**Fixed point analysis**

We analyzed the model’s neural dynamics during the successful completion of a task. Following Driscoll et al., 2024, we used the Fixed Point Finder package [https://github.com/mattgolub/fixed-point-finder] to search for stationary points corresponding to correct model output. We stuck to the default arguments for Fixed Point Finder, following the example of the 3 bit flip-flop, including 5,000 maximum iterations and 1,000 initial states. We hypothesized that an arrangement of these features in state space would characterize the trained models and that this characterization would be invariant to the specific solutions (weight matrices) found during training. However, we did not find evidence of fixed points in our analysis. In five models initialized with different random seeds, we observed that the minimum q values were not sufficiently small (<<1/T^2 where T is the timescale of our task) to be considered fixed points. To confirm that the state vector had velocity while the model output remained sufficiently constant to meet our performance criteria, we calculated the magnitude of the velocity vector over time from our hidden states. One can see an appreciable velocity on the scale of sqrt(q). Cosine similarity between successive velocity vectors revealed ballistic motion during the beginning of the task followed by diffusive motion dominated by the intrinsic noise we include in our equations of motion. While the hidden state vector does not, strictly speaking, settle into a stable fixed point as we hypothesized, it does fluctuate around a constant position. Given that the hidden state dimension is much greater than the output dimension, the output matrix is under-determined and multiple hidden states map to similar enough output states to meet our performance criteria.