Input-dependent Directionality of Interactions Between Cortical Areas

Joana Carmona 1*, Francesca Mastrogiuseppe 1*, Byron M. Yu 2, Adam Kohn 3, and Christian K. Machens 1 (*equal contribution)
1. Champalimaud Foundation, Lisbon, PT
2. Carnegie Mellon University, Pittsburgh, US
3. Albert Einstein College of Medicine, New York, US

A long-standing question in systems neuroscience is how different brain areas communicate. One approach to studying information flow is to analyze the covariation of activity across areas, with the temporal structure of this covariation offering clues about the directionality of signaling [1 - 3]. Recent studies applying this method to neural populations across areas have revealed that the directionality of interactions can shift rapidly, depending on stimuli and task demands [4, 5]. These findings suggest that experimentally measured covariation metrics reflect not only the underlying synaptic connectivity but also dynamic influences such as inputs from other regions. Understanding how these metrics relate to circuit-level mechanisms remains a challenging task that warrants theoretical approaches [6, 7].

Here, we develop a theoretical framework for the emergence of directional interactions in recurrent circuits, based on recurrent neural network models driven by stochastic inputs [8]. Our theory leverages the analysis of network activity along particular directions, given by the eigenvectors of the synaptic connectivity matrix. We apply our framework to study how external inputs can flexibly shape the directionality of inter-area communication in mesoscopic cortical circuits. Our analysis reveals that inputs targeting excitatory (E) and inhibitory (I) populations play different roles in inter-areal interactions. Specifically, inputs to E significantly influence directionality by inducing a systematic bias from the area receiving the strongest input, while inputs to I regulate the amplitude and timescale of activity with minimal impact on directionality. In circuits with feature-specific connectivity and inputs, the effect of inputs to E on directionality is most apparent at the level of latent variables reflecting co-fluctuations between stimulus-selective and unselective units. Overall, our work provides a theoretical foundation for the interpretation of experimentally measured covariation metrics, and advances our understanding on how to link the functional and anatomical substrates of neural interactions.

cross-covariances, directionality, inter-areal interactions, recurrent neural networks