A Biologically Plausible Associative Memory Network | Mohadeseh Shafiei Kafraj

1. Mohadeseh Shafiei Kafraj (Gatsby Computational Neuroscience Unit, University College London)
2. Dmitry Krotov(MIT-IBM Watson AI Lab, IBM Research)
3. Brendan A. Bicknell1 (Gatsby Computational Neuroscience Unit, University College London)
4. Peter E. Latham (Gatsby Computational Neuroscience Unit, University College London)

The Hopfield network (Hopfield1982) has been the leading model for associative memory for over four decades, culminating in the recent 2024 Nobel Prize. However, the vanilla version of the Hopfield network has a capacity that scales with the number of connections per neuron (Roudi2007). In the mammalian brain, that's about 1,000, leading to a capacity of about 50 memories in a spiking network—regardless of its size. Therefore, it cannot possibly account for the capacity of human memory.

To address these limitations, various modifications to the Hopfield network have been proposed. One promising variant, Dense Associative Memory or the Modern Hopfield Network, incorporates a two-layer architecture with memory and feature neurons (krotov2021), which significantly increases storage and recall capacity. However, this model lacks biological plausibility in important ways. Its capacity is bounded by the number of memory neurons, and, critically, recalling a memory requires most neurons in the memory layer to remain silent. This behavior contradicts cortical dynamics, where neurons rarely remain silent for extended periods (Buzsáki2014).

This is not easy to fix: the memory layer contains a large number of neurons, and allowing silent neurons to exhibit even low firing rates can introduce an unacceptable level of noise, preventing the perfect recall of stored memories. To address these challenges, we propose a new biologically plausible model for associative memory. This model supports polynomial capacity while integrating dendritic computations, enabling non-recalled memory neurons to exhibit nonzero firing rates without compromising the perfect recall of a large

number of memories. The proposed architecture adheres to key biological constraints, including the presence of both excitatory and inhibitory populations that obey Dale's law and maintain non-saturated firing rates. These properties enhance the model's biological plausibility while achieving polynomial capacity, bridging the gap between theoretical and biological constraints on associative memory.