

Title: Stochastic neural dynamics during short term memory maintenance

Abstract

Short-term memory (STM) errors can arise from different circuit mechanisms. Recent work in mice performing a two-alternative delayed-choice (2-ADC) task shows that behaviour alternates between an engaged state, where forgetting errors arise from stochastic transitions between memory attractors in the anterolateral motor cortex (ALM), and a disengaged state, where behaviour decouples from the task and errors reflect choice-history biases rather than memory maintenance failures.

Recurrent neural network (RNN) models have successfully captured key features of STM maintenance, such as the emergence of persistence attractor dynamics, but none have addressed the episodic alternation between engaged and history-driven behavioural states, nor the error structure that emerges from each regime.

Here, we present an approach to address this gap using a low-rank RNN to be trained on behavioural data from the same 2-ADC task. The proposed model will incorporate variable delay durations, a cognitive control input to switch the network between engaged and disengaged states, and a choice history input that captures the influence of previous responses and biases choices during disengaged epochs. The low-rank connectivity constraint confines the network to a low-dimensional latent space, allowing interpretability of the resulting dynamics and comparison with the population activity recorded from the anterolateral motor cortex (ALM) during the task. By analysing the resulting dynamical system, we aim to determine whether a unified low-rank circuit can account for both error classes and generate testable predictions about the mechanisms underlying STM.