Robust Working Memory Computation and Entropy Production in Random Neuronal Networks

Thiparat Chotibut

Chula Intelligent and Complex Systems Center of Excellence, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

Training recurrent neuronal networks consisting of excitatory (E) and inhibitory (I) units with additive noise for working memory computation slows and diversifies inhibitory timescales, leading to improved task performance that is attributed to emergent marginally stable equilibria [1]. Yet the link between trained network characteristics and their roles in shaping desirable dynamical landscapes remains unexplored. We first discuss the phenomenology from [1], where trained cortical networks with heterogeneous inhibitory dynamics are more robust to noise in working memory models that compute using discrete attractors (fixed-point equilibria) hopping mechanism. Then, we study the Jacobian matrices describing the dynamics near these equilibria and show that they are sparse, *non-Hermitian rectangular-block* matrices modified by *heterogeneous synaptic decay timescales* and activation-function gains. We specify a random matrix ensemble that faithfully captures the spectra of trained Jacobian matrices, arising from the *inhibitory core-excitatory periphery network motif* (pruned E weights, broadly distributed I weights) observed post-training. An analytic theory of this ensemble is developed using statistical field theory methods: a Hermitized resolvent representation of the spectral density followed by a supersymmetry-based treatment in the style of Fyodorov and Mirlin. In this manner, an analytic description of the spectral edge is obtained, relating statistical parameters of the Jacobians (sparsity, weight variances, E/I ratio, and the distributions of timescales and gains) to near-critical features of the equilibria [2]. These results reveal that the slow, inhibitory core naturally pushes the equilibria of these cortical networks toward marginal stability without requiring parameter fine-tuning. Furthermore, we will show that this stable computation arises when the network dynamics is sustained in a non-equilibrium steady state (NESS). In this context, we will also discuss a newly discovered, *universal entropy production rate* formula for stochastic dynamics in an NESS driven by a variety of directed networks [3].

[1] N. Rungratsameetaweemana, R. Kim, T. Chotibut, and T. J. Sejnowski, Proceedings of the National Academy of Sciences, 122(3) (2025) e2316745122

[2] T. Chotibut, O. Evnin, and W. Horinouchi, arXiv: 2512.12767 (2025) (under review in SciPost Physics)

[3] T. Chotibut, E. Gudowska-Nowak, and M. A. Nowak, (to appear)