1] JPN5@NYU.EDU
Andre Fenton[1] AF95@NYU.EDU
Dora Angelaki[1] DA93@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Ruhr University Bochum

Head-fixation in rodent models has become a prominent method for studying neural circuits, yet its impact on neural dynamics remains unclear. Here, we specifically explored the influence of head-fixation on the dynamics of the mouse thalamic head direction system, revealing impaired population coordination during stationary (immobile) periods. Using a 2D virtual reality (VR) setup, which enables mice to rotate their heads in the horizontal plane, and chronic Neuropixel recordings of head direction cells (HDC) in anterior thalamus, we discovered that stationary periods in head-fixed animals represent a distinct brain state. This was evident through an altered pairwise correlation structure, off-ring manifold activity, and lack of modulation of theta frequency by speed. To isolate contribution of head-fixation from VR, we developed a custom head-fixation apparatus that allows mice to freely explore an open-field arena while maintaining an immobile head relative to their body. Interestingly, despite the mice being exposed to all sensory cues, the HDC exhibited similar alterations during stationary periods as observed in VR . Crucially, we show that in both conditions, the fundamental property of HDC, firing exclusively at their preferred direction, is compromised when the head is restrained. During periods of immobility, HDC with preferred directions close to the current orientation reduce their activity, whereas other non-preferred HDC increase their firing, leading to off-ring manifold activity. Simulations of ring-attractor models indicate that disruption of recurrent connectivity of the HDC attractor, resulting in non-preferred head direction firing, leads to off-ring manifold dynamics, as observed in experimental data. Collectively, these findings suggest that stationary periods in head-fixed mice signify a distinct brain state of HD system, potentially due to the lack of gravity cues. Given the extensive use of head-fixation preparation in systems neuroscience, future studies should consider its potential impact on the specific system under investigation.

2-043. Recurrent Neural Networks Controlled With Maximal Input Entropy Perform Reliably Tasks

Chiara Mastrogiuseppe[1,2] CHIARA.MASTROGIUSEPPE@UPF.EDU
Ruben Moreno-Bote[1] RUBEN.MORENO@UPF.EDU

[1]Universitat Pompeu Fabra - Campus de la Ciutadella
[2]Department of Information and Communication Technology

Natural behaviors, even very stereotyped ones, exhibit variability. The basis of this variability is still not well understood, although spontaneous activity fluctuations in cortical neurons may contribute to it. Variability at both the behavioral and neural levels can facilitate exploration of state space and improve learning. Given the coupling between neural activity and behavior, an important question is what kind and extent of neural variability does not compromise behavioral performance. While previous studies have curtailed variability to enhance performance in neural networks, our approach takes the reversed perspective. We investigate the possibility of generating maximal neural variability while preserving the network's functionality. To address the problem, we build on the maximum occupancy principle (MOP) developed for behavior, which maximizes the cumulative sum of future action-state entropy, and thus it promotes variability in a non-greedy way. We consider a random recurrent neural network (RRNN) of fixed weights interacting with a random input current controller. This setup parallels models employed in songbirds, where specialized circuits serve the role of injecting variability into motor areas. Specifically, our controller is designed to: (i) maximize cumulative future input entropy while (ii) not compromising the network performance in a task. First, we show that large variability can be induced in the RRNN's spontaneous activity with both tanh or threshold-linear transfer functions while either avoiding terminal states of high – saturating or diverging – activities or by satisfying a maximum energy constraint. Second, we show that the input controller can drive the behavior of a chaotic RRNN to perform a context-dependent writing task. Remarkably, our network flexibly switches between stochastic and deterministic modes as needed. These results contribute to a novel theory of neural variability based on future entropy production, reconciling stochastic and deterministic behaviors within a single framework.

2-044. Theory of Mind Computations in Large Language Models Parallel to Single Neurons in the Human Brain

Jing Cai[1,2] JCAI5@MGH.HARVARD.EDU
Mohsen Jamali[3] MJAMALI@MGH.HARVARD.EDU
Ziv Williams[3] ZWILLIAMS@MGH.HARVARD.EDU

[1]Massachusetts General Hospital
[2]Neurosurgery
[3]Massachusetts General Hospital, Harvard Medical School

Theory of Mind (ToM) is among the most complex aspects of human social cognition, requiring us not only to form detailed inferences about the hidden thoughts, beliefs and perspectives of others but to also understand that another's beliefs may be false and distinct from one's own. Human ToM capability is supported by multiple interconnected brain areas in the frontotemporal cortices, including single neurons in the dmPFC that directly changed their firing rates in responding to the perspective of others. However, it remains unknown whether this neuronal organization is unique to humans, and whether the human neuronal activities utilize similar ToM coding to artificial networks, which experience no social interactions similar to ours. In this work, we utilized the recent development of Large Language Models (LLMs) and employed a novel 'parallel task' methodology to investigate the process by which LLMs form representations of others and how these representations may compare to those of single neurons in the human brain. Using this approach, our analysis reveals that hidden embeddings in the LLMs emerged to encode others' beliefs even when tested across widely varying naturalistic scenarios. Further, we show that trial-by-trial variations in the LLMs' embeddings are closely aligned to those of human neuronal activities in participants performing the same ToM tasks and that their ToM capacity can be explained, in part, by sparse single units (artificial embeddings or human neurons) tuned to reflect others' beliefs. Finally, we showed that the organization of the ToM-selective embeddings resembles that of a human brain, suggesting a convergence between these two distinct networks. Together, our findings offer initial evidence for a parallel between the activities of neurons in the human brain and the artificial models underlying ToM capability, which provides an alternative perspective on the neural computation of human social cognition.

2-045. Cholinergic control of cortical circuits for reinforcement learning

Quentin Chevy[1] QUENTIN.CHEVY@GMAIL.COM
Loreen Hertag[2] LOREEN.HERTAEG@TU-BERLIN.DE
Rozsa Balazs[3] ROZSABAL@KUKI.HU
Zoltan Szadai[3] ZSZADAI@FEMTONICS.EU
Rui Ponte Costa[4] RUI.COSTA@DPAG.OX.AC.UK
Adam Kepecs[1] AKEPECS@WUSTL.EDU

[1]Washington University St. Louis
[2]Berlin Institute of Technology
[3]Institute of Experimental Medicine, Hungary
[4]University of Oxford

Reinforcement learning algorithms provide a powerful account of how animals and artificial agents modify their behaviour in response to the consequences of their actions. While neuronal correlates of such algorithms are commonly observed in subcortical areas, how reinforcement signals can drive learning in the cortex remains largely unknown. Here, we show that cortical VIP interneurons are recipients of a global and modality-independent reinforcement prediction signal conveyed by cholinergic inputs. Our results suggest a tight interaction between VIP and neighbouring SST interneurons throughout reward-based learning. Using a computational model we show that this AcetylCholine-VIP-SST triad implements all the required ingredients for a reinforcement learning rule in the local circuit. Overall, our results suggest a novel circuit-based framework for the implementation of reinforcement learning algorithms in the cortex.

2-046. Learning mechanism of octopus vertical lobe and its comparison to the mushroom body and cerebellum

Naoki Hiratani[1] HIRATANI@WUSTL.EDU
Flavie Bidel[2] FLAVIE.BIDEL@MAIL.HUJI.AC.IL
Yaron Meirovitch[3] YARON.MR@GMAIL.COM

[1]Washington University in St Louis
[2]The Hebrew University

[3]Harvard University

To understand the principles of learning and computation in the brain, it is crucial to compare the nervous systems of diverse species. Octopus vulgaris is a compelling model system due to its remarkable intelligence, offering a distinct perspective compared to more traditional laboratory animals. The vertical lobe (VL), the learning center of the octopus, has a three-layered feedforward architecture similar to the insect mushroom body and mammalian cerebellum. However, recent electrophysiological and connectomics studies have revealed several important differences with these systems. Most notably, unlike the mushroom body and cerebellum, synaptic plasticity is mainly expressed in the input-to-middle layer in the VL circuit rather than the middle-to-output layer connections. However, its learning algorithm and functional benefits remain elusive. Here, we propose that VL implements the committee machine solution for the XOR problem, in contrast to the nonlinear expansion solution arguably implemented in the mushroom body and cerebellum. We show that the committee machine solution has a significantly larger classification capacity compared to nonlinear expansion when the output layer width is small, but this advantage disappears under a large output width. Correspondingly, the middle layer neurons of VL project to only a small number of output neurons, while middle layer neurons in the cerebellum and mushroom body project to hundreds and dozen of output neurons, respectively. Moreover, in the VL implementation of the committee machine, the capacity-per-connections is maximized at 1-3 synaptic connections, as opposed to the nonlinear expansion model where 3-5 synapses are optimal. This is consistent with recent data that simple amacrine cells in the VL, the main middle-layer neurons, receive only single projection from the input layer. In conclusion, we provide a unified explanation of two distinctive learning mechanisms implemented in similar neural circuits of different species.

2-047. The feature landscape of visual cortex

Rudi Tong[1] RUDI.TONG@MCGILL.CA
Arna Ghosh[2,3] ARNA.GHOSH@MAIL.MCGILL.CA
Erica Cianfarano[4] ERICA.CIANFARANO@MAIL.MCGILL.CA
Blake Richards[4] BLAKE.RICHARDS@MILA.QUEBEC
Ronan da Silva[1] RONAN.DASILVA@MAIL.MCGILL.CA
Dongyan Lin[5] DONGYAN.LIN@MAIL.MCGILL.CA
James Wilsenach[6] VILSNK@GMAIL.COM
Pouya Bashivan[5] POUYA.BASHIVAN@MCGILL.CA
Stuart Trenholm[1] STUART.TRENHOLM@MCGILL.CA

[1]Montreal Neurological Institute, McGill University
[2]McGill University, Mila - Quebec AI Institute
[3]School of Computer Science
[4]McGill University
[5]Mila, McGill University
[6]The Alan Turing Institute

Understanding computations in the visual system requires a characterization of the distinct feature preferences of neurons in different visual cortical areas. However, we know little about how feature preferences of neurons within a given area relate to that area's role within the global organization of visual cortex. To address this, we recorded from thousands of neurons across six visual cortical areas in mouse and leveraged generative AI methods combined with closed-loop neuronal recordings to identify each neuron's visual feature preference. First, we discovered that the mouse's visual system is globally organized to encode features in a manner invariant to the types of image transformations induced by self-motion. Second, we found differences in the visual feature preferences of each area and that these differences generalized across animals. Finally, we observed that a given area's collection of preferred stimuli ('own-stimuli') drive neurons from the same area more effectively through their dynamic range compared to preferred stimuli from other areas ('other-stimuli'). As a result, feature preferences of neurons within an area are organized to maximally encode differences among own-stimuli while remaining insensitive to differences among other-stimuli. These results reveal how visual areas work together to efficiently encode information about the external world.

2-048. Differential phase shifting via simultaneous excitatory and inhibitory connections in a head direction circuit

Dan Turner-Evans DTURNERE@UCSC.EDU
Kerstin Richter KERICHTE@UCSC.EDU
Jorin Eddy JAEDDY@UCSC.EDU
Ali Shenasa AHSHENAS@UCSC.EDU

University of California, Santa Cruz

Single neurons can make both excitatory and inhibitory connections if the composition of receptors varies at the postsynaptic site of these connections. Yet how circuit-level computations might be enabled by simultaneous excitatory and inhibitory signaling is not broadly understood. For example, phase-based computations could be carried out via heterogeneous signaling pathways. In phase-based computing, the phase of a sinusoidally varying signal carries information. Phase-based computing is ubiquitous in neural circuits for navigation. Head direction cells use phase to encode the animal's head direction in their population activity, and phase sets the properties of grid cells and vector object cells. Different compositions of receptors can alter the phase of a signal as it is passed from one population of neurons to the next. For example, inhibitory signaling could invert a sinusoid, producing a 180° phase shifted version of the original signal. We found strong evidence that a single class of neurons in the head direction (HD) circuit in the fruit fly, Drosophila melanogaster, form excitatory connections to some downstream partners and inhibitory connections to others and that these different types of connections differentially shift the phase of the HD signal as it is passed downstream.

2-049. TiDHy: Timescale Demixing via Hypernetworks to learn simultaneous dynamics from partial observations

Elliott Abe[1,2] EABE@UW.EDU
Bing Brunton[3] BBRUNTON@UW.EDU

[1]University of Washington
[2]Biology
[3]University of Washington, Seattle

Natural behavior and neural dynamics are often characterized by multiple latent dynamical systems that are multiscale and nonstationary, yet they can only be partially observed, making these measurements challenging to model. To understand these datasets, a common approach is to define time points as discrete, mutually exclusive states (e.g., via switching linear dynamical systems (SLDS) or clustering). However, this set of assumptions does not represent real neural and behavioral data, so there is a critical need for a computational method that can demix simultaneously evolving systems. For instance, social behaviors involve interdependent, interacting animals with behaviors that span numerous timescales simultaneously. Motivated by this need, we develop timescale demixing via hypernetworks (TiDHy) to learn from partially observed data how linear combinations of latent dynamics may evolve. TiDHy can be considered a generalization of the SLDS framework and is computed with hypernetworks, a type of neural network architecture that generates the weights of another network. A key feature of TiDHy is that it directly learns the intrinsic timescales of the data (rather than as a hyperparameter) and demixes independent systems, yielding interpretable dynamics. Using synthetic data, we show that TiDHy can demix multiple SLDSs from partially superimposed observations, even when they span orders-of-magnitude different timescales. Additionally, we demonstrate the application of TiDHy on a real-world multi-animal tracking dataset by creating an interpretable latent structure of each animal from key-point observations and recapitulating hand-annotated behavior labels.

2-050. Neural Representations of face recognition in biological and artificial systems: Insights from MEG and CNNs

Hamza Abdelhedi[1,2] HAMZA.ABDELHEDI@UMONTREAL.CA
Shahab Bakhtiari[1] SHAHAB.BAKHTIARI@UMONTREAL.CA
Karim Jerbi[1] KARIM.JERBI@UMONTREAL.CA

[1]Universite de Montreal
[2]Computer Science

Artificial neural networks, inspired by brain structure and function, have surpassed human performance in various tasks, but the link between Artificial Intelligence and neuroscience is still underexplored. Combining these fields

has offered mutual reinforcement, especially in the field of Neuro-AI, where comparing artificial and biological systems in cognitive tasks, such as visual categorization, has yielded promising insights (Kubilius et al., 2019; Yamins et al., 2014). Face recognition, however, is less explored in this context. Do Convolutional Neural Networks (CNNs) trained for face recognition mimic neural dynamics of face recognition in brain circuits? A question addressed only by a handful of studies, which in non-human primate mainly focus on the IT cortex, and in humans, largely rely on fMRI or behavioral data (Chang et al., 2021; Jiahui et al., 2023; Kathrina Dobes et al., 2023). Here we compare human brain activity collected using Magnetoencephalography (MEG) during a face recognition task to activations across seven CNNs trained on the same task. Compared to previous work, we leverage the high temporal resolution of MEG and source reconstruction techniques to compare these models to the brain across time, frequency, and space. Out of the tested models, FaceNet emerged as the most brain-like model during face recognition. Crucially, training on face recognition, rather than on object recognition or both simultaneously, was necessary and sufficient for high model-brain similarity. In terms of temporal alignment, peak similarities were observed around 170ms which corresponds to the M170-component linked with face perception. Examining the Fusiform Face Area (FFA), we observed that, compared to an untrained model, the similarity to FaceNet trained on face recognition significantly increased, from 0.02 to 0.08 in certain FFA regions. Our study provides novel insights into the spatio-temporal similarity patterns between artificial and biological neural responses associated with face recognition.

2-051. Explicit neural codes for subjective beliefs in orbitofrontal cortex support state inference

Shannon Schiereck[1] SS11853@NYU.EDU
Andrew Mah[1,2] AM9056@NYU.EDU
Margaret DeMaegd[1] MLD9131@NYU.EDU
Christine Constantinople[1,2] CMC472@NYU.EDU

[1]New York University
[2]Center for Neural Science

Cognitive functions such as decision-making and planning require a mental representation of the environment and the subject's location within the state space of that environment. When the environment is only partially observable, animals must use inference to determine their current location within a state representation of the environment. Previous studies have shown that lateral orbitofrontal cortex (lOFC) is crucial for learning about task contingencies when they are only partially observable, suggesting a role in state inference. A major hypothesis about the role of lOFC in state inference suggests that lOFC is required for determining the current location within a cognitive "map" of the task. However, it is still unclear how neurons in lOFC might implement state inference. We trained rats on a temporal wagering task with hidden transitions between reward states (blocks of trials with high, intermediate, or low average expected rewards). The amount of time rats were willing to wait provided an explicit behavioral readout of their subjective value. Rat's wait times were modulated by the volume of water offered on a given trial as well as the hidden reward state. Behavioral analysis and modeling using Bayes' rule to predict the identity of the current state showed that rats use knowledge of the task structure to infer the hidden reward state and adjust their wait times accordingly. Bilateral muscimol inactivation of lOFC reduced sensitivity to the hidden reward state by impairing belief updating about hidden states over trials. Neural recordings in lOFC and interpretable dimensionality reduction revealed that subpopulations of neurons encode posterior beliefs at specific time points during the delay period and immediately after the reward availability cue. These results elucidate the local computations performed in lOFC for state inference.

2-052. Convergent processing discriminates value information according to modality in the basal ganglia

Seong-Hwan Hwang[1,2] QJSRODIA@GMAIL.COM
Ji-Woo Lee[1] JWOO12@SNU.AC.KR
Hyoung F Kim[1] HFKIM@SNU.AC.KR

[1]Seoul National University
[2]Department of biological science

There has been a long-standing debate on how the basal ganglia system processes information, given its reception of a vast amount of inputs from various cortical structures with relatively fewer neurons. Although previous studies have proposed that this system employs convergence processing, there is still a lack of actual empirical evidence. We addressed this question by recording putamen neurons while monkeys performed tactile and visual value discrimination tasks, associating reward values with different sensory stimuli (visual fractals or braille

patterns). Our findings revealed three distinct types of value neurons in the putamen: tactile-selective, visual-selective, and bimodal value neurons. Remarkably, over half of the value neurons were identified as bimodal, indicating that the basal ganglia processes diverse value information in a convergent manner. Next, our analyses aimed to explore whether these bimodal value neurons either integrate or discriminate value information from distinct sensory inputs. Notably, the neural trajectories of bimodal neurons were entirely different in value representation based on the associated sensory modality. Further analysis employing optimal targeted dimensionality reduction (oTDR) revealed that both value and modality information were represented in each relevant task dimension, with transitions occurring between these distinct state dimensions over time. Through non-linear decoding analysis, we confirmed again the differentiated representation of value based on modality in bimodal neurons. These findings suggest that the basal ganglia processes diverse information in a convergent manner while retaining the unique characteristics of each input, enabling discrimination within their population responses.

2-053. Mesoscale modules for the control of working memory in primate lateral prefrontal cortex

Xuanyu Wang[1,2] WANGXUANYU0419@GMAIL.COM
Simon N. Jacob[3] SIMON.JACOB@TUM.DE
Daniel Hahnke[1] DANIEL.HAEHNKE@TUM.DE
Andreas Nieder[4] ANDREAS.NIEDER@UNI-TUEBINGEN.DE

[1]Technical University of Munich
[2]Department of Neurosurgery
[3]Klinikum rechts der Isar, TUM
[4]University of Tubingen

Working memory processing in primate lateral prefrontal cortex (lPFC) is characterized by complex spatiotemporal activity patterns. These patterns could arise from a modular functional organization, as has been described in sensory and motor cortices. However, most working memory theories do not incorporate this idea and instead assume that the prefrontal cortical sheet is a large, homogeneous recurrent neural network without spatial parcellation. Here, we sought experimental evidence for a modular organization of lPFC and asked if such modules were governed by working memory content or context (i.e., distinct cognitive operations involved in working memory processing). We obtained extracellular recordings of spiking activity and local field potentials (LFP) from the lPFC and ventral intraparietal area (VIP) of rhesus monkeys performing a delayed match-to-numerosity task. We analyzed the spatiotemporal pattern of LFP bursts in the gamma (60-90Hz) and beta (15-35Hz) bands over multiple recording days and correlated these patterns with neuronal spiking and cortico-cortical oscillatory synchrony. Our results demonstrate that the spatiotemporal dynamics of LFP bursts faithfully parcellate the primate lPFC into distinct functional modules, each supporting different stages of working memory coding. These modules receive temporally synchronized and spatially organized sensory inputs and exhibit distinct local (within PFC) and long-range (PFC-VIP) connectivity, suggesting they are anatomically defined. Our study provides empirical evidence that calls for the inclusion of modular structure in models of prefrontal cognitive operations and emphasizes the importance of considering mesoscale spatiotemporal dynamics in experimental investigations of the primate association cortex.

2-054. Classification of neural excitability types in human and mouse brain circuits

Paul Pfeiffer[1,2] PAUL.PFEIFFER@HU-BERLIN.DE
Robert Gowers[1] ROBERT.GOWERS@HU-BERLIN.DE
Jan-Hendrik Schleimer[1] JH.SCHLEIMER@HU-BERLIN.DE
Susanne Schreiber[1] S.SCHREIBER@HU-BERLIN.DE

[1]Humboldt Universitat zu Berlin
[2]Insitute for Theoretical Biology

Neuronal diversity is a hallmark of complex nervous systems, but its functional implications for computation are not well understood. Recently, it was shown that the dynamical type of the neurons' spike generation (also referred to as excitability type) impacts the network state (e.g., synchronous, asynchronous, or splayed-out). For regularly firing neurons, only four such types, defined by the underlying spike onset bifurcation, exist and have been studied in detail for some canonical neuron models. Yet, an experiment-based classification remains challenging and scarce, so the actual distribution of excitability types in biological networks is unknown.

Here, we profit from the experimental data in the Allen Brain Cell (ABC) and classify the excitability types of over 200 quantitatively precise neuron models that have been derived from intracellular recordings in mice and

humans.

For classification, neuron models are subjected to stimulation protocols specifically designed to detect spike-onset type fingerprints, including, for example, bistability between rest and firing. We find a prevalence of models with a monostable continuous (type I) f-I curve without subthreshold resonance (mouse: 147/180 and human 38/40 cells), corresponding to the saddle node on invariant cycle (SNIC) bifurcation. Yet, the model population encompasses all main spike onset types, including the often neglected yet computationally interesting homoclinic type.

Taken together, we present the first large-scale analysis of excitability types across mouse and human neurons, shedding light on the frequency of computationally different spiking mechanisms. Our work provides an empirical footing for incorporating neuronal heterogeneity at a functionally relevant level in simulations of brain networks and lays the ground for studying how neural circuits maintain a distribution of spike onset types suited for their functional task.

2-055. Vector quantized representations for hierarchical delineation of behavioral repertoires

Tianqing Li[1,2] TIANQING.LI@DUKE.EDU
Ugne Klibaite[3,4] KLIBAITE@FAS.HARVARD.EDU
Jumana Akoad[5] JAKOAD@FAS.HARVARD.EDU
Joshua Wu[1,6] JOSHUA.WU@DUKE.EDU
Timothy Dunn[1,6] TIMOTHY.DUNN@DUKE.EDU

[1]Duke University
[2]Department of Biomedical Engineering
[3]Harvard university
[4]Organismic and Evolutionary Biology
[5]Harvard University
[6]Biomedical Engineering

Understanding animal behaviors and their neural underpinnings requires precise kinematic measurements plus analytical methods to parse these continuous measurements into interpretable, organizational descriptions. Existing approaches, such as Markov models or clustering, can identify stereotyped behavioral motifs out of 2D or 3D keypoint-based data but are limited in their interpretability, computational efficiency, and/or ability to seamlessly integrate new measurements. Moreover, these methods lack the capacity for capturing the intrinsic hierarchy among identified behavioral motifs (e.g., 'turning' $\rightarrow$ subtypes of left/right turning with varying angles), necessitating subjective post hoc annotations by human labelers for grouping. In this paper, we propose an end-to-end generative behavioral analysis approach that dissects continuous body movements into sequences of discrete latent variables using multi-level vector quantization (VQ). The discrete latent space naturally defines an interpretable behavioral repertoire composed of hierarchically organized motifs, where the top-level latent codes govern coarse behavioral categories (e.g., rearing, locomotion) and the bottom-level codes control finer-scale kinematics features defining category subtypes (e.g., sidedness). Using 3D poses extracted from recordings of freely moving rodents (and humans), we show that the proposed framework faithfully supports standard behavioral analysis tasks while enabling new applications stemming from the discrete information bottleneck, including realistic synthesis of animal body movements and cross-species behavioral mapping.

2-056. How connectivity structure shapes rich and lazy learning in neural circuits

Yuhan Helena Liu[1,2] HYLIU24@UW.EDU
Aristide Baratin[3] ARISTIDEBARATIN@HOTMAIL.COM
Jonathan Cornford[4] CORNFORJ@MILA.QUEBEC
Stefan Mihalas[5] STEFANM@ALLENINSTITUTE.ORG
Eric Shea-Brown[1] ETSB@UW.EDU
Guillaume Lajoie[6,7] G.LAJOIE@UMONTREAL.CA

[1]University of Washington
[2]Applied Mathematics
[3]Samsung - SAIT AI Lab
[4]Mila - Quebec AI Institute and McGill University
[5]Allen Institute
[6]Universite de Montreal

[7]Mila

In theoretical neuroscience, it has been shown that structural variations can have a pronounced impact on learning dynamics in humans and animals. One particularly notable aspect of the structure is the initial connectivity before training takes place [1]. Notably, as shown in recent deep learning theory literature, initial weight distributions with small (resp. large) variance can lead to a rich (resp. lazy) regime; in this context, significant (resp. minor) changes to network states and representation are observed over the course of training [2-3]. However, these stand in stark contrast to biological neural circuits that typically display a pronounced low-rank eigenstructure. Given this discrepancy, we probe how the initial weight structure, especially its effective rank, impacts the learning regime, even when the initial weight magnitude remains the same. Through both empirical and theoretical analyses, we discover that high-rank initializations typically yield smaller network changes indicative of lazier learning, a finding we also confirm with experimentally-driven initial connectivity in recurrent neural networks (RNNs). Conversely, low-rank initialization biases learning towards richer learning. Importantly, however, as an exception to this rule, we find lazier learning can still occur with a low-rank initialization that aligns with task and data statistics. Our research highlights the pivotal role of initial weight structures in shaping learning regimes, with implications for metabolic costs of plasticity [4] and risks of catastrophic forgetting [5].

2-057. Brain-wide calcium imaging in zebrafish reveals cell-level functional network properties of seizure

Wei Qin[1,2] WEI.QIN@UNIMELB.EDU.AU
Jessica Beevis[3] J.BEEVIS@UQ.NET.AU
Ellen Hoffman[4] ELLEN.HOFFMAN@YALE.EDU
Andre Peterson[5] PETERSON@UNIMELB.EDU.AU
Ethan Scott[5] ETHAN.SCOTT@UNIMELB.EDU.AU

[1]the University of Melbourne
[2]Department of Anatomy and Physiology
[3]The University of Queensland
[4]Yale University
[5]The University of Melbourne

Epilepsy is a neurological disorder that causes recurrent seizures, but the exact mechanisms that trigger seizure transitions are still unclear. Traditional methods, using data from humans, nonhuman primates, or rodents, have limitations in resolving the activity of single cells that may initiate seizures. An approach that captures the dynamics of individual neurons and their interactions within brain-wide networks could therefore be of great utility in understanding epilepsy. Zebrafish and calcium imaging offer such an approach, as they allow for simultaneous in-vivo recording of neuronal activity across the brain at cellular resolution. Zebrafish share genetic and physiological similarities with humans and can exhibit seizure-like behaviours in response to various drugs. One such drug is Pentylenetetrazol (PTZ), a pharmacological agent that blocks inhibitory GABAergic signalling, causing hyperexcitability and seizure-like activity. Additionally, mutations in the scn1lab gene, encoding a sodium channel, can also cause spontaneous seizures in zebrafish. In this study, we used in-vivo light-sheet calcium imaging, brain-wide and at cellular resolution, on wildtype and scn1lab mutant zebrafish larvae under baseline and PTZ conditions. We apply network analyses, computational modelling, and graph theory to statistically quantify the differences in network topology and dynamics between the two genotypes and the two conditions. Specifically, we investigated the network of active neuronal cells involved in ictogenesis and seizure propagation from microscopic to macroscopic scales. Our study revealed significant and consistent changes in brain network connectivity and information transmission pathways, indicating that scn1lab mutations affect the functional structure of the brain. We also developed a generative model to explain the wiring principles that govern the development of both genotypes, and how the scn1lab mutation affects this brain-wide functional network. Combining experimental data and mathematical modelling, our approach provides a novel perspective on the mechanisms of epileptogenesis at a breadth and resolution that traditional epilepsy studies cannot achieve.

2-058. Geometry-aware inference in the olfactory bulb

Paul Masset[1] PAUL.MASSET@GMAIL.COM
Jacob Zavatone-Veth[2,3] JZAVATONEVETH@G.HARVARD.EDU
William Tong[2,4] WTONG@G.HARVARD.EDU
Joseph Zak[5] JDZAK@UIC.EDU
Venkatesh Murthy[2] VNMURTHY@FAS.HARVARD.EDU
Cengiz Pehlevan[2] CPEHLEVAN@SEAS.HARVARD.EDU

[1]McGill University
[2]Harvard University
[3]Physics
[4]School of Engineering and Applied Sciences
[5]University of Illinois Chicago

Within a single sniff, the mammalian olfactory system can decode the identity and concentration of odorants wafted in turbulent plumes of air. Yet, it must do so given access only to the noisy, dimensionally-reduced representation of the odor world provided by olfactory receptor neurons. As a result, the olfactory system must solve a compressed sensing problem, relying on the fact that only a handful of the millions of possible odorants are present in a given scene. Inspired by this principle, past works have proposed normative compressed sensing models for olfactory decoding. However, these models have not captured the unique anatomy and physiology of the olfactory bulb (OB), nor have they shown that sensing can be achieved within the 100-millisecond timescale of a single sniff. Here, we propose a rate-based Poisson compressed sensing circuit model for the OB. This model maps onto the neuron classes of the OB, and recapitulates salient features of their connectivity and physiology. For circuit sizes comparable to the human OB, we show that this model can accurately detect tens of odors within the timescale of a single sniff. We also show that this model can perform Bayesian posterior sampling for accurate uncertainty estimation. Fast inference is possible only if the geometry of the neural code is chosen to match receptor properties, yielding a distributed neural code that is not axis-aligned to individual odor identities. Importantly, we show that it is the geometry in the space defined by the receptor affinity (or OSN activation) that controls the speed of inference. This view is distinct from previous geometric theories of olfaction, which have focused on the space of odorants. We propose that thinking in terms of the geometry of OSN coding will allow for deeper understanding of early olfactory processing

2-059. Selective convergence of distinct inputs in the visual thalamus reinforces motion processing

Yue Fei[1,2] YUE.FEI@YALE.EDU
Liang Liang[1] LIANG.LIANG@YALE.EDU

[1]Yale University
[2]Neuroscience department

The primary visual thalamus, or the dorsolateral geniculate nucleus (dLGN) of the thalamus, conveys visual signals from the retina to the visual cortex. In addition to receiving diverse channels of inputs from retinal ganglion cells, the dLGN is also innervated by the midbrain superior colliculus1. The highly conserved colliculogeniculate axons possess several synaptic properties that resemble those of retinogeniculate axons, including innervating the proximal dendrites of dLGN neurons and providing strong excitatory synaptic inputs that can trigger neural firing in the target neurons1,2. However, it remains unclear how collicular inputs combine with retinal inputs to impact thalamic visual processing. Using chronic deep-brain dual-color two-photon calcium imaging, we simultaneously recorded visual responses from hundreds of collicular and retinal axonal boutons in the dLGN of awake, head-restrained mice. We observed that collicular axons also exhibited diverse visual response properties and their retinotopic organization closely aligned with that of retinal axons. On a scale of ~6 μm, nearby collicular boutons shared one or several feature preferences in common, following a similar fine-scale functional logic for spatial arrangements of retinal boutons. Notably, neighboring collicular and retinal boutons also displayed similar feature preferences, suggesting that the two distinct sources of excitatory inputs may work in synergy. Finally, inhibiting collicular inputs resulted in the reduction of response magnitudes in the majority of impacted dLGN neurons. In some dLGN neurons, silencing collicular inputs decreased responses to all directions of motion by a similar proportion; in others, it additionally suppressed responses in the preferred direction and reduced their direction selectivity. Specifically, a subset of dLGN neurons lost their preferences for motion along the temporal direction or the horizontal axis. These findings suggest that functional convergence between collicular and retinal inputs in the dLGN can already reinforce select channels of motion signals prior to the cortex.

2-060. Brain-wide Electrophysiological Atlas

Olivier Winter[1] OLIVIER.WINTER@INTERNATIONALBRAINLAB.ORG
Gaelle Chapuis[1] GAELLE.CHAPUIS@INTERNATIONALBRAINLAB.ORG
Han Yu[1] HAN.YU@INTERNATIONALBRAINLAB.ORG
Julien Boussard[2] JULIEN.BOUSSARD@INTERNATIONALBRAINLAB.ORG
Kcenia Bougrova[3] KCENIA.BOUGROVA@INTERNATIONALBRAINLAB.ORG
Yanliang Shi[4] YANLIANG.SHI@INTERNATIONALBRAINLAB.ORG
Renata Proa[2] RENATA.PROA@INTERNATIONALBRAINLAB.ORG
The International Brain Lab The International Brain Lab[1] INFO@INTERNATIONALBRAINLAB.ORG

[1]International Brain Laboratory
[2]Columbia University
[3]Champalimaud Research
[4]Cold Spring Harbor Laboratory

Past decades have seen incredible progress in mapping gene expression, cell types, and connectivity across the entirety of mammalian brains. Yet, no attempt has been made to systematically characterize and quantify basic electrophysiological measurements across the brain. We curated a large dataset comprising 1037 recordings tiling 194 regions of the mouse brain, to build a new brain-wide reference atlas of electrophysiological properties. We developed an architecture to compute electrophysiological features (e.g. local field (LF) power spectra, spike rate or shape) at such a large scale, and have improved the denoising of both LF and spike waveforms. We quantified which features are most informative about brain region identity, and developed a website that allows brain-wide exploration of these features. For instance, we see a strong dependence of spike shape on brain region, with wider spikes more prevalent in cortex and narrow spikes more prevalent in hypothalamus and midbrain. We built interpolation models to construct an atlas of full-brain coverage at high spatial resolution (voxel size 200 um). These models use anatomical and molecular information as priors, and interpolate the full map from sparse data samples. We find that spatial gene expression profiles capture a large fraction of variance of electrophysiological signature variance across the brain. Finally, we are establishing tools that enable automated localization of recording locations. We developed "decoder" algorithms that can predict a brain region label based on the electrophysiological signatures (accuracy 60%). This is the first step towards building on-line tools to be used in real-time experiments. This work embodies the first systematic characterization and quantification of electrophysiological features across the mouse brain. Our electrophysiology atlas will provide an important open resource for neuroscientists to both guide on-line electrophysiological recordings and their post hoc interpretation.

2-061. Natural statistics and stimulus representations in visual working memory

Ivan Tomic[1,2] IVN.TOMIC@GMAIL.COM
Zahara Girones[3] ZG282@CAM.AC.UK
Mate Lengyel[4] M.LENGYEL@ENG.CAM.AC.UK
Paul Bays[3] PMB20@CAM.AC.UK

[1]University of Zagreb
[2]Department of Psychology
[3]University of Cambridge
[4]University of Cambridge; Central European University

Visual working memory (VWM) refers to a limited capacity for the active retention of visual information over time. Between entering the visual system and being retrieved from VWM, memoranda undergo a series of representational transformations that leave an imprint on the recalled information. However, it is unknown at what stage of these transformations stimulus information is stored and maintained in VWM. Here, we leveraged Bayesian and efficient coding accounts of perceptual decision-making to predict how natural scene statistics at different stages of the visual processing pipeline should affect estimation bias and variability in a VWM task. We collected psychophysical data from an analogue report task testing memory for orientation, while varying the number of stored objects and the retention period. Consistent with previous research, increasing load and retention interval each increased variability of recall estimates, and a repulsive bias of estimates from the cardinals was apparent, in line with the over-representation of cardinal orientations in natural scenes. However, we saw no effects of set size or delay on the repulsive bias. This finding challenges previous theoretical accounts that predict a tight coupling between bias and variability. To provide a principled account of these observations, we devised competing ideal observer models. These models incorporated constraints from natural statistics in both encoding and decoding, but differed in the processing stages at which noise related to memory load and maintenance was introduced. Within this framework, the only model that could account for the empirical results was one in which these effects arise after the internal measurement is combined with the natural-statistics prior to form a posterior. These find-

ings suggest that natural statistics shape internal representations at an early processing stage that precedes the capacity and maintenance limitations associated with working memory.

2-062. The Cerebellum Shapes the Preparatory Dynamics of Motor Cortical Neurons in Force Field Adaptation

Hugo Ninou[1,2] HUGONINOU@GMAIL.COM
Sharon Israeli[3] SHARON.ISRAELI1@MAIL.HUJI.AC.IL
Lee Elmaleh[3] LEE.ELMALEH@MAIL.HUJI.AC.IL
Firas Mawase[4] MAWASEF@BM.TECHNION.AC.IL
Yifat Prut[3] YIFAT.PRUT@MAIL.HUJI.AC.IL
Jonathan Kadmon[3] JONATHAN.KADMON@MAIL.HUJI.AC.IL

[1]Ecole Normale Superieure
[2]Cognitive Science department
[3]The Hebrew University of Jerusalem
[4]Technion Israel Institute of Technology

In vertebrates, the cortico-cerebellar circuit plays a crucial role in motor adaptation. However, little is known about how the cortex and cerebellum interrelate during this process. In particular, it remains unclear whether the cerebellum only performs feedback control, by adjusting movement during execution, or whether it also shapes preparatory activity.

To address this question, we trained monkeys to perform a center-out reaching task and adapt to a velocity-dependent force field (FF) while reversibly blocking cerebellar output using high-frequency stimulation (HFS) in the superior cerebellar peduncle. The monkeys could still complete the task, although the cerebellar block impeded adaptation and increased motor noise. Electrophysiological recordings in the primary and premotor cortical areas showed that force field adaptation during the cerebellar block, but not during control trials, led to an angular shift in cortical preparatory activity in a manner akin to re-aiming. These results suggested that cortical activity attempts to compensate for the loss of cerebellar signals. This cortical adaptation, however, came at a computational expense. We found a significant increase in cortical preparatory activity dimensionality under HFS compared to the control condition. The increased dimensionality reduced the cross-condition generalization performances of a linear readout. Consistent with the neural findings, the monkeys' ability to generalize their adaptive response across targets diminished. Furthermore, neural simulations showed that blocking low-dimensional feedback resulted in increased dimensionality and reduced generalization performance.

These results point to a dual role for cerebellar signals in motor adaptation. First, cerebellar signals containing task-related information reduce the dimensionality of preparatory activity and improve generalization. Second, during movement execution, these signals translate preparatory cortical activity into adapted motor output. In the absence of cerebellar control, cortical preparatory activity provides insufficient compensation and a maladaptive behavioral response. Taken together, these results imply that the cerebellum participates both in movement control and movement planning.

2-063. Value-based Decision-making Relying on Uncertain Prior-level Information

Risa Katayama[1,2] KATAYAMA.RISA.44N@ST.KYOTO-U.AC.JP
Shin Ishii[1] ISHII@I.KYOTO-U.AC.JP

[1]Kyoto University
[2]Graduate school of Informatics

To make decisions efficiently in complex environments, our observations need to be supplemented with the prior-level information—environmental states behind the observations. Such environmental states are often hidden and involve some degree of uncertainty. Neuroimaging researches focused on the confidence in the last stage of the decision-making processes. Several human behavioural studies have examined the uncertainty for the higher-level states that contribute to decisions and their confidence in hierarchical perceptual tasks. However, how the brain processes uncertain prior-level information to control the cognitive behaviours has not been well elucidated. Here, we designed a "Twenty-one" game task, in which to make better hands whose total score is closer to but not exceeding twenty-one, participants may or may not take a card from the unknown deck-type that is of either of two types of different score distributions. Applying computational modelling to human task behaviours, we observed that the participants adapted their decisions and confidence evaluation depending on the inference of the environmental state, i.e., the deck-type, and the degree of uncertainty for it. Functional

magnetic resonance brain imaging revealed that the activity in the left insula was scaled with the confidence level of the deck-type specifically during the decision phase. In addition, the neural representation in this area was functionally coupled with the medial prefrontal cortex that was locally associated with the subjective value belief and the decision confidence. These findings suggested that the prefrontal-insula network supports the modulation of decision-making and related metacognition in terms of confidence for the prior-level information under this uncertain circumstance.

2-064. Acetylcholine integrates past reward to guide decision making under uncertainty

Ella Svahn[1,2] ELLA.SVAHN.20@UCL.AC.UK
Jessica Passlack[1] JESSICA.PASSLACK.19@UCL.AC.UK
Athena Akrami[3] ATHENA.AKRAMI@UCL.AC.UK
Andrew MacAskill[4,5] A.MACASKILL@UCL.AC.UK

[1]UCL
[2]NPP, SWC
[3]Sainsbury Wellcome Centre, UCL
[4]University College London
[5]Neuroscience, Physiology and Pharmacology

Making decisions in a noisy and ever-changing world is difficult. To make optimal choices, we must use past experience to continuously track outcomes and uncertainty. This tracking allows us to persevere with optimal choices despite inherent variability in the world; called 'expected uncertainty'. It also helps us detect when circumstances have changed, which should prompt us to flexibly update our choices. An optimal decision-making strategy requires a balance between sticking with learnt choices versus flexibly updating, which relies on the continuous integration of choice feedback. But how neural circuits support the integration of experience to guide uncertain decision making is poorly understood. Neuromodulators have been suggested to be involved in uncertainty estimation and flexible learning as they can both induce plasticity and signal over long time scales. Specifically, acetylcholine has been suggested to play a key role in uncertainty encoding, but it is unclear how this is achieved, particularly on the long timescales required to effectively track uncertainty. To test this, we trained mice in a probabilistic reversal learning task where mice had to integrate past outcomes over tens of seconds across different experimentally controlled uncertainties. We found that performance in this task was dramatically impaired by pharmacological inhibition of cholinergic signalling, through reducing the influence of outcomes history on upcoming choice. To compliment this finding, we used a genetically encoded acetylcholine sensor to record acetylcholine release into a key area for flexible learning; the ventral hippocampus. Consistent with our pharmacological findings, hippocampal acetylcholine levels around choice were scaled by outcome history. Together this suggests a key role for hippocampal acetylcholine in shaping flexible behaviour. Ongoing work is utilising reinforcement learning agents with different strategies to track uncertainty, to more quantitatively investigate the role of acetylcholine signalling in reward integration and uncertainty estimation.

2-065. Weight transport through spike timing for robust local gradients

Timo Gierlich[1,2] TIMO.GIERLICH@KIP.UNI-HEIDELBERG.DE
Andreas Baumbach[3] ANDREAS.BAUMBACH@KIP.UNI-HEIDELBERG.DE
Akos F. Kungl[3] KUNGL.F.AKOS@GMAIL.COM
Kevin Max[1,4] KEVIN.MAX@UNIBE.CH
Mihai Petrovici[1] MIHAI.PETROVICI@UNIBE.CH

[1]University of Bern
[2]Department of Physiology
[3]Heidelberg University & University of Bern
[4]Dept of Physiology

Both in machine learning and in computational neuroscience, plasticity in functional neural networks is frequently expressed as gradient descent on a cost. Often, this imposes symmetry constraints that are difficult to reconcile with local computation. For example, wake-sleep learning in networks characterized by Boltzmann distributions inherently builds on the assumption of symmetric connectivity. Similarly, the error backpropagation algorithm is notoriously plagued by the weight transport problem between the representation and the error stream. Existing solutions tend to circumvent the problem by deferring to the robustness of these algorithms to weight asymmetry. For example, feedback alignment simply observes that learning tends to approximately align forward weights to frozen, random backward weights, thus producing approximate gradients of sufficient quality. However, such

solutions are known to scale poorly with network size and depth and require additional mechanisms to improve their functionality. We suggest that synapses can use spike timing statistics to extract and correct the asymmetry between effective reciprocal connections. Apart from being quintessentially spike-based and fully local, our proposed mechanism takes advantage of a ubiquitous feature of physical neuronal networks: noise. Based on an interplay between Hebbian and anti-Hebbian plasticity, it allows synapses to recover the true local gradient. This implicitly also alleviates discrepancies that arise from neuron and synapse variability - an omnipresent property of physical neuronal networks, both biological and artificial. We demonstrate this generic mechanism in two different network models. In probabilistic spiking networks, we show how a combination of Hebbian and anti-Hebbian plasticity can significantly improve convergence to the target distribution as compared to Hebbian plasticity alone. And in neuronal hierarchies based on cortical microcircuits, we show how our proposed mechanism effectively enables the copying of forward weights to the feedback pathway, thus allowing the backpropagation of correct feedback errors.

2-066. Neurons Tuned to Chaotic States

Chanwoo Chun[1] CC2465@CORNELL.EDU
Sweta Agrawal[2,3] SWETA@VT.EDU
John Tuthill[4] TUTHILL@UW.EDU
Dmitri Chklovskii[5,6] MITYA@FLATIRONINSTITUTE.ORG

[1]Weill Cornell Medicine
[2]Virginia Tech University
[3]School of Neuroscience
[4]University of Washington
[5]Simons Foundation
[6]Flatiron Institute

In the predictive coding literature, the brain is believed to make predictions of the dynamically evolving environment. Because trajectories in the chaotic regions of the phase space are more difficult to predict, the brain is likely to have neurons signaling the approach of chaos. Here, we explore the hypothesis that the receptive fields of leg proprioceptive neurons are tuned to the chaotic regions of the walking cycle orbit characterized by a high maximal finite-time Lyapunov exponent (FTLE), the divergence rate of two neighboring trajectories. Specifically, we compare the dynamics of the femur-tibia joint angle during walking and the activity of the directionally selective proprioceptors. First, we reconstruct the walking limit cycle by time-delay embedding of the joint angle time series in a 4D state space. Second, we compute the maximal FTLE for each explored region of the state space and discover that the FTLE peaks at the end of the stance phase. Finally, we compare these predictions to in vivo calcium imaging of the activity of the leg proprioceptors and find that they are most active in the same region of the state space. Whereas it was known in other species that analogous proprioceptors are tuned to the end of the stance phase, the correlation with the peak FTLE, i.e. chaotic state, is a novel finding. We also show this correlation in artificial systems, by training a neural network that predicts a future state of a chaotic 2D dynamics, and show that the receptive fields of the neurons are tuned to the chaotic regions. As the dynamical system description can be applied to other neural systems as well, our finding may generalize beyond early proprioception and explain neuronal receptive fields in general.

2-067. Training networks of morphologically detailed biophysical neuron models with thousands of parameters

Michael Deistler[1,2] MICHAEL.DEISTLER@UNI-TUEBINGEN.DE
Pedro Goncalves[3] PEDRO.GONCALVES@NERF.BE
Jakob Macke[1] JAKOB.MACKE@UNI-TUEBINGEN.DE

[1]University of Tubingen
[2]Computer Science faculty
[3]VIB-Neuroelectronics Research Flanders, Belgium

Single neurons can have remarkable computational abilities such as nonlinear pattern separation [1]. However, optimizing the parameters of detailed biophysical neuron models is notoriously difficult, and tuning as few as 20 parameters can be a challenging task [2]. Therefore, many studies ignore the computational repertoire that is made possible by the biophysical and morphological complexity of single cells, and instead focus on studying networks of point-neurons and the computational abilities conferred by their synaptic interactions. Networks of point-neurons are computationally cheap and can be efficiently optimized with backpropagation of error (or surrogates thereof) even when they have millions of parameters. Can we also use backpropagation to train detailed,

multicompartment biophysical models, or even networks of such models?

Here, we present a toolbox for simulation and inference in neuroscience, 'Jaxley', which makes this possible: Jaxley provides efficient and numerically stable solutions of the stiff dynamics of networks of biophysically detailed multicompartment neurons. Jaxley can scale simulations to multiple GPUs and, unlike previous toolboxes, can optimize model parameters with backpropagation, thereby allowing to train networks of biophysically-detailed neurons to perform physiologically relevant tasks. We show that Jaxley can simultaneously optimize cell-level parameters (e.g., branch length, radius, and channel conductances) and system-level parameters (e.g., synaptic strengths), with virtually no limit on the number of parameters. We use Jaxley to demonstrate that (1) single neurons can learn to perform nonlinear computations, (2) single neurons can solve binary MNIST, and (3) networks of biophysically-detailed neurons with thousands of parameters can be trained with backpropagation. Jaxley is a flexible and easy-to-use toolbox which will help bridge systems neuroscience and biophysics, allowing new insights and opportunities for multi-scale neuroscience.

2-068. Memories by a thousand rules: Automated discovery of plasticity rules reveals variety and degeneracy at the heart of learning

Basile Confavreux[1] BASILE.CONFAVREUX@GMAIL.COM
Poornima Ramesh[2] POORNIMARAMESH1995@GMAIL.COM
Pedro Goncalves[3] PEDRO.GONCALVES@NERF.BE
Jakob Macke[4] JAKOB.MACKE@UNI-TUEBINGEN.DE
Tim Vogels[5,6] TIM.VOGELS@IST.AC.AT

[1]Institute of Science and Technology Austria
[2]University of Tuebingen
[3]VIB-Neuroelectronics Research Flanders, Belgium
[4]University of Tubingen
[5]IST, Austria
[6]_

Synaptic plasticity is a cornerstone of learning and memory. However, the mechanistic link between individual synaptic changes and emerging network functions remains elusive, due to the difficulty of recording synapses in vivo and because hand-tuning network models with coupled activity and weight dynamics is near-impossible. Here, we apply a numerical method for meta-learning plasticity rules—filter Simulation-Based Inference (fSBI)—to in vivo data. We identify thousands of co-active excitatory(E)-to-E, E-to-inhibitory(I), I-to-E and I-to-I rules that can drive learning and memory functions in large recurrent spiking networks.

We begin with large search spaces of flexibly-parameterized co-active plasticity rules. We then use fSBI to sequentially narrow down these sets of rules with an increasingly fine mesh of ad-hoc constraints: First we filter for rules that establish stable, cortical-like dynamics, i.e., asynchronous irregular activity with low rates. We noticed that many of these rules display additional qualities such as engram formation, recall and graceful forgetting, even though such functionality was not included in the filtering constraints. With such sets of rules in hand, we turn to published experimental data in which learning and engram formation have been observed, but for which the rules that form them are unknown. We find entire subspaces of plasticity rules that reproduce qualitative and quantitative aspects of the data, e.g., faithful population responses to familiar vs novel stimuli. We interrogate these plasticity manifolds with regard to memory lifetime, sensitivity encoding, and temporal dynamics of a learned response, unveiling the impact of different plasticity parameters.

Our results show that thousands of plasticity rules can establish neural function. They need not be too finely tuned nor "orchestrated" beyond constraints for basic function, but small changes within the space of feasible solutions will determine specific aspects of a memory.

2-069. Computation with program operations in replay

Sebastijan Veselic[1,2] SEBASTIJAN.VESELIC@GMAIL.COM
Timothy Muller[1] TIMOTHYMULLER127@GMAIL.COM
Nour Mohsen[3] NOUR.MOHSEN.22@UCL.AC.UK
Lennart Luettgau[3] L.LUETTGAU@UCL.AC.UK
Steve Kennerley[1] STEVEN.KENNERLEY@PSY.OX.AC.UK
Tim Behrens[4] TIMOTHY.BEHRENS@NDCN.OX.AC.UK
Zeb Kurth-Nelson[5] ZEBKURTHNELSON@GMAIL.COM

[1]University of Oxford
[2]Experimental Psychology
[3]University College London

[4]University of Oxford & University College London
[5]Google DeepMind

A series of surprising experimental discoveries about replay have led to the proposal that replay itself could be a general-purpose reasoning mechanism in the brain (Schwartenbeck et al., 2023; Liu et al., 2019; Kurth-Nelson et al., 2023). This theory makes two predictions. First, the results of reasoning should appear in the brain immediately after the replay that performs the reasoning. Second, replay should include information about the operations being performed; thus including rich syntax enabling objects to be composed into sequences. To test these predictions, we developed a magnetoencephalography (MEG) task where participants sequentially execute simple programs. These programs operate on a data structure – a cartoon face with three orthogonal attributes. In total, there are six program operations. Three require swapping attributes of the cartoon face. The other three require branching conditional on the cartoon face attribute. This conditioning defines the next program that needs to be executed. Program operations were represented as images of objects (e.g. an elephant or a key) and randomized across participants. Each operation was instantiated across two different objects to dissociate object from operation representations. During scanning, participants (N = 29) mentally executed these programs with high accuracy. On trials where they ultimately gave correct solutions, we observed neural replay of the correct program sequence, even though these program sequences were never executed in training. On incorrect trials, we observed replay of incorrect program sequences. When a program operation spontaneously activated during reasoning, the result of that operation followed in close temporal alignment, supporting the first prediction. Furthermore, objects and operation representations were replayed together, supporting the second prediction. Our results are aligned with the hypothesis that replay incorporates representations of rich syntax (i.e. operations) to compositionally assemble objects into meaningful structured compounds during reasoning.

2-070. The synaptic locus of song learning

Drew Schreiner[1,2] DREW.SCHREINER@DUKE.EDU
Samuel Brudner[3,4] SAMUEL.BRUDNER@YALE.EDU
John Pearson[5,2] JOHN.PEARSON@DUKE.EDU
Richard Mooney[5] MOONEY@NEURO.DUKE.EDU

[1]Duke University School of Medicine
[2]Neurobiology
[3]Yale University
[4]Quantitative Biology Institute
[5]Duke University

Juvenile songbirds spontaneously copy their tutor's acoustically complex song without external reinforcement. To solve this challenging reinforcement learning problem, the juvenile requires three ingredients: 1. Motor variation (what did I just do?). 2. Timing (when did I do it?). 3. Evaluation (how good was it?). Intriguingly, a song-specialized basal ganglia nucleus (the sBG) is believed to have access to all three of these ingredients. Nonetheless, the synaptic locus within the sBG that drives song learning is unknown. We combined computational analyses of song learning and closed-loop optogenetic manipulations of specific sBG synapses in singing birds to identify the synaptic locus of song learning.

One major challenge in identifying the synaptic locus of learning in the sBG is the high dimensionality of birdsong. We applied a variational autoencoder to compress high-dimensional vocal data into a low-dimensional latent space, then developed a feed forward neural network that predicted the age of production for every syllable rendition based on these latent dimensions. In this way, we identified learning-relevant dimensions in the vocal data. Another major challenge is identifying which sBG synapse is driving learning. Importantly, this process is greatly simplified because, unlike the mammalian striatum, the sBG receives input from only two premotor "cortical" nuclei, known as HVC (the "when" pathway) and LMAN (the "what" pathway). We used closed-loop optogenetic suppression to silence the sBG, HVC terminals in the sBG, or LMAN terminals in the sBG of juvenile finches as they engaged in song learning. Suppression of either the sBG or of HVC-sBG terminals immediately and transiently erased the prior ~2-6 hours of learning. Thus, plasticity at this premotor corticostriatal synapse is necessary to express recently learned changes to song on a rendition-to-rendition basis. Our work is of broad relevance to understanding how the basal ganglia help solve high-dimensional learning problems.

2-071. Improving optimal control in systems with biologically realistic multiplicative and internal noise

Francesco Damiani[1,2] FRANCESCO.DAMIANI@UPF.EDU
Gregory DeAngelis[3] GDEANGELIS@UR.ROCHESTER.EDU
Jan Drugowitsch[4] JAN_DRUGOWITSCH@HMS.HARVARD.EDU
Ruben Moreno-Bote[1] RUBEN.MORENO@UPF.EDU
Akiyuki Anzai[3] AANZAI@UR.ROCHESTER.EDU

[1]Universitat Pompeu Fabra - Campus de la Ciutadella
[2]Department of Information and Communications Technologies
[3]University of Rochester
[4]Harvard Medical School

To act in the world, we integrate sensory information as we move our sensors and body through the external environment, creating perception-action loops. Movements alter sensory inputs, which are then compared with predictions and used in planning future movements to accomplish internal objectives. However, these mechanisms are challenged by different noise-sources, coming from the integration of sensory feedback and the motor output itself. Furthermore, neural representations are subject to internal fluctuations, which affect estimation processes and, consequently, behaviour. Stochastic optimal control theory formalizes these concepts to explain behaviour through optimality principles at the algorithmic level. In this context, having an optimal solution is crucial for assessing the rationality of the observed behavior. Our work is then particularly relevant in the context of inverse optimal control. A control problem involves designing the optimal control law, or state-to-action policy, to minimize a cost-function of a system, determined by task goals and energetic costs. Exact solutions to the control problem can only be derived under linear dynamics, additive Gaussian noise, and a quadratic cost function, exploiting the independence between estimation and control. However, when considering a realistic noise-model of the sensory-motor system (including multiplicative noise at the feedback and motor output levels and internal noise in the estimation process), this independence breaks down, requiring additional assumptions and approximations to derive optimal control laws. In this work, we introduce two algorithms that outperform, in terms of cost minimization, state-of-the-art solutions for stochastic control problems in the presence of internal noise. We provide both heuristic and mathematical explanations for this improved performance, offering a practical application for sensory-motor control. These developments will allow stochastic control theory to be applied a to broader range of problems in systems neuroscience.

2-072. A Lagrangian Perspective on Dual Propagation

Rasmus Kjaer Hoier[1,2] HIER@CHALMERS.SE
Christopher Zach[1] ZACH@CHALMERS.SE

[1]Chalmers University of Technology
[2]Electrical Engineering

The search for "biologically plausible" learning algorithms has converged on the idea of representing gradients as activity differences. However, most approaches require a high degree of synchronization (distinct phases during learning) and introduce high computational overhead, which raises doubt regarding their biological plausibility as well as their potential usefulness for neuromorphic computing. Recently it has been shown that by modelling artificial neurons as dyads with two oppositely nudged compartments, it is possible for a fully local learning algorithm to bridge the performance gap to backpropagation, without requiring separate learning phases. However, the algorithm, called dual propagation, has the drawback that convergence of its inference method relies on symmetric nudging of the output units, which may be infeasible in biological and analog implementations. In this contribution we sketch out the steps required to derive dual propagation in a principled manner inspired by LeCun's Lagrangian approach to backpropagation. This leads to a slightly altered variant of dual propagation, which is robust to asymmetric nudging.

2-073. Zero-Shot Visual Numerical Reasoning in Dual-Stream Neural Networks

Jessica Thompson[1] JESSICATHOMPSON00@GMAIL.COM
Hannah Sheahan[2] SHEAHAN.HANNAH@GMAIL.COM
Christopher Summerfield[1] CHRISTOPHER.SUMMERFIELD@PSY.OX.AC.UK

[1]University of Oxford
[2]Google Deepmind

Zero-shot numerical reasoning is challenging for modern computer vision systems but easy for humans. We present a dual-stream glimpsing recurrent neural network that combines gaze contents ("what") and gaze location ("where") to count the number of target items in a visual array, while ignoring distractors. The network successfully learns to count target items and generalizes to an out-of-distribution (OOD) test set including images with novel items. Through ablations and comparison to control models, we establish the contribution of brain-inspired computational principles to this generalization ability. The model displays several neural response properties and patterns of behaviour that have previously been documented in primate visual enumeration. These results provide a proof-of-principle for a theory of the role of the parallel pathways of the primate visual system and posterior parietal cortex in visual relational reasoning.

2-074. Robust variability of grid cell properties within individual grid module enhances encoding of local space

William Redman[1,2] WREDMAN@UCSB.EDU
Xuexin Wei[3,4] WEIXX@UTEXAS.EDU
Santiago Acosta-Mendoza[1] SANTIAGO_ACOSTA@UCSB.EDU
Michael Goard[1] MICHAEL.GOARD@LIFESCI.UCSB.EDU

[1]University of California, Santa Barbara
[2]Dynamical Neuroscience
[3]UT Austin
[4]Department of Neuroscience

The discovery of grid cells (GCs) in medial entorhinal cortex (mEC) has led to considerable experimental and computational work aimed at identifying their properties and function. The organization of GCs into discrete modules, with particular grid properties (e.g., grid spacing and orientation) being conserved within a module, but not between modules, has fundamentally shaped the way in which these questions have been studied. While previous theoretical work has made the additional corollary assumption that GC properties are identical within module, the experimentally measured distributions of spacing and orientation show non-zero variability. This could be due to measurement noise (e.g., finite recording time, neurophysiological noise) or sensitivity of numerical methods used to fit grid properties. Alternatively, this variability could reflect an underlying inhomogeneity, at a fine-scale, within each module. To date, no detailed characterization of the degree and robustness of variability in grid properties within individual modules has been performed. To address this gap, we perform analyses on state-of-the-art mEC electrophysiological recordings, collected with NeuroPixel probes (Gardner et al, 2022). We characterize the heterogeneity of grid spacing and orientation within individual modules, and find evidence for small, yet robust, differences in grid properties of GCs within the same module. To assess the computational implications of this heterogeneity, we perform spatial decoding on synthetic GC populations, allowing us to demonstrate that variability in grid spacing and orientation of a magnitude present in the data leads to improved decoding of local space, even when using the activity of only a single module. Our results challenge a long-held notion about the nature of the grid code and the necessity of output from multiple modules for spatial localization.

2-075. Dynamic reinforcement learning reveals time-dependent shifts in strategy during a reward-learning task.

Sarah Jo Venditto[1,2] VENDITTO@PRINCETON.EDU
Kevin Miller[3,4] KJMD10@GMAIL.COM
Carlos Brody[1,5] BRODY@PRINCETON.EDU
Nathaniel Daw[1] NDAW@PRINCETON.EDU

[1]Princeton University
[2]Neuroscience
[3]Google DeepMind
[4]Neuroscience Lab
[5]Princeton Neuroscience Institute

Different regions and subcircuits in the brain have been hypothesized to dictate multiple "experts" that compete to generate behavior (O'Doherty et al., 2021). A leading example is the study of reinforcement learning (RL). Two general processes, one model-free (MF) and one model-based (MB), capture different learning strategies that interact to inform decisions (Daw et al., 2005), and are frequently associated with competing control of habitual and goal-directed systems in the brain. Behaviorally, this is modeled as a mixture of both learning rules, denoted a mixture-of-agents (MoA) model. A MoA measures the relative contribution of multiple learning rules or "agents" that compete to generate actions. However, changing environments and motivations can cause shifts in strategy,

which cannot be captured by a static MoA. To address this, we present the mixture-of-agents hidden Markov model (MoA-HMM), which combines a MoA with a recent behavioral modeling technique, the GLM-HMM (Ashwood et al., 2022). The MoA-HMM simultaneously learns inferred action values from a set of agents alongside the temporal dynamics of underlying "hidden" states that dictate unique agent mixtures, capturing shifts in agent contributions over time. We apply this model to a multi-step, reward-guided task in rats in which rats use a predominantly MB strategy (Miller et al., 2017). We find that the degree to which MB learning contributes to decisions changes over time. Specifically, we find a progression of strategies throughout sessions: an initial strategy resembling MB exploration, a secondary strategy resembling MB exploitation, and a tertiary strategy resembling decreased engagement. Additionally, the inferred state dynamics significantly predict changes in both response time and OFC neural encoding during the task. More generally, the MoA-HMM can be applied to arbitrary learning rules and behavioral tasks, and future work can look to include additional behavioral and neural metrics for a richer description of behavior.

2-076. Short-circuiting the Wake-Sleep algorithm to model the effects of classical psychedelics

Colin Bredenberg[1] COLIN.BREDENBERG@MILA.QUEBEC
Blake Richards[2] BLAKE.RICHARDS@MILA.QUEBEC
Guillaume Lajoie[1,3] G.LAJOIE@UMONTREAL.CA

[1]Universite de Montreal
[2]McGill University
[3]Mila

Classical psychedelics (including psilocybin, DMT, and LSD) are a family of hallucinogenic compounds with a common mechanism of action: they are agonists for the 5-HT2A serotonin receptor commonly expressed on the apical dendrites of cortical pyramidal neurons (Jakab et al., 1998). These drugs induce complex visual hallucinations in humans, generating percepts that are coherent at a low-level, but which have 'dream-like' qualities at a high level (Carhart-Harris, 2007). While there are many hypotheses as to how 5-HT2A agonists could induce these effects, there are no concrete mechanistic models that capture the variety of observed effects in humans, while remaining consistent with the known pharmacological effects of classical psychedelics. In this work, we investigate a structural similarity between the mechanism of action of classical psychedelics in cortex and the Wake-Sleep algorithm. By simulating the effects of classical psychedelics through manipulating the relative influence of top-down and bottom-up connections in neural networks trained with the Wake-Sleep algorithm on images, we are able to capture a number of effects observed in experiments in individuals under the influence of psychedelics, including: the emergence of closed-eye hallucinations, increases in stimulus-conditioned variability, and large increases in synaptic plasticity. We further provide a number of testable predictions which could be used to validate or invalidate our model.

2-077. A brainwide sequence of activation by arousal

Agnes Landemard[1,2] A.LANDEMARD@UCL.AC.UK
Charu Reddy[1] C.REDDY@UCL.AC.UK
Michael Krumin[1] M.KRUMIN@UCL.AC.UK
Maxwell Shinn[1] M.SHINN@UCL.AC.UK
Matteo Carandini[1] M.CARANDINI@UCL.AC.UK

[1]University College London
[2]Cortical Processing Laboratory

Mice spontaneously go through episodes of high arousal revealed by locomotion, pupil dilation, and whisking, associated with activation of neuromodulatory systems. These episodes dramatically affect cortical activity, but it is not clear how they affect the rest of the brain, and whether they are associated with specific brainwide patterns of activity. To address this issue, we imaged activity in multiple brain regions with functional Ultrasound imaging while head-fixed mice were free to run on a wheel. Locomotion strongly modulated activity not only in the cortex, but in the whole brain. The amplitude and timecourse of this modulation differed across brain regions. Locomotion onset was associated with a stereotyped brainwide sequence of activation over the course of seconds, from ventral regions such as midbrain and hypothalamus up to the cerebral cortex and hippocampus. To model these effects, we fit linear filters for locomotion onsets and wheel velocity. The filters associated with onsets showed a clear progression of onset across the brain over seconds. On slower timescales, the modulation of activity by wheel velocity showed first an increase of activity (for lags 1-10 s) and then a suppression (for lags > 15 s). We replicated these effects in visual cortex using two-photon imaging in a separate cohort of mice. We conclude

that arousal affects activity in the whole mouse brain, not just in the cortex, and does so at multiple timescales, according to a sequence of onsets that evolves over seconds and to a sequence of activations and suppressions that evolves over tens of seconds.

2-078. Distributed memory engrams underlie flexible and versatile neural representations

Douglas Feitosa Tome[1] DOUGLAS.FEITOSATOME@IST.AC.AT
Tim Vogels[2,3] TIM.VOGELS@IST.AC.AT

[1]Institute of Science and Technology Austria
[2]IST, Austria
[3]-

Early pioneering work hypothesized that neuronal ensembles or engrams encoding a specific memory are distributed across multiple functionally-connected brain regions. This network of engram cell ensembles, referred to as a unified engram complex, recently found comprehensive experimental support enabled by technological breakthroughs. However, the computational functions supported by distributed memory engrams remain unclear. Here we employed a modeling approach to investigate whether distributed engram ensembles differentially regulate memory discrimination and generalization, opposing and complementary computations that must be balanced for adaptive memory-guided behavior. For instance, while animals need to discriminate between threat-predictive and neutral stimuli, they also need to generalize reward-predictive cues to novel ones with shared features. By combining brain state-dependent and brain region-specific synaptic plasticity, our multi-region spiking neural network model captured the emergence of functionally-connected and synaptically-coupled distributed engram ensembles in line with previous experimental findings. Critically, our model also generated two testable predictions. First, our model predicted that while engram ensembles in multiple brain regions promote memory generalization following initial encoding, engrams in some regions switch to memory discrimination over the course of memory consolidation while other regions continue to support memory generalization. Second, our model predicted that engram ensembles in monosynaptically-connected brain regions are dynamic, allowing neurons to drop into and out of engrams in each region. Taken together, our results suggest that distributed engram ensembles collectively form a flexible and versatile unified engram complex that: i) supports switching from memory generalization to discrimination for behavioral memory expression as observed experimentally, and ii) enables simultaneous memory generalization and discrimination at the neural level in distinct brain regions. Thus, our work proposes a quantitative and testable theory that uncovers functional flexibility and versatility as normative computational principles underlying the distributed organization of memory.

2-079. Normalization Drives Optimal-like Visuomotor Integration in Drosophila Premotor Circuits

Andre Marques[1,2] ANDRE.MARQUES@NEURO.FCHAMPALIMAUD.ORG
Tomas Cruz[3] TOMAS.CRUZ@NEURO.FCHAMPALIMAUD.ORG
Terufumi Fujiwara[4] TERUFUMIFUJIWARA@GMAIL.COM
Nelia Varela[1] NELIA.VARELA@NEURO.FCHAMPALIMAUD.ORG
Eugenia Chiappe[1] EUGENIA.CHIAPPE@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Research
[2]Champalimaud Neuroscience Programme
[3]Biozentrum, University of Basel
[4]RIKEN

During navigation, animals integrate multimodal information to estimate their movements, control their gaze and interpret external events properly. Yet, the circuit mechanisms underlying these computations remain incompletely understood. In the context of exploration, adult flies (Drosophila melanogaster) use visual feedback to keep their gaze stable. A population of optic-flow sensitive neurons, the so called lobula plate tangential cells (LPTCs), process visual feedback robustly and contribute to gaze stability during straight line walking. Interestingly, LPTCs combine visual and motor information in a behavioral-context specific manner: during rapid gaze shifts, motor signals cancel visual inputs, effectively functioning as efferent copy signals. In contrast, while maintaining gaze stability, visual and motor signals in LPTCs are combined congruently, but the function of this multimodal interaction remains unclear. In this study, we performed whole-cell patch recordings from LPTCs in self-paced walking flies, used normative modeling and targeted cell-activity perturbations to examine the function and mechanisms of these congruent multimodal interactions. We found that LPTCs combine independent unimodal signals following the principle of inverse effectiveness. Moreover, our findings show that under varying sensory reliability, and

in the presence of perturbations, multimodal signals within LPTCs are compatible with the weighting of the unimodal inputs based on their reliability. Leveraging the fly's compact nervous system and the available genetic and connectivity tools, we were able to establish a causal link between normalization, arriving through an identified population of local inhibitory interneurons, and optimal-like multimodal integration within individual neurons. Previous theoretical work has proposed normalization as a potential mechanism for reliability-weighted multisensory integration in primates. Our data demonstrates that normalization plays a critical role mediating optimal multimodal interactions, fine tuning self-motion estimation by LPTCs for steering control. Altogether, these findings suggest the presence of conserved computational principles across animals for the internal estimation of body movement.

2-080. Connectivity of an electron-microscopic reconstruction reveals specific inhibitory competition between large motifs of neurons

Michael W. Reimann[1] MICHAEL.REIMANN@EPFL.CH
Eilif B. Muller[2] EILIF.MULLER@UMONTREAL.CA

[1]ecole Polytechnique Federale de Lausanne (EPFL)
[2]University of Montreal

Neuronal activity is thought to be structured around the activation of assemblies, or a lower-dimensional manifold describing the activity state. Both views describe neurons acting not independently, but in concert, likely facilitated by strong recurrent excitation between them. The role of inhibition in these frameworks – if considered at all – is often reduced to blanket inhibition with no specificity with respect to which excitatory neurons are targeted. We analyzed the structure of excitation and inhibition in the Microns dataset, an electron-microscopic reconstruction of a piece of cortical tissue. We found that excitation was structured around a feed-forward flow in non-random motifs of seven or more neurons. This revealed a structure of information flow from a small number of sources to a larger number of potential targets that became only visible when larger motifs were considered instead of individual pairs. Inhibitory neurons targeted and were targeted by neurons in specific positions of these motifs. Additionally, disynaptic inhibition was strongest between target motifs excited by the same group of source neurons, implying competition between them. The structure of this inhibition was also highly specific and symmetrical, contradicting the idea of non-specific blanket inhibition. None of these trends are detectable in only pairwise connectivity, demonstrating that inhibition is specifically structure by these large motifs. To further demonstrate the significance of the results, we compared them to connectivity in a recently released, detailed computational model and a distance-dependent control. These findings have important implications for how synaptic plasticity reorganises neocortical connectivity to implement learning and for the specific role of inhibition in this process.

2-081. Locomotor maturation during early development in a small vertebrate

Monica Coraggioso[1,2] MONICA.CORAGGIOSO1996@GMAIL.COM
Georges Debregeas[3] GEORGES.DEBREGEAS@UPMC.FR
Volker Bormuth[3] VOLKER.BORMUTH@UPMC.FR
Ghislaine Morvan-Dubois[4] GHISLAINE.MORVAN-DUBOIS@UPMC.FR

[1]Sorbonne University, Paris Brain Institute (ICM)
[2]Physics
[3]Sorbonne University
[4]Chargee de recherche CNRS

Most animals are born with a minimal repertoire of behaviors, which is gradually expanded as they grow and adapt to their changing needs, environments, and body morphology. The emergence of new functional capacities is associated with changes in the brain architecture, as new neurons become functional and new connections are formed. How brain maturation orchestrates phenotypic transformations during development is still an open question in behavioral neuroscience. Here, we investigate this question in Danionella cerebrum (DC), a novel vertebrate model whose brain remains small and transparent up to the adult stage, thus offering a unique opportunity to perform large-scale monitoring of brain activity with cellular resolution through its entire lifespan. A customized freely-swimming assay was developed to obtain a comprehensive dataset on exploratory behavior throughout development. This setup enables the examination of swimming kinematics with high spatial resolution and across various time scales. On the short timescale of seconds, hydrodynamic arguments are employed to elucidate the behavioral ontogeny by establishing a connection between phenotypic changes in behavior and morphologic growth. On a longer timescale, spanning tens of minutes, a Markovian-based state space model is utilized to uncover the inherent structure of exploration, consistently maintaining a cross-age comparative perspective. Finally, calcium imaging across development will elucidate the functional role of neuronal circuits underlying this

phenotypic transition.

2-082. Inferring internally driven cortical dynamics with Sinkhorn recurrent neural networks

Lucas Pompe[1] LUCASVHSPOMPE@GMAIL.COM
Sepp Kollmorgen[2] SEPP.KOLLMORGEN@UZH.CH
Ariel Gilad[3] ARIEL.GILAD@MAIL.HUJI.AC.IL
Fritjof Helmchen[4] HELMCHEN@HIFO.UZH.CH
Valerio Mante[1] VALERIO@INI.UZH.CH

[1]Institute of Neuroinformatics, UZH and ETH Zurich
[2]URPP Adaptive Brain Circuits in Development and Learning, UZH
[3]Faculty of Medicine, the Hebrew University of Jerusalem
[4]University of Zurich

In a variety of tasks and brain areas, neural activity is highly variable across trials, even when all task variables are fixed. The origin and computational significance of this variability remain a matter of debate. Here we propose a modelling approach to distinguish components of the variability that are generated internally, through recurrence in the recorded areas, from components reflecting external inputs. One promising computational approach addressing this problem relies on training an artificial, recurrent neural network (RNN) to reproduce measured neural dynamics. However, current approaches typically suffer from non-identifiability, as many different combinations of recurrent dynamics and external inputs explain the measured activity equally well. To sidestep the fundamental issue of non-identifiability, we bias RNNs to infer only the component of the neural activity that could be explained by internal dynamics in the recorded areas. To this end, the RNNs have no hidden units (each unit reproduces activity at one recording location); are driven by unstructured noise, rather than structured inputs; and are trained to reproduce the full distribution of measured activity, by minimizing the smoothed, optimal transport "Sinkhorn divergence" between predicted and measured activity. We validated this approach on simulations of multi-area RNNs, and find that it retrieves the key features of the ground-truth RNN connectivity and reproduces the resulting recurrent dynamics. We applied this approach to wide-field calcium imaging data from a full cortical hemisphere of mice engaged in a texture-discrimination task. Our Sinkhorn RNNs explain about half the variance in the cortex-wide trial-by-trial variability and correctly predict the functional connectivity between areas. Notably, model performance varies strongly across areas, potentially revealing distinct relative contributions from internal and external influences. Our Sinkhorn RNNs are broadly applicable, and provide a new approach to infer internal dynamics from large-scale neural recordings and imaging in animals and humans.

2-083. Optimal predictive coding in a population of retinal ganglion cells

Kyle Bojanek[1] KBOJANEK@GMAIL.COM
Jared Salisbury[1] JARED.SALISBURY@GMAIL.COM
Baptiste Lefebvre[2] BAPTISTE.LEFEBVRE@ENS.FR
Olivier Marre[2] OLIVIER.MARRE@INSERM.FR
Stephanie E Palmer[1] SEPALMER@UCHICAGO.EDU

[1]University of Chicago
[2]Institut de la Vision, Sorbonne Universite, INSERM, CNRS

Many tasks necessary for an organism's survival require reactions on time scales less than 100 ms. The integration time scale for light in rods and cones ranges from 50 - 120 ms, depending on the organism. This processing lag must be overcome if the organism is to be successful in its particular niche. Previous work has established that in larval tiger salamander retina, information about the stimulus is compressed so as to retain maximal information about the future of the stimulus. Such encodings are essential for overcoming the processing lag in the retina. What should be retained by an optimal predictive compression depends on the the specific dynamics of stimulus, and so must change with the input statistics. We present five different moving bar stimuli to the retina, each with different time constants and deterministic (predictable) and stochastic (unpredictable) components, to investigate how retinal encodings change with the time constant of the moving bar. We use the information bottleneck (IB) technique to find optimal compressions of the stimulus in each context and compare the IB optimal encodings to the retinal ganglion cell (RGC) encoding. We find that for time scales near 100 ms, the retina encodes a stimulus representation very similar to the optimal IB solution for a predictive compensation equal to the processing lag of the retina. For longer time scales, the encoding overshoots the processing lag of the retina and becomes anticipatory, while for shorter time scales, the encoding lags behind. This suggests that IB may be a powerful framework for studying how sensory populations prioritize information.

2-084. Integrating allocentric and egocentric representations for flexible navigation

Daniel Shani DANIEL.SHANI23@GMAIL.COM
Peter Dayan DAYAN@TUE.MPG.DE
Max Planck Institute for Biological Cybernetics

Much work on spatial behaviour equates allocentric representations with strategies based on cognitive maps, and egocentric representations with taxon-like habits (Geerts et al., 2020). This has led to a focus on the hippocampus and the medial entorhinal cortex (MEC), which exhibit allocentric coding for aspects of space in rodents and beyond. However, egocentric representations are of particular value for aspects of policies defined relative to the self, and such deictic notions have been exploited in reinforcement learning (RL; Agre & Chapman, 1987, Finney et al., 2012). The lateral entorhinal cortex (LEC), which is involved in associative learning (Suter et al., 2018, Wilson et al., 2013, Tsao et al., 2018) and spatial processing (Hales et al., 2014), encodes the bearing of external items and boundaries in egocentric coordinates (Wang et al., 2018). This suggests that it might encode a similar sort of cognitive map as the MEC, but in an egocentric reference frame. Here, we build a reinforcement learning agent that combines a putatively LEC-based egocentric successor representation (SR; Dayan, 1993) with a conventional allocentric SR to navigate complex 2D environments. We demonstrate that the agent learns generalisable egocentric and allocentric value functions which can be composed additively to learn policies efficiently and to adapt to new environments quickly. Our work shows the benefit for the hippocampal formation to capture egocentric as well as allocentric relational structure.

2-085. Internal state dependent control of feeding behaviour via hippocampal ghrelin signalling.

Andrew MacAskill[1,2] A.MACASKILL@UCL.AC.UK
Ryan Wee[1] RYAN.WEE.13@UCL.AC.UK
[1]University College London
[2]Neuroscience, Physiology and Pharmacology

Hunger is an internal state that not only invigorates feeding, but also acts as a contextual cue for the higher-order control of anticipatory feeding-related behaviour. The ventral hippocampus is a brain region crucial for differentiating optimal behaviour across different contexts, but how internal context such as hunger influence hippocampal circuits to define behaviour is not known. Pyramidal neurons in the ventral hippocampus, including the ventral CA1/subiculum border (vS) express the receptor for the peripheral hunger hormone ghrelin, and ghrelin is known to cross the blood brain barrier and directly influence hippocampal circuitry. But how ghrelin influences vS has not been directly investigated. In this study, we used a combination of behavioural modelling, electrophysiology, optogenetics and in vivo calcium imaging in mice to investigate the role of vS during feeding behaviour across different states of hunger. We first analysed mouse behaviour around food as a discrete time Markov chain to show that increases in peripheral ghrelin specifically increase the transition from approaching and investigating food, to eating. We then found that activity of a unique subpopulation of vS neurons that project to the nucleus accumbens (vS-NAc) increased when animals approached and investigated food, and this activity inhibited the transition to begin eating. Increases in peripheral ghrelin reduced vS-NAc activity during this anticipatory phase of feeding behaviour by increasing the postsynaptic influence of inhibition, and promoted the initiation of eating. This effect could be reversed by artificial activation of vS-NAc neurons. Furthermore, this peripheral ghrelin-induced inhibition required postsynaptic expression of the ghrelin receptor GHSR1a in vS-NAc neurons, and removal of GHSR1a from vS-NAc neurons impaired ghrelin-induced changes in feeding-related behaviour. Together, these experiments define a ghrelin-sensitive hippocampal circuit that informs the decision to eat based on internal state.

2-086. Sparse and robust memory storage reproduces signatures of synaptic and systems consolidation

Georgios Iatropoulos[1] GEORGIOS.IATROPOULOS@EPFL.CH
Johanni Brea[1] JOHANNI.BREA@EPFL.CH
Wulfram Gerstner[2] WULFRAM.GERSTNER@EPFL.CH
[1]EPFL
[2]ecole Polytechnique Federale de Lausanne

Memory consolidation is believed to involve a two-stage process of engram stabilization and reorganization that

primarily occurs during sleep. On the neural level, recent evidence suggests that this process is underpinned by a combination of memory replay, homeostatic plasticity, and synaptic pruning, together with selective strengthening of task-relevant synapses. It remains unclear, however, how these different aspects of consolidation can be explained within a single computational framework, and how they can be incorporated into existing mathematical models of synaptic plasticity. Here, we propose a normative account of consolidation by deriving a synaptic learning rule that stores memories with maximal noise-tolerance and sparse connectivity in a recurrent neural network. The model includes several mechanisms linked to sleep-based consolidation, such as replay, tagging, and multiplicative hetero- and homosynaptic plasticity, as well as bipartite synaptic connections. Simulations of the model reproduce various hallmarks of consolidation, such as task-driven, self-supervised synaptic pruning and potentiation, increased neural stimulus selectivity, and preferential strengthening of weak memories. A distinct prediction of the model is that intrinsic synaptic noise scales sublinearly with synaptic strength. This is supported by a meta-analysis of multiple published datasets on synaptic fluctuations.

2-087. Multiplexing action selection and learning in the striatum

Jack Lindsey[1] JACKWLINDSEY@GMAIL.COM
Ashok Litwin-Kumar[1,2] A.LITWIN-KUMAR@COLUMBIA.EDU

[1]Columbia University
[2]Neuroscience

Spiny projection neurons (SPNs) in the dorsal striatum are thought to serve as a locus of reinforcement learning (RL) in the basal ganglia, as they process diverse signals from cortex encoding state information, influence action selection, and undergo reward-modulated plasticity mediated by dopamine activity. In this work, we identify and resolve a fundamental inconsistency between standard RL models of the striatum and known plasticity rules in SPNs. SPNs come in two main classes — direct-pathway (dSPNs) and indirect-pathway (iSPNs), which are thought to promote / suppress actions, respectively, and which exhibit different plasticity rules. Broadly, dSPN / iSPN plasticity reinforces activity associated with elevated / suppressed dopamine release, respectively. We show that the iSPN plasticity rule prevents learning from taking place when paired with standard models of striatal activity during action selection, as it reinforces iSPN activity patterns associated with negative outcomes. However, we show that this pathological behavior is reversed if opponent dSPN and iSPN neurons receive correlated input encoding the animal's present action, i.e. an efference copy of the chosen action. We find strong support for this prediction in experimental recordings of dSPNs and iSPNs in dorsolateral striatum (DLS) during spontaneous behavior. Thus, our model provides an explanation for the surprising observation of correlated dSPN and iSPN activation associated with movement onset. Next, we show that our model allows efferent learning signals to be multiplexed with feedforward SPN activity without interference. This enables the striatum to implement so-called off-policy RL algorithms, which allow the cortico-striatal pathway to learn from the outcomes of actions driven by other neural pathways. Pursuing this model further, we note that off-policy RL predicts different responses in dopamine neurons than standard temporal difference learning models, and we find support for the off-policy model of dopamine activity in recent recordings from DLS.

2-088. Graph neural network guided in silico deorphanization technique for olfactory receptors

Grant McConachie[1,2] GDMAC@BU.EDU
Meg A Younger[1] MYOUNGER@BU.EDU
Brian DePasquale[1,3] BDDEPASQ@BU.EDU

[1]Boston University
[2]Biomedical Engineering
[3]BME

Identifying ligands that bind to receptors on olfactory sensory neurons (OSNs) that elicit maximal excitation is known as deorphanization. Deorphanization is imperative for a fundamental understanding of odor encoding, however experiments prove to be both challenging and time-intensive. Performing deorphanization experiments in animal models lacking robust genetic tools further amplifies the complexity of this endeavor. Recently, graph neural networks (GNNs) have been shown to perform well analyzing chemical odors to solve complicated olfactory tasks1-3. We developed a GNN architecture that can accurately predict the OSN response given an odorant, providing an in silico deorphanization technique. We applied this GNN to a large dataset of odorants tested on Drosophila melanogaster receptors called DoOR4. We transformed each odorous molecule in the dataset into a computational graph where each node in the graph is representative of an atom and each edge is representative of a bond between atoms. We then built one graph attention network5 per odorant receptor with the ability to

transform these molecular graphs into predicted OSN responses. When we trained this model on the dataset, we were able to 1) accurately predict the response of an OSN to a held out subset of odors and 2) use the attention mechanism to give insight into what, within a molecule, is responsible for eliciting large responses in the OSN. These findings suggest that GNNs are powerful models to guide and expedite deorphanization efforts, enabling a rapid computational method for characterizing the responses of these critical early sensory neurons. Furthermore, the models offer an avenue for understanding how odor information is encoded in the olfactory system, and could further extend into synthetic molecule generation designed to specifically target certain olfactory neurons.

2-089. Pulvinar interactions with visual cortical areas V1 and V2

Alison Xu ALISON.XU@EINSTEINMED.EDU
Anna Jasper ANNA.JASPER@EINSTEINMED.EDU
Adam Kohn ADAM.KOHN@EINSTEINMED.EDU

Albert Einstein College of Medicine

Visual cortical areas communicate via cortico-cortical (CC) connections, but it has been proposed that they also relay signals to each other via the pulvinar, a thalamic nucleus extensively connected to visual cortex. How this cortico-pulvino-cortical (CPC) pathway contributes to signaling between visual areas has received limited experimental exploration. We sought to determine the properties of the CPC pathway and how they compare to CC signaling utilizing population spiking activity recorded simultaneously in response to drifting gratings in three brain areas: the ventrolateral pulvinar and areas V1 and V2 of anesthetized macaques. To probe the relative timing of activity between pulvinar and cortex, we cross-correlated pulvinar and cortical neuronal responses. V1-V2 and V1-pulvinar pairs were more likely to fire in close relation to each other when their RFs were overlapping, indicating retinotopic specificity of interactions. The V2-pulvinar spike timing relationship was less retinotopically specific and asymmetric, with pulvinar spiking tending to lead V2 activity. To assess signaling between V1-V2 and between these areas and the pulvinar, we fit multivariate regression models to predict trial-to-trial fluctuations of activity in one area using population activity in another. Visual cortical activity was predictive of activity in the pulvinar and the performance of these cortico-pulvinar (CP) models was similar to CC models. The mapping relating cortical to pulvinar activity was low dimensional, suggesting that CP interactions utilize a communication subspace, as previously observed for CC interactions. However, CP and CC communication subspaces relied on different patterns of V1 activity. Our results suggest that the CP and CC pathways are similar in strength and structure, but that the two pathways relay distinct patterns of population activity. These results are consistent with proposals that the CPC represents a distinct functional pathway for routing signals between visual areas.

2-090. Asynchronous Derivative-Free Learning Solving Synaptic Credit Assignment in Recurrent Neural Networks

Saranraj Nambusubramaniyan[1,2] SARANRAJ.NAMBUSUBRAMANIYAN@INFORMATIK.TU-CHEMNITZ.DE
Andreas Knoblauch[3] KNOBLAUCH@HS-ALBSIG.DE
Florian Rohrbein[4] FLORIAN.ROEHRBEIN@INFORMATIK.TU-CHEMNITZ.DE

[1]Technische Universitat Chemnitz
[2]Professorship Neurorobotics
[3]Department of Computer Science, Albstadt-Sigmaringen University
[4]Professorship Neurorobotics, Technische Universitat Chemnitz

In the quest for biologically faithful, derivative-free algorithms, the primary aim is to mimic the efficiency and adaptability observed in neural systems, while also attaining performance comparable to gradient-based methods. However, traditional gradient-based backpropagation (BP) algorithms like Back-Propagation Through Time (BPTT) face criticism for relying on i) synaptic credit assignment that relies on non-local synaptic information, as well as issues such as the ii) update locking and iii) weight symmetry problems.

To address these concerns, we introduce an optimizer, Dopamine, for Weight Perturbation (WP) learning—a reward-based learning algorithm. Dopamine mitigates the aforementioned issues by i) leveraging a global learning signal to acquire "local error information" while adhering to locality constraints, ii) being gradient-free; it updates the parameters asynchronously during the backward pass, eliminating the need to freeze weights, and iii) ensuring that feedforward and feedback weights are not identical by perturbing the feedforward weights to get reward signal.

We conducted a comparative evaluation of our Dopamine powered WP learning against the standalone WP, Stochastic Gradient Descent (SGD) and Adam Optimizers. The tests were conducted on synthetic time series generated by solving the Rossler and Lorenz equations and also on classical "Figure 8" task for time series prediction and system modeling. The results showcase the accelerated convergence of Dopamine model compared to the models trained with SGD and standard WP models. However, Adam exhibits a slight advantage in

convergence speed.

Moreover, Dopamine's parameter update operations are parallelizable and independent of time T, resulting in a total computational complexity of O(2NT+N) given a sequence of length T for N neural units. In contrast, the computational complexity of BPTT is O(NT+TN^2) due to the process of unrolling the network over time. Additionally, by bypassing the activation function, this algorithm provides an alternative to surrogate gradient descent in the training of spiking neural networks.

2-091. Uncovering motifs of concurrent signaling across multiple neuronal populations

Evren Gokcen[1] EGOKCEN@CMU.EDU
Anna Jasper[2] ANNA.JASPER@EINSTEINMED.EDU
Alison Xu[2] ALISON.XU@EINSTEINMED.EDU
Adam Kohn[2] ADAM.KOHN@EINSTEINMED.EDU
Christian Machens[3] CHRISTIAN.MACHENS@NEURO.FCHAMPALIMAUD.ORG
Byron Yu[1] BYRONYU@CMU.EDU

[1]Carnegie Mellon University
[2]Albert Einstein College of Medicine
[3]Champalimaud Research

Modern recording techniques now allow us to record from distinct neuronal populations in different brain networks. However, especially as we consider multiple (more than two) populations, new conceptual and statistical frameworks are needed to characterize the multi-dimensional, concurrent flow of signals among these populations. Here, we develop a dimensionality reduction framework that determines (1) the subset of populations described by each latent dimension, (2) the direction of signal flow among those populations, and (3) how those signals evolve over time within and across experimental trials. We illustrate these features in simulation, and further validate the method by applying it to previously studied recordings from neuronal populations in macaque visual areas V1 and V2. Then we study interactions across select laminar compartments of areas V1, V2, and V3d, recorded simultaneously with multiple Neuropixels probes. Our approach uncovered signatures of selective communication across these three areas that related to their retinotopic alignment. This work advances the study of concurrent signaling across multiple neuronal populations.

2-092. Correcting cortical output: a distributed learning framework for motor adaptation

Leonardo Agueci[1,2] LEONARDO.AGUECI@ENS.PSL.EU
N. Alex Cayco Gajic[1] NATASHA.CAYCO.GAJIC@ENS.FR

[1]Ecole Normale Superieure Paris
[2]Departement d'etudes cognitives

Learning is fundamental for interacting with a changing environment. Critical to this process is the ability to rapidly adapt previously acquired skills to external perturbations while avoiding catastrophic forgetting of previous tasks. Classic behavioral studies of motor control posited that the cerebellum encodes and continuously updates an internal model of the environment in order to adapt motor cortical command signals to external perturbations. However, the underlying mechanism is not well understood. On the other side, recent works have shown how cortical dynamics generate temporally patterned motor outputs and learn from feedback. Despite their distinct architectures and learning rules, the neocortex and cerebellum are densely interconnected through looped pathways. However, it is not clear how to integrate the view of the motor cortical pattern generator with classic cerebellar internal models. To address this gap, we propose that the cerebellum and motor cortex together form a distributed learning system, with supervised cerebellar learning guiding rapid adaptation, and slower consolidation of the new dynamical memory in the motor cortex through local plasticity. We test this hypothesis by training a modular cerebello-cortical recurrent neural network (CCNet) to control a motor plant while performing a visuomotor perturbation to a classic center-out reaching task. We show that CCNet is able to improve task performance in a few trials through rapid adaptation of the cerebellar module. On a slower timescale, cerebellar feedback guides local plasticity in the motor cortex, enabling the adapted dynamics to be stored in the cortical module. This two timescales approach allows CCNet to mitigate catastrophic forgetting by significantly modifying its dynamical memory only in the case of persistent changes of the environment. These results provide new insight into how distinct plasticity rules in different neural circuits could collaborate as part of a brain-wide distributed learning strategy.

2-093. One nose but two nostrils: bilateral alignment of cortical odor representations using sparse connections

Bo Liu[1,2] BOLIU@G.HARVARD.EDU
Shanshan Qin[1,3] SSQIN@SEAS.HARVARD.EDU
Venkatesh Murthy[1] VNMURTHY@FAS.HARVARD.EDU
Yuhai Tu[4] YUHAI@US.IBM.COM

[1]Harvard University
[2]Center for Brain Science, Department of Molecular and Cellular Biology
[3]SEAS
[4]yuhai@us.ibm.com

Bilateral integration of cortical representations is a general problem in neuroscience, which has been mainly studied in vision and audition but less well explored in olfaction. Recent studies reveal that odor representations in two olfactory cortices can be aligned, presumably by structured inter-hemispheric projections, but how this structure forms is unclear. We hypothesized that the continuous exposure to environmental odors shapes these projections, and modeled it as online learning with local Hebb's rule. We found that Hebbian learning with sparse connections achieves bilateral alignment, and identified an inverse scaling between cortical neuron number and the minimum inter-hemispheric projection density for desired alignment accuracy, meaning that more cortical neurons allow sparser inter-hemispheric projections. To clarify this inverse scaling, we obtained the analytical expressions of the inter-hemispheric projection matrices and the alignment accuracy. To compare with local Hebb's rule, we used global stochastic gradient descent (SGD) rule which has higher alignment accuracy but the same inverse scaling. Our work showed that a biologically plausible mechanism with sparse connections suffices for bilateral alignment, and quantitatively compared local Hebb's rule and global SGD rule, which may inspire sparse learning algorithms that are more efficient.

2-094. Feature-dependent mechanisms of reshuffling in cortical circuits

Tuan Nguyen[1,2] THN2112@COLUMBIA.EDU
Alessandro Sanzeni[3] ALESSANDRO.SANZENI@GMAIL.COM
Junxiang Luo[4] JULUO@SALK.EDU
Jonathan Nassi[4] JNASSI@INSCOPIX.COM
John Reynolds[4] REYNOLDS@SALK.EDU
Nicolas Brunel[5] NICOLAS.BRUNEL@DUKE.EDU
Kenneth Miller[1] KENDMILLER@GMAIL.COM
Agostina Palmigiano[1] AP3676@COLUMBIA.EDU

[1]Columbia University
[2]Department of Physics
[3]Bocconi University
[4]Salk Institute for Biological Studies
[5]Duke University

Recent results showed that adding optogenetic excitation to visual stimulation of macaque V1 excitatory (E) cells yields a broad distribution of responses, with a non-significant mean response and no change in the overall distribution of firing rates of cells whose preferred features are matched to the visual stimulus ("rate reshuffling"). Similar broad distributions mixing facilitation and suppression arise in responses to other perturbations, notably, locomotory signals in the visual cortex of the primate. Rate reshuffling matching the data was shown to arise in randomly-connected E/I network models with strong recurrent connectivity through a "tight balance" mechanism: recruiting recurrent inhibition to tightly cancel the mean optogenetic input. Using structured, feature-dependent connectivity allowed reshuffling to occur with looser balance (less precise cancellation), but the specific mechanisms enabling this remain elusive. Here we numerically simulate and use mean-field techniques to analytically study randomly connected rate neurons whose connection probability depends on the orientation preference of cells. Assuming the firing rate moments decay as a Gaussian of the difference between stimulus and preferred visual features to a baseline value, the model's equations reduce to an effective two-location E/I model representing the activity at the visual input peak and of the visual baseline. Under this approximation, we can develop an efficient dynamical mean-field theory approach to predict the responses of a feature-dependent network by only tracking rate moments at three locations. Feature-nonspecific E perturbations cause increases in mean baseline firing which in turn can both suppress the peak response and increase response heterogeneity, enabling reshuffling at a looser level of balance. We fit this model to the distributions of responses of cells to both visual and optogenetic stimuli and demonstrate that feature-dependent connectivity robustly strengthens reshuffling by reducing changes to firing rate moments and increasing response heterogeneity across a wide parameter regime.

2-095. Hierarchical modeling of latent dynamics of subjective value over different timescales

Danilo Trinidad Perez-Rivera[1,2] DP2650@NYU.EDU
Shannon Schiereck[1] SS11853@NYU.EDU
Christine Constantinople[1,2] CMC472@NYU.EDU
Cristina Savin[1,3] CS5360@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]Center for Neural Science, Center for Data Science

Nearly all cognitive processes require computations that span multiple timescales. For instance, subjective values of offers or goods reflect aspects of their objective value, as well as context, which can vary on longer timescales. How do neural population dynamics at multiple timescales reflect the computation of subjective value and drive behavior? We used a Hierarchical Kalman Filter (HKF) to fit large-scale electrophysiological recordings from the orbitofrontal cortex (OFC) of rats performing a value-based temporal wagering task. Critically, the task included explicit reward offers on each trial, but also hidden reward contexts that varied across blocks of many trials, both of which influenced how long rats were willing to wait for a given reward. Neural recordings from OFC revealed responses at the single neuron level that reflected both reward offers and context. We fit these data with a hierarchy of latent processes to explicitly separate slow latent dynamics that varied across trials from faster within-trial dynamics. The HKF model successfully captured low-dimensional latent dynamics operating at both fast and slow timescales, demonstrating superior predictive performance over non-hierarchical models. Analyses of the extracted latents revealed that fast dynamics captured event-aligned responses on single trials. Meanwhile, slow dynamics consistently represented reward contexts, with significant mutual information between the slow latent and hidden reward contexts in the majority of recording sessions. The HKF therefore provides a computational framework for extracting dynamical signatures of computations spanning multiple timescales, a key feature of cognition.

2-096. Representational Geometric Measures for Correlated Neural Manifolds

Chi-Ning Chou[1] CCHOU@FLATIRONINSTITUTE.ORG
Luke Arend[2] LAA9607@NYU.EDU
Albert Wakhloo[3,4] AJW2232@CUMC.COLUMBIA.EDU
SueYeon Chung[5,6] SCHUNG@FLATIRONINSTITUTE.ORG

[1]Flatiron Institute
[2]New York University
[3]Columbia University; Flatiron Institute
[4]Center for Theoretical Neuroscience; Center for Computational Neuroscience
[5]New York University; Flatiron Institute
[6]Center for Computational Neuroscience

Neurons collectively represent task-relevant information in the brain. These representations can be organized into neural manifolds, where a manifold is defined as the collection of neural response vectors to a given task condition or input stimulus. Each manifold occupies only a small fraction of the available dimensions in neural state space, and comes with its own set of measurable geometric properties such as radius, dimensionality and alignments with the others. Manifold geometry is intricately coupled with task performance, and the nature of this interaction furthermore depends on correlations among object manifolds. This invites investigation of the interplay between manifold geometry, correlation and task performance. How do we move beyond traditional dimensionality reduction methods and tease out the underlying task-relevant geometry?

In this work, we adopt the manifold capacity approach (Chung et al., 2018; Wakhloo et al., 2023) to develop several geometric measures that reflect how the correlations between manifolds affect their linear read-out. We then apply these to a diverse range of datasets obtained through different recording methods from various model organisms. For example, we study an electrophysiology dataset on a monkey reaching task requiring rapid learning (Perich et al., 2018). Our effective geometric measures unveil rich nonlinear dynamics for manifolds in motor areas. We find that neural activity during the action stage exhibits ring and center-surround shaped correlations, and that manifold alignments in motor areas are influenced during learning while premotor area representations stay unperturbed. Our work also presents the first application of manifolds capacity theory to motor systems.

2-097. A mechanosensory feedback signal that uncouples external and self-generated sensory responses in the olfactory cortex

Filippo Michelon[1] FILIPPO.MICHELON@IIT.IT
Alireza A. Dehaqani[2] ALIREZA@HMS.HARVARD.EDU
Paola Patella[1] PAOLA.PATELLA@IIT.IT
Luigi Petrucco[1] LUIGI.PETRUCCO@IIT.IT
Eugenio Piasini[3] EPIASINI@SISSA.IT
Giuliano Iurilli[1,4] GIULIANO.IURILLI@IIT.IT

[1]Istituto Italiano di Tecnologia
[2]Istituto Italiano di Tecnologia - Harvard Medical School
[3]International School for Advanced Studies (SISSA)
[4]Systems Neurobiology

Sampling behaviors have sensory consequences that can hinder the stable representation of the external world. In olfaction, sniffing alters the odor concentration encountered by the olfactory sensory neurons, mimicking a sudden increase in odor concentration. Here, we examined how the inhalation speed impacts the representation of odor concentration in the main olfactory cortex. To this aim, we analyzed the spiking activity of hundreds of cortical neurons on a sniff-by-sniff basis while presenting odors at different environmental concentrations. We observed that neurons combined the odor input with a mechanosensory feedback signal generated by the air passage through the nose during inhalation. Thus, the odor responses of individual neurons significantly varied with the inhalation speed. Still, the population representation of concentration was remarkably sniff-invariant. This was because the mechanosensory and olfactory inputs were uncorrelated within and across cortical neurons. Thus, a faster odor inhalation and an increase in environmental concentration changed the cortical activity pattern in distinct ways. A simulation showed that this encoding strategy affords tolerance to potential concentration fluctuations inside the nasal cavity caused by varying inhalation speeds. Mechanosensory reafferences are widespread across sensory systems. The coding scheme described here may be a canonical strategy that complements inhibitory efferent motor corollary discharges to mitigate the sensory ambiguities caused by the movement of the sensors.

2-098. A theory of memory stability in hippocampal area CA3

Uri Cohen[1,2] UC231@CAM.AC.UK
Roland Mason Rodriguez[1] RMR50@CAM.AC.UK
Ole Paulsen[1] OP210@CAM.AC.UK
Mate Lengyel[3] M.LENGYEL@ENG.CAM.AC.UK

[1]University of Cambridge
[2]Engineering
[3]University of Cambridge; Central European University

Hippocampal area CA3 is widely considered to play a key role in memory. Providing experimental support for this idea requires a concrete computational account of how memories can be stored and recalled robustly in hippocampal circuits. Classical theories of auto-associative memory (i.e. the Hopfield model and its extensions) described conditions for successful memory recall but considered circuit dynamics with limited biological plausibility and thus provided few specific experimentally testable predictions. Here, using the replica method from statistical physics, we derive a new analytical theory for the recall of distributed patterns of sparse, graded activities by a recurrent neural network with rate-based dynamics. We assume the existence of a plasticity rule that stores such patterns as steady states of the network's dynamics and, unlike previous theories, we also specifically analyse the dynamical stability of these steady states – a critical prerequisite of robust memory recall. Surprisingly, we find that the stored patterns tend to be all stable or all unstable, and establish formal mathematical conditions for the stable recall of a large number of memories. Characterizing this stable regime provides experimentally testable predictions. First, neurons should be spontaneously firing even in the absence of recurrent inputs. Second, the optimal form of the single-neuron input-output nonlinearity depends on the statistics of memory patterns. To further constrain our predictions, we re-analysed publicly available data from CA3 place cells and found pattern statistics for which the theory specifically predicts that rectified linear neural activation maximises memory storage. We provide preliminary evidence for such f-I curves in CA3 pyramidal cells. Our theory provides a new analytical framework for understanding the interplay of single-neuron properties and circuit-level dynamics in memory recall and makes concrete predictions that will be readily testable in future experiments.

2-099. A posture subspace in primary motor cortex

Patrick Marino[1] PMARINO162@GMAIL.COM
Lindsay Bahureksa[2] LBAHUREK@ALUMNI.CMU.EDU
Carmen Fernandez Fisac[2] C.F.FISAC@GMAIL.COM
Emily Oby[1] EMILYOBY@GMAIL.COM
Adam Smoulder[2] ASMOULDE@ANDREW.CMU.EDU
Asma Motiwala[2] AMOTIWALA@CMU.EDU
Alan Degenhart[1] ALANDEGENHART@GMAIL.COM
Erinn Grigsby[1] ERINN.GRIGSBY@GMAIL.COM
Wilsaan Joiner[3] WMJOINER@UCDAVIS.EDU
Steven Chase[2] SCHASE@ANDREW.CMU.EDU
Byron Yu[2] BYRONYU@GMAIL.COM
Aaron Batista[1] AARON.BATISTA@GMAIL.COM

[1]University of Pittsburgh
[2]Carnegie Mellon University
[3]University of California, Davis

To generate movement commands, the brain must combine information about the goal of the movement with information about the posture of the body. Primary motor cortex (M1) is a key node for the convergence of these information streams and is the primary source of descending motor commands for volitional movement. How are posture and goal information organized within M1's neural activity to permit the rapid and flexible generation of movement commands? To answer this question, we recorded neural population activity in M1 while Rhesus monkeys performed brain-computer interface (BCI), isometric force, and reaching tasks with the forearm in a range of postures. We observed a clear and simple population-level organization of posture information in M1. First, posture- and goal-related components of neural population activity were separable and resided in nearly orthogonal subspaces. Second, a single subspace could be used to decode the arm's posture across tasks, even though arm posture had very different implications for the performance of each task. Third, within each task, neural trajectories for a given goal had similar shapes across postures. Together, these findings indicate the existence of a "posture subspace" in M1 activity that is stable across tasks. This posture subspace might allow for posture and goal information to be flexibly combined, as they likely are in the service of our broad repertoire of everyday actions.

2-100. Dynamic fading memory and prior expectancy effects in monkey primary visual cortex

Yiling Yang[1] YILING.YANG@ESI-FRANKFURT.DE
Johanna Klon-Lipok[2] JOHANNA.KLON@BRAIN.MPG.DE
Katharine Shapcott[1] KATHARINE.SHAPCOTT@ESI-FRANKFURT.DE
Andreea Lazar[1] ANDREEA.LAZAR@ESI-FRANKFURT.DE
Wolf Singer[2] WOLF.SINGER@BRAIN.MPG.DE

[1]Ernst Strungmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society
[2]Ernst Strungmann Institute

Neuroimaging studies on human subjects suggest that working memory (WM) content can be read out from primary visual cortex (V1) and that activity-silent WM can be revived by delivering a magnetic or visual impulse ("pinging"). This WM reactivation in the impulse response agrees with theoretical prediction that a nonspecific input to the whole network can reactivate WM stored in synapses. To evaluate the decodability and reactivation of WM in V1 at neuronal level, we recorded multiunit activity from awake monkey V1 engaged in a delayed match-to-sample (DMS) task. During the WM delay, V1 neurons maintained a persistent trace of the sample stimulus which could be reactivated by an intervening impulse stimulus that enhanced neuronal firing. This lingering and revivable stimulus trace was also present in a passive viewing control task without WM, and therefore likely resulted from the fading memory of the reverberating local recurrent network dynamics. Optimal perceptual decisions integrate external sensory input with internal priors of stimulus probability. To investigate whether predictive priors stored in WM affect stimulus encoding, we introduced an implicit prior of transition probability between sample and test stimuli, such that the sample would predict the upcoming test stimulus with variable probability. Test stimuli with high predictive probability evoked lower firing rates in V1 than those with low probability. The effect was only visible when we gradually ramped up stimulus intensity, to dampen the sharp onset transient produced by abrupt stimulus onset. The probability-dependent firing rate difference emerged in late sessions, implying gradual learning of the prior. Visual system used prior probability in WM to improve coding efficiency.

2-101. State-dependent Spatiotemporal Dynamics of Noradrenergic Release in the Neocortex

Clayton Barnes[1,2] CLAYTON.BARNES@YALE.EDU
Jessica Cardin[1] JESS.CARDIN@YALE.EDU

[1]Yale University
[2]Neuroscience

Animals cycle through multiple waking brain-states that profoundly influence patterns of neuronal activity, perception, and behavior. A growing body of research suggests that arousal, reflected in elevated motor activity, is associated with distinct alterations in local circuit operations within the neocortex. Moreover, changes in behavioral state are linked to reorganization of cortical networks underlying cognitive and motor processes. This state-dependent neural variation reflects the coordinated, brain-wide influence of ascending neuromodulatory systems. Noradrenergic neurons in the locus coeruleus (LC) send projections throughout the neocortex that contribute to arousal, sleep-wake state transitions, attention, and state-dependent cognitive processing. Noradrenergic projection neurons in the LC make highly divergent projections across the brain. The canonical model of the noradrenergic (NE) system posits phasic and tonic firing modes of the locus coeruleus (LC) that cause spatially homogeneous release and region-specific effects through the differential expression of adrenergic receptors, consequently regulating the engagement of distinct cognitive elements and coordinating noradrenergic. Noradrenergic signaling may critically regulate cortical firing and modulate local functional connectivity. Using single and multichannel mesoscopic imaging of the dorsal cortex, paired with counterbalanced, double-blind in vivo local pharmacological infusions, we measured the spatiotemporal dynamics and impact of NE signaling in the neocortex of awake behaving mice. In a subset of experiments, we simultaneously measured either neural calcium (jRCaMP1b, red) and NE (Ne2m, green) signals or cholinergic (ACh1.4, red) and noradrenergic (Ne2m) signals across the cortex. We assessed the relationship between behavioral state and neuromodulation in spontaneously behaving mice and following moderate stress. We find spatiotemporally heterogeneous release of NE that regulates long-range functional connections between cortical regions. NE and ACh exhibit distinct spatiotemporal patterns. Finally, the spatial distribution of NE release associated with arousal is altered by a moderate restraint paradigm, suggesting long-term changes in neuromodulatory regulation of cortical networks during periods of heightened stress.

2-102. Gradient-based updates in hierarchical sensory models mimic category learning effects in macaque IT

Lynn Soerensen[1,2] LYNNKA@MIT.EDU
James J. DiCarlo[3] DICARLO@MIT.EDU
Kohitij Kar[4,5] K0H1T1J@YORKU.CA

[1]Massachusetts Institute for Technology
[2]McGovern Institute for Brain Research
[3]Massachusetts Institute of Technology
[4]York University
[5]Biology

Like humans, adult non-human primates can learn to categorize visual objects. Prior work shows that neurons in the inferior temporal (IT) cortex, critical for object recognition, modestly increase their selectivity to objects from learned categories. Do these neural changes underlie the behavioral performance gains (“learning”), and if so – how? While the field now has relatively accurate models of image-driven IT responses, we still lack a similar computational understanding of adult IT plasticity. To address this, we measured neural activity across the IT cortex in two groups of monkeys: one group ("naive") was only trained to fixate passively on images; the other group ("trained") also learned to discriminate object categories. First, consistent with previous studies, we observed a significant increase (63%) in object-category selectivity in IT responses of trained compared to naive monkeys. Next, this selectivity increase led to a more (37%) categorical representation at the population level (as assessed by an RDM analysis), and also enhanced (19%) the IT population activity-based linear decoding accuracy for the learned object categories. Lastly, these changes in trained responses also improved the predictions of image-level behavioral error patterns. How do these observed changes in IT lead to improvements in behavior? We present a systems-level perspective by casting the monkey’s category training as an extension of contemporary artificial neural networks (ANNs). Interestingly, we observed that for various finetuned ANNs (with different architectures, pre-training objectives, and category learning schemes), the untrained IT-matched ANN-layer showed macaque-IT-like increases in category information after training. Akin to IT, specific ANN-IT representations were also more predictive of monkey behavior after training. In sum, we provide empirical evidence of moderate, behaviorally relevant plasticity in adult IT upon category learning and introduce a computational framework to simulate these changes, enabling us to formulate testable hypotheses about the representational reconfigurations induced by category learning.

2-103. Adaptive coding efficiency with fast gain modulation and slow synaptic plasticity

David Lipshutz[1] DLIPSHUTZ@FLATIRONINSTITUTE.ORG
Lyndon R. Duong[2] LYNDON.DUONG@GMAIL.COM
Dmitri B Chklovskii[3] DCHKLOVSKII@FLATIRONINSTITUTE.ORG
Eero P. Simoncelli[2] EERO.SIMONCELLI@NYU.EDU

[1]Flatiron Institute
[2]NYU and Flatiron Institute
[3]Flatiron Institute and NYU Medical School

Efficient transmission of information from dynamic environments necessitates sensory systems that rapidly adapt to changes in sensory statistics. Neurons in early sensory areas rapidly adapt to changes in sensory statistics to both normalize their response variances and reduce between-neuron response correlations, which together can be viewed as a form of adaptive whitening. While it is well-established that single neurons can modulate their gains to normalize their response variance, the mechanism underlying adaptive decorrelation of neural populations remains unclear. Existing mechanistic models of adaptive whitening exclusively use either synaptic plasticity or gain modulation as the biological substrate for decorrelation; however, on their own, each of these models has significant limitations.

We unify these approaches in a multi-timescale mechanistic model that adaptively whitens its responses using a combination of fast gain modulation and slow synaptic plasticity. Our model is derived from a novel whitening objective that factorizes the whitening transformation into context-independent basis vectors, which correspond to synaptic weights, and a context-dependent diagonal matrix, which corresponds to neuronal gains. We test our model on synthetic and natural datasets and find that the synapses learn to match shared structure of the stimulus inputs over long timescales that enable adaptive whitening on short timescales using gain modulation.

2-104. Replay constructs compositional maps in hippocampus

Jacob Bakermans[1,2] JACOB.BAKERMANS@GMAIL.COM
James Whittington[3] JCRWHITTINGTON@GMAIL.COM
Joseph Warren[4] JOSEPH.WARREN@UCL.AC.UK
Timothy Behrens[5] BEHRENS@FMRIB.OX.AC.UK

[1]University of Oxford
[2]Clinical Neurosciences
[3]University of Stanford
[4]University College London
[5]University of Oxford & University College London

Hippocampal replay, the activation of remote representations while an animal is at rest, has been implicated in a wide range of mental processes, from planning and credit assignment to consolidation and memory. We will further complicate this picture by providing evidence for yet another (though intimately related to all of the above) function: the construction of compositional maps. This function was proposed in recent theoretical work, which hypothesises that hippocampus composes state spaces from structural elements, like space and walls and rewards. The idea that replay carries out this composition leads to precise neural predictions, for example "some landmark cells will appear in replay (...) before they appear in physical navigation". Here, we test this prediction in two datasets, and find that upon replay, changes in the ratemap of hippocampal neurons 1) closely align to the neuron's replayed location, 2) generalise in a compositional way, and 3) reflect structural changes beyond just reward. Together, these results support a role for replay in building maps (1) through composition (2) from structural elements (3). We demonstrate how replay changes maps in hippocampal recordings during an alternating home-away well task. We find many replays while the animal sits at the home well, where planning is not very useful (the away well is in a random location) but mapping how to return home later is. Indeed, when we decode these replay trajectories, we discover that ratemap changes occur exactly where a cell fires in replay. Moreover, for certain cells these ratemap changes generalise across home well locations, affording compositional reconfiguration when the environment changes. Finally, such environment changes may include structure beyond reward: in a four-room maze task where doors lock halfway through the experiment, we find new place fields that emerge after replay at a recently locked door.

2-105. Contextual reasoning in the primate hippocampal-prefrontal circuit

Thomas Elston[1,2] TELSTON@NURHOPSI.ORG
Joni Wallis[1] WALLIS@BERKELEY.EDU

[1]University of California, Berkeley
[2]Helen Wills Neuroscience Institute

What is good in one scenario might be bad in another. Despite the ubiquity of such contextual reasoning in everyday choice, how the brain flexibly utilizes different valuation schemes across contexts remains unknown. We addressed this question by monitoring neural activity from the hippocampus (HPC) and orbitofrontal cortex (OFC) of two monkeys performing a state-dependent choice task. We found that HPC neurons encoded state information as it became available and then, at the time of choice, relayed this information to OFC via theta synchronization. During choice, OFC represented value in a state-dependent manner: many OFC neurons uniquely coded for value in only one state but not the other. This suggests a functional dissociation whereby HPC encodes contextual information that is broadcast to OFC via theta synchronization to select a state-appropriate value subcircuit, thus allowing for contextual reasoning in value-based choice.

2-106. Cortical representation of economic values independent from actions

Oliver Gauld[1] O.GAULD@UCL.AC.UK
Joseph Tutt[1] JOSEPH.TUTT@UCL.AC.UK
Joseph Warren[2] JOSEPH.WARREN@UCL.AC.UK
Jingjie Li[1] JINGJIE.LI.21@UCL.AC.UK
Jeffrey Erlich[2,3] J.ERLICH@UCL.AC.UK
Chunyu A. Duan[2,3] ANNDUAN2@GMAIL.COM

[1]Sainsbury Wellcome Centre, UCL
[2]University College London
[3]Sainsbury Wellcome Centre

In economic decision-making, individuals make multi-attribute choices by computing and comparing expected values of different offers. Decisions can be formed in value space while spatial choices are planned in action space before enactment in the physical world. Dissecting the neural mechanisms that delineate these processes is challenging as value representations are often correlated with other ongoing cognitive and preparatory signals. To address this issue, we designed a novel head-fixed mouse behaviour to dissociate in time decisions in value space from action planning. On each trial, mice chose between a small but safe reward (surebet) and a risky, probabilistic reward (lottery). Mice were first cued with how good the lottery was on that trial before the location of the lottery was revealed, separating in time the choice between offers and the selection of spatial actions. To identify value-based versus action-based valuation networks, we conducted widefield imaging of GCaMP6s fluorescence across the dorsal cortex in expert mice during engaged task performance (identified by GLM-HMM). We then used a linear encoding model to map task-related (expected value, 'risky' and spatial choices) and task-unrelated (e.g. orofacial movements) variables onto spatiotemporal cortical activity. We found widespread coding of value and risky choice with distinct temporal dynamics across brain regions that correspond to value- versus action-based valuation. In particular, posterior association areas such as the retrosplenial cortex (RSC) preferentially encoded lottery value during the value epoch compared to the motor planning epoch. In comparison, in frontal areas such as the secondary motor cortex (M2), value signals persisted into the action planning period, possibly instantiating value-to-action transformation. Our ongoing multi-region two-photon calcium imaging and laser-scanning optogenetic perturbations will further identify specific neural circuits and causal mechanisms underlying the transformation of abstract economic decisions to embodied choices.

2-107. Inhibition-stabilized supralinear memory ensembles

Samuel Eckmann[1,2] EC.SAM@OUTLOOK.COM
Mate Lengyel[3] M.LENGYEL@ENG.CAM.AC.UK
Yashar Ahmadian[1] YA311@ENG.CAM.AC.UK

[1]University of Cambridge
[2]Computational and Biological Learning Lab
[3]University of Cambridge; Central European University

The hippocampus plays a central role in memory formation and retrieval. In order to avoid deleterious interference between stored and ongoing experience, theoretical considerations require a separation between phases

of memory encoding and recall within the same neural circuit, putatively controlled by hippocampal theta oscillations. However, the neural mechanisms subserving this separation remain unknown. Classical models either remain mute about these mechanisms, or assume purpose-built neuromodulatory interactions that are in conflict with biologically realistic timescales and specificity of synaptic modulation. In addition, computational models of memory recall typically do not consider inhibitory neurons at all, or only for stabilizing the network globally. In contrast, recent experiments suggest that structured inhibitory connections are crucial for memory retrieval. Here, we develop an excitatory-inhibitory network model with structured connectivity between units conforming to a canonical circuit motif, the inhibition-stabilized supralinear network. This network naturally gives rise to a separation between phases that are ideal for either the recall or the storage of memories, solely determined by the input strength of an external memory cue. For weak input, the cued memory is recalled, and neurons are strongly stabilized by inhibition. For strong input, the external cue is encoded, while inhibition stabilization is paradoxically weaker. Our model only requires a Hebbian and an anti-Hebbian form of biologically plausible plasticity that respectively store the positive and negative parts of the pattern covariance matrix. Specifically, patterns are stored in an input-dominated encoding regime via synapse-type-specific competitive Hebbian plasticity, for the positive part, and—motivated by experimental results—via anti-Hebbian plasticity at excitatory-to-inhibitory synapses, for the negative part. The resulting recurrent connectivity is highly structured and consistent with Dale's law. In summary, we present a model of hippocampal memory recall that meets key biological constraints and reveals a novel mechanism for alternating between storage and recall within the same circuit.

2-108. Continual learning using dendritic modulations on view-invariant feedforward weights

Viet Anh Khoa Tran[1,2] V.TRAN@FZ-JUELICH.DE
Emre Neftci[3] E.NEFTCI@FZ-JUELICH.DE
Willem Wybo[4] W.WYBO@FZ-JUELICH.DE

[1]Forschungszentrum Julich
[2]PGI-15 - Neuromorphic Software Ecosystems
[3]Juelich Research Center
[4]Julich Research Center

The brain is remarkably adept at learning from a continuous stream of data without significantly forgetting previously learnt skills. Conventional machine learning models struggle at continual learning, as weight updates that optimize the current task interfere with previously learnt tasks. A simple remedy to catastrophic forgetting is freezing a network pretrained on a set of base tasks, and training task-specific readouts on this shared trunk. However, this assumes that representations in the frozen network are separable under new tasks, therefore leading to sub-par performance. To continually learn on novel task data, previous methods suggest weight consolidation – preserving weights that are most impactful for the performance of previous tasks – and memory-based approaches – where the network is allowed to see a subset of images from previous tasks. For biological networks, prior work showed that dendritic top-down modulations provide a powerful mechanism to learn novel tasks while initial feedforward weights solely extract generic view-invariant features. Therefore, we propose a continual learner that optimizes the feedforward weights towards view-invariant representations while training task-specific modulations towards separable class clusters. In a task-incremental setting, we train feedforward weights using a self-supervised algorithm, while training the task-specific modulations and readouts in a supervised fashion, both exclusively through current-task data. We show that this simple approach avoids catastrophic forgetting of class clusters, as opposed to training the whole network in a supervised manner, while also outperforming (a) task-specific readout without modulations and (b) frozen feedforward weights. This suggests that (a) top-down modulations are necessary and sufficient to shift the representations towards separable clusters and that (b) the SSL objective learns novel features based on the newly presented objects while maintaining features relevant to previous tasks, without requiring specific synaptic consolidation mechanisms.

2-109. Segregated neuronal populations in prefrontal cortex encode task variables during working memory

Klaus Wimmer[1] KWIMMER@CRM.CAT
Bijan Pesaran[2] BIJAN.PESARAN@PENNMEDICINE.UPENN.EDU
Nicolas Pollan Hauer[1] NPOLLAN@CRM.CAT

[1]Centre de Recerca Matematica
[2]University of Pennsylvania

Non-linear mixed selectivity, with neurons responding to diverse combinations of task-relevant variables, has

been proposed as a key mechanism to enable flexible behavior and cognition. However, it is debated whether the structure of neural population responses in fronto-parietal cortices is better described as random mixed-selective or as non-random, that is, in terms of multiple subpopulations with stereotypical response profiles. Here, we show that neural activity in macaque prefrontal cortex during a working memory and a visual response task is organized into subpopulations that provide a comprehensive description of the low-dimensional population dynamics. First, analysis of the demixed Principal Components shows that the neural code faithfully represents stimulus identity, task condition and elapsed time during the trial. Second, a model-free analysis of the population structure reveals a significant degree of clustering, implying a non-random distribution of feature selectivity that is incompatible with random mixed selectivity. Closer inspection shows stimulus-selective neurons also tend to be task-selective. Third, examining the contribution of stimulus-selective neurons to task condition-related variance reveals two contrasting activity profiles that correspond to functionally different populations. One population responds during visual stimulation while the other activates during memory maintenance. Finally, the observed neural geometry explains how stable task and stimulus information can be read out from the population response using a linear decoder. Our results highlight that despite the heterogeneity of prefrontal responses during working memory, neurons do not represent random mixtures of task features but are structured according to neural subpopulations.

2-110. Disentangling the roles of distinct cell classes with cell-type dynamical systems

Aditi Jha[1,2] ADITIJHA@PRINCETON.EDU
Diksha Gupta[3] DIKSHA.GUPTA@UCL.AC.UK
Carlos Brody[1,4] BRODY@PRINCETON.EDU
Jonathan Pillow[1] PILLOW@PRINCETON.EDU

[1]Princeton University
[2]Electrical and Computer Engineering
[3]Sainsbury Wellcome Centre
[4]Princeton Neuroscience Institute

Latent dynamical systems are widely used to characterize the low-dimensional dynamics of neural population activity and gain insight into the computations underlying behavior. However, these models often ignore the fact that the brain contains multiple cell types. This prevents them from shedding light on the functional roles of distinct cell classes, and from accounting for experimental data involving optogenetic perturbations of particular cell types. To overcome this limitation, we introduce "cell-type dynamical systems" (CTDS). We extend latent linear dynamical system (LDS) models by placing structural constraints on the model parameters so that each cell class has a distinct set of latent variables. To illustrate our approach, we focus on networks with two distinct cell types: excitatory (E) and inhibitory (I) neurons. Standard analyses would characterize the dominant modes of network activity without considering the distinct roles played by E and I cells. By contrast, our CTDS model defines separate latents for E and I cells, and constrains the dynamics matrix so that E and I latents have a positive or negative (respectively) effect on the other latents. We applied CTDS to recordings from rat frontal orienting fields (FOF) and anterior dorsal striatum (ADS) during an auditory decision-making task. We find that CTDS predicts neural activity better on held out trials than a standard LDS model. We also show that the latents extracted from both E and I populations can be used to decode the animal's choice, revealing that choice-related information is not restricted to a single cell class. Finally, we simulate perturbed neural activity by inhibiting E neurons during part of the trial and show that our model is able to accurately infer the dynamical effects of the perturbations in the latent space. Overall, CTDS can be broadly useful for accurately identifying the causal neural dynamics underlying behavior.

2-111. Time cells contribute to working memory through value-based recurrent dynamics

Dongyan Lin[1] DONGYAN.LIN@MAIL.MCGILL.CA
Blake Richards[2] BLAKE.RICHARDS@MILA.QUEBEC
Ann Huang[3] ZIXIANG.HUANG@MAIL.MCGILL.CA

[1]Mila, McGill University
[2]McGill University
[3]Harvard University

Neuroscientists have observed both cells in the brain that fire at specific points in time, known as “time cells”, and cells whose activity steadily increases or decreases over time, known as “ramping cells”. It is speculated that time and ramping cells support temporal computations in the brain and carry mnemonic information. However,

we still have little understanding whether and how these cells really contribute to behavior. Here, by training deep reinforcement learning (DRL) models on the simulated Delayed Nonmatch-to-Stimulus (DNMS) task, we show that time cells and ramping cells naturally emerge in recurrent neural networks trained to estimate expected rewards in the future. Then, leveraging our ability to selectively manipulate activity in specific units of the RNNs at specific times, we show that these cells indeed carry information about stimuli, but they contribute to behavior in large part via their contributions to recurrent dynamics. Moreover, by separating the circuits for value calculation and action execution, we provide evidence that value learning gives rise to the temporally informative recurrent dynamics, which helps the network learn to compute the policy. Our results suggest that time cells and ramping cells could contribute to temporal and mnemonic calculations, but the way in which they do so may be complex and unintuitive to human observers.

2-112. The entrainment power of the external environment on chimera states: a computational stochastic model.

Jacopo Epifanio[1] JACOPO.EPIFANIO@UPF.EDU
Ralph Gregor Andrzejak[2] RALPH.ANDRZEJAK@UPF.EDU

[1]Universitat Pompeu Fabra - Campus de la Ciutadella
[2]Universitat Pompeu Fabra

Chimera states are intriguing phenomena characterized by the coexistence of synchronous and asynchronous motion in networks of coupled identical oscillators. Accordingly, the oscillators segragate into two groups: a high-coherence group and a low-coherence one. This presence of partial synchronization in systems of identical oscillators makes chimeras good models for studying real-world complex systems, such as the brain. Therefore, chimeras have also been extensively studied in multi-layer networks. For example, Andrzejak et al. demonstrated the possibility of obtaining synchronization between a driving single-layer chimera network and a driven one. Nevertheless, an extensive research on the possibility of achieving synchronization between the external environment and chimeras is still lacking. Here, we address this open problem by driving chimeras with a multivariate stochastic model representing the external environment. In particular, we focus on measuring the entrainment power that these external signals exert on chimeras, where by entrainment power we mean the external signals' capacity to impose their own rythm on the driven system. In this context, we have found an increment of the entrainment power by increasing not only the driving strength, as it must be expected, but also the autocorrelation of the multivariate stochastic driving process. A possible future application of this result could lie in the Seizure Onset Zone (SOZ) localization for pharmaco-resistant focal epilepsy patients who need surgical resection of the SOZ to be epilepsy-free. Indeed, we believe that by driving chimeras with phases from electroencephalography (EEG) signals of epilepsy patients we would be able to recognize SOZ channels because they would show a higher entrainment power. This belief is rooted in the observation that signals from the SOZ exhibit a greater autocorrelation compared to the others.

2-113. Cell-type-specific plasticity shapes neocortical dynamics for motor learning

Shouvik Majumder[1,2] SHOUVIK.MAJUMDER@MPFI.ORG
Koichi Hirokawa[1] KOICHI.HIROKAWA@MPFI.ORG
Zidan Yang[1] ZIDAN.YANG@MPFI.ORG
Ronald Paletzki[3] RON.PALETZKI@NIH.GOV
Charles Gerfen[3] GERFENC@MAIL.NIH.GOV
Lorenzo Fontolan[4] LORENZO.FONTOLAN@UNIV-AMU.FR
Sandro Romani[5] ROMANIS@JANELIA.HHMI.ORG
Anant Jain[1] ANANT.JAIN@MPFI.ORG
Ryohei Yasuda[1] RYOHEI.YASUDA@MPFI.ORG
Hidehiko Inagaki[1] HIDEHIKO.INAGAKI@MPFI.ORG

[1]Max Planck Florida Institute for Neuroscience
[2]Neural Dynamics and Cognitive Functions
[3]National Institute of Mental Health
[4]Turing Centre for Living Systems
[5]Janelia Research Campus

Neocortical spiking dynamics control many aspects of behavior, yet how these dynamics emerge during learning remains elusive. Activity-dependent synaptic plasticity is likely a key mechanism, as it reconfigures network architectures that govern neural dynamics. Yet, how plasticity in specific neocortical cell types contributes to learning

and how it influences spiking dynamics that drive behavior is unknown. Here, we examined how the mouse premotor cortex acquires neural dynamics that control movement timing, specifically lick timing. To probe the role of synaptic plasticity, we have performed a series of transient and cell-type-specific genetic manipulations of proteins essential for major forms of synaptic plasticity, Ca2+/calmodulin-dependent protein kinase II (CaMKII) and Cofilin, in a region and cell-type-specific manner. We discovered that plasticity-related proteins in a small subpopulation of glutamatergic cell type, pyramidal tract (PT) neurons, but not in a major subpopulation, intratelencephalic (IT) neurons, in the premotor cortex are essential for learning new lick timing. High-density electrophysiology in the premotor cortex uncovered that neural dynamics anticipating licks are progressively reconfigured during learning, which explains the change in lick timing. CaMKII manipulation in PT neurons stops the reconfiguration of behaviorally relevant dynamics, and therefore disrupts learning. To explain learning-induced changes, we developed a network model implementing a simple, reward-based plasticity rule. The model captured lick timing behavior and neural dynamics during learning. Altogether, by combining electrophysiology, manipulation of synaptic plasticity, and computational modeling, we have identified a key cell type required for sculpting neocortical dynamics to learn new behavior.

2-114. Towards end-to-end cell-typing in large-scale recordings

Simone Azeglio[1,2] SIMONE.AZEGLIO@GMAIL.COM
Thomas Buffet[3] THOMAS.BUFFET@INSERM.FR
Gabriel Mahuas[4] GABRIEL.MAHUAS@INSERM.FR
Chiara Boscarino[5] CHIARA.BOSCARINO@MAIL.POLIMI.IT
Ulisse Ferrari[6] ULISSE.FERRARI@INSERM.FR
Olivier Marre[7] OLIVIER.MARRE@INSERM.FR

[1]Sorbonne University & Ecole Normale Superieure
[2]Vision Institute & Laboratoire des Systemes Perceptifs
[3]Sorbonne Universite, Institut de la Vision
[4]Vision Institute (Sorbonne University) & Ecole Normale Superieure
[5]Vision Institute (Sorbonne University) & Politecnico di Milano
[6]Vision Institute (Sorbonne University)
[7]Institut de la Vision, Sorbonne Universite, INSERM, CNRS

Reliably identifying cell types is crucial for understanding neural systems (Zeng & Sanes, 2017; Weis et al., 2023). However, classifying the neurons recorded in large-scale population recordings into well-defined types remains a challenge. While some methods rely on genetic targeting or anatomy for classification, another approach employed in sensory systems is to analyze the responses of neurons to stimuli that elicit discriminable responses in different types (Vlasits et al, 2019; Baden et al., 2016). Previous approaches of this nature have been limited by a restricted set of stimuli and scalability issues. In this study, we present a novel, end-to-end approach for the functional classification of retinal ganglion cells (RGCs), the retinal output, that combines recording responses to standardized stimuli with modeling using Convolutional Neural Networks (CNNs). Initially, we recorded the responses of a large set of ganglion cells to a novel noise stimulus that incorporated multiple spatial frequencies (MSF). These data were then used as a training set to learn and predict how each recorded neuron would respond to a variety of standardized stimuli, which can be used to determine cell types. Notably, previous research has demonstrated that surround properties can distinguish many cell types (Farrow & Roska, 2013; Goetz et al., 2022). Our method successfully predicted surround suppression in several cell types and significantly outperformed previous methods used for this purpose. Therefore, our approach serves as a powerful tool for determining RGC types in large-scale recordings. While classifications based on transcriptomic analysis (Tran et al., 2019), anatomical features (Bae et al., 2020), and functional characteristics (Baden et al., 2016) have been proposed in the retina, our method suggests a general approach for performing classification in population recordings by comparing the models learned for each cell.

2-115. From microcircuits to behavior: dense putative monosynaptic connections in the zebra finch HVC

Jeong Woo Kim[1,2] JK7334@NYU.EDU
Margot Elmaleh[3] MARGOT.ELMALEH@BRAIN.MPG.DE
Ellie Hozhabri ELLIE.HOZHABRI@GMAIL.COM
Michael A. Long[4] MICHAEL.LONG@NYULANGONE.ORG

[1]New York University
[2]Neuroscience Institute
[3]Max Planck Institute for Brain Research

[4]New York University Langone Medical Center

Neural sequences have been observed in a wide range of brain regions across species and are central to many behavioral processes. However, we have relatively little understanding of the cell-type specific wiring that generates such sequential activity. To gain traction on this issue, we investigate HVC (proper name), a cortical structure in the zebra finch brain where premotor neurons fire in a stereotyped sequence during song performance. We analyzed the population activity of neurons from HVC of behaving zebra finches using high-density silicon probes. We found that a large fraction of spike train cross-correlograms (CCGs) between well-isolated pairs had a significant peak or dip at a short latency (~1 ms), likely representing putative excitatory (E) or inhibitory (I) monosynaptic connections, respectively. This analysis provides the opportunity to uncover cell-type connectivity motifs underlying the precise sequential activity of HVC, and we use tools from graph theory to further explore these connections. In our preliminary analysis, we found that some single interneurons receive many inputs from and project to many local excitatory cells, consistent with 'hub' neurons that have been described in other systems. In addition, we observed a higher-than-expected number of reciprocally connected E and I neuron pairs, highlighting a role for fast inhibitory feedback within this network. By uncovering such specific connectivity motifs, we can decompose a complex network into building blocks, enabling cross-species/behavioral comparisons.

2-116. Matching the spatial properties of V1 neuronal receptive fields improves robustness in CNNs

Ruxandra Barbulescu[1] RUXI.BARBULESCU@GMAIL.COM
Tiago Marques[2] TIAGO.MARQUES@RESEARCH.FCHAMPALIMAUD.ORG
Arlindo L. Oliveira[1] AML@INESC-ID.PT

[1]INESC-ID Lisboa
[2]Champalimaud Foundation

While some convolutional neural networks (CNNs) have achieved great success in object recognition, they struggle to identify objects in images corrupted with different types of common noise patterns. Recently, it was shown that simulating computations in early visual areas at the front of CNNs leads to improvements in robustness to image corruptions. Here, we further explore this result and show that precisely matching the distributions of receptive field (RF) properties found in primate V1 is key for this improvement in robustness. We built two variants of a model with a front-end modeling the primate primary visual cortex (V1): one sampling RF properties uniformly and the other sampling from empirical biological distributions. The model with the biological sampling is considerably more robust to image corruptions than the uniform variant (relative difference of 8.72%). While similar neuronal sub-populations across the two variants have similar response properties and learn similar downstream weights, their impact on downstream processing is strikingly different. This result sheds light on the origin of the improvements in robustness observed in some biologically-inspired models, pointing to the need of precisely mimicking the neuronal representations found in the primate brain.

2-117. Memory consolidation facilitated by burst-induced late-phase plasticity

Kathleen Jacquerie[1,2] KATHLEEN.JACQUERIE@GMAIL.COM
Danil Tyulmankov[3,4] DT2586@COLUMBIA.EDU
Pierre Sacre[1] P.SACRE@ULIEGE.BE
Guillaume Drion[1,5] GDRION@ULIEGE.BE

[1]University of Liege
[2]Neuromorphic engineering
[3]Columbia University
[4]Center for Theoretical Neuroscience
[5]Department of Electrical Engineering

How do alternating periods of learning and rest contribute to memory consolidation? While it is recognized that learning relies on synaptic plasticity triggered by the spiking activity correlation between neurons, the role of rest periods and their biophysical mechanisms remain elusive. In this work, we leverage the interaction between the brain state fluctuations, reflecting changes in neuronal excitability, and memory, relying on synaptic plasticity occurring at different phases. Our approach involves a neural network model capable of transitioning between learning periods characterized by fast low-amplitude oscillations, and rest periods marked by slower large-amplitude oscillations. At the neuronal level, it is characterized by biophysical neurons capable of switching between input-driven tonic firing and the less-explored collective bursting.

In our model, synapses exhibit calcium-based early-phase plasticity, as studied in previous work. Here, we propose a new additional burst-induced late-phase plasticity mechanism. During learning, the early-phase plasticity forms new memories, as traditionally observed. During rest, the early-phase plasticity resets, returning to its baseline set point. It provides a physiological trace to drive the late-phase plasticity facilitating memory consolidation.

Validating our model through a memory task utilizing the MNIST dataset, we demonstrate that switching from tonic to burst, combined with early- and late-phase plasticity enables the network to acquire new information while preserving existing memories. The collective bursting activity during rest, combined with late-phase plasticity, represents the generation of new postsynaptic proteins and morphological synapse changes (termed structural plasticity). We find that substituting rest with an additional learning period impedes memory consolidation, rendering it susceptible to noise.

These findings propose a potential biological mechanism for unsupervised memory consolidation during rest and explain how the brain balances synaptic homeostasis and memory processes. Moreover, they suggest the utility of incorporating rest periods into machine learning models, highlighting the importance of including collective bursting and structural plasticity.

2-118. Deciphering intermediate neural representations in whole-brain sensorimotor circuits

Shuhong Huang[1] SHUHONG.HUANG@TUM.DE
James E Fitzgerald[2] FITZGERALDJ@JANELIA.HHMI.ORG
Ruben Portugues[1] RUBEN.PORTUGUES@TUM.DE

[1]Technical University of Munich
[2]HHMI Janelia Research Campus

A sensorimotor circuit is a neural network that connects sensory inputs to motor outputs, allowing an organism to perform coordinated movement within its environment. Sensorimotor circuits, therefore, should contain both general-purpose sensory representations that encode the external world and motor-relevant representations that meet the demands of specific behaviors. Most neuroscience studies assume that motor-relevant representations first emerge in the motor system, permitting sensory systems to focus on the general encoding of sensory environment. However, learning in artificial neural networks relies upon sequential nonlinear processing that gradually sculpts input representations into the output. Whether the brain similarly contains intermediate sensory representations that are best understood in terms of the statistics of the motor outputs is largely unknown. Here we develop and demonstrate a theoretical framework for comparing how well neuronal activity matches the statistics of the representations at sensory and motor peripheries. First, in silico experiments with this framework validate its ability to measure the stage of each hidden unit within the entire transformation of a trained artificial neural network. We then focus on analyzing visuomotor stabilization behaviors in zebrafish larvae. Using in vivo whole-brain imaging, our results indicate that some midbrain retina-targeted regions accurately capture the stimulus correlations needed for motor coding (e.g., pretectum). This motor reshaping of visual stimuli is specific to behaviorally indicated brain regions, and it is absent from the optic tectum's general purpose visual representation. Our framework further reveals the that the representations of hindbrain and telencephalon are more closely associated with motor output compared to midbrain. Overall, our framework presents a novel and general perspective for understanding sensorimotor transformations, providing fresh insight into the complexities of neural processing.

2-119. Higher-order thalamic contributions to flexible spatial navigation

Xintong (Cindy) Yuan[1,2] XYUAN@G.HARVARD.EDU
Joshua Stern[1] STERNJ@G.HARVARD.EDU
Justin Bucalo[1] BUCALO.J@NORTHEASTERN.EDU
Christopher Harvey[1] CHRISTOPHER_HARVEY@HMS.HARVARD.EDU

[1]Harvard Medical School
[2]Neurobiology

Spatial navigation in dynamically changing environments requires behavioral flexibility. Previous studies have characterized the representation of navigation tasks in cortex and found conjunctive coding of sensory, movement, and decision-related variables throughout cortical areas. Beyond cortex, many subcortical areas likely contribute to decision-making during navigation. In particular, the posterior parietal cortex (PPC) and prefrontal cortex (PFC) are heavily interconnected with higher-order thalamic nuclei LP (lateral posterior) and MD (mediodorsal), respectively. However, the activity patterns and contributions of MD and LP during flexible navigation have yet to be characterized.

We trained mice to perform a rule-switching task in virtual reality in which the rewarded cue-action mapping switched multiple times within a session without explicit signals. Well-trained mice were able to adapt behaviorally to the rule switch, and we recorded their cortical and thalamic activity simultaneously using Neuropixels 1.0. To understand the relationship between neural activity and task variables, decision dynamics and movement, we used a Generalized Linear Model (GLM) to extract encoding profiles for individual neurons. We find that LP and MD show distinct encoding profiles during the rule-switching task, which are also different from those of the posterior cortices. Specifically, LP has strong encoding of visual cues and choice in the maze, while MD shows the strongest outcome encoding during periods with reward and inter-trial-interval (ITI). Our results therefore suggest that LP is involved in the sensory-motor interface to convert visual cues into behavioral choices, whereas MD is important for evaluating the correctness of actions to potentially update internally stored task rules. Optogenetic inhibition experiments are currently underway to assess if MD and LP activity are causal for performance during different task epochs. Our results provide a first step towards a comprehensive understanding of how cortical and higher-order thalamic nuclei may function together in a network to enable flexible behavior.

2-120. Memory storage and retrieval in recurrent neural networks with behavioral timescale synaptic plasticity

John Briguglio[1] BRIGUGLIOJ@JANELIA.HHMI.ORG
Yiding Li[2] YIDING.LI@BCM.EDU
Jeffrey Magee[2] JEFFREY.MAGEE@BCM.EDU
Sandro Romani[3] ROMANIS@JANELIA.HHMI.ORG

[1]HHMI Janelia Research Campus
[2]Baylor College of Medicine
[3]Janelia Research Campus

One-shot learning is critical for survival, allowing animals to minimize exposure to danger, remember locations of food stores, and return home from a newly explored place. Hippocampus is known to be involved in the formation and recall of episodic memories, often memories of events experienced only once, but the macroscopic properties of memories inherited from the synaptic plasticity rules in the brain area remains poorly understood. It has been recently discovered that behavioral timescale synaptic plasticity (BTSP) in the CA3 hippocampal region is temporally symmetric and only operates at CA3 recurrent connections. This region has extensive recurrent connectivity and is thought to be important for the rapid storage of memories. One implication of the temporal symmetry is that it could allow the formation of stable attractor states in recurrent neural networks. We analytically studied the properties of a memory system endowed with BTSP using a recurrent neural network with binary activity patterns in discrete time with a simplified plasticity rule that captures the fundamental aspects of BTSP. We demonstrated that the memory storage capacity of the network scales optimally with population size, the network stabilizes patterns that are robust to temporal correlations in the activity while simultaneously improving pattern separation capabilities, and we elucidated the structural motifs underlying the memories formed in the network. Thus, we established that BTSP has optimal properties for forming stable attractors in a one-shot learning paradigm and provided predictions for the macroscopic structure of memories formed by BTSP.

2-121. State-dependent population dynamics control the speed and stability of sensory encoding in mouse V1

Edward Horrocks[1,2] EABHORROCKS@GMAIL.COM
Fabio Rodrigues[1] FABIO.RODRIGUES@UCL.AC.UK
Aman Saleem[1] AMAN.SALEEM@UCL.AC.UK

[1]University College London
[2]Institute for Behavioural Neuroscience

Sensory inputs can vary at millisecond timescales and neural activity must operate at similar timescales to encode these inputs. This is particularly true during active behavioural states, where sensory inputs change faster and behavioural demands are more immediate. Yet state-dependent neural coding has largely been investigated in seconds-long time windows, leaving fundamental gaps in our understanding of real-time sensory encoding. We therefore investigated state-dependent neural population responses to visual stimuli at fast, behaviourally-relevant timescales (10s of milliseconds), by recording 100s of neurons simultaneously with Neuropixel 2.0 probes (n=5 mice; 235-464 'good' units/session).

Locomotion reshaped single-neuron response dynamics, primarily by reducing transient stimulus onset responses. These changes were consistent across electrophysiologically-defined cell-types and analysis of the Allen Institute's 'Visual Coding' dataset revealed comparable changes across visual cortical areas. Reshaped response

dynamics enabled tuning for visual speed to emerge twice as quickly during locomotion, resulting in faster decoding of visual speed. Population temporal dynamics reorganised during locomotion such that population activity stabilised faster. Visual stimuli triggered a reorganisation of pairwise correlations which re-stabilised faster during locomotion. Consequently, the optimal linear decoding readout for visual speed stabilised faster. Changes in the dynamics of noise correlations also increased the efficiency of population encoding by reducing noise correlations in the direction of signal correlations.

Latent population trajectories were dominated by oscillatory dynamics during stationary states which were dampened during locomotion. Moreover, population trajectories exhibited reduced neural tangling during locomotion, indicating that mouse V1 transitions into a more stable dynamical system during active behavioural states. As a result of these changes in temporal dynamics, population trajectories made more direct transitions between baseline and stimulus-encoding steady-states during locomotion.

Our study establishes a novel principle of how population coding adapts to behavioural state, wherein changes in temporal response dynamics control the speed and stability of sensory encoding.

2-122. Trans-saccadic integration for target recognition peters out with pre-saccadic target eccentricity

Junhao Liang[1] JH.LIANG93@GMAIL.COM
Li Zhaoping[1,2] LI.ZHAOPING@TUEBINGEN.MPG.DE

[1]Max Planck Institute for Biological Cybernetics and University of Tubingen
[2]Sensory and Sensorimotor Systems

Bringing a visual object at a peripheral visual location to fovea by a saccade helps recognize this object. Human observers can integrate pre- and post-saccadic visual inputs for the recognition. To our knowledge, this integration has only been studied by an instructed saccade to a target at a prescribed and known location. Furthermore, the target is typically the only meaningful object in visual display, and the post-saccadic viewing duration of the target is often fixed by experimental design. For the first time, we study trans-saccadic integration in visual exploratory behavior, when observers decided themselves when and to which locations to make their saccades. We ask whether the pre-saccadic perception may be too ambiguous to contribute to the trans-saccadic integration when the pre-saccadic target is too peripheral from fovea in a crowded scene. To answer this question, we study the trans-saccadic integration in visual search behavior to find and report as soon as possible a target among 404 non-targets. Distributed locations of the search items, each 1.32° × 0.6° in visual angle, spanned 57.3° × 33.8° in visual angle. We measured (1) the pre-saccadic retinal eccentricity, e, of the target, and (2) the post-saccadic foveal viewing duration, T, which ended when the target was reported. We found that T increased with increasing e and eventually saturated at e around 10°–20°. We model our data by a decision-making process that integrates the pre-saccadic peripheral visual inputs with post-saccadic central visual inputs. Due to crowding, the pre-saccadic contribution to this integration diminishes with e. Meanwhile, the post-saccadic contribution increases with T. To reach a threshold for decision-making, our model predicts the experimental observation that trans-saccadic integration peters out with increasingly peripheral pre-saccadic object. This observation should apply to general visual exploratory behavior, at least when object perception is highly vulnerable to crowding.

2-123. Visual generalization from one exemplar in mice

Miguel Nunez[1,2] NUNEZM@HHMI.ORG
Fengtong Du[1] FENGTONGD@HHMI.ORG
Lin Zhong[1] ZHONGL@HHMI.ORG
Scott Baptista[1] BAPTISTAS@HHMI.ORG
Carsen Stringer[1] STRINGERC@HHMI.ORG
Marius Pachitariu[1] PACHITARIUM@HHMI.ORG

[1]HHMI Janelia Research Campus
[2]Mechanistic Cognitive Neuroscience, Computation & Theory

Visual object recognition relies on high-level visual features to identify objects despite changes in viewpoint, deformations, and other transformations. Neurons that encode invariant, high-level features have been found along the primate ventral stream, but it is not known how these invariances are obtained at the circuit level, whether they are innate or learned from experience, and how these neurons are used to support generalization behaviors. Here we investigated these properties in the mouse visual system and trained mice to perform behaviors that required visual generalization. Using simultaneous recordings of over 50,000 neurons from primary (V1) and higher-order visual areas (HVA), we found that the areas encoding invariant visual features were the medial HVAs, rather than the lateral HVAs as expected from its similarities to the primate ventral stream. Furthermore, we found that the

most superficial cortical neurons achieved the highest invariance, which we hypothesize is due to their proximity to the dense horizontal processes in the most superficial parts of layer 2/3, allowing neurons to integrate information from more input neurons and over larger areas. Finally, we found that equivalent invariance properties already existed in dark-reared mice, suggesting an innate origin for the high-level visual features. Next, we asked whether mice could perform tasks that require high-level visual invariance, such as generalization from a single learned exemplar to an entire visual category. We found that while mice always learned to distinguish individual texture exemplars, their ability to generalize varied according to the specific category pair used. Further, the amount of behavioral generalization was directly related to the neural invariance in visual areas on a stimulus-by-stimulus basis. Our findings establish mice as a viable model for studying high-level visual invariance, opening avenues for employing diverse neuroscience techniques to uncover the complex computations giving rise to invariance.

2-124. Learning cortical hierarchies with local, calcium-dependent synaptic plasticity

Sander de Haan[1] KIANDERDEHAAN@GMAIL.COM
Pau Vilimelis Aceituno[1] PAU@INI.ETHZ.CH
Reinhard Loidl[1] RLOIDL@STUDENT.ETHZ.CH
Benjamin Grewe[2] BGREWE@ETHZ.CH

[1]Institute of Neuroinformatics, University of Zurich and ETH Zurich
[2]ETH Zurich

The mammalian neocortex possesses a remarkable ability to translate complex sensory inputs into abstract representations by processing them in a hierarchical manner. While the hierarchical architecture of the neocortex is well described in terms of sensory bottom-up processing and top-down feedback, how learning is orchestrated across the spatial scales ranging from large neuronal networks to neurons and their individual synapses is unknown. To address this gap we start at the lowest level by modelling the membrane potential and calcium dynamics of individual pyramidal neuron synapses. We complement our model with a molecular-scale calcium dependent plasticity rule which comprehensively details the effects of various synaptic calcium levels on plasticity [Graupner & Brunel (2012)]. We then extend our synaptic model with somatic and apical dendritic compartments to integrate the diverse intracellular somatic and dendritic processes that affect synaptic dynamics and plasticity. We implement somatodendritic coupling via backpropagating action potentials, dendritic calcium spikes, as well as bursting. We show that, when working together, these mechanisms can supervise plasticity at the basal synapses by relating inputs arriving at the apical dendrite. We validate our model through, in vitro electrophysiology experiments in L5 pyramidal neurons of the mouse neocortex. Employing bio-plausible architectures for deep learning, we then finally demonstrate on a standard computer vision benchmark that our neuron model is suited for hierarchical network learning, and evaluate how different proposed bio-plausible architectures are compatible with our model and experiments. Our interdisciplinary work bridges the gap between neuroscience and deep learning, offering a comprehensive perspective on cross-scale mechanisms and their integration to enable hierarchical learning.

2-125. The unified framework for multi-dimensional distributional neural learning

Daniel McNamee DANIEL.MCNAMEE@RESEARCH.FCHAMPALIMAUD.ORG
Joseph Paton JOE.PATON@RESEARCH.FCHAMPALIMAUD.ORG
Margarida Sousa MARGARIDA.SOUSA@RESEARCH.FCHAMPALIMAUD.ORG

Champalimaud Research

Efficient neural coding (ENC), a long-established predictive framework for computational neuroscience, has been re-energized by novel theoretical developments and high-dimensional neural population recordings. In particular, recent work has generalized classical analyses for single units sensitive to one-dimensional stimulus features, to populations of neurons encoding multi-dimensional stimulus spaces. However, these theories only describe the offline optimization of efficient population codes and lack a model of learning in this context. In a distinct area of computational neuroscience overlapping with machine learning, many advances in state-of-the-art reinforcement learning (RL) are variants of the novel distributional RL (DistRL) suite of algorithms for which there is evidence of its implementation in the brain. DistRL facilitates the online learning of distributional population codes (each "unit" corresponding to a quantile of the reward distribution) however only for one-dimensional feature spaces (i.e. reward magnitude). We contribute a unifying framework that integrates efficient coding theory and distributional RL as special cases and combines the best of both of these theoretical paradigms. Algorithmically, our novel framework results in online learning equations for dynamically updating multi-dimensional distributional popula-

tion codes, stimulus sample by stimulus sample. We show that these population codes converge on the globally optimal (efficient) codes asymptotically and analytically characterize the non-asymptotic finite-sample suboptimality associated with learning variability. Finally, we demonstrate our theory in the RL context by learning a distributional code for a non-factorizable reward distribution which enables a simple softmax layer to readout dynamical behavioral strategies in an environment with complex reward dynamics. More generally, we suggest that our unified theory provides a complete framework for learning optimal neural distributional population codes in the brain.

2-126. Inferring ring-attractor structure in the zebrafish head-direction system from single-cell activity

Siyuan Mei[1] MEISIYUAN1@GMAIL.COM
Hagar Lavian[2] HAGAR.LAVIAN@TUM.DE
You Kure Wu[2] YK.WU@TUM.DE
Martin Stemmler[3] STEMMLER@BIOLOGIE.UNI-MUENCHEN.DE
Ruben Portugues[4] RUBEN.PORTUGUES@TUM.DE
Andreas V. M. Herz[3] HERZ@BCCN-MUNICH.DE

[1]LMU Munich
[2]Institute of Neuroscience, TUM
[3]Faculty of Biology, LMU; Bernstein Center for Computational Neuroscience Munich
[4]Technical University of Munich

Head direction (HD) cells, found in both vertebrate and non-vertebrate species, fire when the animal is facing a particular direction in the horizontal plane (1). Arranging cells with different preferred HDs onto a circle results in a bump-like activity profile across the population. When an animal turns its head, the activity bump rotates around the circle, providing a continuously updated representation of the head direction. The persistence of the activity bump in the dark and the maintenance of correlations across HD cells even when the HD cells' preferred directions undergo global remapping support the hypothesis that recurrent connections among HD cells generate a continuous ring-attractor (2,3), with one attractor state corresponding to one HD. A key feature of the hypothesis is that the attractor state is not unique but can be shifted easily and continuously to other equivalently stable states that represent different HDs. In Drosophila, there is strong evidence for the existence of a three-ring HD system with distinct bumps for each ring that interact in a push-pull fashion to update the representation of HD (3). But it has remained an open question whether the zebrafish HD system can similarly be subdivided into multiple, potentially intermingled rings or whether it consists of one ring only. Here, we develop a method to infer the ring structure from the movement-dependent HD cell activity, test its validity on simulated data, and re-analyze zebrafish data (4). We provide evidence that, overall, the functional connectivity of the HD system is consistent with a three-ring structure, similar to what has been reported for flies. This finding suggests that surprisingly similar mechanisms underlie the representation of angular orientation in species as distinct as flies and fish.

2-127. Contribution of Prefrontal Neural Dynamics to Reward History Integration and Strategy Selection

Zsombor Ungvarszki[1] UNGVARSZKI.ZSOMBOR@GMAIL.COM
Anna Szekely[2,3] SZEKELY.ANNA.95@GMAIL.COM
Gergo Orban[4,5] ORGERGO@GMAIL.COM
Matteo di Volo[6] MATTEO.DI-VOLO@UNIV-LYON1.FR
Emmanuel Procyk[6] EMMANUEL.PROCYK@INSERM.FR

[1]Inserm, Stem Cell and Brain Research Institute; Universite Lyon 1
[2]Wigner Research Centre for Physics // Budapest University of Technology and Economics
[3]Department of Computational Sciences // Department of Cognitive Science
[4]Wigner RCP
[5]Dep Computational Sciences
[6]Inserm, Stem Cell and Brain Research Institute

Flexible decision making involves complex coordination of cognitive processes to decide, act, detect outcomes, and adapt response policies. Prefrontal cortical areas, in particular the midcingulate cortex (MCC) and lateral prefrontal cortex (LPFC) were shown to contribute differentially yet coherently to these processes: neural substrates to seek new information were identified in the LPFC and MCC and shifting between different behavioral strategies was identified in switches between different neural states in the MCC. This work aims to understand the neural basis of recruiting competing behavioral strategies. Macaques performed a stochastic 3-armed bandit

task, while we simultaneously recorded population activity in LPFC and MCC. Critically, efficient harvesting of rewards required the integration of the history of rewards across trials. Behavioral analysis revealed that monkeys monitored their performance by computing Strategy value, a recency-weighted average of the previous outcomes, as shown by accurate prediction of strategy switching (between exploration and exploitation). We found that the Strategy value calculated from behavior can predict neural activity: linear decoding revealed strong neural correlates of the Strategy value in MCC populations, while this was largely absent from LPFC. By projecting the high-dimensional neural activity to the subspace defined by the decoder normal vector, we found that switching between strategies could be predicted by the activity in this low-dimensional subspace. Similar to the behavior, the neural correlate of the Strategy value reflected a weighted history of rewards and not only the most recent trial outcome. PCA analysis revealed that activity representing Strategy value emerges in MCC shortly before action selection, highlighting its role in decision making, and vanishes shortly after the feedback signal, indicating its role in the temporal integration of environmental information. Taken together, our analyses highlight the role of MCC neurons and their population dynamics during the decision making processes in higher-level cognition.

2-128. Neuron-level prediction, adaptation, and noise can implement reward-seeking behavior

Chenguang Li[1,2] CCLI.3896@GMAIL.COM
Adam Boesky[1] ABOESKY@COLLEGE.HARVARD.EDU
Jonah Brenner[1] JBRENNER@COLLEGE.HARVARD.EDU
Gabriel Kreiman[3] GKREIMAN@GMAIL.COM

[1]Harvard University
[2]Biophysics
[3]Harvard Medical School

Predictive coding plays an important role in nervous systems. But for prediction to serve as a unifying principle for neural computations, it must be linked not only to learning but also to flexible reward-seeking behavior. Prior work has attempted to make this connection, but there does not yet exist a self-contained predictive neural network model that can 1. seek and exploit reward, 2. balance exploratory and exploitative states, 3. learn associations, 4. adjust to environmental changes, 5. switch between goals in a self-directed manner, 6. act autonomously, that is, with no external control signals such as those required in standard reinforcement learning algorithms, and 7. do all of the above with local learning rules. These are behavioral features that even animals with only a few hundred neurons possess but which are still challenging for many machine learning algorithms. Thus, a model that can achieve these desiderata is an important open problem in computational neuroscience. Here we show that by combining predictive coding with two known physiological mechanisms—adaptation to sensory inputs and noise at the neuronal level—a model exhibits these desired behaviors. We test our model in simple reinforcement learning environments, showing how exploration frequency changes with noise, and how adaptation allows for reward maximization and goal-switching. The model's ability to simultaneously accomplish the desired biologically relevant behaviors using only prediction and physiologically plausible mechanisms makes it a novel normative model of intelligent computations in small organisms.

2-129. Sequential predictive learning accounts for hippocampal representation and replay

Daniel Levenstein[1] DANIEL.LEVENSTEIN@MILA.QUEBEC
Aleksei Efremov[1] ALEKSEI.EFREMOV@MILA.QUEBEC
Roy Eyono[2] ROY.EYONO@MILA.QUEBEC
Blake Richards[1] BLAKE.RICHARDS@MILA.QUEBEC
Adrien Peyrache[1] ADRIEN.PEYRACHE@MCGILL.CA

[1]McGill University
[2]Mila

The mammalian hippocampus represents an animal's position in the environment during active behavior, and generates "replay" simulations of plausible trajectories in the environment during offline periods of behavioral quiescence. While continuous attractor network models can produce both online representation and offline simulations, they require specific wiring between units with pre-assigned spatial locations or learning from signals with pre-existing spatial tuning, and it's unclear how such a network can be learned from sensory information alone. Recently, it's been found that recurrent neural networks trained to predict sensory inputs develop spatially tuned cells, aligning with predictive theories of hippocampal function. However, whether predictive learning can also account for the replay capacities of the hippocampus is unknown.

Here, we show that learning to predict sensory inputs can produce hippocampal-like representation and replay. However, we find that spatially tuned cells, which robustly emerge from all forms of predictive learning, do not guarantee the ability to produce offline simulations. Offline simulations only emerged in networks that used recurrent connections and head-direction information to predict multi-step observation sequences, which promoted the formation of a continuous attractor manifold reflecting the spatial geometry of the environment. When networks were trained to predict future observation sequences, they were able to rapidly learn cognitive maps, reactivate recently explored locations, and form new position-object associations from a limited number of exposures. This produced representations that cyclically sweep from immediately behind to multiple steps ahead of the agent, reminiscent of hippocampal theta sweeps. Our results demonstrate how continuous attractors can emerge in neural networks engaged in sequential predictive learning, and suggest that hippocampal theta sequences reflect a data-efficient algorithm for rapid sequential predictive learning. Together, this provides a unifying theory for hippocampal physiology, hippocampal functions, and hippocampal-inspired approaches to artificial intelligence.

2-130. Single-cell optogenetics reveals attenuation-by-suppression in visual cortical neurons

Paul LaFosse[1] LAFOSSEPK@NIH.GOV
Zhishang Zhou[1] ZHISHANG.ZHOU@NIH.GOV
Victoria Scott[1] VICTORIA.SCOTT@NIH.GOV
Yanting Deng[1] YANTING.DENG@NIH.GOV
Mark Histed[2,3] MARK.HISTED@NIH.GOV

[1]National Institute of Mental Health
[2]NIH
[3]NIMH

A central aspect of brain computation is the relationship between neurons' input and spiking output. Theoretical work confirms the shape of input-output (activation) functions is important to network operation, impacting both learning efficiency and performance of artificial neural networks (Hendrycks and Gimpel, 2016; Moskovitz et al., 2018). To understand how neurons' input-output (IO) functions affect network function, it is essential to measure during awake states, because IO function nonlinearities can emerge as ongoing network activity changes (Anderson et al., 2000; Hansel and van Vreeswijk, 2002; Miller and Troyer, 2002). In addition to IO function shape, the operating point of neurons on these curves during the awake state is not well known.

Here, we characterize cortical principal neurons' average input-output functions in awake mice using two-photon optogenetic stimulation and imaging. We optogenetically stimulated neurons during spontaneous activity and while neurons' activity was increased or suppressed by visual stimuli, and measured differences in responses to the same optogenetic input between conditions. We find responses to optogenetic input are unchanged as neurons are excited, reflecting a linear response regime. In contrast, responses are strongly attenuated by suppression.

Thus, neurons in vivo operate with a supralinear-to-linear IO function. There is only weak saturation at the top of neurons' operating range, with strong visual stimuli. The function we measure is best described not by a ReLU or power law (as in the stabilized supralinear network, Ahmadian et al., 2013), but by the Riccardi function created by integrate-and-fire networks (Sanzeni et al., 2020). Also, in the awake state, neurons rest near the top of a supralinearity in their IO function. This supralinearity means that as cells are suppressed, their responses to input are dramatically decreased. Our results thus show a mechanism for how single neurons can act to filter inputs as their firing rates change: attenuation-by-suppression.

2-131. Shared articulatory representations drive a bilingual speech neuroprosthesis

Alex Silva[1,2] ASILVAALEX4@GMAIL.COM
Edward F. Chang[1] EDWARD.CHANG@UCSF.EDU
Jessie Liu[3] JESSIE.LIU@BERKELEY.EDU
Sean Metzger[3] SEAN.METZGER@UCSF.EDU
Ilina Bhaya-Grossman[3] ILINA.BHAYA-GROSSMAN@UCSF.EDU
Maximilian Dougherty[3] MAXIMILIAN.DOUGHERTY@UCSF.EDU
Margaret Seaton[3] MARGARET.SEATON@UCSF.EDU
Kaylo Littlejohn[4] KAYLO_LITTLEJOHN@BERKELEY.EDU
Adelyn Tu-Chan[3] ADELYN.TU@UCSF.EDU
Karunesh Ganguly[3] KARUNESH.GANGULY@UCSF.EDU
David Moses[3] DAVID.MOSES@UCSF.EDU

[1]University of California, San Francisco
[2]Neurosurgery
[3]University of California at San Francisco
[4]University of California at Berkeley

A core goal of speech neuroprosthetics is to restore naturalistic communication to all people living with paralysis. Over half of the world is bilingual with proficiency in at least two languages. However, advancements in decoding speech from the brain have focused on monolinguals, leaving generalizability to bilinguals an open question. Further, it is unclear whether unique or shared cortical activity underlies bilingual speech production in each language. This has important implications for basic neuroscience and training decoders that may rapidly generalize across languages, without significant additions of training data. Here, we leverage electrocorticography (ECoG) to directly record cortical activity in speech-motor cortex from a Spanish-English bilingual with vocal-tract and limb paralysis. Using recurrent neural networks, along with statistical natural-language models of English and Spanish, we flexibly decode cortical activity into bilingual sentences without the participant needing to manually select the target language. By exploring model confusability between different bilingual words, we demonstrate that bilingual decoding relies on shared articulatory representations. More specifically, we demonstrate a shared bilingual syllable representation on speech-motor cortex. Finally, we show that cortical activity recorded in one language can be used to improve decoding in the other language via transfer learning, expediting training of a bilingual decoder. Overall, this work leverages techniques in machine learning to demonstrate a multilingual cortical articulatory representation that persists after paralysis and enables flexible decoding of multiple languages.

2-132. Striatal dopamine reflects individual long-term learning trajectories

Samuel Liebana Garcia[1,2] SAMUEL.LIEBANAGARCIA@MERTON.OX.AC.UK
Aeron Laffere[1] AERON.LAFFERE@DPAG.OX.AC.UK
Chiara Toschi[1] CHIARA.TOSCHI@DPAG.OX.AC.UK
Louisa Schilling[1] LOUISA.SCHILL@GMAIL.COM
Jacek Podlaski[1] JACEK.PODLASKI@SJC.OX.AC.UK
Matthias Fritsche[1] MATTHIAS.FRITSCHE@DPAG.OX.AC.UK
Peter Zatka-Haas[1] PETER@PETERZH.COM
Yulong Li[3] YULONG@GMAIL.COM
Rafal Bogacz[1] RAFAL.BOGACZ@NDCN.OX.AC.UK
Andrew Saxe[4] A.SAXE@UCL.AC.UK
Armin Lak[1] ARMIN.LAK@DPAG.OX.AC.UK

[1]University of Oxford
[2]Department of Physiology, Anatomy and Genetics
[3]Peking University School of Life Sciences
[4]Gatsby Computational Neuroscience Unit & Sainsbury Wellcome Centre, UCL

Learning from naive to expert occurs over long periods of time, entailing extensive changes in the brain's neuronal signals. The principles governing behavioral and neuronal dynamics during long-term learning remain unknown. We developed a psychophysical visual decision task for mice that allowed for studying learning trajectories from naive to expert. Mice adopted sequences of strategies that became more stimulus-dependent over time, showing substantial diversity in the strategies they transitioned through and settled on. Despite this diversity, these transitions were systematic; the initial strategy of naive mice predicted their strategy several weeks later. Longitudinal imaging of dopamine release in dorsal striatum demonstrated that dopamine signals emerged over learning, reflecting stimulus-choice associations determined by each individual's strategy in using stimuli to make decisions. A deep neural network model trained on the task with reinforcement learning captured behavioral and dopamine

dynamics. The model's learning dynamics, derived analytically, accounted for the diverse and systematic learning trajectories of mice. The model's prediction errors, used to update its parameters, mirrored dopamine signals, offering a concrete account of striatal dopamine's role in long-term learning. Our results demonstrate that long-term learning is governed by diverse yet systematic transitions through behavioral strategies, and that dopamine signals exhibit key characteristics to support such learning.

2-133. Towards Generalizable Neural Decoding: Simultaneous Goal- and Data-driven Modeling of Motor Cortex

Muhammad Noman Almani[1,2] MUHAMMADNOMAN.ALMANI@YALE.EDU
Shreya Saxena[1,3] SHREYA.SAXENA@YALE.EDU

[1]Yale University
[2]Electrical Engineering
[3]Wu Tsai Institute

Can the interaction among functionally and anatomically distinct entities of the brain, body and environment help us understand and predict the neural dynamics and single-unit firing rates underlying movement generation? The brain is evolved to interact with the body and the environment under suitable neural constraints, such as energy-minimization, and behavioral constraints, such as physical laws governing the musculoskeletal dynamics, which existing goal- or data- driven models of the motor cortex (MC) lack [1-3]. Here, we develop a dynamical systems model of MC termed Musculo-RNN that is both goal- and data- driven: using a recurrent neural network (RNN) -based controller, it transforms sensory feedback into muscle excitations. The Musculo-RNN controls an anatomically-accurate musculoskeletal model to produce the resulting kinematics from muscle excitations, thus capturing the physical laws underlying musculoskeletal dynamics. We also implement high-level neural constraints, i.e., minimization of neural firing rates and simple low-dimensional neural trajectory evolution, that have been shown to govern neural dynamics [1, 4]. To further enable goal- and data- driven modeling, we constrain a subset of the RNN's nodes to the experimentally observed MC single-unit activity for training conditions. We show that Musculo-RNN trained using deep reinforcement learning (DRL) successfully reproduces both the experimental MC single-unit activity and the resulting kinematics, and is generalizable to unseen conditions that are significantly different than the training conditions (out-of-distribution generalization). The trained controller outperforms existing goal-driven or data-driven models of MC in single-unit encoding accuracy. Moreover, jPCA analysis reveals that the network dynamics of Musculo-RNN mimic the MC neural population dynamics uncovering the underlying neural strategies for task representation (distinct initial fixed-points) and movement generation (oscillatory dynamics). Lastly, ablation studies show that proprioceptive feedback plays the most important role in maintaining MC neural dynamics followed by visual feedback, suggesting its role in online control of movements [6].

2-134. Predictions enable top-down pattern separation in the macaque face-processing hierarchy

Tarana Nigam[1,2] TNIGAM@DPZ.EU
Caspar Schwiedrzik[3] C.SCHWIEDRZIK@ENI-G.DE

[1]German Primate Center
[2]Perception and Plasticity Group
[3]German Primate Center; European Neuroscience Institute, Gottingen

Telling people apart is crucial to thrive in a social environment & crucially depends on visual information derived from faces. Distinguishing faces constitutes a significant computational challenge that relies on well separable neural activity patterns which are also invariant to changes like pose. Our recognition abilities can be enhanced by additional context, e.g., from prior encounters. On a neural level, this contextual enhancement may be driven by predictive information that separates representations. Here, we investigate how predictions derived from learnt prior information affect the separability of neural population activity in the macaque-monkey face-processing system, a 3-level processing hierarchy in ventral visual cortex, using electrophysiology & functional magnetic resonance imaging. We find that in the presence of predictions, early stages of this hierarchy exhibit well separable neural geometries otherwise characteristic only of the top of the hierarchy. This is driven by an expansion of representational dimensionality & population correlations. These changes are accompanied by a systematic shift of the neural representational space from higher to lower areas, endowing lower areas with higher-order, invariant representations instead of their feedforward tuning properties. Thus, top-down signals dynamically transform lower-order neural representations of faces towards high-order, separable & high-dimensional neural geometries. Our results provide evidence how predictive context provided through feed-

back flexibly transform representational spaces to optimally use the computational resources provided by cortical processing hierarchies for better & faster distinction of facial identities.

2-135. Extraction and recovery of spatio-temporal structure in neural alignment via diffusion models

Yule Wang[1,2] YULEWANG@GATECH.EDU
Zijing Wu[1] ZWU381@GATECH.EDU
Chengrui Li[1,3] CNLICHENGRUI@GATECH.EDU
Anqi Wu[4] ANQIWU@GATECH.EDU

[1]Georgia Institute of Technology
[2]School of Computational Science and Engineering
[3]Computational Science & Engineering
[4]georgia institute of technology

In the field of behavior-related brain computation, it is crucial to align raw neural recordings against the drastic domain shift spanning sessions, subjects, and experimental settings. A foundational framework within neuroscience research posits that trial-based neural population activities rely on low-dimensional latent neural trajectories, thus focusing on these trajectories significantly facilitates the alignment process. However, despite the turnover and low signal-to-noise ratio of recorded neurons, the spatio-temporal structure of latent trajectories largely remains persistent across domains. Existing methods often overlook such intrinsic structure during the alignment phase, leading to poor preservation of latent trajectory structures and overall performance. To address this issue, our proposed alignment method ERDiff first leverages the expressivity of diffusion models to extract the spatio-temporal structure of latent trajectories in the source domain. Subsequently, under the guidance of the diffusion model, these structures are precisely recovered in the target domain through a maximum likelihood alignment process. Comprehensive experiments conducted on both simulated and real-world neural datasets demonstrate that ERDiff notably outperforms existing methods in alignment goodness-of-fit and downstream neural decoding tasks.

2-136. Neuronal mechanisms of flexible decision-making during foraging

Lyle Kingsbury[1,2] LYLEKINGSBURY@GMAIL.COM
Naoshige Uchida[1] NUCHIDA@MCB.HARVARD.EDU

[1]Harvard University
[2]Molecular and Cellular Biology

The ability to flexibly adapt cognition and behavior to different situations is a fundamental feature of animal intelligence. In the natural world, animals exhibit flexible behavior when foraging for resources, as food and water are often distributed differently across different environments, encouraging distinct foraging strategies. In systems neuroscience, there is some understanding of the basic mechanisms underlying simple perceptual and value-based decisions; but it is still unclear how neural circuits can be “reconfigured” across different contexts to implement flexible computational functions. Foraging provides an ecological window into the neural circuit mechanisms that support context-dependent decision-making and behavioral flexibility. In this study, we use a virtual foraging task to investigate flexible decision-making in mice, in which animals navigate an environment to find patches that give water rewards. While deterministic patches always yield a fixed reward sequence, stochastic patches give probabilistic rewards that diminish over time. Mice learn to adjust their decision process to the statistical structure of different environments, becoming more patient when expecting deterministic rewards. Building on previous theoretical and experimental work, we found that a flexible drift-diffusion model captures context-driven changes in decision strategy. High-density neural recordings show that different foraging contexts recruit distinct modes of brain activity in the dorsal frontal cortex, as well as neuronal populations that encode context-specific foraging decision variables.

2-137. Signatures of generalised spatial representations in frontal cortex

Adam Harris[1] ADAM.HARRIS2@MERTON.OX.AC.UK
Mohamady El-Gaby[1] MOHAMADY.EL-GABY@NDCN.OX.AC.UK
Ben Pendry[1] BEN.PENDRY@PSY.OX.AC.UK
Arya Bhomick[2] ARYA.BHOMICK.17@UCL.AC.UK
Mark E. Walton[1] MARK.WALTON@PSY.OX.AC.UK
Thomas Akam[1] THOMAS.AKAM@PSY.OX.AC.UK
Tim Behrens[3] TIMOTHY.BEHRENS@NDCN.OX.AC.UK

[1]University of Oxford
[2]University College London
[3]University of Oxford & University College London

Spatially tuned neurons have been reported in medial frontal cortex (mFC), a region critical for flexible behaviour. But the nature of spatial representations in frontal cortex is disputed, with evidence for both concrete and abstracted representational forms. Here we designed a behavioural paradigm in which behavioural structure and spatial structure are disentangled, and use a representational geometry approach to query the format of mFC spatial representations. We find that decomposing the neural state space into subspaces did not preserve the concrete spatial relationships between locations of the maze. Instead, the subspaces split locations by their abstract features: the centre, corners, and locations on the cardinal axes - with these location types being linearly discriminable in the discovered subspace. Additionally, patterns of selectivity for these abstract location types can be found in the firing rates of individual mFC neurons. When comparing the angles of subspaces for different behavioural step, we found no evidence of subspace orthogonality. Lastly, we present example place maps from neurons recorded as the physical environment is expanded by the introduction of new locations, thus changing the connectivity of the maze. These findings support a generalised form of spatial representations in mFC and will permit further understanding of the interactions between such abstracted spatial and behavioural representations in frontal cortex.

2-138. The split-trial analysis: efficient and reliable inference of information-limiting noise from neural population recordings

Dylan Le[1,2] DYLANLE@UTEXAS.EDU
Xuexin Wei[3,4] WEIXX@UTEXAS.EDU

[1]University of Texas Austin
[2]Institute for Neuroscience
[3]UT Austin
[4]Department of Neuroscience

Understanding the functional implications of correlated neural noise represents one of the central problems in neural coding. Recent theories suggest that shared noise along the stimulus encoding direction is the primary factor that limits information encoding (i.e. information-limiting noise, which leads to information-limiting noise correlations). While a crucial theoretical insight, this idea can be difficult to apply to experiments because of challenges in inferring information-limiting noise from neural data. To overcome this major limitation, we have developed a new method based on partitioning of the neural population. Specifically, we split the neurons into subpopulations, then perform decoding analysis on each subpopulation. Intuitively, the decoding error for each subpopulation contains both information-limiting noise and private noise along the encoding direction that is specific to that sub-population. We devise an appropriate statistical procedure to infer the shared noise, which serves as an estimator of the information-limiting noise. Notably, our method goes beyond popular single-trial analysis because it crucially relies on the splitting of a single-trial to create repeated measurements. We thus refer to our method as the “split-trial analysis". We have performed theoretical analyses and extensive numerical validation using synthetic datasets. We find that our method has several key advantages: (i) compared to prior methods, the split-trial analysis achieves a given accuracy with a much smaller number of neurons; (ii) it is reliable across a wide range of parameter regimes; (iii) importantly, it does so without directly estimating the noise covariance matrix, which is a major hurdle to existing techniques. We further apply our methods to publicly available datasets recorded from rodent Hippocampus to infer the spatial encoding precision during navigation. Our method is simple yet general, and should be generally applicable for understanding neural coding in many brain regions

2-139. Integrating actions and their sensory consequences relative to an internal goal in larval zebrafish

Emanuele Paoli[1,2] EMANUELEP.EP@GMAIL.COM
Virginia Palieri[3] VIRGINIA.PALIERI@RESEARCH.FCHAMPALIMAUD.ORG
Ruben Portugues[1] RUBEN.PORTUGUES@TUM.DE

[1]Technical University of Munich
[2]Institute of Neuroscience
[3]Champalimaud Research, Champalimaud Centre for the Unknown

To navigate effectively, animals must interpret sensory information relative to internal goals and past choices. This integration process shapes adaptive behavioral strategies, with successful decisions likely to be repeated. However, how the vertebrate brain integrates actions and their consequences in a navigational context remains unclear. To address this gap, we use larval zebrafish thermal navigation as a tractable system. As a cold-blooded animal, zebrafish rely solely on navigational strategies to reach their preferred temperature, with temperature changes, sensed during swim bout events, used as the primary navigational cue. We found that the mapping between temperature cues and actions is dynamically adjusted based on experience. Larval zebrafish integrate the action identity with its sensory consequences relative to their internal goal. If a specific action previously brought the fish closer to its preferred temperature, that choice will likely be repeated in the future when the animal faces the same conditions. Our closed-loop functional imaging experiments in the Interpeduncular Nucleus (IPN) revealed neurons that, in a neutral context, exhibited stereotyped activity patterns in response to actions with a slow time constant. Interestingly, actions were simultaneously mapped onto two IPN sub-compartments (shell and core) with inverted orientation. When thermal cues were introduced in a mildly aversive temperature environment, neurons in the shell sub-compartment selectively tracked actions relative to the animal's internal goal. Finally, we will discuss how this code is further processed in the IPN and the serotonergic Raphe Nucleus to guide adaptive behavioral strategies.

2-140. Flexible and generalizable representations of cognitive maps

Sarah Sweigart[1,2] SSWEIGART@UCDAVIS.EDU
Nam Nguyen[1] NAANGUYEN@UCDAVIS.EDU
Charan Ranganath[1] CRANGANATH@UCDAVIS.EDU
Seongmin Park[3] APARK@UCDAVIS.EDU
Erie Boorman[1] EDBOORMAN@UCDAVIS.EDU

[1]University of California, Davis
[2]Psychology
[3]Institute of cognitive science, CNRS

Cognitive maps allow the brain to infer new relationships and support flexible decision making by efficiently representing multidimensional relationships in a map-like structure. While the entorhinal cortex (ERC), hippocampus (HPC), and medial prefrontal cortex (mPFC) have all been implicated in containing cognitive map representations, their respective roles and contributions are still unclear. We designed a novel 'wine space' task to test to what extent these regions represent a context-dependent (specific and flexible) code or context-invariant (abstracted and generalizable) code. Participants learned a 2D wine attribute space and then used this knowledge to flexibly select the best wine for different "market" contexts while fMRI data was collected. Blood oxygen-level-dependent (BOLD) activity for a context-dependent decision-relevant rank difference was identified in the mPFC and HPC, pointing to the flexible use of task-relevant relational values. Whole-brain representational similarity analysis (RSA) of activity elicited by wine stimulus presentations revealed a specific context-relevant rank representation in the mPFC. Interestingly, further analyses revealed both a context-invariant 2D Euclidian representation of wine "position" in the 2D attribute space in the ERC, and an abstracted context-dependent rank "position" representation that generalized over wines and contexts. Finally, we detected a 6-fold modulation of ERC activity with respect to decision vectors over the 2D wine attribute space, consistent with a grid-like code for novel inferences between wines in the ERC. These findings suggest a transformation between abstract, generalizable representations in ERC to context-dependent representations for inference decisions in HPC and mPFC.

2-141. Value Pop-out Results from Spatial Enhancement of Object Processing in Prefrontal Cortex

Mojtaba Abbaszadeh[1,2] ABBASZADEH989@GMAIL.COM
Kiomars Sharifi[3] KIO.SHARIFI@GMAIL.COM
Ali Ghazizadeh[4] ALIEGHAZIZADEH@GMAIL.COM

[1]University of Montreal
[2]Department of Neuronscience, School of Medicine
[3]Institute for Research in Fundamental Sciences (IPM)
[4]Sharif University of Techonology & Institute for Research in Fundamental Sciecnes (IPM)

Recent evidence demonstrates that long-term object value association causes efficiency in visual search. This phenomenon known as value pop-out in which the first saccade goes to high-value (good) objects among varying low-value (bad) objects. Ventrolateral prefrontal cortex (vlPFC) neurons can detect the presence of over-trained good objects in less than 150ms. We do not know how long-term value training shapes the spatial tuning of vlPFC neurons. The question is how this early selectivity of good objects, following object-value association training, leads to search efficiency in vlPFC neurons. We hypothesized that this effect is due to expansion of spatial processing of good objects after over-training. To test this hypothesis, two monkeys participated in a value-based visual search task, wherein they determined the presence of a good target among a variable number of bad objects. The computation modeling results revealed an increase in the parameter representing the visual processing area, affirming the expansion of the processing area for good objects. Furthermore, our neural analyses indicated that the spatial tuning curve of vlPFC neurons was enlarged (having both a higher amplitude and being wider) leading to efficient detection of good objects. Together, our behavioral and neural results suggest that an enhancement of spatial processing of good objects might be a reason for value pop-out effect.

2-142. Tuned cell-type specific inhibition refines pattern completion in mouse V1

Ho Yin Chau[1] HC3190@CUMC.COLUMBIA.EDU
Mora Ogando[2] MORAOGANDO@BERKELEY.EDU
Lamiae Abdeladim[2] LABDELADIM@BERKELEY.EDU
Savitha Sridharan[2] SAVI.SRIDHARAN@BERKELEY.EDU
Karthika Gopakumar[2] KARTHIKAKARTHIK@BERKELEY.EDU
Silvio Temprana[2] STEMPRANA@BERKELEY.EDU
Hyeyoung Shin[2] SHINEHYEYOUNG@GMAIL.COM
Hillel Adesnik[2] HILLEL.ADESNIK@GMAIL.COM
Kenneth Miller[1] KENDMILLER@GMAIL.COM
Agostina Palmigiano[1] AP3676@COLUMBIA.EDU

[1]Columbia University
[2]University of California, Berkeley

The visual cortex is able to both “fill in" missing visual information, indicating pattern completion capabilities, while having distinct representations of almost identical images, suggesting pattern separation. Here we investigate how the mouse visual cortex performs these fundamental operations. We combined two-photon calcium imaging and patterned optogenetics for high-accuracy visual activity recreation without visual input, and developed novel recurrent network theory to model and interpret the data. With this toolkit, we optogenetically activated iso-tuned pyramidal (Pyr) cell ensembles of different sizes while recording the activity of Pyr cells, Parvalbumin-expressing (PVs) and Somatostatin-expressing interneurons (SSTs). Aligned with previous findings, we find net inhibition in cells surrounding the perturbed pattern and like-to-like connectivity between pyramidal cells, where similarly oriented cells are less suppressed.

Our theory allowed us to express the network’s response to holographic perturbations as a function of orientation and space analytically. By fitting these expressions to each cell-type calcium data, the model provided a key prediction: like-to-like SST->Pyr (but not PV->Pyr) connectivity. We tested this prediction by directly activating co-tuned SST ensembles, and found that indeed SST cells primarily suppressed co-tuned Pyr cells. We reasoned that, while the like-to-like connectivity between excitatory cells could serve for pattern completion, the inhibitory counterpart might play the opposite role. Indeed, pattern completion was observed in the data, as it was possible to decode visual stimuli from incomplete holographic stimulation. To test the role of like-to-like SST->Pyr connections in pattern separation and completion, we replaced tuned connections with untuned ones of the same strengths. This manipulation led to a pattern completion bias, which impaired the discriminability of similarly tuned inputs. These results suggest that ensemble-level input/output transformations in this network are shaped by two opposing forces: pattern completion and pattern separation carried out by excitatory and inhibitory sub-networks.

2-143. Rapid switching of cross-modal selective attention and its neural correlate in the basal forebrain

Szwen Liu[1,2] SZWENLIU.LS06@NYCU.EDU.TW
Shih-Chieh Lin[1] SHIHCHIEH.LIN@NYCU.EDU.TW

[1]National Yang Ming Chiao Tung University
[2]Insititute of Neuroscience

Cross-modal selective attention is the ability to focus attention on one sensory modality while ignoring stimuli from other modalities. This cognitive function is essential for animals and humans to survive in environments with multi-modal and competing streams of sensory inputs. Despite its importance, much remains unknown about how the focus of attention is rapidly switched between different sensory modalities, and whether this process involves subcortical neural circuits. Here we develop a novel cross-modal selective attention task in rats and show that the rapid and dynamic switching of attention between sensory modalities drives basal forebrain (BF) neurons to respond to sensory stimuli in the attended modality only. Rats were trained in a concurrent audio-visual oddball task, in which one auditory and one visual stimulus were simultaneously presented once every two seconds. The infrequent oddball stimuli in each modality were the potential targets, and the relevant sensory modality that predicted reward switched across trial blocks without overt warning. Animals learned to switch their behavioral response patterns toward auditory and visual oddball stimuli across trial blocks, and this switch occurred within a few trials of block transitions. Concurrently, a subset of BF neurons were selectively excited by the oddball stimulus in the attended sensory modality, while ignoring the same stimulus when the modality became irrelevant. BF responses to multi-modal stimuli were equivalent to their responses toward the single-modal stimulus in the attended modality. Shifts in BF response patterns were rapid and tightly coupled with behavioral switching during block transitions. Such coordinated behavioral and BF activity shifts were also observed during spontaneous shifts of attention toward the non-rewarded modality initiated by the animal. Together, these results establish a new approach to study the dynamic shifts of cross-modal selective attention at single trial resolution, and highlight the important role of BF neural circuits.

2-144. Geometry of relational knowledge in the macaque posterior parietal cortex

Somang Paeng[1] SOMANGPAENG@GMAIL.COM
Hansem Sohn[1] HANSEM@MIT.EDU
Mehrdad Jazayeri[2] MJAZ@MIT.EDU

[1]Sungkyunkwan University
[2]Massachusetts Institute of Technology

Abstract relational reasoning is key to many cognitive functions. However, it remains unknown how the brain represents abstract relational information and integrates it with sensory information to guide behavior. To address this question, we designed a relation-based memory task in which monkeys remember the spatial relation between a cue and a target, and later apply the relation to a new cue location to make a saccade to the final target location. Animals learned the task successfully and were also able to rapidly generalize to new combinations of relations and cue locations. Recent AI models have proposed that such rapid generalization requires a representation of abstract relations independent from sensory information (Kerg et al., 2022, Battaglia et al., 2018). We, therefore, hypothesized that spatial relations would be represented in an independent subspace from the spatial location subspace in the neural state space. We tested this hypothesis by analyzing Neuropixels recordings from area 7 of the posterior parietal cortex, which has been strongly implicated in relational reasoning (Lyu et al., 2022). Single neurons were highly heterogeneous and showed different degrees of selectivity to the cue locations, the cue-target relations, or both. However, the population activity had the requisite features needed for abstract relational reasoning. First, the memorized relation and the subsequent cue location were encoded within independent low-dimensional subspaces of population activity. Second, the embeddings within these subspaces retained the spatial layout of cues and cue-target relations. Together, these results indicate that abstract relational reasoning in the brain may rely on a factorized representation of spatial locations from spatial relations using independent low-dimensional subspaces, consistent with the AI models.

2-145. A unifying framework for neural computations of motion and depth by a moving observer

Brian Xu[1,2] BRIAN.ZX.XU@GMAIL.COM
Jiayi Pang[1] JPANG8@U.ROCHESTER.EDU
Akiyuki Anzai[1] AANZAI@UR.ROCHESTER.EDU
Gregory DeAngelis[1] GDEANGELIS@UR.ROCHESTER.EDU

[1]University of Rochester
[2]Brain and Cognitive Sciences

Humans and other animals constantly move in the environment. These movements impose a critical challenge for our brain: the visual consequences of our actions blend with those of objects moving independently in the world. Therefore, accurately perceiving the dynamic 3D environment requires the brain to infer the causes of visual inputs and compensate for self-motion (body and eye movements). Traditionally, this is thought to be a simple vector operation that subtracts the self-generated optic flow vector from retinal image motion. However, by analyzing the 3D projective geometry, we show that this idea of vector subtraction does not generalize to more naturalistic viewing scenarios. Here, we propose a geometric-derived framework that unifies the computations of motion and depth in various 3D viewing geometries. Specifically, we examine two well-known perceptual phenomena, coordinate transformation and depth from motion parallax, as special cases of our framework. Quantitative predictions made by our framework are confirmed by psychophysical evidence. Humans exhibit opposite patterns of biases in their perceived motion and depth when presented with optic flow that simulates different viewing geometries. Importantly, these perceptual biases occur naturally without any feedback or training. Electrophysiological recordings in nonhuman primates show that neural responses in the middle temporal (MT) area are differentially modulated by different viewing geometries simulated by large-field background motion. At the population level, these modulations enable better decoding performance of the corresponding perceptual variables in each simulated geometry. This indicates that the observed neural modulations in MT are compatible with the perceptual biases observed in humans. Our study provides a comprehensive account of the computations of motion and depth by a moving observer, along with psychophysical evidence from humans and neural correlates in the macaque area MT. These findings extend our understanding of the interactions between motion and depth perception in more naturalistic viewing geometries.

2-146. A causal test of the interactions between brain areas

Sam Snyder[1,2] SES253@PITT.EDU
Emily Oby[1] EMO22@PITT.EDU
Matthew A. Smith[3] MATTSMITH@CMU.EDU
Steven Chase[3] SCHASE@ANDREW.CMU.EDU
Byron Yu[3] BYRONYU@CMU.EDU
Aaron Batista[1] AARON.BATISTA@PITT.EDU

[1]University of Pittsburgh
[2]Center for Neuroscience
[3]Carnegie Mellon University

How is flexible behavior supported by the nervous system? Modular structure in the brain suggests an organization whereby individual regions are specialized for specific processes, and behavior is implemented through selective brain-wide interactions. In support of this view, previous work has shown how limited is the ability of a single brain area to rapidly reconfigure and generate novel neural patterns on short timescales (Oby et al., 2019). Here, we sought to quantify whether interactions between brain areas possess flexibility beyond that seen in the interactions within single brain areas. We hypothesize that interactions between brain regions are more flexible than interactions within single regions. To probe the limits of flexibility in the interactions between populations of neurons, we developed a novel brain-computer interface (BCI) paradigm using populations of neurons in the dorsal premotor cortex (PMd) and primary motor cortex (M1). In order to characterize functional interactions between separate populations, we used canonical correlation analysis (CCA) to identify the most strongly correlated dimensions of neural activity between the two populations. Then, we used our BCI paradigm to present a novel decoder that directly challenged the animal to break the naturally-occurring correlation, and thus administer a causal test of the flexibility of the interactions between neural populations. Here, to break a correlation means to express a distribution of neural activity patterns that decorrelate across the two populations. To test interactions within single areas, two populations were made by randomly splitting the neurons from that area into two groups and then applying an identical procedure to challenge correlations. We found that correlations between two brain regions were easier to break than correlations identified within either single brain area. These results provide causal evidence that the correlations between brain regions are flexible. Thus, these interactions could be a mechanism to support flexibility in behavior.

2-147. Structural localization is embedded in the spike trains of neurons

Gemechu Bekele Tolossa[1,2] G.TOLOSSA@WUSTL.EDU
Aidan Schneider[3,4] AIDANMSCHNEIDER@GMAIL.COM
Keith Hengen[3] KHENGEN@WUSTL.EDU

[1]Washington University in St Louis
[2]Neuroscience
[3]Washington University, St. Louis
[4]Biology

Neurons within a brain region have long been observed to exhibit correlated dynamics while neurons from distinct regions receive diverse inputs and display different connectivity patterns. As a result, the electrophysiological features of neurons may vary across functional brain regions, potentially possessing information about their spatial location. However, the extent to which anatomical location is reliably embedded in spike timing, i.e., the neural code, remains uncertain. Further, the degree to which such signatures in neural activity emerge at the ensemble level or are manifest within the spike timing of individual neurons is unknown. In this study, we investigated the differences in spike trains among neurons across various functional brain regions using machine learning. Using the visual coding neuropixels data from the Allen Brain Observatory, we extracted spike train features and trained machine learning models to classify neuronal units into their functional brain regions, subregions, and layers. Our findings demonstrate that discernible patterns exist within the spike time series generated by individual neurons, which support their localization at arbitrary levels of structural hierarchy such as functional visual cortex subareas, hippocampal subregions, thalamic nuclei, and cortical layers. Crucially, trained classifiers were robustly capable of identifying single neuron anatomical location when tested against neurons from withheld animals. This suggests the presence of universal spatial fingerprints within neuronal spike trains. Beyond demonstrating latent information streams in neural spiking, these findings have practical implications for neuron-level brain-region localization directly from electrophysiological recording.

2-148. Latent representation learning for extracellular waveform clustering

Xiang Wang[1] XIANG.WANG@YALE.EDU
Mitchell Morton[1,2] MITCHELL.MORTON@YALE.EDU
Sachira Denagamage[1,2] SACHIRA.DENAGAMAGE@YALE.EDU
Nyomi Hudson[1] NYOMI.HUDSON@YALE.EDU
Anirvan Nandy[1] ANIRVAN.NANDY@YALE.EDU
Monika Jadi[1] MONIKA.JADI@YALE.EDU

[1]Yale University
[2]Neuroscience

Cortical Neurons are classified into a diversity of cell types that exert distinct functions in cortical circuits. Most in vivo recordings identify putative neuron types by applying unsupervised clustering algorithms on extracellular action potential waveforms. However, previous methods either require specification of waveform features or suffer from the curse of dimensionality. Furthermore, the lack of a principled way of selecting the optimal number of clusters hampers the comparisons among different clustering efforts. To address the problem of dimensionality, we developed a new clustering pipeline (latentMAP) that deploys a latent representation learning and a network clustering algorithm to identify putative cell types from extracellular waveforms. Additionally, we propose a systematic way of choosing the optimal cluster number by using a soft-voting strategy with multiple quality metrics. Compared with previous methods, our pipeline reveals a more diverse collections of waveform types that are robust over a range of hyperparameter values. We applied our clustering pipeline on multi-area recordings from both monkeys and mice and compared them with excitatory and inhibitory tagging obtained from either cross-correlograms (CCG) or optogenetics. We found consistent distribution patterns of clusters across the two functional cell-types in both species. Our study provides a novel framework that can scale up to electrophysiological features beyond those describing waveform shape, and can discover a large number of stable cell clusters in high dimensional extracellular data.

2-149. Higher D1R density on dlPFC PV cells increases distractibility in marmoset versus macaque

Tsvetoslav Ivanov[1] PREMURAX.SCIENCE@GMAIL.COM
Mary Kate Joyce[2] MARYKATE.JOYCE@YALE.EDU
Fenna Krienen[3] FK5009@PRINCETON.EDU
Jude Mitchell[4] JMITCHELL@BCS.ROCHESTER.EDU
Jay Ma[2] J.MA@YALE.EDU
Wataru Inoue[5] WINOUE@UWO.CA
Anirvan Nandy[2] ANIRVAN.NANDY@YALE.EDU
Dibyadeep Datta[2] DIBYADEEP.DATTA@YALE.EDU
Alvaro Duque[2] ALVARO.DUQUE@YALE.EDU
Jon Arellano[2] JON.ARELLANO@YALE.EDU
Rahul Gupta[1] XV20319@BRISTOL.AC.UK
Guillermo Gonzalez-Burgos[6] GBURGOS@PITT.EDU
David Lewis[6] LEWISDA@UPMC.EDU
Nenad Sestan[2] NENAD.SESTAN@YALE.EDU
Steven McCarroll[7] MCCARROLL@GENETICS.MED.HARVARD.EDU
Julio Martinez-Trujillo[8] JULIO.MARTINEZ@ROBARTS.CA
Sean Froudist-Walsh[1,9] SEAN.FROUDIST-WALSH@BRISTOL.AC.UK
Amy Arnsten[2] AMY.ARNSTEN@YALE.EDU

[1]University of Bristol
[2]Yale University
[3]Princeton University
[4]University of Rochester
[5]Western University
[6]University of Pittsburgh
[7]Harvard University
[8]University of Western Ontario
[9]Bristol Computational Neuroscience Unit

The New World monkeys, marmosets, are emerging as a promising non-human primate model of cognition due to their expanded dorsolateral prefrontal cortex (dlPFC) and, like rodents, amenability to gene editing paradigms (Oikonomidis et al., 2017). However, marmosets are more distractible and harder to train than macaques (Prins et al., 2017), impeding their experimental use. What causes such distractibility is unclear. Given the critical role of dopamine and inhibition in distractor resistance (Williams & Goldman-Rakic, 1995), we assessed distractibility in a passive visual fixation task and performed two independent assays of dopamine 1 receptor (D1R) expression in inhibitory (I) parvalbumin (PV) cells in the dlPFC of marmosets and macaques. We then designed a spiking network model of visuospatial working memory to link microcircuit anatomy and behaviour (Compte et al., 2000). The model featured two functional I cell types – INEAR and IOPP – with near- and opposite-feature selectivity (Kuan et al., 2022). We probed the biological plausibility of cell-type-specific D1R expression in three networks where excitatory (E) cells and, respectively, (1) INEAR, (2) IOPP, or (3) both expressed D1Rs. Our behavioural results confirmed higher distractibility in marmosets versus macaques, while the molecular analyses revealed higher expression of the D1R-coding genes and proteins in the marmoset dlPFC PV cells. The modelling demonstrated that the network where D1Rs exclusively modulate E and INEAR cells robustly reproduces the classic inverted-U relationship between D1R stimulation and persistent activity. By integrating our molecular findings into this network, we discovered that higher D1R density on marmoset dlPFC PV cells contributes to greater distractibility by reducing the D1R stimulation range that supports distractor resistance. This provides a plausible mechanistic explanation for the greater distractibility in marmosets. Our cross-species, interdisciplinary study provides insights into species differences in fundamental cognitive abilities and how dopaminergic modulation of inhibitory cells affects working memory.

2-150. Switching motor cortical dynamical rules during dexterous movements

Ahmet Arac[1,2] AARAC@MEDNET.UCLA.EDU
Sanjay Shukla[1] SANJAY.SHUKLA398@GMAIL.COM
Erica Nagase[1] ENAGASE@UCLA.EDU
Alan Yao[1] ALANYAO314@GMAIL.COM
Kate Santoso[1] KSANTOSO@STUDENTS.LLU.EDU
Emily Stenzler[1] EJSTENZLER@GMAIL.COM
David Lipkin[1] DAVIDLIPKIN@G.UCLA.EDU
Angela Kan[1] ANGELAKAN2002@GMAIL.COM

[1]UCLA
[2]Neurology

The complex behaviors of animals, such as reach-and-grasp movements, arise from the coordinated activity of multiple brain regions. To understand the neural principles underlying this behavior, we recorded spiking activity in 10 brain regions while the mice performed a skilled reach-and-grasp task. We found kinematically distinct behavioral phases of this behavior. Our neural data analyses with dynamical system modeling demonstrated that each of these behavioral phases were governed by different discrete neural dynamics with regional differences. Data-driven, multi-regional recurrent neural network (RNN) modeling combined with in silico electrophysiological perturbation experiments, and supported by in vivo optogenetic perturbation experiments, showed that the motor cortex neural activity is shaped by different cortical and subcortical brain regions during different behavioral phases. Overall, we show that different movement patterns underlie the reach-and-grasp behavior, and the dynamical rules of the motor cortex switch and are influenced by different brain regions during different phases of the skilled motor task.

2-151. Stochastic gene expression drives correlated synaptic noise causing representational drift

Oleg Senkevich[1] O.SENKEVICH@ULSTER.AC.UK
Cian O'Donnell[1,2] C.ODONNELL2@ULSTER.AC.UK

[1]University of Ulster
[2]Intelligent Systems Research Centre

Individual synapses in the brain fluctuate in size on time scales of hours-days even in the absence of electrical activity. These fluctuations are likely caused by local translation bursts since all mRNAs are stochastically delivered from the soma along highly elongated and branched axon and dendrites. Although a number of works studied the network-level implications of uncorrelated synaptic noise, the particular noise definitions in such works lack a solid foundation. Here, we attempt to close this gap and obtain the exact statistics of activity-independent synaptic fluctuations under the premise of the standard model of gene expression on tree-like neuronal morphologies. This is achieved using the method of generating functions, and the resulting statistics depend not only on the interplay of the temporal scales of gene-mRNA-protein dynamics but also on the spatial properties of the cells. One of the key findings from the model is that, due to resource sharing, the protein noise in nearby synapses is highly correlated presumably implying high correlations in the corresponding synaptic weights. We tested the implications of such correlations in a simple two-layer neural circuit model with Hebbian learning and found that the correlated synaptic weight fluctuations tend to increase the rate of representational drift compared to the independent noise with the same variance. Furthermore, apart from spatial correlations, our model predicts high correlations in time on the typical for representational drift experiments time scales of weeks. Such temporal correlations may, in principle, lead to misinterpretation of experimental data making an ultimately undirected diffusion of representations appear on short time scales as directional drift. Overall, this work provides an efficient tool for finding realistic synaptic noise statistics and demonstrates their importance for systems neuroscience.

2-152. Sensory experience aligns feedforward-recurrent networks to drive reliable cortical representations

Augusto Lempel[1,2] AUGUSTO.LEMPEL@MPFI.ORG
Sigrid Tragenap[3] TRAEGENAP@FIAS.UNI-FRANKFURT.DE
Clara Tepohl[1] CLARA.TEPOHL@MPFI.ORG
Matthias Kaschube[3] KASCHUBE@FIAS.UNI-FRANKFURT.DE
David Fitzpatrick[1] DAVID.FITZPATRICK@MPFI.ORG

[1]Max Planck Florida Institute for Neuroscience
[2]Fitzpatrick Lab
[3]Frankfurt Institute for Advanced Studies

Endogenous developmental mechanisms provide an initial cortical network organization that evolves with experience into a reliable stimulus representation critical for behavior. To identify circuit-level changes underlying this process, we investigate the development of layer 2/3 (L2/3) of the ferret primary visual cortex (V1), where an endogenous modular structure of highly variable responses is transformed by experience into a reliable representation of stimulus orientation1. We use calcium imaging, electrophysiology, and computational modeling to resolve the functional state of feedforward and recurrent networks and analyze how their interaction affects representation reliability, i.e. stimulus discriminability across trials, before and after sensory experience. Our results reveal a network-alignment process underlying the emergence of reliable stimulus representations. At eye opening, modular responses to oriented stimuli in L2/3 are highly correlated with single-unit activity in L2/3 but poorly correlated with single-unit activity in layer 4 (L4), the main provider of feedforward inputs to L2/3. While both L2/3 and L4 responses are highly variable, stimulus discriminability is higher in L2/3 than L4 at eye opening. Computational modeling shows that this visually naive functional organization can be explained as resulting from a well-organized modular recurrent network in L2/3 that is driven by misaligned, poorly selective feedforward inputs from L4. After experience, experimental data shows increased L4 to L2/3 correlation and improved stimulus discriminability across layers. Computational modeling shows that an alignment of feedforward inputs with amplifying recurrent subnetworks in L2/3 can contribute to stimulus discriminability and predicts improved temporal stability of orientation preference following experience. In-vivo whole-cell experiments confirm that orientation preference stability improves in L2/3 after experience, providing strong evidence of increased alignment. In conclusion, our experimental and theoretical work indicate that sensory experience aligns feedforward inputs and recurrent networks to facilitate reliable stimulus representations.

2-153. Optimal flexible inference for behavior without generative world models

Francesco Trapani[1,2] FRANCESCO.TRAPANI@RESEARCH.FCHAMPALIMAUD.ORG
Carlos Stein[1] CARLOS.STEIN@RESEARCH.FCHAMPALIMAUD.ORG
Daniel McNamee[3] DANIEL.MCNAMEE@RESEARCH.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Natural Intelligence Lab
[3]Champalimaud Research

All descriptive and quantitative formalisms regarding brain function posit that the brain encodes an internal generative model of the world to some degree of approximation. That is, the brain receives sensory information about the current state of the world, uses an internal representation of the world's structure to predict future states e.g. via simulation, and leverages this information to flexibly induce rational behavior. This perception-model-action cycle is predicated on the axiomatic assumption that generative world models are a necessary intermediate step for inducing flexible behavior via inference and planning. We present a theoretic analysis that deconstructs this assumption and shows that (1) internal world models are not necessary for planning behavior, and (2) counterintuitively, internal world models are irrational for planning behavior due to their fundamentally inefficient use of the most important resource of all for biological agents, namely time. We describe the implications of this in particular suggesting that the brain does not encode internal world models and hypothesize an alternative relational formalism for internally reflecting dynamic structure in the world which is time-optimal for inducing flexible behavior. For initial support of this hypothesis, we compare predictions of our model to neural phenomena in particular rerouting hippocampal replay.

2-154. Dynamics of decision variable in prefrontal cortex predict the impact of prior during perceptual decision-making

Julie Charlton[1,2] JULIECHARLTON94@GMAIL.COM
Thomas Langlois[3] THOMAS.LANGLOIS@AUSTIN.UTEXAS.EDU
Robbe Goris[3] ROBBE.GORIS@UTEXAS.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute
[3]University of Texas at Austin

Perceptual decisions are informed by the present sensory input and by expectations, or "priors", that reflect previously experienced statistical regularities in the environment. The neural mechanisms that integrate sensory signals with priors are unknown. To shed light on this matter, we developed an analysis that relates the dynamics of population activity in the prefrontal cortex to perceptual decisions on a trial-by-trial basis. Multi-electrode arrays were used to record neural population activity in the prearcuate gyrus of two macaque monkeys as they performed a perceptual orientation discrimination task. The animals judged the orientation of drifting grating stimuli under two contexts, each associated with a different distribution of stimulus orientation. The choices of both monkeys were biased in a context-specific manner, reflecting the impact of their priors on their perceptual interpretations. Importantly, our task design decoupled the monkeys' choice from the saccade they used to report their choice. This allowed us to study the component of choice formation that was uncontaminated by motor planning, or a prior over saccade direction. To study the neural correlate of this behavior, we decoded a time-varying decision variable (DV) from jointly recorded neural responses. We found that the initial value of the DV on average differed across the task contexts but did not predict choice outcome. In contrast, the dynamic range on average did not differ across the two task contexts, yet it accounted for systematic differences in the choice behavior. On trials in which the DV exhibited a dynamic range smaller than the median, the animals tended to be more biased by the context-specific expectation. These findings reveal how prefrontal circuits integrate prior stimulus expectations and incoming sensory signals at the behaviorally relevant timescale of the single trial and suggest a role for the dynamics of the DV in choice formation.

2-155. Hyperpolarization-activated currents drive neuronal activation sequences in sleep

Dhruv Mehrotra[1,2] DHRUV.MEHROTRA@MAIL.MCGILL.CA
Daniel Levenstein[1] DANIEL.LEVENSTEIN@MILA.QUEBEC
Adrian Duszkiewicz[3] ADRIAN.DUSZKIEWICZ@ED.AC.UK
Sofia Skromne Carrasco[1] SOFIA.SKROMNECARRASCO@MAIL.MCGILL.CA
Sam Booker[4] SBOOKER@EXSEED.ED.AC.UK
Angelika Kwiatkowska[4] ANGELAKW68@GMAIL.COM
Adrien Peyrache[1] ADRIEN.PEYRACHE@MCGILL.CA

[1]McGill University
[2]Integrated Program in Neuroscience
[3]University of Stirling
[4]University of Edinburgh

Cortical activity is characterized by state-specific dynamics arising from the interplay between connectivity, cellular diversity, and intrinsic properties. Specifically, during sleep, computations must be carried out by internally organized activity patterns that constrain and coordinate neuronal spiking across thalamocortical networks. During non-Rapid Eye Movement (NREM) sleep, cortical population activity alternates between periods of neuronal firing ("UP" states) and neuronal silence ("DOWN" states). Neuronal activity patterns at DOWN-to-UP (DU) transitions have functional relevance beyond sleep: they are related to sensory coding during wakefulness and support homeostatic processes and memory consolidation. Despite their functional importance, the factors that organize these spiking patterns remain unknown, but mechanisms that rely on network connectivity or intrinsic excitability have been proposed.

To elucidate the origin of sequential activity at UP onset, we recorded populations of neurons in the head-direction cortex (HDC, i.e., post-subiculum), where the behavioral correlates of most neurons are well established. While UP-DOWN (UD) transitions were synchronous along the dorsoventral (DV) axis, we observed sequential activation of neurons at UP onset. The decoded virtual HD during UP onset rapidly converged to a stable direction before full reinstatement of the population activity.

To understand the mechanisms underlying the sequential activation, we modelled UP/DOWN alternations using an Adapting Wilson-Cowan model, and found that unlike gradients in local connectivity, excitability/input, and adaptive current, a gradient in rectifying current (Ih) was able to uniquely reproduce the experimental observations.

Subsequent ex vivo intracellular recordings confirmed the predicted DV gradient of Ih in HDC. In addition, the model was also able to predict a yet-unobserved relationship between UP onset and post-DOWN rebound activity which was able to explain UP onset timings across various cortical regions.

Our findings suggest that the sequential activation at UP onset potentially activates the downstream entorhinal-hippocampal formation by querying the network with specific directional information.

2-156. Adaptive recurrent visual inference with learnt top-down attention

Eivinas Butkus[1,2] EB3407@COLUMBIA.EDU
Nikolaus Kriegeskorte[1] NK2765@COLUMBIA.EDU

[1]Columbia University
[2]Psychology

Feedforward convolutional neural networks can account for the neural and behavioral data associated with the initial feedforward sweep through the primate visual hierarchy. However, such networks lack recurrent connectivity and mechanisms that could explain top-down attentional effects. Here we present a convolutional neural network model of human object detection that performs visual inference using adaptive, recurrent computation with learnt top-down attention. The task for both humans and models was to determine the class and location of a digit (the "what" and the "where"). Top-down attention in our model can modulate the gain of feature maps ("what gain"), locations ("where gain"), or both. We found that recurrent models with top-down attention outperform the feedforward and recurrent models without attention at both "what" and "where" judgments. Recurrent models with attention also qualitatively capture the effect of presentation duration on human performance. This work is a step towards understanding the adaptive dynamics of human visual inference.

2-157. Distinguishing modular and distributed computations using population-level neural variability

Francesco Massari[1,2] FMASSARI@STUDENT.ETHZ.CH
Aniruddh R. Galgali[3] A.GALGALI@UCL.AC.UK
Diogo Peixoto[4] DIOGO.RP5@GMAIL.COM
William T. Newsome[4] BNEWSOME@STANFORD.EDU
Maneesh Sahani[3] MANEESH@GATSBY.UCL.AC.UK
Valerio Mante[5] VALERIO@INI.UZH.CH

[1]ETH Zurich
[2]Institute of Neuroinformatics
[3]University College London
[4]Stanford University
[5]Institute of Neuroinformatics, UZH and ETH Zurich

Complex behaviors are thought to emerge from the coordinated activity across multiple brain areas. During perceptual decision-making, for instance, similar decision- and movement-related signals are typically represented concurrently across many areas. It remains mostly unknown whether such signals arise in individual brain areas and are then broadcast across cortex (modular computation) or whether they emerge through the interactions between multiple areas (distributed computation). Here, we distinguish between these scenarios by analyzing the dynamics of population-level, single trial residuals from Utah arrays in pre-motor (PMd), primary motor, (M1) and prefrontal cortex (PFC) of macaque monkeys engaged in a random-dots motion discrimination task. Average neural trajectories show similar choice-predictive activity and movement-related transients in all areas. Residual dynamics, however, reveal differences in the underlying computations. During decision-formation, residual dynamics support evidence integration that is perfect in PMd/M1, but leaky in PFC. During movement, residuals are transiently amplified in PMd/M1, but quickly decaying in PFC, suggesting that movement transients may be locally generated in PMd/M1, but inherited from upstream areas in PFC. In simultaneously recorded PMd and M1 responses, we show that residual dynamics can disentangle local, within-area contributions to the dynamics from contributions reflecting across-area interactions. These analyses suggest that dynamics during decision-formation result from a distributed computation involving recurrent interactions between PMd and M1, whereas movement dynamics emerge locally in M1 and are then broadcast to PMd. Overall, we find that the dynamics of residuals distinguish between modular and distributed multi-area computations and can thus reveal differences in the local computations that are not otherwise apparent from the average dynamics.

2-158. Unsupervised discovery of nonlinear and interpretable communication submanifolds

Sai Koukuntla[1,2] SAI.KOUKUNT@GMAIL.COM
Timothy Harris[1] THARR116@JHU.EDU
Adam Charles[1,3] ADAMSHCH@GMAIL.COM
Carlos Brody[4,5] BRODY@PRINCETON.EDU

[1]Johns Hopkins University
[2]Biomedical Engineering
[3]Department of Biomedical Engineering
[4]Princeton University
[5]Princeton Neuroscience Institute

A fundamental goal of systems neuroscience is to understand interactions between neural populations across the many regions involved in any given behavior. Previous work has proposed modeling inter-brain region interactions through low-dimensional linear communication subspaces. Vari- ability along the linear communication subspace encodes information shared between the brain regions while variability along the nullspace encodes private information. However, as many neural manifolds (e.g., place cell codes) can be highly nonlinear, shared representations between these nonlinear neural manifolds will in general also be nonlinear. Here, we introduce communication submanifolds, a nonlinear generalization of communication subspaces. We develop an unsupervised neural network-based method to obtain an interpretable, generalizable low-dimensional embedding of the communication submanifold between two brain regions. Our approach consists of two stages. For the first stage, we develop Subman- ifold Partitioning via Least-variance Informed Channel Estimation (SPLICE), a novel neural network- based approach that disentangles shared information from private information, and thus finds nonlinear communication submanifolds. For the second stage, we build in interpretability with an autoencoder variant of Isomap that learns a low-dimensional embedding of the communication submanifold while preserving its intrinsic geometry. Using simulated responses of lateral geniculate nucleus (LGN) and V1 populations to visual stimuli, we show that our method 1) successfully disentangles the contributions of shared and unshared latent variables and 2) finds a low-dimensional embedding of the communication submanifold whose coordinates naturally correspond to stimulus parameters.

2-159. Biologically plausible neural decoder ensembles are robust to overfitting and noise

Benjamin Ruben[1,2] BENRUBEN@G.HARVARD.EDU
Cengiz Pehlevan[1] CPEHLEVAN@SEAS.HARVARD.EDU

[1]Harvard University
[2]Biophysics PhD Program

A population of neurons that subserve a common function often encodes information in collective patterns of activity. Many works interpret neural population codes in terms of the information available to a single linear decoder with access to the full code. However, biological neural circuits are sparsely connected and noisy, with broad heterogeneity in the degrees of synaptic connectivity. Further, the brain's internal representations are interpreted collectively by populations of neurons comprising downstream brain regions. Motivated by these observations, we study the effect of subsampling, heterogeneity, and noise on the accuracy of biologically plausible decoder ensembles trained to predict the angle of a visual stimulus from recordings of mouse V1. First, we demonstrate that single-neuron decoders with a critical synaptic connectivity may over-fit to their training set. Then, comparing decoder ensembles with homogeneous and heterogeneous degrees of synaptic connectivity, we find a clear benefit to heterogeneity as an implicit regularization against overfitting. Finally, we show that under noisy conditions, ensembles of sparsely connected decoders outperform a single fully-connected decoder. We also demonstrate these phenomena in an analytical theory of ensembled ridge regression. Together, our results suggest that the sparse, heterogeneous connectivity of biological neural networks allows for robust computations by preventing overfitting and reducing the effect of noise inherent to biological neural networks.

2-160. Neuronal avalanches support cognitive processes during speech and music listening

Matteo Neri[1] MATTEBLACKS98@GMAIL.COM
Claudio Runfola[2] CLAUDIO.RUNFOLA@UNIV-AMU.FR
Noemie te Rietmolen[3] NOEMIETER@GMAIL.COM
Pierpaolo Sorrentino[2] PIERPAOLO.SORRENTINO@UNIV-AMU.FR
Daniele Schon[2] DANIELE.SCHON@UNIV-AMU.FR
Benjamin Morillon[2] BENJAMIN.MORILLON@UNIV-AMU.FR
Giovanni Rabuffo[2] GIOVANNI.RABUFFO@UNIV-AMU.FR

[1]Institut des neurosciences de la Timone
[2]Aix Marseille Universite, INSERM, INS, Institut de Neurosciences des Systemes, Marseille, France
[3]Language and Computation in Neural Systems Group, Max Planck Institute for Psycholinguistics, Nijme

Neuronal avalanches consist of collective network events propagating across the brain in short-lived and aperiodic instances. These salient events have garnered a great interest in the study of cortical dynamics. They have been observed across different imaging modalities and scales, and have been used to successfully distinguish between healthy and pathological conditions or between resting wakefulness and sleep states. While a growing body of literature investigated neuronal avalanches in task-free conditions, whether they index cognitive functions or purely reflect physiological states remains an open question. In this work we investigated neuronal avalanches to index cognition, analyzing an intracranial stereo electroencephalography (sEEG) dataset collected during speech, music listening (naturalistic stimuli of about 10 minutes) and resting state in 19 epileptic patients. Firstly, we observed that avalanches relate to cognitive processes insofar as they are similarly distributed in time across patients while listening to speech and music, but not during rest. Furthermore, this result is not only driven by avalanches in auditory regions, but extends to more integrative networks. Secondly, we found that there are time windows in which avalanches are particularly coordinated across participants, and this is the case for both music and speech in partially overlapping distributed networks involving auditory and non-auditory regions. Finally, we analyzed avalanche propagation patterns in the brain and we found that the directed functional connections that differ the most across conditions (speech, music and resting state) are the ones between auditory and non auditory regions. This underlines the importance of feedforward and feedback mechanisms in shaping brain dynamics during speech and music perception. Our work supports neuronal avalanches as a valuable and computationally advantageous framework for the study of cognition in humans, that can be adopted using different imaging modalities and at diverse spatio-temporal scales.

2-161. Deep learning-driven characterization of single cell tuning in primate visual area V4 unveils topological organization

Konstantin Willeke[1] KONSTANTIN.WILLEKE@GMAIL.COM
Kelli Restivo[2] KELLIRESTIVO@GMAIL.COM
Katrin Franke[2] KATRIN.FRANKE@BCM.EDU
Arne Nix[3] ARNE-FABIAN.NIX@UNI-TUEBINGEN.DE
Santiago Cadena[1] SA.CADENA721@GMAIL.COM
Tori Shinn[2] TORIW@BCM.EDU
Cate Nealley[2] CATE.NEALLEY@BCM.EDU
Gabrielle Rodriguez[2] GABRIELLE.RODRIGUEZ@BCM.EDU
Saumil Patel[2] SPATEL@BCM.EDU
Alexander Ecker[4] ECKER@CS.UNI-GOETTINGEN.DE
Fabian Sinz[3] SINZ@UNI-GOETTINGEN.DE
Andreas S. Tolias[2] ASTOLIAS@BCM.EDU

[1]University of Tubingen
[2]Baylor College of Medicine
[3]University of Gottingen
[4]University of Goettingen

Deciphering the brain's structure-function relationship is key to understanding the neuronal mechanisms underlying perception and cognition. The cortical column, a vertical organization of neurons with similar functions, is a classic example of primate neocortex structure-function organization. While columns have been identified in primary sensory areas using parametric stimuli, their prevalence across higher-level cortex is debated. A key hurdle in identifying columns is the difficulty of characterizing complex nonlinear neuronal tuning, especially with high-dimensional sensory inputs. Here, we asked whether area V4, a mid-level area of the macaque visual system, is organized into columns. We combined large-scale linear probe recordings with deep learning methods to systematically characterize the tuning of >1,200 V4 neurons using in silico synthesis of most exciting images

(MEIs), followed by in vivo verification. We found that the MEIs of single V4 neurons exhibited complex features like textures, shapes, or even high-level attributes such as eye-like structures. Neurons recorded on the same silicon probe, inserted orthogonal to the cortical surface, were selective to similar spatial features, as expected from a columnar organization. We quantified this finding using human psychophysics and by measuring MEI similarity in a non-linear embedding space, learned with a contrastive loss. Moreover, the selectivity of the neuronal population was clustered, suggesting that V4 neurons form distinct functional groups of shared feature selectivity, reminiscent of cell types. These functional groups closely mirrored the feature maps of units in artificial vision systems, hinting at shared encoding principles between biological and artificial vision. Our findings provide evidence that columns and functional cell types may constitute universal organizing principles of the primate neocortex, simplifying the cortex's complexity into simpler circuit motifs which perform canonical computations.

2-162. Cue-specific neuronal ensembles span intermittent rate coding of working memory

Matt Panichello MATT.PANICHELLO@GMAIL.COM
Donatas Jonikaitis DJONIKAITIS@GMAIL.COM
Yu Jin Oh JIN.OH.OFFI@GMAIL.COM
Ethan Trepka TREPKA@STANFORD.EDU
Tirin Moore TIRIN@STANFORD.EDU

Stanford University

Persistent, memorandum-specific neuronal spiking activity has long been hypothesized to underlie working memory (Fuster & Alexander, 1971; Funahashi et al., 1989; Wang, 2021). However, recent theoretical and experimental work suggests a possible role for 'activity-silent' synaptic mechanisms (e.g., Mongillo et al., 2008; Stokes, 2015; Lundqvist et al., 2018; Barbosa et al., 2020). This issue remains controversial because evidence for either view has largely depended on datasets that fail to capture single-trial population dynamics or on indirect measures of neuronal spiking. We addressed this by examining the dynamics of mnemonic information on single trials obtained from large, local populations of prefrontal neurons recorded simultaneously in monkeys performing a working memory task. We show that mnemonic information does not persist in the spiking activity of neuronal populations, but instead alternates between coordinated 'On' and 'Off' states during memory delays. At the level of single neurons, Off states are driven by a coordinated loss of selectivity for memoranda and a return of firing rates to baseline levels. Further exploiting the large-scale recordings, we show that mnemonic information is available in the pattern of functional connections among neurons throughout the memory delay, even during Off states. Therefore, intermittent periods of memoranda-specific spiking coexist with activity-silent mechanisms to span memory delays.

2-163. Some and Done? Temporally extended decisions with very few rollouts

Sixing Chen[1,2] SIXING.CHEN@NYU.EDU
Kristopher Jensen[3] KTJ21@CAM.AC.UK
Marcelo Mattar[1,4] MARCELO.MATTAR@NYU.EDU

[1]New York University
[2]Department of Psychology
[3]University College London
[4]Psychology

It has been suggested that humans mentally simulate the outcomes of their actions when making decisions. However, this process becomes challenging in the context of human decision-making, which typically involves navigating through decision trees that extend over time and encompass numerous potential outcomes. This complexity leads to the assumption that effective forward simulation would demand an impractical level of resources. Given the constraints of time and computational power in human cognition, such extensive forward simulation seems at odds with our ability to make rapid and accurate decisions in these extended scenarios. To address this apparent discrepancy, we propose a model where optimal planning can be achieved with very few forward simulations, or rollouts. Our approach expands upon an influential theory where, in bandit settings, humans make optimal decisions based on very few samples. We demonstrate that under tight computational constraints, conducting numerous partial (shallow) rollouts can yield more favorable outcomes than a smaller number of comprehensive (deep) rollouts. Additionally, our model aligns with the human tendency to prefer paths with low variance in rewards, offering an alternative explanation for behaviors traditionally attributed to reflexive pruning strategies. In conclusion, our research offers a plausible mechanism for efficient human action planning and provides insights into the neural underpinnings of this process.

2-164. Large Scale Study of Human Memory for Narratives using Large Language Models

Antonios Georgiou[1] ANTOINE.GEORGIOU@GMAIL.COM
Tankut Can[2,3] TANKUT.CAN@GMAIL.COM
Mikhail Katkov[4] MIKHAIL.KATKOV@GMAIL.COM
Misha Tsodyks[1] MISHA@WEIZMANN.AC.IL

[1]Weizmann Institute of Science
[2]Institute for Advanced Study
[3]School of Natural Sciences
[4]Institute for Advanced Study and Weizmann Institute of Science

Large-scale experiments in human memory have traditionally relied on stimuli which are amenable to automated analysis, such as random lists of words. In contrast, the study of memory for meaningful material, such as narratives, has resisted automation, consequently limiting the scope of such investigations. In this work, we develop a pipeline for designing large scale memory experiments and analyzing the obtained results using large language models (LLMs) as a scientific tool. We performed online memory experiments with a large number of participants and collected recognition and recall data for narratives of different lengths. We found that both recall and recognition performance scale linearly with narrative length. Furthermore, in order to investigate the role of narrative comprehension in memory, we repeated these experiments using scrambled versions of the presented stories. We found that even though recall performance declined significantly, recognition remained largely unaffected. Interestingly, recalls in this condition seem to follow the original narrative order rather than the scrambled presentation, pointing to a contextual reconstruction of the story in memory.

2-165. Animal vocalizations can be discriminated because of their slow, predictable features

Ron DiTullio[1,2] RON.W.DITULLIO@GMAIL.COM
Linran Wei[1,3] LINRAN@SAS.UPENN.EDU
Vijay Balasubramanian[1] VIJAY@PHYSICS.UPENN.EDU

[1]University of Pennsylvania
[2]Computational Neuroscience Initiative
[3]Physics

Many animals use vocalizations to communicate. While the form of vocalizations varies significantly across species, animals learn to successfully discriminate vocalizations within and between species. We propose that this universal learnability is possible because vocalizations share an underlying structure – temporal regularity, or the tendency of stimulus properties to change smoothly over time. We tested this proposal by using Slow Feature Analysis (SFA) to find the most temporally regular components of vocalizations from multiple species including birds (blue jay, house finch, American yellow warbler, and great blue heron), macaques, and humans (English speakers). To assess learnability, we projected vocalizations into the learned SFA space, and tested both intra-class (two different vocalizations from the same speaker or species for birds) and inter-class (different speakers, or species for birds) auditory discrimination by a trained network classifier. We found that: 1) the classifier achieves high performance in vocalization discrimination tasks for all species 2) performance depends on the $\sim$10 slowest (most temporally regular) features, 3) a classifier trained on this subset of features has nearly the same performance as using all features, 4) Most vocalizations are comprised of a small number ($\sim$ 10) of features with high temporal regularity, as indicated by an inflection in the SFA eigenvalue distribution, and 5) These slowest SFA features are highly correlated with the most predictable components of the vocalizations. Taken together, our findings shows that vocalizations across diverse species contain a small number of slow, predictable components that are sufficient to support auditory discrimination. Our results suggest that auditory neuroscientists should revive research on circuits tuned to temporal regularities in natural sounds [2], which are not captured by the dominant paradigm of deriving Spike Triggered Receptive Fields to study neural activity as only a function of frequency content and delays.

2-166. Unveiling the circuitry mechanism of novelty coding in the mouse visual system

Renzimo Zhang[1,2] ZIMO004@OUTLOOK.COM
Ruilin Zhang[3] ZHANGRUILIN_JERRY@PKU.EDU.CN
Disheng Tang[1] DISHENGTANG3@GMAIL.COM
Xiaoxuan Jia[1] JXIAOXUAN@GMAIL.COM

[1]Tsinghua University
[2]School of life sciences
[3]Peking University

Novelty-induced changes in neural activity have been observed across various brain regions in many species, with predictive coding being the prevailing theory. However, debates persist on where and how the error signals are computed, which urges more experimental evidence and computational efforts. Here we provide a cell-type specific computational model for novelty coding mechanism by analyzing the electrophysiological dataset collected from the mouse visual cortex in response to changes of novel or familiar images. Using unsupervised clustering, we identified functional clusters which respond differently to novelty. Further decoding analysis on individual neurons revealed distinct joint-distributions of coding accuracy for novelty and image identity between novel and familiar sessions, suggesting the existence of novelty encoding neurons independent from the encoding of image identity. Anatomical distribution of these novelty coding neurons shows bias towards superficial layers and a significant higher proportion in area VISal. To evaluate the role of these neurons in a network, we compared the functional signal flow with frequency-based Granger causality before and after image change. Topological analysis of the signal flow during image change showed significantly more diverging motifs in the novel condition, with novelty encoding neurons as the main driving node. As the previous studies failed to completely explain the response changes induced by novelty in different cell types, we propose a new model with distinct cell types initialized with the known anatomical connectivity. Incorporating Hebbian-like learning rule, our model reproduces the dynamics of different clusters with image coding and novelty coding that matches the experimental observations in different cell types. Overall, our work identified novelty encoding neurons by evaluating the mixed selectivity in individual neurons across different dimensions, and proposed a possible circuitry mechanism that explains the novelty induced changes, thereby expanding our understanding of novelty coding in the brain.

2-167. Neural network decoding of concept recall from human intracranial recordings

John Sakon[1,2] JOHNSAKON@GMAIL.COM
Yuanyi Ding[3] SEMISWIET@G.UCLA.EDU
Yipeng Zhang[3] ZYP5511@G.UCLA.EDU
Soraya Dunn[4] SLDUNN@MEDNET.UCLA.EDU
Nathan Wei[4] NJWEI@G.UCLA.EDU
Inesh Chakrabarti[3] INESH33@G.UCLA.EDU
Anthony Rangel[4] AJRANGEL@MEDNET.UCLA.EDU
James Bruska[4] BRUSKATECH@PROTONMAIL.COM
Vwani Roychowdhury[3] VWANI@G.UCLA.EDU
Itzhak Fried[4] IFRIED@MEDNET.UCLA.EDU

[1]UCLA
[2]Neurosurgery
[3]UCLA Electrical and Computer Engineering
[4]UCLA Neurosurgery

Intracranial encephalogram (iEEG) recordings of single neurons in the human medial temporal lobe (MTL) have revealed concept cells: neurons that selectively increase their firing rates to semantic concepts (e.g. Rachel from Friends). These cells respond to their selected concept regardless of sensory type (e.g. image, audio, text) or whether the concept is externally or internally generated (e.g. cued vs. remembered). The presence of neurons with such invariant responses suggests neural signals can be used to decode whenever participants internally process a given concept. For any given concept, however, odds are low that a concept cell will be recorded by chance, limiting the practicality of using such signals to decode memory for new experiences. Here we show that using transformer neural network models trained on whole brain iEEG recorded as participants watch audiovisual episodes we can successfully decode when participants subsequently recall concepts prior to their vocalization. Successful decoding occurs seconds prior to vocalization and ceases at time of vocalization, suggesting our models reveal reinstatement of concepts from internally-generated memories. We achieve significant decoding of concepts whether models input spectral powers of local field potential (LFP) from macrowires or clusterless spikes in voltage from microwires. Notably, these two models, which use inputs from macro-scale LFP vs. micro-

scale neuron amplitudes, typically decode different concepts, suggesting distinct information exists at each neural scale. The ability to decode internal representations during recall evidences that our models can determine when people reinstate memories—providing key windows for interventional memory therapies including during sleep.

2-168. Projection-specific cortical processing of vocalizations.

Amy LeMessurier[1,2] AMY.LEMESSURIER@GMAIL.COM
Ayat Agha[1] AYAT.AGHA@NYULANGONE.ORG
Robert Froemke[1] ROBERT.FROEMKE@NYULANGONE.ORG

[1]NYU Grossman School of Medicine
[2]Neuroscience Institute

Perception of vocalizations is one of the most important functions of the auditory system in social animals. Responding to vocalizations requires segmentation and categorization of complex sounds, as well as communication between auditory and motor circuits during behavior. In mammals, the central auditory pathway consists of both feedforward circuits from the ears to higher cortical areas, and descending cortical projections that innervate much of the feed-forward pathway. In particular, corticofugal projections to the auditory midbrain are numerous, and key for computations that may support understanding speech, such as separating salient sounds from noise, and facilitating temporal integration in feed-forward neurons. These projections may be critical for behaviorally responding to vocalizations. Direct projections from auditory cortex to striatum, which are plastic with training on audio-motor tasks, may also link vocalizations with appropriate motor responses. We tested this in the context of an ethological mouse behavior that depends on left auditory cortex. Mothers find and retrieve isolated pups into the nest when pups emit ultrasonic vocalizations (USVs). Virgin females don't usually retrieve pups, but begin retrieving pups after several days of co-housing with a mom and litter. We chemogenetically silenced neurons projecting to inferior colliculus (corticocollicular) or striatum (corticostriatal) during pup retrieval in experts, and found that corticocollicular, but not corticostriatal neurons were key for successful retrieval. We compared encoding of USVs over days of co-housing using 2-photon Ca2+ imaging in awake virgins. Spontaneous activity increased dramatically in corticocollicular neurons during blocks of USV playback compared to pure tones, but not in corticostriatal neurons. We corroborated this with in vivo electrophysiology in optotagged projection neurons. In both groups, we observed delayed responses to USVs, which were larger in magnitude on expert days. These findings may reflect network-level plasticity upregulating activity in recurrent auditory circuits to support perception of USVs.

2-169. Uncovering dynamic internal states in mice learning a new decision-making task

Lenca Cuturela[1] CUTURELA@PRINCETON.EDU
The International Brain Lab The International Brain Lab[2] INFO@INTERNATIONALBRAINLAB.ORG
Jonathan Pillow[1] PILLOW@PRINCETON.EDU

[1]Princeton University
[2]International Brain Laboratory

Recent work has shown that animals performing a perceptual decision-making task alternate frequently between different internal states or strategies [1, 2]. However, it remains unknown how these states emerge over the course of learning. Does an animal alternate between multiple strategies from the very onset of training? Or do the states observed in fully-trained animals emerge only after extensive exposure to the task? Here we address this problem by introducing a novel dynamic latent state model, which we call "dynamic GLM-HMM". This model extends previous work that identified internal states using a Hidden Markov Model (HMM) with state-dependent Bernoulli generalized linear model (GLM) observations [1, 2]. However, the standard GLM-HMM has static parameters, making it inapplicable to a wide range of nonstationary phenomena. To overcome this limitation, we add a dynamic prior that allows both the HMM transition probabilities and GLM weights to evolve over time. This extension is critical for capturing nonstationary phenomena like learning, which is characterized as an increase in the stimulus weights over time [3]. After validating our approach on simulated data, we applied it to animal training data from the International Brain Lab (IBL) [4]. We found that animals switch between three states that can last for tens to hundreds of trials: an "engaged" state, in which task performance is high, and two "biased" states, in which performance is lower. Remarkably, we show that these states are present and identifiable even in the early training periods. During learning, animals improve their accuracy on the task through a combination of two changes: the stimulus weight in the engaged state grows larger, and the state transition probabilities evolve so that animals spend more time in the engaged state. Thus, we offer a novel method for uncovering changes in behavioral strategies during learning.

2-170. Bipartite invariance in mouse primary visual cortex

Zhiwei Ding[1] ZHIWEID@BCM.EDU
Santiago Cadena[2] SA.CADENA721@GMAIL.COM
Saumil Patel[1] SPATEL@BCM.EDU
Katrin Franke[1] KATRIN.FRANKE@BCM.EDU
Alexander Ecker[3] ECKER@CS.UNI-GOETTINGEN.DE
Dat Tran[1] DAT.TRAN@BCM.EDU
Kayla Ponder[1] KAYLA.JONESPONDER@BCM.EDU
Erick Cobos[1] EMCOBOST@GMAIL.COM
Paul Fahey[1] PAUL.FAHEY@BCM.EDU
Zhuokun Ding[1] ZHUOKUN.DING@BCM.EDU
Eric Wang[1] ERIC.WANG2@BCM.EDU
Taliah Muhammad[1] TALIAH.MUHAMMAD@GMAIL.COM
Jiakun Fu[1] JIAKUN.FU@BCM.EDU
Stelios Papadopoulos[1] STELIOS.PAPADOPOULOS@BCM.EDU
Fabio Anselmi[4] FABIO.ANSELMI@UNITS.IT
Edgar Walker[5] EYWALKER@BCM.EDU
Jacob Reimer[1] REIMER@BCM.EDU
Fabian Sinz[6] SINZ@CS.UNI-GOETTINGEN.DE
Xaq Pitkow[7] XAQ@CMU.EDU
Andreas Tolias[1] ATOLIAS@BCM.EDU

[1]Baylor College of Medicine
[2]University of Tubingen
[3]University of Goettingen
[4]University of Trieste
[5]University of Washington
[6]University of Gottingen
[7]Carnegie Mellon University

A key challenge sensory systems have to solve is to robustly extract specific features despite large variations in their natural visual input. To understand how brains achieve this generalization, it is crucial to identify features that neurons exhibit selectivity and invariance towards. However, the high-dimensional nature of ecological visual inputs makes it challenging to systematically characterize neuronal tuning. As a result, our knowledge of the invariances encoded by neurons is restricted to a handful of examples, such as phase invariance demonstrated by complex cells in the primary visual cortex when presented with grating stimuli. Here, we extended "inception loops" — a paradigm that iterates between large-scale recordings, neural predictive models, and in silico experiments followed by in vivo verification — to characterize neuronal invariances in the mouse primary visual cortex. Using a model trained to predict responses to arbitrary visual stimuli we synthesized Diverse Exciting Inputs (DEIs), a set of inputs that differ substantially from each other while each driving a target neuron strongly, and verified these DEIs' efficacy in vivo. We discovered a novel bipartite invariance: one portion of the receptive field encoded phase-invariant texture-like patterns, while the other portion encoded a fixed spatial pattern. Our analysis revealed that the division between the fixed and invariant subfields matched object boundaries defined by spatial frequency differences in highly activating natural image patches, suggesting that bipartite invariance contributes to segmentation. Overall, our work marks the first attempt to systematically characterize single-neuron invariances in an unbiased and interpretable manner, setting the stage for more comprehensive approaches that blend selectivity and invariance for a deeper understanding of neuronal tuning. By applying this method across various cell types, sensory modalities, and species, we can achieve an enriched understanding of how the brain decodes sensory inputs to isolate behaviorally relevant latent variables.

2-171. Mouse olfactory bulb encodes breathing rhythms and place

Scott Sterrett[1,2] SCOTT0@UW.EDU
Morgan Brown[3] MORGANALLEN@GMAIL.COM
Reese Findley[3] TFINDLEY@UOREGON.EDU
Aldis Weibel[3] AWEIBLE@UOREGON.EDU
Sid Rafilson[3] SIDR@UOREGON.EDU
Mike Wehr[3] WEHR@UOREGON.EDU
James Murray[3,4] JMURRAY9@UOREGON.EDU
Adrienne Fairhall[5] FAIRHALL@UW.EDU
Matt Smear[3] SMEAR@UOREGON.EDU

[1]University of Washington, Seattle
[2]Physiology and Biophysics

[3]University of Oregon
[4]Institute of Neuroscience
[5]University of Washington

Sensory areas in the rodent brain have recently been shown to exhibit widespread activity that is driven by movement and other contextual signals. How is this information integrated with incoming sensory signals and what role might it play in neural computation? We investigated these questions in the olfactory bulb (OB) of freely moving mice in the absence of explicit odor cues, task, or reward structure. We find that units in the OB track long timescale structure of breathing rhythms, as well as the animal's allocentric location. However, because mice often breathe differently in different parts of an arena, it is important to determine whether sniff and place make unique contributions. To tease correlated variables apart, we use a nested generalized linear model to quantify the mixed encoding of olfactory bulb neurons to sniffing and place. These analyses reveal that both make unique contributions to OB activity, so place selectivity in the OB is not merely a reflection of sniffing behavior. Further, we can decode place from populations of OB neurons with errors comparable to hippocampus. Why should the olfactory bulb, in which the nose connects to the brain, encode context? To better understand the function of this mixed encoding, we train recurrent neural networks (RNNs) to perform a simulated olfactory navigation task. We find that the network activity of task-trained RNNs displays mixed encoding, which provides normative insights into the role of contextual modulations of odor encoding. Our results show that the OB of freely moving mice encodes contextual information alongside odor cues, suggesting that the integration of odor information into cognitive maps of the environment begins as soon as olfactory information enters the brain.

2-172. Prioritized dynamical learning of shared dynamics across brain regions

Trisha Jani[1] TJANI@USC.EDU
Bijan Pesaran[2] BIJAN.PESARAN@PENNMEDICINE.UPENN.EDU
Maryam Shanechi[1] SHANECHI@USC.EDU

[1]University of Southern California
[2]University of Pennsylvania

Most tasks rely on interactions among several brain regions. Prior works studying these interactions have often utilized static models that do not consider the dynamic nature of the data. Recent methods have also made progress in accounting for dynamics, but a remaining challenge is that shared dynamics between brain regions may be mistaken for or masked by within-region dynamics. We address this challenge by prioritizing the learning of shared dynamics over other dynamics in a dynamical system. We provide a new formulation for prioritizing inter-regional dynamics by extending the ideas in preferential subspace identification (PSID). Within a 2-stage learning approach, we learn the shared states in a first stage, i.e., learn them first and with priority. This stage finds the shared dynamical subspace between the two regions via projection-based subspace identification. Then, in a second stage, we learn any residual within-region dynamics. To show that multi-stage learning, and the prioritization enabled by it, are important for accurate learning of inter-regional dynamics, we formulate a second block-structured dynamical modeling method. This method numerically optimizes the joint log-likelihood of both shared and within-area dynamics in a single stage, i.e., without prioritizing the shared dynamics. In simulations and neural data from non-human primates (NHP), we show that prioritized dynamical learning: i) predicts the shared dynamics better than both static methods and dynamic methods without prioritization, including a recent state-of-the-art dynamic method; ii) more accurately finds the dimensionality of shared dynamics, and iii) can quantify the dominant interaction pathways between bilateral premotor and primary motor cortical areas, which serves as its further validation. Our results address the challenge of learning shared dynamics between brain regions with priority to mitigate them being masked by, confounded by, or mistaken for within-region dynamics.

2-173. Distilling decision-making dynamics with low-dimensional architectures

Huadong Xiong[1,2] SAKIMARQUIS@GMAIL.COM
Li Ji-An[3] JIAN.LI.ACAD@GMAIL.COM
Robert Wilson[1] BOB@ARIZONA.EDU
Marcelo Mattar[4,5] MARCELO.MATTAR@NYU.EDU

[1]University of Arizona
[2]Psychology Department
[3]University of California, San Diego
[4]New York University

[5]Psychology

Recent advances in examining biological decision-making behaviors have increasingly favored recurrent neural networks (RNNs) over traditional cognitive models grounded in normative principles such as reinforcement learning. This shift owes to RNN's superior predictive performance on behavioral data, achieved with minimal manual engineering. To glean insights into biological decision-making through these networks, this approach identifies a compact set of latent dynamical variables representing key underlying factors by limiting the bottleneck size in the recurrent layer. Yet, the differences and practical effectiveness of these low-dimensional RNN architectures in capturing behavioral patterns of biological decision-making remain largely unknown. Here, we offer a comprehensive comparison of these low-dimensional RNN architectures, evaluating their ability to predict human behavior in an explore-exploit task and to yield insights into the underlying cognitive mechanisms. Remarkably, our findings highlight the efficacy of low-rank RNNs and disentangled RNNs over alternatives like gated recurrent units in this task setting. Moreover, these low-dimensional RNNs reveal diverse strategies that individuals employ across different decision-making phases, advancing our understanding of intricate human decision-making dynamics. Our approach offers a powerful framework for discerning individual cognitive nuances.

2-174. A residue-number attractor neural network model of error-correcting updates among grid cell modules

Christopher Kymn[1] CJKYMN@BERKELEY.EDU
Connor Bybee[1] BYBEE@BERKELEY.EDU
Sonia Mazelet[2] SONIA.MAZELET@BERKELEY.EDU
Denis Kleyko[3] DENIS.KLEYKO@GMAIL.COM
Bruno Olshausen[2] BAOLSHAUSEN@BERKELEY.EDU

[1]University of California, Berekley
[2]University of California, Berkeley
[3]Orebro University

Attractor neural networks are a prominent circuit model for how the brain stores memories and represents variables such as spatial position. It is commonly believed that multiple attractor neural networks co-exist within mammalian hippocampus and entorhinal cortex, jointly encoding cognitively relevant variables such as spatial position and environmental context. Because these networks both represent spatial position, an open question is how these networks interact with each other to update information in a self-consistent manner. Here we report a new type of attractor network, based on a residue number system, that sheds light on this question by providing a minimalistic, mechanistic hypothesis for how grid cell modules interact with inputs from place cells. A residue number system is a framework for representing and computing over integers in terms of their remainder values with respect to a set of base moduli. Prior work suggested that grid cells implement a residue number system, and that this defines an error-correcting code with resolution scaling exponentially in the number of neurons. We extend this line of work in two important ways: first, by defining an attractor neural network achieving an exponential (exp(n ln n)) number of stable fixed points in the number of stored patterns (n), and second, by defining this system in terms of non-negative hexagonal coordinates, connecting residue number systems to frame theory in signal processing. Remarkably, the proposed framework generalizes to sub-integer encodings, enabling continuous path integration. We provide theoretical arguments and experimental sim- ulations confirming that this attractor network has high capacity for a fixed number of neurons and that the dynamics enforce robust error-correction even in the face of high levels of noise. Finally, we discuss experimental predictions of our model, showing it provides a normative framework for understanding computations in grid cells.

2-175. Unraveling the Geometry of Visual Relational Reasoning

Jiaqi Shang[1] JIAQISHANG@G.HARVARD.EDU
Gabriel Kreiman[2] GABRIEL.KREIMAN@TCH.HARVARD.EDU
Haim Sompolinsky[3] HSOMPOLINSKY@MCB.HARVARD.EDU

[1]Harvard University
[2]Boston Children's Hospital, Harvard Medical School
[3]Harvard University, Hebrew University

When perceiving visual scenes, we recognize individual objects and their relations. Deciphering the neural mechanisms underlying relational reasoning from high-dimensional neural activities remains challenging. Our study focuses on a simplified version of the Raven's Progressive Matrices (RPM) task [1], which tests relational reasoning by asking subjects to identify a relational rule (e.g., constant shape) from rows of images and complete a

test row with the same rule. We posit that the brain encodes visual relational information in rule-specific neural manifolds. We reframe RPM as a few-shot learning task and apply a geometrical theory [2] to study what neural representation is conducive to relational reasoning.

The geometrical theory directly links RPM task errors to four key representation geometrical properties: signal, bias, dimensionality, and signal-noise overlap. Using relational networks incorporating convolutions and relational encoding of rows of visual panels, trained on RPM, we demonstrate that the theory accurately predicts model task performance. Further analysis of the model representations reveals two distinct strategies for effective relational reasoning: one prioritizing signal maximization and the other focusing on high dimensionality. These potential mechanisms can be easily tested with experimental neural data. Furthermore, our model, through trained solely to classify different rules, reflects the relational semantics of rules in its error patterns. The model also shows higher task errors for abstract relations that humans find difficult. Tracing the contributing geometrical terms to the differences in abstract relation difficulties provides insights into how the brain processes various abstract relations. In conclusion, our research provides new insight into the neural mechanism of relational reasoning through the lens of representation geometry.

2-176. Modeling Full-Body Human Musculoskeletal System and Locomotion Neural Control with Hierarchical Low-Dimensional Representation

Kaibo He HKB21@MAILS.TSINGHUA.EDU.CN
Chenhui Zuo ZUOCH22@MAILS.TSINGHUA.EDU.CN
Jing Shao SHAO-J20@MAILS.TSINGHUA.EDU.CN
Yanan Sui YSUI@TSINGHUA.EDU.CN

Tsinghua University

Building a controllable human musculoskeletal system is important for the investigation of motor functions. However, current open-source models are restricted to a limited range of body parts and often with a reduced number of muscles. There is also a lack of algorithms capable of controlling over 600 muscles to generate reasonable human movements. To fill this gap, we build a comprehensive musculoskeletal model with 90 body segments, 206 joints, and 700 muscle-tendon units, allowing simulation of full-body neural excitation driven dynamics and interaction with various devices. We develop a novel algorithm using low-dimensional representation and hierarchical deep reinforcement learning to achieve neural-driven full-body motion control. We validate the effectiveness of our model and algorithm in simulations and on real human locomotion data. The musculoskeletal model, along with its control algorithm, will promote a deeper understanding of human motor control.

2-177. A collicular mechano-sensorimotor map of touch events on the tongue surface in mice

Brendan Ito[1,2] BSI8@CORNELL.EDU
Yongjie Gao[1] YG499@CORNELL.EDU
Brian Kardon[1] BMK27@CORNELL.EDU
Jesse Goldberg[1] JHG285@CORNELL.EDU

[1]Cornell University
[2]Neurobiology and Behavior

Successful actions require the use of endpoint feedback to guide future movements. While reaches can be guided by visual feedback, tongue movements are unique in that visual feedback is often unavailable, and instead require tactile feedback. To examine how touch guides tongue movements, we tracked tongue kinematics in 3D as head-fixed mice retrieved water from a moving spout. We detected the offset of tongue-spout contact on the first lick (L1) in a lick bout and randomly displaced the spout to the left or right such that the next lick (L2) would touch the spout with the left or right side of the tongue. Mice rapidly used contact location to re-aim the next lick in the bout: left contacts guided the third lick (L3) left, and right contacts guided L3 right. We used closed-loop photoinhibition to screen brain regions implicated in directional licking. Surprisingly, bilateral inactivation of cortical and cerebellar regions during L3 did not impair touch-guided re-aiming. Interestingly, bilateral lateral superior colliculus (latSC) inactivation halted ongoing licking, and unilateral latSC inactivation abolished contralateral re-aiming, analogous to unilateral SC inactivation during visually-guided saccade tasks. For visuomotor control, SC contains a retinotopic map of visual space that directs saccades to novel visual targets. We wondered whether latSC contains an analogous mechano-sensorimotor map that supports directing licks to tactile targets. Viral tracing revealed that latSC receives input from tongue sensory neurons. Photostimulation of anterior-latSC guided licks more laterally than posterior-latSC. Finally, latSC electrophysiological recordings revealed that neural activity was rapidly modulated by contact location on the tongue at both the single neuron and population level. Our

findings demonstrate for the first time that latSC contains a mechano-sensorimotor map of touch events on the tongue used to guide licks to tactile targets, analogous to what has been observed for visuomotor control of saccades.

2-178. From connectome to effectome: learning the causal interaction map of the fly brain

Dean Pospisil[1,2] DEAN.ABRAHAM.POSPISIL@GMAIL.COM
Max J Aragon[1] MJARAGON@PRINCETON.EDU
Sven Dorkenwald[1] SVENMD@PRINCETON.EDU
Arie Matsliah[1] ARIE@PRINCETON.EDU
Amy R Sterling[1] AMY@EYEWIRE.ORG
Philipp Schlegel[3] PMS70@CAM.AC.UK
Szi-chieh Yu[1] SZICHIEH@PRINCETON.EDU
Claire E McKellar[1] CLAIREMCKELLAR@GMAIL.COM
Marta Costa[3] MMC46@CAM.AC.UK
Katharina Eichler[3] KE327@CAM.AC.UK
Gregory SXE Jefferis[3] JEFFERIS@MRC-LMB.CAM.AC.UK
Mala Murthy[1] MMURTHY@PRINCETON.EDU
Jonathan Pillow[1] JPILLOW@GMAIL.COM

[1]Princeton University
[2]Princeton Neuroscience Institute
[3]University of Cambridge

A long-standing goal of neuroscience is to obtain a causal model of the nervous system. This would allow neuroscientists to explain animal behavior in terms of the dynamic interactions between neurons. The recently reported whole-brain fly connectome specifies the synaptic paths by which neurons can affect each other but not whether, or how, they do affect each other in vivo. To overcome this limitation, we introduce a novel combined experimental and statistical strategy for efficiently learning a causal model of the fly brain, which we refer to as the "effectome". Specifically, the effectome is the dynamics matrix, Wr, specifying the strength and sign of the causal effect each neuron has on every other neuron. We propose a consistent estimator of this dynamics matrix that uses stochastic optogenetic perturbation data to accurately estimate causal effects and the connectome as a prior on the effectome to drastically improve estimation efficiency. Even so, it would be infeasible in the fly—and most organisms—to independently stimulate and record from all neurons at once. The effectome thus would need to be gradually constrained across many experiments. It is unclear how to order experiments such that insight into whole brain dynamics are achieved efficiently. We take the approach of analyzing the connectome to propose circuits that have the greatest total effect on the dynamics of the fly brain. We discover that, fortunately, the dominant circuits significantly involve only relatively small populations of neurons. Intriguingly, we find that this approach also re-discovers known circuits and generates testable hypotheses about their dynamics. Overall, our analyses of the connectome provide evidence that global dynamics are generated by a large collection of small circuits. This in turn implies that a causal model of a brain can be feasibly obtained in the fly.

3-001. Neural encoding of eye-head gaze shifts by single cells in monkey Superior Colliculus

John van Opstal[1,2] VANOPSTALJOHN@GMAIL.COM

[1]Donders Centre for Neuroscience, Radboud University
[2]Neurophysics

The midbrain superior colliculus (SC) is a crucial sensorimotor stage for programming and generating saccadic eye-head gaze shifts. Although it is well established that SC cells encode a neural command that specifies amplitude and direction of the gaze-shift vector, there is controversy about the role of the firing-rate dynamics of these neurons during saccades. Earlier work has proposed a simple quantitative model that explains how the recruited SC population may specify the detailed kinematics (trajectories and velocity profiles) of head-restrained saccadic eye movements[1,2]. We here show that similar principles may apply to a wide range of saccadic eye-head gaze shifts with strongly varying kinematics, despite the substantial nonlinearities and redundancy in programming goal-directed eye-head gaze shifts. Our findings provide additional evidence for an important role of the SC in the optimal control of eye-head saccades.

3-002. Specification curve analysis of representational similarity findings using fMRI and EEG processing pipelines

Satwick Sen Sarma[1,2] SATWICK22@IITK.AC.IN
Gouravmoy Boruah[1,2] GOURAVMOY22@IITK.AC.IN
Nisheeth Srivastava[1] NSRIVAST@IITK.AC.IN

[1]Indian Institute of Technology Kanpur
[2]Department of Cognitive Science

Neuroimaging data analysis increasingly relies on complex processing pipelines, with processing steps implemented using specific parametric or conceptual specifications. Specification curve analysis (SCA) has been recently introduced as a way of inferring the robustness of claimed results to alternative specifications of analysis decisions (Simonsohn, Simmons & Nelson, 2020). In this work, we devised alternate specifications for EEG and fMRI analyses and conducted specification curve analysis. For fMRI, we devised 13 theoretically reasonable alternative specifications to a particular analysis (Hebart et al. 2023) to identify whether Representational Dissimilarity Matrices (RDMs) calculated using these analyses were informative of condition labels, as was the case for the original specification. SCA was conducted by shuffling the labels for the dataset resulting from each analysis, and testing whether the RDM obtained from the unshuffled data was significantly different from RDMs obtained from 20 shuffled samples of the obtained dataset. SCA results indicated statistical non-significance across all specifications except the original specification, indicating that the RDMs obtained from all other analyses did not contain the condition-identifying representational structure contained in the analysis following the original specification. We likewise compared the representational dynamics for EEG by computing time-varying decoding accuracies derived from RDMs for 10 different specifications to a particular analysis (Grootswagers, T. et al. 2022). We found a high correlation across k-values, indicating that EEG RDMs, unlike fMRI, are relatively robust to specification changes. Thus, we show that, for representational similarity analyses, for the representative analyses we tested, EEG-based findings are robust while fMRI-based findings are not.

3-003. Machine learning of functional network and molecular mechanisms in autism spectrum disorder subtypes

Amanda Buch[1,2] AMB2022@MED.CORNELL.EDU
Petra Vertes[3] PV226@CAM.AC.UK
Jakob Seidlitz[4] SEIDLITZJ@CHOP.EDU
So Hyun Kim[5] SOK2015@MED.CORNELL.EDU
Logan Grosenick[1] LOG4002@MED.CORNELL.EDU
Conor Liston[1] COL2004@MED.CORNELL.EDU

[1]Weill Cornell Medicine, Cornell University
[2]Psychiatry
[3]University of Cambridge
[4]University of Pennsylvania and Children's Hospital of Philadelphia
[5]Weill Cornell Medicine and Korea University

Autism Spectrum Disorder (ASD) describes a diverse group of neurodevelopmental disorders encompassing a wide range of clinical impairments. The two core symptoms that define ASD are social communication impairments and restricted and repetitive behaviors, and there is a wide range of cognitive and language abilities. How distinct neurobiological substrates give rise to differing clinical symptoms in subsets of ASD patients is unknown. Using a large, publicly available neuroimaging dataset comprising resting state functional magnetic resonance imaging (rsfMRI) scans from N=299 subjects with ASD and N=907 neurotypical controls, we identified three latent dimensions of functional brain network connectivity that predict individual differences in ASD symptoms and behaviors. We show that patients with ASD can be grouped into distinct neurophysiological subgroups based on patterns of dysfunctional connectivity and clinical behaviors. In this cohort, functional connectivity features were extracted from rsfMRI data, regularized canonical correlation analysis was used to identify associations between connectivity features and behavioral data, and ASD subjects were clustered along these dimensions. Cross-validation analyses showed high stability in the brain-behavior dimensions, with replicable clusters in held-out data and a second out-of-sample dataset (National Database for Autism Research; NDAR). Next, we integrated neuroimaging data with gene expression data from the Allen Human Brain Atlas, and found that within each subgroup, ASD-related functional connectivity was explained by regional differences in the expression of distinct gene sets. We replicated these findings using the BrainSpan developmental atlas. These were enriched for transcriptionally-regulated and ASD-associated genes along with immune and synaptic signaling pathways. In sum, our results identify discrete ASD subgroups associated with specific ASD behaviors and neurophysiological signatures, and these different forms of ASD implicate distinct genetic mechanisms. The results of this study suggest a promising new approach for understanding the neurobiological substrates of ASD.

3-004. Dynamics of learning in the non-linear perceptron

Christian Schmid[1] CHRISTIAN1.SCHMID@GMAIL.COM
James Murray[1,2] JMURRAY9@UOREGON.EDU

[1]University of Oregon
[2]Institute of Neuroscience

The ability of a brain or a neural network to efficiently learn depends crucially on both the task structure and the learning rule. Previous works have analyzed the dynamical equations describing learning in the relatively simplified context of the perceptron under assumptions of a student-teacher framework or a linearized output. However, in addition to being highly questionable on biological grounds, these assumptions have precluded a detailed understanding of the roles of the nonlinearity and input-data distribution in determining the learning dynamics. Here, we use a stochastic-process approach to derive flow equations describing learning, applying this framework to the case of a nonlinear perceptron performing binary classification. We characterize the effects of the learning rule (supervised or reinforcement learning, SL/RL) and input-data distribution on the perceptron's learning curve, the fixed point to which learning converges, and the forgetting curve as subsequent tasks are learned. In particular, we quantify how the input-data noise differently affects the learning speed under SL vs. RL, as well as determines how quickly learning of a task is overwritten by subsequent learning. More generally, this approach points a way toward analyzing learning dynamics for more-complex circuit architectures.

3-005. Koopman Spectral Analysis Uncovers the Temporal Structure of Spontaneous Neural Events

Kaidi Shao[1] KAIDI.SHAO@ICPBR.AC.CN
Yuanchao Xu[2] YUANCHAO@UALBERTA.CA
Nikos Logothetis[1] NIKOS.LOGOTHETIS@ICPBR.AC.CN
Zhongwei Shen[2] ZHONGWEI@UALBERTA.CA
Michel Besserve[3] MICHEL.BESSERVE@TUEBINGEN.MPG.DE

[1]International Center for Primate Brain Research (ICPBR), CEBSIT, Chinese Academy of Science (CAS)
[2]University of Alberta
[3]Max Planck Institute for Intelligent Systems

Neural information processing requires precise coordination of brain activities across various spatiotemporal scales, reflected by the occurrence of neural events. A paradigmatic case is the emergence of hippocampal Sharp Wave-Ripples (SPW-Rs), synchronizing widely distributed cellular assemblies during memory consolidation. Reliably defining and detecting such events in multichannel neural signals with minimal prior assumptions is therefore key for understanding brain function, but remains challenging. Here, we rely on the detection of "salient" transient dynamics in nonlinear dynamical systems based on Koopman Operator Theory. In this framework, a nonlinear dynamical system underlying neural signals can be transformed into a linearly evolving system governed by the Koopman operator, thus benefiting from the well-established linear system theory. Through spectral decomposition, we can approximate dominant eigenvalues and associated eigenfunctions of the operator, uncovering dynamics across a wide range of timescales, where the salient transient values of the dominant eigenfunctions can then be used to define events, thereby avoiding prior frequency range-based detection. We implement this decomposition using Extended Dynamic Mode Decomposition (EDMD), with bases learned through a multi-layer neural network. Applying these methods to simulations of a regime-switching neural mass model reveals activities of three distinct timescales representing regimes, transitions, and fast oscillations intrinsic to the system. The transition timescale naturally marks the occurrence of oscillatory events. When applying our method to hippocampal electrophysiological signals exhibiting three types of transient events, we successfully detect these events across different regions, outperforming classical methods such as bandpass filtering. We believe that Koopman Operator-based spectral analysis not only introduces a novel approach to multi-source neural event detection but also presents a promising theoretical framework for disentangling various timescales in neural signals and gaining insights into their temporal structures. This framework holds the potential to advance our understanding of neural event dynamics on a mechanistic level.

3-006. A disinhibitory basal forebrain to cortex projection supports sustained attention

Shu-Jing Li[1] SHUJING@WUSTL.EDU
Balazs Hangya[2] HANGYA.BALAZS@KOKI.HU
Unmukt Gupta[3] G.UNMUKT@WUSTL.EDU
Adam Kepecs[4] AKEPECS@WUSTL.EDU

[1]Washington University, St. Louis
[2]Institute of Experimental Medicine, Hungarian Academy of Sciences, Budapest, Hungary
[3]Departments of Neuroscience and Psychiatry, Washington University School of Medicine, St. Louis
[4]Washington University St. Louis

Sustaining attention over time requires cognitive effort, a limited resource, which is subject to ongoing fluctuations. The neural origins of these fluctuations and their underlying resource constraints remain unknown. Here we combine behavioral, neural circuit, and computational approaches to examine how sustained attention fluctuates and what drives its allocation. We identified a neural circuit with key attributes relevant to sustained attention: cortex-projecting basal forebrain parvalbumin-expressing GABAergic neurons (BF-PV) that preferentially target cortical inhibitory neurons. Optogenetic activation of BF-PV neurons disinhibits and selectively amplifies sensory activity in the auditory cortex. In a mouse sustained attention task, we found that the activity of BF-PV neurons predicted performance metrics such as reaction time and accuracy. Furthermore, optogenetic stimulation of BF-PV neurons enhanced task performance, indicating a causal role in modulating attention. BF-PV neurons also responded to motivationally significant events in cued-outcome tasks: predictive cues, rewards, punishments, and surprises. To interpret the complex response properties of BF-PV neurons, we developed a model employing a partially observed Markov decision process (POMDP) within a reinforcement learning framework to simulate key cognitive variables. This model accurately predicts detection performance, reaction times, and BF-PV neuronal activity, reflecting a combination of unsigned value and prediction error. The model encapsulates the concept that diverse BF-PV neuron responses reflect motivational salience, which guides attention allocation and provides a comprehensive framework for understanding attentional resource constraints. In summary, our study reveals a disinhibitory basal forebrain-to-cortex projection that modulates cortical gain based on a motivational salience computation, thereby promoting sustained attention.

3-007. Influence of working memory limitations and dopamine on evidence accumulation

Cina Aghamohammadi[1] CA6941@PRINCETON.EDU
Christopher Langdon CL1704@PRINCETON.EDU
Tatiana Engel[2] TATIANA.ENGEL@PRINCETON.EDU
Jochem van Kempen[3] JOCHEMVANKEMPEN@GMAIL.COM
Molly Stapleton[3] MOLLY.STAPLETON@HOTMAIL.CO.UK
Alwin Gieselmann[3] ALWIN.GIESELMANN@NEWCASTLE.AC.UK
Alexander Thiele[3] ALEX.THIELE@NEWCASTLE.AC.UK

[1]PNI, Princeton University
[2]Princeton University
[3]Biosciences Institute, Newcastle University

Complex decisions involve evidence accumulation over long periods, which requires maintaining and updating the representation of the cumulative evidence in the working memory. Both working memory and decision-making depend on the prefrontal and parietal cortical areas and are sensitive to dopamine. Despite shared behavioral demands and neural circuitry, the impact of working memory on decision-making is unappreciated. Standard decision-making tasks have low working memory demands with short trials. Thus, how working memory constraints affect decision-making performance remains unknown. We trained two monkeys to perform a task that required integrating sequential pieces of evidence with varying reliability over long periods. The monkeys made decisions based on a sequence of shapes, each providing probabilistic evidence for two choice alternatives. Despite increasing cumulative evidence with a larger number of observed shapes, monkeys' rewards decreased, suggesting that they did not use all available evidence to guide their choices. To identify the processes causing evidence loss, we modeled choice behavior, exploring multiple hypotheses for working memory limitations: finite capacity, temporal decay, primacy, recency, and priming. Memory decay explained the choice behavior and inter-animal differences. We tested the role of dopamine in regulating the working memory constraints on decision-making using systemic drug applications. The model revealed that dopamine modulates both sensory evidence gain and memory decay rate. During the task, we recorded spiking activity from dorsolateral prefrontal (dlPFC) and lateral intraparietal (LIP) cortical areas. We found that responses of single LIP neurons clustered according to their tuning to cumulative evidence, and dlPFC neurons clustered according to their tuning to the

weight of the displayed shape. Differ- ent clusters of LIP and dLPFc neurons encoded distinct information about the cumulative evidence and weight, respectively. Our work identifies mechanisms by which working memory limitations affect decision-making, indicates that dopamine regulates these limitations, and reveals the associated neural representations.

3-008. Temporal information encoding in isolated cortical networks

Yevgeny Berdichevsky[1,2] YEB211@LEHIGH.EDU
Zubayer Ibne Ferdous[1] ZUI216@LEHIGH.EDU

[1]Lehigh University
[2]Bioengineering and Electrical and Computer Engineering

Many functions of the brain rely on the ability to tell time. Sensory information has both spatial and temporal organization; and cortical circuits must be able to work with past information, and to know exactly what point in the past this information describes. Examples where this ability may be critical include processing of moving visual information and processing of speech. It has been proposed that time-dependent neuronal and synaptic properties may contribute to the cortex's ability to tell time on the order of hundreds of milliseconds1. The computational mechanism by which this may occur has been described as 'reservoir computing'2. Essentially, a network of recurrently connected neurons represents a system whose state depends on the inputs as well as the previous state of the network, represented by activity of different neurons. In the experimental study reported here, we determine whether the state of an isolated cortical network can be used to accurately determine the timing of occurrence of an input pattern - in a sense, to tell time. We used an experimental system based on patterned optogenetic stimulation of dissociated cortical cultures, and read out activity via fluorescent calcium indicator. We delivered inputs organized into sequences of patterns such that a pattern of interest occurred at different times. We developed a readout function for network state based on a support vector machine (SVM) with recursive feature elimination and custom error correcting output code. We found that the state of these experimental networks contains information about inputs up to at least 900 msec. Timing of input pattern occurrence could be determined with 100 msec precision. Accurate classification required the use of a large number of neurons, suggesting that timing information was encoded via population code. Our overall conclusion was that an isolated cortical network could indeed 'tell time'.

3-009. Hippocampal sequences span experience relative to rewards

Mari Sosa[1,2] MSOSA2@STANFORD.EDU
Mark Plitt MARKPLITT@GMAIL.COM
Lisa Giocomo[1] GIOCOMO@STANFORD.EDU

[1]Stanford University
[2]Neurobiology

Positive experiences must be remembered strongly to reinforce rewarding behaviors, and these memories must be able to update when knowledge of reward changes. How does the brain flexibly amplify memories of events surrounding reward while maintaining a stable representation of the external world? The hippocampus provides a potential neural circuit for this memory process. Hippocampal place cells fire in sequences that span spatial environments, and "remap", or change their preferred firing locations, across different environments1. This sequential firing is common to multiple modalities beyond space2, suggesting that hippocampal activity can anchor to the most behaviorally relevant or salient aspects of experience. As reward is a highly salient event, we hypothesized that broad sequences of hippocampal activity can likewise become anchored relative to reward, such that moving the reward within a constant environment should induce remapping even at locations far from reward. To test this hypothesis, we performed two-photon (2P) imaging of calcium activity in hippocampal area CA1 as mice navigated virtual linear environments with multiple changing hidden reward locations. We found that when the reward moved, a subpopulation of cells remapped to the same relative position with respect to reward, including cells with fields distant from reward. These "reward-relative" cells constructed sequences that spanned the task structure irrespective of spatial stimuli. The density of the reward-relative sequences increased with task experience as additional neurons were recruited to the reward-relative population, putatively providing a more precise code for estimated distance from reward as mice learned each reward update. In contrast, a largely separate subpopulation of cells maintained a stable place code. This work provides insight into how separate hippocampal ensembles may flexibly encode the most behaviorally relevant aspects of experience, by amplifying a learned representation of events surrounding reward in parallel to a stable code for the external environment.

3-010. Non-Hebbian rewiring of olfactory cortex by experience

Andrew Fink[1,2] AF2243@COLUMBIA.EDU
Samuel Muscinelli[1,3] SPM2176@COLUMBIA.EDU
Shuqi Wang[1] SW3894@COLUMBIA.EDU
Marcus Hogan[4] MIH2121@COLUMBIA.EDU
Richard Axel[5] RA27@COLUMBIA.EDU
Ashok Litwin-Kumar[1] AK3625@COLUMBIA.EDU
Carl Schoonover[1] CES2001@COLUMBIA.EDU

[1]Columbia University
[2]Neuroscience
[3]Zuckerman Institute
[4]University of California, Berkeley
[5]Columbia University / HHMI

The Hebbian model of learning posits the enrichment of synaptic coupling between neurons that have similar response properties, resulting in like-to-like connectivity (Hebb, 1949). Observing this phenomenon in the brain has been challenging because of the difficulty of estimating both synaptic connectivity and response properties in vivo. The mouse olfactory cortex (piriform) presents an opportunity to directly test Hebb's postulate and measure whether experience produces like-to-like connectivity. We developed a deep neural network approach that takes advantage of the piriform's ultra sparse recurrent connectivity to infer, at low error rate, the excitatory synaptic connections between pairs of piriform neurons from recordings of their spike time cross-correlograms (Perkel et al., 1967; Csicsvari et al., 1999; Bartho et al., 2004; English et al., 2017). We discovered that in naive animals neurons with similar odor response properties are more likely to be connected. This organization, however, may reflect local spatial structure in the network, rather than the formation of cell assemblies by a Hebbian process. We therefore asked whether experience with odors enhances this like-to-like motif. We observed that sensory experience profoundly alters the connectivity of the piriform network, but not in a manner consistent with the Hebbian postulate: experience does not enrich connectivity between excitatory neurons with similar response properties. Rather, this reorganization of the connectivity is independent of the tuning of the presynaptic neuron. Instead, we observe plasticity of connections onto inhibitory neurons that is determined by the degree of odor selectivity of the postsynaptic neuron. Specifically, synapses onto odor-selective neurons are enriched, whereas synapses onto nonselective neurons are depleted. This novel plasticity rule may serve to decorrelate and expand representations (Recanatesi et al., 2019) and balance network activity in a non-topographic network (Rupprecht 2018, Mackwood et al., 2021).

3-011. A new approach for testing specific hypotheses about probabilistic representations

Adam Koblinger[1,2] KOBLINGER_ADAM@PHD.CEU.EDU
Theoklitos Amvrosiadis[3,4] TA533@CAM.AC.UK
Nathalie L. Rochefort[3] N.ROCHEFORT@ED.AC.UK
Mate Lengyel[5] M.LENGYEL@ENG.CAM.AC.UK

[1]Central European University
[2]Cognitive Science Department
[3]University of Edinburgh
[4]Centre for Discovery Brain Science
[5]University of Cambridge; Central European University

There are multiple competing theoretical proposals for how neural activities may represent posterior probability distributions about relevant latent variables. However, no empirical approach exists for systematically arbitrating between these proposals. Here, we develop an approach for systematically evaluating candidate representations that differ along two axes: the neural code for posteriors, and the latent variables for which posteriors are computed. We also present proof-of-concept results from applying our analysis approach to mouse primary visual cortex (V1) calcium imaging data in a perceptual decision making task. Our approach combines model-based analyses of behavioral data and population decoding analyses of simultaneously recorded neural data. Specifically, we develop a Bayesian ideal observer that in each trial computes posterior distributions over two latent variables: a low level perceptual variable (the orientation of the stimulus), and a high level decision variable (which of two responses is correct). We fit this model to the trial-by-trial behavior of individual subjects and use the posteriors it computes on each trial as the target for a population decoder algorithm that is applied to neural responses recorded in the same trial. We develop two different decoders: one that uses the within-trial temporal patterns of responses, and the other that uses the spatial pattern of responses across the population. Critically, for a fair comparison, the two decoders have closely analogous structure and the same number of parameters.

We use the normalized Kullback-Leibler divergence between behaviorally and neurally decoded posteriors to compare 2-by-2 hypotheses differing in whether posteriors are computed over perceptual or decision variables, and whether these posteriors are represented by a spatial or a temporal code. After extensively validating our approach on synthetic data, we find preliminary evidence for temporal codes for perceptual posteriors in mouse V1. Our approach should be widely applicable to a large variety of paradigms.

3-012. Top-down computations in a hierarchical generative model of primate V1 and V2

Ferenc Csikor[1,2] FERENC.CSIKOR@GMAIL.COM
Balazs Meszena[1] MESZENAB@GMAIL.COM
Gergo Orban[3,4] ORGERGO@GMAIL.COM

[1]Wigner Research Centre for Physics
[2]Department of Computational Sciences, Computational Systems Neuroscience Lab
[3]Wigner RCP
[4]Dep Computational Sciences

Interpreting computations in the visual cortex as unconscious inference in a generative model of the statistics of the environment has sparked long term interest in the neuroscience community. Encouragingly, generative models of low level vision gave promising insights into the means and variances of stimulus induced neural responses in V1, but extending these to higher cortical areas has proven methodologically challenging. Here we overcome some of the problems by creating a two-latent-layer hierarchical Variational Autoencoder (VAE) and adapting it to known properties of the primate visual cortex through mild inductive biases. We find that after training on natural images the model learns a hierarchical image representation with the two latent layers corresponding to V1 simple cells and texture selective neurons in V2 in the macaque visual cortex. Prominently, the architecture of the model includes a top-down path in its recognition model (hence its name, Top-Down VAE, TDVAE). We show that this top-down path learns computations that display known top-down effects in the primate visual cortex (Kanizsa illusion, specificity of noise correlations to high level stimulus structure) and fulfill general desiderata of hierarchical inference (increased reliance on high-level contextual priors for noisy stimuli), making TDVAE a normative account of top-down computations in the visual cortex. In perspective, TDVAE can be scaled to larger image sizes (by introducing parameter sharing, e.g., convolutional) and to more latent layers, turning it into a new methodology for studying inductive biases shaping the primate visual cortex as well as the rich contextual effects present in the visual cortex.

3-013. Simulation-based behavioral profiling by model-guided task optimization and task-guided data generation

Jae Hoon Shin[1,2] SKALCLRPTSP@GMAIL.COM
Sang Wan Lee[1] SANGWAN.KAIST@GMAIL.COM

[1]KAIST
[2]Bio and Brain Engineering

Computational neuroscience has benefitted from task design and data analyses. However, the design is often not optimized enough to elicit desired responses from the brain functions of interest, and behavior analyses or computational modeling can also be undermined by individual variability. To tackle these issues, we propose a novel framework that combines computational modeling, task optimization, and data generation: (1) Computational models are fitted to accommodate individual variability; (2) Tasks are optimized to elicit the desired responses from an individually fitted computational model; (3) Simulated behavior data are obtained from the model of step 1 performing the entire task set of step 2 (Figure 1). As a proof of concept, we applied this to the problem of model-based (MB) and model-free (MF) reinforcement learning (RL), whose behavior patterns are known to be difficult to separate. Using the computational models of MB x MF RL [1] fitted to individual subject's data of 2-stage Markov decision tasks [2], the task structure was optimized to minimize the model's long-term prediction errors (called task-optimizer; Figure 2). To reduce any bias introduced by individual variability, we obtained the model's behavior data (called synthetic data) using task-model permutation, in which each model performs all the individually optimized tasks (Figure 3A). The behavior clustering analyses based on simple dimension reduction confirmed that from the synthetic data, it is easier to read out various characteristics of human RL, including the long-term prediction error (Figure 3B), its sensitivity as parameterized by the computational model (Figure 3C), and the bias towards MB/MF RL (Figure 3D). The proposed framework helps us understand the underlying characteristics of the brain from the potentially biased data obtained with suboptimal task design.

3-014. A comprehensive large-scale model of primary visual cortex (V1)

Shivang Rawat[1,2] SR6364@NYU.EDU
David Heeger[1] DAVID.HEEGER@NYU.EDU
Stefano Martiniani[1,3] SM7683@NYU.EDU

[1]New York University
[2]Mathematics
[3]Physics

We introduce a comprehensive retinotopic model of V1 based on ORGaNICs, a stochastic recurrent circuit framework implementing divisive normalization via recurrent amplification. Specifically, we simulate the membrane potentials and firing rates of simple and complex V1 neurons driven by the outputs of a steerable pyramid, thus capturing the retinotopy, spatial frequency, receptive field size, and orientation-tuning selectivity of the neurons. Then, using the theory of stochastic LTI systems, we demonstrate that, for a grating response, the circuit oscillates at the observed (gamma) frequency and accurately captures the contrast dependence of gamma activity and LFP coherence, measured across neuronal populations at different spatial locations, with different orientation tuning and receptive field size.

We further design a modified Gaussian-Rectification (GR) model for the generation of spiking activity that takes into account the time-correlations of the membrane potentials. We demonstrate that this framework accurately captures the dependence of the Fano factor and noise correlations as a function of stimulus contrast and provides an analytical expression, depending on circuit parameters, for deviations from the Poisson-like behavior of the spiking activity. This spiking activity is then filtered (simulating synaptic filtering) and fed back as input to the dynamical variables simulating the membrane potential of neurons. Finally, we predict these quantities in the context of realistic stimuli: Gabor, plaids, and natural images. Therefore, our framework offers a versatile tool for understanding the dynamics and noise correlation of V1 activity.

3-015. Brain-grounding of Semantic Vectors Improves Neural Decoding of Visual Stimuli

Shirin Vafaei[1,2] SHIRINVAFAEE@GMAIL.COM
Ryohei Fukuma[1] R-FUKUMA@NSURG.MED.OSAKA-U.AC.JP
Huixiang Yang[1] YANGHUIXIANG@NSURG.MED.OSAKA-U.AC.JP
Takufumi Yanagisawa[1] TYANAGISAWA@NSURG.MED.OSAKA-U.AC.JP

[1]Osaka University
[2]Medicine

Developing algorithms for precise neural decoding of visual stimuli is pivotal for advancing brain-machine interfaces (BMIs) and understanding neural representations of objects. Traditionally, neural decoding entails training machine learning models to map neural activity patterns to a "semantic vector representation" of stimuli [1]. Previous studies leveraging vector representations from neural networks trained for object or text processing showcased potential in decoding visual stimuli but faced limitations in generalizability for decoding new, untrained categories [1]. Our study introduces a novel approach to improve decoders' learning capabilities by refining the "vector representations of stimuli" to mimic brain-like features. This refinement allows improved mapping from neural activity patterns to their corresponding vectors, even for novel categories and neuroimaging types not included in creating this integrated semantic space. Our model, "brain-grounding of semantic vectors," employs an autoencoder reconstructing original vectors, while constraining its latent space's representational similarity matrix (RSM) [2] to mirror the RSM of their corresponding brain activity patterns. We trained this autoencoder by utilizing functional magnetic resonance imaging (fMRI) data from 1200 images across 150 categories (8 images/category) and employing two distinct original semantic vector spaces: an image-based and a text-based. Decoding analyses on 50 categories that were not part of the autoencoder's training data revealed increased accuracies when using brain-grounded vectors compared to original vectors. We further extended our analyses to magnetoencephalographic (MEG) brain activity that was obtained while watching the same stimuli and observed that brain-grounded vectors that were trained on fMRI data can significantly increase the MEG decoding accuracies. Overall, our study presents a novel method for creating vector representations of images that (1) significantly improves neural decoding accuracy across both seen and unseen data, and (2) is applicable for neural data recorded by a different neuroimaging modality; hence, the proposed method showcases promising potential in advancing accurate BMIs.

3-016. Bias-corrected synaptic plasticity is essential for capacity in mushroom body circuits

Zhanmiao Huang[1,2] ZHUANGDJ@CONNECT.UST.HK
Yu Hu[1] MAHY@UST.HK

[1]The Hong Kong University of Science and Technology
[2]Department of Mathematics

In fruit flies, the Mushroom Body plays a crucial role in learning odors associated with positive or negative reinforcement signals. Despite extensive experimental and computational studies, the precise mechanisms of learning and plasticity remain open. In particular, the classic Hebbian rule introduces a bias in the readout of Mushroom Body Output Neuron (MBON). There are two solutions to this bias issue: implementing a bias correction within connectivity levels or removing the bias in the readout layer. Although appearing similar in effect, we show that the bias-corrected plasticity yields significantly higher and stable memory capacity. Using a simplified model of the MB circuit, we rigorously derived the error rate and capacity analytically, which also clarifies the parameter regime under which the aforementioned capacity gap can be seen. Our findings extend to more biologically realistic models, including correlated Kenyon Cells (KCs) patterns as the overlaps exist in odor representations. Under realistic parameters corresponding to adult and larval Drosophila, our predicted capacity numbers suggest that the bias-corrected plasticity is likely being implemented in the MB circuit. This predicts a small increase related to activity mean in synaptic strengths between non-activated KCs and the MBON under unconditional stimulus (US)/DAN activation that supplements established plasticity in the circuit and can be tested experimentally. Our work offers theoretical insights into the computational benefit and biologically plausible implementations of adjusting for bias in circuit connectivity.

3-017. Generalized attention benefits that outlast neurofeedback training

Vishesh Choudhary[1,2] VISHESHC@IISC.AC.IN
Devarajan Sridharan[3,4] SRIDHAR@IISC.AC.IN

[1]Centre for Neuroscience, Indian Institute of Science, Bangalore, India
[2]Center for Neuroscience
[3]Indian Institute of Science
[4]Centre for Neuroscience

Paying sustained attention produces rapid and specific changes in neural state [1]. Whether sustained attention abilities can be improved by making participants aware of these changes in neural state – "neurofeedback" – remains an open question [2]. Previous neurofeedback studies in humans employing electro/magnetoencephalography (EEG/MEG) have found limited, and often null, evidence for attentional benefits that outlast training, or for learning effects that generalize beyond the trained spatial location [3].

Here, we tracked, in real-time, EEG steady-state visually evoked potentials (SSVEPs), whose power is reliably modulated by attention [4]. Participants were provided a continuous readout of their momentary attention levels, and lapses, at distinct (attended, unattended) locations, by tagging gratings with two distinct frequencies in the "trained" hemifield (Fig. 1A-B). To test if the learning effect outlasted neurofeedback training, participants performed a cued continuous performance "search" task (CPST) (Fig. 1C), both "pre"- and "post"- neurofeedback. To test if the neurofeedback effect generalized beyond the trained location, CPST performance was tested in both hemifields (Fig. 1C).

In the neurofeedback session, participants (n=8) successfully increased SSVEP power at the attended location while concurrently reducing SSVEP power at the unattended location (Fig. 2A). By contrast, control participants (n=8) who received fake ("yoked") neurofeedback failed to achieve such modulations (Fig. 2A). Behavioral accuracy (d') improved significantly following even a single neurofeedback training session, as compared to control participants (Fig. 2B, p<0.05). The improvement in attention occurred globally, across both trained and untrained hemifields (Fig. 2B). A linear classifier trained on evoked neural responses decoded the attended location significantly better post-neurofeedback (Fig. 2C,D), indicating better separability of attention-related neural states.

Lapses of attention, momentary or prolonged, can have adverse consequences for survival. Our neurofeedback paradigm demonstrates attentional benefits that generalize across space and time, with potential therapeutic applications for attention disorders.

3-018. Arousal as a universal embedding for spatiotemporal brain dynamics

Ryan Raut[1] RYAN.RAUT@ALLENINSTITUTE.ORG
Zachary Rosenthal[2] ZPROSENTHAL@GMAIL.COM
Xiaodan Wang[3] XIAODANWANG@WUSTL.EDU
Hanyang Miao[3] BENMIAO@WUSTL.EDU
Zhanqi Zhang[4,5] ZHZ091@UCSD.EDU
Jin-Moo Lee[3] LEEJM@WUSTL.EDU
Marcus Raichle[3] MRAICHLE@WUSTL.EDU
Adam Bauer[6] AQBAUER@WUSTL.EDU
Steven Brunton[7] SBRUNTON@UW.EDU
Bing Brunton[7] BBRUNTON@UW.EDU
J. Nathan Kutz[7] KUTZ@UW.EDU

[1]Allen Institute & University of Washington, Seattle
[2]University of Pennsylvania
[3]Washington University in St. Louis
[4]University of California, San Diego
[5]Computer Science
[6]Washington University in Saint Louis
[7]University of Washington, Seattle

Neural activity in awake organisms shows widespread and spatiotemporally diverse correlations with behavioral and physiological measurements. We propose that this covariation reflects in part the dynamics of a unified, arousal-related process that regulates brain-wide physiology on the timescale of seconds. In this study, we test a central prediction of this unified perspective: specifically, that a scalar index of arousal (namely, pupil diameter) suffices to accurately reconstruct multimodal measurements of large-scale spatiotemporal brain dynamics. This hypothesis is inspired from Takens' embedding theorem from dynamical systems theory, which permits the (topology-preserving) reconstruction of a full-state attractor manifold from a single observable and its time history. To test this hypothesis, we perform multimodal, cortex-wide optical imaging and pupillometry in awake mice. We further introduce a data-driven procedure to parsimoniously link observables that evolve according to a common but unknown dynamical mechanism. Specifically, we train neural networks to approximate the hypothesized mapping from the time history ("delay embedding") of pupil diameter to widefield optical measurements of neuronal calcium, metabolism, and blood-oxygen. We show that these multimodal spatiotemporal measurements can be accurately and parsimoniously modeled from a low-dimensional state-space reconstruction based on the time history of pupil diameter. Notably, this state space reconstruction enabled us to account for multidimensional variance across the spatiotemporal measurements, while further enabling frame-by-frame decoding of spatial topographies. Our results support the hypothesis that spontaneous, spatially structured fluctuations in brain-wide physiology — widely interpreted to reflect regionally-specific neural communication — are in large part reflections of an arousal-related process. This enriched view of arousal dynamics has broad implications for interpreting observations of brain, body, and behavior as measured across modalities, contexts, and scales.

3-019. Dopamine as a Sensory Prior Prediction Error in the Sensory Striatum

Eleonora Bano[1,2] ELEONORA.BANO@WUSTL.EDU
Amelia Christensen[1] AMELIAC@WUSTL.EDU
Fengrui Zhang[3] FENGUI.ZHANG@WUSTL.EDU
Heejae Choi[3,4] C.HEEJAE@WUSTL.EDU
Adam Kepecs[3] AKEPECS@WUSTL.EDU

[1]Departments of Neuroscience and Psychiatry, Washington University School of Medicine, St. Louis
[2]Neuroscience
[3]Washington University St. Louis
[4]Department of Neuroscience

To behave optimally in a dynamic environment, perceptual priors should adapt when environment statistics change. To study this process, we trained mice on an auditory detection task, with a latent block structure that governed the probability of cue presence on each trial. Mice were able to dynamically update their decision thresholds to track cue probability. Performance of similar auditory decision-making tasks relies on monosynaptic input from the auditory cortex to the auditory striatum (TS); thus, to study how mice learned and used this cue probability prior, we focused on the auditory striatum. We reasoned that dopaminergic input to TS might be ideally positioned to update auditory priors. When we measured dopamine release in the TS during the auditory detection task, we observed large dopamine transients upon cue delivery, precisely time-locked only to perceived cues, but not to report initiation. These cue evoked dopamine transients reflected the difference between prior expectation and sound volume, reminiscent of a prediction error signal comparing the perceived sound to a prior

expectation of its presence. Importantly, if this prediction error serves as a teaching signal for updating of auditory priors, this signal should adjust expectations for future events. Indeed, we found that after adjusting for tone volume, cue evoked dopamine release was more predictive of cue response probability on the next trial than it was of cue detection on the current trial. To test whether mouse cue probability tracking behavior in our task was consistent with online reinforcement learning of statistical priors, we built and fit a hierarchical belief-state reinforcement learning model. This model, with just two parameters, effectively captures the trial-to-trial variations in detection threshold as a function of past perceptual experience. Our findings add support to an emerging view that the cortical-striatal-basal ganglia circuit plays a key role in performing statistical inference.

3-020. A Game of Memory: Learning in Spiking Networks with Preserved Weight Distributions

Maayan Levy[1] MAAYAN.LEVY@IST.AC.AT
Tim Vogels[2,3] TIM.VOGELS@IST.AC.AT

[1]Institute of Science and Technology Austria
[2]IST, Austria
[3]-

How changes in synaptic connections underlie learning and memory is a central question in Neuroscience. Previous modeling efforts focused on biologically realistic learning rules and dynamics, but these rules usually produce bimodal weight distributions that don't align with experimentally observed lognormal distributions of synaptic weights. How biological learning rules preserve such unimodal distributions and at the same time successfully retain information, i.e., learn, remains unknown. Here, we take a game theoretical approach to learning, asking whether learning is possible under the constraint of retaining a particular weight distribution. In other words, we have omitted mechanistic learning rules as a means to changing weights locally. Instead we assume a fixed set of weights and open a game of trading synaptic weights to achieve learning. Towards this goal we set up a spiking neural network with asynchronous and irregular dynamics and a number of fixed input stimuli that produce increased firing rates in their target neurons when active. The aim of the game is to patten-complete network responses to incomplete stimuli, as a proxy for learning. We find that both functional swapping of weights between existing synapses and structural swapping, in which entire connections can be destroyed or created de novo can achieve robust pattern completion, even for small and severely degraded stimuli, but overall network stability can become compromised. To explore the constraints of how a network learns multiple, overlapping stimuli while remaining robust in its dynamics, we expand to a multiplayer game in which each memory is a player and their strategies are the fraction of synaptic connection allowed to change. We identify combinations of strategies of swapping that are posed at equilibria and have the potential to increase network capacity without adding strong synapses. In turn, these constraints can inform the boundaries of biological action in mechanistic rules.

3-021. Fast, sparse, and local learning in motor cortex

Mathew Bull[1] MATTHEW.BULL@ALLENINSTITUTE.ORG
Marton Rozsa[2] MARTON.ROZSA@ALLENINSTITUTE.ORG
Lu Mi[3] LU.MI@ALLENINSTITUTE.ORG
Peter Humphreys[4] PETERHUMPHREYS@DEEPMIND.COM
Maria Eckstein[4] MARIAECKSTEIN@DEEPMIND.COM
Kimberly Stachenfeld[5] STACHENFELD@GOOGLE.COM
Zeb Kurth-Nelson[4] ZEBK@DEEPMIND.COM
Timothy Lillicrap[4] COUNTZERO@DEEPMIND.COM
Claudia Clopath[6] CLAUDIA.CLOPATH@GMAIL.COM
Matthew Botvinick[7] BOTVINICK@GOOGLE.COM
Karel Svoboda[3] KAREL.SVOBODA@ALLENINSTITUTE.ORG
Kayvon Daie[2] KAYVON.DAIE@ALLENINSTITUTE.ORG
Matthew D. Golub[8,9] MGOLUB@CS.WASHINGTON.EDU

[1]Allen Institute + University of Washington
[2]Allen Institute for Neural Dynamics
[3]Allen Institute
[4]Google DeepMind
[5]Google DeepMind; Columbia University
[6]Imperial College London + Google DeepMind
[7]Deep Mind
[8]University of Washington

[9]Computer Science & Engineering

How does the motor cortex change as we learn new behaviors? To address this question, we studied learning in mouse motor cortex using an optical brain-computer interface (BCI) learning task that explicitly maps the activity of an experimenter-selected “conditioned neuron” to movements of a motorized reward port. Following each learning session, we used optical photostimulation to measure changes in the causal connectivity of each neuron onto all others in the imaged population. We found that learning was fast, sparse, and preferential. Fast learning occurs behaviorally within 10s of trials (~5 min). Sparse learning strengthens the trial-start-tuning of a small subpopulation of ’enhanced neurons’. Preferential learning manifests as enhanced neurons increasing causal connectivity onto other enhanced neurons. Next, to investigate whether these phenomena are consistent with changes to the upstream inputs to motor cortex versus local changes to motor cortex connectivity, we created two-region recurrent neural network (RNN) models and simulated both BCI-learning and optical photostimulation. Both upstream and local learning in task-trained networks learn to solve the task with high reward rate, but only local updates recapitulate preferential learning. Further, we found that biophysical constraints (E-I cell type and distance-dependent connectivity) in these networks were crucial to recapitulate the sparse learning observed experimentally. Taken together, this work integrates closed-loop BCI, targeted photostimulation, and biophysical network modeling to characterize population-level changes in motor cortex activity during learning and provides evidence for local connectivity changes in motor cortex during learning.

3-022. Representational drift as the consequence of ongoing memory storage

Alex Roxin[1] ALEXANDER.ROXIN@GMAIL.COM
Federico Devalle[2] FEDERICO.DEVALLE@GMAIL.COM
Licheng Zou[2] LICHENGZOU0509@GMAIL.COM
Gloria Cecchini[2] GCECCHINI@CRM.CAT

[1]Centre De Recerca Matematica
[2]Centre de Recerca Matematica

Memory systems with biologically constrained synapses have been the topic of intense theoretical study for over thirty years. Perhaps the most fundamental and far-reaching finding from this work is that the storage of new memories implies the partial erasure of already-stored ones. This overwriting leads to a decorrelation of sensory-driven activity patterns over time, even if the input patterns remain similar. Representational drift (RD) should therefore be an expected and inevitable consequence of ongoing memory storage. We tested this hypothesis by fitting a network model to data from long-term chronic calcium imaging experiments in mouse hippocampus. Synaptic turnover in the model inputs, consistent with the ongoing encoding of new activity patterns, accounted for the observed statistics of RD. This mechanism also provides a parsimonious explanation for the recent finding that RD in CA1 place cells has two distinct components: one which depends only on the passage of time, and another which depends on the time spent exploring a given environment. Furthermore, in the context of ongoing learning, the drift rate of any one memory depends on its repetition rate, a mechanism which can reproduce the diverse effects of experience on drift found in experiment. Our results suggest that RD should be observed wherever neuronal circuits are involved in a process of ongoing learning or memory storage.

3-023. A predictive coding model of cortical interneuron responses during change detection

Abdelrahman Sharafeldin ABDO.SHARAF@GATECH.EDU
Hannah Choi HANNAHCH@GATECH.EDU

Georgia Institute of Technology

Recent studies have shown that inhibitory interneurons, specifically vasoactive intestinal polypeptide (VIP)-expressing and Somatostatin (SST) neurons, exhibit context-dependent response dynamics shaped by experience and reflecting familiarity with the sensory environment. In a change detection task, Garrett et al. showed that both VIP and excitatory neurons in the mouse visual cortex respond more strongly to novel images than to familiar ones. Furthermore, VIP activity increases leading to familiar image presentations, followed by suppression upon onset—an effect absent for novel images. Similar effects were noted previously, but the underlying computational mechanisms remain unclear. We hypothesize that this interneuron circuit implements a predictive coding strategy that is modulated by uncertainty in the context of familiar and novel stimuli. To test this, we develop a dynamic circuit model based on Kalman filtering while accounting for biological connectivity and cell-type constraints. Our model consists of three populations corresponding to pyramidal cells, VIP cells, and SST cells in layer 2/3 of the mouse visual cortex. Using the same change detection task paradigm, we train the model with a predictive coding

objective on image sequences featuring stimuli used in experiments. When comparing population responses to familiar and novel images, we find that our model, despite being only optimized for perception, reproduces some experimental observations. For example, VIP neurons fire more at change image onsets relative to pre-change images while responding more strongly to novel images relative to familiar images. A similar effect is observed for excitatory neural responses to pre-change images, with prediction error neurons firing more during novel sequences. Taken together, these preliminary results suggest that the VIP-SST-Excitatory circuit may employ a predictive coding strategy akin to Kalman filtering during change detection.

3-024. A theory of thalamocortical loops and decision-making

Michael Berry[1,2] BERRY@PRINCETON.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute

We have developed a novel theory of the function of thalamocortical loops based on known anatomy and biophysics. We focus on higher-order thalamic nuclei that receive driver-type inputs from layer 5 of the neocortex and matrix-type thalamic relay cells that synapse primarily in layers 1 and 5 of neocortex, forming a reciprocal excitatory loop. As layer 5 pyramidal tract neurons send motor commands to the brainstem and spinal cord, we assume that ensembles of L5 neurons encode distinct movements. Descending cortical inputs into layer 1 set the gain of these L5 neurons. At a critical gain, excitatory feedback from the thalamus exactly cancels the leak in L5 neurons resulting in an ideal neural integrator. We explore the dynamics of single thalamocortical loops, showing an underamplified regime with steady-state activity, an overamplified regime with exponentially rising activity, and a regime with good integrator properties.

Thalamocortical loops compete with each other via lateral inhibition in the thalamus. When competing loops are sufficiently amplified, the more active loop suppresses the less active, thereby ramping up to the threshold for activating the striatum and releasing the movement encoded by that loop. Together, multiple competing loops constitute a global system of movement selection and/or decision-making.

When sensory inputs drive L5 neurons and its corresponding loop is in the integrator regime, the activity of the loop represents the integration of sensory evidence in favor of making the encoded movement. We show that over a wide range of parameters, competing thalamocortical loops closely resemble the drift-diffusion model of decision-making. Finally, we explore the effects of thalamic feedback to layer 1 that significantly alters the gain of L5 neurons. We demonstrate a regime where the time for competing loops to reach a decision becomes very long.

3-025. Hyperbolic geometry of spatial representation in medial entorhinal cortex neurons

Ruixin Qian[1,2] RUIXIN.QIAN.0.0@GMAIL.COM
Tatyana Sharpee[3] TANYA.SHARPEE@GMAIL.COM

[1]Nanjing University
[2]Physics
[3]Salk Institute for Biological Studies

Neurons located in the medial entorhinal cortex (MEC) exhibit complex firing patterns, including mixed selectivity and irregular tuning curves for spatial variables. Past research has often focused on only on certain neuron types with regular firing patterns, such as grid cells and head direction (HD) cells. However, those criteria often overlook many other MEC neurons. In contrast, our study examines all MEC neurons as population, revealing the hyperbolic geometry of spatial representation and the intrinsic hierarchical structure within the neuron population. Neurons could be classified as “root” or “branch” depending on their assigned radius was small or large, respectively, within hyperbolic embedding. The root neurons were the most active and displayed more mixed selectivity than branch neurons. Furthermore, root and branch neurons together formed a core-periphery network, where root neurons are the core within the firing correlations matrix. Notably, this network exhibited small-world properties, promoting efficient inter-neuron communication. Our research approach, which views neurons from a population perspective, does not employ rigid patterns to exclude neurons. As a result, we can fully account for the complexity of MEC neuron firing, fostering a more comprehensive understanding of MEC neuron coding mechanisms.

3-026. Rhythmically Structured Predictive Coding Enables Invariant Semantic Recovery

Olesia Dogonasheva[1,2] ODOGONASHEVA@GMAIL.COM
Olesia Platonova[3] OLESIA.PLATONOVA@PASTEUR.FR
Denis Zakharov[4] DGZAKHAROV@HSE.RU
Anne-Lise Giraud[5] ANNE-LISE.GIRAUD-MAMESSIER@PASTEUR.FR
Boris Gutkin[6] BORIS.GUTKIN@ENS.FR

[1]Ecole normale superieure
[2]Laboratoire de Neurosciences Cognitives et Computationnelles
[3]Institut Pasteur
[4]HSE University
[5]Institut Pasteur, University of Geneva
[6]ecole Normale Superieure PSL*

Our ability to perceive speech is resilient to variations in voice, rate, and temporal interruptions. While inference models addressed the first two [1], computational principles by which speech understanding remains impervious to temporal restructuring are largely unexplored. Previous studies indicated intriguing recoveries in comprehension when speech interruptions were over 1 Hz [2] and temporally compressed speech was segmented with silences of specific durations [3, 4]. Here, we show that predictive coding, constrained by endogenous rhythms, accounts for these quizzical results and enables robust speech recovery.

We build upon the major hypotheses that the rhythmic structure of speech establishes temporal windows, allowing the brain circuits to effectively process auditory signals. Moreover, rhythmic activity is hierarchically structured in line with the structure of speech [5] and modulates predictive coding so that successful comprehension relies on actively minimizing contextual uncertainty and surprise [6]. These in turn modulate theta and delta rhythms, respectively [7].

Integrating this evidence, we propose a predictive coding framework (BRyBI), which implements a hierarchy of rhythms and actively minimizes both uncertainty and surprise. The theta rhythm in the BRyBI reduces uncertainty in the subsequent phoneme distribution. Theta rhythm entrainment by speech minimizes errors in the gamma code of phonemes. On the other hand, the delta rhythm enables temporally-structured semantic prediction error minimization, thereby implementing online word-context inference.

BRyBI allows for robust speech recognition under temporal perturbations such as compression, interruption, and segmentation. Furthermore, behaviors observed experimentally that so far have escaped explanation, such as error-related potentials that emerge naturally in BRyBI; speech-rhythms coherence decreases for theta and grows for delta with increased uncertainty and surprise. In sum, we suggest that oscillation-constrained predictive coding generically explains the results of multiple experiments with temporal scale alterations and provides a new view of the speech recognition process in the brain.

3-027. Choice-wide behavioral association study: reliable and interpretable differences across learning

David Kastner[1,2] DBKASTNER@GMAIL.COM
Cristofer Holobetz[1] CRISTOFER.HOLOBETZ@BERKELEY.EDU
Nicole Yokota[1] NYOKOTA@BERKELEY.EDU
Greer Williams[1] GREERWILLIAMS28@GMAIL.COM
Christina Lee[1] DUN05137@BERKELEY.EDU
Jane Ton[1] NHUNGTON18@G.UCLA.EDU
Joseph Romano[3] ROMANO@STANFORD.EDU
Peter Dayan[4] DAYAN@TUE.MPG.DE

[1]University of California, San Francisco
[2]Department of Psychiatry and Behavioral Sciences
[3]Stanford University
[4]Max Planck Institute for Biological Cybernetics

Animal behavior contains rich structure across many timescales, but there is a dearth of methods for characterizing the acquisition and improvement of task competence over the long run. Inspired by the goals and techniques of genome-wide association studies, we report a data-driven method–the choice-wide behavioral association study: CBAS–that systematically identifies differences between the choices of groups of subjects. CBAS compares the frequencies of many sequences of choices between two groups, then uses powerful, resampling-based, multiple comparisons methods to identify the sequences that differ significantly between the groups. We illustrate CBAS by comparing the behavior of wild-type (WT) to genetically modified rats and two structurally different reinforcement-

learning (RL)-based computational agents as they learn a series of contingencies in a spatial alternation task. Each contingency requires subjects to visit arms of a track in a repeating pattern. When we apply CBAS to all sequences up to six arm visits long made by two RL agents that learn with alternative strategies, it identifies distinct sequences that can be related to the different algorithms used by the agents. When we use CBAS to compare WT rats to those haploinsufficient for a high-confidence autism spectrum disorder risk gene (Scn2a), it identifies specific and consistent ways that Scn2a haploinsufficient rats differ throughout all phases of learning. Through identifying the choices that differ between groups of subjects, CBAS provides a uniquely informative framework to interpret neural function and its changes in the context of disease processes.

3-028. Identifying representational structure in CA1 to benchmark theoretical models of cognitive mapping

J. Quinn Lee[1] JQUINNLEE@GMAIL.COM
Alexandra Keinath[2] ATK@UIC.EDU
Erica Cianfarano[1] ERICA.CIANFARANO@MAIL.MCGILL.CA
Mark Brandon[1] MARK.BRANDON@MCGILL.CA

[1]McGill University
[2]University of Illinois Chicago

The hippocampus and associated regions of the neocortex are thought to support diverse learning and memory processes through cognitive maps instantiated in the activity of principal neurons. Growing evidence suggests that such maps are formed across metric spaces to support flexible, goal-directed navigation and long-term memory; functions that are impaired following hippocampal damage. These findings have recently motivated an increasing number of theories expressed in computational models to explain how the hippocampal system instantiates cognitive maps. However, there has been no consensus on how to compare predictions across models and, importantly, against empirical observation. To address this need, we recorded from large populations in hippocampal subregion CA1 in a condition-rich geometric deformation paradigm (5,413 unique neurons recorded across 207 sessions in 10 geometries, forming 69,744 rate maps). Leveraging a similarity-based framework, we observed that a constrained range of allocentric, BVC-based models accurately predict CA1 spatial representation (within maximum theoretical limits). The success of such models to predict CA1 representation suggest that an allocentric vector-based code strongly determines population-level neuronal dynamics in CA1. These are the first results to our knowledge demonstrating that large-scale predictions from neurobiological models of cognitive mapping can be directly evaluated against population dynamics in freely behaving animals. The present dataset and framework provides a benchmark (the first) for theoretical innovations in cognitive mapping research and, more generally, establishes a novel approach to compare representational structure across brain regions, assays, species, and theoretical models.

3-029. A new family of statistical tests for responses in point-event and time-series data for one- and two-sample comparisons

Jorrit Montijn[1] JSMONTIJN@GMAIL.COM
Guido Meijer[2] GUIDO.MEIJER@DONDERS.RU.NL
Alexander Heimel[1] A.HEIMEL@NIN.KNAW.NL

[1]Netherlands Institute for Neuroscience
[2]Donders Centre for Neuroscience

Quantifying whether and when signals are modulated by autonomous or external events is ubiquitous in the field of neuroscience. Existing statistical approaches, however, are not ideally suited to do this, especially when the signals under scrutiny show temporal autocorrelations. For example, a standard approach in the analysis of calcium imaging data is to use a t-test on predetermined time-windows to quantify whether neurons respond (differently) to an event of interest. While this is attractive because of its simplicity, only average signal differences can be detected. In practice, neurons often show complex response dynamics which are missed by conventional statistical tests. Bin-wise ANOVAs can detect more complex patterns, but these require an arbitrary choice of bin-size and fail when the underlying data show temporal autocorrelations. To solve this issue, we present three extensions of the recently developed ZETA-test (Montijn et al., 2021): 1) a modification for analysing time-series data; 2) a two-sample variant to detect a difference in neural responses between two conditions; and 3) a two-sample variant for time-series data.

In all data we investigated, the ZETA-tests showed improved statistical sensitivity compared to established and powerful statistical techniques. The performance of all ZETA-tests was without exception equal or superior to those of t-tests and ANOVAs. Moreover, the family of ZETA-tests is easier to use than established methods, as

it can be applied directly to raw spike times or time-series data and stimulus onsets, and the lack of parameter selection naturally lends itself to the bulk-analysis of large numbers of cells. We expect that our procedures may be of interest to statisticians and theoreticians, but we have created programmatic implementations (available on GitHub) specifically with experimentalists in mind, who can use our methods to obtain higher cell yields and more reliable results from their neurophysiological data.

3-030. Optimal modulation of sensory neurons during locomotion

Jonathan Gant[1,2] GANT@BIO.LMU.DE
Wiktor Mlynarski[1] MLYNARSKI@BIO.LMU.DE

[1]LMU Munich
[2]Faculty of Neurobiology

With the advent of neural recordings in behaving animals, it became clear that the behavioral state of an organism strongly modulates sensory neurons. A prominent example of such modulations is the locomotion-induced gain change of visual neurons observed in rodents and insects. Although the influence of behavior on sensory coding is well established, its precise computational purpose remains unknown.

Here, we aim to provide a candidate, normative account of locomotion-induced modulation of sensory systems. Our starting point is the efficient coding hypothesis, which postulates that neurons adapt to stimulus statistics in order to efficiently transmit sensory information. Such adaptation can occur rapidly in response to a change in the stimulus distribution due to external factors that are independent of the organism. However, locomotion–a self-generated behavior–also exerts a significant impact on stimulus statistics. We hypothesize that internal modulation induced by locomotion adapts sensory neurons to efficiently encode stimuli whose statistics change with movement.

To understand possible principles of such modulation, we simulated an agent moving in spatially varying environments. At each position along its trajectory, the agent observed a stimulus drawn from a position-specific distribution, corresponding to e.g. spatial variations of light intensity. The resulting temporal sequences of stimuli were strongly affected by the agent's speed. To determine how locomotion should modulate sensory coding, we optimized a model neuron to efficiently encode stimuli experienced during movement at different speeds. We found that, in agreement with experimental observations, speed systematically modulated the gain and slope of optimal neural tuning curves to match movement-induced changes in the dynamic range of stimuli. Analysis of temporal stimulus statistics generated novel predictions about the dynamics of sensory adaptation and modulation of temporal receptive fields during locomotion, suggesting new directions for experimental exploration.

3-031. Moving speed dilates the toroidal structure of population activity in grid cells

Zeyuan Ye[1,2] ZEYUANYE@HOTMAIL.COM
Ralf Wessel[1] RW@PHYSICS.WUSTL.EDU

[1]Washington University in Saint Louis
[2]Department of Physics

Calculating one's location based on previous motion is an important ability in navigation. This ability is especially important when the agent is moving at high speed where the location is constantly changing. It has been shown that the animal's moving speed modulates individual grid cell activities (a key neuron for navigation which has a characteristic hexagonal firing pattern) in various ways [1]. However, an understanding from a population level is lacking. The implication of speed modulation is also unclear.

In this work, we used a geometric framework to interpret the speed modulation of grid cell population. The geometric framework assumes that high-dimensional neural population activities (dimensionality equals to the number of neurons) are constrained by a low-dimensional smooth subspace called manifold. We introduced the Gaussian Process Regression (GPR) to infer the smooth manifold from noisy grid cell populational activities (data sourced from Richard et al. [2]). We found that from the populational level, speed modulation is straightforward: speed dilates the toroidal structure underlying grid cell population activity. This dilation coexists with better location decoding accuracy (decoded from grid cell activity). In all, we introduced GPR as a tool to infer smooth manifold from scattering data, and using it discovered (i) speed dilation of grid cell population activity manifold (ii) better location decoding accuracy with manifold dilation. Our results suggest that the neural system employs neural population representation strategies to help navigation when the agent is moving in high speed.

3-032. Binge feeding promotes appetite via modulating olfactory flavor representation.

Hung Lo[1,2] HUNG.LO.ECN@GMAIL.COM
Malinda L.S. Tantirigama[3] MALINDA.TANTIRIGAMA@NTNU.NO
Anke Schoenherr[1] ANKE.SCHOENHERR@CHARITE.DE
Laura Moreno-Velasquez[1] LAURA.MORENO-VELASQUEZ@CHARITE.DE
Lukas Faiss[1] LUKAS.FAISS@CHARITE.DE
Benjamin R. Rost[1] BENJAMIN.ROST@CHARITE.DE
Matthew E. Larkum[3] MATTHEW.LARKUM@HU-BERLIN.DE
Benjamin Judkewitz[1] BENJAMIN.JUDKEWITZ@CHARITE.DE
Katharina Stumpenhorst[3] KATHARINA.STUMPENHORST@HU-BERLIN.DE
Marion Rivalan[1] MARION.RIVALAN@CNRS.FR
York Winter[3] YORK.WINTER@HU-BERLIN.DE
Dietmar Schmitz[1] DIETMAR.SCHMITZ@CHARITE.DE
Friedrich W. Johenning[1] FRIEDRICH.JOHENNING@CHARITE.DE

[1]Charite – Universitatsmedizin Berlin
[2]Neuroscience Research Center
[3]Humboldt Universitat zu Berlin

Binge eating commonly leads to overeating, but the exact mechanism is unclear. While it is known that experiencing flavor contributes to satiety, the interactions between flavor, feeding rate, and food intake remain unknown. Here, we demonstrate a novel feeding rate-dependent modulation of food intake by olfactory flavor representation in the anterior olfactory (piriform) cortex (aPC). We developed a liquid food delivery system that enables food consumption at different feeding rates. Using miniscopes for in vivo calcium imaging in freely foraging mice, we identified specific excitatory neuronal responses to food and water during slow feeding. Switching to binge feeding transformed these specific responses into unspecific global suppression of neuronal activity. In the gustatory cortex (GC) and the olfactory bulb, we observed similarities in flavor representation during binge and slow feeding. Food consumption was predicted by the degree of suppression of neuronal activity in the aPC during binge feeding, and food deprivation enhanced neuronal activity suppression. We confirmed the hypothesis that aPC suppression promotes food intake with closed-loop optogenetics experiments. Together, our results show that olfactory sensory representation in the aPC reciprocally interacts with consummatory behavior to enhance food intake, which may suggest an interesting neuronal system for future computational modeling.

3-033. Barcoding of episodic memories in the hippocampus of a food-caching bird

Selmaan Chettih[1,2] SC4551@COLUMBIA.EDU
Emily Mackevicius[1] EM3406@COLUMBIA.EDU
Stephanie Hale[1] SMH2279@COLUMBIA.EDU
Dmitriy Aronov[1] DA2006@COLUMBIA.EDU

[1]Columbia University
[2]Zuckerman Institute

The hippocampus is critical for episodic memory. Although hippocampal neurons represent place and other variables useful for memory, it is unclear how their activity encodes memories of specific, rapidly occurring episodes. To study episodic coding, we leveraged the specialized food-caching behavior of chickadees – birds that use hippocampal-dependent memory to retrieve previously cached food. They thus experience well-defined moments of memory storage each time they cache a food item for later retrieval. We developed methods for 3D postural reconstruction and automated behavioral annotation of this natural behavior, along with silicon probe electrophysiology, enabling the first characterization of hippocampal activity during food caching. Our recordings in the chickadee hippocampus revealed sparse, “barcode-like” patterns of firing across cells during caching, which contrasted with conventional place coding during navigation. Barcodes were orthogonal to the place code and to other barcodes, even for caches at nearby locations with similar place cell activity. Barcodes were even distinct for different caching events at the same location. Most importantly, the barcode for a specific cache was transiently reactivated during subsequent retrieval of that cache. Thus, barcode activity encoded a single memorable episode, which could be reactivated later, during memory recall. We also found location-independent hippocampal coding for the presence of cached seeds, and observed a precise temporal sequence of place-, barcode, and seed-coding. We propose that barcodes are signatures of episodic memories, which become associated with place and seed codes during memory formation. These patterns assign a unique identifier to each event and may be a mechanism for rapid formation and storage of many non-interfering memories.

3-034. Anesthesia fragments cortical activity within a hemisphere, but synchronizes it across hemispheres

Alexandra Bardon[1,2] ABARDON@MIT.EDU
Jesus Ballesteros[3] JESUS.BALLESTEROSCARRASCO@RUB.DE
Scott Brincat[1] SBRINCAT@MIT.EDU
Emery Brown[1] ENB@NEUROSTAT.MIT.EDU
Earl Miller[1] EKMILLER@MIT.EDU

[1]Massachusetts Institute of Technology
[2]Brain and Cognitive Sciences
[3]Ruhr-University Bochum

Many different anesthetics cause unconsciousness, despite having diverse underlying molecular mechanisms. To explore the convergent effects of these drugs, we examined the effects of ketamine, an N-methyl-D-aspartate (NMDA) receptor antagonist, and dexmedetomidine, an α2 adrenergic receptor agonist, on neural oscillations in the prefrontal cortex of nonhuman primates. Previous work has shown that anesthesia may disrupt cortical communication by misaligning 'on' and 'off' spiking states. On the other hand, anesthesia causes an increase in slow oscillation coherence across cortex, which would seem to enhance cortical communication, albeit abnormally. However, coherence does not necessarily improve communication. It also depends on aligning oscillatory phases, so neurons are in states of excitation simultaneously. We found that both ketamine and dexmedetomidine increased slow wave power in the ventrolateral and dorsolateral prefrontal cortex, as well as phase locking between these regions. Yet, while these regions all became more phase locked, there was a difference in their phase offsets. We observed that different subregions within a hemisphere were locked anti-phase, while homologous prefrontal subregions across hemispheres were more often locked in-phase. Within each subregion, phase offsets increased with distance between recording sites; the cross-hemisphere in-phase locking was thus unexpected based solely on distance effects. At sub-anesthetic doses of each drug, we found qualitatively similar, but weaker, changes in phase offsets, with no significant increase in phase locking. We show that anesthesia acts through complex changes in phase alignment of oscillatory dynamics that are common across drugs despite their different molecular actions. These dynamics suggest that appropriate levels of interregional communication may be necessary for the computations involved in conscious cognition.

3-035. Cultivation of cosine-tuning in both artificial spiking and cortical neural networks during training

Tengjun Liu[1] TENGJUNLIU@ZJU.EDU.CN
Yansong Chua[2] CAIYANSONG@CNAEIT.COM
Yiwei Zhang[3] ZHANGEVEY@ZJU.EDU.CN
Yuxiao Ning[3] NINGYUXIAO@ZJU.EDU.CN
Pengfu Liu[3] PL1690@NYU.EDU
Zijun Wan[3] 11715005@ZJU.EDU.CN
Shaomin Zhang[3] SHAOMIN@ZJU.EDU.CN
Weidong Chen[3] CHENWD@ZJU.EDU.CN

[1]Zhejiang University, Friedrich Miescher Institute for Biomedical Research (FMI)
[2]China Nanhu Academy of Electronics and Information Technology (CNAEIT)
[3]Qiushi Academy for Advanced Studies, Zhejiang University

Goal-driven rate-based artificial neural networks (ANN) have exhibited numerous promising bio-functional similarities with the biological neural system, rendering them suitable models of neural computation for neuroscientific predictions, especially for the sensory system. However, they are far less reported to model the motor system. Furthermore, the extent to which their goal-driven spiking counterparts demonstrate such bio-functional similarities in the motor cortex remains even more unclear, let alone their predictive capability. In this study, we introduced the motorSRNN, a recurrent spiking neural network (SNN) inspired by the primate neural motor circuit. It decoded cortical spike trains (CST) from the primary motor cortex (M1) in two monkeys conducting 4-direction center-out tasks, with an average performance of 89.4% and 79.9% classification accuracy respectively. Validating its bio-functional similarity with the motor cortex, we noticed that the distributions of preferred directions (PD) of significantly cosine-tuned neurons (SCtN) in motorSRNN closely resemble those observed in neuron recordings from biological motor cortex. Moreover, motorSRNN communicated with the input by capturing and cultivating more SCtNs and maintained the stable existence of cultivation during training. We regard this finding as a prediction that should also be observed in a learning biological neural network of the motor cortex. To examine this hypothesis, we designed mind control experiments in which two monkeys were trained to modulate the cortical neural network activities to control a new decoder in 4 sessions. An interval of ~1 month between consecutive sessions was considered to ensure a large compositional difference of recorded neuronal groups for better gen-

eralizability. Our results confirmed that new task training can indeed induce the stable existence of cultivation of SCtNs in the primate motor cortex. In short, the goal-driven motorSRNN demonstrated bio-functional similarity and neuroscientific predictive capacity, offering a potential framework for building neural computation models of the motor circuit.

3-036. Hippocampal representations in a complex route planning task

Beatriz Godinho[1,2] BEATRIZ.GODINHO@NDCN.OX.AC.UK
Chongyu (Xiao) Qin[3] CHONGYU.QIN.20@UCL.AC.UK
Francesca Pozzolo[3] FRANCESCA.POZZOLO.23@UCL.AC.UK
Peter Doohan[1,4] PETER.DOOHAN@CCC.OX.AC.UK
Mark E. Walton[1] MARK.WALTON@PSY.OX.AC.UK
Tim Behrens[5] TIMOTHY.BEHRENS@NDCN.OX.AC.UK
Thomas Akam[1] THOMAS.AKAM@PSY.OX.AC.UK

[1]University of Oxford
[2]Nuffield Department of Clinical Neurosciences
[3]University College London
[4]Nuffield Department of Clinical Neuroscience
[5]University of Oxford & University College London

Behavioural flexibility is a hallmark of human and animal intelligence. It relies on internal models of the world - cognitive maps - which allow the behavioural consequences of new information to be inferred. Rodent spatial navigation is a powerful model system for studying the neural basis of cognitive maps, due to its ethological validity and our precise knowledge of how neurons represent space. However, most recordings come from simple behaviours and environments. A major open question is how spatial representations support model-based behaviours such as planning, in complex and dynamic environments. We recorded CA1 neurons as mice performed a complex route planning task, navigating to visually cued goal locations in an elevated maze (Fig.A), with different start and goal locations on each trial. We added or removed one link from the maze each day such that its structure gradually evolved over time (30 structures, 506 sessions, 21,162 trials, 8,869 cells) (Fig.D). Choices' analysis revealed mice use knowledge of the maze structure - often taking optimal paths even when they required heading away from the goal, and showed rapid learning of the structural change within the first session (Fig.C). Further behavioural analysis (in another anonymous submission to Cosyne24) provides preliminary evidence that mice plan compositionally with a set of commonly used routes. Cellular data analysis reveals that CA1 cells on the maze are not typical place cells. Few cells have a single localised firing field. Most firing fields are directionally selective, and many are spatially extended - entirely (rarely) (Fig.F-cell 2) or partially (commonly) (Fig.F-cell 1,3) - as if representing behavioural trajectories (routes). In preliminary analysis, dimensionality reduction of the neural space reveals neural factors mirroring the compositional elements of behaviour (Fig.G). Ongoing analysis will examine how these representations change with the maze changes, and whether sleep replay supports these updates.

3-037. Synaptic wiring motifs in posterior parietal cortex support decision-making

Aaron Kuan[1,2] AARON.KUAN@YALE.EDU
Giulio Bondanelli[3] GIULIO.BONDANELLI@IIT.IT
Laura Driscoll[4] LNDRISCOLL@GMAIL.COM
Julie Han[5] JULIEHAN5296@GMAIL.COM
Minsu Kim[5] MINSU_KIM@HMS.HARVARD.EDU
David Hildebrand[6] DAVID@DAVID.HILDEBRAND.NAME
Brett Graham[5] BRETTGRAHAM@GMAIL.COM
Daniel Wilson[5] DANIEL_WILSON@HMS.HARVARD.EDU
Logan Thomas[7] THOMAS.LO@HUSKY.NEU.EDU
Stefano Panzeri[8] STEFANO.PANZERI@IIT.IT
Christopher Harvey[5] CHRISTOPHER_HARVEY@HMS.HARVARD.EDU
Wei-Chung Lee[9] WEI-CHUNG_LEE@HMS.HARVARD.EDU

[1]Yale School of Medicine
[2]Neuroscience
[3]Istituto Italiano di Tecnologia
[4]Allen Institute for Neural Dynamics
[5]Harvard Medical School
[6]The Rockefeller University

[7]University of California Berkeley
[8]Istituto Italiano di Tecnologia, University Medical Center Hamburg-Eppendorf
[9]Boston Children's Hospital, Harvard Medical School

The posterior parietal cortex (PPC) exhibits choice-selective activity during perceptual decision-making tasks. However, it is not known how this selective activity arises from the underlying synaptic connectivity. Here, we combined virtual reality behavior, two-photon calcium imaging, high throughput electron microscopy, and circuit modeling to analyze how synaptic connectivity between neurons in PPC relates to their selective activity. We found that excitatory pyramidal neurons preferentially target inhibitory interneurons with the same selectivity. In turn, inhibitory interneurons preferentially target pyramidal neurons with opposite selectivity, forming an opponent inhibition motif. Using circuit models, we show that opponent inhibition amplifies selective inputs and induces competition between neural populations with opposite selectivity, thereby improving the encoding of trial-type information. These results provide evidence for how synaptic connectivity in cortical circuits supports a learned decision-making task.

3-038. Mean field theory of representation learning in large RNNs

Blake Bordelon[1,2] BLAKE_BORDELON@G.HARVARD.EDU
Jacob Zavatone-Veth[1,3] JZAVATONEVETH@G.HARVARD.EDU
Cengiz Pehlevan[1] CPEHLEVAN@SEAS.HARVARD.EDU

[1]Harvard University
[2]Applied Mathematics
[3]Physics

Flexible learning of recurrent dynamics is believed to underlie rapid acquisition of skilled behaviors. Yet, though pattern generation using trained readout from randomly-connected recurrent reservoirs has been the subject of exhaustive theoretical study, an analytical understanding of the effect of training the recurrent weights themselves is lacking. Here, we study a tractable continuous sequence learning setting in which the dynamics of a large trained RNN can be characterized using dynamical mean field theory (DMFT). Unlike previous analytical studies of learning in RNNs, we consider end-to-end task optimization of all parameters to produce desired output. In the case of a linear network, the model is fully solvable, but for nonlinear networks the DMFT is not analytically tractable and must be solved numerically. Focusing on the linear case, we characterize how feature learning shapes the network transfer function: the model learns to enhance frequencies required to generate the target signal. We show through numerical experiments that nonlinear networks display some qualitatively similar behavior, while also being able to learn to nonlinearly transfer frequencies. In total, our work provides the first step towards an analytical theory of how learning to generate temporal patterns shapes the internal dynamics of RNNs.

3-039. Analytic model of response statistics in noisy neural populations with divisive normalization

Daniel Herrera-Esposito[1,2] DHERRERA1911@GMAIL.COM
Johannes Burge[1] JBURGE@PSYCH.UPENN.EDU

[1]University of Pennsylvania
[2]Department of Psychology

Divisive normalization is an essential neural computation. However, normalization has mostly been studied without accounting for neural noise, which is crucial for neural coding. Here, we derive analytic formulas describing how normalization shapes, and is shaped by, noise in a model neural population. Like classical normalization models, there is an initial unnormalized population drive (e.g. the outputs of linear receptive fields). Then, the population drive is divided by a normalization signal that is a function of the pooled drives (e.g. L2-norm), resulting in the normalized population response. The key difference with classic models is that the current model incorporates noise in the population drive, and hence the normalization signal. We model the unnormalized drive as multivariate Gaussian, allowing neural noise to affect both the numerator (drive) and the denominator (normalization signal) of the normalization formula. We derive analytic formulas for the mean and covariance of the normalized responses, given the statistics of the unnormalized drive. The formulas are flexible and can incorporate arbitrary noise correlations and different types of normalization (e.g. broadband, feature-specific). They are differentiable, making them suitable for optimization routines. Like classic normalization models, responses of the current model saturate at high contrasts. Several non-obvious behaviors are predicted. First, independent noise in the drive is transformed by normalization into noise correlations in the normalized responses. Second, normalization-induced correlations are stimulus-dependent. Third, increasing noise in fixed-mean unnormalized

drives reduces mean normalized responses. In sum, we derive analytic formulas characterizing a model that incorporates neural noise into divisive normalization. Our model includes elements (e.g. large neural populations, principled dependence of normalization on population drive) missing from other recent attempts to analytically model how noise and normalization interact. The model displays behaviors characteristic of real neural systems, some of which have not been linked analytically to normalization.

3-040. Environmental dynamics affect whether matching is optimal for foraging

Yipei Guo[1] GUOY2@JANELIA.HHMI.ORG
Ann Hermundstad[2] HERMUNDSTADA@JANELIA.HHMI.ORG

[1]HHMI Janelia Research Campus
[2]HHMI Janelia

When foraging for resources across different environments, animals must often sample many options that yield reward with different probabilities. In such scenarios, many animals have been shown to exhibit “matching”, an empirical behavioral observation in which the fraction of successful samples is the same across all options [1]. While previous work has shown that matching can be optimal in environments with diminishing returns [2], this condition is not sufficient to determine optimality. Furthermore, while diminishing returns naturally arise when resources in the environment deplete and take time to be replenished, the specific form of diminishing returns depends on the temporal structure and statistics of the replenishment process. Here, we explore how these environmental properties affect whether matching is optimal. We consider a class of policies in which an agent samples different options with fixed sampling rates. We derive the observed reward probability as a function of these sampling rates for different types of replenishment structures, and analytically determine the conditions under which the optimal sampling-rate policy exhibits matching. When all options are governed by the same replenishment dynamics, we find that optimality gives rise to matching across a wide range of environments. However, when these dynamics differ across options, the optimal policy can deviate from matching. In such cases, the rank-ordering of the observed reward probabilities depends only on the qualitative nature of the replenishment process, but not on the specific replenishment rates. As a result, the optimal policy can exhibit under- or over-matching depending on the relative replenishment rates across options. We use this result to identify conditions under which the difference between matching and optimality is large. Together, these findings provide testable experimental predictions across a wide range of environmental setups.

3-041. The Role of Inhibition in Shaping Dendritic Synaptic Arrangement

Nikos Malakasis[1] MALAKASISNIKOS@GMAIL.COM
Julijana Gjorgjieva[2,3] GJORGJIEVA@TUM.DE

[1]School of Life Sciences, TUM
[2]Technical University of Munich
[3]School of Life Sciences

Dendrites receive a vast array of inputs, integrating and processing information before it reaches the cell body. Active dendritic mechanisms amplify synaptic inputs at dendritic locations far from the cell body, supported by synaptic clusters. These clusters consist of neighboring synapses on the dendritic arbor, which encode information about a shared stimulus feature or a specific memory. Hence, how these synapses are spatially organized profoundly influences neuronal functional properties and dynamics. Despite their influence on dendritic processing, the mechanisms governing the spatial clustering of synapses on dendritic branches remain unclear. Extensive evidence points to inhibition playing a substantial role in influencing synaptic organization, by modulating Ca2+ currents on dendritic branches (Hayama et al., 2013). Here, we aimed to unravel how inhibition and inhibitory plasticity sculpt the arrangement of synapses on dendritic branches. We constructed a dendritic branch model that integrates inputs through excitatory and inhibitory synapses, each subject to functional and structural plasticity. Excitatory synapses undergo activity-dependent plasticity, stabilizing coactive synapses while depressing inactive ones in a distance-dependent fashion. Additionally, we modeled the formation of inhibitory synapses at dendritic sites exhibiting increased activity, inspired by an experimentally observed retrograde signaling mechanism (Hu et al., 2019). After their formation, inhibitory synapses also undergo activity-dependent plasticity dependent on local excitatory activity, leading to the emergence of co-tuned clusters comprising both inhibitory and excitatory synapses selective to the same stimulus feature. We found that inhibition constrains excitatory synaptic clusters within distinct spatial domains. This dynamic interaction of excitation and inhibition at the subcellular level substantially improves the dendritic branch’s capacity to store multiple stimuli. These findings suggest that inhibition and inhibitory plasticity can powerfully shape the arrangement of dendritic synapses, and influence neuronal

computations, further adding to their already known roles in regulating dynamics and cotuning at the network level.

3-042. Hippocampus is necessary for implicit statistical learning: insights from mouse and human pupillometry

Adedamola Onih[1,2] A.ONIH@UCL.AC.UK
Athena Akrami[3] ATHENA.AKRAMI@UCL.AC.UK

[1]University College London
[2]Sainsbury Wellcome Centre
[3]Sainsbury Wellcome Centre, UCL

Understanding the mechanisms of statistical learning, where organisms infer structure from sensory observations without explicit instruction, is crucial in comprehending how the brain processes and responds to complex environments. This work explores the innate capabilities of both humans and mice to discern environmental regularities, focusing on the role of the hippocampus in this implicit learning processes. We use pupillometry to measure whether animals can detect and track patterns in rapidly unfolding sounds. The study leverages a novel comparative behavioural paradigm, termed X-detection, to manipulate the statistics of embedded tone sequences (patterns) in an auditory cover task: subjects should respond to a target sound, X, that appears at a random time. We show pupil responses in mice and humans track statistics of presented patterns, with limited exposure. Importantly, these embedded patterns are decoupled from the X; therefore, from the reward. We show that the performance on X-detection is not affected by the patterns. This paradigm allows probing of implicit learning of sound statistics without reinforcing the learning via reward. Our inactivation results, from bilateral infusion of dCA1 with GABA-A agonist muscimol, shows hippocampal inactivation, despite sparing the performance on X-detection, diminishes the pupil response tracking the statistics. To our knowledge, these results provide the first set of evidence, in rodents, for causal role of hippocampus in fast statistical learning of sound statistics.

3-043. Visual cortex reformats task-specific information to facilitate sensation, cognition, and action

Ramanujan Srinath[1] RAMSRINATH@UCHICAGO.EDU
Martyna Czarnik[2] MC5494@PRINCETON.EDU
Marlene Cohen[1] MARLENECOHEN@UCHICAGO.EDU

[1]University of Chicago
[2]Princeton University

Visually guided tasks have three hallmark components: perception (sensing the external world), cognition (making an inference or decision in service of a goal while ignoring irrelevant information), and action (planning a behavior based on that inference). Traditionally, these computations have been assumed to be performed by separate neural populations that communicate flexibly to achieve cognitive flexibility. But evidence that interactions between brain areas are flexible enough to enable the wide range of visually guided behaviors performed by humans and animals has been limited. Here, we consider a non-modular possibility: signals related to perception, cognition, and action are integrated in a single neural population in visual cortex, and flexibility is instantiated via modest modulations of neural responses that rotate population activity to drive appropriate inferences and actions. To test this hypothesis, we trained monkeys to make veridical (continuous) estimates of the 3D curvature of objects that varied in task-relevant and irrelevant parameters while we recorded simultaneously from populations of neurons in primary visual cortex (V1) and mid-level visual area (V4). We discovered that populations of V1 and V4 neurons have two ingredients necessary to instantiate cognitive flexibility: they 1) robustly encode both task-relevant and irrelevant visual information and 2) contain a stable representation of the subject's curvature inference amid irrelevant stimulus variation. The same V4 population has a third ingredient: it encodes which of several possible actions the subject will use to report their inference by rotating or reformatting the representation of the stimulus to align with a putative readout axis. This reformatting can be accomplished via modest gain changes, such as those associated with surround modulation, attention, or task switching. Our results suggest that visual cortex contains all the ingredients to mediate cognitively flexible visually guided behavior. These results support a new, non-modular mechanism for cognitive flexibility.

3-044. Ketamine Enhances Performance on a Perceptual Evidence Accumulation Task

Cristina Delgado Sallent[1,2] CDS4@BU.EDU
Benjamin Scott[1] BBS@BU.EDU
Sanaa A. Ahmed[1] SANAAAHM@BU.EDU
Anosha Khawaja-Lopez[1] AKHWJA@BU.EDU
Juliana Gomez[1] JULIEGOM@BU.EDU
Steve Ramirez[1] DVSTEVE13@GMAIL.COM
Arula Ratnakar[1] ARULA@BU.EDU

[1]Boston University
[2]Department of Psychological and Brain Sciences

Psychedelics are effective in treating psychiatric disorders, often producing sustained effects after a single dose. However the neural mechanisms underlying their action is poorly understood. Perceptual changes are frequent and robust features of the psychedelic experience and therefore investigation into how they alter perceptual integration is crucial in order to understand their mechanism of action. Here we characterize ketamine's therapeutic properties by combining a perceptual evidence accumulation task in mice, and brain-wide mapping of drug-responsive populations by using activity-dependent gene expression. Specifically, we trained mice on a free-response, pulse-based perceptual integration task. We first find that after ketamine administration at medium-high doses (30-50 mg/kg) mice slowed their response times (RT) and increased accuracy in the task. Drift diffusion model (DDM) fits suggest that these behavioral changes were due to an increased boundary separation and non-decision time. RTs were well fit by the DDM, and parameter fits suggested that mice use a multi-flash accumulation strategy. To further characterize which brain regions are activated by ketamine administration (30 mg/kg), we combined whole-brain immuno-staining of the activity-dependent immediate early gene, c-fos, with the theoretical framework of graph theory. First, we observed that ketamine produced a more interconnected network topology by applying Leiden clustering algorithms to the c-fos density correlation networks. By measuring the difference in cfos expression between ketamine and saline and its correlations with the task performance, we found that several regions correlated with the task performance and the DDM parameters, particularly the superior colliculus (SC), that we identify as a key region (hub) for ongoing experiments seeking to bi-directionally perturb SC's activity. Taken together, our results indicate that subanesthetic doses of ketamine improves perceptual evidence accumulation by raising the decision threshold and allowing the integration of more evidence before the decision and this may be facilitated by the SC.

3-045. The role of input synchrony in the generation of dendritic representations of sensorimotor behavior

Jacob Gable GABLE098@UMN.EDU
Zachary Newman NEWMANZA@UMN.EDU
Sarah Young YOUN2941@UMN.EDU
Jackson Scheib SCHEI399@UMN.EDU
Savannah Bliese BLIES035@UMN.EDU
Nicole Simco SIMCO017@UMN.EDU
Aaron Kerlin AKERLIN@UMN.EDU

University of Minnesota

Individual cortical neurons transform information from thousands of presynaptic inputs conveying diverse functional information. At any given moment, individual neurons may participate in one of the many distinct functional subnetworks, but the mechanisms that control this multiplexing remain poorly understood. It has been suggested that synchronous spiking of upstream neuronal ensembles – driven by behavioral context – may engage distinct subnetworks by facilitating the generation of strongly depolarizing regenerative events (dendritic spikes) in dendrites of downstream neurons. However, direct evaluation of this model requires measurements of the spatiotemporal patterns of input to cortical dendrites with millisecond precision and during behavior. We achieve this by co-expressing a recently developed fast glutamate sensor (iGluSnFR3) with a red calcium sensor (jRGECO1a) in the apical dendrites of layer 5 neurons in premotor cortex of mice. While mice engaged in a delayed-response directional licking task, glutamatergic input and dendritic calcium transients were imaged simultaneously across up to 50 dendritic spines and multiple dendrite branches. Correlations between nearby inputs (< 5 μm separation) targeting the same branch exhibited significantly higher correlation than inputs at longer distances or to neighboring dendritic branches, particularly at timescales less than 5 ms. Synchronized input to dendrite branches was also significantly correlated with the generation of dendritic calcium transients in the dendritic shaft of those branches. Finally, by analyzing input patterns that distinguish trial types (i.e., coding directions) at different spatial and temporal timescales, we have begun to determine if synchronized input to dendritic branches

facilitates the engagement of distinct subnetworks during different behavioral conditions (i.e., motor preparation, motor execution and performance outcome).

3-046. More Diffusive Replay Sequences Correlate with Longer Theta Sequences in the Hippocampus

Zilong Ji[1,2] ZILONG.JI@UCL.AC.UK
Neil Burgess[1,2] N.BURGESS@UCL.AC.UK

[1]University College London
[2]Institute of Cognitive Neuroscience

Replay dynamics plays an important role in hippocampal-dependent cognitive functions, including learning, generalization, and planning. While previous studies primarily focused on statistically identifiable replay sequences, recent research has unveiled diverse replay dynamics such as stationary, Brownian diffusion, and superdiffusion. Despite these advancements, the precise mechanisms governing the variety of replay dynamics remain elusive. In this study, we found that the diffusivity of replay dynamics correlates with the length of theta sequences. Specifically, longer theta sequences result in more diffusive replay dynamics. Notably, no such correlation was observed between replay dynamics and behaviorally relevant movement dynamics, indicating a specific link between online dynamics at the "theta" timescale and replay dynamics. To unravel the circuit mechanisms underlying this covariance, we constructed a hippocampal circuit model capable of generating both theta sequences and replay sequences. Our model suggested that the mobility of intrinsic dynamics in the hippocampus governs both sequential dynamics. Higher intrinsic mobility leads to longer theta sequences and more diffusive replay sequences. These findings reconciled disparate views on sequential dynamics in the hippocampus and establish a principled foundation for understanding information processing in this crucial brain region.

3-047. Tracking neurons across days with high-density probes

Enny van Beest[1,2] ENNYVANBEEST@GMAIL.COM
Celian Bimbard[1] C.BIMBARD@UCL.AC.UK
Julie Fabre[1] J.FABRE@UCL.AC.UK
Flora Takacs[1] FLORA.TAKACS.15@UCL.AC.UK
Pip Coen[1,3] P.COEN@UCL.AC.UK
Anna Lebedeva[1] ANNA.LEBEDEVA.17@UCL.AC.UK
Kenneth Harris[1] KENNETH.HARRIS@UCL.AC.UK
Matteo Carandini[1] M.CARANDINI@UCL.AC.UK

[1]University College London
[2]Institute of Ophthalmology
[3]Cell and Developmental Biology

Neural activity spans multiple scales from milliseconds to months. Its evolution can be recorded with chronic electrodes, and especially with high-density arrays such as Neuropixels probes, which measure each spike at tens of sites and record hundreds of neurons. These arrays often record units with consistent spike waveforms over time, but produce vast amounts of data that require new approaches for tracking neurons across many recordings. To meet this need, we developed UnitMatch, an open-access pipeline that operates after spike sorting, and is based only on parameters extracted from each neuron's spatiotemporal spike waveform for each half of every recording. We computed the similarity of parameters for all pairs of units to identify putatively tracked neurons across days, based on the distribution of similarity scores between halves of a recording within day. Putatively tracked neurons were then used to build probability distributions of being the same unit given the different similarity scores, which were fed into a Naive Bayes classifier. The resulting posterior probability was used to identify which neurons were the same across days. We tested UnitMatch in Neuropixels recordings from many different areas in the mouse brain, where it tracked neurons across weeks. We validated successful tracking with functional properties that were remarkably stable across weeks, yet distinctive across neurons. One such fixed property is a neuron's correlation with other neurons, which remained stable over weeks. Another such fixed property is the selectivity for visual stimuli in the visual cortex, which remained stable over weeks. UnitMatch can also be used to measure neural properties that change over time. We used it to track striatum neurons across days while mice learned a task, and found diverse changes across neurons over days. UnitMatch is thus a promising tool to reveal invariance or plasticity in neural activity across days.

3-048. Bayesian Inference of Nonlinear Neural Manifolds Made Easy

Isabel Garon[1,2] ISABELGARON@NYU.EDU
Stephen Keeley[3] SKEELEY1@FORDHAM.EDU
Alex Williams[1] ALEX.H.WILLIAMS@NYU.EDU

[1]New York University
[2]Neuroscience Institute
[3]Fordham University

Computational neuroscientists have invested substantial effort into extracting nonlinear latent dynamics and manifolds from high-dimensional neural recordings. Neural datasets are noisy and limited in duration, so it is desirable for these methods to be data-efficient and amenable to uncertainty quantification. Bayesian models are appealing in both these regards, and several methods in this category have been developed for neural data. However, inference in these settings is complicated by the nonlinear mapping of latent variables to firing rates, and non-Gaussian distributions of noise. Past work has partially overcome these challenges with approximate inference schemes (e.g. variational methods or other posterior approximations) which are often mathematically technical and inflexible (e.g. specific to Poisson likelihood or manifold structure of the latent space). The resulting models can also have many hyperparameters—e.g. inducing points schemes or deep networks for amortized inference. Our work revisits these approaches and shows that they can be both conceptually simplified and algorithmically improved. We exploit the fact that most neural manifolds have low intrinsic dimension, leading to two important consequences. First, the integral over latent variables—which is invariably required for this variety of probabilistic model—is numerically tractable. Second, nonparametric functional priors (e.g. Gaussian process priors) can be well-approximated by finite sets of basis functions. Together, these qualities enable us to approximate inference through numerical integration techniques like Quasi and Sequential Monte Carlo. Importantly, the approximation error incurred by these methods is small in typical neuroscience applications, and disappears entirely as one devotes more computational resources to the inference algorithm. Furthermore, aside from specifying the prior distribution, essentially no hyperparameters need to be tuned for the inference procedure (e.g. inducing points). We demonstrate the efficacy of this method on synthetic data, head direction cell data, and primate motor cortical data.

3-049. Frontal cortical circuit dynamics for sensorimotor associative learning

Jianxiang Zhou[1,2] JXZHOU@ION.AC.CN
Ninglong Xu[1,3] XUNL@ION.AC.CN

[1]Chinese Academy of Sciences
[2]Institute of Neuroscience
[3]Institute of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology

The ability to form arbitrary associations between sensory stimuli and voluntary actions based on outcomes is essential for behavioral flexibility. While much has been known on instrumental learning, the neuronal-level mechanisms through which cortical circuits dynamically engage in learning new associations remain elusive. Here, we investigate the role of the anterior lateral motor (ALM) cortex in learning new stimulus-action-outcome contingencies during a perceptual decision-making task. After acquiring an auditory-based choice task, mice were trained to learn the reversed association between sound categories and motor choices. Using trial-interleaved optogenetic inhibition, we established that ALM neurons play a causal role in choice behavior during learning phase, but gradually become dispensable with training. Using chronic in vivo two-photon imaging, we longitudinally tracked the activity of the same population of excitatory ALM neurons throughout the learning process. This revealed a gradual recruitment of task-related neurons and a drifting of the task representations across different neurons over time. Intriguingly, the activity of the tracked neuron populations progressively advanced in time, leading to more synchronized choice-selective population activity aligned to stimulus onset, suggesting the emergence of a population representation of learned stimulus-action association. The temporal dynamics of population activity and learning behavior were both recapitulated by a recurrent spiking neural network (RSNN) model incorporating asymmetric spike-timing dependent plasticity (STDP) rules gated by global reward signals. Moreover, directly measuring dopamine signals in ALM using a fluorescent dopamine sensor revealed task-related global dopamine signals. Together, our findings suggest that frontal cortical populations contribute to associative learning with evolving population dynamics shaped by STDP learning rules and global reward signals.

3-050. Inhibitory plasticity supports consolidation of generalizable memories

Zhenrui Liao[1,2] ZHENRUI.LIAO@COLUMBIA.EDU
Satoshi Terada[1] ST3166@COLUMBIA.EDU
Darian Hadjiabadi[3] DHH@STANFORD.EDU
Ivan Raikov[3] IRAIKOV@STANFORD.EDU
Ivan Soltesz[3] ISOLTESZ@STANFORD.EDU
Attila Losonczy[1] AL2856@COLUMBIA.EDU

[1]Columbia University
[2]Department of Neuroscience
[3]Stanford University

The hippocampus reinstates learned sequences in temporally compressed "replay" epochs inside sharp-wave ripples (SWRs), which occur while an animal is at rest. Replay is required for the long-term consolidation of memories. However, the content of these replay episodes, and in particular whether specific experiences or a generalized representation of the world are consolidated, remains controversial. Using two-photon calcium imaging of CA3 outputs combined with simultaneous local field potential (LFP) recordings, we show that the statistics of replay deviate significantly from those of experience: environmental features which are highly salient in experience may be either selected or suppressed in replay depending on their statistical generalizability. We propose a parsimonious novel mechanism for this phenomenon: the reuse of the symmetric spike-time dependent plasticity rule (sSTDP), previously reported at CA3 excitatory synapses, to remodel inhibitory synapses as well. We show using three levels of modeling—spiking network, detailed biophysical, and abstract—that this mechanism enables efficient inference of the latent statistical structure of the world given noisy observations. We analyze mathematically how sSTDP shapes sequence dynamics in a recurrent network, and show that replay is a consistent estimator of a latent sequence under an injection noise model. Finally, using sparse optogenetic manipulation, we experimentally test the model prediction that artificially created nongeneralizable representations will be suppressed from SWRs. Our experimental and theoretical work here outlines a potential direct link between the synaptic and cognitive levels of memory consolidation.

3-051. Dynamics of a neuronal central pattern generator to control the REM/non-REM sleep cycle in lizards

Juan Luis Riquelme JUAN.RIQUELME@BRAIN.MPG.DE
Lorenz Fenk LORENZ.FENK@BRAIN.MPG.DE
Gilles Laurent GILLES.LAURENT@BRAIN.MPG.DE

Max Planck Institute for Brain Research

Sleep is ubiquitous in the animal kingdom, yet our understanding of its control and evolution remains rudimentary. Biphasic sleep is widespread, existing in mammals, birds, non-avian reptiles and fish, suggesting a common origin and conserved function. Two main sleep states are recognized: rapid eye movement (REM) sleep and slow-wave (non-REM) sleep. Early mammalian studies and computational models hinted at a central pattern generator (CPG) circuit in the brainstem alternating between these phases.

Recent studies in the Australian dragon Pogona vitticeps, have identified a reptilian homolog of the mammalian claustrum, which autonomously generates non-REM sleep patterns. In contrast, during REM, the bilateral claustra exhibit tightly coordinated activity, likely a reflection of midbrain input characterized by winner-take-all dynamics.

This study combines new experiments and computational analysis to explore the dynamics of biphasic sleep. By investigating the responses to stimulation of the circuits that underlie this rhythm, our results shed light on the existence of a CPG-type circuit in lizards. Specifically, we exploit the accessibility of claustrum and the extreme regularity and short duration of the Pogona sleep cycle (120 seconds) and reveal hallmarks of CPG circuits described in other systems. Firstly, we demonstrate phase resetting in response to a phasic light input, coupled with phase-dependent dynamics resulting in either phase delays or advances. Secondly, unilateral stimulation reveals traits of coupled bilateral oscillators. Finally, we show the entrainment of an identical rhythm in the awake animal when sensory statistics match the natural sleep rhythm. Altogether, our experiments and analysis provide evidence for the presence of a CPG-type circuit driving the biphasic sleep rhythm in lizards, contributing to the broader understanding of sleep mechanisms in the animal kingdom.

3-052. A novel framework for social learning in male Drosophila

Frederic Roemschied[1,2] FREDERIC.ROEMSCHIED@GMAIL.COM
Osama Ahmed[3,4] OMAHMED@UW.EDU
Elise Ireland[5] ELISE.C.IRELAND@GMAIL.COM
Adam Calhoun[6] ADAM.CALHOUN@GMAIL.COM
Minseung Choi[7] MINSEUNG@STANFORD.EDU
Mala Murthy[5] MMURTHY@EXCHANGE.PRINCETON.EDU

[1]University Medical Center Goettingen
[2]European Neuroscience Institute
[3]University of Washington, Seattle
[4]Psychology
[5]Princeton University
[6]Meta Reality Labs
[7]Stanford University School of Medicine

Mammals, birds, and insects learn from social experience to increase their social status, reproductive success, and chances of survival. Inability to adjust to social experience (as in autism spectrum disorder) can severely impair life in a social world. Yet, we do not fully understand how social experience shapes behavioral strategies through learning, because in most systems we lack i) tools to perturb social interactions, and ii) a general framework to measure behavioral strategies and their dependence on social experience. We aimed at closing both gaps: we first developed an analytical framework that quantitates social experience and behavioral strategies by decomposing data from automated pose estimation in freely behaving animals into a set of base experiences and base strategies. We then applied this framework to show that the innate courtship strategy of male Drosophila is shaped by past social experience. Specifically, we developed a novel assay for social learning, comprising two experiments: in a first training experiment, we used closed-loop optogenetics to control female feedback to the song males produce (via wing vibration) during courtship, to systematically perturb male social experience around the time of song. In a subsequent test experiment, we applied our analytical framework to evaluate the trained males' song strategies towards females providing unperturbed feedback. Compared to controls, males that had experienced perturbed feedback during training used distinct song strategies during testing, suggesting that the innate courtship strategy is shaped through learning from social experience. Known learning mutants and males with genetically downregulated dopamine receptor expression lacked social learning in our assay, corroborating our findings. These results demonstrate a surprising flexibility of fly courtship behavior, and they open the door for a circuit-level understanding of this type of social learning. Our analytical framework is applicable in any system to quantitate behavioral strategies and their dependence on social experience.

3-053. Using information bottleneck methods to understand parallel channels in visual motion detection

Tianzhi Lambus Li[1,2] TLI@G.HARVARD.EDU
Siwei Wang[3] SIWEIW@UCHICAGO.EDU
James E Fitzgerald[4] FITZGERALDJ@JANELIA.HHMI.ORG
Damon A Clark[5] DAMON.CLARK@YALE.EDU

[1]Harvard Medical School
[2]Department of Neurobiology
[3]University of Chicago
[4]HHMI Janelia Research Campus
[5]Yale University

Animals have evolved visual circuits that detect motion, a crucial task for survival. Canonical models compute motion from spatiotemporal correlations in visual stimuli. However, early visual systems of Drosophila divide motion signals into ON/OFF channels with opposing direction selectivity. Canonical models have not explained why these parallel channels have their specific representation of motion. Here we used the information bottleneck (IB) framework and its associated compression-retention tradeoff to reveal principles underlying parallel channels encoding visual motion.

Our work was inspired by the anatomical bottleneck in the eye: numerous photoreceptors converge into fewer channels that represent motion. Such systems face a fundamental tradeoff between compressing the stimulus and retaining velocity information. We therefore aimed to understand the structure of motion channels in Drosophila via this tradeoff. First, we used the information bottleneck (IB) method to transform visual inputs into optimized encodings under various tradeoff conditions. We found that IB-optimized encodings employ both direction-selective (DS) and non-direction-selective (non-DS) components. Non-DS components contribute substantially to encoding the velocity, hinting at the potential role of non-DS signals in Drosophila's visual circuit. Then, comparing

canonical models to IB, we found the former neither maximally compressed the stimulus nor maximally retained velocity information due to neglect of higher-order statistics in the stimulus. Solving IB problems typically requires discretized quantities that are inconvenient for modeling continuous channels. We therefore used the variational information bottleneck (VIB), a deep learning approximation to IB, to find interpretable motion channels. We found that motion channels resembling Drosophila's arise naturally at a "diverging region" in the compression-retention tradeoff—where obtaining extra velocity information requires disproportionally more information about the stimulus. This suggests motion channels in Drosophila may favor compression over obtaining marginally more velocity information. In summary, the IB framework and the compression-retention tradeoff revealed principles underlying motion-encoding channels.

3-054. A biologically-constrained model of area MT neurons predicts perceptual tradeoffs between object motion and depth

Yelin Dong[1,2] YDONG33@UR.ROCHESTER.EDU
Zhe-XIn Xu[1] ZXU53@UR.ROCHESTER.EDU
Gregory DeAngelis[1] GDEANGELIS@UR.ROCHESTER.EDU

[1]University of Rochester
[2]Department of Brain and Cognitive Sciences

The ability to decompose retinal image motion into its causes is essential for estimating real-world object motion. When objects move relative to the world, the depth computed from retinal velocity (based on motion parallax cues) may differ from the depth computed from binocular disparity. A previous study (Kim et al., 2022) found that monkeys use cue conflicts between disparity and motion parallax cues to detect object movement (relative to the world). Previous behavioral work in human subjects, using a dual-report task, found that perceptual estimates of depth are biased by inference about object motion when the physical stimulus remains fixed. We want to understand whether these perceptual tradeoffs between object motion and depth might arise in a feedforward manner from a population of neurons with response properties similar to those found in MT. To do this, we extended a previous model (Xu & DeAngelis, 2022) of how MT neurons combine eye velocity and retinal velocity to signal depth from motion parallax. We incorporated disparity tuning into the new MT model, and we demonstrated that it performs well in fitting data from the Kim et al. (2022) study. Based on this model, we constructed a population of neurons using tuning parameters drawn from real MT neurons, and we simulated neural population responses to stimuli very similar to those used in the dual-report task with humans. We then trained a joint linear decoder to decode object motion (relative to the world) and depth simultaneously from this population of MT neurons. Our purely feedforward model qualitatively reproduces the behavioral results of the human experiment, in which the depth perception is biased by motion estimation. This suggests that causal inference computations underlying the perception of object motion and depth during self-motion could be implemented in a feedforward manner in the brain.

3-055. Data-constrained and generative RNN models of mouse cortical dynamics during navigation

Siyan Zhou[1,2] SZHOU@G.HARVARD.EDU
Ryan Badman[3,2] RYAN_BADMAN@HMS.HARVARD.EDU
Charlotte Arlt[1] CHARLOTTEARLT@GMAIL.COM
Kanaka Rajan[3] KANAKA_RAJAN@HMS.HARVARD.EDU
Christopher Harvey[1] CHRISTOPHER_HARVEY@HMS.HARVARD.EDU

[1]Harvard Medical School
[2]Neurobiology
[3]1. Harvard Medical School 2. Kempner Institute, Harvard University

Complex behaviors, such as goal-directed spatial navigation, emerge from neural activity in many brain regions. Even with advances in multi-site recording technologies, a theoretical framework is lacking for how functional ensembles of neurons across multiple brain areas dynamically coordinate the diverse computations underlying navigation. Here, we developed a generative multi-area recurrent neural network (RNN) model constrained by behavioral and neural data from mice performing a navigation task in virtual reality (VR). The model is trained by fitting each RNN unit to the instantaneous activity of a single neuron from a calcium imaging dataset containing thousands of neurons imaged simultaneously across multiple cortical areas (primary visual cortex (V1), secondary motor cortex (M2), posterior parietal cortex (PPC), and retrosplenial cortex (RSC)). Further, the model is trained to perform the same navigation task as mice, whereby the model generates locomotor outputs that match those of mice and these outputs update the model's location in a maze. Once trained, the model autonomously produces

single-trial neural dynamics and moment-by-moment behavioral outputs that recapitulate key features of mouse data, even on test data not used to train the model. The model also captures the encoding profiles of individual neurons and cortical areas. Even though the model makes no a priori assumptions about connectivity, interacting functional ensembles emerge within and between cortical areas and between units encoding specific features of the navigation task. Based on in silico perturbations and analyses of neural dynamics in the network, our model identifies a circuit that transforms navigational choices into moment-by-moment locomotor outputs through bidirectional interactions between recurrently connected RSC neurons encoding navigational trajectories and M2 movement-tuned neurons. Together, our work establishes new approaches for data-constrained and generative multi-area RNNs that recapitulate single-trial neural and behavioral dynamics and yield testable predictions about across-cortex circuits underlying navigation.

3-056. Flow-field inference from neural data using deep recurrent networks

Timothy Kim[1] TDKIM@PRINCETON.EDU
Thomas Luo[1,2] ZHIHAOL@PRINCETON.EDU
Tankut Can[3] TANKUT@IAS.EDU
Kamesh Krishnamurthy[1] KAMESHK@PRINCETON.EDU
Jonathan Pillow[1] PILLOW@PRINCETON.EDU
Carlos Brody[1,2] BRODY@PRINCETON.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute
[3]Institute for Advanced Study

Computations involved in processes such as decision-making, working memory, and motor control are thought to emerge from the dynamics governing the collective activity of neurons in large populations. But the estimation of these dynamics remains a significant challenge. Here we introduce Flow-field Inference from Neural Data using deep Recurrent networks (FINDR), an unsupervised deep learning method that can infer low-dimensional nonlinear stochastic dynamics underlying neural population activity. Using population spike train data from frontal brain regions of rats performing an auditory decision-making task, we demonstrate that FINDR outperforms existing methods in capturing the heterogeneous responses of individual neurons. We further show that FINDR can discover interpretable low-dimensional dynamics when it is trained to disentangle task-relevant and irrelevant components of the neural population activity. Importantly, the low-dimensional nature of the learned dynamics allows for explicit visualization of flow fields and attractor structures. We suggest FINDR as a powerful method for revealing the low-dimensional task-relevant dynamics of neural populations and their associated computations.

3-057. Dissecting a circuit for pain behavior under competing needs

Amadeus Maes[1] AMADEUS.MAES@GMAIL.COM
Ann Kennedy[1,2] ANN.KENNEDY@NORTHWESTERN.EDU
Nitsan Goldstein[3] NITSAN.GOLDSTEIN@GMAIL.COM
Nick Betley[4] JNBETLEY@SAS.UPENN.EDU

[1]Northwestern University
[2]Neuroscience
[3]Massachusetts Institute of Technology
[4]University of Pennsylvania

Animals often face multiple survival needs that compete for their attention. How does the brain balance competing needs to produce adaptive behavior? New technologies can record from deep hypothalamic nuclei thought to encode need-related states like aggression [1], hunger [2] and fear [3]. However, we lack models that predict how these states might interact to inform behavior. We investigated pain-coping behavior of mice, and its inhibition by the competing needs of thirst, hunger [4], and fear. To do so, we performed fiber photometry recordings in the parabrachial nucleus (PBN), a known sensory hub for pain, during formalin or spared nerve pain assays.

We related PBN activity and the animals' actions to a reinforcement learning model of pain-coping behavior. Our model starts from the assumption that mice behave to reduce pain with minimal effort. We define a two-dimensional state space of pain vs effort expenditure, and a two-action output space allowing the model to either expend effort to reduce pain, or to replenish effort. We find that our model can recapitulate pain-coping behavior, and that observed PBN activity can be decomposed into fast and slow components that correlate with the model's effort and pain states, respectively.

Monosynaptic rabies tracing revealed two distinct brain regions that release neuropeptide Y (NPY) onto PBN in response to hunger and predation, inhibiting pain-coping behavior. To ask what computational role NPY might

play, we manipulated various parts of our model to simulate its effect and asked which manipulation best matched observed data. Only a reduction in ascending pain input was consistent with the behavioral and neural data, thus we propose that NPY deprioritizes pain-coping in the presence of a competing need by gating pain processing in PBN. We believe our approach can be broadly used to delineate circuit mechanisms of behavioral prioritization under competing needs.

3-058. Decoding Stable Hippocampal Tasks in Contextual Learning via Dimensionality Reduction

Hannah Wirtshafter[1,2] HSW@NORTHWESTERN.EDU
Sara A. Solla[1] SOLLA@NORTHWESTERN.EDU
John Disterhoft[1] JDISTERHOFT@NORTHWESTERN.EDU

[1]Northwestern University
[2]Neuroscience

This study uses advanced dimensionality reduction techniques to determine if representations of task-relevant stimuli found in the hippocampus (HPC) can be maintained against the background of remapped place cells. We trained rats in trace eyeblink conditioning, an HPC-dependent conditioning task, in two different environments, while recording HPC cellular activity using calcium imaging. We then used CEBRA, a dimensionality reduction technique, to analyze the neural data and assess the consistency of task representations across contexts. We found a remarkable consistency in HPC task representation despite environmental place cell remapping: The HPC effectively preserves task-related information even as place cells remap. We found similar embedding geometries of cellular representations in both environments, and established the ability of models trained in environment A to decode task parameters in environment B. This consistency highlights the HPC's adaptability and robustness in learning and memory processes. This novel finding research addresses a critical aspect of neuroscience: how the HPC adapts its encoding strategies to facilitate learning and memory that generalize across contexts.

3-059. Simple synaptic modulations implement diverse novelty computations

Kyle Aitken KYLE.AITKEN@ALLENINSTITUTE.ORG
Luke Campagnola LUKEC@ALLENINSTITUTE.ORG
Marina Garrett MARINAG@ALLENINSTITUTE.ORG
Shawn Olsen SHAWNO@ALLENINSTITUTE.ORG
Stefan Mihalas STEFANM@ALLENINSTITUTE.ORG

Allen Institute

Since environments are constantly in flux, the brain's ability to identify novel stimuli that fall outside its own internal representation of the world is crucial for an organism's survival. Within the mammalian neocortex, inhibitory microcircuits are proposed to regulate activity in an experience-dependent manner and different inhibitory neuron subtypes exhibit distinct novelty responses. Discerning the modulation of neural circuits by experience can be daunting unless one has a biologically plausible mechanism to detect and learn from novel experiences that is both understandable and flexible. Here we introduce a learning mechanism, *familiarity modulated synapses* (FMSs), through which a network response encoding novelty emerges from unsupervised local synaptic modifications depending only on the presynaptic or both the pre- and postsynaptic activity. FMSs stand apart from other familiarity mechanisms in their simplicity: they operate under continual learning, do not require specialized architecture, and can distinguish novelty rapidly without requiring feedback. Implementing FMSs within an experimentally-constrained model of a visual cortical circuit, we demonstrate the generalizability of FMSs by reproducing three distinct novelty effects recently observed in experiments: absolute, contextual (or oddball), and omission novelty. Additionally, our model results in a set of diverse physiological responses across cell subpopulations, allowing us to analyze how their connectivity and synaptic dynamics influences their distinct behavior, leading to predictions that can be tested experimentally. Altogether, our findings demonstrate how simple plasticity mechanisms within the cortical circuit structure can give rise to qualitatively distinct novelty responses and the emergence of functional cell subclasses. The flexibility of FMSs opens the door to computationally and theoretically investigating how distinct synapse modulations can lead to a variety of experience-dependent responses in a simple, understandable, and biologically plausible setup.

3-060. The role of nucleus accumbens spiny projection neurons in action reinforcement at short- and long-time scales

Wan Chen Lin[1,2] WANCHENLIN@BERKELEY.EDU
Lung-Hao Tai[1] LTAI@BERKELEY.EDU
Albert Qu[1,3] ALBERT_QU@BERKELEY.EDU
Moses Lee[4] ANDREWMOSES.LEE@UCSF.EDU
Linda Wilbrecht[1] WILBRECHT@BERKELEY.EDU

[1]University of California, Berkeley
[2]Helen Wills Neuroscience Institutue
[3]Psychology; Center for Computational Biology
[4]University of California, San Francisco

The nucleus accumbens (NAc) core has long been associated with reinforcement of action. Multiple studies show activation of dopamine and glutamatergic terminals in the NAc leads to positive reinforcement of an action. Yet, the role of activity in NAc core spiny projection neurons (SPNs) in the reinforcement of past actions is still unclear. To address this gap, we used fiber photometry (FP) to record NAc SPN calcium (Ca2+) signals and optogenetic manipulation methods in adult D1-cre, D2-cre, and A2a-cre mice while they performed a 2-arm bandit task (2ABT). Our FP and multiple linear regression results suggested that D1 and D2/A2A-SPNs encoded current outcomes (in a manner consistent with prediction error signals) and carried a trace of past outcomes and actions over multiple trials. In optogenetic experiments, we found that transient stimulation of D1-SPNs was reinforcing at the time of reward outcome in the 2ABT (promoting repeating the most recent action) and supported nose poking in an intracranial self-stimulation assay (ICSS). D1-SPN inhibition in the 2ABT promoted switching to the alternative action. Stimulation of D2/A2A-SPNs had variable results but, on average, promoted repeating the most recent action in the 2ABT and had no significant impact on ICSS. We next redesigned the 2ABT to have center reward so it could not be solved using place preference. With the center reward 2ABT, we found that D1- and D2/A2A-SPNs stimulation still reinforced repeating the most recent action. Our results together show that NAc core SPNs maintain a trace of past actions and outcomes and that increases in SPN activity (either D1R+ or D2R+) at the time of outcome can contribute to reinforcement of past action.

3-061. Dynamic functional connectome reveals novel neural function in C. elegans

Luciano Dyballa[1,2] LUCIANO.DYBALLA@YALE.EDU
Samuel Lang[3] SAMUEL.LANG@UMASSMED.EDU
Eviatar Yemini[3] EVIATAR.YEMINI@UMASSMED.EDU
Steven Zucker[1,2] STEVEN.ZUCKER@YALE.EDU

[1]Yale University
[2]Computer Science
[3]University of Massachusetts Chan Medical School

We seek to infer a weighted functional connectome for C. elegans from dynamic brain-wide neural activity measurements, predictive of neural communities underlying a given behavior. Our unsupervised algorithm organizes putative connections—estimated similarities between dynamic neural activity—rather than static anatomical synaptic connections, and predicts behavioral communities that can be tested experimentally. One surprising prediction is confirmed.

The inference occurs in three steps. First, pairwise affinities between neuronal traces (brain-wide calcium activity) are computed in a differential fashion, providing an instantaneous measure of putative coordination across time. We develop a non-linear computation that allows for both excitatory and inhibitory interactions. Second, non-negative tensor factorization (NTF) organizes pairwise affinities directly, instead of individual neurons. The learned factors specify which groups of neurons are most likely interacting during which intervals in time, and for which animals. Since these factors imply functional motifs active at different time intervals, each affinity factor can be interpreted as a weighted network. Third, to realize the functional connectome for different intervals, we apply a generative model-based community detection algorithm to the functional motifs. Since time codes different experimental variables or behaviors, this provides a dynamic functional connectome: an atlas of neural motifs active during separate stages of an experiment.

Applied to data from sensory-avoidance experiments, we infer a community of known, putative salt-sensing neurons (among others), including AWB, the worm's sole aversive olfactory neuron. This surprising prediction of a role for AWB in the salt-sensing circuit was confirmed by comparing salt avoidance in worms where AWB was functional versus worms where AWB was silenced, thus validating the predictive power of our algorithm. It also suggests, more broadly, the potential for complex interactions between salt sensation and olfactory aversion cir-

cuits.

3-062. Representational sparsity determines representational stability in sensory cortices

Shanshan Qin[1,2] SSQIN@SEAS.HARVARD.EDU
Cengiz Pehlevan[1] CPEHLEVAN@SEAS.HARVARD.EDU

[1]Harvard University
[2]SEAS

Recent advancements in large-scale neural activity recordings have revealed a continuous evolution in neural population activity associated with familiar tasks, percepts, and actions over extended periods. The underlying mechanisms and functional implications of such “representational drift” remain poorly understood. In many sensory cortices, the stability of representation significantly depends on the nature of sensory stimuli. For instance, in mouse primary visual cortex, representational drift occurs for natural movie stimuli but not for drifting gratings. To understand the mechanism behind such stimulus-dependent representational drift in visual cortex, we propose that natural stimuli often elicit denser responses compared to their artificial counterparts, and synaptic noise has larger effect on the stability of denser representations. We evaluated this hypothesis by training a sparse coding network with continually updating synaptic weights. We found that representations for more complex image patches are denser and also exhibited more drift compared to simpler ones. This hypothesis aligns with experimental observations. To further comprehend the correlation between drift speed and representational sparsity, we constructed a mean-field model to discern how different noise sources cumulatively contribute to the drift. Our model offers a plausible explanation for stimulus-dependent representational drift and proposes several testable predictions.

3-063. Biologically plausible credit assignment without weight symmetry

Li Ji-An[1] JIAN.LI.ACAD@GMAIL.COM
Marcus Benna[2,3] MBENNA@UCSD.EDU

[1]University of California, San Diego
[2]UC San Diego
[3]Neurobiology

Backpropagation (BP), a cornerstone algorithm for training artificial neural networks, predominates in contemporary deep learning. Although highly successful, it is often deemed biologically implausible. A significant limitation arises from the requirement for exact symmetry between connections in the backward and forward pathways to backpropagate gradient signals accurately, as such symmetric connectivity patterns are not observed in biological brains. Researchers have proposed several algorithms to mitigate this symmetry constraint, such as feedback alignment (FA; Lillicrap et al. 2016) and direct feedback alignment (DFA; Nokland 2016). These algorithms employ fixed backward weights B, demonstrating alignment of the forward weights W with B during training (e.g., W ~B^T in FA). However, because of this divergence from backpropagation dynamics, they face challenges in deeper networks and especially in convolutional layers. In this study, we introduce the Product Feedback Alignment (PFA) algorithm. Our findings demonstrate that PFA can closely approximate BP, and achieve performance akin to BP in deep convolutional networks, surpassing FA and DFA. In PFA, the forward weights W align with the product of a pair of backward weights R and B (such that W ~(RB)^T) rather than with B. Our results offer a novel resolution to the longstanding weight symmetry problem, leading to substantially improved learning in deep convolutional networks compared to previous proposals.

3-064. The role of the cerebellum in fluid intelligence: An fMRI study

Anat Leibovici[1,2] LEI.ANAT@GMAIL.COM
Reut Raizman[3] REUTIM007@GMAIL.COM
Itzhaki Nofar[1] ITZHAKI.NOFAR@GMAIL.COM
Tik Niv[4] NIVTIK@GMAIL.COM
Sapir Maayan[1] MAAYAN.SAPIR16@GMAIL.COM
Tsarfaty Galia[5] GALIA.TSARFATY@SHEBA.HEALTH.GOV.IL
Livny Abigail[6] ABIGAIL.LIVNYEZER@GMAIL.COM

[1]Sheba Medical Center, Tel Hashomer, Israel
[2]Division of Diagnostic Imaging
[3]Sheba Medical Center, Tel Hashomer, Israel; Sackler Faculty of Medicine, Tel-Aviv University
[4]Sackler Faculty of Medicine, Tel-Aviv University; Sagol School of Neuroscience, Tel Aviv University
[5]Sackler Faculty of Medicine, Tel-Aviv University; Sheba Medical Center, Tel Hashomer, Israel
[6]Sheba Medical Center Tel Hashomer, Israel; Sackler Faculty of Medicine, Tel-Aviv University

Fluid intelligence has been defined as the capacity to solve novel problems through adaptive reasoning, relational reasoning, and goal-directed decision. The importance of these cognitive abilities has motivated efforts to identify neural correlates of fluid intelligence. Traditionally, neuroimaging studies of fluid intelligence have focused on brain activation in frontal-parietal regions. In the past decade evidence has accumulated regarding the involvement of the cerebellum in higher cognitive function. In the current study we aimed to further investigate the role of the cerebellum in processing of fluid intelligence. We scanned 39 healthy participants (13 females, 26 males), recruited from the general population. Participants performed a novel abstract reasoning functional Magnetic Resonance Imaging task, modeled after stimuli from the advanced Raven's Progressive Matrices test. Analyses of both brain function and neural circuit architecture, focusing on hubness, were performed. We demonstrate activation in frontal-parietal well-known regions, together with an extensive activation in several cerebellar sub-regions. Moreover, four cerebellar regions served as crucial hubs in the neural topographical architecture circuit, indicating that while performing a fluid reasoning process, the cerebellum served as a central hub with high connectivity and centrality. Therefore, we are first to provide evidence for the cerebellum's role in fluid intelligence both by means of task brain activation and graph theory topology of the neural network. Previous studies have positioned the cerebellum as a "supervised learning machine" of spatiotemporal information. Our findings suggest that it is also crucial for controlling the reasoning process. This may advance machine learning models of fluid intelligence, by establishing the cerebellum as a component responsible for the fine tuning of the cognitive system. Moreover, pathologies of neural circuits can cause cognitive impairments. Future studies should further assess cerebellar contribution to cognitive processing in different brain disorders involving neural network plasticity, allowing a better understanding of cognitive deficits.

3-065. Distinct policy identification via model-based belief update

Zhe Li[1,2] ZHEL@BCM.EDU
Panos Alefantis[3] PA77@NYU.EDU
Noushin Quazi[4] NQUAZI@ANDREW.CMU.EDU
Dora Angelaki[3] DA93@NYU.EDU
Xaq Pitkow[4] XPITKOW@ANDREW.CMU.EDU

[1]Baylor College of Medicine
[2]Neuroscience
[3]New York University
[4]Carnegie Mellon University

Both artificial and biological agent face the challenge of interacting with the world based on incomplete and noisy observations. In general the agent relies on a latent representation that integrates past experience, and makes decisions based on it. When the agent has an internal model of the environment, its latent representation can be a distribution about the unknown state, i.e. belief. A major challenge in using this framework is that the model's belief dynamics are generally difficult to derive by hand, and this approach is certainly not scalable to complex environments. We developed an algorithm that automates this process, requiring only a simulator of the environment dynamics that the agent assumes. Our approach uses this simulator to estimate dynamics of the internal model beliefs. We applied our algorithm to explain behavioral data of a monkey foraging experiment, where the monkey needs to navigate between three food boxes and decide when to open the box based on noisy color cues. We found the beliefs span on a low dimensional manifold in the belief space, and they are well organized by behaviorally relevant variables. We also developed an algorithm to identify distinct modes of behavior, accommodating potentially non-stationary policies. We verified the algorithm on simulated behavior data, and show an initial application to real behavior, revealing changes in strategy over time. In summary, our tools allow us to infer latent representation of model-based agents and the policy defined over this representation, which will enable a

search for neural evidence of possible internal model properties that are not directly measurable.

3-066. Functional Connectivity in Area V1: Identifying Neuronal Modules Within and Across Layers

Maria Papadopouli[1,2] MGP@ICS.FORTH.GR
Manos Koniotakis[3] MANOSKONIOTAKIS@GMAIL.COM
Ioannis Smyrnakis[4] ISMYRNAKI@GMAIL.COM
Marios Alexios Savaglio[1] SAVAGLIO@ICS.FORTH.GR
Christina Brozi[1] BROZICHRISTINA@GMAIL.COM
Eleftheria Psilou[3] PSILOU.ELEFTHERIA.8485@GMAIL.COM
Ganna Palagina[5] GPALAGINA@BWH.HARVARD.EDU
Andreas S. Tolias[6] ASTOLIAS@BCM.EDU
Stelios Manolis Smirnakis[5] SMSMIRNAKIS@BWH.HARVARD.EDU

[1]University of Crete and Foundation for Research and Technology-Hellas
[2]Computer Science
[3]Foundation for Research and Technology-Hellas
[4]Hellenic Mediterranean University
[5]Brigham and Women' s & Jamaica Plain VA Hospitals, Harvard Medical School
[6]Baylor College of Medicine

Ensembles of neurons firing in-synchrony are likely to be more efficient at relaying information to downstream targets [1] and to belong to networks subserving similar functions [2,3]. Functional connectivity analysis performed during epochs of spontaneous activity (absence of visual-stimulation) allows us to identify groups of neurons that participate in synchronous-activity patterns, and thus, have the potential to coordinate in processing information. Identifying functional-connectivity patterns under spontaneous conditions and studying potential implications they may have for visual stimulus encoding is important: for example, it has been suggested that spontaneous activity patterns may span a “vocabulary space” shared with population activity patterns elicited during sensory responses. Nevertheless, the relationship between stimulus-driven and spontaneous activity patterns remains obscure. Here we used the Spike Time Tiling Coefficient (STTC) metric of inter-neuronal correlation strength [4] to measure pairwise functional-connectivity between pyramidal neurons in granular (L4) and supra-granular layers (L2/3) of mouse area V1 recorded via large field-of-view 2-photon microscopy. Approximately 15-25% of neuronal pairs are functionally-connected at high statistical-significance (z-score>4) both within and across layers 2/3 and 4. Compared to L2/3, L4 exhibits higher percentage of strong pairwise-correlations, higher clustering-coefficients as well higher and more uniformly distributed degrees-of-connectivity, suggesting a less hierarchical, more robust hub-structure. Both layers' architecture exhibits small-worldness and robustness. Of particular interest is the set of L4-pyramidal neurons that are functionally-connected to a single L2/3 neuron, constituting its “putative-input” group. The probability of L2/3 neuron firing depends strongly on the level of activity in its L4 putative-input group and is modulated strongly by brain-state as reflected by the pupil size and overall V1-population activity. Interestingly, L2/3 neurons with small L4 putative-input groups behave differently than those with large ones both with respect to the slope of the prediction (fig. 4) and their modulation by behavioral state. In contrast, they do not seem to markedly differ along dimensions related to visual properties such as orientation- or direction-of-motion preference. In sum, resting-state functional connectivity analysis in area V1 allows us to classify pyramidal neurons into different hub classes depending on their functional connectivity, with interesting implications for stimulus encoding.

3-067. Human and rodent perceptual biases emerge in a recurrent neural network with ongoing Hebbian plasticity

Francesca Schonsberg FSCHONSB@SISSA.IT
Davide Giana DGIANA@SISSA.IT
Yukti Chopra YCHOPRA@SISSA.IT
Mathew E. Diamond DIAMOND@SISSA.IT
Sebastian Goldt GOLDT.SEBASTIAN@GMAIL.COM

International School for Advanced Studies (SISSA)

Perceiving the magnitude of a stimulus is a complex brain function that arises from the interplay of working memory and experience. This interplay results in two well-known perceptual biases in memory tasks: 1. In a series of vibrational stimuli of varying strength, both humans and rodents tend to overestimate the strength of a stimulus after a series of weak stimuli (and vice versa) due to the repulsive bias of representations. 2. The contraction bias

instead shifts the representation of a stimulus held in working memory towards the average of stimuli observed in the past. While a series of experiments have yielded a detailed phenomenological description of both biases, the neural mechanisms underlying these biases remain poorly understood.

Here, we show that the representations learnt by recurrent neural networks with ongoing Hebbian plasticity quantitatively reproduce (i) the contractive bias we find in experiments with human subjects, and (ii) the repulsive effect found in rodents by Hachen et al. (Nat. Comm. 2021). In our model, a fully-connected network of rate-based units is driven by external inputs modeled after the experimental protocol, while its connectivity is continuously evolving due to Hebbian plasticity. We do not use gradient descent nor do we fine-tune the model to different experimental paradigms. We finally design a new behavioural paradigm where contraction and repulsive bias interact and find again that the model predicts salient features of the performance of our human participants.

Our results show that a single recurrent neural network with ongoing Hebbian plasticity reproduces two perceptive biases observed across three experimental paradigms. The striking match between experimental data and theoretical predictions supports the hypothesis that perceptual biases arise from simple Hebbian plasticity within a unique recurrent subregion of the brain, e.g. vM1 in rats, which acts as a plastic platform that filters perception based on context.

3-068. Cerebellar encoding of temporal prior knowledge

Julius Koppen[1,2] J.KOPPEN@ERASMUSMC.NL
Ilse Klinkhamer[1] ILSE.KLINKHAMER@HOTMAIL.COM
Marit Runge[1] M.RUNGE@ERASMUSMC.NL
Devika Narain[1,3] D.NARAIN@ERASMUSMC.NL

[1]Erasmus University Medical Center
[2]Department of Neuroscience
[3]Neuroscience

Behavior in the natural world is rife with feats of temporal precision but laboratory measurements of such behaviors reveal surprising biases in temporal perception. Previous theoretical work attributes some of these biases to Bayesian inference processes that increase their reliance on prior knowledge of time intervals under uncertainty. Using theory, electrophysiology, and optogenetics, we investigated whether cerebellar circuits could provide a substrate for encoding prior knowledge of temporal statistics. Mice learned time interval associations, which were drawn from different temporal prior distributions in a modified eyelid conditioning task. We found that as these mice switched from a narrow to a wide temporal prior distribution, the predictive eyelid behavior changed in a manner consistent with model-predicted changes in temporal statistics. To examine whether the same was true for the activity of cerebellar cortical neurons, we used a variational autoencoder-based technique, LFADS, to decode trial-by-trial activity from cerebellar cortical neurons to show correspondence between the changing temporal statistics and the activity of Purkinje cells and putative molecular layer interneurons. We found that the activity of individual cerebellar cortical neurons and the population dynamics also changed concomitantly with variations in the temporal statistics of the stimuli. Calibrated optogenetic perturbations to cerebellar Purkinje cells within the duration of the temporal distribution caused a complete suppression of the prior-related response. Furthermore, we found prior-related signaling in cerebellar Purkinje cell complex spike activity that was time-locked to the onset of prior distributions. Together these results provide support for a mechanism for encoding prior knowledge of time intervals in the cerebellar cortex.

3-069. Modeling conditional distributions of neural and behavioral data with masked variational autoencoders

Auguste Schulz[1,2] AUGUSTE.SCHULZ@UNI-TUEBINGEN.DE
Daniel Morales[3] DANIEL.MORALES@NATURE.COM
Victor Lobato Rios[3] VICTOR.LOBATORIOS@EPFL.CH
Pavan Ramdya[4] PAVAN.RAMDYA@EPFL.CH
Pedro Goncalves[5] PEDRO.GONCALVES@NERF.BE
Jakob Macke[1] JAKOB.MACKE@UNI-TUEBINGEN.DE

[1]University of Tubingen
[2]Excellence Cluster Machine Learning, Tubingen AI Center
[3]ecole Polytechnique Federale de Lausanne
[4]Ecole Polytechnique Federale de Lausanne
[5]VIB-Neuroelectronics Research Flanders, Belgium

One of the main goals in systems neuroscience is to understand how neural activity gives rise to complex behavior. Flexible dimensionality reduction methods, such as variational autoencoders (VAEs), have become commonly used tools for making sense of high-dimensional and multi-modal datasets. However, most VAE-based methods, even when trained jointly on two modalities, cannot adequately deal with a ubiquitous analysis task in neuroscience: calculating conditional distributions. For example, encoding studies in neuroscience require calculating the conditional distribution of neural activity given behavior (or other observations). Conversely, for decoding analyses, one needs to calculate the conditional over behavior, given neural activity. Finally, conditional distributions also arise when dealing with missing data or when we want to assess interactions of purely behavioral variables or between brain regions. Here, we present a VAE-based approach for accurately calculating such conditional distributions. We modify standard VAEs to model the distribution over a subset of the data (e.g., behavior) conditioned on the remaining data (e.g., neural activity) by adjusting the VAE loss and training through subset masking. We first validate our approach on a tractable task on which we have access to the ground-truth distributions and show how one can assess the quality of uncertainty estimates also for data without ground-truth. Next, we demonstrate applicability to high-dimensional behavioral time series by retrieving the conditional distributions over masked body parts of walking fruit flies. Finally, we show that we can obtain realistic samples from behavioral decoding distributions in a monkey reach task. In summary, we focus on learning conditional distributions using masked variational autoencoders. We show how to evaluate the models' uncertainty estimates in this context, a crucial but commonly neglected aspect in deep-learning-based dimensionality reduction. Our approach will allow for addressing central questions in neuroscience while coping with the scale and complexity of advancing datasets.

3-070. One-hot Generalized Linear Model for Switching Brain State Discovery

Chengrui Li[1,2] CNLICHENGRUI@GATECH.EDU
Soon Ho Kim[1,3] SOONHOKIM@GATECH.EDU
Chris Rodgers[4,5] CHRISTOPHER.RODGERS@EMORY.EDU
Hannah Choi[1] HANNAHCH@GATECH.EDU
Anqi Wu[6] ANQIWU@GATECH.EDU

[1]Georgia Institute of Technology
[2]Computational Science & Engineering
[3]School of Mathematics
[4]Emory University
[5]Neurosurgery
[6]georgia institute of technology

Exposing meaningful and interpretable neural interactions is critical to understanding neural circuits. Inferred neural interactions from neural signals primarily reflect functional interactions. In a long experiment, subject animals may experience different stages defined by the experiment, stimuli, or behavioral states, and hence functional interactions can change over time. To model dynamically changing functional interactions, prior work employs state-switching generalized linear models with hidden Markov models (i.e., HMM-GLMs). However, we argue they lack biological plausibility, as functional interactions are shaped and confined by the underlying anatomical connectome. Here, we propose a novel prior-informed state-switching GLM. We introduce both a Gaussian prior and a one-hot prior over the GLM in each state. The priors are learnable. We will show that the learned prior should capture the state-constant interaction, shedding light on the underlying anatomical connectome and revealing more likely physical neuron interactions. The state-dependent interaction modeled by each GLM offers traceability to capture functional variations across multiple brain states. Our methods effectively recover true interaction structures in simulated data, achieve the highest predictive likelihood with real neural datasets, and render interaction structures and hidden states more interpretable when applied to real neural data.

3-071. How noise sources shape cortical inter-areal communication

Joana Carmona[1,2] JOANA.CARMONA@RESEARCH.FCHAMPALIMAUD.ORG
Francesca Mastrogiuseppe[3] FRAN.MASTROGIUSEPPE@GMAIL.COM
Byron Yu[4] BYRONYU@CMU.EDU
Adam Kohn[5] ADAM.KOHN@EINSTEINMED.EDU
Christian Machens[3] CHRISTIAN.MACHENS@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Research, Champalimaud Neuroscience Programme
[3]Champalimaud Research
[4]Carnegie Mellon University
[5]Albert Einstein College of Medicine

Hierarchical visual processing relies on the directional propagation of signals across brain areas to subserve computational goals. Recent work leveraged simultaneous multi-area recordings to investigate the strength and directionality of inter-areal signaling, revealing that V1-V2 communication shifts from feedforward-dominated following stimulus onset to feedback-dominated during spontaneous activity. Importantly, this shift does not simply reflect a change in the directionality of the mean drive but rather results from trial-to-trial activity fluctuations predominantly traveling in different directions during evoked and spontaneous periods. What hierarchical circuit architectures support the observed directional flow of activity fluctuations? And what are the network mechanisms involved in the stimulus-dependent shift in the directionality of interactions? Here, we investigated these questions by using two-area recurrent neural networks (RNNs). We first investigated the conditions that lead to a directional flow of activity fluctuations, as observed in neural data, in linearized circuit models. Through mathematical analysis and simulations we concluded that, for arbitrary network connectivity, communication directionality is strongly determined by the difference in noise variance injected in the two areas. Based on this finding, we designed a model for noise sources in V1 and V2 that causes a shift in communication directionality, and is also consistent with biology. This includes a feedforward source associated with variability in the stimulus, and a persistent feedback source linked to slow global signals arising in higher-order areas. To evaluate this mechanism in a larger and functional network, we focused on a predictive coding network, an established normative model of V1-V2 dynamics. We found that simulated activity could qualitatively recapitulate the main features of inter-areal communication observed in the data. Our approach allowed us to gain insight into possible network mechanisms underlying bidirectional and selective signaling as well as to constrain the space of models of inter-areal communication that are consistent with experimental observations.

3-072. Rapid implicit learning of temporal context in a cerebellar task

Luca Mangili[1] L.MANGILI@ERASMUSMC.NL
Charlotte Wissing[1] C.C.WISSING@STUDENT.TUDELFT.NL
Devika Narain[2,3] D.NARAIN@ERASMUSMC.NL

[1]Dept. of Neuroscience, Erasmus University Medical Center, Rotterdam, Netherlands
[2]Erasmus University Medical Center
[3]Neuroscience

In motor neuroscience, implicit learning is characterized as a slow process, whereas explicit learning is expected to occur over more rapid timescales. In cognitive neuroscience, however, implicit processes are believed to be rapid and automatic, and explicit processes are considered slow and deliberate. Implicit processes are often believed to rely on associative learning and have been ascribed to subcortical regions like the cerebellum, whereas, explicit processes are generally ascribed to the frontal cerebral cortex. Recent work has revealed that these areas are connected in disynaptic functional loops, raising the question of how implicit and explicit processes are acquired in interconnected brain-wide circuits. We investigated a novel human context-dependent timing paradigm, which contained both explicit and implicit learning readouts that were previously shown to originate in frontal cortical and cerebellar circuits, respectively. Participants learned to associate different contextual cues with delayed interval responses. Our task used a periocular airpuff as error feedback, linking learning to eyeblink conditioning, widely studied in cerebellar circuits. In general, implicit behaviors such as predictive eyeblink closure are considered inflexible and their underlying circuitry is believed to consist of rudimentary cerebellar circuits. Here, we found that predictive eyeblink responses were able to develop a flexible context-dependent character such that the eye would preemptively close at different time intervals based on the context cue at the start of each trial. Moreover, we found, contrary to reports from motor neuroscience, that well-timed implicit responses were acquired faster than well-timed explicit manual press responses. Finally, we provide a computational model to explain how rapid and flexible implicit learning can be implemented in cerebellar circuits. Our model and behavioral findings generate testable predictions for cortical and cerebellar systems neuroscientists and suggest that implicit learning associated with cerebellar circuitry may be more sophisticated than previously believed.

3-073. Dual neuromodulator control of rapid synaptic plasticity

Mark Plitt[1,2] MARK.PLITT@BERKELEY.EDU
Dan Turner-Evans[3] DTURNERE@UCSC.EDU
Mark Eddison[4] EDDISONM@JANELIA.HHMI.ORG
Robert Ray[4] RAYR@JANELIA.HHMI.ORG
Tanya Wolff[4] WOLFFT@JANELIA.HHMI.ORG
Gerald Rubin[4] RUBING@JANELIA.HHMI.ORG
Vivek Jayaraman[4] VIVEK@JANELIA.HHMI.ORG
Yvette Fisher YFISHER@BERKELEY.EDU

[1]University of California, Berkeley
[2]Molecular and Cell Biology
[3]University of California, Santa Cruz
[4]Janelia Research Campus

The head direction (HD) circuit in fruit flies is a biological continuous attractor network. EPG neurons form the core of this HD ring attractor. Their dendrites form wedges that tile a ring-shaped structure called the ellipsoid body (EB). A localized bump of increased activity in the EPG dendrites represents the animal's heading, swinging around the ring like the needle of a compass as the fly turns. In order for EPGs to flexibly track an animal's orientation across diverse environments, they must tether to visual cues, i.e. the synapses between visual neurons (ER4ds, which are inhibitory) and EPGs must be plastic. In particular, ER4d synapses that physically coincide with the EPG activity bump should be weakened, disinhibiting the learned HD representation, while ER4d synapses that are away from the EPG activity should be strengthened, suppressing competing HD representations. This anti-Hebbian plasticity allows visual inputs to push the EPG activity bump to repeatable locations when the fly revisits particular orientations. This plasticity rule is computationally straightforward, but a physiological implementation for coincidence detection at an inhibitory synapse is not understood. We identified a novel mechanism that may achieve this learning rule by utilizing two distinct neuromodulatory signals. We believe a local octopaminergic signal from EL neurons instructs active ER4d synapses that coincide with the EPG activity bump to depress. We propose that a separate global dopamine signal, provided by ExR2 neurons, potentiates the remaining active ER4d synapses. ExR2 activity is correlated with the fly's rotational velocity, restricting plasticity to periods that contain more salient visual information. Through computational modeling, we demonstrate that decoupling synaptic potentiation and depression via this dual neuromodulator circuit endows the compact HD circuit of the fly with additional flexibility, capable of both rapid learning and resetting under different conditions.

3-074. Synaptic Theory of Working Memory for Serial Order

Gianluigi Mongillo[1] GIANLUIGI.MONGILLO@GMAIL.COM
Misha Tsodyks[2] MISHA@WEIZMANN.AC.IL

[1]Sorbonne Universite, Institut de la Vision
[2]Weizmann Institute of Science

Working Memory (WM) enables the temporally-ordered encoding of a sequence of stimuli and its maintenance over short periods of time. This ability is critical for the performance of many cognitive and behavioral tasks. In spite of its importance, however, how WM encodes, stores and retrieves information about serial order remains a major outstanding problem. As a matter of fact, the existing neurophysiologically-grounded theories of WM are unable to account for the encoding of serial order. Here, we extend the synaptic theory of WM to include synaptic augmentation besides synaptic depression and facilitation. Augmentation consists in an enhancement of the synaptic efficacy that builds up slowly with repetitive pre-synaptic activity and that, once induced, persists over tens of seconds in the absence of activity. Experimentally, augmentation is observed at the same synapses that exhibit significant short-term facilitation. We find that the level of synaptic augmentation within the neuronal population coding for an item naturally encodes the time during which that item has been maintained in WM. This time, in turn, correlates with the serial position of the item in the sequence: Items earlier in the sequence spend longer times in WM than items later in the sequence. The resulting "primacy gradient", carried by the augmentation levels, can be readily read-out by an appropriate control of the background input to the network, resulting in the items being recalled following the order of presentation. The model reproduces prominent features of the behavior of human subjects recalling lists of items and makes a series of experimentally-testable predictions. Finally, the model with synaptic augmentation exposes an intriguing trade-off between "storage" and "retrieval" capacity, suggesting that limited WM capacity could result from the inability to retrieve the information, rather than from the inability to encode and/or maintain it.

3-075. An efficient coding theory for cortical connectivity

Isabel Maria Cornacchia[1,2] ISABEL.CORNACCHIA@ED.AC.UK
Angus Chadwick[1,3] ANGUS.CHADWICK@ED.AC.UK

[1]University of Edinburgh
[2]Institute for Adaptive and Neural Computation
[3]School of Informatics

Cortical circuits transform sensory inputs into distributed neural firing patterns via the interactions of excitatory and inhibitory cell types. However, the principles relating cortical connectivity to efficient sensory codes are poorly understood. Two fundamental properties of cortical connectivity shape representations of sensory stimuli: 1) neurons with shared stimulus preferences connect more strongly than those with disparate preferences (stimulus-specific connectivity), and 2) excitatory and inhibitory synaptic inputs to each neuron are co-tuned and approximately equal in magnitude (E-I balance). A substantial literature has investigated the functional properties of circuits endowed with these properties and the learning rules that generate them. This effort has yielded mechanistic explanations for various phenomena observed experimentally in visual cortex, including contrast-invariant tuning curves and cross-orientation suppression. However, a normative, first principles explanation for cortical connectivity and the response properties it generates is currently lacking.

Here, we asked whether these properties could emerge from an efficient coding objective. To test this hypothesis, we developed a method to adjust the recurrent weights of an E-I network to maximise the Fisher information of the network response for a given ensemble of inputs. We found that networks optimised to encode stimulus orientation exhibit co-tuned E-E and E-I connectivity (stimulus-specific connectivity) and co-tuned E/I synaptic currents (E-I balance). Strong recurrent excitatory connectivity selectively amplifies input patterns, while recurrent inhibition maintains dynamical stability (inhibitory-stabilisation). Although the network was not directly incentivised to encode stimulus contrast, both E and I cells exhibited contrast-invariant tuning curves. Finally, although the network was optimised to encode a single stimulus orientation, the network exhibited cross-orientation suppression when two orientations were presented simultaneously. Taken together, we show that cortical connectivity and response properties can be accounted for by an efficient coding principle.

3-076. Continuous encoding of intent and error in the human motor cortex

Camille Gontier CAG329@PITT.EDU
Nicolas Kunigk NIK94@PITT.EDU
William Hockeimer WIH38@PITT.EDU
Edgar Canario EDC48@PITT.EDU
Brian Dekleva DEKLEVA@PITT.EDU
Jennifer Collinger COLLINGER@PITT.EDU

University of Pittsburgh

The motor cortex is known to encode future motor commands. Decoding motor intents from population activity in the M1 area is at the basis of brain-computer interfaces (BCIs). However, how overall activity in M1, and more specifically the mapping between population activity and motor commands, are perturbed by feedback signals during a motor task remains unknown. To study this, we analyze data from 2 human participants with tetraplegia controlling a BCI to perform a continuous 2D cursor control task and receiving visual feedback of the cursor movement. Epochs of correct and erroneous control feedback are defined based on whether the controlled cursor appears to move on or off-target. First, we highlight significant differences in the population activity of M1 between periods of correct and erroneous control: we show that the subspaces of neural activity spanned during these periods are significantly misaligned, and that neural activity during erroneous epochs is characterized by a dimensionality collapse. But whether these population activity changes are the cause of control errors, or the consequence of receiving feedback of the task output, is still unclear. We show that dimensionality collapse tends to precede the onset of control error, supporting the hypothesis that these errors are caused by endogenous changes in the population activity of M1 rather than by a sensitivity to task feedback. Finally, using these activity changes as a neural signature of ongoing error, we trained a decoder to detect them: this allows to perform online error detection during BCI control to stop ongoing errors without any specific action from the participant. Overall, our results highlight the role of the motor cortex as an optimal controller, leveraging separately upstream target and feedback loop signals to compute downstream control commands.

3-077. CA1 engram cell dynamics before and after learning

Amy Monasterio[1,2] AMONAST@BU.EDU
Caitlin Lienkaemper[1,3] CLIENKAEMPER.PSU@GMAIL.COM
Gabriel Ocker[1,3] GKOCKER@BU.EDU
Steve Ramirez[1] DVSTEVE@BU.EDU
Benjamin Scott[1] BBS@BU.EDU

[1]Boston University
[2]Psychological and Brain Sciences
[3]Mathematics and Statistics

Memories are believed to be stored in the strengthened synaptic connectivity between distributed networks of neurons. This neural memory trace is often referred to as an engram and recent work has identified c-Fos expressing populations in the hippocampus as putative cellular engram assemblies. The current model for engram formation predicts that learning drives synaptic potentiation between engram cells, resulting in their increased propensity to be reactivated together during memory retrieval, which has yet to be demonstrated in vivo. We began by modeling engram formation in a spiking neural network of CA3 and CA1. In this model we simulated engram formation with a network of CA3 and CA1 excitatory and inhibitory populations in which a subpopulation of excitatory engram cells in each subregion was directly imposed by increasing a subset of excitatory synaptic weights. Spontaneous firing of CA1 engram cells was increased and correlations between engram cells increased, in agreement with previous models of recurrent networks. Next, to test this model in vivo, we combined two-photon calcium imaging and viral tet-Tag labeling to identify CA1 engram cells and record spontaneous network activity before and after fear learning in mice. Surprisingly, we found no significant increase in spontaneous activity or correlations between engram cells after learning. We leveraged these findings to update our proposed model of excitatory synaptic strengthening between engram cells and tested the effect of including inhibitory plasticity. We found that the strengthening of inhibitory weights onto engram cells suppressed any increase in their spontaneous firing rates and correlations, more closely resembling our in vivo findings. Together, these approaches suggest that a model based only in excitatory synaptic changes during learning is inaccurate for engram formation. Finally, these results predict that engram cell dynamics may be sensitive to ongoing inhibition and future experiments may test this approach directly.

3-078. Neural basis and functions of flexible undulation in the head of C. elegans

Heng Zhang[1] HENGZHANGPHY@GMAIL.COM
Yifan Su[1] SU_0414@PKU.EDU.CN
Pinjie Li[2] LIPINJIE@MAIL.USTC.EDU.CN
Louis Tao[1] TAOLT@MAIL.CBI.PKU.EDU.CN
Quan Wen[2] QWEN@USTC.EDU.CN

[1]Peking University
[2]University of Science and Technology of China

Caenorhabditis elegans displays complex head exploratory behavior in natural environments. Here, we quantitatively analyzed the head movements and investigated the motor circuit that generates their intricate dynamics. Using variational mode decomposition (VMD), we distinguished between fast casts and slow bends in head movements. The slow bends uniquely propagate along the body, whereas the fast casts, which exhibit a phase-specific temporal distribution, are responsible for the phase lock between the head and body and contribute to the directional bias during a forward run. Combinatorial ablations of three types of cholinergic motor neurons and subsequent phase space analysis revealed their distinct roles in head movements. RMD motor neurons contribute to head casts; SMD maintain bending states; SMB and SMD neurons together facilitate slow rhythmic bending and head-body coupling. Our findings underscore that different types of excitatory motor neurons, despite sharing similar anatomical connectivities, control distinctly different aspects of exploratory head dynamics, thereby contributing to the overall complexity of motor neuron function and behavioral control.

3-079. Behavioral state regulates the role of somatostatin interneurons in stabilizing network activity.

Celine Cammarata[1,2] CELINE.CAMMARATA@DUKE.EDU
Yingming Pei[1] YINGMING.PEI@DUKE.EDU
Tammy Hawley[1] TAMMY.HAWLEY@DUKE.EDU
Shaun Sze-Xian Lim[1] SZE.XIAN.LIM@DUKE.EDU
David St-Amand[1] DAVID.ST-AMAND@DUKE.EDU
Michael Tadross[1] MICHAEL.TADROSS@DUKE.EDU
Nicolas Brunel[3] NICOLAS.BRUNEL@DUKE.EDU
Lindsey Glickfeld[1] GLICKFELD@NEURO.DUKE.EDU

[1]Duke University School of Medicine
[2]Neurobiology
[3]Duke University

Normalization is a key function of sensory cortices, allowing detection of weak stimuli without overexcitation to strong stimuli. A potential mechanism for normalization is the inhibition stabilized network (ISN), in which recurrent inhibition balances rising recurrent excitation. While there is evidence that sensory cortex can act as an ISN, how and when different types of interneurons contribute to inhibition stabilization is poorly understood. Somatostatin+ interneurons (SST) are strongly recurrently connected with neighboring Pyramidal cells (Pyr) and contribute to normalization and shaping Pyr output in high arousal states, but their specific role in ISN dynamics remains unclear. The tight coupling of inhibition and excitation in an ISN produces characteristic effects that can be used to detect ISN dynamics: exogenous suppression of inhibitory cells yields a paradoxical elevation of their activity through increased recurrent excitation. While past work has relied on optogenetic control of interneurons' excitability to test for ISN dynamics, the efficacy of this approach is likely to change as a function of input and state. Thus, we sought to impact the excitatory input that recruits interneurons into an ISN. We combined cell type-specific pharmacology with 2-photon calcium imaging to record activity of SST cells and neighboring Pyr cells in mouse primary visual cortex (V1) before and after antagonizing AMPA receptors on SST cells. This antagonism suppressed SST responses to weak visual stimuli, but the suppressive effect was attenuated by strong stimuli or locomotion. Analysis of a computational model including Pyr, SST and Parvalbumin+ (PV) cells revealed that this data is consistent with a network that is stabilized purely by PV cells in stationary conditions and with weak stimuli, but that requires SST stabilization during locomotion and with strong sensory stimuli. These results elucidate the conditions under which SST cells are necessary to stabilize cortical circuits.

3-080. Learning Successor Features the Simple Way

Raymond Chua[1,2] RAY.R.CHUA@GMAIL.COM
Blake Richards[1] BLAKE.RICHARDS@MILA.QUEBEC
Christos Kaplanis[3] CHRISTOS.KAPLANIS@GOOGLEMAIL.COM
Doina Precup[4] DPRECUP@CS.MCGILL.CA

[1]McGill University
[2]Computer Science
[3]Google DeepMind
[4]McGill University, Mila & Google DeepMind

Biological agents like primates exhibit lifelong learning and adaptation, processes that are mirrored in the continual Reinforcement Learning (RL) frameworks of computational models. In RL, Successor Representations (SRs) enable RL agents to exhibit flexible behaviour by capturing the world's transition dynamics using discounted state visitations, akin to how animals learn from experiences over time. Successor Features (SFs) build upon SRs by using artificial neural networks to extract features, an area where SRs, with their discrete representations, fall short, especially in complex environments. However, learning SFs from scratch risks representation collapse and poor performance due to missed key features. To address this, two strategies have been explored: pre-training, unsuitable for continual RL scenarios due to out-of-distribution challenges, and auxiliary losses, including constraints like reconstruction on basis features. While the latter may reduce mismatches, it introduces inductive biases that hamper learning, as shown by our experiments. To tackle this, we devised a novel method that streamlines learning SFs from pixels, using a simplified modification of Temporal-difference (TD) loss, thus removing the necessity for pre-training and auxiliary losses. In single-task scenarios within both 2D (Minigrid) and 3D (Miniworld) mazes, our model matches with the standard RL model: Deep Q Network (DQN). More notably, in a continual RL setup involving two tasks, our model exhibits superior transferability when re-encountering tasks, outperforming both DQN and other SF models trained with auxiliary constraints. Additionally, through dimensionality reduction and geospatial color mapping, we visually observed that our SFs effectively capture the statistical structure of the environments. Interestingly, our findings also hint at similarities between our computational model

and hippocampal predictive representations in dynamic contexts. This suggests new avenues for exploring diverse neural functionalities, in particular, how reward-driven representations differ from others, thereby enriching our understanding of neural adaptability and learning.

3-081. Controlling behavioral strategy by constraining dynamics in RNNs

Manuel Molano-Mazon[1] MANUELMOLANOMAZON@GMAIL.COM
Yuxiu Shao[2,3] IVYEROSION@GMAIL.COM
Jaime de la Rocha[1] JROCHAV@RECERCA.CLINIC.CAT
Srdjan Ostojic[4] SRDJAN.OSTOJIC@ENS.FR

[1]IDIBAPS
[2]Beijing Normal University
[3]School of Systems Science
[4]Ecole Normale Superieure

Cognitive tasks can be solved using various strategies, as shown by experiments and computational models. The strategies found by Recurrent Neural Networks (RNNs) to solve a given task vary widely depending on training methods and network characteristics. This highlights the importance of theoretical frameworks that, for a given task, can map and describe these strategies in relation to network dynamics. Here we use low-rank RNNs (lrRNNs) to investigate how constraining the dynamics of networks through their connectivity affects their behavioral strategies in a context-dependent perceptual categorization task. We show that lrRNNs developed diverse strategies depending on their connectivity rank.

We consider a previously studied categorization task in which the stimulus sequence included across-trial sequential correlations. Rats leverage the context information from recent trial history but only after correct trials. This suboptimal behavior, called reset, is observed in all rats (n=56). We found that rank-one lrRNNs (r1) base their decisions solely on current stimuli, ignoring context. Rank-two networks (r2) emulate the reset strategy observed in rats and give away the context information after an error trial. Networks with rank-three (r3) or higher display a more optimal reverse strategy, fully exploiting the task statistics.

Analysis of the networks' latent dynamics revealed that r2 networks represent context using a limit cycle and a fixed point confined to the same plane. In contrast, r3 networks use an additional dimension to create two fixed points orthogonally from the limit cycle. This configuration allows r3 networks to utilize the error signal as a cue that reverses the context-dependent dynamics, and ultimately outperform both r1 and r2.

Our findings demonstrate that varying RNN connectivity systematically produces a range of identifiable strategies, including those found in experiments, that balance between structural simplicity and behavioral complexity.

3-082. Modelling contour integration with corners in a convolutional neural network

Udo Ernst[1,2] UDO@NEURO.UNI-BREMEN.DE
David Rotermund[1] DAVROT@NEURO.UNI-BREMEN.DE
Katharina Korb[1] KKORB@UNI-BREMEN.DE

[1]University of Bremen
[2]Institute for Theoretical Physics

The human visual system effortlessly recognizes complex objects, even when these objects are fragmented or occluded. Since neurons in the early stages of the visual cortex have access to only a small region of the visual field, a major task of the brain is to integrate distributed local information into coherent global percepts. An important paradigm for elucidating the neural mechanisms of visual information integration is contour integration, which requires subjects to connect oriented image segments into curves and outlines of potential objects. While previous work suggests that the principles of contour integration are well understood and can be formulated in terms of two Gestalt rules (co-circularity, co-linearity), recent psychophysical experiments show a different picture: Corner elements allow contour integration even at points of strong angular discontinuity. This was previously thought to be impossible. These findings imply a fundamental rethinking of presumed neural mechanisms of contour integration. Here, we investigate how the visual system might integrate contours with and without corner elements, as well as differences that this might introduce. Instead of designing specific mechanisms, we train networks to find their own solutions using identical stimuli, as in psychophysics. We chose a shallow feedforward convolutional network with a hierarchical structure similar to the visual system (areas V1, V2, V4) and realistic receptive field sizes. When trained only on contours with corners, it quantitatively reproduces human behavior and generalizes well to all other stimulus conditions. This is not the case when the network is trained on classical

contours without corners. Integration can be performed with minimal resources, requiring about 4-5 feature maps in each layer. Our approach not only provides a novel mechanistic explanation for unexplained experimental findings, but also contributes to a long-standing debate by showing that feedforward integration is sufficient for contour integration and does not require recurrent computations.

3-083. Why is everything everywhere? Broad mixing leads to increased reliability in neural representations.

Jeff Johnston[1,2] WJJ@POSTEO.NET
Stefano Fusi[3] SF2237@COLUMBIA.EDU

[1]Columbia University
[2]Zuckerman Institute
[3]Columbia University, Zuckerman Institute

Recent experimental work has shown that representations of sensory and motor variables are broadly intermixed across the whole brain. For instance, activity in the primary visual cortex of the mouse has been shown to correlate with numerous task-irrelevant motor actions. This experimental work has lent further credence to the longstanding idea that a significant fraction of the apparent noise in the brain may actually reflect signals that are related to unobserved sensory, motor, and cognitive variables. To understand the advantages of such broad mixing, we compare two alternative theories for how all these different variables could be represented. In modular coding, each neuron represents only a single variable – so, the population is split into distinct subpopulations for each variable. In superposition coding, each neuron represents every variable, with randomly centered receptive fields for each. By construction, the different variables represented in superposition coding interfere with each other; however, despite this interference, we show that superposition coding is ultimately more reliable – that is, it has a lower mean squared reconstruction error for a maximum likelihood decoder. Next, we show that superposition coding is orders of magnitude more reliable when only random subsets of the variables are present at a given time. The response rescaling exhibited by superposition codes in this context leads to a quenching in response variance after the onset of a new stimulus, consistent with experimental observations. Finally, we show that the relative magnitudes of relevant and irrelevant noise under superposition coding are consistent with the pattern observed in population recordings from mouse V1, while the pattern predicted by modular coding is not. Overall, this work shows that widespread mixing between representations of different variables provides both more reliable and more flexible neural computation than representations that are strictly segregated into either subpopulations or subspaces of population activity.

3-084. Investigating mechanisms of visual cortical involvement in working memory using stochastic resonance

Noa Krause[1,2] NOA.NOELLE.KRAUSE@GMAIL.COM
Rosanne Rademaker[1] ROSANNE.RADEMAKER@GMAIL.COM

[1]Ernst Strungmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society
[2]Rademaker Lab

The sensory recruitment hypothesis posits that sensory circuits, specialized for processing relevant sensory information, are recruited by anterior cortical regions for the retention of high-precision working memory (WM). Support for sensory recruitment comes from the ability to decode mnemonic contents from patterns of BOLD responses in early visual cortex (EVC; Harrison & Tong, 2009, Nature). However, such delay-period decoding is challenging to detect with magneto-/electroencephalography (MEEG; e.g., Wolff et al., 2017, Nat. Neurosci.). Presumably, this discrepancy emerges because feedback-based WM recruitment of EVC is spike-silent, changing only synaptic thresholds. Resultant modulations of the local field potential might give rise to decodable BOLD patterns (Mendoza-Halliday et al., 2014, Nat. Neurosci.), but do not leave the electrophysiological trace needed for decoding from the MEEG signal. In this study, we aim to make spike-silent mnemonic traces detectable in MEG by exploiting stochastic resonance. Stochastic resonance refers to a noise-benefit in nonlinear systems (e.g., a neuron), where sub-threshold signals can be made supra-threshold (i.e., translated to firing) by injecting a small "resonant" amount of noise (Simonotto et al., 1997, Phys. Review Letters). To test this, we asked participants to remember orientation stimuli while continuously activating their visual cortex with noise images of varying intensity over a 2s delay. We expect to find a resonant noise intensity that provides sustained decoding of WM contents. Indeed, in preliminary data, we observe a trend towards better delay-period decoding when presenting noise of intermediate intensities. With more data, we hope to reconcile MEEG and fMRI findings on sensory recruitment, present a viable method to improve readout of WM information using MEEG, and elucidate WM mechanisms in sensory cortex.

3-085. The ups and downs of visuo-tactile processing in the mouse cortex

Sami El-Boustani[1,2] SAMI.EL-BOUSTANI@UNIGE.CH
Giulio Matteucci[1] GIULIO.MATTEUCCI@UNIGE.CH
Maelle Guyoton[1] MAELLE.GUYOTON@UNIGE.CH
Charlie Foucher[1] GRIHMALKIN.FOUCHER@ETU.UNIGE.CH

[1]University of Geneva
[2]Department of Basic Neurosciences

Flexible and effective behaviors in natural environments rely on the ability to generalize sensorimotor associations across different sensory modalities. The spatial congruency of stimulation across different modalities can be a significant driver of such cross-modal generalizations. In this study, we investigated whether and how mice can transfer learned spatial sensorimotor associations between touch and vision. We found that mice trained to discriminate between stimulations of two vertically aligned whiskers could seamlessly switch to discriminating between two vertically arranged visual cues, but only when the reward contingencies remained spatially congruent across modalities (i.e., if the top or bottom stimulus continued to be the rewarded one after the switch). Through multi-scale calcium imaging over the dorsal cortex, we identified two distinct associative domains within the ventral and dorsal streams that exhibited signatures of visuo-tactile integration. Within these visuo-tactile areas, we observed spatially congruent multimodal responses potentially enabling cross-modal transfer learning. Additionally, we confirmed the topographically ordered convergence of visual and tactile inputs to/from these regions by tracing feedforward and feedback projections to/from primary sensory areas. Finally, we employed a loss-of-function approach to causally assess the involvement of these visuo-tactile associative areas in cross-modal transfer. By locally expressing tetanus toxin light chain (TeNT), which enables the long-term silencing of synaptic transmission, we demonstrated the necessity of the dorsal stream areas during training for successful transfer. Our results highlight a pivotal cortical pathway essential for visuo-tactile multisensory integration and goal-directed cross-modal transfer learning. Moreover, our findings lay the groundwork for developing computational models of transfer learning, where overlapping sensory maps of stimuli features across modalities are crucial for enabling rapid knowledge transfer across the senses.

3-086. Beyond Individuals: Comparing spontaneous whole-brain dynamics across zebrafish larvae

Matteo Dommanget-Kott[1,2] MATTEO.DOMMANGET-KOTT@SORBONNE-UNIVERSITE.FR
Georges Debregeas[3] GEORGES.DEBREGEAS@SORBONNE-UNIVERSITE.FR
Volker Bormuth[3] VOLKER.BORMUTH@SORBONNE-UNIVERSITE.FR
Jorge Fernandez-de-Cossio-Diaz[4] JORGE.COSSIO@PHYS.ENS.FR

[1]Sorbonne Universite
[2]Laboratoire Jean Perrin
[3]Sorbonne University
[4]ecole Normale Superieure

Comparing neuronal activity across individuals is experimentally challenging, despite animals of the same species having stereotypical behaviors. Functional MRI allows one to compare coarse-grained brain activity, but this technique lacks the resolution to understand the emergence of whole-brain dynamics from neurons. Small animals like Zebrafish larvae offer the possibility of whole-brain single-neuron resolution imaging, but current methods for cross-individual comparisons are limited to experimental paradigms in which activity is sensory driven. Thus, analyzing the spontaneous brain activity across Zebrafish larvae remains elusive. Introducing a novel approach, we utilize a single Restricted Boltzmann Machine (RBM) to model neuronal statistics from multiple Zebrafish larvae, addressing challenges in comparing spontaneous, neuron-level, whole-brain activity. Our method extends previous work where RBMs were used to model whole-brain neuronal statistics in single Zebrafish larvae. In this novel approach, a single RBM captures the neuronal activity of multiple larvae, projecting it to a shared latent space. We introduce two methods: one using voxelized brain activity, showing conserved pairwise correlations between hidden units despite variable voxel-voxel correlations across individuals; the other, working at the single-neuron scale, involves training an RBM on one fish and adapting it to others. This provides a neuron-to-cell-assemblies scale transition common to multiple individuals. Adding a classification layer enables the assignment of brain states to each recording time point. Markovian transition probabilities between states are partially conserved between individuals, indicative of our method's ability to capture stereotypical Zebrafish brain dynamics. Our approach combines spontaneous neuronal activity from multiple Zebrafish larvae into a shared and interpretable latent space, opening avenues for a generic model of Zebrafish neuronal dynamics. It also provides a tool to shorten RBM training time, offering practical applications for online experiments.

3-087. Novelty detection by density estimation in the fruit fly olfactory circuit

Kathryn Simone[1,2] KPSIMONE@UWATERLOO.CA
P. Michael Furlong[3] MICHAEL.FURLONG@UWATERLOO.CA
Jeff Orchard[3] JORCHARD@UWATERLOO.CA
Terrence C. Stewart[4] TERRENCE.STEWART@NRC-CNRC.GC.CA

[1]University of waterloo
[2]Cheriton School of Computer Science
[3]University of Waterloo
[4]National Research Council of Canada

The capacity to recognize an unfamiliar situation is essential for survival. The fruit fly olfactory novelty detection circuit (Aso et al., 2014) affords a tractable model for discerning computational principles of this ubiquitous ability, and, as this circuit has been extensively characterized, theorists have well-understood constraints to inform development of biologically-plausible algorithms. Stewart et al. (2023) hypothesize that in insects, synapses between mushroom body Kenyon Cells and the MBON α'3 neuron assess probability of an input given some history, by means of an intermediate projection to a high-dimensional vector space. However, it is unclear whether this algorithm would work when subject to the extreme constraints of the brain of the fruit fly specifically, one of the most widely-used animal models in neuroscience. To address this uncertainty, here we characterize a variant of the model with size, sparsity, and connectivity corresponding to the Drosophila melanogaster circuit. We observe that incorporating these biological details reproduces aspects of Kenyon Cell response patterns as seen in actual recordings. Moreover, there is a monotonic relationship between probability and network output, which would support its application as a novelty detector. Characterizing model performance on a novelty detection task as a function of network size and sparsity, we find that the fruit fly circuit model acts as an ideal classifier, as assessed with the area under the curve of the receiver operator characteristic. These findings simultaneously uphold the interpretation of the fruit fly mushroom body as performing novelty detection via probability estimation and invites the application of the approach to other species, towards a deeper understanding of novelty detection across the Arthropod phylum.

3-088. Dimensionality of familiarity spectrum in medial prefrontal cortex representations

Meghan Cum[1,2] MCUM@UFL.EDU
Sequioa Smith[1] SEQUIOASMITH@UFL.EDU
Aidan Higgs[1] HIGGSA@UFL.EDU
Albert Li[1] ALBERT.LI@UFL.EDU
Ryo Iwata[1,2] RYOI360@UFL.EDU
Elizabeth Illescas-Huerta[1] E.ILLESCASHUERTA@UFL.EDU
Nancy Padilla-Coreano[1] NPADILLACOREANO@UFL.EDU

[1]University of Florida
[2]Neuroscience

Long-term relationships are essential for all social species. Our understanding of long-term familiarity is limited because previous studies have focused on short-term social memories of individuals with little familiarity. The medial prefrontal cortex (mPFC) plays a major role in social cognition. Recent work shows that mPFC activity is necessary for long-term but not short-term social memory, suggesting that the mPFC could encode short- and long-term familiarity differently. Previous work in the hippocampus showed that littermates were represented with higher dimensionality than novel individuals3, providing a theoretical framework to test in the mPFC. We hypothesize that the same is true in the mPFC and that extended familiarity increases the dimensionality of representation in mPFC. To test this, cells were recorded in the mPFC in mice during social interactions with agents of different familiarity levels: novel, short-term familiar (10 min of prior exposure), and long-term familiar (cagemate). We found that mPFC population dynamics during interactions with the different agents were separable and consistent in low dimensional space across two distinct assays. A classifier decoded familiarity level from neural activity during investigation bouts and long-term familiar agents showed strongest encoding. The dimensionality needed for identity decoding varied with familiarity levels. As hypothesized, long-term familiarity decoding improved with more dimensions while novelty decoding plateaued. Decoders trained to distinguish two agents did the best for novel vs long-term individuals when mice interacted with one agent at a time. Short-term vs long-term decoding had higher accuracy than predictions of novel vs short-term familiar agents, suggesting that mPFC encoding of short-term familiar and novel agents is more similar than its encoding for short-term and long-term familiarity. As we begin to disentangle the transition from novelty to familiarity, our results show that the mPFC has distinct encoding for different familiarity levels and potentially encoding familiarity as a spectrum.

3-089. Strong but not weak noise correlations benefit coding in sensory systems

Gabriel Mahuas[1] MAHUAS.GABRIEL@GMAIL.COM
Olivier Marre[2] OLIVIER.MARRE@GMAIL.COM
Ulisse Ferrari[3] ULISSE.FERRARI@GMAIL.COM
Thierry Mora[4] THIERRY.MORA@GMAIL.COM

[1]Laboratoire de Physique de l'ecole Normale Superieure
[2]Sorbonne Universite, Institut de la Vision
[3]Institut de la Vision, Sorbonne Universite, INSERM, CNRS
[4]Laboratoire de Physique de l'ecole Normale Superieure

Sensory neurons encode information about incoming stimuli in their collective activity. This population activity is shaped by two main sources of correlations: first, stimuli tend to contain strong spatiotemporal correlations that correlate neural activity. Second, neurons receive shared noise from upstream cells and interact with each other such that neural noise is correlated across the network. While a persistent topic in neuroscience, the impact of noise correlations on the encoding of stimulus information remains a major open question. Noise correlations are thought to be beneficial for stimulus encoding whenever they are opposed to stimulus-induced correlations (an effect known as the "sign-rule"). However, this is at odds with the structure of noise correlations observed in many sensory systems, where positively noise correlated cells also tend to be positively stimulus correlated. We built a concise analytical picture revealing how correlations affect stimulus encoding by using an expansion of mutual information between stimulus and response in terms of means and correlations of the population activity. We found that noise correlations are not only beneficial when following the "sign-rule", but also when they are larger than a critical value. This approach can be easily applied to neural data to quantify the information carried by neural responses about stimuli. Applying our approach to retinal electrophysiological recordings, we validated our prediction and found that strong and positive noise correlations are beneficial to mutual information and enhance the encoding of incoming stimuli. These beneficial effects arise because short ranged noise correlations favor fine details of the stimulus to the detriment of long ranged features, which are however already well encoded by the network. Our approach uncovers and characterizes a new biologically relevant regime where noise correlations are beneficial to stimulus information encoding, bridging the gap between contradicting theoretical and experimental work from the literature.

3-090. Reconciling optimality in reinforcement learning with suboptimal behavior

Anja Tamara Zai[1] ZAIA@ETHZ.CH
Corinna Lorenz[2] CORINNA@INI.ETHZ.CH
Nicolas Giret[3] NICOLAS.GIRET@UNIVERSITE-PARIS-SACLAY.FR
Richard H.R. Hahnloser[2] RICH@INI.ETHZ.CH

[1]Institute of neuroinformatics
[2]ETH Zurich
[3]Universite Paris Saclay

Reinforcement learning (RL) is a promising computational theory for understanding learned behavior and its underlying neural mechanisms. Yet, most natural behaviors are too diverse to be compatible with optimal action policies that strictly maximize reward. To frame motor skill learning within a flexible but normative RL theory, we postulate that optimality in RL applies not to behavior itself but to a latent source of pure motor variability, from which an animal learns. All other aspects of behavioral variability we attribute to non-ideal variability sources from which an animal cannot directly learn (akin to computational noise). We test this latent RL hypothesis in songbirds subjected to a reinforcement task where they must adapt the pitch (fundamental frequency) of their song to escape an aversive reinforcer. Our model agrees with a wealth of published data and produces excellent fits to new behavioral data even when learning trends defy optimality due to circadian fluctuations. When we lesion the cortical output area of the pitch-learning circuit, we find an excellent linear relationship between lesion volume and reduction of variability attributed to the latent learner. Estimation of learner variance is possible even from spontaneous unreinforced behavior, which provides a convenient and non-invasive access to the function of this neural circuit. Our work suggests that, as an evolutionary driving force, learning from randomized trials is more important than achieving behavioral optimality.

3-091. Selective amplification of recurrent subnetworks in the developing visual cortex

Haleigh Mulholland[1] HMULHOLLAND.NIH@GMAIL.COM
Sigrid Tragenap[2] TRAEGENAP@FIAS.UNI-FRANKFURT.DE
Matthias Kaschube[2] KASCHUBE@FIAS.UNI-FRANKFURT.DE
Gordon Smith[1] GBSMITH@UMN.EDU

[1]University of Minnesota
[2]Frankfurt Institute for Advanced Studies

Prior to visual experience, spontaneous activity in the primary visual cortex forms low-dimensional, modular activity patterns that reveal functional networks with millimeter-scale correlations that are precursors to the functional architecture of mature cortex (1). Computational modeling suggests that these patterns emerge out of short-range intracortical interactions through a recurrent network with heterogenous connectivity (1). In recurrent models, stimulus inputs that align with the structure of these endogenous subnetworks would be recurrently amplified, leading to more reliable evoked responses (2). Thus, the alignment of feedforward inputs onto the existing cortical network structure could be crucial for building and refining stimulus specific representations (3), but whether developing cortical networks selectively amplify aligned inputs has not been empirically tested. Using spatially structured optogenetic stimulation and in vivo calcium imaging in young ferret visual cortex before eye opening, we tested whether input patterns consistent with endogenous networks are selectively amplified relative to artificial inputs. We found that optogenetic stimuli derived from endogenous patterns of spontaneous activity evoked stronger average responses in targeted areas than artificial random patterns made from bandpass white noise. In addition, endogenous stimuli evoked responses that were more reliable from trial-to-trial and were more similar to their input pattern compared to artificial stimuli that had similar spatial frequencies. Furthermore, the degree of reliability and stimulus similarity was predicted by the amount of overlap the stimulus input pattern had with the leading principal components of spontaneous activity. Together, these results indicate that already before eye opening, early cortical activity is organized into preferentially connected millimeter-scale subnetworks that recurrently amplify inputs that align with the endogenous network architecture. Over development, these amplified responses could become preferentially reinforced through activity-dependent plasticity mechanisms, providing insight into how reliable responses to sensory input might be built over the course of early sensory experience.

3-092. Weak behavior supervision for latent dynamics is all you need to capture motor corrections

Nina Kudryashova[1,2] NKUDRYAS@ED.AC.UK
Cole Hurwitz[3] CH3676@COLUMBIA.EDU
Robyn Greene[1] ROBYN.GREENE@ED.AC.UK
Matthias Hennig[1] MHENNIG@ED.AC.UK

[1]University of Edinburgh
[2]Informatics
[3]Columbia University

Neural activity in motor cortical areas is well-explained by latent neural population dynamics: the motor preparation phase sets the initial condition for the movement while the dynamics that unfold in the motor execution phase orchestrate the sequence of muscle activations. This preparatory activity explains a large fraction of both neural and behavior variability during the execution of a planned movement [Churchland, 2006]. Yet, accounting for unplanned corrections during the movement requires inferring inputs from other brain areas.

Here we provide evidence that these inputs can cause a small transient deviation from an autonomous neural population trajectory that can be missed by unsupervised inference methods. We propose Behavior-Aligned Neural Dynamics (BAND) model: a latent dynamics model weakly supervised with behavior. We show that BAND achieves consistently higher behavior reconstruction in monkey reaching tasks from the Neural Latents Benchmark (NLB) [Pei, 2021], while neural reconstruction remains similar to the unsupervised baseline (LFADS [Pandarinath, 2018]). These results suggest that additional behavioral variability captured with BAND is encoded in a relatively small neural variability.

To reveal the neural correlate of unplanned behavioral variability, we analyzed neural recordings from monkeys performing a center-out reaching task with a force field perturbation. We found that an unsupervised baseline (LFADS) captures instructed reach direction, but can not explain the uninstructed behavioral variability related to online motor corrections. In contrast to LFADS, a recently proposed supervised embedding method CEBRA [Schneider, 2023] accounted for some motor corrections, yet struggled to decode the overall reaching trajectory. Finally, BAND accounts for both the planned movement towards the instructed target and the motor correction in perturbed trials.

These results show that 1) transient motor corrections are encoded in small neural variability; 2) combining latent dynamical modeling with weak behavior supervision allows for capturing both the movement plan and corrections.

3-093. The neuromechanical basis for goal-directed antennal grooming through multiple body part coordination

Pembe Gizem Ozdil[1] PEMBE.OZDIL@EPFL.CH
Pavan Ramdya[2] PAVAN.RAMDYA@EPFL.CH
Auke Ijspeert[1] AUKE.IJSPEERT@EPFL.CH

[1]Swiss Federal Institute of Technology Lausanne (EPFL)
[2]Ecole Polytechnique Federale de Lausanne

Goal-directed movements, such as self-grooming, are ubiquitous across limbed animals. However, it remains poorly understood how these movements arise from an interplay between sensory feedback, central processing, and musculoskeletal dynamics. The adult fly, Drosophila melanogaster, performs goal-directed reaching during antennal grooming. Although some neural elements of the antennal grooming circuit have been identified[1], we lack a clear and comprehensive picture of how extensive brain networks coordinate multiple body parts during grooming.

Here, we uncover neuromechanical mechanisms for multi-body part coordination during antennal grooming by combining behavioral experiments, 3D kinematic analyses, data-driven neuromechanical simulations, and connectome-constrained neural network modeling. We found that antennal grooming contains coordinated head, antennae, and foreleg movements. When grooming one antenna, the fly shifts its forelegs toward the targeted antenna, lifts the contralateral antenna, dips and rotates its head to lower the ipsilateral antenna. Notably, this complex coordination pattern persists without the forelegs, indicating independent body part control. By simulating the perturbed kinematics[2] in a physics-based fly simulation[3], we then investigated how attenuating each body-part movement affects grooming. We found that each body part movement uniquely contributes to the efficacy of antennal cleaning. For example, attenuating the head pitch decreases the amount of tibial contact with the antenna while increasing that of the tarsal. This impairs antennal cleaning because the tarsus is less stiff than the tibia. We next constructed a comprehensive antennal grooming network containing the neck, and antennal motor neurons using the adult fly brain connectome[4]. Based on the connectivity of this network, we built an artificial neural network[5] with inputs mimicking what a real animal experiences during an experiment. By silencing different sets of artificial neurons, we found that interneurons with high betweenness scores are crucial for orchestrating body part movements, providing insights for future neural silencing experiments in real animals.

3-094. Latent circuit models reveal line attractor dynamics across visual cortical areas

Mitra Javadzadeh[1] MITRA.NO.17@UCL.AC.UK
Marine Schimel[2,3] MARINE.SCHIMEL@HOTMAIL.FR
Sonja Hofer[1] S.HOFER@UCL.AC.UK
Yashar Ahmadian[2,4] YA311@CAM.AC.UK
Guillaume Hennequin[2] G.HENNEQUIN@ENG.CAM.AC.UK

[1]Sainsbury Wellcome Centre, UCL
[2]University of Cambridge
[3]Department of Engineering
[4]Engineering

Sensory perception involves many brain areas whose dynamic interactions remain poorly understood. Recent technological advances have facilitated (i) the data-driven identification of nonlinear dynamics from multi-area population recordings during behaviour and/or optogenetic perturbations, and (ii) the anatomical and physiological circuit dissection required to appropriately constrain multi-area models. Here, we show that these two powerful approaches can be integrated to infer dynamical motifs from neural recordings and relate them to specific circuit features. Specifically, we propose to construct “latent circuit models” (LCMs) which are structurally constrained by known physiology and whose parameters are directly learned from neural recordings. We apply this approach to the joint dynamics of two mouse visual areas, V1 and LM, with a LCM in which the two areas are reciprocally connected and have local excitation-inhibition (E- I) structure. We use iLQR-VAE, a nonlinear system identification method, to learn the model parameters from simultaneous population recordings in both areas during a visually-guided go/no-go task. Sensory stimulation was occasionally paired with photoinhibition of either V1 or LM. Because the exact time course of sensory- or photoinhibition-driven external inputs to the latent circuit is unknown, we constrain them only loosely via statistical priors reflecting the temporal structure of each trial and

let iLQR-VAE infer their details. Our model correctly infers the presence and step-like nature of laser stimulation during photoinhibition, and assigns an E/I identity to each recorded neuron in a way that appears consistent with their spike waveforms. For each mouse, the learned dynamics are organized around a line attractor that enables persistent sensory responses despite the transience of the inferred sensory inputs. This attractor does not exist in each area separately but arises from the long-range V1<->LM connections. Overall, our work suggests a testable cross-area mechanism underlying the maintenance of sensory information in visual cortex.

3-095. Mice dynamically adapt to opponents in multiplayer games

Chunyu A. Duan[1,2] ANNDUAN2@GMAIL.COM
Ivana Orsolic[1,2] I.ORSOLIC@UCL.AC.UK
Qianbo Yin[1] GRAYSON.YIN@UCL.AC.UK
Mehul Rastogi[3] M.RASTOGI@UCL.AC.UK
Tom Hagley[1] TOM.HAGLEY.16@UCL.AC.UK
Bruno Cruz[4,5] BRUNO.CRUZ@ALLENINSTITUTE.ORG
Andre Almeida[6] A.ALMEIDA@NEUROGEARS.ORG
Jeffrey Erlich[1,2] J.ERLICH@UCL.AC.UK

[1]University College London
[2]Sainsbury Wellcome Centre
[3]Sainsbury Wellcome Centre, UCL
[4]Allen Institute
[5]Neural Dynamics
[6]Neurogears

Competing for resources in a dynamic and social environment is a fundamental decision process. Reward maximisation in this environment requires animals to continuously integrate information of 'self' and 'other', and rapidly adapt their strategies. To study the behavioural and computational mechanisms underlying this decision process, we developed a novel foraging arena, the Octagon, where freely-moving mice harvest reward from eight possible reward patches in a solo or competitive social context. On each trial, visual stimuli are presented to indicate the location and value of two active reward patches, each associated with either a high or low reward. In the solo context, mice gradually developed a value-based preference for the high reward patch that generalised across patch locations. We then introduced competitive social sessions, where two 'solo'-trained mice foraged together, but only the first mouse to report its choice would be rewarded. As a result, the expected values of the two reward options dynamically changed on a trial by trial basis, modulated by the ongoing behaviours of 'self' and 'other'. We compared animals' strategies in solo versus social conditions, and found that mice shifted their preference towards safer, low-payout options when competing against another mouse. Behavioural analyses and reinforcement learning models revealed that this preference shift was reward-maximising; depended on the position and speed of their opponent; and could not be simply explained by learning the changing reward statistics in the social context. Ongoing bistable attractor modelling of patch selection suggests plausible biophysical mechanisms for how opponent information can be used to shift attractor dynamics and guide choices in a multi-agent decision context; generating testable predictions for future experimental investigation. Together, we propose new behavioural and computational frameworks suited for quantitative and systematic characterisation of value-based decision-making in a fast-changing, social world.

3-096. A systematic approach to unravel causal interactions in large neural systems and resting-state human brain networks with reservoir computing

Joan Falco-Roget[1] J.ROGET@SANOSCIENCE.ORG
Adrian I. Onicas[1] A.ONICAS@SANOSCIENCE.ORG
Felix Akwasi-Sarpong[2] FSARPONG@AIMS.EDU.GH
Alessandro Crimi[1] A.CRIMI@SANOSCIENCE.ORG

[1]Sano - Centre for Computational Personalised Medicine
[2]African Institute for Mathematical Science

Accurate assessment of causal interactions within large neural systems remains an open problem despite substantial efforts. Recent proposals offer an interesting compromise between specific model-based and simple model-free approaches by exploiting the state space of the system. In cross convergence mapping, causality is defined based on the state space properties of the system. Reservoir computing networks partially automatize the process thus reducing some of its arbitrary heuristics. However, its application to neural systems faces practical challenges that need to be addressed. In this work, we revisited the sufficient conditions to establish directed

interactions within this reservoir computing framework. We defined a reduced set of scores encoding causal evidence between pairs of variables based on the ability of the reservoir networks to find homeomorphic mappings between their state spaces. We validated our solution in small-sized in-silico systems with known interactions. Subsequently, we studied causal interactions present in networks of the human brain. We found significant asymmetry between anterior and posterior interactions suggesting a unidirectional flow of information from posterior and parietal areas to frontal regions of the brain. Notably, we did not observe a strong dependence on the length of the time series, which is beneficial considering the high cost of data acquisition and the scarcity of longer-duration recordings. Lastly, we note how the usage of reservoir computing causality seems to be particularly suited to neural systems where growing evidence suggests that their irregular and nearly chaotic dynamics - both in spiking and brain networks - present a rather regular and stable state space structure. Importantly, with our systematic approach, we were able to surpass the limitations that large-scale systems imposed upon earlier attempts, thus paving the way to further analyses in several subfields of computational and systems neuroscience.

3-097. A population geometry view of hippocampal remapping

Guillermo Martin-Sanchez[1,2] GUILLERMO.MARTIN@RESEARCH.FCHAMPALIMAUD.ORG
Christian Machens[3] CHRISTIAN.MACHENS@NEURO.FCHAMPALIMAUD.ORG
William Podlaski[1,2] WILLIAM.PODLASKI@RESEARCH.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Neuroscience Programme
[3]Champalimaud Research

Place cells in the hippocampus are thought to represent space, with localized, spatial tuning curves that remap across different environments or contexts. However, recent work has shown that the firing of place cells also depends on various sensory and cognitive variables, providing evidence that hippocampal population activity includes both spatial and non-spatial information. Such findings suggest that remapping could be partially or even fully reinterpreted as an effect of place cell mixed-selectivity. To bridge this gap, we study the problem of remapping from the point of view of neural population geometry, rather than single-cell tuning. Specifically, we consider that hippocampal population activity is confined to a low-dimensional subspace which is shared across different contexts, and from which spatial position can be linearly decoded. We then explore two competing perspectives on remapping in this common framework. The first is analogous to the classical view and assumes a context-dependent change in the neural representation of space, requiring a different decoder for position in each context. The second explains remapping through context-dependent changes in the sensory and cognitive variables encoded by mixed-selective cells, which allows for a universal decoder across contexts. We show that both models are consistent with remapping, but that they can be differentiated through tuning properties across contexts. Finally, we confirm our theoretical results using a mechanistic spiking network model that exhibits low-dimensional population activity, localized place fields, and remapping. Intriguingly, our network model also accounts for remapping induced by the experimental inhibition of place cells, arguing that remapping can also occur as a consequence of compensation in a network with redundancy. Overall, our work unifies classic models of remapping with more recent cognitive and latent state space models, illustrating their differences, and suggesting ways to arbitrate between them.

3-098. The impact of persistent accumbal dopamine transients on the preference between natural and “drug-like” reward

Laurena Python LAURENA.PYTHON@UNIGE.CH
Alex Pouget ALEXANDRE.POUGET@UNIGE.CH
Vincent Pascoli VINCENT.PASCOLI@UNIGE.CH
Agnes Hiver AGNES.HIVER@UNIGE.CH
Christian Luscher CHRISTIAN.LUSCHER@UNIGE.CH

University of Geneva

Drug-choice at the expense of natural reward has been suggested to be a hallmark of addiction. It has been hypothesised that a bias towards drug rewards might be due to the pharmacological power of the drugs on the dopamine (DA) system, which results in an overvaluation of the reward to come. However, these influential theories have not been supported experimentally. Here, we suggest an alternative hypothesis: the value update carried out by the DA system is not done based on a constant learning rate. More precisely, we propose an implementation of the temporal difference learning rule where the learning rate decreases as a function of the amount of DA recently released into the nucleus accumbens (NAc). The decrease in learning rate could contribute to the lack of behavioral flexibility when a risk of punishment is introduced, hence playing a role in an important symptom

of drug addiction: compulsivity. This hypothesis is based on DA recordings in mice during the learning phase of an operant choice task between a natural (food) and an artificial "drug-like" reward (optogenetic dopamine neuron self-stimulation oDASS). These data show that the choices are not necessarily biased towards the "drug-like" reward in this task, and that a stronger decrease in the inferred learning rate is linked to more persistent reward seeking when oDASS is paired with a risk of foot shock. This research contradicts the current explanation for the addictive nature of drugs and suggests that the impact of dopamine on value learning is not constant. A model taking into account a possible homeostatic regulation in the efficiency of DA in the NAc can explain these surprising data and could be predictive of compulsivity in mice.

3-099. Gradient-based methods for spiking physical systems

Julian Goeltz[1,2] JULIAN.GOELTZ@KIP.UNI-HEIDELBERG.DE
Sebastian Billaudelle SEBASTIAN.BILLAUDELLE@KIP.UNI-HEIDELBERG.DE
Laura Kriener[3,4] LAURA.KRIENER@UNIBE.CH
Luca Blessing LUCA.BLESSING@KIP.UNI-HEIDELBERG.DE
Christian Pehle CHRISTIAN.PEHLE@KIP.UNI-HEIDELBERG.DE
Eric Muller MUELLER@KIP.UNI-HEIDELBERG.DE
Johannes Schemmel SCHEMMEL@KIP.UNI-HEIDELBERG.DE
Mihai Petrovici[3] MIHAI.PETROVICI@UNIBE.CH

[1]Heidelberg University & University of Bern
[2]Kirchhoff-Institute for Physics & Department of Physiology
[3]University of Bern
[4]Institute for Physiology

Recent efforts have fostered significant progress towards deep learning in spiking networks, which accelerated the research in the field of neuromorphic computing. Novel substrates, especially physical devices, promise high energy efficiency and computation speed but come with the disadvantage of reduced controllability, precision and reproducibility. Here, we contrast several different methods to train spiking neural networks (SNNs), including a comparison of the results on the mixed-signal neuromorphic platform BrainScaleS-2. Physical computation directly exploits the intrinsic dynamics of a given substrate to efficiently process and propagate information. In contrast to numerical computers, physical computers implicitly obey the dynamics required by certain models of information processing (e.g., neuronal integration) rather than calculating the dynamics explicitly by arithmetically manipulating binary representations thereof. Neuromorphic computers represent a prominent class of physical systems, drawing inspiration from the nervous system by mimicking the dynamics of neurons and synapses. Due to their massively parallel and time-continuous implementation, as well as an asynchronous event-based propagation of signals, they can emulate SNNs in a highly efficient and fast manner. For adjusting an emulated SNNs to a task, we focus on gradient-based optimization schemes, which can be grouped into two categories: surrogate and exact gradients. The former utilize a surrogate for the derivative of the neuronal activation to approximate the gradients and directly include the membrane voltage in the calculation. In contrast, the second method uses the differentiability of the timing of a spike, thereby computing exact gradients which are also sparse in time. After a detailed comparison of these approaches, we discuss their application on the mixed-signal neuromorphic system BrainScaleS-2. Our results have implications for all currently-used gradient-based training methods of SNNs and highlight important characteristics of the algorithms depending on the used substrate.

3-100. Context-dependent Nonlinear Classification of Neural Representations

Francesca Mignacco[1,2] FMIGNACCO@PRINCETON.EDU
Chi-Ning Chou[3] CCHOU@FLATIRONINSTITUTE.ORG
SueYeon Chung[4,5] SCHUNG@FLATIRONINSTITUTE.ORG

[1]CUNY Graduate Center
[2]Center for the Physics of Biological Function
[3]Flatiron Institute
[4]New York University; Flatiron Institute
[5]Center for Computational Neuroscience

Understanding the neural population code that underlies efficient representations is crucial for neuroscience and machine learning. Approaches focused on the geometry of task structures in neural population activities have recently emerged as a promising direction for understanding information processing in the brain. In particular, analytical advances linking the geometry of representations to the capacity of downstream readout have shown a promise as a normative theory and data analysis tool, providing a pathway for explicitly connecting the structure of neural representations and emergent information. However, these approaches have so far been limited to

the capacity of linear decoders. Therefore, interpreting nonlinear separation of neural manifolds poses key challenges, such as capturing representation reformatting in early sensory processing stages before the emergence of linearly separable information. We develop a classification theory of neural manifolds for non-linear separations, particularly focused on context-dependence, a ubiquitous phenomenon in neural systems, from circuit-level mechanisms to behavior. We explore a prototypical model that derives its decision-making rules from a collection of input-dependent neurons, associated with distinct contexts through half-space gating based on a set of random or pre-trained hyperplanes in representation space. In this setting, the non-linearity of a task can be quantified by the number of contexts required for separability - an upper bound on the number of piece-wise linear components of the decision boundary. First, we derive an analytic expression for the context-dependent capacity. This formula elucidates the interplay between stimulus statistics and correlations within context hyperplanes, showing that maximal capacity is achieved for orthogonal contexts. Then, we explore the structure of memorization across layers of deep neural networks by estimating numerically the nonlinear capacity. The increased expressivity of our framework captures quantitatively that the layer hierarchy progressively untangles representations. Our method is data-driven and widely applicable across datasets and models.

3-101. Sufficient conditions for offline reactivation in recurrent neural networks

Nanda H Krishna[1] NANDA.HARISHANKAR@GMAIL.COM
Colin Bredenberg[2] COLIN.BREDENBERG@MILA.QUEBEC
Daniel Levenstein[3] DANIEL.LEVENSTEIN@MILA.QUEBEC
Blake Richards[3] BLAKE.RICHARDS@MILA.QUEBEC
Guillaume Lajoie[2,4] G.LAJOIE@UMONTREAL.CA

[1]Universite de Montreal / Mila – Quebec AI Institute
[2]Universite de Montreal
[3]McGill University
[4]Mila

During periods of quiescence, such as sleep, neural activity in many brain circuits resembles that observed during periods of task engagement. However, the precise conditions under which task-optimized networks can autonomously reactivate the same network states responsible for online behavior are poorly understood. In this study, we develop a mathematical framework that outlines sufficient conditions for the emergence of neural reactivation in circuits that encode features of smoothly varying stimuli. We demonstrate mathematically that noisy recurrent networks optimized to track environmental state variables using change-based sensory information naturally develop denoising dynamics, which, in the absence of input, cause the network to revisit state configurations observed during periods of online activity. We validate our findings using numerical experiments on two canonical neuroscience tasks: spatial position estimation based on self-motion cues, and head direction estimation based on angular velocity cues. Overall, our work provides theoretical support for modeling offline reactivation as an emergent consequence of task optimization in noisy neural circuits.

3-102. Dynamic gating of perceptual flexibility by diverse cortical responses

Tiange Hou[1,2] TH2498@NYU.EDU
Jade Toth[1,3] JACKMTOTH@PITT.EDU
Blake Sidleck[1] BES120@PITT.EDU
Olivia Lombardi[1] OCL4@PITT.EDU
Abraham Eldo[1] ABE24@PITT.EDU
Danyall Saeed[1] DSAEED@SOM.GEISINGER.EDU
Madelyn Kerlin[1] MSK85@PITT.EDU
Xiangjian Zeng[4] XIANGJIZ@ANDREW.CMU.EDU
Priya Agarwal[1] PRIYAGARWAL83@GMAIL.COM
Dylan Leonard[1] DYLAN.LEONARD@UMASSMED.EDU
Luz Andrino[4] LANDRINO@ANDREW.CMU.EDU
Tal Inbar[1] TAI30@PITT.EDU
Michele Insanally[1] MNI@PITT.EDU

[1]University of Pittsburgh School of Medicine
[2]Departments of Otolaryngology, Neurobiology
[3]Otolaryngology, Neurobiology
[4]Carnegie Mellon University

Flexible responses to sensory cues in dynamic environments are essential for adaptive auditory-guided behaviors such as navigation and communication. How do neural circuits flexibly gate sensory information to select appropriate behavioral strategies based on sensory input and context? Auditory neural responses during behavior are diverse, ranging from highly-reliable 'classical' responses (i.e. robust, frequency-tuned cells) to irregular or seemingly random 'non-classically responsive' firing patterns (i.e., nominally non-responsive cells) that fail to demonstrate any significant trial-averaged responses to sensory inputs or other behavioral factors. While classically responsive cells have been extensively studied for decades, the contribution of non-classically responsive cells to behavior has remained underexplored despite their prevalence. Previous work has shown that non-classically responsive cells in auditory cortex (AC) and secondary motor cortex (M2) contain significant stimulus and choice information and encode flexible task rules. Additionally, both classically and non-classically responsive units are essential for asymptotic task performance, however their role during learning is unknown. Recent work presented at Cosyne 2023 demonstrated that non-classically responsive neurons in AC are preferentially recruited during learning. Here, we expand this investigation to explore how task variables are encoded and evolve during learning both at the single-cell and population levels. Surprisingly, single-cell decoding performance was highest for classically responsive cells during early and expert learning phases but significantly decreased during late learning when the greatest gains in behavioral performance occur. Population- level decoding revealed that during late learning mixed ensembles comprised of both classically and non- classically responsive cells encode significantly more task information than homogenous ensembles of either type and emerge as a functional unit critical for learning. Optogenetically silencing inputs from M2 selectively modulated non-classically responsive cells in AC, prevented mixed ensemble recruitment, and impaired reversal learning. Thus, top-down inputs recruit non-classically responsive neurons into diverse ensembles in auditory cortex to enable behavioral flexibility.

3-103. Cortical Column model of Predictive Coding

Kwangjun Lee[1,2] K.LEE@UVA.NL
Cyriel Pennartz[1] C.M.A.PENNARTZ@UVA.NL
Jorge Mejias[1,3] J.F.MEJIAS@UVA.NL

[1]University of Amsterdam
[2]Swammerdam Institute for Life Sciences
[3]Cognitive and Systems Neuroscience

Predictive coding (PC) is an influential theory in neuroscience, which gives rise to a perceptual architecture of the cortex by constantly generating and updating predictive representations of sensory inputs. While its computational goal is clearly defined as prediction error minimization via hierarchical interactions between representation and prediction error neurons in cortical areas, a neural implementation that maps such a program onto cortical columns remains a challenge. In this work, we incorporate anatomical projections [1], laminar organizations [2], and the diversity of neurons [3] in the sensory cortex to the theory of PC and develop a cortical column model of PC (CoCo-PC). This involves converting the error-computing and prediction units of the classic PC network [4] based on point neuron assumption to microcircuits with four major cortical neuron types (pyramidal, PV, SST, and VIP cells), placing them in superficial and deep layers, respectively, and devising connection patterns between them to facilitate prediction error minimization circuits. The model successfully learned various image datasets (MNIST, fashion-MNIST, and grayscale CIFAR-10) via Hebbian learning, as it could reconstruct novel instances. Beyond the learning of image representations, our model captured mismatch negativity signals given a sequence of repetitive stimuli interrupted by a deviant stimulus (i.e., oddball paradigm). Interestingly, we observed oscillatory dynamics emerging in excitatory firing rates for both image reconstruction and oddball paradigms. Our simulated optogenetic silencing experiment suggests that PV cells are important for sensory suppression and generating oscillations and that SST and VIP cells have specialized roles in different prediction error circuits, revealing SST/VIP optogenetic silencing as a reliable procedure to experimentally identify positive and negative prediction errors circuits.

3-104. Layer 1 NDNF interneurons are specialized top-down master regulators of cortical circuits

Jan Hartung[1,2] JAN.HARTUNG@PHYSIOLOGIE.UNI-FREIBURG.DE
Anna Schroeder[1] ANNA.SCHROEDER@PHYSIOLOGIE.UNI-FREIBURG.DE
Rodrigo Alejandro Perez Vazquez[1] RODRIGO.ALPEVA@GMAIL.COM
Rogier Poorthuis[3] R.B.POORTHUIS@UMCUTRECHT.NL
Johannes Letzkus[1] JOHANNES.LETZKUS@PHYSIOLOGIE.UNI-FREIBURG.DE

[1]University of Freiburg
[2]Insitute for Physiology I

[3]Utrecht University

Cortex harbors a diverse range of inhibitory interneuron (IN) types, each of which has specific connectivity patterns and intrinsic properties, enabling it to shape cortical activity in a unique way. Collectively, these INs control and orchestrate neuronal activity, thereby bestowing cortex with its computational power. Whereas genetic markers have greatly facilitated the study of various cortical IN types, a selective marker for INs in L1 has long been missing. Here, we use the recently identified neuron-derived neurotrophic factor (NDNF) to characterize the properties of L1 INs in vitro. First, using in vitro patch-clamp recordings in combination with optogenetics and genetic identification of postsynaptic interneuron types, we study outbound connectivity of L1 NDNF INs and find that they broadly control cortical L2/3, including parvalbumin (PV) and vasoactive intestinal peptide (VIP) expressing interneurons. Second, we describe two major electrophysiological clusters within L1 NDNF INs and validate an intersectional genetic approach to specifically access neuropeptide Y (NPY) expressing L1 NDNF INs using in situ hybridization. Employing this approach, we report very similar electrophysiological properties and connectivity patterns of L1 NDNF and NDNF/NPY INs. Third, we demonstrate that, after sufficient stimulation, L1 NDNF INs are almost uniquely capable of persistent firing (PF), a mode of activity which uncouples outputs from current inputs. PF comes with a transition into a highly excitable electrophysiological state, which is maintained for the duration of stimulation and reverts only after a protracted period of rest. Taken together, the data we provide will be valuable to build models of cortical circuit function and, due to the abundant top-down afferents to NDNF INs, of biologically informed models of predictive coding. In conclusion, in combination with their previously described control of distal pyramidal cell (PC) dendrites and long-range innervation, we propose L1 NDNF INs as top-down master regulators of superficial cortical circuits.

3-105. An Analytical Theory of Multi-Task Representation Learning and Disentanglement

Albert Wakhloo[1,2] AJW2232@CUMC.COLUMBIA.EDU
Will Slatton[3] WSLATTON@FLATIRONINSTITUTE.ORG
SueYeon Chung[3,4] SCHUNG@FLATIRONINSTITUTE.ORG

[1]Columbia University; Flatiron Institute
[2]Center for Theoretical Neuroscience; Center for Computational Neuroscience
[3]New York University; Flatiron Institute
[4]Center for Computational Neuroscience

Humans and animals are able to efficiently learn diverse tasks that depend on common, underlying latent structures. Recent work argues that neural representations enable this ability by forming representations that facilitate efficient downstream decoding across tasks. Several works argue that neural populations accomplish this by separating unrelated latent variables into orthogonal subspaces in the neural representation space–i.e., by forming a disentangled representation. However, in these works, disentanglement is typically assessed using empirical generalization-error-based metrics or using heuristic metrics that have no direct connection to the multi-task learning problem. Here, we develop an analytical theory for learning multiple tasks that depend on a common latent structure. In this way, we analytically describe how disentanglement depends on representational geometry and why these representations are useful for such problems. Using a popular model for generating diverse tasks from a common latent space, we find that the test error on these tasks can be analytically decomposed into three scalar statistics: (1) the dimensionality of the neural representations, (2) the total correlation between neural population responses and latent directions of variation, and (3), the projection of neuronal noise onto the coding directions. Using this theory, we study the properties of neural representations that achieve the optimal error across tasks. We show that optimal neural codes disentangle independent latent variables into orthogonal subspaces. Moreover, at small training budgets, optimal representations compress less informative latent variables in the state space, whereas at large training budgets, optimal representations expand these variables. These features can be traced back to measurable properties of the eigenspectrum of the neural population responses. Finally, we show that our theory accurately predicts the performance of macaque V4 and IT representations on decoding tasks that are formed from latent variables in the stimuli, demonstrating the applicability of the theory.

3-106. Modulation of perceived task difficulty impacts neuronal variability in visual cortex

Patricia Stan PSTAN@ANDREW.CMU.EDU
Matthew A. Smith MATTSMITH@CMU.EDU

Carnegie Mellon University

Differences in global cognitive factors – such as effort, motivation, or arousal – impact behavioral performance on a variety of perceptual tasks. One common source of shifts in these internal cognitive processes of the brain is task difficulty. What changes in neuronal populations can impact perceptual performance in response to task demands? To address this overarching question, we created a task paradigm where monkeys performed a natural scene change detection task under different task contexts (i.e., different changes occurred in the natural scene) while we recorded the activity of populations of neurons in visual cortical area V4. The task contexts varied in relative difficulty with each of the images used (a different image each day), which naturally produced shifts in the animals' behavioral performance within each session and across sessions. In data from two monkeys we found that when the task was more difficult the animals improved their perceptual sensitivity for identical visual changes. Overall, our recordings indicated that better ability to detect a change in an image was associated with decreased firing rates to the image, leading to increased signal separation between the image and the target (i.e. the changed image). Similar to other manipulations of cognitive state such as attention or task switching, dimensionality reduction revealed that improved performance was associated with a decrease in population trial-to-trial shared variability. Together, our results support the notion that shifts in variability early in visual processing are a key common mechanism linked to improved perceptual abilities.

3-107. Modeling the flexibility of cortical control of motor units

William Surmeier[1,2] CLAY-SURMEIER@NORTHWESTERN.EDU
Elom Amematsro[3] EAA2164@CUMC.COLUMBIA.EDU
Najja Marshall[3] NAJJA.MARSHALL@GMAIL.COM
Mark Churchland[3] MC3502@COLUMBIA.EDU
Josh Glaser[1] J-GLASER@NORTHWESTERN.EDU

[1]Northwestern University
[2]Neurology
[3]Columbia University

Motor units, which compose spinal motor neurons and their innervated muscle fibers, are the fundamental unit of motor control. For nearly one hundred years, researchers have studied how MUs are recruited during movement, and the textbook understanding is that a simplified control strategy is used. A single 'common drive' is sent to a muscle's motor neuron pool, which controls the amount of force that muscle will produce. This understanding has drawn primarily from experiments involving subjects generating static or gradually changing forces. However, recent results, using experiments with a much wider range of force profiles, have shown otherwise. To explain motor unit activity across a wide range of movement conditions, multiple degrees of freedom were needed. A large question remains however – what underlies this flexibility of motor unit control? Is it driven by the cortex? Here, to answer that question, we analyzed simultaneous recordings of neurons in primary motor cortex (M1) and motor units (MUs) in a monkey making a wide range of force profiles. We developed a model framework in which MUs were predicted from M1 neurons through a low-dimensional bottleneck, in order to test how many degrees of freedom within M1 were beneficial for predicting MU activity. If M1 is simply sending a 'common drive' to control all the motor units within the motor neuron pool, then a single degree of freedom within M1 (a 1-d bottleneck) would explain the motor unit activity just as well as using more degrees of freedom. However, models with a one-dimensional bottleneck were not able to accurately predict all motor units, particularly those selectively active during high-frequency oscillatory conditions. Multiple dimensions within the bottleneck were beneficial. This strongly suggests cortical flexibility over motor unit control.

3-108. Neural signatures of stress susceptibility and resilience in the amygdala-hippocampal network

Frances Xia[1,2] FRANCES.XIA@UCSF.EDU
Valeria Fascianelli[3] VF2266@COLUMBIA.EDU
Nina Vishwakarma[1] NINA.VISHWAKARMA@UCSF.EDU
Frances Grace Ghinger[1] FGGHINGER@GMAIL.COM
Stefano Fusi[4] SF2237@COLUMBIA.EDU
Mazen Kheirbek[1] MAZEN.KHEIRBEK@UCSF.EDU

[1]University of California, San Francisco
[2]Department of Psychiatry and Behavioral Sciences
[3]Columbia University
[4]Columbia University, Zuckerman Institute

Mood disorders are characterized by behavioral changes arising from network-level alterations, with the baso-

lateral amygdala (BLA)-ventral CA1 (vCA1) pathway being a central component. Anhedonia is a core aspect of major depressive disorder. Traditionally viewed as a blunted emotional state in which individuals are unable to experience joy, anhedonia also diminishes the drive to seek rewards and the ability to value and learn about them. The neural underpinnings of anhedonia and how this emotional state drives related behavioral changes remain unclear. Here, we investigated these questions by taking advantage of the fact that when mice are exposed to traumatic social stress, susceptible animals become socially withdrawn and anhedonic, where they cease to seek high-value rewards, while others remain resilient. By performing high density electrophysiological recordings and comparing neural activity patterns of these groups in the BLA and vCA1 of awake behaving animals, we identified neural signatures of susceptibility and resilience to anhedonia. When animals actively sought rewards, BLA activity in resilient mice showed stronger discrimination between upcoming reward choices. In contrast, susceptible mice displayed a rumination-like signature, where BLA neurons encoded the intention to switch or stay on a previously chosen reward. When animals were at rest, the spontaneous BLA activity of susceptible mice was higher dimensional than in controls, reflecting a greater number of distinct neural population states. Notably, this spontaneous activity allowed us to decode group identity and to infer if a mouse had a history of stress better than behavioral outcomes alone. Finally, targeted manipulation of vCA1 inputs to the BLA in susceptible mice rescued dysfunctional neural dynamics, amplified dynamics associated with resilience, and reversed their anhedonic behavior. This work reveals population-level neural signatures that explain individual differences in responses to traumatic stress, and suggests that modulating vCA1-BLA inputs can enhance resilience by regulating these dynamics.

3-109. The impact of biomechanical actuators on neural embodied control

Eric Leonardis[1] ELEONARDIS@SALK.EDU
Dan Butler[1] DBUTLER@SALK.EDU
Adam Lee[2] A3LEE@UCSD.EDU
Scott Yang[1] YUY004@UCSD.EDU
Diego Aldarondo[2] DIEGOALDARONDO@G.HARVARD.EDU
Bence Olveczky[2] OLVECZKY@FAS.HARVARD.EDU
Eiman Azim[1] EAZIM@SALK.EDU
Talmo Pereira[1] TALMO@SALK.EDU

[1]Salk Institute for Biological Studies
[2]Harvard University

The brain has evolved to effectively control the body, and in order to understand the relationship we need to model the sensorimotor transformations underlying embodied control. As part of a coordinated effort, we are developing a general-purpose platform for behavior-driven simulation modeling high fidelity behavioral dynamics, biomechanics, and neural circuit architectures underlying embodied control. The appropriate level of biomechanical detail required to train realistic simulations, and their associated tradeoffs, are highly debated. A central question examined with this research platform is to investigate the appropriate level of abstraction for biomechanical actuation during complex motor behaviors. Specifically, we investigate the impact of utilizing torque actuators, akin to those found in robots, in contrast to biologically-inspired muscle actuators when training models to control rodent bodies. We implement a mouse model in our VirtualNeuroLab (VNL) framework using each type of actuator to perform a dextrous control task with the rodent forelimb using reinforcement learning. We present results indicating that both types of biomechanical models adhere to observed trade-offs between speed and accuracy in real animal movements, as explained by Fitts' Law (Fitts, 1954). We find differences, however, in the learned neural representations between the torque- and muscle-based actuator systems, suggesting that the neural coding for motor control is dependent on the type of biomechanical actuator used. This work provides early evidence to support the use of faithful biomechanics in embodied NeuroAI modeling.

3-110. Stimulation allows for reshaping network connectivity through plasticity: a training protocol for rate models

Francesco Borra[1,2] FRANCESCO.BORRA@PHYS.ENS.FR
Simona Cocco[3] SIMONA.COCCO@PHYS.ENS.FR
Remi Monasson[3] MONASSON@LPT.ENS.FR

[1]CNRS
[2]Physics
[3]Laboratoire de Physique de l'Ecole Normale Superieure, PSL Research, CNRS

Biological neural system possess remarkable computational features which have inspired artificial neural net-

works. Still, computation with actual biological cells has remained an abstract prospect until recent developments in organoid technologies. The possibility to cultivate, stimulate cells and record their activity with high precision provides interfacing to potentially operate with neural organoids as one would do with miniature computers. Attempts are being made to exploit their computational abilities experimentally but a systematic approach is still missing. Is it possible to compile a generic computational task into a network of biological neurons? This question is extremely challenging since neural connections in biological networks cannot be altered arbitrarily (contrary to artificial neural nets), but change due to neural plasticity featuring both associative and homeostatic effects. Our work directly addresses and solves, under certain assumptions, this question by providing a proof of concept for a flexible framework of task implementation into a neural structure. Our protocol consists in controlling plasticity by inducing a succession of neural activity states through adequate time- and neuron-specific stimulations. This task is hard because plasticity imposes strong constraints (e.g. no adaptive learning rates) and the number of connections to be learned greatly exceeds the number of variables one can control, typically N squared vs. N , where N is the number of neurons. We propose two in silico applications to rate-based networks in stationary conditions. The first task aims at reconnecting a random network into a specific structure (with continuous attractor properties). The second is an input-to-output task, i.e. a subregion in the network (called output) should produce specific images (patterns of activity) when the input region receives specific stimulations.

3-111. Capacity of Networks with Arbitrary Topologies and Neuron Activation Probabilities

Kaining Zhang[1,2] KAINING@WUSTL.EDU
Gaia Tavoni[1] GAIA.TAVONI@WUSTL.EDU

[1]Washington University in Saint Louis
[2]Department of Neuroscience

We use a method pioneered by Gardner to compute the capacity (maximal number of patterns that can be classified or stored) of a general class of perceptrons and Hopfield networks. The method is based on the calculation of the expected volume of the weights that satisfy all the classification/memory constraints: the capacity is maximized when this volume shrinks to zero, i.e., no pattern can be added without breaking the constraints. Here, we extend previous results, which apply to networks with homogeneous connectivity and neuron activation probabilities, to the more general case of (a) perceptrons with arbitrary and heterogeneous unit activation probabilities, and (b) Hopfield models with both heterogeneous unit activation probabilities and arbitrary network topologies (i.e., any arrangements of connections between the units). For perceptrons, we find that the capacity depends on the activation probability of the output unit but is independent of the distribution of input states. For Hopfield networks, we find that the capacity depends on each neuron's activation probability and number of input connections. We use this analytical result to compute the capacity of a two-layer network modeled after the hippocampal-cortical memory system, and to derive a prediction that links connectivity and activation probabilities of the neurons in networks with maximum capacity.

3-112. Improving Temporal Credit Assigment in Recurrent Networks using Dynamical Systems Theory

Rainer Engelken RE2365@COLUMBIA.EDU
Larry F Abbott LFA2103@COLUMBIA.EDU

Columbia University

Discovering and responding to associations between temporally distant cues is crucial for an animal's survival. Solving this temporal credit assignment problem allows organisms to bridge the gap between when a stimulus cue arrives and when its effect unfolds. Gradient-based training of recurrent neural circuit models for temporal tasks with long time horizons presents challenges, potentially leading to vanishing or exploding gradients due to issues with gradients across many time steps. Recent research has connected this issue to the Lyapunov exponents of the forward dynamics, describing how nearby perturbations grow or shrink in forward passes. Here, we propose a novel approach, termed gradient flossing, to address the gradient instability in recurrent spiking and firing rate networks by controlling the Lyapunov exponents of the forward dynamics throughout learning. More specifically, we regularize one or several Lyapunov exponents towards zero, ensuring that the corresponding directions in tangent space grow or shrink only slowly, which means that the information of learning signals are more robustly propagated over long time horizons. We achieve this by backpropagating through the Lyapunov exponents, which enables us to 'floss' the gradients and improve network training. Our method enhances the success rate of RNNs on typical neuroscience tasks involving bridging task events across many time steps, and we demonstrate that applying gradient flossing during training further improves trainability for challenging temporal credit assignment

tasks. We establish a connection between the Lyapunov exponents and the dimensionality of the gradient signal in backpropagation. Additionally, we demonstrate the effectiveness of our approach both on spiking and firing rate networks. Our results suggest that flossing gradients through dynamic control of Lyapunov exponents can significantly improve the stability and effectiveness of RNN training. We speculate on the optimization of neural dynamics in animals over evolutionary timescales to bridge long time horizons.

3-113. Dissecting Local Circuit Mechanisms of Cortical Plasticity in a Multi-Layer Spiking Neural Network

Tea Tompos[1] T.TOMPOS@NEUROPHYSIOLOGY.NL
Fleur Zeldenrust[1] F.ZELDENRUST@NEUROPHYSIOLOGY.NL
Tansu Celikel[2] CELIKEL@GATECH.EDU

[1]Donders Centre for Neuroscience, Radboud University
[2]Georgia Institute of Technology

Sensory cortices have been extensively studied to unravel the mechanisms underlying experience-dependent plasticity, a crucial process by which the brain adapts its neural representations in response to sensory, motor, and perceptual experiences. Despite significant advancements, the anatomical underpinnings of neural representational plasticity remain largely unexplored. To address this critical gap in our understanding, we developed a downscaled (N=350 neurons) multi-layer spiking neural network (SNN) model of the rodent primary somatosensory (S1) cortex. This biologically constrained model, employing Hodgkin-Huxley neurons and conductance-based synapses, accurately simulates realistic neural interactions and network responses. By systematically removing projections (sparsifying the network) in a synapse-type-specific manner, we selectively disrupted neural representations and evaluated the resulting firing-rate stability (homeostasis) at each sparsification stage for each cell type across all layers. Our findings reveal that global alterations in thalamocortical connectivity have larger effects on overall firing rates than projection-specific changes. Furthermore, intralaminar projection sparsification was found to disrupt cortical homeostasis to a greater extent than translaminar sparsification. Intriguingly, the contribution of inhibitory synapses to the disruption of homeostasis was found to be significantly larger than that of excitatory ones. Moreover, similar changes in synaptic strength applied to excitatory synapses resulted in substantially smaller modulations compared to manipulations of inhibitory ones. These findings suggest that the effects of cortical plasticity are layer- and synapse-type specific, proposing a large dynamic range in the mechanisms to adapt to sensory experience. They also provide a framework for designing new experiments to systematically investigate the cellular mechanisms underlying network plasticity.

3-114. Balancing accuracy and diversity : principles of model-driven active sampling in the brain

Hamza Oueld[1] HAMZA.OUELD-KADDOUR-EL-HALLALOUI@UNIV-AMU.FR
Andrea Brovelli[1] ANDREA.BROVELLI@UNIV-AMU.FR
Emmanuel Dauce[2,3] EMMANUEL.DAUCE@CENTRALE-MARSEILLE.FR

[1]CNRS / Aix-Marseille Univ
[2]Ecole Centrale Mediterranee / CNRS
[3]Institut des neurosciences de la Timone

Understanding our environment requires not only passively observing sensory samples but also acting to seek out useful relationships between our actions and their possible outcomes. Inspired by the principle of visual salience, we introduce the concept of an "ideal participant," which is the active counterpart of a Bayesian ideal observer. During learning, an ideal participant would build a statistical model of its environment while updating its policy (action selection) after each new observation. This action selection requires a tight balance between making correct predictions (accuracy) and providing diverse data to feed the models (diversity).

We assess these principles in a knowledge-oriented action selection task called the "volleyball" task, where participants estimate the causal influence of a player on the outcome of a volleyball game. The behavioral data we have collected suggest that such active sampling strategies do occur in the brain and improve the accuracy of action/outcome models. We show that the balance between accuracy and diversity objectives can lead to specific action selection biases, reflected both in the model and in the experiments.

3-115. Organization of mitochondria within a connectome

Garrett Sager[1,2] GARRETT.SAGER@YALE.EDU
Fabian Pallasdies[3] FABIANPALLASDIES@GMAIL.COM
Robert Gowers[4] ROBERT.GOWERS@HU-BERLIN.DE
Snusha Ravikumar[1,5] SNUSHA.RAVIKUMAR@YALE.EDU
Daniel Colon-Ramos[1] DANIEL.COLON-RAMOS@YALE.EDU
Susanne Schreiber[4] S.SCHREIBER@HU-BERLIN.DE
Damon A Clark[1] DAMON.CLARK@YALE.EDU

[1]Yale University
[2]Molecular, Cellular, and Developmental Biology
[3]Humboldt University of Berlin
[4]Humboldt Universitat zu Berlin
[5]Neuroscience

Understanding how networks of neurons compute requires an understanding of how strongly neurons are connected. Beyond the number of synapses, many mechanisms control how well signals are transmitted over synapses (neurotransmitter dynamics, Hebbian learning, etc.). Mitochondria also impact synaptic strength by collecting near synapses to provide ATP, buffer calcium, and perform other functions. However, much remains unknown about how their morphology and positioning relate to neuron identity, neural function, and the connectivity of the neuronal network they are embedded in the brain. Here, to address this gap, we analyze the morphology and positioning of mitochondria across 21 neuron types, with approximately 100 neurons within each type, in a large FIB-SEM connectomics dataset. We show that mitochondria obey cell-type specific morphology and positioning rules that are correlated with the connectivity and functional properties of their respective neuron type. Since mitochondria are thought to have local effects on synaptic function, we also asked directly how the connectivity of mitochondrially-associated synapses differs from the full connectome. We find that the mitochondrially weighted connectome is systematically different from the full connectome, since mitochondria associate with presynapses that connect onto specific downstream neurons. This shows that mitochondrial placement is coordinated across the connectome. Further, this work shows how high-resolution data acquired to identify connectomes can be used to ask questions about cell biology and relate subcellular structures to the millimeter length scales of neural circuits

3-116. Neuronal timescales across development and brain areas

Irina Pochinok[1,2] IRINA.POCHINOK@ZMNH.UNI-HAMBURG.DE
Henrik ostby[1] HENRIK.OSTBY@ZMNH.UNI-HAMBURG.DE
Johanna K. Kostka[1] JOHANNA.KOSTKA@ZMNH.UNI-HAMBURG.DE
Guoming Tony Man[1] GUOMING.MAN@ZMNH.UNI-HAMBURG.DE
Mattia Chini[1] MATTIA.CHINI@ZMNH.UNI-HAMBURG.DE
Ileana L. Hanganu-Opatz[1] ILEANA.HANGANU-OPATZ@ZMNH.UNI-HAMBURG.DE

[1]University Medical Center Hamburg-Eppendorf
[2]Center of Molecular Neurobiology, Institute of Developmental Neurophysiology

The ability to engage in complex cognitive tasks relies on the brain's capability to integrate and store information over different temporal scales. This ability is potentially facilitated by the heterogeneity of intrinsic timescales (ITs), estimated by fitting the exponential decay of the neuronal signal autocorrelation function. ITs form a brain-wide gradient, with sensory regions exhibiting fast ITs supporting rapid processing, and associative areas with slower dynamics supporting prolonged integration. The ITs gradient is observed in different recording modalities and conserved across species. However, it is not yet known how ITs develop, and whether the IT gradient is preconfigured or experience-dependent. We address these open questions in a large-scale in vivo electrophysiological dataset, recorded in several brain areas of mice of postnatal day (P) 4-12, in combination with neural network modeling. We infer ITs from single unit activity (SUA) and local field potentials (LFP). We show that, already during the first postnatal weeks, ITs form a gradient like the adult brain. While they differ in absolute values, SUA- and LFP-derived ITs are broadly consistent with each other, except for the striatum. Moreover, while LFP-derived ITs decrease with age, SUA-derived timescales are either constant or increase. Notably, SUA-derived ITs exhibit a remarkable within-brain area variability with a right-skewed and heavy-tailed distribution. To gain mechanistic insights into the structural underpinnings of ITs heterogeneity, we used a biophysically constrained two-population spiking network of leaky integrate-and-fire units. Heavy-tailed ITs distributions were observed exclusively in networks with skewed synaptic properties. Additional factors influencing ITs included the excitation-inhibition ratio and the proportion of excitatory and inhibitory cells. Our study shows that ITs are among the numerous heavy-tailed distributed properties in the brain and that they are organized in a brain-wide gradient from an early age, thereby contributing to the idea of a preconfigured brain.

3-117. Metabolic dynamics shapes neural activity: a framework for control of epilepsy

Richard Sebastian Eydam[1,2] RICHARD.EYDAM@RIKEN.JP
Igor Franovic[3] FRANOVIC@IPB.AC.RS
Louis Kang[1] LOUIS.KANG@RIKEN.JP

[1]RIKEN Center for Brain Science
[2]Neural Circuits and Computations Unit
[3]Institute of Physics Belgrade

Epilepsy affects around 50 million people worldwide, with approximately 30% of cases proving resistant to conventional pharmacological interventions. This resistance arises from the heterogeneous nature of epilepsies underlying etiologies, rendering it a complex spectrum of disorders. An alternative to pharmacological treatment is achieved by a specialized diet, an approach that was pioneered in the early twentieth century. It was found that the impact of fasting on epilepsy can be successfully mimicked by adopting the ketogenic diet (KD), which is a structured diet that consists of low carbohydrate and high-fat contents and is nowadays renowned for its efficacy in reducing epileptic seizures [1]. Nonetheless, with the emergence of modern antiepileptic drugs (AEDs), dietary interventions received limited research attention for several decades, only experiencing a resurgence in recent years. We introduce a model of quadratic integrate and fire neurons coupled to a global energy reservoir. In this model, we can mimic the transition from a standard diet to a ketogenic diet and study its impact on neural excitability. Our proposed mechanism for seizure control hinges on the presence of adenosine triphosphate (ATP)-dependent potassium channels [2], whose activity, in the absence of ATP, results in neuronal hyperpolarization. Our findings support the viability of this mechanism for regulating epileptic activity, demonstrating that adherence to a KD can restore normal neuronal activity. We substantiate this relationship by investigating bifurcations within a corresponding mean-field model [3] and presenting three distinct pathways for transitioning between normal and seizure-like activity: ATP concentration shocks, parametric perturbations in ATP production rates, and external current stimulation.

[1] Lutas, Andrew & Yellen, Gary. DOI: 10.1016/j.tins.2012.11.005 (2013). [2] Joo, Pangyu et al. DOI: 10.3389/fncom.2021.738362 (2021). [3] Montbrio, Ernest; Pazo, Diego; Roxin, Alex . DOI: https://doi.org/10.1103/PhysRevX.5.02 (2015).

3-118. Unveiling Movement Patterns as Cognitive Nodes: Exploratory Decisions Embodied in Macaque Actions

Hildie Leyser[1,2] HILDELITH.LEYSER@MAIL.MCGILL.CA
Taku Hasegawa[3] TAKU.HASEGAWA@RIKEN.JP
Ningyi Zhou[4] NINGYI.ZHOU@TRINITY.OX.AC.UK
Tomomi Watanabe[3] TOMOMI.WATANABE@RIKEN.JP
Akane Nagano[3] ANAGANO.PSYCHOLOGY.JAPAN@GMAIL.COM
Kentaro Miyamoto[3] KENTARO.MIYAMOTO.WG@RIKEN.JP

[1]McGill University
[2]Neurology and Neurosurgery
[3]RIKEN Center for Brain Science
[4]University of Oxford

The decision between whether to Explore vs Exploit is characterised as foraging and has been studied from many different approaches in the psychological and cognitive sciences. Most previous research was conducted in a small testing room with body restraints. However, a variety of body movements is accompanied by decision making for foraging in ecological environments, and thus, we hypothesized that explorative and exploitative decisions were embodied also in animals' movements and the movements' influence on their decision making. Our study aimed to elucidate the intricate nexus between cognition and movement, with a specific focus on decision-making processes. To test our hypothesis, we devised a 3D paradigm 2-armed bandit task for three Japanese Macaque Monkeys in a natural, unrestrained environment. The task presented choices for exploration (white capsule with uncertain reward) and exploitation (coloured capsule which indicates the content of reward), allowing us to discern distinctive movement signatures corresponding to different decision states within the context of their spontaneous movements and behaviours.For analysis, the monkeys' choice behaviour was modelled by multiple regression model with the capsule location and Bayesian and non-Bayesian reward difference, which considers learning of uncertain reward capsule value and does not consider it, respectively. The utilization of DeepLabCut, a markerless motion pose algorithm, allowed for the collection of body part locations, subsequently converted into egocentric coordinates. We found that monkeys manifested exploitation-exploration switch behaviour captured by our multiple regression model even in an unrestrained ecological environment. Moreover, Uniform Manifold Approximation and Projection (UMAP) method to categorise movement patterns with dimensionality reduction

successfully captured the exploitation-exploration switch. Our findings both complement and complicate established literature on explore/exploit decision-making, shedding light on the interconnectedness of movement and cognition. These insights have implications for developing embodied therapeutic approaches, particularly in the evolving landscape of the metaverse.

3-119. Nucleus accumbens glutamatergic afferents integrate outcomes across time

Eshaan Iyer[1,2] ESHAAN.IYER@MAIL.MCGILL.CA
Peter Vitaro[1] PETER.VITARO@MAIL.MCGILL.CA
Serena Wu[1] SERENA.WU@MAIL.MCGILL.CA
Jessie Muir[1] JESSIE.MUIR@MAIL.MCGILL.CA
Vedrana Cvetkovska[1] VEDRANA.CVETKOVSKA@MCGILL.CA
Rosemary Bagot[1] ROSEMARY.BAGOT@MCGILL.CA

[1]McGill University
[2]Integrated Program in Neuroscience

To maintain flexible, adaptive behavior, information about outcomes must be integrated across time. In the brain, the nucleus accumbens (NAc) integrates a variety of glutamatergic inputs with dopaminergic input to direct motivated behavior. How and what information is integrated by distinct glutamatergic inputs remains unclear. To probe this, we used dual-site in vivo fiber photometry in mice performing a two-armed bandit task to examine reward integration in two monosynaptic inputs that converge in NAc medial shell, the medial prefrontal cortex (mPFC) and ventral hippocampus (vHip). We find that both mPFC-NAc and vHip-NAc encode information about outcome with reward driving suppression of neural activity and loss gradually restoring activity. Reward-mediated suppression continues into subsequent trials, tracking prior reward history as a graded function of loss. Despite similar encoding, conditional entropy analyses reveal that, while mPFC-NAc consistently encodes outcome, vHip-NAc encoding is sensitive to uncertainty. Sequentially degrading task requirements, removing first choice and then all response requirement, revealed that action-outcome pairings drive outcome history encoding, suggesting a key role for these pathways in integrating outcomes to guide action. In line with this, higher mPFC-NAc and vHip-NAc activity associates with decreased task engagement and optogenetic stimulation of either pathway suppressed engagement, with an additive effect of simultaneous dual-pathway stimulation. We identify coordinated mPFC-NAc and vHIP-NAc activity as a neural mechanism for propagating reward-associated information across time to dynamically modulate task engagement in rewarding environments.

3-120. Parallel movement planning via an optimal preparatory state in motor cortex

Nicolas Meirhaeghe NMRGHE@GMAIL.COM
Alexa Riehle ALEXA.RIEHLE@UNIV-AMU.FR
Thomas Brochier THOMAS.BROCHIER@UNIV-AMU.FR

Institut des neurosciences de la Timone

An influential hypothesis in motor neuroscience proposes that motor cortex implements a neural dynamical system whose initial state is optimized during the preparatory phase of movement. Despite its conceptual impact on the field, the so-called “initial condition” (IC) hypothesis has remained difficult to test in experimental conditions that go beyond those that motivated its formulation. It is thus crucial to identify and test novel predictions of the IC hypothesis in more complex behavioral settings. Here, we focused on the case of parallel planning, whereby multiple movements are planned simultaneously. Parallel planning poses an interesting challenge to the motor system, and can be used to expose new predictions of the IC hypothesis. In particular, the hypothesis predicts that preparatory activity associated with two movements planned in parallel should settle in an intermediate state between the states associated with the two movements planned individually. To test this prediction, we analyzed motor cortical activity of non-human primates trained on an instructed, delayed reach-to-grasp task. In this task, following a Go signal, animals had to grab an object using one of two possible hand grips. Critically, the desired grip was either cued in advance, or left uncertain until the time of Go, thus providing the means to compare planning of a single grip versus two simultaneous grips. As predicted by the IC hypothesis, we found that preparatory activity associated with parallel planning reached an intermediate state between the states associated with the two grips planned separately. Moreover, fine fluctuations at the single trial level of this intermediate state were directly reflected in the animals’ readiness to execute one grip versus the other, as assessed by reactions times. These results offer strong empirical support to the IC hypothesis, and generalize its validity to the important case of parallel motor planning.

3-121. Network mechanisms for statistical learning and place field formation in the hippocampus

Margaret Lane[1] MARGARET.LANE@EPFL.CH
Merkourios Simos[1] MERKOURIOS.SIMOS@EPFL.CH
James Priestley[2,3] JAMES.PRIESTLEY@EPFL.CH

[1]EPFL
[2]Ecole polytechnique federale de Lausanne, CH
[3]Brain Mind Institute

Sparse coding is a ubiquitous model for brain computation in both early sensory and higher-order systems. In the hippocampus, a critical locus for memory formation, neurons produce sparse, spatially selective firing fields that are distributed throughout the animal's physical environment. Artificial neural networks have similarly been shown to produce localized, manifold-tiling responses under a variety of computational objectives and constraints by processing sensory or action information sampled during navigation. These include networks that explicitly encourage sparse representations, either by random projections or learned features, or networks that optimize alternative objectives such as prediction of future observations. Here we show that in a simple simulation of 1D navigation, these models often fail to learn localized place fields. We introduce a novel autoencoder model that explicitly optimizes the dimensionality of neural representations, and find that it robustly learns local place representations across a spectrum of navigation tasks. Through analysis of the learning dynamics, we find highly divergent representational geometry across different model classes. In particular, the genesis of place fields is driven by the compression of correlated, nearby sensory inputs and effective competition between neurons. We show that this solution naturally arises in non-negative networks when balancing faithful reconstruction of the inputs and dimensionality expansion, as long as the underlying input manifold is smooth. Using large-scale neural recordings of mice navigating in virtual reality, we connect these learning processes to the evolution of place fields in novel environments, where we identify congruent signatures of pattern decorrelation and competitive dynamics. These features are concomitant with increased independence between neurons and increased dimensionality as a function of experience. Overall, we suggest that place coding may arise as a consequence of generalised computations that optimize high dimensional representations from the statistics of ongoing experience, ameliorating both computational flexibility and robust memory storage.

3-122. Revealing effects of nonlinear response properties on visual perception using temporal divisive normalization

Amber Brands[1,2] A.M.BRANDS@UVA.NL
Nikolina Vukšić[1] NIKAVUKSIC@GMAIL.COM
Paulo Ortiz[1] PAULO.ORTIZFONSECA@GMAIL.COM
Iris Groen[1] I.I.A.GROEN@UVA.NL

[1]University of Amsterdam
[2]Informatics Institute

Our perception of sensory inputs depends heavily on nonlinear computations evident in various neural response properties, including contrast gain (the sigmoidal relationship between stimulus contrast and the neural response) and repetition suppression (reduced responses to repeating stimuli). While these phenomena have each been studied extensively in isolation, their joint impact on perception is unclear. Here, we collected neural and behavioural measurements while humans performed an object classification task with temporally repeated noise patterns, whereby object contrast was varied. Results show a performance increase as a result of adaptation to the noise pattern, with higher recognition scores for similar compared to different noise preceding the object. This benefit of adaptation was most pronounced for intermediate contrast levels. Consistently, neural responses exhibited increased response magnitudes for intermediate contrast objects when adapting to the same noise. To elucidate the computational mechanisms mediating the neural responses and perception we employ deep convolutional neural network (DCNN) modeling. DCNNs have recently emerged as promising models of sensory processing in the human and primate brain and offer an approach to link neural phenomena to behavioural outcomes. We endow the DCNN with temporal divisive normalization, a biophysically-realistic canonical computation that has been shown to capture nonlinear neural dynamics. We find that feedforward DCNNs with divisive normalization capture human behaviour and neural responses more effectively than other types of networks, including Recurrent Neural Networks, by incorporating both contrast gain and suppression effects. Examining the network representations, we reveal that adapting to the same noise results in increased activations for the object due to noise suppression. Overall, this work demonstrates perceptual benefits of nonlinear response properties and offers a modeling framework to study neural mechanisms involved in object recognition in temporal sequences.

3-123. A data reduction method for on-board unit analysis in Next-Gen active CMOS-based BCIs

Matteo Vincenzi[1] MATTEO.VINCENZI@IIT.IT
Alberto Perna[2] ALBERTO.PERNA@IIT.IT
Gabor Orban[3] GABOR.ORBAN@IIT.IT
Christine Stubbendorff[3] CHRISTINE.STUBBENDORFF@IIT.IT
Joao Filipe Ribeiro[3] JOAO.RIBEIRO@IIT.IT
Gian Nicola Angotzi[4] GIANNICOLA.ANGOTZI@IIT.IT
Luca Berdondini[3] LUCA.BERDONDINI@IIT.IT

[1]Istituto Italiano di Tecnologia (IIT)
[2]Fondazione Istituto Italiano di Tecnologia(IIT); Open University Affiliated Research Centre(ARC@IIT)
[3]Fondazione Istituto Italiano di Tecnologia (IIT)
[4]Fondazione Istituto Italiano di Tecnologia (IIT); Corticale Srl, Genova, Italy

Brain Computer Interfaces (BCIs), which enable communication links between brain and external artificial machines, are getting the attention of the scientific and biotech communities due to their suitability for chronic experiments and practical applications. When combined with active CMOS-based implantable probes, high data rates are required for the transmission of the full-resolution neural data provided by up to thousands of closely spaced microelectrodes made available by such probes. Due to bandwidth limitations typical for on-body transmission technologies, BCIs require on-board data reduction strategies, limiting streaming to only relevant neural activity, thereby relaxing other design constraints (e.g., form factor, battery duration and transmission range). This is even more important in wireless scenarios. The current state-of-the-art in data reduction includes dimensionality reduction, data decimation, and source compression. As the latter may be viewed as an additional step for the other methods, this study exclusively concentrates on the first two approaches. Although single unit sorting is a good candidate for offline dimensionality reduction, the high-complexity (growing with the channel count) and the absence of a standardized and unsupervised solution, makes it an over-dimensioned solution for on-board data reduction. In this study, we investigate the space-time redundancy provided by CMOS active dense neural probes to implement a hybrid data reduction method based on electrodes decimation and clustering. Results are given for a chronic recording performed over 29 days and compared with those of standard sorting tools. Our novel approach opens the path for the design of innovative architectures for low-complexity and real-time data reduction.

3-124. Massive impact of isoflurane anesthesia on sound representations in the auditory brainstem

Etienne Gosselin ETIENNE.GOSSELIN@PASTEUR.FR
Sophie Bagure SOPHIE.BAGUR@PASTEUR.FR
Brice Bathellier BRICE.BATHELLIER@PASTEUR.FR

Institut Pasteur

Increasing evidence indicates that anesthetics strongly affect sensory representations and network dynamics in the thalamo-cortical circuits. By contrast, processing in the early sensory systems is thought to be mostly preserved under anesthesia, based on stimulus tuning properties like those observed in wakefulness. However, in the absence of recordings tracking the activity of the same neurons across states, there is no evidence that anesthesia preserves one-to-one peripheral sensory representations. Here, tracking for the first time the same cochlear nucleus neurons across wakefulness and isoflurane anesthesia, we show that the amplitude, sign, and tuning width of single neuron responses to sounds are massively affected by anesthesia in the first relay of the auditory system. Hence, the functional identity of cochlear nucleus neurons is tightly linked to the network state and is ill-determined under anesthesia. Moreover, using dimensionality reduction and classification techniques, we show that sound representations during wake and under anesthesia lie in different neural sub-spaces. In line with this, representation of sounds in anesthesia cannot be decoded based on classifier trained on awake sound representations. Together, these results demonstrate a much larger effect of anesthesia than previously thought on peripheral auditory processing and indicate that anesthesia will strongly perturb the normal integration of sound information in downstream targets of the cochlear nucleus. This supports the use of awake animals to study the evolution of sound representations in any area of the sensory system. Interestingly, when performed separately on awake or anesthesia data, single cell tuning measures and population decoding showed similar tuning width properties and information levels. Hence, this explains why the massive changes observed when tracking the same neurons stay unnoticed when recording separately in the two states.

3-125. Geometry of anisotropic contextual interactions in the visual cortex places fundamental limits on spatial vision.

Mitchell Morton[1,2] MITCHELL.MORTON@YALE.EDU
Sachira Denagamage[1,2] SACHIRA.DENAGAMAGE@YALE.EDU
Nyomi Hudson[1] NYOMI.HUDSON@YALE.EDU
Anirvan Nandy[1] ANIRVAN.NANDY@YALE.EDU

[1]Yale University
[2]Neuroscience

Crowding, the impaired ability to accurately recognize a target stimulus among distractors, is a major bottleneck in visual perception. The spatial configuration of distractors in relation to the target profoundly influences perceptual fidelity. Notably, when a distractor is placed at a more eccentric point on the radial axis (termed 'radial-out crowding'), it exerts the strongest impairment. Despite the pronounced perceptual anisotropy, the prevalent assumption underlying our understanding of contextual interactions in the visual cortex assumes isotropy. We investigated how distractor stimuli in different spatial configurations impacted the representation of a target stimulus in laminar microcircuits in the primary visual cortex (V1). Our study reveals that radial-out crowding more strongly impacts the ability to decode the target orientation from V1 population activity compared to other spatial configurations. This effect was strongest among putative excitatory neurons in the superficial and input layers, which are the primary neural populations involved in feed-forward information propagation. Remarkably, the feedback pathway involving the deep cortical layers does not exhibit anisotropy. Mechanistically, the anisotropy is explained by a tuned suppression and untuned facilitation of orientation responses, leading to an anisotropic broadening of tuning curves in the feedforward pathway, but not in the feedback pathway. These results underscore the non-uniform spatial integration of information by neurons in the visual cortex, establishing the presence of anisotropic contextual interactions in the earliest stages of cortical processing. By elucidating the distinct roles of feed-forward and feedback pathways in the context of crowding, this study advances our understanding of the intricate interplay between spatial arrangement, neural circuitry, and the constraints on perceptual fidelity during early visual processing.

3-126. Ambiguity aversion arises via distributional sampling of nonlinear future reward states

Kenway Louie[1,2] KL837@NYU.EDU

[1]New York University
[2]Center for Neural Science

Empirical decision-making in biological choosers depends markedly on outcome uncertainty. Such uncertainty can differ in the degree of knowledge held by a chooser: in decisions under risk, outcomes are probabilistic but those probabilities are known; in decisions under ambiguity, the probabilities themselves are unknown or uncertain. While human choosers generally exhibit aversion to both risk and ambiguity, these two types of uncertainty preferences differ in fundamental ways. Risk aversion can be rationally derived as a normative tradeoff between magnitude and probability according to individual chooser utility functions. In contrast, ambiguity aversion is normatively irrational and cannot be explained by existing models of valuation and choice. Here, we show that ambiguity aversion arises naturally in agents employing (1) a normalized value representation and (2) a distributional sampling of possible outcome states under ambiguity. This nonlinear forward sampling model replicates known characteristics of empirical ambiguity aversion, including preference for risky over ambiguous options, a quasi-linear relationship between valuation and ambiguity, and a dissociation between risk and ambiguity preferences. At the behavioral level, this model makes the counterintuitive (and testable) prediction that the degree of ambiguity aversion varies inversely with the density of forward sampling. At the neural level, the state sampling inherent in the model suggests a mechanism for ambiguity preference via distributional reinforcement learning, a process recently linked to normalized value representations. Together, these results offer a simple explanation for ambiguity preferences based on normalized value coding and forward inference, and argue for an incorporation of biologically valid value functions in computational models of decision-making.

3-127. Conservation of sensory coding in the auditory cortex of mice between wakefulness and sleep

Allan Muller ALLAN.MULLER@PASTEUR.FR
Sophie Bagure SOPHIE.BAGUR@PASTEUR.FR
Brice Bathellier BRICE.BATHELLIER@PASTEUR.FR

Institut Pasteur

Sleep is a physiological state often associated with a cutoff from the external world and especially loss of conscious perception of sensory stimuli . Even though brain signatures of sleep have been studied for years, the mechanisms that prevent sensory information from reaching consciousness remains to be elucidated . Recent findings suggest that the loss of perception experienced in anesthesia may be due to neurons population activity coding of stimuli in the cortex collapsing into the spontaneous activity subspace, therefore reflecting internally generated dynamics rather than evoked sensory responses (Filipchuk et al. 2022). In this study, we examined whether this property is also at play during sleep. We used two photon calcium imaging to record the responses to sounds of hundreds of neurons simultaneously in the primary auditory cortex of awake and naturally asleep mice. The analysis of the population activity revealed that, contrary to anesthesia, sounds were encoded with highly similar population activity patterns between the awake and NREM states. This conservation of sounds representations between wakefulness and sleep suggests that sensory processing is impacted by qualitatively different mechanisms in anesthesia and in sleep. Moreover, while under anesthesia sounds representations in the cortex are highly deteriorated, our result indicates that high level cortical processing of sounds is intact during NREM sleep, a property that is at odd with the classical view that the thalamus gates sensory information to the cortex during sleep, but which may however be critical for survival.

3-128. High-dimensional communication and gating of behavioral information across cortical areas

Lee Susman[1,2] LEE.SUSMAN@GMAIL.COM
Johnatan Aljadeff[3] ALJADEFF@UCSD.EDU
Tom Kern[4] KERNTOM00@GMAIL.COM
Karel Svoboda[5] KAREL.SVOBODA@ALLENINSTITUTE.ORG
Arseny Finkelstein[4] ARSENYF@GMAIL.COM

[1]Princeton University
[2]Physics
[3]University of California San Diego
[4]Tel Aviv University
[5]Allen Institute

Many cognitive processes unfold on a wide range of spatial and temporal scales and involve multiple brain areas. Recent studies have begun to characterize spatiotemporal dynamics across the neocortex. However, little is known about how information is processed and propagated within and across cortical areas – at the level of individual cells. To address this gap, we used mesoscale calcium imaging to record the activity of up to ~30,000 neurons simultaneously (~1,000,000 neurons total) from over 10 cortical areas, including motor, somatosensory, and high-order visual areas. Cortical activity was recorded while mice performed a goal-directed task involving multidirectional tongue-reaching for water rewards. Global cortical activity (across areas) and activity within each cortical area were high-dimensional (Stringer et al., 2019 Nature), and the dimensionality of neural interactions changed during the behavioral task in an area-specific way. The spatial scale of activity along the cortical sheet exhibited a broad spectrum. Specifically, dimensions of neural activity with high variance were spatially extended, spanning multiple areas, whereas dimensions with intermediate variance were local. Do these extended spatial correlations reflect internal communication between brain areas or shared behavioral signals? Contrary to previous reports, we found that the communication subspace (Semedo et al., 2019 Neuron) between brain areas was high-dimensional. Furthermore, it was largely non-overlapping with the subspace of neural activity that captured movement-related information. Our results suggest that neural activity differentially gates movement-related signals conveyed between cortical areas, such that other information (e.g., cognitive, motivational) can be conveyed via communication channels that are independent of movement representations.

3-129. Fear conditioning reduces the influence of external stimulation on network reorganization

Thomas Lai[1,2] LAITHOMA@UNI-MAINZ.DE
Jens-Bastian Eppler[3] EPPLER@FIAS.UNI-FRANKFURT.DE
Dominik F. Aschauer[1] DASCHAUER@UNI-MAINZ.DE
Simon Rumpel[1] SIRUMPEL@UNI-MAINZ.DE
Matthias Kaschube[3] KASCHUBE@FIAS.UNI-FRANKFURT.DE

[1]Johannes Gutenberg University Mainz
[2]Institute for Physiology
[3]Frankfurt Institute for Advanced Studies

Learning is a fundamental process in neuroscience, yet the intricate relationship between behavioural learning and the underlying mechanisms at the neuronal circuit level remains elusive. We aim to address this knowledge gap by investigating how fear conditioning influences neuronal activity and the associated neural network in the mouse auditory cortex. In this study, we examined signal - and noise correlations in pairwise neuronal activity data collected during chronic imaging experiments in the mouse auditory cortex [1]. Under basal conditions, where explicit learning is absent, we found substantial changes across days, both in signal correlations, reflecting the similarity of neuronal responses to sensory stimuli, and in noise correlations, which serve as a measure of effective connectivity. Furthermore, we observed that signal correlations predicted noise correlations at a later time. Specifically, neuron pairs with high signal correlation had increased noise correlations two days later, compared to neuron pairs with low signal correlation. However, following auditory cued fear conditioning, this relationship was altered. The predictive power of signal correlations on noise correlations decreased. To interpret these results, we employed a minimal network model with random- or Hebbian-like plasticity. In this model, a Hebbian-like learning mechanism in the recurrent weights was required to obtain the observed predictive power of signal correlations on noise correlations. This effect could not be reproduced by random drift alone. Furthermore, the decrease in predictive power following fear conditioning could be attributed to a reduction in synaptic learning rates. This finding suggests that fear conditioning may act to decelerate the ongoing process of statistical learning, where new information continually overwrites old information. Our results propose that in the absence of behavioural learning, inputs to sensory cortex constantly overwrite the network structure. Fear conditioning appears to slow down this process, potentially facilitating the transfer of information to long-term memory storage.

3-130. Relating network heterogeneity to the dimension of population covariability

Gengshuo Tian[1,2] GTIAN@UCHICAGO.EDU
Oliver Zhu[1] ZHUOU@UCHICAGO.EDU
Vinay Shirhatti[1] VINAYS@UCHICAGO.EDU
David Freedman[1] DFREEDMAN@UCHICAGO.EDU
Brent Doiron[1] BDOIRON@UCHICAGO.EDU

[1]University of Chicago
[2]Department of Statistics

Both trial-averaged population activity and its trial-to-trial fluctuations are critical aspects of neural representation. A heterogeneous trial-averaged response among neurons confers a myriad of benefits on neural computation, which has been extensively studied under the framework of sparse coding (Olshausen & Field, 2004). At the same time, trial-to-trial fluctuations of neural activity that are concentrated around a low-dimensional subspace orthogonal to the informative subspace have been shown to facilitate linear decoding downstream (Averbeck et al., 2006). Here we link these two aspects of neural representation using a recurrent circuit model and derive the following relation: the more heterogeneous the distribution of trial-averaged responses, the lower the effective dimension of population trial-to-trial covariability. This surprising prediction is tested and validated using multiple population datasets from numerous brain areas in both non-human primates and mice. We verify in simulation that indeed a sparser code leads to better fine discrimination performance through a lowering of the dimension of population covariability. In line with this result, we show that neural populations across the brain exhibit both more heterogeneous mean responses and lower-dimensional fluctuations when the brain is in more heightened states of information processing. In sum, we present a key organizational principle of neural population response that is widely observed across the nervous system and acts to synergistically improve population representation.

3-131. Recurrent networks under constraint of sparse reward learn interacting belief state dynamics

John Schwarcz[1] JOHNSCHWARCZ@GMAIL.COM
Jan Bauer[1] JAN.BAUER@RWTH-AACHEN.DE
Eran Lottem[2] ERAN.LOTTEM@MAIL.HUJI.AC.IL
Jonathan Kadmon[2] JONATHAN.KADMON@MAIL.HUJI.AC.IL
Gabrielle Marmur[1] GABRIELL.MARMUR@MAIL.HUJI.AC.IL
Haneen Rajabi[1] HANEEN.RAJABI1@MAIL.HUJI.AC.IL

[1]Hebrew University of Jerusalem
[2]The Hebrew University of Jerusalem

Animals must develop adaptive strategies to rapidly optimize decision-making in dynamic environments, often without the benefit of immediate rewards. Existing literature posits that animals use internal "belief" states as the foundation for their decision policy. However, the neural mechanism implementing reward-independent belief updates is poorly understood. This is a critical gap, especially in scenarios where trial-and-error approaches are inefficient and potentially perilous. To address this problem, we take a multidisciplinary approach integrating theoretical derivation, training artificial neural networks, and behavioral experiments in rodents to explore potential neural mechanisms.

A belief state is a joint probability distribution over all relevant latent variables of the environment. Updating the joint distributions is computationally demanding and nontrivial, particularly when the agent averts acting. To tackle these challenges, we develop a change-detection task to capture the complexity of partially observed dynamic environments while being accessible to experimental and analytical treatment. We formulate a Bayesian theory for sequentially updating joint probabilities and demonstrate that neural networks can accomplish the task near optimally, even without immediate rewards. In particular, we show that the network dynamics mirror the joint update of the Bayesian latent state estimators. Furthermore, the behavior of rodents trained on this task aligns with our theoretical model and neural network simulations, suggesting that mice utilize dynamic internal state representation and inference to solve this task. Overall, our findings elucidate the computational principles behind zero-shot adaptation to environmental changes without the need for trial and error.

3-132. Mental representations of latent states in the human brain

Flora Bouchacourt[1,2] FLORA.BOUCHACOURT@GMAIL.COM
Linda Yu[1] LINDA_YU@BROWN.EDU
Avinash Vaidya[3] AVINASH_VAIDYA@BROWN.EDU
Aigerim Akhmetzhanova[4] AIGERIM_AKHMETZHANOVA@BROWN.EDU
Sienna Bruinsma[1] SIENNA_BRUINSMA@BROWN.EDU
Matt Nassar[1] MATTHEW_NASSAR@BROWN.EDU

[1]Brown University
[2]Neuroscience
[3]NIH, Natl. Inst. on Drug Abuse (NIDA), Baltimore, MD
[4]UC Davis

Behavioral policies need not always be relearned from scratch when the environment changes. When appropriate, we build mental representations of the hidden contingencies ("latent states"), and flexibly use these representations to retrieve a previous behavior more rapidly. The prefrontal cortex and the hippocampus have been argued to be the locus of the neural representations of latent states, however with a limited characterization of their implementation. To examine the formation and reuse of latent states in the brain, we designed a human task that dissociates latent states from action-outcome contingencies to quantify behaviorally when a latent state is being reused. Participants were required to learn two latent states, each composed of five color-location associations. Looking at behavior, we found evidence for gradual learning of the latent states. We also found evidence for single trial reuse of the latent states, after an initial learning phase of which length varied across subjects. To explain the behavior and the individual variability, we developed a model in which latent states emerged through chunking of conjunctive neurons selective to colors and locations activated in temporal contiguity between trials. When repeated, a latent state could be retrieved from a single color-location association by reactivation of the whole chunk. The model predicted abrupt changes in behavioral responses after switches in latent states, which were borne out by the data. Ongoing analyses examine whether and how fMRI signals in the orbitofrontal cortex and dorsolateral prefrontal cortex relate to learning and inference in the model, with the end goal of better characterizing the computational roles of these regions.

3-133. A role for hippocampal CA1 in structural learning in mice

Svenja Nierwetberg[1,2] SVENJA.NIERWETBERG.18@UCL.AC.UK
David Orme[1,3] DAVID.ORME.17@UCL.AC.UK
Andrew MacAskill[1,3] A.MACASKILL@UCL.AC.UK

[1]University College London
[2]Sainsbury Wellcome Centre
[3]Neuroscience, Physiology and Pharmacology

The meaning of individual events or cues in the environment is often dependent on other cues surrounding them, for example their order in time, or location in space. The ability to learn about relationships between such ambiguous cues – often called relational, or structural learning – enables us to recognise common underlying structures of events and is thought to form the basis of episodic memory. One area implicated in structural learning is the hippocampus. Specifically, neurons in the CA1 area of the hippocampus have been shown to represent variables essential for constructing a relational structure of the environment, such as cue configurations and their order in space and time. To investigate the neural basis of structural learning, we designed an odour-based task that requires mice to learn not only about sets of odour cues, but about their relative order in time. Importantly, the task design allows for manipulation of the temporal structure and identity of cues separately, allowing dissociation of their neural mechanisms. Using this task, we found that mice can flexibly use previously learnt relational structures, and adapt to both changes in the temporal pattern as well as in cue identity. In line with a role for hippocampal circuitry, optogenetic inactivation of ventral CA1 (vCA1) markedly impaired task performance. Using in vivo calcium imaging, we found that vCA1 neurons encode a wide variety of task-relevant information, including maintaining odour identity across the delay and exhibiting context-specific responses to odours. Together, we show that mice can rapidly learn and flexibly perform an olfactory structural learning task, and that vCA1 plays an important role in connecting the individual cues of the structure across a delay. Ongoing work aims to understand how the activity of vCA1 cells at the population level supports the integration of sensory and internal variables to support structural learning.

3-134. Combinatorial architecture of circuit neuromodulation

Nikolas Karalis NIKOLASKARALIS@GMAIL.COM
Andreas Luthi ANDREAS.LUTHI@FMI.CH

Friedrich Miescher Institute for Biomedical Research

Internal states and behavior are associated with brain-wide changes in the activity of neuronal circuits. Such global shifts are believed to be implemented by the changes in the neuromodulatory tone. Neuromodulators originate from few neurons and are released widely throughout the brain where they alter the activity and plasticity of target neural circuits.

The amygdala is a core brain region that plays an essential role in the processing of emotional stimuli. Salience and synaptic alterations in this region are critical for emotional processing and learning. The amygdala is innervated by all major neuromodulators, yet we know little about the dynamics of neuromodulator release under physiological conditions and how neuromodulator combinations regulate amygdala circuit activity during different behavioral states and learning.

To address this question we characterized the neuromodulatory inputs to the amygdala across behavioral states, including exploration, sleep, reward, and aversion learning. Using novel sensors that report the dynamics of neuromodulator release, paired with multi-site, multi-color fiber photometry, we simultaneously measured the activity of all major neuromodulators (dopamine, acetylcholine, serotonin, and norepinephrine) in the amygdala of behaving mice. Using this data, we characterized their release patterns during distinct behaviors and identified the interactions between different neuromodulators.

Further, using simultaneous optical and electrophysiological recordings from the amygdala using high-density silicon probes, we have identified distinct patterns of large-scale neuronal activity that are associated with differential neuromodulator combinations during distinct behavioral states. In summary, using this novel approach and multi-modal recordings, we are able to jointly characterize the activity of these two core systems and to generate predictions about the causal influence of simultaneous neuromodulation on the activity of downstream circuits during behavior.

This data will enable the design and validation of new models about the combinatorial action of neuromodulators in defining internal states and organizing circuit activity.

3-135. Actionable Neural Representations: Optimal Representations of Internal Models

William Dorrell[1,2] DORRELLWEC@GMAIL.COM
Peter E Latham[1] PEL@GATSBY.UCL.AC.UK
Mohamady El-Gaby[3] MOHAMADY.EL-GABY@NDCN.OX.AC.UK
Tim Behrens[4] TIMOTHY.BEHRENS@NDCN.OX.AC.UK
James Whittington[5] JCRWHITTINGTON@GMAIL.COM

[1]University College London
[2]Gatsby Unit
[3]University of Oxford
[4]University of Oxford & University College London
[5]University of Stanford

Understanding neural tuning is a fundamental problem in neuroscience. Most theories of neural representations descend from the Efficient Coding Hypothesis. This argues that neural populations are informatively encoding certain variables under biological efficiency constraints. While broadly successful for sensory encodings, it's been difficult to apply the same principles to cognitive representations (e.g. grid cells). We argue this is because many cognitive tasks require an internal model which is not captured by just efficiently encoding variables. To remedy this, we present a normative theory of neuronal computations using an internal model (i.e. answering: "if I take this action how will this variable change?"). We show this theory introduces additional constraints on the representation ('actionability') that alter the optimal representation, and we'll apply this theory in three settings. First, the theory accounts for recently recorded prefrontal representations of task structure (the 'music box' representation of action sequences), and additionally predicts changes when the task rules change. Second, our theory provides a complete normative theory of multi-modular grid cells, and makes predictions we verify in data, such as the relative inter-modular angle. Third, for latent state representations of arbitrary sequence tasks, our theory prescribes when additional neurons (splitter cells) are required versus differential activity of the same neurons (rate-remapping). Neural activity is the substrate of the computations that undergird behaviour. To understand these computations we need theories that comprehensibly link (i.e. more than neural networks) computations and neural responses. Our theory hopes to build one such link.

3-136. Low dimensional representations of schema and value in the monkey hippocampus

Sofia Landi[1] SLANDI@UW.EDU
Solana Fernandez[2] SOLANAF@UW.EDU
Kevin Lu[2] KLU19@JHU.EDU
Ellen Bakotich[2] ESTREITW@UW.EDU
Vinay Shirhatti[3] VINAYS@UCHICAGO.EDU
David Freedman[3] DFREEDMAN@UCHICAGO.EDU
Adrienne Fairhall[1] FAIRHALL@UW.EDU
Elizabeth Buffalo[2] EBUFFALO@UW.EDU

[1]University of Washington
[2]Physiol. and Biophysics, University of Washington; Washington Natl. Primate Res. Ctr., Seattle, WA
[3]University of Chicago

A concept bridging spatial representation and memory that may be a powerful principle for the primate hippocampus is that of the cognitive map, a form of internal model which allows one to organize knowledge gained from experience, assimilate new information and plan. Using high-density laminar probes, we recorded the activity of hippocampal neurons in nonhuman primates carrying out a set of naturalistic foraging tasks in a virtual reality environment. Foraging is a compelling and rapidly learned behavior that allows us to evaluate hippocampal activity in a complex goal-directed navigation task. In this paradigm, the value of foraged items is governed by an underlying schema, defined as an abstract relational map that captures a task structure and generalizes across different instantiations. Unlike other behavioral tasks that can take many months for monkeys to learn, our paradigm enables efficient and robust learning within a single recording session. We probed multiple exemplars of a single schema, as well as multiple schema variations, i.e., tasks governed by different abstract relationships, within single sessions. We examined representations of sensory inputs, their associated values, and the abstract task structure at both the level of single units and the simultaneously recorded neural ensemble. We used principal component analysis (PCA) to identify the leading covariance patterns in the population. Although individual neurons showed significant mixed selectivity, we identified a low dimensional subspace in the hippocampus that captures information about reward value and task schema. Value maps emerged in a schema-dependent manner, i.e., spatial value maps were elicited only during learning under a spatial schema. We observed a similar trend at

the single unit level, with hippocampal neurons evolving schema-selective maps. The hippocampus thus emerges as a key node in linking crucial aspects of ongoing experiences with their outcomes, weighting value maps in a schema-dependent manner.

3-137. Hierarchical, structure-yoked integration spontaneously emerges with self-supervised training on speech

Samuel Norman-Haignere[1,2] SAMUEL_NORMAN-HAIGNERE@URMC.ROCHESTER.EDU
Pierre Orhan[3,4] ORHAN.PIERRE.FRANCE@GMAIL.COM
Jean-Remi King[3] JEANREMI@META.COM
Yves Boubenec[3] YVES.BOUBENEC@ENS.FR

[1]University of Rochester Medical Center
[2]Biostatistics & Computational Biology, Neuroscience
[3]Ecole Normale Superieure
[4]Laboratoire des systemes perceptifs

Biological sensory systems learn to flexibly integrate across hierarchically organized structures such as phonemes, syllables, and words in speech without explicit supervision, a challenging task that has been difficult to replicate in machine hearing systems. Recent work has shown that deep neural networks (DNNs), trained on a "self-supervised" prediction task, learn representations that are predictive of perceptually important speech structures, as well as brain responses to speech in high-level regions of the auditory cortex, providing a candidate model of structure discovery. However, little is known about how these networks learn to integrate across the complex multiscale structure of natural stimuli. To address this issue, we used the temporal context invariance [1] paradigm to measure temporal integration windows from a popular, self-supervised DNN trained on speech [2]. We show that self-supervised training causes DNNs to integrate hierarchically across speech with substantially longer integration windows in deeper network layers. Moreover, we find that the integration windows in deeper layers dynamically expand and contract as the structures that compose speech are stretched and compressed, suggesting they are partially yoked to structure duration rather than absolute time. We show that these integration dynamics are absent from an untrained model and are learned hierarchically during training, with later layers achieving their final integration characteristics later in training. Finally, we show that the learned integration strategies vary substantially with the type of sounds they are trained on, with distinct strategies learned for music. These findings provide new insights into how self-supervision shapes integration dynamics, and provides testable metrics and predictions for comparing the integration dynamics and learning strategies between artificial and biological neural systems.

3-138. RatInABox: A unified Python framework for modelling spatial behaviour and neural data

Caswell Barry[1] CASWELL.BARRY@UCL.AC.UK
Mehul Rastogi[2] M.RASTOGI@UCL.AC.UK
William de Cothi[1] WILLIAM.COTHI.15@UCL.AC.UK
Claudia Clopath[3] C.CLOPATH@IMPERIAL.AC.UK
Kimberly Stachenfeld[4] STACHENFELD@DEEPMIND.COM
Tom George[2] TOMGEORGE1@BTINTERNET.COM

[1]UCL
[2]Sainsbury Wellcome Centre, UCL
[3]Imperial College London
[4]Google DeepMind

Studying the brain's role in spatial navigation often necessitates creating synthetic behaviour and/or neural data. Without a common framework, this time consuming process raises the entry barrier to computational research and makes it difficult to reproduce or compare research which generates data in different ways. Here we present RatInABox, an open source Python toolkit for modelling locomotion and synthetic neural data from spatially modulated cell types. RatInABox provides users with (i) the ability to construct environments with configurable barriers, holes, visual cues and rewards, (ii) a physically realistic random motion model fitted to experimental data, (iii) fast online calculation of neural data for most of the known self-location or velocity selective cell types in the hippocampal formation (including place cells, grid cells, boundary vector cells, head direction cells etc.) and (iv) a framework for constructing bespoke complex cell types, multi-layer network models and data- or policy-controlled motion trajectories. Motion and neural models are spatially and temporally continuous and topographically sensitive to boundary conditions and walls. RatInABox was built to be intuitive, visual and easy to learn, supported

by a set of diverse and well-documented tutorials. We hope this tool will significantly streamline computational research into spatial navigation.

3-139. Impact of Anxiety and Depression on Decision-Making: Insights from Readiness Potentials and Drift-Diffusion Modeling

Mrugsen Gopnarayan[1] MRUGSENNG@GMAIL.COM
Feng Sheng[2] FSHENG@ZJU.EDU.CN
Wenjia Zhao[2] ZHAO.3463@OSU.EDU
Michael Platt[2] MPLATT@PENNMEDICINE.UPENN.EDU
Arjun Ramakrishnan[3] ARJUNR@IITK.AC.IN

[1]University of Hamburg
[2]University of Pennsylvania
[3]Indian Institute of Technology Kanpur

It has long been recognized that experiencing anxiety and depression significantly alters our perception of the world. However, the extent to which these conditions affect decision-making remains incompletely understood. While previous studies have often utilized prospect theory to model economic decision-making, yielding mixed results regarding the impact of anxiety or depression on such decisions[1], prospect theory itself does not fully elucidate the underlying decision-making process. Sequential sampling models, such as the Drift Diffusion Model (DDM), inspired by neuronal evidence accumulation, provide a more comprehensive process model of decision-making [2]. In our research, we employed both the drift-diffusion model and readiness potentials (RPs) derived from EEG to delve into how anxiety or depression influences decision-making processes. Our investigation comprised two distinct studies: i) a behavior-only study with 202 participants, and ii) an EEG study with 42 participants. In the behavioral study, Prospect Theory analysis yielded inconclusive results. However, the DDM analysis revealed critical insights: despite no significant change in response bias, there was a noticeable decrease in evidence accumulation for gains (gain drift) and losses (loss drift) with increasing depression scores. This trend suggests a diminished motivation for rewards in individuals experiencing higher levels of depression, a finding consistent across various experimental blocks. In the EEG study, the behavioral analysis mirrored the earlier results; heightened levels of anxiety and depression correlated with a decrease in both gain and loss-seeking behaviors. An analysis of the EEG data across all trials and participants indicated that readiness potentials typically commenced at around 837 ms. This temporal marker provides a valuable neurophysiological correlate to the behavioral changes observed. We are currently in the process of integrating and comparing these neural findings with behavioral data to establish a more comprehensive neural basis for these claims.

3-140. RLegans: using reinforcement learning to model C. elegans neural activity during complex behavior

Andrew Warrington[1,2] AWARRING@STANFORD.EDU
Vladislav Susoy[3] VLADISLAV_SUSOY@FAS.HARVARD.EDU
Aravinthan Samuel[3] SAMUEL@G.HARVARD.EDU
Scott Linderman[1,4] SCOTT.LINDERMAN@STANFORD.EDU

[1]Stanford University
[2]Department of Statistics
[3]Harvard University
[4]Statistics Department and Wu Tsai Neurosciences Institute

A key goal of systems neuroscience is understanding how recurrent neural circuits control behavior. Studying simulated embodied agents, so-called virtual neuroethology, is an exciting new approach toward this goal [1, 2, 3]. We apply these concepts to study the C. elegans mating behavior. This choice is important for two reasons: Firstly, direct correspondence can be drawn between simulated and real agents due to the known and simple C. elegans biology. Secondly, mating is the most complex behavior worms perform [4], requiring the male to integrate multiple noisy sensory cues in real-time to select context-dependent actions. These factors mean our approach can offer insights into sensory processing and integration, internal state dynamics, decision-making, and motor programs in real specimens. We first develop a mechanistic simulation of the courtship phase of mating. The simulated male "brain" is parameterized by a recurrent neural network policy, which combines an internal neural state with environmental stimuli to update the internal state and select motor actions. These stimuli are partial and biophysically plausible observations of the environment. The network is trained using reinforcement learning.

We then examine the learned simulations. Learned behaviors resemble real behaviors, including searching, reorientation, and backward scanning. Dimensionality reduction and clustering reveal that simulated neural activity is

low-dimensional and that behavior within inferred discrete states is stereotyped, as is observed in real neural activity. Transitions between inferred states recapitulate the sequence of discrete behaviors exhibited by real worms. Looking forward, as we imbue synthetic worms with more biological realism, the simulated activity will become more realistic. This will facilitate simultaneous quantitative in silico experimentation and direct comparison to real data. The goal-oriented nature means concrete, behavioral-level questions can be quantitatively answered by analyzing the statistics of simulations under different conditions. Finally, in silico modifications can be used to investigate in vivo testable hypotheses.

3-141. Explainable Artificial Intelligence (xAI) to Identify Biomarker for Deep Brain Stimulation for Depression

Sankar Alagapan[1] SANKAR.ALAGAPAN@GATECH.EDU
Christopher Rozell[1,2] CROZELL@GATECH.EDU
Elif Ceren Fitoz[1] CERENFITOZ@GATECH.EDU
Martijn Figee[3] MARTIJN.FIGEE@MSSM.EDU
Mosadoluwa Obatusin[3] MOSADOLUWA.OBATUSIN@MSSM.EDU
Ki Sueng Choi[3] KISUENG.CHOI@MSSM.EDU
Stephen Heisig[3] STEPHEN.HEISIG@MSSM.EDU
Tanya Nauvel[3] TANYA.NAUVEL@MSSM.EDU
Allison Waters[3] ALLISON.WATERS@MSSM.EDU
Robert Butera[1] RBUTERA@GATECH.EDU
Patricio Riva Posse[4] PRIVAPO@EMORY.EDU
Helen Mayberg[3] HELEN.MAYBERG@MSSM.EDU

[1]Georgia Institute of Technology
[2]Electrical and Computer Engineering
[3]Icahn School of Medicine at Mount Sinai
[4]Emory University

Deep brain stimulation (DBS) of the subcallosal cingulate cortex (SCC) has been effective in treating patients with treatment-resistant depression (TRD). However, there is limited understanding of the effect of DBS on local circuit activity that accompanies recovery. Identifying these changes in brain activity that lead to stable recovery will provide an objective marker in a disorder that lacks such biological endpoints to target. The main requirements for the biomarker are 1) to track the clinically relevant changes in symptom severity and 2) to respond to changes in DBS settings. Local field potential (LFP) from SCC was acquired from 14 participants across 2 cohorts during 24 weeks of SCC DBS therapy using implanted DBS devices. Spectral features extracted from the recordings of 6 participants in cohort 1 were used to train a neural network classifier of 'sick' and 'stable response' states. Generative causal explainer (GCE), a novel xAI approach to decode a biomarker intervening weeks. This LFP-state marker was compared against a state marker derived from a gold-standard clinical assessment (Hamilton Depression Rating Scale; HDRS). The LFP classifier was able to differentiate the 'sick' and 'well' stages of DBS treatment (Area under ROC curve: 0.87$\pm$ 0.09) in cohort 1. The LFP-state marker predicted the corresponding HDRS-state with high accuracy (0.82$\pm$ 0.13) in cohort 1 and with lower accuracy in cohort 2 (0.69$\pm$ 0.23). In addition, the LFP marker predicted relapse in a participant whose LFP data was not used for training the classifier or the xAI model, suggesting the marker generalizes beyond the training dataset. Notably, the marker was significantly affected by changes in stimu-lation dose. The results suggest that LFP-marker satisfies two important requirements for a biomarker. Thus, it may be a po-tential biomarker for SCC DBS, which could enhance the scalability of DBS and reduce variability in outcomes.

3-142. Optimal control of spiking neural networks

Tiago Costa[1,2] TIAGO.COSTA@NEURO.FCHAMPALIMAUD.ORG
Juan R. Castineiras[1] JUAN.CASTINEIRAS@NEURO.FCHAMPALIMAUD.ORG
Alfonso Renart[1] ALFONSO.RENART@NEURO.FCHAMPALIMAUD.ORG

[1]Champalimaud Foundation
[2]Champalimaud Research

Control theory provides a natural language to describe flexible cognitive tasks such as covert attention or brain-machine interface (BMI) experiments, which require finding adequate inputs to a local circuit in order to steer its dynamics in a context-dependent manner. In optimal control, the target dynamics should maximize a notion of long-term value along trajectories, possibly subject to control costs. Because this problem is, in general, not tractable, current approaches to the control of networks mostly consider simplified settings (e.g., the Linear-

Quadratic Regulator). Here, we present a mathematical framework for optimal control of recurrent networks of spiking neurons with low-rank connectivity. An essential ingredient is a control-cost that penalizes deviations from the default dynamics of the network (specified by its recurrent connections) through a regularization parameter beta. We derived a Bellman Equation that specifies a Value function over the low-dimensional network state (LDS), and a corresponding optimal control input. The optimal control law takes the form of a feedback controller that provides external excitatory (inhibitory) synaptic input to neurons in the recurrent network if their spiking activity tends to move the LDS towards regions of higher (lower) Value. As an initial challenge, we study the problem of steering the state of the network towards particular terminal regions which can lie either in or out of regions in the LDS with slow dynamics, in analogy to a standard BMI experiment. The regime where control-costs are large is the most interesting. Here, the optimal control attempts to exploit the default dynamics as much as possible, leading to idiosyncratic trajectories that don't go directly to the target, and which would be difficult to interpret from the performance costs of the problem alone. Our results provide the foundation of a novel approach with broad applicability that unifies bottom-up and top-down perspectives on neural computation.

3-143. Encoding and decoding of semantic content during language comprehension at single cell resolution

Mohsen Jamali[1] MJAMALI@MGH.HARVARD.EDU
Benjamin Grannan[2,3] BGRANNAN@UW.EDU
Jing Cai[4,5] JCAI5@MGH.HARVARD.EDU
Ziv Williams[1] ZWILLIAMS@MGH.HARVARD.EDU

[1]Massachusetts General Hospital, Harvard Medical School
[2]University of Washington Medical Center
[3]Neurological Surgery
[4]Massachusetts General Hospital
[5]Neurosurgery

Humans can communicate exceptionally diverse and complex meanings through language. Yet despite a growing understanding of the frontotemporal brain regions that supports semantic processing, the precise derivation of linguistic meaning in neural tissue at the cellular level and over the temporal scales of action potentials remains largely unknown. Little is also known about how sentence context influences word meaning representations or what their dynamic may be. Here, by recording the activities of prefrontal neurons as participants listened to diverse linguistic materials, we find a fine-scaled parcellation of semantic information and language-selectivity by individual neurons. Specifically, utilizing word embedding vectors and by employing a modeling-and-decoding approach on both single- and population-level neuronal data, we find neurons displaying preferential responses to specific word meanings. The responses of these neurons were dynamic, reflecting the meaning of words based on the specific linguistic context in which they were heard and independently of their phonetic form. Modeled collectively, the neuronal ensembles reliably predict the words' meanings based on their corresponding domains. Importantly, these semantic domains can be accurately decoded from neural activities obtained from entirely different story narratives; together suggesting that these meaning representations are robustly reflected within the population's response patterns. Using a pretrained recurrent neural network we find a direct relationship between the decoding accuracies and the predictability of words based on surprisal analysis, further supporting the context-dependency of the semantic representations. Finally, we show that the combined activity patterns also reflect the hierarchical structure of meaning representations. Together, these findings reveal a remarkably detailed organization of semantic information by prefrontal neurons in humans and begin to illuminate the cellular-level representation and processing of linguistic meaning during language comprehension.

3-144. An opponent striatal circuit for distributional reinforcement learning

Adam Lowet[1] ADAM.LOWET@GMAIL.COM
Melissa Meng[1] MELISSAMENG@COLLEGE.HARVARD.EDU
Sara Matias[1] SARAMATIAS@G.HARVARD.EDU
Qiao Zheng[2] ZHENGQIAO86@GMAIL.COM
Jan Drugowitsch[2] JAN_DRUGOWITSCH@HMS.HARVARD.EDU
Naoshige Uchida[1,3] UCHIDA@MCB.HARVARD.EDU

[1]Harvard University
[2]Harvard Medical School
[3]Department of Molecular and Cellular Biology

Machine learning research has achieved large performance gains on a wide range of tasks by expanding the

learning target from mean rewards to entire probability distributions of rewards — an approach known as distributional reinforcement learning (RL). The mesolimbic dopamine system is thought to underlie RL in the mammalian brain by updating a representation of mean value in the striatum. However, little is known about whether, where, and how neurons in this circuit encode information about higher-order moments of reward distributions. To fill this gap, we used Neuropixels probes to acutely record striatal activity from well-trained, water-restricted mice performing a classical conditioning task in which reward mean, reward variance, and stimulus identity were independently manipulated. In contrast to traditional RL accounts, we found robust evidence for abstract encoding of variance in the striatum, which was disrupted by chronic ablation of dopamine inputs. Two-photon calcium imaging revealed that the two major classes of striatal medium spiny neurons — D1 and D2 MSNs — contributed to this code by preferentially encoding the right and left tails of the reward distribution, respectively. We synthesize these findings into a new model of the striatum and mesolimbic dopamine that harnesses the differential specializations of D1 and D2 MSNs to reap the computational benefits of distributional RL. Overall, this work provides a novel, normative perspective on widespread observations of opponency within the striatum while drawing closer connections between artificial and natural intelligence.

3-145. Optimistic and pessimistic beliefs define choice values under ambiguity.

Willa Kerkhoff[1,2] WGK11@PITT.EDU
William R Stauffer[1] WRS@PITT.EDU

[1]University of Pittsburgh
[2]Neurobiology

Adaptive decision making requires management of uncertainty. Economic theory recognizes two forms of uncertainty: risk, when potential outcomes and probabilities are known, and ambiguity, where the probabilities are not known. We and others have described the behavioral and neural computations that determine risk attitudes. However, the computational mechanisms driving ambiguity attitudes remain poorly understood because they require fundamentally distinct computational elements. Whereas risky uncertainty can be calculated, from experience or instruction, ambiguous uncertainty cannot be calculated. Thus, decisions under ambiguity rely on subjective probability judgements, i.e., beliefs. Here, we present a nonhuman primate (NHP) task for estimating those beliefs: the method of distributional equivalents. Symbolically informative visual stimuli used a two-dimensional scale to independently indicate reward magnitudes and probabilities (Fig. A). We created conditions of ambiguity by masking the probability dimension (Fig. A). Each trial paired ambiguous prospects in one of three reward magnitude ranges (low, medium, and high) with risky prospects from the same magnitude range (Fig. B, C). We define the distributional equivalent as the risky prospect that is chosen with equal frequency as an ambiguous gamble of the same reward magnitude range. Using this method, we find that distributional equivalents diverge significantly across the tested reward magnitude ranges. Specifically, NHPs pessimistically estimate probability distributions when the reward range is high, and they optimistically estimate probability distributions when the range is low. In other words, the NHPs hold inconsistent beliefs about ambiguous probabilities depending on the reward magnitudes on offer (Fig. D, E). These inconsistent beliefs drive diverse ambiguity attitudes. We are currently investigating how pharmacological manipulation of sympathetic and parasympathetic drive modulates these beliefs and trying to discover the neurophysiological correlates of irrational economic beliefs. These efforts will reveal the biological and neurocomputational basis of adaptive and maladaptive choice behaviors.

3-146. A link between the memory trace in motor cortex and savings

Juliana Couras[1,2] JCS237@PITT.EDU
Emily Oby[1] EMO22@PITT.EDU
Asma Motiwala[3] AMOTIWALA@CMU.EDU
Sam Snyder[1,2] SES253@PITT.EDU
Darby Losey[3] LOSEYDM23@GMAIL.COM
Jay Hennig[4] JAY.A.HENNIG@GMAIL.COM
Byron Yu[3] BYRONYU@CMU.EDU
Steven Chase[3] SCHASE@ANDREW.CMU.EDU
Aaron Batista[1] AARON.BATISTA@PITT.EDU

[1]University of Pittsburgh
[2]Center for Neuroscience
[3]Carnegie Mellon University
[4]Harvard University

Relearning a motor skill is faster than learning from scratch. This phenomenon, called savings, is a hallmark of motor learning, but its underlying neural mechanisms are not well understood. A candidate explanation for savings would be that after a bout of learning, neural activity has been altered to retain that learning. Here, we report a neural population correlate of savings in the motor cortex. Losey et al., 2022 recently detected a memory trace in the motor cortex. A memory trace is a change in neural activity, following learning, that encodes the learning experience. The memory trace is a result of the neural activity moving to a region in neural space that might facilitate control of the newly learned mapping. This led us to ask: does the memory trace lead to savings? We leveraged a brain-computer interface (BCI) learning paradigm to investigate the neural basis of savings. Each session started with a block of “intuitive control” in which the monkey’s natural neural population activity patterns allowed it to proficiently perform an eight-target center-out task. Then, for the first learning block, a novel BCI mapping was instantiated. After learning, the monkey performed a few hundred trials of “intuitive control” before the newly learned mapping was re-instantiated for a relearning block. We found that the early control during the relearning block showed significantly better cursor movements towards the target than the early trials of the learning block, i.e., "savings" occurred. Furthermore, we found the size of the memory trace was correlated with the magnitude of savings. Overall, our results suggest that the recently reported memory trace may be a neural substrate of savings.

3-147. Comparing neural representations with a metric that is sensitive to single-neuron tuning

Meenakshi Khosla[1,2] MEENAKSHIK1993@GMAIL.COM
Alex Williams[3] ALEX.H.WILLIAMS@NYU.EDU

[1]Massachusetts Institute for Technology
[2]Department of Brain and Cognitive Sciences
[3]New York University

Neuroscientists recurrently debate the importance of individual neural response profiles (“tuning curves”). In higher-order brain regions, neurons respond to complex combinations of sensory features (“mixed selectivity”), making the significance of tuning curves less clear. This has led to greater focus on population-level geometry– i.e. the structure of responses in N-dimensional firing rate space. But even if it is difficult to interpret mixed selective neurons, such tuning could be important to underlying computations. Thus, it is unclear whether “tuning matters” or “geometry is all you need." If tuning matters, then tuning profiles should be reproducible across subjects (even when they are mixed selective). There are a multitude of measures to compare population geometry across subjects—e.g., centered kernel alignment, Representational Similarity Analysis, and shape distance—but common approaches are invariant to rotations of the firing rate space and are therefore unable to distinguish different neural tuning profiles (which are rotation-sensitive). Thus, to resolve the debate, new metrics are needed that are sensitive to rotations but still invariant to the arbitrary ordering of neuron indices. The “permutation Procrustes” distance is one such metric, but it only works for networks with the same number of neurons due to its strict matching requirements. Here, we generalize this approach to networks of different sizes by identifying a “soft permutation" of tuning curves. We show this novel approach has desirable mathematical properties (the distance between representations is symmetric and obeys the triangle inequality), and avoids counter-intuitive behavior suffered by alternative approaches. Using this new metric, we demonstrate that tuning is preserved above chance levels across networks, challenging the stance that “geometry is all you need." Furthermore, in comparisons between artificial and biological networks, our new metric provides a more stringent test of similarity and we show that it differentiates between competing models better than existing approaches.

3-148. Chemotopy and chemical tuning in mouse olfactory bulb

Nikola Milicevic[1,2] NQM5625@PSU.EDU
Shawn Burton SDBURTON@GMAIL.COM
Michael Schmuker M.SCHMUKER@HERTS.AC.UK
Matt Wachowiak MATT.WACHOWIAK@UTAH.EDU
Vladimir Itskov[1] VLADIMIR.ITSKOV@PSU.EDU

[1]Pennsylvania State University
[2]Mathematics

Understanding the representation of chemical features in the olfactory bulb (OB) remains an elusive task. Unlike the stimuli of the other sensory modalities, the vast space of chemical features is poorly understood. This significantly complicates characterizing the receptive fields (tuning curves), analogous to those in the other sensory modalities. In addition to the lack of a natural coordinate system, the organization of chemical sensitivity across

the OB, often referred to as chemotopy, is also not known. To this end, we introduced an interpretable method of Shapley fields, an olfactory analog to retinotopic receptive fields. We used these computational tools to test the chemotopy hypothesis as well as probing diverse sets of chemical features using epifluorescence recordings of the mouse dorsal OB in response to stimuli across a wide range of the chemical space.

We found that Shapley fields reveal a weak ("mosaic") chemotopic organization of the chemical feature sensitivity on the surface of the dorsal OB. This organization is consistent across animals. Moreover, this chemotopic organization agrees across vastly different chemical feature sets: (i) the expert-curated PubChem database features and (ii) features derived from a Graph Neural Network trained on human olfactory perceptual tasks. Perhaps unsurprisingly, this suggests the presence of a "universal" chemical space between different mammalian species.

3-149. Here there be dragons no more - Mapping retinal cells with spatial biology and neural networks

Samuel Budoff[1,2] SABUDOFF@GMAIL.COM
Alon Poleg-Polsky[1] ALON.POLEG-POLSKY@CUANSCHUTZ.EDU

[1]University of Colorado
[2]Physiology and Biophysics

We address a critical gap in retinal neuroscience: the precise mapping of murine retinal cell types and their functional implications. Spurred by recent revelations in single-cell sequencing, which have clarified the array of retinal subtypes, and historic physiological insights, our study develops an innovative neural network strategy and methodologic approach to fill this gap in knowledge. The burgeoning understanding of retinal cell subtype spatial diversity and their influence on visual behavior shows that different retinal cell subtypes vary in their retinal distribution as well as their contribution to specific behaviors. For instance, the optic flow minimization observed in mice during hunting correlates with the high-density spatial cluster of Alpha-On sustained ganglion cells, implying a computational strategy encoded in the spatial distributions of such cells. Insights such as this underscore the necessity of detailed spatial maps to understand the retina's computational framework fully, and to predict other such behaviors. Moreover, knowledge of such distributions can be leveraged to develop novel CNN-like architectures optimized for naturalistic tasks. Unfortunately, the spatial distribution of most retinal cell subtypes is unknown. To address this need, we have utilized spatial sequencing technology, specifically the 10X Xenium platform, enabling us to map cellular distributions with unprecedented precision. Complementing this, our novel neural network approach employs gradient-based gene selection to minimize genes needed for cellular classification from transcriptomic data. The integration of spatial biology and machine learning in our study paves the way for a comprehensive understanding of retinal function. By creating detailed maps of cellular subtypes, our research illuminates how spatial variations contribute to the retina's complex computations and behavioral adaptations. This work not only fills a significant knowledge gap in retinal neurobiology, but also sets forth a new strategy for studies exploring the spatial properties of other neural tissues.

3-150. Dopamine cue responses encode a heading direction error with varying representations across the striatum

Eleanor Brown[1,2] EHBROWN@BU.EDU
Mai-Anh Vu[1] MAIANHVU@BU.EDU
Yihan Zi[1] DARCYZYH@BU.EDU
Chinyere Godfrey-Nwachukwu[1] CGODFRN@BU.EDU
Brian DePasquale[1,3] BDDEPASQ@BU.EDU
Mark Howe[1] MWHOWE@BU.EDU

[1]Boston University
[2]Graduate Program for Neuroscience
[3]BME

Dopamine release in the striatum is critical for learning associations between cues, actions and outcomes. Dopamine is well known to encode reward prediction errors, which compute changes in the expectation of an upcoming reward based on external stimuli. Recently however, dopamine signaling across the striatum has been found to vary in its encoding of locomotion, rewards, and stimuli. While these signals hint at distinct functional roles of dopamine across striatal subregions, how spatially varying dopamine release across the striatum contributes to learning and executing context-specific instrumental actions is poorly understood. Here, we used a novel multi-optical fiber array to record high-resolution spatiotemporal dopamine dynamics during a behavioral task which requires mice to run in a particular direction in response to arbitrary visual cues to receive a water reward. We found that dopamine release across the striatum encodes a heading direction error, which compares

the current heading-direction with the goal heading-direction, as indicated by the cue. When the cue indicates a turn in the same direction as the animal's current heading, there is an increase in dopamine release, as the mouse continues in its current direction. However, when the cue indicates a turn in the opposite direction, there is a dip in dopamine release as the mouse switches its running direction. Further, this dopamine heading direction error signal scales with the magnitude of the error. Using a finite impulse response model, we isolated this novel signal from a heading direction error independent cue response. A linear model, which identifies spatial gradients in their magnitude and timing, revealed that these two components are separated in both time and space. Together, these results expand upon our current understanding of what dopamine encodes, its potential behavioral functions, and how its 3-dimensional spatiotemporal dynamics are coordinated across the striatum.

3-151. Interplay between reset and nonlinearity drives metastability in networks of stochastic spiking neurons

Siddharth Paliwal[1,2] SIDDHARTHPALIWAL55@GMAIL.COM
Braden Brinkman[1,2] BRADEN.BRINKMAN@STONYBROOK.EDU
Gabriel Ocker[3,4] GKOCKER@BU.EDU

[1]Stony Brook University
[2]Neurobiology and Behavior
[3]Boston University
[4]Mathematics and Statistics

The firing rates as a function of mean membrane potential of single neurons in the primary visual cortex are fit well by power law nonlinearities. Observed exponents fall in the range of 2 ~5, larger than the rectified linear units often used to model neural firing in mathematically tractable models. Moreover, models that have investigated superlinear firing rates have done so only at a mean-field level, neglecting spiking fluctuations. A comprehensive understanding of how superlinear rectified neural firing rates couple with stochasticity impact the overall network activity is missing. In this work, we determine the phase diagram of the spiking network as a function of its synaptic connectivity and input current, and show that superlinear firing rates, a hard reset of the membrane potential after spiking, and stochastic firing together enable metastability between i) a quiescent and active state for subthreshold currents and ii) two active states for super-threshold currents. In the latter case the networks stochastically switch between these two active states. Using a stochastic field theory formalization of the spiking network model we can go beyond mean-field theory ("self-consistent 1-loop") to elucidate how the exponent of the firing rate nonlinearity, spiking fluctuations, and membrane potential reset influence the firing rate dynamics beyond the predictions of a mean-field treatment of the network dynamics.

3-152. Seizure susceptibility is related to task computations in recurrent neural networks

Ismaeel Ramzan[1,2] ISMAEELR@GMAIL.COM
Richard Sebastian Eydam[1,2] RICHARD.EYDAM@RIKEN.JP
Hannah Nemeth[3] NEMETH.RPI@GMAIL.COM
Louis Kang[1] LOUIS@LOUISKANG.GROUP

[1]RIKEN Center for Brain Science
[2]Neural Circuits and Computations Unit
[3]Rensselaer Polytechnic Institute

Recurrent brain circuits are thought to initiate seizures by pathologically amplifying neural activity. Undernormal conditions, they perform many different cognitive functions that require information and neuralactivity to be propagated in different ways. We hypothesize that these computational requirements in-herently influence the risk for a recurrent neural network to develop excessive activity under excitatoryperturbations. Networks trained on tasks with discrete, well-separated target values may be more resistantto seizure-like behavior because such separation requires robustness against perturbation and limits thespread of excessive activity. Those trained on tasks that produce continuous, smoothly-varying targets maybe more susceptible to seizure-like behavior because perturbations may easily drive activity amplificationacross closely spaced network states. To test this hypothesis, we train spiking recurrent neural networkswith identical architectures and similar baseline activity statistics to perform either a discrete classificationtask or a continuous autoencoding task. We then promote epileptogenesis in the trained networks by eitheradding a positive value to their recurrent weights or injecting excitatory currents to their neurons. We findthat the continuous-trained networks are more sensitive to these perturbations and exhibit higher firingrates at higher frequencies compared to the discrete-trained networks. This observed relationship betweentask attributes and excessive responses upon perturbation contributes a fundamentally computational

per-spective to our understanding of seizure initiation in biological networks.

3-153. Simultaneous brainwide recordings reveal a cortico-striatal subnetwork mediating perceptual choice

Adrian Bondy[1,2] ABONDY@PRINCETON.EDU
Julie Charlton[1] JC2988@PRINCETON.EDU
Thomas Luo[1,2] ZHIHAOL@PRINCETON.EDU
Sarah Jo Venditto[1,3] VENDITTO@PRINCETON.EDU
Wynne Stagnaro[1] WYNNES@PRINCETON.EDU
Charles Kopec[1,2] CKOPEC@PRINCETON.EDU
Carlos Brody[1,2] BRODY@PRINCETON.EDU

[1]Princeton University
[2]Princeton Neuroscience Institute
[3]Neuroscience

Making a choice based on noisy sensory inputs is thought to depend on neural activity across the brain, and widespread brain activity correlates with the decision process. While interactions between such brain regions are presumably critical for choice formation, the nature of these interactions is obscure. A major obstacle has been the difficulty of observing neural activity simultaneously across many brain regions at behaviorally relevant timescales. Here, we recorded spikes simultaneously from over 3,000 neuronal units bilaterally across 12 brain regions/hemisphere while rats performed a freely moving auditory evidence accumulation task. To measure the decision-related interactions among brain regions, we quantified each region's evolving representation of the internal decision process (the "decision variable"; DV), and examined their correlations. Importantly, we presented repeat trials with identical stimulus sequences, which enabled us to subtract the choice- and stimulus- conditioned mean DV. This allowed us to isolate moment-to-moment co-fluctuations in the decision process not directly driven by the sensory stimulus or the behavioral choice. The structure of the residual correlations revealed three anatomically interconnected frontal cortical and striatal brain regions–dorsomedial frontal cortex, primary motor cortex, and anterior dorsal striatum–whose decision dynamics are highly coupled. Along with findings from causal perturbation, this result suggests that these three frontal cortical and striatal regions form a subnetwork that mediate choice formation. To test this possibility, we used a recent method to infer from the neural activity and behavioral choice the moment within each trial at which the animal stops accumulating evidence and commits to a choice. We found that the correlations amongst the three regions abruptly diminished after the inferred moment of decision commitment. Our results suggest a frontal cortico-striatal subnetwork mediates the formation of perceptual choices and highlight the promise of large-scale simultaneous neural recording for uncovering brain-wide interactions underlying cognition.

3-154. Context shifts the geometry of representations to bias choices during perceptual decision-making

Ramon Nogueira Manas[1,2] RN2446@COLUMBIA.EDU
Saleh Esteki[3] ME1168@NYU.EDU
Stefano Fusi[4] SF2237@COLUMBIA.EDU
Roozbeh Kiani[3] RK97@NYU.EDU

[1]Columbia University
[2]Center for Theoretical Neuroscience
[3]New York University
[4]Columbia University, Zuckerman Institute

Decision-making in the natural world relies on swiftly combining sensory input with contextual information to optimize choices. The underlying computations involve learning a model of the world, inferring present context based on that model, and adjusting decisions in a context-dependent manner. The geometry of representations of sensory and contextual variables in a population of neurons can facilitate or hinder those computations, influencing generalization and flexible decision-making. Here, we investigate how sensory and context variables are jointly represented in the neural population state space during and after learning. We recorded large populations of lateral prefrontal neurons as monkeys performed a direction discrimination task with uncued context changes. In one context left choices were more rewarded than right choices, and in the other context the reward asymmetry was reversed. Monkeys adapted choice biases to maximize reward in each context. The geometry of the representations in each context could be described as a curve made of the points that represent different coherence levels. The two curves corresponding to the two contexts were shifted with respect to each other. The geome-

try also reflected a factorized code that enabled monkeys to perform task inference following the uncued context switches. By tracking behavior and neural activity during learning, we show increasingly faster transitions between context manifolds that matched the experience-dependent improvement in the speed of choice bias adjustments after context switches. Finally, RNNs trained to perform a simplified version of the task showed similar behavioral biases and captured the representational shifts observed in the experiments. Our findings indicate a factorized neural code in which context inference shifts decision variable manifolds to induce beneficial choice biases.

3-155. Beyond pulsed inhibition: alpha oscillations modulate awake resting-state network excitability

Fabrizio Lombardi[1] FABRIZIO.LOMBARDI@UNIPD.IT
Liborio Parrino[2] LIBORIO.PARRINO@UNIPR.IT
Silvia Scarpetta[3,4] SSCARPETTA@UNISA.IT
Anna Vaudano[5] ANNAVAUDANO@GMAIL.COM
Hans J. Herrmann[6] HANS.HERRMANN@ESPCI.FR
Dietmar Plenz[7] PLENZD@MAIL.NIH.GOV
Lucilla de Arcangelis[8] LUCILLA.DEARCANGELIS@UNICAMPANIA.IT
Oren Shriki[9] SHRIKIO@BGU.AC.IL

[1]University of Padova
[2]University of Parma, Italy
[3]University of salerno
[4]Dept of Physics
[5]University of Modena and Reggio Emilia, Italy
[6]Universitade Federal do Ceara
[7]NIH
[8]University of Campania "Luigi Vanvitelli"
[9]Ben-Gurion University of the Negev

Alpha oscillations are considered to reflect cortical inhibition. However, the nature of this inhibition, as well as its effects on collective neural dynamics, is still poorly understood. To address this question, we investigate collective neural activity during resting wake and NREM sleep, a physiologic state with marginal presence of alpha oscillations. We show that, during resting wake, alpha oscillations drive an alternation of attenuation and amplification bouts of neural activity. Our analysis indicates that inhibition is activated in pulses that last a single alpha cycle and gradually suppress neural activity, while excitation is successively enhanced over a few alpha cycles to amplify neural activity. Furthermore, we show that long-term alpha amplitude fluctuations—known as the “waxing and waning” phenomenon—are associated with an attenuation-amplification mechanism acting over the timescales of several seconds and described by a power law decay of the activity rate in the “waning” phase. Importantly, we do not observe such dynamics during NREM sleep with marginal alpha oscillations. The results suggest that alpha oscillations control the alternation of inhibition and excitation bouts across multiple timescales, the “waxing and waning” being a long-term control mechanism of cortical excitability. The amplification regime observed beyond the timescales of the individual alpha cycle indicates in turn that alpha oscillations might modulate the intensity of neural activity not only through pulses of inhibition, as proposed in the pulsed inhibition hypothesis, but also by timely enhancing excitation (or disinhibition).

3-156. Neurogenesis enhances olfactory coding efficiency in changing environments

Ryan McGee[1,2] RYANSMCGEE@GMAIL.COM
Gaia Tavoni[3] GAIA.TAVONI@WUSTL.EDU

[1]Washington University in St Louis
[2]Neuroscience
[3]Washington University in Saint Louis

In a few areas of the brain, such as the olfactory bulb and the dentate gyrus, neurons continue to be generated throughout the lifespan, leading to the replacement of older neurons within mature circuits. The role of neuronal replacement in the computations performed by these areas remains unclear. Here we focus on the effects of adult neurogenesis on sensory processing in the olfactory bulb (OB). The OB processes complex stimulus patterns and integrates them with contextual feedback from other brain regions involved in multi-sensory association, memory, and attention. Odor discrimination depends in part on the quality of pattern separation in the OB, which is modulated by local inhibitory interneurons (granule cells), many of which are generated in adulthood. New

neurons in the OB are believed to contribute to this pattern separation function, but the impact of neurogenesis has not been assessed (a) in terms of optimal coding benchmarks, or (b) in relation to synaptic plasticity. Through a combination of theory and computational modeling, we show that ongoing neurogenesis enhances the coding efficiency of olfactory circuits beyond what can be achieved with Hebbian plasticity alone, particularly in changing environments. That is, plausible mechanisms of neuronal replacement and integration facilitate population codes that better maximize the decodability of neural representations. In addition, we find that neuronal replacement itself can act as a learning mechanism, even in the absence of other forms of plasticity. Together, the effects of neurogenesis significantly improve the OB's ability to maintain high coding efficiency when stimuli statistics vary over time. These results connect normative predictions derived from optimal coding theory with biologically plausible processes, providing a foundation for further study of the role of neurogenesis in sensory computation.

3-157. Hard-wired early visual pathway enables continual learning under dynamic environments

Minjun Kang[1,2] 4401KMJ@KAIST.AC.KR
Gwangsu Kim[1] KGSPIANO@KAIST.AC.KR
Hyeonsu Lee[1,3] HSLEE9305@KAIST.AC.KR
Se-Bum Paik[1] SBPAIK@KAIST.AC.KR

[1]KAIST
[2]Department of Brain and Cognitive Sciences
[3]Department of Bio and Brain Engineering

Continuously learning new information is a crucial ability of animals but a challenging problem for deep neural networks (DNNs), which often undergo catastrophic forgetting (McCloskey, 1989) when trained on new data distributions. However, the specific mechanisms enabling continuous visual learning of brains remain unclear. Unlike DNNs, for which all layers are trained, the brain's early visual pathway is hard-wired and possesses innate selectivity for specific visual features (Godecke, 1997), implying that these inherent circuits are the general basis for processing various image domains during lifelong learning. Here, inspired by this notion, we demonstrate that fixing early layers of DNNs using filters resembling the receptive fields of the primary visual cortex (V1) enables networks to learn objects continuously in environments that change dynamically. We fixed the first layers using V1-inspired Gabor filters and sequentially trained networks with various domains. First, we found that networks with fixed layers maintain the performance of the initial domain even after being trained on a completely different domain with a distinct frequency spectrum, while conventional DNNs entirely forgot previous information. Notably, fixing randomly initialized early layers offered no such benefit, suggesting that fixing Gabor-like filter structures is the key. To assess general applications of fixed filters, we designed more realistic tasks whose domains change continuously without explicit boundaries (Zhang, 2022). We observed that fixing Gabor filters significantly improved training performance in dynamic environments. Finally, by measuring the correlations among encoded features before and after training the new domain, we confirmed that fixing early-layer filters enables higher layers to encode features that remain largely unchanged throughout continual learning. Taken together, our results suggest that fixed Gabor filters in early layers could be key architectures for continual learning in dynamic environments, highlighting the functional significance of hard-wired early visual pathways in higher mammals.

3-158. Neurocomputational characterisation of differences in multisensory processing in Autism and Schizophrenia

Amirreza Nadimi Shahraki[1,2] SCANS@LEEDS.AC.UK
Ioannis Delis[1] I.DELIS@LEEDS.AC.UK
Maida Toumaian[3] MAIDA.TOUMAIAN@GMAIL.COM
Nikolaos Smyrnis[3] SMYRNIS@MED.UOA.GR

[1]University of Leeds
[2]Faculty of biological sciences
[3]University Mental Health, Neurosciences and Precision Medicine Research Institute "COSTAS STEFANIS"

Multisensory integration, yielding faster and less variable responses than uni-sensory cues, is impaired in individuals with developmental differences compared to neurotypicals. Neuroimaging, particularly electroencephalogram (EEG) and event-related potential (ERP) analysis, are employed to recognize these disparities. However, inefficient signal processing during transformation and interpretation hinders optimal information preservation thus as a non-invasive, cost-effective tool we must reconsider our approach. This study explores computational approaches to elucidate neural representations in Autism Spectrum Disorder (ASD) and Schizophrenia (SZ) compared to neurotypicals (CN). Due to the lack of absolute evidence based on past literature in characterizing these differences,

innovative strategies are employed here to enhance understanding and address gaps in distinguishing ASD and SZ. Significant effects were observed for population (p < 0.01) and sensory condition (p << 0.01) on RTs, with CN displaying the fastest responses, followed by ASD. AV conditions consistently yielded faster responses across groups, indicating multisensory behavioural gain. EEG tensor decomposition revealed audiovisual integration components in the central parietal lobe and parieto-occipital regions, showing consistent linear correlations with audio-visual (AV) cues during [300-500] ms across trials. In CN, the anterior right temporal lobe uniquely correlated with AV integration, particularly when presented with the least uni-sensory complementary information, activating parieto-occipital regions when most uni-sensory information was presented at [300-400] ms. Deficits in central AV integration in ASD and SZ did not appear in early electro-cortical responses, indicating late processing stage involvement, potentially in response selection. However, reduced uni-sensory cue involvement led to slower and wider AV peaks, and in the presence of uni-sensory cues, central parietal lobe activation appeared non-existent. This phenomenon, identified as a lack of AV integration facilitation, is common in ASD and SZ populations. This provides computational evidence in line with the behavioural analysis to complement each population's sensory profile.

3-159. The representational geometry of emotional states in the basolateral amygdala

Pia-Kelsey O'Neill[1] PTO2102@COLUMBIA.EDU
Lorenzo Posani[1,2] LP2878@COLUMBIA.EDU
Jozsef Meszaros[1] JM3648@COLUMBIA.EDU
Phebe Warren[1] PW2490@COLUMBIA.EDU
Carl Schoonover[1] CES2001@COLUMBIA.EDU
Andrew Fink[1,3] AF2243@COLUMBIA.EDU
Stefano Fusi[4] SF2237@COLUMBIA.EDU
C. Daniel Salzman[1] CDS2005@COLUMBIA.EDU

[1]Columbia University
[2]Center for Theoretical Neuroscience, Zuckerman Institute
[3]Neuroscience
[4]Columbia University, Zuckerman Institute

Emotional responses to salient events can generate multiple behaviors, such as freezing or fleeing in response to threats. The question of whether the brain uses distinct neural mechanisms to represent these different behaviors is an open problem. The basolateral amygdala (BLA) is a fundamental region for connecting emotionally relevant stimuli to the appropriate behavioral response [1]. Thus, a natural question is whether and how neurons in the amygdala encode external stimuli, the associated emotional valence, and the behavioral states reflecting these emotions. To address these questions, we analyzed the coding properties of BLA neurons in mice engaged in a behavioral paradigm, the virtual burrow assay [2], which allows detailed characterization of these multiple variables: subjects are presented with aversively conditioned or neutral stimuli, to which they can respond with different defensive behaviors, including trembling (freezing) and ingressing into the burrow (fleeing to safety). We used a linear decoding approach to analyze the population activity of BLA neurons in response to (1) stimulus identity, (2) stimulus valence, (3) trembling behavior, and (4) ingressed state. First, we found that these variables are decodable from population activity with high accuracy. We then analyzed the selectivity properties of individual neurons by comparing their contribution to population decoding across different variables. We found that stimulus identity, valence, tremble, and ingress were encoded with mixed selectivity, challenging the notion that the amygdala represents different variables with dedicated, specialized subpopulations. Finally, we found that the representational geometry of valence during tremble and after ingressing were markedly different. While valence and trembling were represented in a low-dimensional, disentangled geometry [3], valence was not even decodable in the ingressed state. In fact, ingressing in the burrow modulated the conjunctive coding of valence and state, forming a geometry that indicates an abstract representation of safety.

3-160. Motion adaptation induced object position bias in macaque IT and SlowFast video recognition models

Elizaveta Yakubovskaya[1,2] EYAKUB@MY.YORKU.CA
Hamidreza Ramezanpour[1] HAMIDRAM@YORKU.CA
Sara Djambazovska[1] DJAMBAZOVSKA@GMAIL.COM
Kohitij Kar[1,2] K0H1T1J@YORKU.CA

[1]York University
[2]Biology

To efficiently interact with their environment, primates excel in not just recognizing objects ('what') but also discerning their spatial attributes ('where'). This dual capacity, traditionally attributed to functional segregation of ventral and dorsal visual processing pathways, is currently being reexamined in light of emerging evidence. Recent work of Hong et al. (2016) revealed the macaque inferior temporal (IT) cortex's role in encoding object positions. Our study further ventures into this relatively uncharted territory with three main objectives: firstly, to extend the findings of Hong et al., assessing how scaling of neural recording sites in IT influences object-position estimates; secondly, to investigate the impact of motion adaptation (a phenomenon typically associated with dorsal stream) on these estimates; and thirdly, to evaluate whether existing ventral stream models align with our observations. We performed large-scale recordings across IT cortex of 3 monkeys (~500 sites). Monkeys passively fixated Test images (640; 1 of 8 objects, varying latent parameters, embedded in naturalistic backgrounds). Indeed, we observed highly accurate (Pearson R >0.7) IT-population based linear decodes of object positions. Next, to test whether motion-direction adaptation biases position estimates, we preceded the Test image presentation by prolonged (3000 ms) oriented gratings moving in one of four directions. Remarkably, IT-based (192 sites) position decodes showed a significant bias (p<0.0001; permutation test) in the direction opposite to the preceding motion. These biases align with perceptual reports, suggesting that the IT cortex represents perceptual rather than ground-truth positions. Interestingly, simulating the experiments in-silico on SlowFast networks (video recognition model with ResNet-50 backbone) demonstrated a similar bias (absent in vanilla ResNet-50 with scaled activation mimicking neural fatigue). Our findings introduce a framework for probing how dorsal-ventral interactions could generate adaptation after-effects and a model-based hypotheses space to guide the exploration of computational mechanisms critical for dynamic scene perception.

3-161. Emergence of illusory contours in robust deep neural networks by accumulation of implicit priors

Tahereh Toosi TAHERE.TOOSI@GMAIL.COM
Kenneth D. Miller KEN@NEUROTHEORY.COLUMBIA.EDU

Columbia University

Deep neural networks (DNNs), trained for object recognition, exhibit similarities to neural responses in the monkey visual cortex and are currently considered the best models of the primate visual system. It remains unclear whether psychophysical effects, such as illusory contours perceived by humans, also emerge in these models. Utilizing the invertibility properties of robustly trained feedforward neural networks, we demonstrated that illusory contours and shapes emerge when the network integrates its learned implicit priors. It is believed that our visual system stores perceptual priors, with visual information learned and embedded in neural connections across all visual areas. This stored information is harnessed when required, for instance, during occlusion resolution or visual imagination generation. While the significance of feedback connections in these processes is well recognized, the precise neural mechanism that aggregates dispersed information throughout the visual cortex remains elusive. In this study, we leverage a ResNet50 neural network, conventionally used in image recognition, to shed light on the neural basis of illusory contour perception through its inherent feedback mechanism during error backpropagation. By iteratively accumulating the gradients of the loss with respect to an input—a Kanizsa Square—within an adversarially trained network, we observed the emergence of edge-like patterns in the area of the perceived 'white square'. This process, which unfolds over multiple iterations, echoes the time-dependent emergence of illusory contours in the visual cortices of both rodents and primates as seen in experimental studies. Notably, the ResNet50 employed in this study was neither specifically enhanced with feedback capabilities nor optimized to detect or decode these illusory contours; it was merely trained for robust object recognition against adversarial examples. These findings highlight a compelling parallel, suggesting that the ability to perceive contours might be an incidental consequence of the network's ability to handle adversarial noise during its training regime.

3-162. XFADS: Predicting single-trial cued behavior solely from preparatory activity

Matthew Dowling[1] MATTHEW.DOWLING@STONYBROOK.EDU
Yuan Zhao[2] YUAN.ZHAO@NIH.GOV
Memming Park[3,4] MEMMING.PARK@RESEARCH.FCHAMPALIMAUD.ORG

[1]Stony Brook University
[2]National Institute of Mental Health
[3]Champalimaud Foundation
[4]Champalimaud Centre for the Unknown

Consider a classic neuroscience dataset — comprising recordings from monkey's motor cortex during a center-

out reaching task. A prevailing hypothesis elucidating the coordinated neural population dynamics in this task suggests that the initial condition of a dynamical system, steering muscle movement, is set during the preparatory phase of the trial. The motor cortex, released from the initial condition and governed by the dynamical system, controls the muscles orchestrating the monkey's movement. However, existing cutting-edge tools often rely on extraneous-input-driven dynamical systems, which tend to learn control inputs anti-causally or lack a dynamic systems model entirely. This led us to explore the feasibility of learning the hypothesized autonomous dynamical system, exclusively from neural data, that is capable of causally predicting, initiated before movement onset, monkey's reaching. Consequently, we developed a novel Bayesian inference framework, eXponential FAmily Dynamical System (XFADS), aimed at learning highly expressive dynamical systems from neural data. Remarkably, we confirm the long standing autonomous nature of motor cortex activity associated with the planned reach behavior through single-trial causal forecasting. Specifically, our inferred stochastic dynamics revealed that the causally-inferred latent states, derived without any prior knowledge of the target or future behavior, were nearly as effective in predicting the monkey's arm velocity as the latent states utilizing the entire trial data.

3-163. The variability of representations in mice and humans changes with learning, engagement, and attention

Praveen Venkatesh[1] PRAVEEN.VENKATESH@ALLENINSTITUTE.ORG
Corbett Bennett[2] CORBETTB@ALLENINSTITUTE.ORG
Sam Gale[2] SAMG@ALLENINSTITUTE.ORG
Juri Minxha[3] JMINXHA@GMAIL.COM
Hristos Courellis[3] HRISTOS.COURELLIS@CSHS.ORG
Greggory Heller[4] GREGGH@MIT.EDU
Tamina Ramirez[5] TKR2116@CUMC.COLUMBIA.EDU
Severine Durand[2] SEVERINED@ALLENINSTITUTE.ORG
Ueli Rutishauser[3] UELI.RUTISHAUSER@CSHS.ORG
Shawn Olsen[2] SHAWNO@ALLENINSTITUTE.ORG
Stefan Mihalas[2] STEFANM@ALLENINSTITUTE.ORG

[1]Allen Institute + University of Washington
[2]Allen Institute
[3]Cedars-Sinai Medical Center
[4]Massachusetts Institute of Technology
[5]Columbia University

In responding to a visual stimulus, cortical neurons exhibit a high degree of variability, and this variability can be correlated across neurons. How this variability changes across multiple conditions has not been systematically studied. Such a systematic study can provide insights toward computational roles of variability. Here we use recordings from both mice and humans to systematically characterize how the variability in the representation of visual stimuli changes with learning, engagement and attention. We observe that in mice, familiarization with a set of images over many weeks reduces the variability of responses, but does not change its shape. Switching engagement, from passive to an active task which has been over-trained, changes the overall shape by shrinking the neural variability only along the task-relevant direction, leading to a higher signal-to-noise ratio. In a selective attention task in humans wherein multiple distributions are compared, a higher signal-to-noise ratio is obtained via a different mechanism, by mainly increasing the signal of the attended category. Overall, these findings show that representation variability changes with learning, and can be adjusted with task needs for over-trained tasks. A potential speculative role for variability, consistent with these findings, is that it helps control when to be specific or generalize, and which directions to generalize over.

3-164. Uncovering neural mechanisms of mental simulation by symbolically programming RNNs.

Daniel Calbick[1,2] DANIEL.CALBICK@YALE.EDU
Hansem Sohn[3] HANSEM@MIT.EDU
Mehrdad Jazayeri[4] MJAZ@MIT.EDU
Jason Kim[5] JK2557@CORNELL.EDU
Ilker Yildirim[1] ILKER.YILDIRIM@YALE.EDU

[1]Yale University
[2]Interdisciplinary Neuroscience Program
[3]Sungkyunkwan University
[4]Massachusetts Institute of Technology

[5]Cornell University

How is it that through the distributed and dynamic activity in our brain's neural circuits, we think thoughts about objects, mentally simulate how they will move and react to forces, and plan actions? We present a multi-level modeling framework that natively inter-operates in both cognitive hypotheses (simulatable object representations) and neural mechanisms (distributed codes and dynamic attractors). Using this new framework, we reverse-engineer a "mental simulation circuit" in prefrontal cortex populations of macaques playing the video game pong.

In computational neuroscience, deep neural networks (DNNs) have made strides by providing models mappable to neural data. DNNs offer scientific hypotheses with respect to a model's architecture, training data, and training objective, while obscuring the algorithms and representations at play. We introduce a different approach that directly encodes explicit algorithmic and representational hypotheses into the weights and dynamics of recurrent neural networks (RNNs) and test them against neural data.

We do so by extending a recent approach in the physics of dynamical systems that allows us to specify a cognitive hypothesis—the approximate dynamics of the game of pong, in terms of objects and surfaces— and program a biologically plausible RNN to encode these dynamics and perform goal-directed control.

We compare prefrontal cortex activity recorded in a pair of macaques to the dynamics of these RNNs under the same set of pong trials. We find that the hidden states of programmed RNNs explain substantial variance in neural data. Beyond a quantitative alignment, programmed RNNs recapitulate a key non-linearity in prefrontal populations that machine-learning style trained RNNs don't. These results provide support for the multi-level hypothesis of an object-based mental simulation system implemented in the neural machinery of the prefrontal cortex.

This multi-level modeling framework holds significant promise for computational neuroscience, offering new ways to reverse-engineer fundamental neural mechanisms underpinning mental life.

3-166. Causal inference in a target interception task: optimal strategies and neural circuits

John Vastola[1,2] JOHN_VASTOLA@HMS.HARVARD.EDU
Valentina Vencato[3] VV2247@NYU.EDU
Jean-Paul Noel[3] JPN5@NYU.EDU
Gregory DeAngelis[4] GDEANGELIS@UR.ROCHESTER.EDU
Dora Angelaki[3] DA93@NYU.EDU
Jan Drugowitsch[1] JAN_DRUGOWITSCH@HMS.HARVARD.EDU

[1]Harvard Medical School
[2]Neurobiology
[3]New York University
[4]University of Rochester

Whether motion across the retina is due to object motion, self-motion, or a mix of both is often ambiguous. Resolving this ambiguity is important, since animals may have to react differently depending on whether, e.g., a predator is moving towards them or not. We study the required disentangling of self- and object-motion, a form of dynamic causal inference, in the context of a novel navigation task for both humans and macaque monkeys. In each trial, subjects in a virtual reality environment first briefly observe a target that may or may not be moving, and they may also experience self-motion. After that, the target becomes invisible and subjects must attempt to intercept it by using a joystick to steer through the virtual environment. We find that the behavior of humans and one monkey is consistent with the predictions of a normative model which couples (i) a Bayesian strategy for target trajectory inference with (ii) near-optimal navigation that minimizes a certain set of costs (including movement costs and proximity-to-target costs). In particular, major qualitative features of the behavioral data, like a biased assessment of target motion given self-motion, are consistent with the model, and quantitative features like subjects' trajectories can be fit reasonably well. Given that the model appears to describe behavioral data well, we also construct a biologically plausible neural circuit model that suggests how our proposed strategy might be implemented by the brain. The circuit model, which couples several probabilistic population codes, is consistent with ideas about how causal inference might unfold in cortex: primary sensory cortices reflect estimates according to each possible causal structure, which are then combined in intermediate areas (e.g., posterior parietal cortex), and used to drive decisions in downstream areas.

3-167. Tuning diversity creates efficient neural representations

Sonica Saraf[1,2] SS6786@NYU.EDU
J. Anthony Movshon[3] MOVSHON@NYU.EDU
SueYeon Chung[1,4] SCHUNG@FLATIRONINSTITUTE.ORG

[1]New York University; Flatiron Institute
[2]Center for Neural Science
[3]New York University
[4]Center for Computational Neuroscience

Neurons exhibit a wide variety of selective tuning properties. Why this diversity is advantageous for neural coding has intrigued researchers for decades. One question is whether heterogeneous neuronal populations carry more information than homogeneous ones. Here, we develop a theory of how tuning heterogeneity affects decoding efficiency. Our theory connects the distribution of neuronal tuning properties to the representational geometry that governs the information capacity of the neuronal population. We then test the theory with simulated and real neuronal data. Specifically, we consider heterogeneity in the tuning curve's peak to trough amplitude, and in its bandwidth – the range of stimulus values for which the response exceeds half maximum. Our theory shows that diversity in amplitude and bandwidth increases decoding efficiency when compared to homogeneous populations. Heterogeneity in both properties leads to increased capacity in the population via higher separability of the representations. Amplitude heterogeneity serves to push the centers of the representations further apart, while bandwidth heterogeneity decreases the correlations of the centers. Supporting our theory, we observed this effect in simulated populations and biological neural populations recorded from macaque primary visual cortex (V1). By studying representational geometry, we overcome limitations of previous analytical studies of diversity's effect on coding. For example, with Fisher information one can measure discrimination between pairs of stimuli by analyzing tuning curves, whereas our measures can also apply to one versus many discrimination tasks and do not require knowledge of tuning curves. Lastly, this belongs to our broader effort to connect neuronal tuning properties to representational geometry. Representational geometry is generally used to study population coding in areas where single neuron properties are not well understood. Our work bridges the gap between the analytical approaches that rely on single neuron properties and those utilizing population-level geometry.

3-168. Neural correlates of multidimensional feature tracking and choice behavior in macaque

Patrick Zhang[1,2] PQZ317@UW.EDU
Elizabeth Buffalo[3] EBUFFALO@UW.EDU
Edgar Walker[1] EYWALKER@UW.EDU
Adrienne Fairhall[4] FAIRHALL@UW.EDU
John Ferre[4] JBFERRE@UW.EDU
Michael Jutras[4] MJUTRAS@UW.EDU

[1]University of Washington, Seattle
[2]Neuroscience
[3]Physiol. and Biophysics, University of Washington; Washington Natl. Primate Res. Ctr., Seattle, WA
[4]University of Washington

To explore how monkeys use internal models for complex tasks to draw inferences via trial and error, we modeled behavioral data from animals performing a multidimensional choice task. Subjects must learn to identify which of several visual features is rewarded, while the rewarded feature switches without cue. We compared a suite of published models and found that a reinforcement learning model, with a novel extension that accounts for a specific anomalous behavior, provides the best fit to behavior. We used this model to guide analysis of neural spiking data recorded from over 200 electrodes across the monkey's brain during trained task performance. We found representations of visual features and of reward prediction error predicted by the model, and we show that during feedback, the population encodes features and RPE in an overlapping subspace.

3-169. Neural Circuit Underlying Economic Decisions: Insights from a Computational Model

Aldo Battista[1,2] AB9710@NYU.EDU
Xiao-Jing Wang[1] XJWANG@NYU.EDU
Camillo Padoa-Schioppa[3] CAMILLO@WUSTL.EDU

[1]New York University
[2]Center for Neural Science
[3]Washington University in St. Louis

Economic choice is a fundamental aspect of neuroeconomics, encompassing decisions based on subjective preferences for goods varying in type, quantity, and probability. The Orbitofrontal Cortex (OFC) is crucial for encoding subjective values that underlie these choices. We propose a novel model using excitatory-inhibitory vanilla recurrent neural networks (RNNs), trained via reinforcement learning to perform a spectrum of real-world choice tasks, surpassing traditional models in task complexity and biological plausibility. Our networks replicate behavioral outcomes consistent with experimental observations, including biases such as risk attitude and order preference, providing a platform to investigate neural circuits responsible for these biases. Single-neuron analysis within our RNNs identifies neuron types – offer value, chosen value, choice – mirroring those found in OFC, with temporal dynamics reflecting the decision process. At the connectivity level, we predict a low-rank structure. This suggests low-dimensional population dynamics where decision variables explain most of the variance, with the chosen value encoded as a line attractor. The analysis of the input connectivity shows that the offer value is computed upstream of the recurrent network by approximating the product of the goods' features. A reduced model of the recurrent circuit reveals a Winner-takes-all mechanism to implement the decision-making process. Task variance analysis of individual units shows functional specialized and shared neuron clusters for different decision tasks. Notably, a distinct group for the sequential rule suggests a dedicated neural mechanism for working memory functions. Furthermore, the analysis of rule subspaces indicates a shared neural substrate for all but the sequential task, supporting the concept of a common decision-making schema. Indeed, we observed that a curriculum learning protocol speeds up the training, confirming a shared computational mechanism across tasks. These results set the groundwork for future empirical validations with neural data and advance our understanding of the neural circuits involved in economic choices.

3-170. Learning robust neural representations by straightening natural videos

Xueyan Niu[1,2] XN314@NYU.EDU
Cristina Savin[1] CSAVIN@NYU.EDU
Eero P. Simoncelli[3] EERO.SIMONCELLI@NYU.EDU

[1]New York University
[2]Center for Neural Science
[3]NYU and Flatiron Institute

It has long been debated whether biological vision is optimized for predicting incoming sensory stimuli, or for discriminating patterns and objects. Are the two views really incompatible? In this work we propose "straightening" as a simple and biologically plausible learning objective that achieves both goals at the same time. This objective is motivated by recent experimental findings that neural representations evolve over time in straighter temporal trajectories than their initial photoreceptor encoding, measured both at V1 population level (Henaff et al. 2021) and in terms of perceptual discriminability (Henaff et al. 2019). Such a neural representation automatically fulfills the first goal: for straight representations, prediction is simply a matter of linearly extrapolating past responses. Recent results show that robust training leads to straighter representations (Toosi & Issa 2023; Harrington et al. 2023). Here, we show that the converse also holds: straightening makes a discrimination model more immune to noise. In particular, we show that optimizing a network for straightening can give representations that not only achieve near state-of-the-art object recognition performance on several datasets, but also are more robust to two forms of degradation: 1) independent white noise injected into each channel throughout the network (mimicking neural noise); 2) adversarial perturbations (known to yield perceptual differences between artificial and biological systems). Our results suggest that straightness and robustness are two sides of the same coin. The dual successes of straightening as a learning mechanism provide an important step towards bridging the predictive and discriminative perspectives in vision.

3-171. Bipolar Disorder and the Dentate Gyrus: Effects of Lithium Therapy on Pattern Separation in silico

Selena Singh[1,2] SINGHS11@MCMASTER.CA
Suzanna Becker[1] BECKERS@MCMASTER.CA
Anouar Khayachi[3] ANOUAR.KHAYACHI@MCGILL.CA
Thomas Trappenberg[4] TT@CS.DAL.CA
Abraham Nunes[4] NUNES@DAL.CA

[1]McMaster University
[2]Psychology, Neuroscience and Behaviour
[3]McGill University
[4]Dalhousie University

Individuals with bipolar disorder (BD) experience cognitive and memory impairments that contribute to illness-related disability. Induced pluripotent stem cell (iPSC) derived hippocampal dentate granule cell-like neurons from individuals with BD are hyperexcitable and more spontaneously active relative to healthy control (HC) neurons. These abnormalities are normalized after the application of lithium in neurons derived from lithium responders (LR) only. How these abnormalities impact hippocampal microcircuit computation, contributing to BD-associated cognitive and memory impairments, is not understood. We aim to investigate the impacts of BD-associated abnormal dentate granule cell (GC) activity on the neural computation called pattern separation (PS) using a computational model of the hippocampal dentate gyrus (DG). PS involves mapping highly overlapping inputs onto less overlapping representations, which may aid the hippocampus with precise memory encoding.

We first fitted the parameters of biophysically realistic GC models using parameter optimization to electrophysiological data from iPSC GCs derived from LRs and non-responder (NR) BD patients and HCs. These models were incorporated into a DG network to assess performance on a spatiotemporal PS task. Relationships between GC hyperexcitability, lithium and spontaneous activity were analyzed using a linear mixed effects model.

By normalizing spontaneous activity levels, our results demonstrate that lithium improves PS performance in LR models only, while impairing PS in HC and NR models, paralleling clinical observations of cognitive functioning in BD patients after lithium exposure. Interestingly, although HC neurons show a slight reduction in excitability after lithium administration, lithium still impairs PS in the HC models, suggesting that reductions in cellular excitability do not necessarily equate to better PS performance in all cases, contradicting classical theories of PS. Therefore, excitation/inhibition balance in hippocampal microcircuits is critical for neural computations such as PS; this balance may be disrupted in individuals with BD, leading to disease-related cognitive and memory impairments.

3-172. Temporally-precise inference of neural dynamics from slow-sampling rate calcium imaging

Anjali Agarwal[1,2] ANJALI.AGARWAL2@EMORY.EDU
Feng Zhu[3] FENG.ZHU@EMORY.EDU
Yiyi Yu[4] YIYIY@UCSB.EDU
Harrison Grier[5] HARRISONGRIER@UCHICAGO.EDU
Asaph Zylbertal[6] A.ZYLBERTAL@UCL.AC.UK
Isaac Bianco[6] I.BIANCO@UCL.AC.UK
Spencer Smith[4] SLS@UCSB.EDU
Matthew Kaufman[7] MATTKAUFMAN@UCHICAGO.EDU
Chethan Pandarinath[1,8] CHETHAN.PANDARINATH@EMORY.EDU

[1]Emory University & Georgia Tech
[2]Wallace H. Coulter Department of Biomedical Engineering
[3]Emory University
[4]University of California Santa Barbara
[5]The University of Chicago
[6]University College London
[7]University of Chicago
[8]Biomedical Engineering

Two-photon calcium imaging allows monitoring the activity of vast populations of neurons in the brain, but a fundamental tradeoff exists between imaging more neurons and sampling at higher speed. A recent deep learning method, RADICaL, uses neural population dynamics and subframe timing to improve inference of higher frequencies at fast imaging rates (~30Hz), but relies on deconvolution. Deconvolution fails at slow imaging rates where sparse sampling misses peaks in the fluorescence. Here we present Free-RADICaL, a novel solution that infers neural population dynamics without deconvolution, enabling high temporal resolution inference at low imaging

rates. Using synthetic datasets, we demonstrate that Free-RADICaL accurately separates neural population dynamics from calcium indicator dynamics, with a novel regularization strategy: neuron-Coordinated Dropout. This enables accurate inference of high-frequency features (7-10Hz) from synthetic data with low imaging rates (2-4Hz). We next tested Free-RADICaL on fast 2p recordings from motor+somatosensory cortical areas in mice performing a water-grab task (31Hz, 400-500 neurons), which were previously used to evaluate RADICaL. Free-RADICaL achieved high-fidelity inference of neuronal firing rates and accurate decoding of hand kinematics, comparable to RADICaL, without deconvolution.We then applied Free-RADICaL to 2p data collected from mouse visual cortex using a Diesel2p mesoscope. Its dual scan engines allowed simultaneous imaging of two fields of view (FOVs) at independent imaging rates: 2150 neurons at 5Hz, and a 483-neuron subset at 95Hz. When applied to the 5Hz data, Free-RADICaL recovered features of neuronal responses originally missed in the 5Hz data but present in the 95Hz data, and inferred single-trial population trajectories with high temporal fidelity. Finally, we characterize Free-RADICaL's memory dependence and scaling using whole brain imaging data from larval zebrafish (5Hz, ~100k neurons). Currently, large GPUs (A100) support application to ~18k neurons simultaneously. Thus, Free-RADICaL enables greatly improved temporal resolution for calcium imaging of vast neural populations.

3-173. Functional connectomics reveals general wiring rule in mouse visual cortex

Zhuokun Ding[1] ZHUOKUND@BCM.EDU
Paul Fahey[1] PAUL.FAHEY@BCM.EDU
Stelios Papadopoulos[1] STELIOS.PAPADOPOULOS@BCM.EDU
Eric Y. Wang[1] ERICWANG919@GMAIL.COM
Brendan Celii[1] BRENDAN.CELII@BCM.EDU
Christos Papadopoulos[1] CHRISTOS.PAPADOPOULOS@BCM.EDU
Alexander Kunin[2] ALEXKUNIN@CREIGHTON.EDU
Andersen Chang[1] ANDERSEN.CHANG@BCM.EDU
Jiakun Fu[1] JIAKUN.FU@BCM.EDU
Zhiwei Ding[1] ZHIWEID@BCM.EDU
Saumil Patel[1] SPATEL@BCM.EDU
Kayla Ponder[1] KAYLA.JONESPONDER@BCM.EDU
MICrONS Consortium[3] INFO@MICRONS-EXPLORER.ORG
Emmanouil Froudarakis[4] FROUMAN@IMBB.FORTH.GR
Fabian Sinz[5] SINZ@CS.UNI-GOETTINGEN.DE
H. Sebastian Seung[6] SSEUNG@PRINCETON.EDU
Forrest Collman[7] FORRESTC@ALLENINSTITUTE.ORG
Nuno Macarico da Costa[7] NUNOD@ALLENINSTITUTE.ORG
R. Clay Reid[7] CLAYR@ALLENINSTITUTE.ORG
Edgar Walker[8] EYWALKER@UW.EDU
Xaq Pitkow[9] XAQ@CMU.EDU
Jacob Reimer[1] REIMER@BCM.EDU
Andreas S. Tolias[1] ASTOLIAS@BCM.EDU

[1]Baylor College of Medicine
[2]Creighton University
[3]MICrONS Consortium
[4]Foundation for Research and Technology Hellas
[5]University of Gottingen
[6]Princeton University
[7]Allen Institute for Brain Science
[8]University of Washington, Seattle
[9]Carnegie Mellon University

To understand how the brain computes, it is important to unravel the relationship between circuit connectivity and function. Previous research has shown that excitatory neurons in layer 2/3 of the primary visual cortex of mice with similar response properties are more likely to form connections. However, technical challenges of combining synaptic connectivity and functional measurements have limited these studies to few, highly local connections. Using the millimeter scale and nanometer resolution of the MICrONS dataset, we studied the connectivity-function relationship in excitatory neurons of the mouse visual cortex. Across interlaminar and interarea projections, we assessed connection selectivity at the coarse axon trajectory and fine synaptic formation levels. We found that neurons with highly correlated responses to natural videos tended to be more connected with each other, not only within the same cortical area but also across multiple layers and visual areas, including feedforward and feedback connections. Next, we comprehensively characterized the neuronal function using a digital twin model of this mouse that accurately predicted responses to arbitrary video stimuli. The digital twin enabled us to factorize each neuron's tuning into a feature component (what the neuron responds to) and a spatial component (where

the neuron's receptive field is located). We found that, across multiple projection types, the spatial components predicted connectivity at the coarse axonal scale, while the feature components were indicative of connectivity at the fine synaptic scale. This discovery highlights a multi-scale wiring rule of the neocortex, where feature and spatial tuning play crucial roles at their respective spatial scales. Such insights fundamentally reshape our understanding of the intricate relationship between neural network organization and function.

3-174. Acute and chronic isolation promote diverse behaviors and modifies mPFC responses to social contact

Christopher Lee[1] CRLEE@UCSD.EDU
Gates Schneider[1] GATESSCH@UNC.EDU
Felix Taschbach[2] FTASCHBA@UCSD.EDU
Tristan Tuazon[1] TTUAZON@SALK.EDU
Alexandra Garcia[1] ALG013@UCSD.EDU
Dexter Tsin[1] TTSIN@UCSD.EDU
May Chan[1] MCHAN@SALK.EDU
Kanha Batra[3] KBATRA@ENG.UCSD.EDU
Anousheh Bakhti-Suroosh[1] ABAKHTIS@HEALTH.UCSD.EDU
Romy Wichmann[1] RWICHMANN@SALK.EDU
Talmo Pereira[1] TALMO@SALK.EDU
Marcus Benna[4,5] MBENNA@UCSD.EDU
Kay Tye[1] TYE@SALK.EDU

[1]Salk Institute for Biological Studies
[2]University of California, San Diego
[3]The Salk Institute for Biological Studies
[4]UC San Diego
[5]Neurobiology

Divergent social behaviors emerge from different durations of social isolation (Lee et al., 2021). However, the exact time course of isolation impacting social behavior and the neural systems and circuits maintaining social homeostasis remain elusive. Previously, Matthews et al. found that dorsal raphe nucleus dopamine (DRNDA) neurons mediate a loneliness-like state and innervate the medial prefrontal cortex (mPFC). To explore how the mPFC encodes social information and undergoes a state change following isolation, we used calcium imaging, neuropeptide sensors, computer vision, and machine learning tools. To first determine how different durations (2hr, 6hr, 24hr, 7d, 14d, and 28d) of isolation impact social behavior, we performed a juvenile intruder task after isolating adult male mice. We performed pose estimation using SLEAP (Pereira et al., 2022) and developed a custom pipeline for behavioral feature extraction and unsupervised clustering for behavioral motif discovery. We found an inverse correlation between isolation duration and interaction time and that 2 and 6 hours of isolation promotes social interaction with juvenile mice. Despite not detecting differences in interaction time in chronically isolated mice, unsupervised clustering of behavior features revealed changes in social behavior repertoire compared to group-housed mice, promoting face sniffing and reducing chasing. Next, we performed calcium imaging using miniature endoscopes in the mPFC of mice engaged in social behavior after group-housing and isolation. We found that isolation (24hr and 7d) increased the responsiveness of mPFC neurons to social contact by promoting excitatory responses. Additionally, we found an increase in mPFC population trajectory length following isolation compared to a group housed session prior, an effect reversible followed by re-housing. Finally, we found that dopamine release in the mPFC increases following social isolation during social behavior. Overall, our findings support a role for mPFC in promoting features of the response to novel social stimuli following social isolation.

3-175. Underlying tristability explains sharp waves with two pyramidal components

Stefano Masserini[1,2] STEFANOMASSE@GMAIL.COM
Richard Kempter[3,4] R.KEMPTER@BIOLOGIE.HU-BERLIN.DE

[1]Humbold Universitat zu Berlin
[2]Faculty of Life Sciences
[3]Humbold-Universitat zu Berlin
[4]Department of Biology, Institute for Theoretical Biology

Hippocampal sharp waves (SPWs) are massive firing barrages co-occurring with ripple oscillations and involved in memory consolidation. It was recently shown that in area CA3 these waves have two distinct components:

so-called athorny and thorny pyramidal neurons fire at different times, with a considerable delay. We investigated this phenomenon in a network model comprising, in addition to these pyramidal populations, two interneuronal populations with opposite pro-SPW and anti-SPW roles. While competition between the two kinds of interneurons regulates the switches between SPW and non-SPW states, the sequential activation of the two pyramidal types defines two distinct phases of the SPW. We found that a long delay between the peaks can be found when the first population contributes to the second one more inhibition - mediated by the interneurons - than excitation. We confirmed these observations in a simplified Wilson-Cowan rate model directly derived from the spiking network. The rate model displays nested oscillations, in which the two pyramidal components clearly activate in anti-phase, within the broader alternation between SPW and non-SPW states. Removal of adaptation, the driving force of all these oscillation, reveals an underlying tristability. Our findings call for further experimental investigations on the connectivity among the two pyramidal populations and to the pro-SPW interneurons.

3-176. Learning a predictive representation of control in meta-RL

Kai Sandbrink[1,2] KAI.SANDBRINK@LMH.OX.AC.UK
Laurence Hunt[1] LAURENCE.HUNT@PSY.OX.AC.UK
Christopher Summerfield[1] CHRISTOPHER.SUMMERFIELD@PSY.OX.AC.UK

[1]University of Oxford
[2]Department of Experimental Psychology

Recent work has increasingly suggested roles for dopamine in reinforcement learning that go beyond signaling reward to signaling general sensory prediction error or driving habit-driven learning. However, it remains unclear how slow reinforcement learning could enable organisms to flexibly adapt their learning between behaviors in response to environmental changes. In this study, we propose an additional role for action-prediction errors to allow credit assignment and rapid adaptation across different levels of control by learning a predictive representation of an agent's efficacy using a novel temporal-difference learning algorithm. We formulate two observe-or-bet tasks that require trading off purely exploratory actions that only yield information with purely exploitative actions that only yield reward but no feedback across different efficacy levels. We show that meta-reinforcement learning networks from whose hidden layers we train a predictive efficacy readout modulate their behavior flexibly, whereas those trained purely using reward adapt static policies. Investigating the emergent representations in the network, we find that only the prediction-trained meta-reinforcement learning agents learn task representations that are differentiated across levels of control, suggesting interference in credit assignment in the regular algorithms hinders learning. We test human performance on the task and find that people also dynamically adapt their behavior. When we perturb networks' sense of efficacy, we generate behavioral traces of neural networks that correspond to individual variation in human behavior across different efficacy levels. By suggesting a role for dopaminergic error-driven signals in driving the interpretation of environmental feedback, this work introduces a framework for understanding behavioral adaptation and learning across levels of control.

3-177. Task-relevant information is enriched in mouse PPC but not selectively propagated to M1

Poojya Ravishankar[1,2] POOJYA@UCHICAGO.EDU
Harrison Grier[3] HARRISONGRIER@UCHICAGO.EDU
David Sabatini[1,4] DAVIDASABATINI@GMAIL.COM
Matthew Kaufman[1] MATTKAUFMAN@UCHICAGO.EDU

[1]University of Chicago
[2]Computational Neuroscience
[3]The University of Chicago
[4]Organismal Biology and Anatomy

To understand how brain areas work together to transform sensation into action, we must determine how relevant information propagates forward while irrelevant information is discarded. Some cortical projections are known to preferentially carry general classes of information that are more relevant to the downstream area: e.g., projection neurons from V1 carry higher spatial frequencies to object recognition areas than to motion-sensitive areas, and barrel cortex routes whisking information to S2 but not M1. However, it is unclear whether selective routing can develop when information relevance is particular to a learned task. Here, we ask whether mouse Posterior Parietal Cortex can selectively route task-relevant information to M1. PPC preferentially engages in vision-to-movement tasks and exhibits complex task-feature selectivity. To ask whether PPC can learn to selectively route task-relevant information, we developed a novel, closed-loop, visuomotor 2D-joystick task that included both relevant and irrelevant visual information in the stimulus. We performed two-photon calcium imaging of pyramidal cells in

contralateral PPC of expert mice (>14,000 neurons, 5 mice, 2-3 months experience), with PPC-M1 projection neurons labeled via retrograde tracing (>500 neurons). Linear encoding models identified what task features modulated each neuron's activity. PPC neurons were, on average, most strongly modulated by joystick movement; a slight majority was more modulated by relevant vs. irrelevant movement. Most visually responsive neurons were more strongly modulated by task-relevant than task-irrelevant visual drift (57%). These results argue that PPC activity more strongly encodes task-relevant than task-irrelevant information. Surprisingly, however, encoding in labeled PPC-M1 neurons was virtually indistinguishable from unlabeled neurons; they were not more biased toward task-relevant information than other pyramidal cells were. This argues that the full representation present in PPC is sent downstream to M1, and that, at least for this pathway, half a lifetime of training is insufficient to induce selective information routing.

3-178. Identifying Distinct Neural Dynamics using Switching Recurrent Neural Networks

Yongxu Zhang[1,2] YONGXU.ZHANG@YALE.EDU
Shreya Saxena[1,3] SHREYA.SAXENA@YALE.EDU

[1]Yale University
[2]Department of Biomedical Engineering
[3]Wu Tsai Institute

Neural population activity often exhibits distinct dynamical features across time, which may correspond to distinct internal processes or behavior. Linear methods and variations thereof, such as Hidden Markov Model (HMM) and Switching Linear Dynamical System (SLDS), are often employed to identify discrete states with evolving neural dynamics. However, these techniques may not be able to capture the underlying nonlinear dynamics associated with neural propagation. Recurrent Neural Networks (RNNs) are commonly used to model neural dynamics thanks to their nonlinear characteristics. In our work, we propose Switching Recurrent Neural Networks (SRNN), combining HMMs and RNNs with time varying weights to reconstruct the switching dynamics of neural time-series data. We apply these models to simulated data as well as cortical neural activity across mice and monkeys, which allows us to automatically detect discrete states that lead to the identification of varying neural dynamics. We show that SRNNs are able to capture the switches between different states of the simulated data, and reconstruct the property of the true dynamic systems. In a monkey reaching dataset with electrophysiology recordings, and a mouse self-initiated lever pull dataset with widefield calcium recordings, SRNNs are able to automatically identify discrete states with distinct nonlinear neural dynamics. The inferred switches are aligned with the behavior, and the reconstructions show that the recovered neural dynamics are distinct across different stages of the behavior.

Author Index