CellPress
REVIEWS

*Correspondence:
l.genzel@donders.ru.nl (L. Genzel).

https://doi.org/10.1016/j.tins.2020.04.011

References

1. Genzel, L. (2020) Memory and sleep: brain networks, cell dynamics and global states. *Curr. Opin. Behav. Sci.* 32, 72–79
2. Maingret, N. *et al.* (2016) Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nat. Neurosci.* 19, 959–964
3. Peyrache, A. *et al.* (2011) Inhibition recruitment in prefrontal cortex during sleep spindles and gating of hippocampal inputs. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17207–17212
4. Cowan, E. *et al.* (2020) Sleep spindles promote the restructuring of memory representations in ventromedial prefrontal cortex through enhanced hippocampal–cortical functional connectivity. *J. Neurosci.* 40, 1909–1919
5. Hennies, N. *et al.* (2016) Sleep spindle density predicts the effect of prior knowledge on memory consolidation. *J. Neurosci.* 36, 3799–3810
6. Tse, D. *et al.* (2007) Schemas and memory consolidation. *Science* 316, 76–82
7. Skelin, I. *et al.* (2019) Hippocampal coupling with cortical and subcortical structures in the context of memory consolidation. *Neurobiol. Learn. Mem.* 160, 21–31
8. Vogel, J.W. *et al.* (2020) A molecular gradient along the longitudinal axis of the human hippocampus informs large-scale behavioral systems. *Nat. Commun.* 11, 960
9. Buzsáki, G. and Tingley, D. (2018) Space and Time: the hippocampus as a sequence generator. *Trends Cogn. Sci.* 22, 853–869
10. Adamczyk, M. *et al.* (2015) Automatic sleep spindle detection and genetic influence estimation using continuous wavelet transform. *Front. Hum. Neurosci.* 9, 624

Spotlight

Network Dynamics Governed by Lyapunov Functions: From Memory to Classification

Merav Stern[1,*] and
Eric Shea-Brown[1,2]

In 1982, John Hopfield published a neural network model for memory retrieval, a model that became a cornerstone in theoretical neuroscience. In a recent paper, Krotov and Hopfield built on these early studies and showed how a network that incorporates a biologically plausible learning rule governed by a Lyapunov function can effectively perform classification tasks.

The Hopfield model [1] has proved to be one of the most influential theoretical models in neuroscience; for many researchers, it also served as a cornerstone in their introduction to the computational neuroscience world. At its core lies the idea that a pattern of sustained neural activity can represent a memory, which is retrievable via a network dynamics that converges into that memory. The model is defined by a connected network of neurons whose activity follows a dynamical update equation. To allow convergence of the network into stable memories, Hopfield used symmetric connectivity. This property facilitated analytical calculation of the memory capacity of the network, an approach that paved the way for many other results (e.g., [2]).

In the model, a memory is represented as a pattern of sustained activity in the network. The activity of each element in the network, a neuron, can have two possible values representing it being either active (spiking) or silent (not spiking). The ongoing activity of each neuron depends on its input, which results from the network activity at the previous time step and the weights of connections from other neurons. The memories are imprinted into the network via its symmetric connectivity structure. As a result, each memory in this model is a stable state of the network, and by following its dynamics the network has a better chance to converge to a certain memory as the correlation between the network state and the memory increases.

A crucial insight made by Hopfield is that the resulting model dynamics are governed by a Lyapunov function (also referred to as the energy function) – a mathematical function of the network activity that monotonically decreases whenever the state of the network changes and yet is bounded from below. Hence, a network with a Lyapunov function descends along the function's value to a local minimum. In the Hopfield model, these minima are governed by the memories imprinted into the network, as illustrated in Figure 1A. Importantly, this property is shared by more general continuous networks as well [3,4].

A recent paper by Krotov and Hopfield [5] presents another possible advantage for neural network dynamics that follows a Lyapunov function, this time to solve a classification task. In their new study, Krotov and Hopfield built a three-layer feedforward neural network, as illustrated in Figure 1B. The network connectivity between the input layer and the hidden (middle) layer dynamically changes in response to the activity driven by the inputs (unsupervised learning), following dynamics governed by a Lyapunov function. Later, the network connectivity between the hidden layer and the output layer is trained to classify the input patterns according to their labels (supervised learning). The network incorporates biologically plausible learning (e.g., as in [6–8]). Specifically, each of its connection weights between the input layer and the hidden layer follows a Hebbian-like plasticity rule [9,10] while being bounded by constraining the vector length of all weights.

Hopfield and Krotov's choice for the boundary between activity strength that drives potentiation and depression in their implementation of Hebbian plasticity follows an intriguing rule. This boundary is chosen by examining the activity of the hidden layer as driven by all of the input examples, leaving potentiation to drive only a few hidden units at a time while depressing others (and not updating connections where signs of inputs and weights are opposite).

In the original Hopfield model discussed earlier, a large correlation value between the network and a memory state pushes the network farther toward that memory while suppressing other, uncorrelated memories. The more recent model by Krotov and Hopfield uses a similar mechanism. Their choice to decrease weights (synapses) that are connected to a weakly

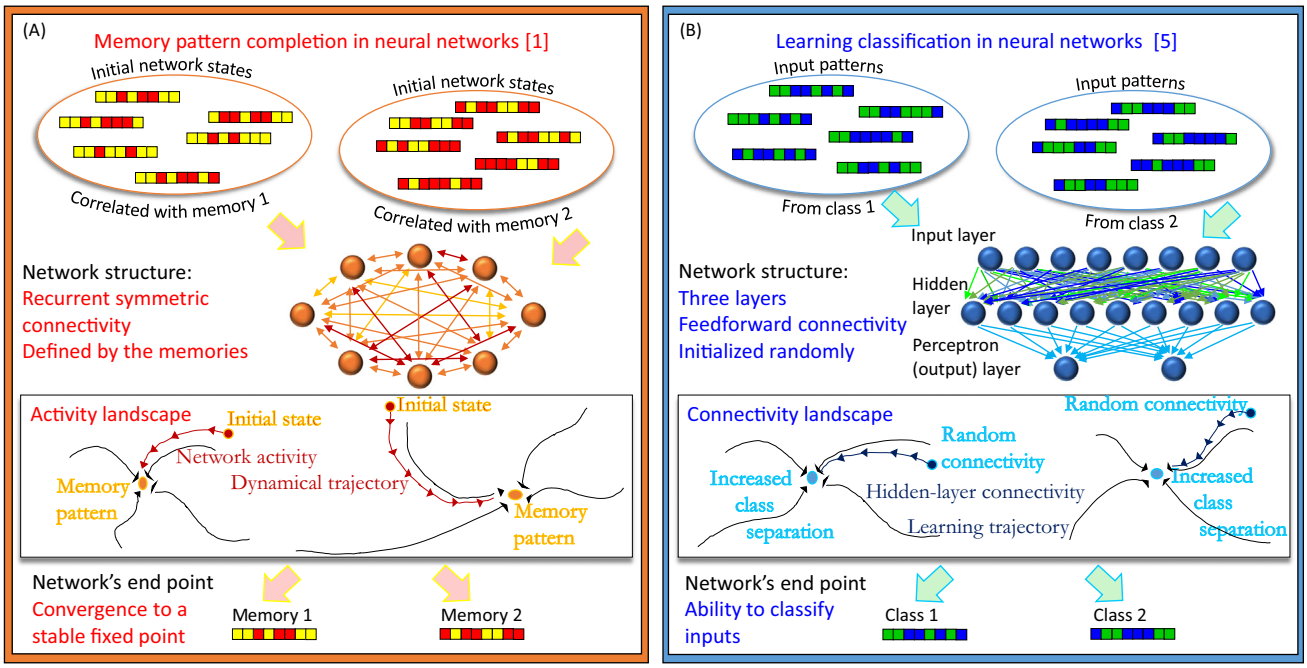

Figure 1. Structure and Activity of Neural Networks with Lyapunov Functions. (A) Left orange box: Hopfield neural network model [1] for memory retrieval. (B) Right blue box: Krotov and Hopfield neural network model [5] for classification. Top ellipsoids: Network initializations. The schematic illustrates two sets of six possible initializations for each of the networks. Each initialization is given by a vector with binary values, indicated by two possible colors and representing each neuron being in an active (spiking) or a silent (not spiking) state. Left, Hopfield network: Each element in the vector defines one neuron's initial state for each of the eight neurons in the network. All initial states in the right and left ellipsoids are highly correlated with memory state 1 and memory state 2, respectively (drawn at the bottom). Right, Krotov and Hopfield network: Each element in the vector defines an input pattern to one of the eight neurons in the input layer of the network. Here, input values are continuous rather than discrete and can be any real number. All input patterns in a set are highly correlated with an example that represents their class (these examples are drawn at the bottom). Middle: Network structures. Each node represents a neuron in the network. Each arrow represents an input from one neuron to another, forming together the network connectivity structure. Left, Hopfield network: The network connectivity is symmetric (evident by the dual-direction arrows). The connectivity structure is fully defined by the memories. Right, Krotov and Hopfield network: The network comprises three layers with feedforward connectivity between the layers, as the single-direction arrows indicate. Middle box: Network dynamics. The Lyapunov function landscape is illustrated with two attractors, each a local minimum of the function. The dynamics are attracted toward these, which serve as local stable states of the network. Left, Hopfield network: Each local minimum corresponds to a memory state. The network starts in an initial state and follows the dynamics. Here, each neuron is updated according to the input it receives from other neurons, defined by the network activity at the previous time step and the weights of connections. This dynamic descends along the Lyapunov function's value, as indicated by the red examples of possible trajectories, to local minima, indicated by orange circles, representing memories or stable states in the network. Right, Krotov and Hopfield network: Each local minimum corresponds to a set of connection weights between the input and the hidden layer. The network starts in a given connectivity structure and the connectivity weights between the input layer and the hidden layer follow the dynamics at hand. Here, weights change their values in response to the activity driven by the inputs, according to a Hebbian-like plasticity rule that is constrained by the vector length of all weights. This dynamic ascends along the Lyapunov function's value, as indicated by the blue examples of possible trajectories, to a local maximum, indicated by light-blue circles, representing attracting connectivity structures. These are stable states of the network connectivity. Bottom: End result of the network. Left, Hopfield network: The network final steady state can be memory 1 or memory 2, according to the initial state with which it had higher correlation. Right, Krotov and Hopfield network: The network final structure allows the classification of inputs according to the class with which the input had higher correlation.

activated neuron, for the majority of neurons in the hidden layer that do not cross their chosen boundary, causes patterns that strongly activate neurons in the hidden layer to push these neurons' synapses even farther toward these patterns while depressing synapses of neurons that were only weakly activated by these patterns.

The network follows a Lyapunov function defined by the integral of the function that sets the boundary between potentiation and depression. In the context of the Krotov and Hopfield network, this means that the dynamics lead this function to monotonically change until a local maximum has been reached, as illustrated in Figure 1B. Hence the network maximizes the responses of the hidden layers to inputs within the normalization constraints, a productive distribution of its weights.

When the training described earlier is complete, the activity of the hidden units is fed into an output layer of a perceptron, which

is trained using supervised learning (via stochastic gradient descent). Krotov and Hopfield show that this scheme can perform classification – of handwritten digits (MNIST), for example – with high accuracy.

The success of Krotov and Hopfield in showing how a biologically plausible learning rule can be governed by a Lyapunov function opens new doors to understanding how networks learn to perform classification, among other fundamental tasks. This is important, as efforts to analyze the underlying mechanisms are central to both neuroscience and artificial intelligence. It is also inspiring, as it comes almost 40 years after Hopfield's original discovery of how Lyapunov functions can govern the dynamics of memory. Noting how this laid the groundwork for so many of the studies that were to follow, we are keenly looking forward to a deepened understanding of classification and other computations in biological plausible networks in the years ahead.

Acknowledgments

We acknowledge support from the Sackler Foundation, the Swartz Foundation via the Swartz Center for Theoretical Neuroscience at the University of Washington, the NSF ERC Center for Sensorimotor Neural Engineering, and NSF DMS Grant #1514743. We thank the Allen Institute founders, Paul G. Allen and Jody Allen, for their vision, encouragement, and support.

[1]Department of Applied Mathematics, University of Washington, Seattle, WA 98195-3925, USA
[2]Allen Institute for Brain Science, Seattle, WA 98109-4307, USA

*Correspondence:
ms4325@uw.edu (M. Stern).

https://doi.org/10.1016/j.tins.2020.04.002

References

1. Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.* 79, 2554–2558
2. Amit, D.J. *et al.* (1985) Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.* 55, 1530
3. Cohen, M. and Grossberg, S. (1983) Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Trans. Syst. Man Cybern.* 13, 815–826
4. Hopfield, J.J. (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U. S. A.* 81, 3088–3092
5. Krotov, D. and Hopfield, J.J. (2019) Unsupervised learning by competing hidden units. *Proc. Natl. Acad. Sci. U. S. A.* 116, 7723–7731
6. Lillicrap, T.P. *et al.* (2016) Random synaptic feedback weights support error backpropagation for deep learning. *Nat. Commun.* 7, 13276
7. Lee, D.-H. *et al.* (2015) Difference target propagation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (Appice, A. *et al.*, eds), pp. 498–515, Springer
8. Pehlevan, C. and Chklovskii, D.B. (2014) A Hebbian/anti-Hebbian network derived from online non-negative matrix factorization can cluster and discover sparse features. In *2014 48th Asilomar Conference on Signals, Systems and Computers* (Matthews, M.B., ed.), pp. 769–775, IEEE
9. Hebb, D.O. (2005) *The Organization of Behavior: A Neuropsychological Theory*, Psychology Press
10. Oja, E. (1982) Simplified neuron model as a principal component analyzer. *J. Math. Biol.* 15, 267–273

Spotlight

Remyelination-Promoting Inflammation: Novel Role for MyD88 Signaling in Microglia/Macrophages

Gopal Murugaiyan,[1,*] Mai Fujiwara,[1] and Lucien P. Garo[1]

Inflammation in the central nervous system (CNS) has been linked to demyelination and remyelination. Using zebrafish and mouse models of demyelination and remyelination, Cunha *et al.* now describe a novel role for myeloid differentiation factor 88 (MyD88) signaling in supporting remyelination by promoting myeloid cell-mediated inflammatory responses via TNF-α, which are essential for phagocytic myelin debris clearance and for oligodendrogenesis.

Multiple sclerosis (MS) is a neurodegenerative disease triggered by inflammatory demyelination. Demyelinated MS lesions are caused by both adaptive and innate immune responses, which include infiltrating T cells and phagocytic myeloid cells (e.g., macrophages), as well as resident microglia [1]. The generation of new myelin sheaths around axons in the CNS is critical for counteracting the pathological loss of myelin. When remyelinating processes during the early stages of MS fail, patients present with permanent axonal loss and progressive disease [2]. Currently approved therapies can be effective in relapsing forms of MS, and work by inhibiting the inflammatory functions of immune cells and/or their trafficking into the CNS. However, current therapies do not directly promote remyelination and, thus, are ineffective in reversing progressive forms of MS [2]. Understanding the pathways that modulate remyelination is of immense clinical importance for MS and other demyelinating diseases.

Although various groups (including our own) have shown that inflammation drives demyelination [3], emerging evidence suggests that certain inflammatory events are necessary for remyelination after injury, particularly innate immune responses [4]. Upon myelin injury, peripheral macrophages and resident microglia are recruited to demyelinated lesions [4]. These phagocytes then create a conducive environment for remyelination by clearing myelin debris and by secreting soluble factors that promote the generation and maturation of myelinating oligodendrocytes [4]. However, the molecular pathways mediating this regenerative response in phagocytes remained unclear. Using zebrafish and mouse models of demyelination and remyelination, a recent study by Cunha *et al.* [5] describes a novel role for MyD88 signaling in supporting remyelination by promoting myeloid cell-mediated inflammatory responses via TNF-α, which are essential for phagocytic myelin debris clearance and oligodendrogenesis (Figure 1).