Recurrent neural networks with diverse intrinsic timescales

Aidan Higgs[1*], Raiyyan Siddiqui[2*], Angela Langdon[1]
[1]Unit on the Neural Computations in Learning, [2]Section on Critical Brain Dynamics,
National Institute of Mental Health, NIH, Bethesda, MD, USA
*Authors contributed equally, names alphabetized

National Institute of Mental Health

How does the distribution of **intrinsic time constants** in a **recurrent neural network** influence training and performance of tasks that require **integration and maintenance of information over delays**?

Background

Recurrent Neural Networks (RNNs) are computational models in which artificial neurons are connected in loops and are used to simulate brain computations that require integration and maintenance of information across time.

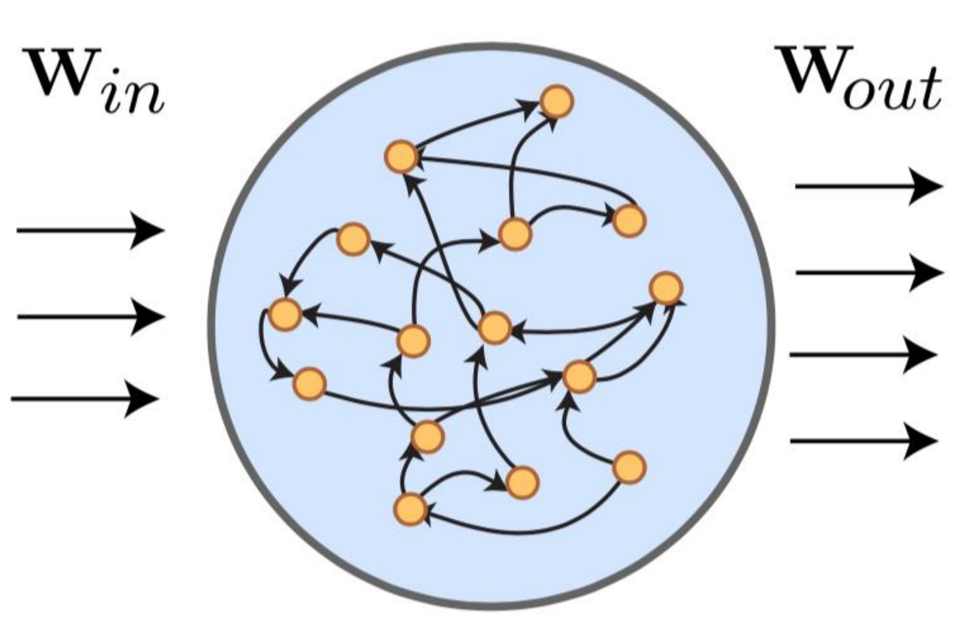
Most RNNs assume all neurons share a fixed intrinsic time constant, unlike the brain, where neurons express diverse dynamics[1].

Prior work[2,3] has shown that time constant heterogeneity can improve training performance of some RNN models.

Methods and model

- Train rate-based RNNs with heterogeneous intrinsic time constants ("multi-tau" RNNs) on tasks requiring integration across multiple timescales.
- Compare training and performance across:
  - Single-tau ("vanilla") RNNs
  - Multi-tau RNNs with different distributions of time constants:
    uniform, normal, bimodal

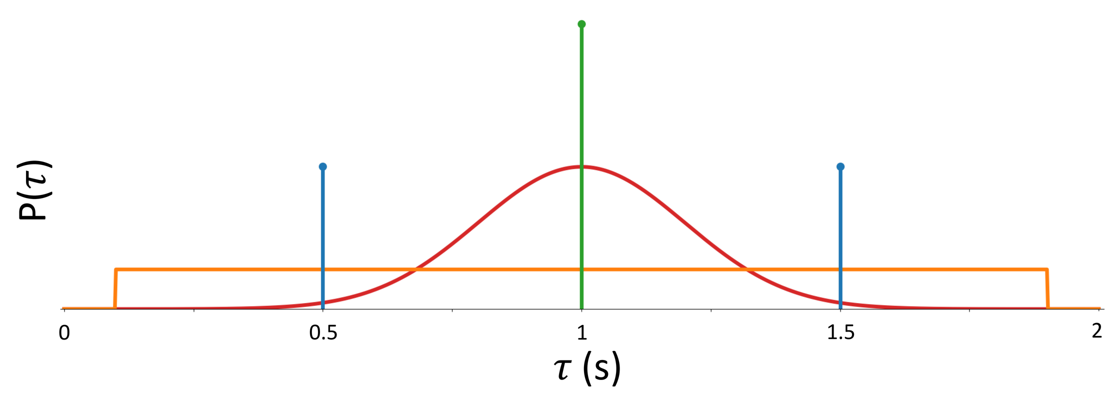
$$\tau_i \frac{dx_i}{dt} = -x_i + \sum_j W_{ij}\,\phi(x_j) + \sum_k W^{\text{in}}_{ik} u_k$$



RNNs with distributed intrinsic time constants τ

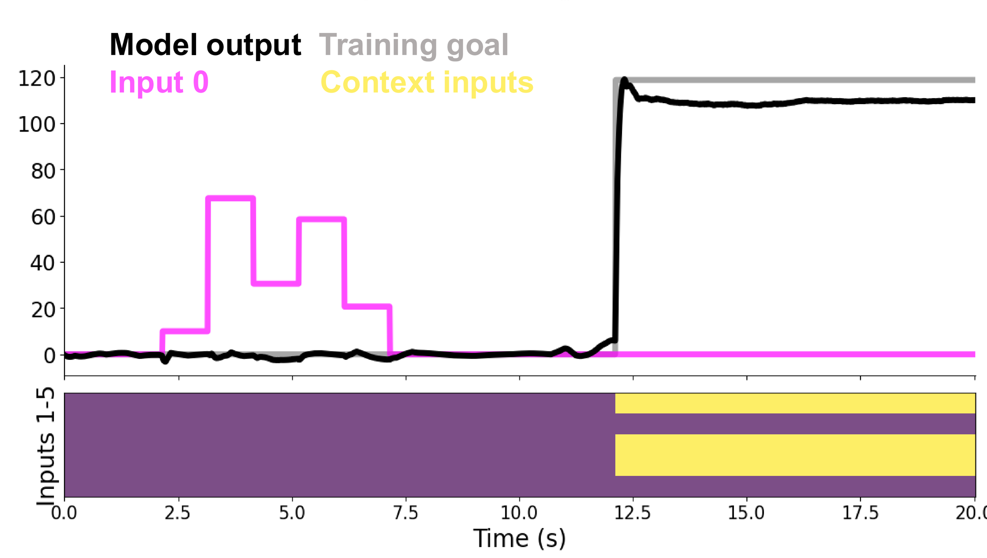
For each RNN, draw single unit τ from:

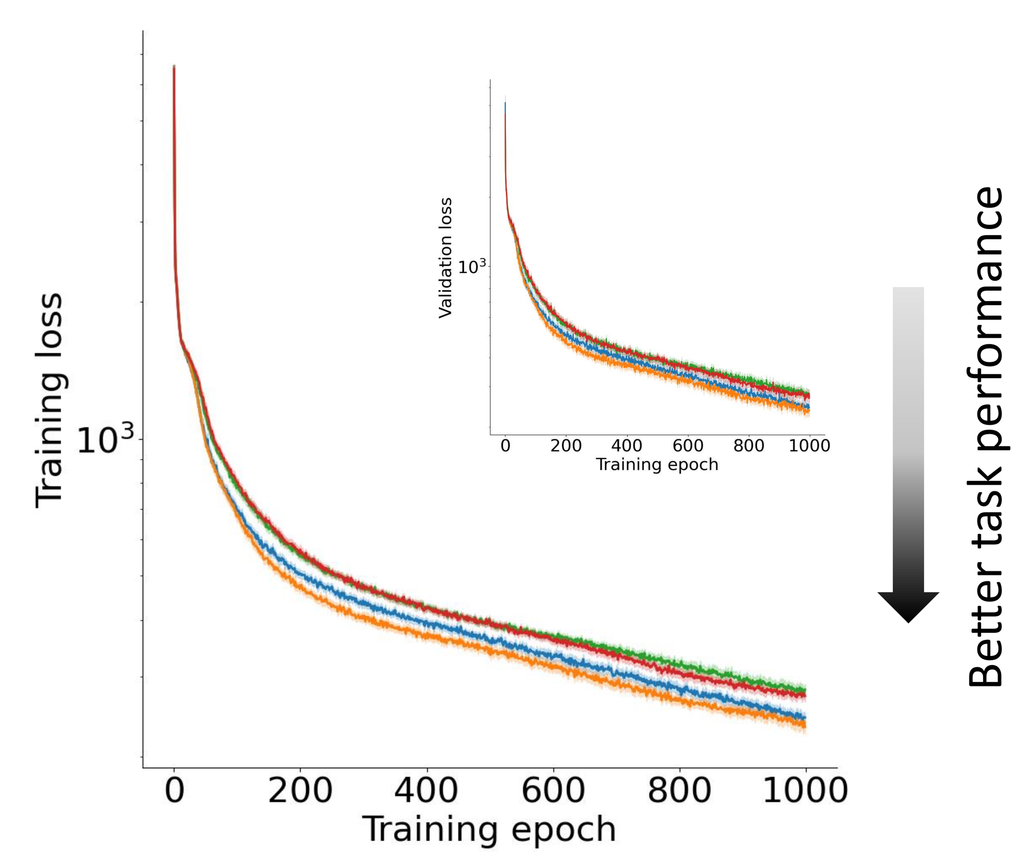
Vanilla: τ = 1s
Normal: $\tau \sim \mathcal{N}(1, 0.2)$
Uniform: $\tau \in [0.1,1.9]$
Bimodal: τ = 0.5s or 1.5s

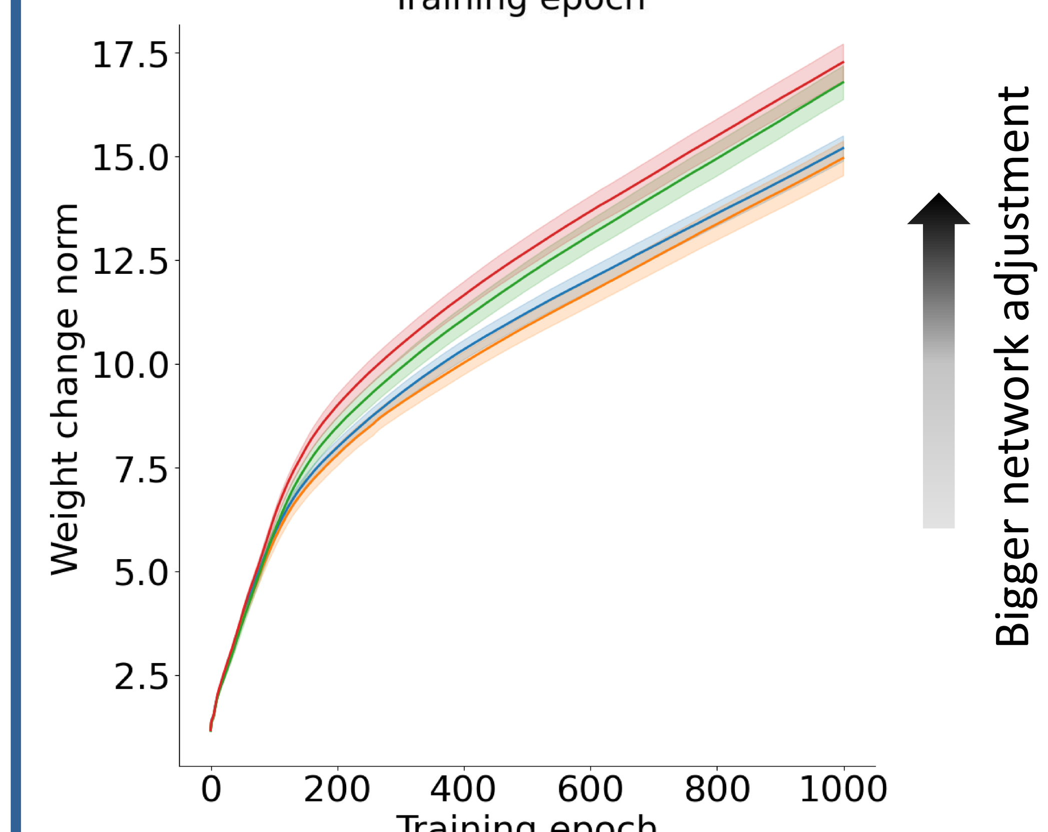


Multi-timescale working memory task

Integrate a pattern of inputs, maintain the result then output the signaled decision:

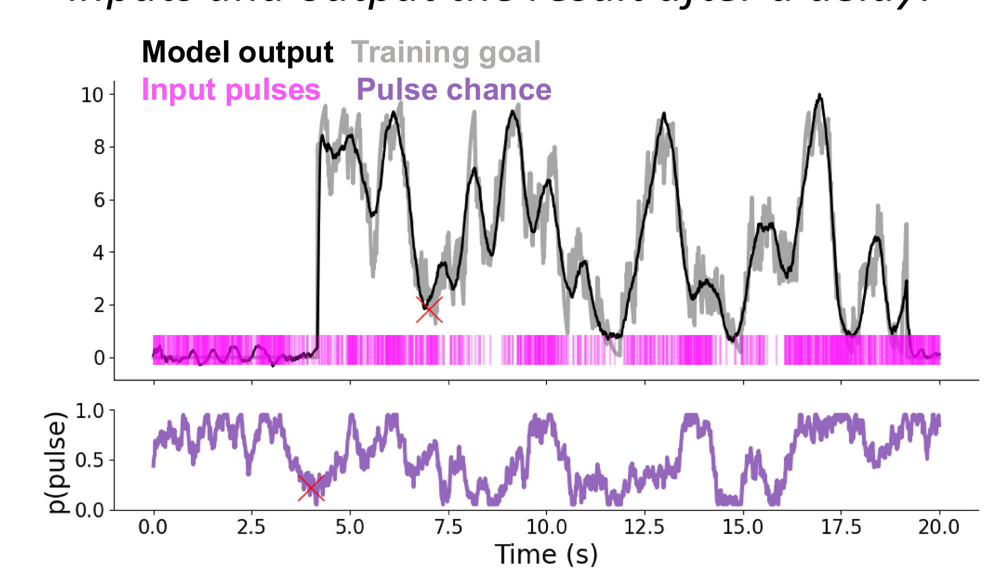


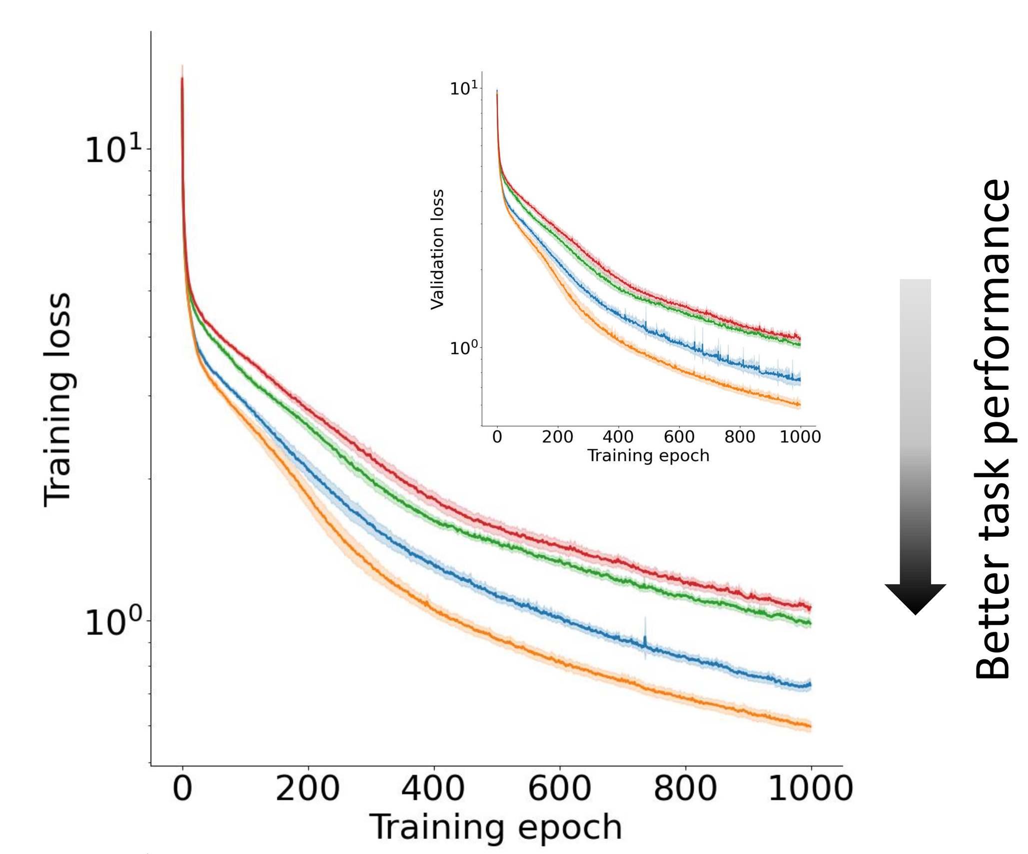


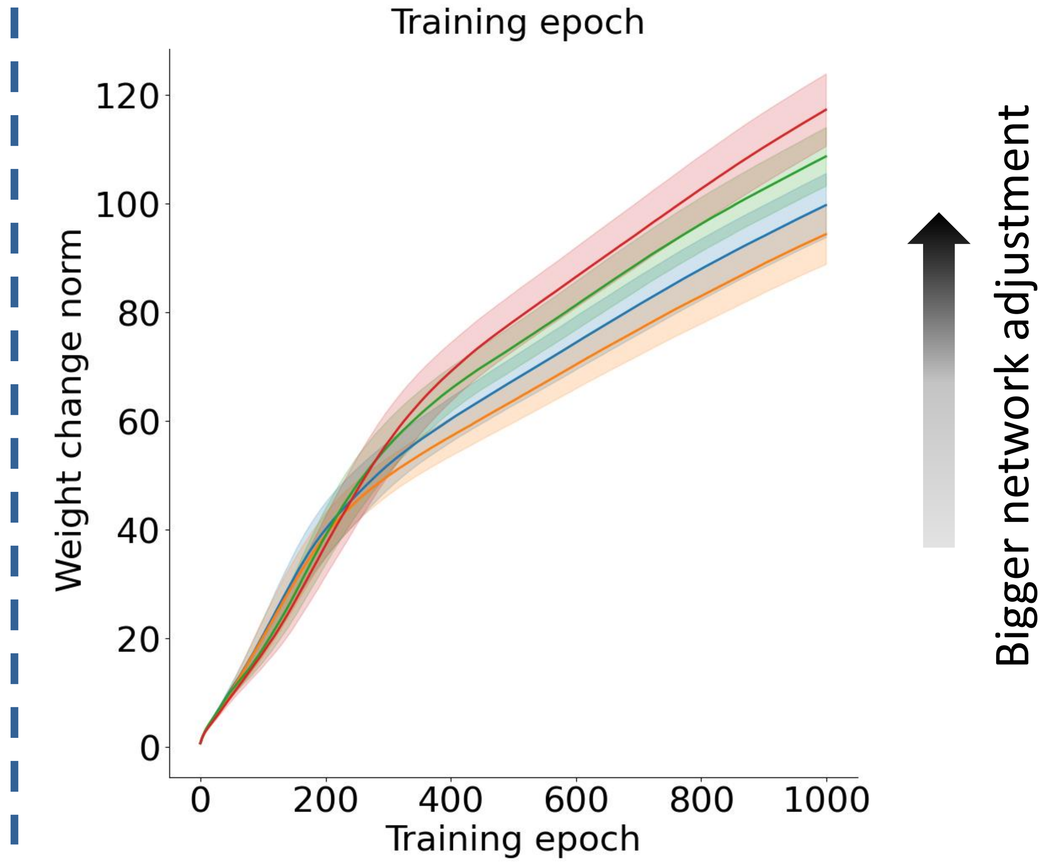


Simultaneous integration and working memory task

Simultaneously integrate and remember inputs and output the result after a delay:



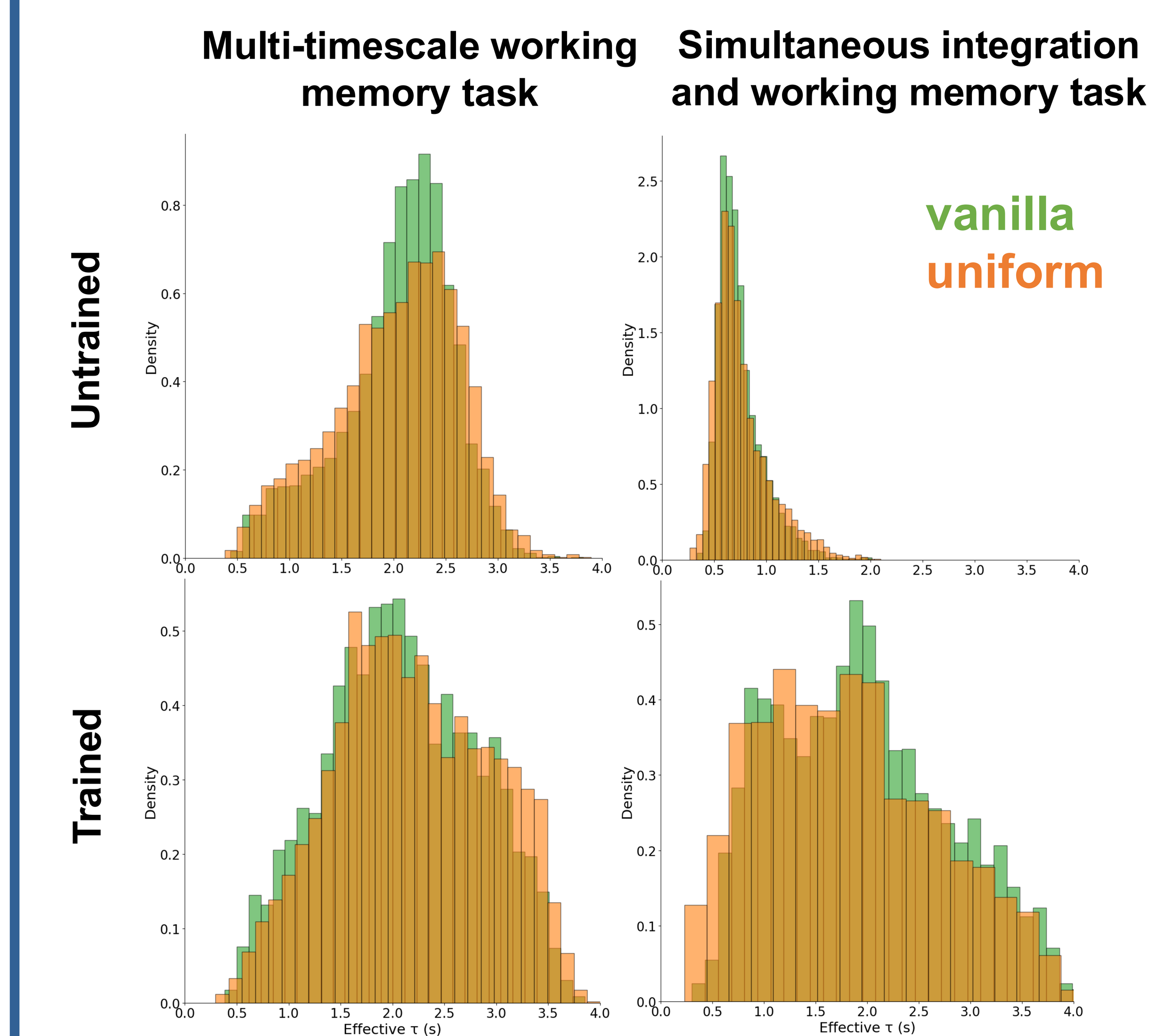




Greater heterogeneity of intrinsic timescales in an RNN led to **greater training efficiency** and **more accurate task performance**.

Smaller changes to recurrent weights suggest the **uniform** and **bimodal** multi-tau RNNs are intrinsically "closer" to the dynamic regime of sustained activity needed for these integration and memory tasks.

Trained RNNs have distributed effective timescales



Decay of autocorrelation in firing rate across time lags shows the effective dynamic timescale of a unit.

RNN training **distributes the effective timescales in the network**, even for homogenous intrinsic-timescale RNNs.

Acknowledgements

This work utilized the computational resources of the NIH HPC Biowulf cluster (https://hpc.nih.gov).

References

1. Murray, J., Bernacchia, A., Freedman, D. et al. A hierarchy of intrinsic timescales across primate cortex. Nat Neurosci 17, 1661–1663 (2014). https://doi.org/10.1038/nn.3862
2. Perez-Nieves, N., Leung, V.C.H., Dragotti, P.L. et al. Neural heterogeneity promotes robust learning. Nat Commun 12, 5791 (2021). https://doi.org/10.1038/s41467-021-26022-3
3. Kurikawa, T. (2025). Slow and fast neurons cooperate in contextual working memory through timescale diversity (arXiv preprint). https://doi.org/10.48550/arXiv.2506.07341
4. Mastrogiuseppe, F., & Ostojic, S. (2018). Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. Neuron, 99(3), 609-623.e29. https://doi.org/10.1016/j.neuron.2018.07.003
5. Liu, Y. H., Baratin, A., Cornford, J., Mihalas, S., Shea-Brown, E., & Lajoie, G. (2024). How connectivity structure shapes rich and lazy learning in neural circuits. In Proceedings of the International Conference on Learning Representations (ICLR 2024).