Markov Transitions between Attractor States in a Recurrent Neural Network

Jeremy Bernstein*[1], Ishita Dasgupta*[2,4], David Rolnick*[3], Haim Sompolinsky[4,5]

[1] Computation and Neural Systems, California Institute of Technology, USA, [2] Department of Physics, Harvard University, Cambridge, MA, USA
[3] Department of Mathematics, Massachusetts Institute of Technology, USA, [4] Center for Brain Science, Harvard University, USA,
[5] The Edmond and Lily Safra Center for Brain Sciences, Hebrew University of Jerusalem, Israel

*These authors contributed equally to the work

Introduction

Why stochastic transitions?

Probabilistic models of cognition (see e.g. [1])

- Enable stochastic **simulations**, such as tracking noisy movements of objects [2] or reasoning with uncertainty about sequences of events [3].
- Allow for **statistical inference** using MCMC methods.
- Model observed processes in birdsong etc. [4]

In a **Markov chain,** the future depends only on the present; this captures many natural applications.

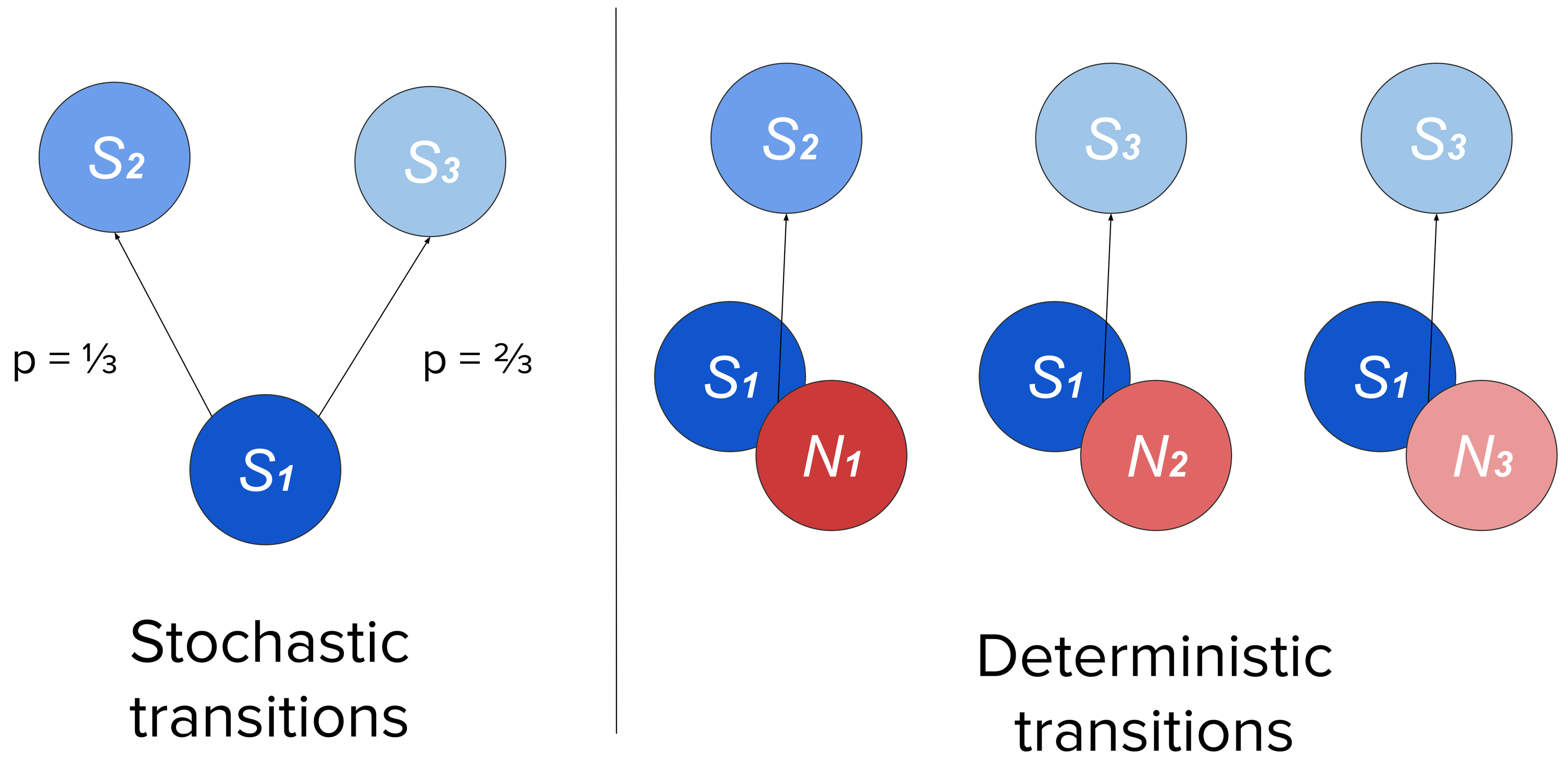
Why attractor networks?

- A **Hopfield network** [5] is a fully connected network of N neurons with update rule:

$$x_i \leftarrow \text{sign}\left(\sum_{j=1}^{n} J_{ij} x_j\right)$$

- For appropriate J_{ij}, the network state will fall into one of several chosen **attractor states, robust to noise**.
- Attractor states are a common model for memories.

How to go from deterministic to stochastic?

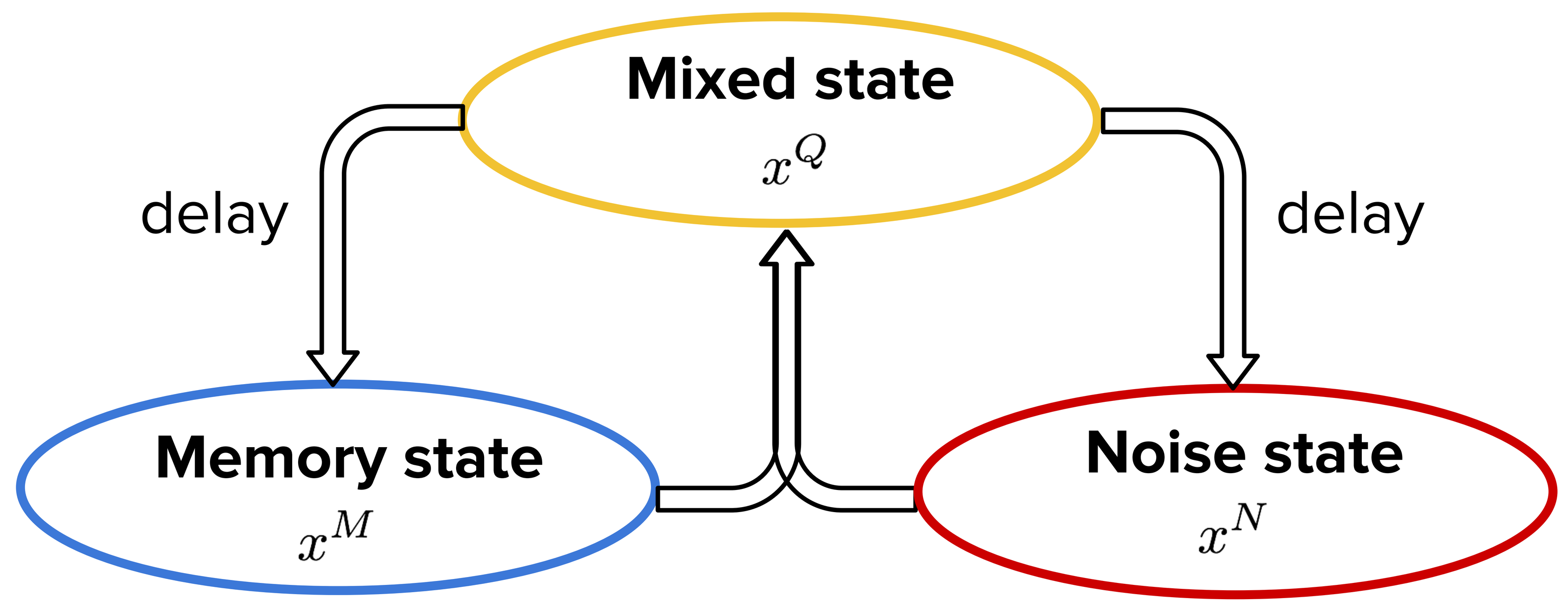


Acknowledgments

We would like to thank the Center for Brains, Minds, and Machines and the Marine Biological Laboratory at Woods Hole for their support during this research. D.R. is additionally supported by NSF grant no. 1122374.

Network design

Structure of the network



$$x_i^M \leftarrow \text{sign}\left(\sum_{\ell=1}^{n_M} J_{i\ell}^M x_\ell^M + \delta_\tau^{\text{mod}}(t)\sum_{k=1}^{n_Q} J_{ik}^{MQ} x_k^Q(t-\tau)\right)$$

$$x_j^N \leftarrow \text{sign}\left(\sum_{\ell=1}^{n_N} J_{j\ell}^N x_\ell^N + \nu(t,\tau)\right)$$

$$x_k^Q \leftarrow \text{sign}\left(\sum_{\ell=1}^{n_Q} J_{k\ell}^Q x_\ell^Q + \sum_{i=1}^{n_M} J_{ki}^{QM} x_i^M + \sum_{j=1}^{n_N} J_{kj}^{QN} x_j^N\right)$$

- The function $\nu(t,\tau)$ denotes a noise function that is resampled uniformly at random at intervals of τ.
- The function $\delta_\tau^{\text{mod}}(t)$ is 1 (else 0) if $t \equiv 0 \pmod{\tau}$.

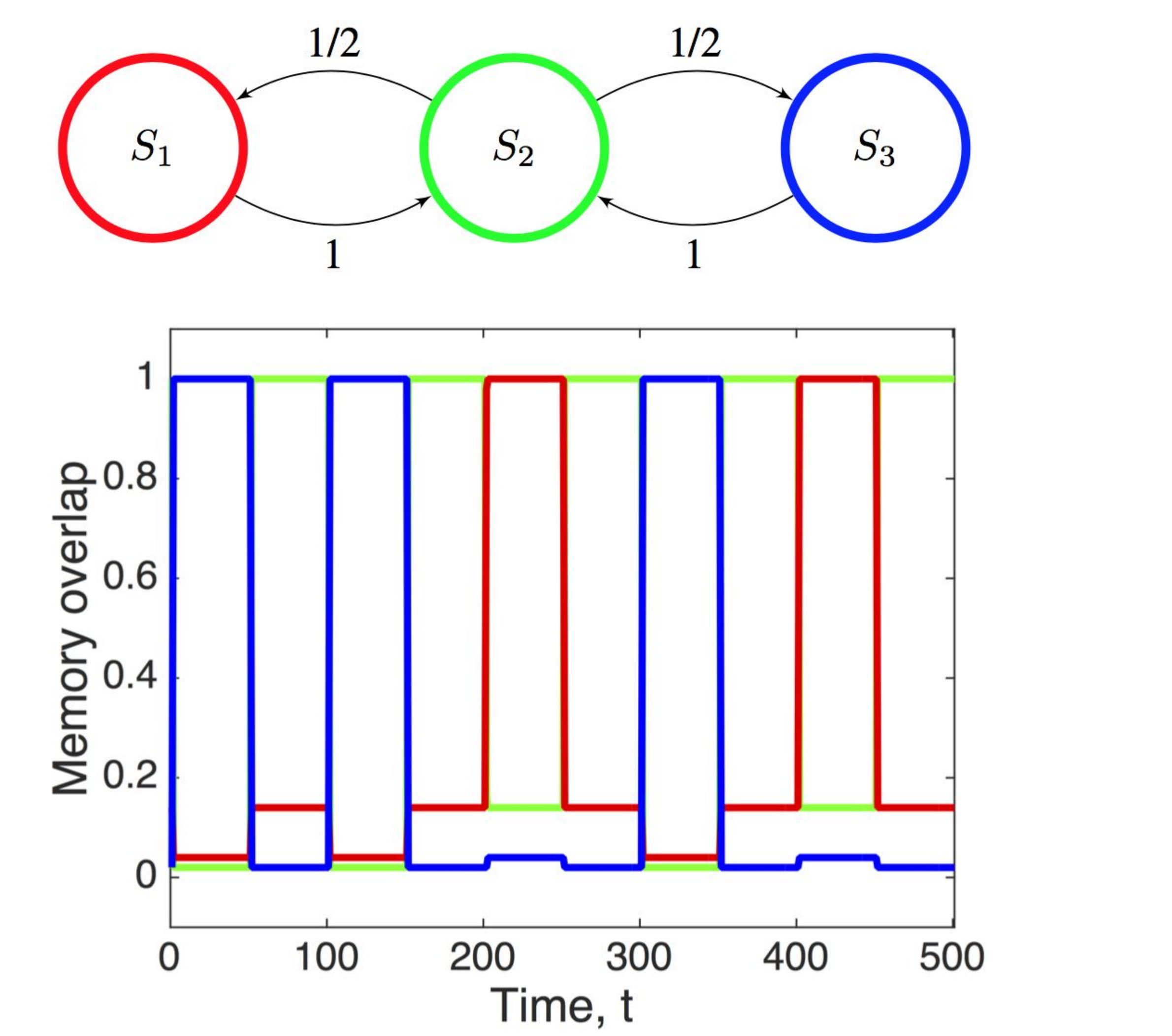
Why the mixed state?

- The next memory state is determined by the **concatenation** of the current state and a noise state. (memory(t), noise(t)) → memory(t+1)
- There is **strong linear dependence** between such concatenations, so Cover's theorem breaks down [6].
- We use a **random matrix** to project (memory, noise) pairs to a higher-dimensional space, where the corresponding mixed states are linearly separable.

Structure of noise states

- The noise state network is a **ring attractor**, with symmetric attractor states. When resampled, it falls into an attractor state **uniformly at random**.
- Attractor states ensure the network is robust to noise, compared with an ensemble of random neurons.

Example output



Learning rules

- The weight matrix J^M is learned using **Hebb's rule**, J^Q using the **perceptron rule**, and J^{MQ} as in [7].
- The weight matrices J^{QM} and J^{QN} are random, while J^N is chosen to instantiate a ring attractor network.

Future directions

- Determine network **capacity** and **noise robustness**.
- Construct statistical inference model for attractor networks using **Markov Chain Monte Carlo** methods.
- Improve **clock gating mechanism** for transitions.
- Develop biologically plausible **on-line learning rule**.

References

[1] Griffiths, T. L.; Chater, N.; Kemp, C.; Perfors, A.; Tenenbaum, J. B. 2010. Trends in cognitive sciences 14(8):357–364.
[2] Vul, E.; Alvarez, G.; Tenenbaum, J. B.; Black, M. J. 2009. Advances in Neural Information Processing Systems, 1955–1963.
[3] Denison, S.; Bonawitz, E.; Gopnik, A.; Griffiths, T. L. 2013. Cognition 126(2):280–300.
[4] Jin, D. Z.; Kozhevnikov, A. A. 2011. PLoS Comput. Biol 7(3), 1–9.
[5] Hopfield, J. J.; Tank, D. W. 1986. Science, Vol. 233, No. 4764.
[6] Cover, T. M. 1965. IEEE Trans Electron Comput, 14:326-334.
[7] Sompolinsky, H.; Kanter, I. 1986. Phys. Review Letters, 57(22):2861.