Global stability of a Hebbian/anti-Hebbian network for principal subspace learning

David Lipshutz* and Robert J. Lipshutz

January 19, 2026

Abstract

Biological neural networks self-organize according to local synaptic modifications to produce stable computations. How modifications at the synaptic level give rise to such computations at the network level remains an open question. Pehlevan et al. [1] proposed a model of a self-organizing neural network with Hebbian and anti-Hebbian synaptic updates that implements an algorithm for principal subspace analysis; however, global stability of the nonlinear synaptic dynamics has not been established. Here, for the case that the feedforward and recurrent weights evolve at the same timescale, we prove global stability of the continuum limit of the synaptic dynamics and show that the dynamics evolve in two phases. In the first phase, the synaptic weights converge to an invariant manifold where the 'neural filters' are orthonormal. In the second phase, the synaptic dynamics follow the gradient flow of a non-convex potential function whose minima correspond to neural filters that span the principal subspace of the input data.

1 Introduction

Biological neural networks self-organize according to local synaptic interactions that produce stable computations at the network-level. A challenge in theoretical neuroscience is to link these local interactions to stable network-level computations.

*Department of Neuroscience, Baylor College of Medicine and Neuroengineering Initiative, Rice University.

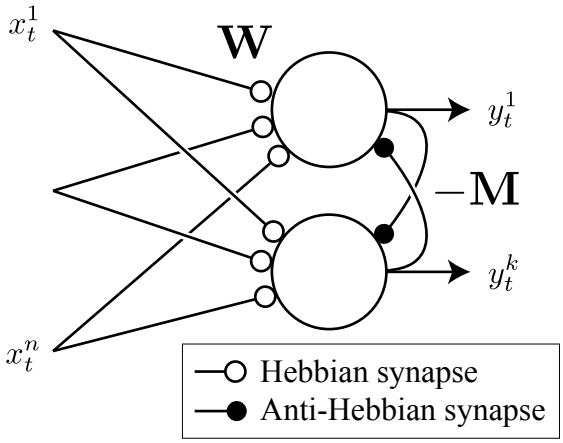

Figure 1: Hebbian/anti-Hebbian network for PSA. Single layer network with k neurons that receives n inputs. Feedforward Hebbian synapses $\mathbf{W}$ connect the n inputs to the k neurons and recurrent anti-Hebbian synapses $-\mathbf{M}$ connect the k neurons.

In a seminal work, Oja [2] proposed a computational model of a neuron as implementing an online algorithm for learning the top principal component of its input data using local, Hebbian synaptic updates. Oja's algorithm thus establishes a link between local, Hebbian interactions at the synaptic level and principal components analysis (PCA) at the neuron level. Moreover, the online (stochastic) algorithm is strikingly stable and data efficient—it is globally stable with a convergence rate that matches the information theoretic lower bound [3]—suggesting that neural networks with local synaptic updates may also be relevant in machine learning and neuromorphic computing applications.

In the multi-channel setting, much less is known about the stability of neural networks with local synaptic updates. Following Oja's work [2], there were several extensions to principal subspace analysis (PSA) algorithms that can be implemented in neural networks with local Hebbian and anti-Hebbian synaptic updates [1, 4–8]; however, these multi-channel networks include recurrent interactions that complicate their analyses and global stability of the networks' dynamics has not been established. In another line of work, networks with *non-local* synaptic updates for PSA have been proposed and analyzed [9–15], but these networks do not explain how local interactions give rise to stable network level computations.

The focus of this work is to analyze the global synaptic dynamics of a multi-channel network for

principal subspace analysis with Hebbian feedforward synapses and anti-Hebbian recurrent synapses introduced by Pehlevan et al. [1, 8] and depicted in Figure 1. Specifically, we show that the ordinary differential equation (ODE) associated with the synaptic dynamics is globally stable in the sense that for almost every initialization, the synaptic weights converge to an optimal configuration associated with PSA. We further show that the dynamics evolve in two phases. In the first phase, the synaptic weights converge to an invariant manifold characterized by orthonormality of the neural filters. In the second phase, the synaptic dynamics follow the gradient flow of non-convex potential function.

2 Hebbian/anti-Hebbian network model

In this section, we first review the Hebbian/anti-Hebbian neural network model for PSA proposed by Pehlevan et al. [1]. A detailed derivation of the network from a so-called "similarity matching" objective can be found in [1, 8]. We then introduce the ODE associated with the synaptic dynamics, which will be the focus of our analysis in this work.

Consider a network with k primary neurons that receives $n > k$ inputs, as illustrated in Figure 1. Feedforward synapses $\mathbf{W}$ connect the inputs to primary neurons, and recurrent synapses $-\mathbf{M}$ connect the primary neurons. At each timestep $t = 1, 2, \ldots$, the network receives an input vector $\mathbf{x}_t = (x_t^1, \ldots, x_t^n)$. The network operates on two timescales: a fast neural dynamics timescale and a slow synaptic update timescale.

In the first phase, the neural activities of the k output neurons, which are denoted by the k-dimensional vector $\mathbf{y}_t = (y_t^1, \ldots, y_t^k)$, evolve according to fast linear neural dynamics

$$\dot{\mathbf{y}}_t(\gamma) = \mathbf{W}_t \mathbf{x}_t - \mathbf{M}_t \mathbf{y}_t(\gamma),$$

which converge to $\mathbf{y}_t = \mathbf{F}_t \mathbf{x}_t$, where $\mathbf{W}_t$ and $\mathbf{M}_t$ respectively denote the states of the feedforward and recurrent synaptic weights at time t, and the row vectors of the $k \times n$ matrix $\mathbf{F}_t := \mathbf{M}_t^{-1} \mathbf{W}_t$ are referred to as the 'neural filters'. After the neural activities converge, the synaptic weights are

updated according to the slow synaptic plasticity rules

$$\mathbf{W}_{t+1} = \mathbf{W}_t + 2\eta(\mathbf{y}_t\mathbf{x}_t^\top - \mathbf{W}_t) \tag{1}$$

$$\mathbf{M}_{t+1} = \mathbf{M}_t + \frac{\eta}{\tau}(\mathbf{y}_t\mathbf{y}_t^\top - \mathbf{M}_t), \tag{2}$$

where $\eta > 0$ is the learning rate for the feedforward synapses $\mathbf{W}$ and τ denotes the ratio between the learning rates for the feedforward synapses $\mathbf{W}$ and the lateral synapses $-\mathbf{M}$. The plasticity rule for the feedforward synapses $\mathbf{W}$ includes a term proportional to the product of the pre- and postsynaptic activities, $\mathbf{y}_t\mathbf{x}_t^\top$, so it is referred to as 'Hebbian'. The plasticity rule for the lateral synapses $-\mathbf{M}$ includes a term inversely proportional to the product of the pre- and postsynaptic activities, $\mathbf{y}_t\mathbf{y}_t^\top$, so it is referred to as 'anti-Hebbian'.

To analyze the stability of their algorithm, Pehlevan et al. [8] considered the continuum limit of the updates. Formally, when the input data $\{\mathbf{x}_t\}$ are independent and identically distributed samples with zero mean and fixed $n \times n$ covariance matrix $\mathbf{A}$ and the step size $\eta > 0$ is infinitesimally small, the synaptic dynamics can be approximated by the ODE

$$\frac{1}{2}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t} = \mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A} - \mathbf{W}(t) \tag{3}$$

$$\tau\frac{\mathrm{d}\mathbf{M}(t)}{\mathrm{d}t} = \mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A}\mathbf{W}(t)^\top\mathbf{M}(t)^{-1} - \mathbf{M}(t). \tag{4}$$

The relationship between the online algorithm and the ODE can be made precise for certain time-dependent learning rates under appropriate regularity conditions (e.g., the spectrum of $\mathbf{M}$ is uniformly bounded away from zero) [16, 17].

Pehlevan et al. [8] proved that every equilibrium point of the ODE corresponds to an eigen-subspace of $\mathbf{A}$, and, when $\tau \leq \frac{1}{2}$, all linearly stable equilibrium points correspond to the *principal* (eigen-)subspace of $\mathbf{A}$. While their theoretical analysis is informative about the synaptic dynamics near the equilibrium points, it is not informative about the synaptic dynamics away from the equilibrium points, which corresponds to most random initializations. Pehlevan et al. [1, 8] provide empirical evidence that both the online algorithm and ODE are globally stable, and they benchmark both against comparable algorithms; however, a theoretical analysis of the global dynamics remains

an open problem. The focus of this work is to prove global stability of the ODE.

The ODE (3)–(4) is naturally viewed as the gradient descent-ascent flow (with timescale separation τ) for solving the nonconvex-concave minimax problem

$$\min_{\mathbf{W}} \max_{\mathbf{M}} f(\mathbf{W}, \mathbf{M}), \qquad f(\mathbf{W}, \mathbf{M}) := \operatorname{Tr}\left(-\mathbf{M}^{-1}\mathbf{W}\mathbf{A}\mathbf{W}^\top + \mathbf{W}\mathbf{W}^\top - \frac{1}{2}\mathbf{M}^2\right). \tag{5}$$

where the minimization is over the set of $k \times n$ matrices $\mathbb{R}^{k\times n}$ and the maximization is over the set of $k \times k$ positive definite matrices $\mathcal{S}^k_{++}$. Analogously, the discrete algorithm (1)–(2) is naturally interpreted as a stochastic gradient descent-ascent algorithm for solving the minimax problem. In general, proving convergence of gradient descent-ascent algorithms for nonconvex-concave minimax problems is challenging and existing results [16, 18, 19] rely on a separation of time-scales (i.e., letting $\tau \to 0$).[1] However, numerical experiments indicate that the optimal convergence rate occurs around $\tau = \frac{1}{2}$, see [8, Figure 4], suggesting that a separation of time-scales is unnecessary for proving global convergence and that $\tau = \frac{1}{2}$ may be a parameter of particular interest.

3 Global stability of the synaptic dynamics

For the case $\tau = \frac{1}{2}$, we prove global convergence of the ODE (3)–(4) to the desired principal subspace. We assume $\mathbf{A}$ is a positive definite $n \times n$ matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n > 0$ and $k < n$ is fixed. The following theorem is our main result.

Theorem 1. *Let $\mathbf{A} \in \mathcal{S}^n_{++}$. For every $\tau > 0$ and $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} := \mathbb{R}^{k\times n} \times \mathcal{S}^k_{++}$, there exists a unique solution $(\mathbf{W}(t), \mathbf{M}(t))$ of the ODE (3)–(4) with initial condition $(\mathbf{W}_0, \mathbf{M}_0)$ for all $t \geq 0$. Moreover, suppose $\lambda_k > \lambda_{k+1}$ and $\tau = \frac{1}{2}$. Then there is a set $\mathcal{Z}$ with Lebesgue measure zero such that if $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} \setminus \mathcal{Z}$, then the solution $(\mathbf{W}(t), \mathbf{M}(t))$ converges, as $t \to \infty$, to the set of equilibrium points $(\mathbf{W}_*, \mathbf{M}_*)$ of the ODE such that the neural filters (i.e., the row vectors of $\mathbf{F}_* := \mathbf{M}_*^{-1}\mathbf{W}_*$) are orthonormal and span the principal subspace of $\mathbf{A}$.*

The next two sections and appendix A are devoted to the proof of Theorem 1. In appendix

[1] During the preparation of this manuscript, we became aware of the preprint [19] that also analyzes the stability of the ODE (3)–(4). However, the analysis considers the regime $\tau \to 0$ whereas we treat the case $\tau = \frac{1}{2}$.

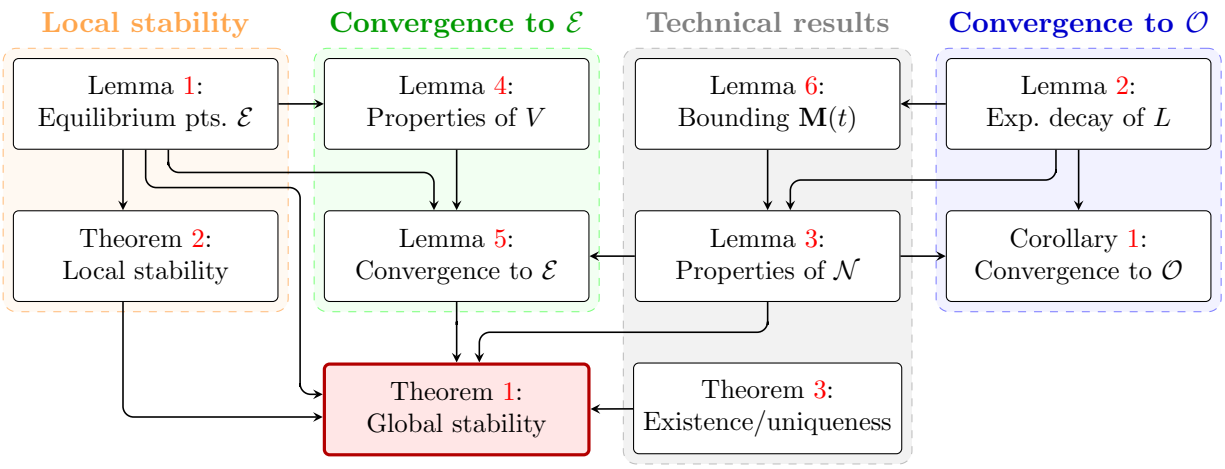

Figure 2: Dependency diagram for our results. Local stability of equilibrium points (orange region) is established in section 4. Convergence of solutions to the invariant manifold $\mathcal{O}$ (blue region) is shown in section 5.1. Convergence starting in or near the invariant manifold to the equilibrium points $\mathcal{E}$ (green region) is shown in section 5.2. Finally, global stability of the ODE (red box) is shown in section 5.3. Technical results (gray region) are proved in appendices A and C.

A we prove existence and uniqueness of solutions, which ensures that solutions that are initialized in $\mathcal{D}$ remain in $\mathcal{D}$ for all $t \geq 0$. In section 4, we review results from [8] on local linear stability of equilibrium points, denoted $\mathcal{E}$. Global convergence is proved in section 5. We show that the synaptic weights evolve in two phases. In the first phase, for any initialization outside of a null set $\mathcal{N}$, the synaptic weights first converge to the invariant manifold $\mathcal{O}$ corresponding to orthonormal neural filters; that is, the set of $(\mathbf{W}, \mathbf{M}) \in \mathcal{D}$ such that $\mathbf{F} := \mathbf{M}^{-1}\mathbf{W}$ has orthonormal row vectors (section 5.1 and Figure 3). This convergence is captured by a convex Lyapunov function $L(\mathbf{W}, \mathbf{M})$. Then, in the second phase, starting on (or near) the invariant manifold $\mathcal{O}$, the synaptic dynamics are approximated by the gradient flow of a non-convex potential function $V(\mathbf{W})$ (section 5.2 and Figure 4). As a result, for almost any initialization, the synaptic weights converge to an equilibrium point such that the neural filters correspond to the desired principal subspace projection. A dependency diagram for our main result is shown in Figure 2.

In section 6, we provide empirical evidence that the synaptic weights also evolve in two phases for the discrete online algorithm (1)–(2). In section 7, we conjecture that Theorem 1 can be

generalized to hold for all $0 < \tau \leq \frac{1}{2}$.

4 Local stability of equilibrium points

Next, we characterize the equilibrium points of the ODE (3)–(4) and recall results by Pehlevan et al. [8] on their linear stability. To this end, let

$$\mathcal{E} := \{(\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \mathbf{G}(\mathbf{W}, \mathbf{M}) = \mathbf{0}\}$$

denote the set of equilibrium points, where $\mathbf{G} : \mathcal{D} \mapsto \mathcal{D}$ is the vector field defined by

$$\mathbf{G}(\mathbf{W}, \mathbf{M}) := \left(2\mathbf{M}^{-1}\mathbf{W}\mathbf{A} - 2\mathbf{W}, \frac{1}{\tau}(\mathbf{M}^{-1}\mathbf{W}\mathbf{A}\mathbf{W}^\top\mathbf{M}^{-1} - \mathbf{M})\right).$$

4.1 Characterization of equilibrium points

The following lemma, whose proof is given in appendix B, characterizes the equilibrium points in terms of their singular value decompositions (SVDs).

Lemma 1. *Assume $\tau > 0$ and $\mathbf{A} \in \mathcal{S}^n_{++}$. A pair $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}$ if and only if $\mathbf{W}_* = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ and $\mathbf{M}_* = \mathbf{U}\mathbf{S}\mathbf{U}^\top$, where $\mathbf{U}$ is any $k \times k$ orthogonal matrix, $\mathbf{V}$ is a $d \times k$ matrix whose column vectors are orthonormal eigenvectors of the covariance matrix $\mathbf{A}$, and $\mathbf{S}$ is a $k \times k$ diagonal matrix whose diagonal entries are the eigenvalues of $\mathbf{A}$ corresponding to the column vectors of $\mathbf{V}$.*

As a consequence of Lemma 1, if $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}$, then $\mathbf{W}_*$ is full rank (since we assume the eigenvalues of $\mathbf{A}$ are positive) and $\mathbf{F}_* := \mathbf{M}_*^{-1}\mathbf{W}_* = \mathbf{U}\mathbf{V}^\top$, where $\mathbf{U}$ is a $k \times k$ orthogonal matrix and the column vectors of $\mathbf{V}$ are orthonormal eigenvectors of $\mathbf{A}$. In other words, the row vectors of $\mathbf{F}_*$ (i.e., the neural filters) are orthonormal and span an eigen-subspace of $\mathbf{A}$. In addition, if $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}$, then $(\mathbf{Q}\mathbf{W}_*, \mathbf{Q}\mathbf{M}_*\mathbf{Q}^\top) \in \mathcal{E}$ for every $k \times k$ orthogonal matrix $\mathbf{Q}$. In particular, each equilibrium point is an element of a $\frac{k(k-1)}{2}$ dimensional manifold of equilibrium points corresponding to an eigen-subspace of $\mathbf{A}$.

4.2 Local linear stability analysis

For our purposes, we say an equilibrium point $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}$ is 'linearly stable' if all of the eigenvalues of the Jacobian of $\mathbf{G}$ evaluated at $(\mathbf{W}_*, \mathbf{M}_*)$ have nonpositive real part and it is 'linearly unstable' if it is not linearly stable; that is, at least one of the eigenvalues of the Jacobian of $\mathbf{G}$ evaluated at $(\mathbf{W}_*, \mathbf{M}_*)$ has positive real part. Let

$$\mathcal{E}_0 := \{(\mathbf{W}, \mathbf{M}) \in \mathcal{E} : \text{the rows of } \mathbf{W} \text{ span the principal subspace of } \mathbf{A}\}.$$

The next result follows immediately from [8, Theorem 1].

Theorem 2. *Suppose* $0 < \tau \leq \frac{1}{2}$ *and* $\mathbf{A} \in \mathcal{S}^n_{++}$. *An equilibrium point* $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}$ *is linearly stable if and only if* $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}_0$.

5 Global convergence analysis

We now prove our main results on global stability of the ODE (3)–(4) for $\tau = \frac{1}{2}$, which is fixed throughout this section. Figure 3 shows the vector field $\mathbf{G}(W, M)$ in the scalar case $k = n = 1$. In section 5.1, we show that the neural filters are asymptotically orthonormal—in Figure 3, this corresponds to the initial convergence of trajectories to the blue line. Then, in section 5.2, we show that starting from (or near) an orthonormal initialization, the ODE (3) governing the dynamics of the feedforward weights $\mathbf{W}$ can be approximated as gradient flow of a potential function whose minima correspond to the principal subspace—in Figure 3, this corresponds to the convergence of trajectories with initial conditions on (or near) the blue line to the red dots. Finally, in section 5.3, we combine these results to prove Theorem 1.

5.1 Asymptotic orthonormality of the neural filters

In this section, we show that for almost every initialization, the solution $(\mathbf{W}(t), \mathbf{M}(t))$ to the ODE (3)–(4) converges to the following invariant subset of matrices in $\mathcal{D}$ that correspond to orthonormal

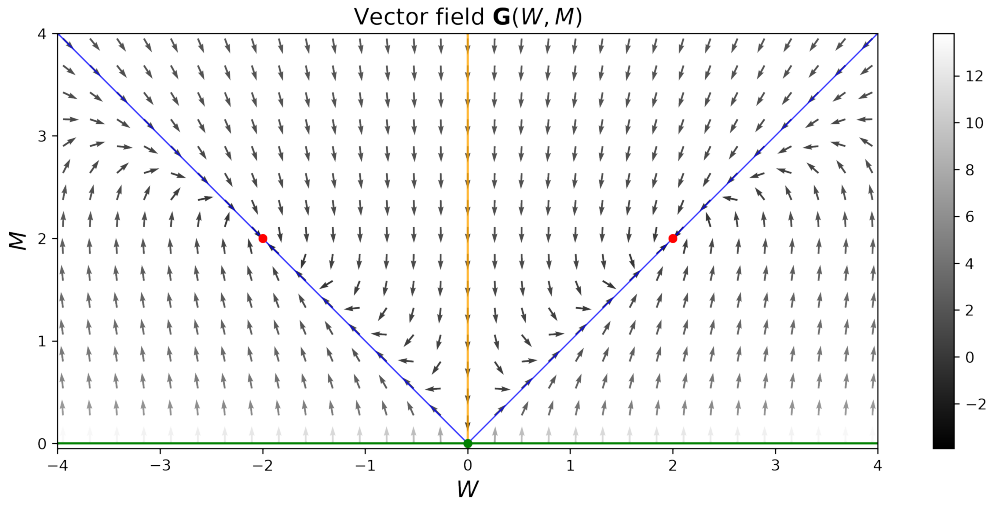

Figure 3: Plot of the vector field $\mathbf{G}(W, M)$ in the case $k = n = 1$ and $\lambda_1 = 2$. The grayscale indicates the *logarithm* of the vector magnitude. The blue lines denote the set $\mathcal{O}$, the orange vertical line denotes the set $\mathcal{N}$, the 2 red dots denote the set $\mathcal{E}_0$ (which is equal to $\mathcal{E}$ in this case), and the green line indicates that the line $M = 0$ does not belong to $\mathcal{D} = \mathbb{R} \times (0, \infty)$.

neural filters $\mathbf{F} = \mathbf{M}^{-1}\mathbf{W}$:

$$\mathcal{O} := \left\{ (\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \mathbf{M}^{-1}\mathbf{W}\mathbf{W}^\top\mathbf{M}^{-1} = \mathbf{I}_k \right\}. \tag{6}$$

To show this convergence, we define the following convex Lyapunov function on $\mathcal{D}$:

$$L(\mathbf{W}, \mathbf{M}) := \|\mathbf{W}\mathbf{W}^\top - \mathbf{M}^2\|^2 = \mathrm{Tr}\left[(\mathbf{W}\mathbf{W}^\top - \mathbf{M}^2)^2\right]. \tag{7}$$

Note that $L(\mathbf{W}, \mathbf{M})$ is nonnegative everywhere and equal to zero if and only $(\mathbf{W}, \mathbf{M}) \in \mathcal{O}$.

Lemma 2. *Suppose* $\mathbf{A} \in \mathcal{S}^n_{++}$ *and* $(\mathbf{W}(t), \mathbf{M}(t))$ *is a solution of the ODE* (3)–(4) *with* $\tau = \frac{1}{2}$ *and initial condition* $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}$. *Then*

$$L(\mathbf{W}(t), \mathbf{M}(t)) = L(\mathbf{W}_0, \mathbf{M}_0)e^{-8t}, \qquad t \geq 0. \tag{8}$$

Proof. First, note that the ODE (3)–(4) implies

$$\begin{aligned}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}\mathbf{W}(t)^\top - \frac{\mathrm{d}\mathbf{M}(t)}{\mathrm{d}t}\mathbf{M}(t) &= 2\mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A}\mathbf{W}(t)^\top - 2\mathbf{W}(t)\mathbf{W}(t)^\top \\ &\quad - \left[2\mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A}\mathbf{W}(t)^\top\mathbf{M}(t)^{-1}\mathbf{M}(t) - 2\mathbf{M}(t)^2\right] \\ &= -2\left[\mathbf{W}(t)\mathbf{W}(t)^\top - \mathbf{M}(t)^2\right].\end{aligned}$$

Therefore, by the chain rule, the product rule, the cyclic property of the trace rule, and the previous display,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}L(\mathbf{W}(t),\mathbf{M}(t)) &= 4\,\mathrm{Tr}\left[\left(\mathbf{W}(t)\mathbf{W}(t)^\top - \mathbf{M}(t)^2\right)\left(\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}\mathbf{W}(t)^\top - \frac{\mathrm{d}\mathbf{M}(t)}{\mathrm{d}t}\mathbf{M}(t)\right)\right] \\ &= -8L(\mathbf{W}(t),\mathbf{M}(t)).\end{aligned}$$

Solving the differential equation by separation of variables yields equation (8). □

To prove that the neural filters are asymptotically orthonormal, we need a technical lemma that states conditions under which the solution $(\mathbf{W}(t),\mathbf{M}(t))$ does not converge to zero. To this end, define the set

$$\mathcal{N} = \{(\mathbf{W},\mathbf{M}) \in \mathcal{D} : \exists\,\mathbf{v} \in \mathbb{R}^k,\ \lambda > 0,\ \text{such that } \mathbf{W}^\top\mathbf{v} = \mathbf{0} \text{ and } \mathbf{M}\mathbf{v} = \lambda\mathbf{v}\}. \tag{9}$$

Then $\mathcal{N}$ is the set of pairs $(\mathbf{W},\mathbf{M}) \in \mathcal{D}$ such that $\mathbf{W}\mathbf{W}^\top$ is singular *and* there is an eigenvector of $\mathbf{M}$ in the null space of $\mathbf{W}\mathbf{W}^\top$. Note that if $\mathbf{W}$ is full rank and $\mathbf{M}$ is positive definite, then $(\mathbf{W},\mathbf{M}) \notin \mathcal{N}$. The set $\mathcal{N}$ corresponds to the orange vertical line in Figure 3. The following technical lemma establishes that if for any initialization not in $\mathcal{N}$, the solution to the ODE (3)–(4) remains bounded away from zero and infinity. The proof is provided in appendix C.

Lemma 3. *The set $\mathcal{N}$ has Lebesgue measure zero. Suppose $\mathbf{A} \in \mathcal{S}^n_{++}$ and $(\mathbf{W}(t),\mathbf{M}(t))$ is the solution of the ODE (3)–(4) with $\tau = \frac{1}{2}$ and starting at $(\mathbf{W}_0,\mathbf{M}_0) \in \mathcal{D}$. If $(\mathbf{W}_0,\mathbf{M}_0) \notin \mathcal{N}$, then*

$(\mathbf{W}(t), \mathbf{M}(t)) \notin \mathcal{N}$ *for all* $t \geq 0$ *and*

$$\limsup_{t\to\infty} \left\{ \|\mathbf{M}(t)^{-1}\| + \|\mathbf{W}(t)\| \right\} < \infty.$$

On the other hand, if $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{N}$*, then* $(\mathbf{W}(t), \mathbf{M}(t)) \in \mathcal{N}$ *for all* $t \geq 0$ *and*

$$\lim_{t\to\infty} \det(\mathbf{M}(t)) = 0.$$

The following corollary of Lemmas 2 and 3 states that almost every solution of the ODE converges exponentially to the invariant manifold $\mathcal{O}$.

Corollary 1. *Suppose* $\mathbf{A} \in \mathcal{S}^n_{++}$ *and* $(\mathbf{W}(t), \mathbf{M}(t))$ *is a solution of the ODE* (3)–(4) *with* $\tau = \frac{1}{2}$ *and initial condition* $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} \setminus \mathcal{N}$*. Then* $(\mathbf{W}(t), \mathbf{M}(t))$ *converges to* $\mathcal{O}$ *as* $t \to \infty$*.*

Proof. By Lemma 3, $K := \sup\{\|\mathbf{M}(t)^{-1}\| : t \geq 0\} < \infty$. Thus,

$$\begin{aligned}\|\mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{W}(t)^\top\mathbf{M}(t)^{-1} - \mathbf{I}_k\|^2 &\leq \|\mathbf{M}(t)^{-2}\|^2\|\mathbf{W}(t)\mathbf{W}(t)^\top - \mathbf{M}(t)^2\|^2 \\ &\leq K^4 e^{-8t}\|\mathbf{W}_0\mathbf{W}_0^\top - \mathbf{M}_0^2\|^2,\end{aligned}$$

where the first inequality is due to the Cauchy-Schwarz inequality, and the second inequality follows from Lemma 2. □

5.2 Convergence to equilibrium points

Having shown that $(\mathbf{W}(t), \mathbf{M}(t))$ converges to $\mathcal{O}$ as $t \to \infty$, we now analyze the dynamics when $(\mathbf{W}(t), \mathbf{M}(t))$ is near the set $\mathcal{O}$. To begin, consider the case $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{O}$ so that $(\mathbf{W}(t), \mathbf{M}(t)) \in \mathcal{O}$ for all $t \geq 0$. Then we can rewrite the right-hand side of the ODE (3) as a function of $\mathbf{W}(t)$ only:

$$\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t} = 2(\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\mathbf{W}(t)\mathbf{A} - 2\mathbf{W}(t) = -2\nabla V(\mathbf{W}(t)), \tag{10}$$

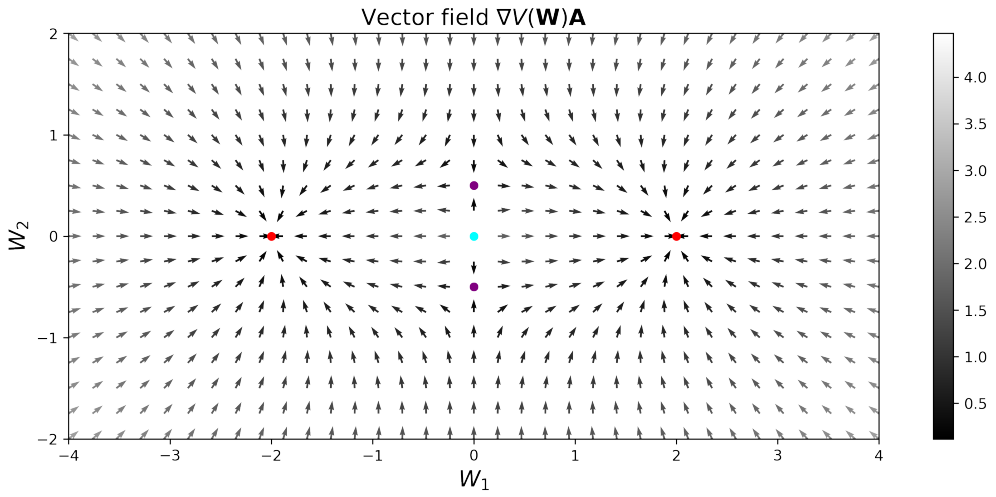

Figure 4: Plot of the vector field $-\nabla V(\mathbf{W})$ in the case $n = 2$, $k = 1$ and $\mathbf{A} = \mathrm{diag}(2, \frac{1}{2})$. The grayscale indicates the vector magnitude. The red dots denote the global minima of V, the purple dots denote the saddle points of V, and the cyan dot at the origin denotes the set $\{\mathbf{W} : \det(\mathbf{W}\mathbf{W}^\top) = 0\}$.

where $V : \mathbb{R}^{k\times n} \mapsto \mathbb{R}$ is the nonconvex potential function

$$V(\mathbf{W}) := f\left(\mathbf{W}, (\mathbf{W}\mathbf{W}^\top)^{\frac{1}{2}}\right) = \mathrm{Tr}\left[-(\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\mathbf{W}\mathbf{A}\mathbf{W}^\top + \frac{1}{2}\mathbf{W}\mathbf{W}^\top\right], \tag{11}$$

where $f(\mathbf{W}, \mathbf{M})$ is defined in (5). Therefore, we can interpret the $\mathbf{W}(t)$ dynamics on $\mathcal{O}$ as the gradient flow of the potential function V. (Note that V is only differentiable on the subset of full-rank matrices $\mathbf{W}$ in $\mathbb{R}^{k\times n}$.) In Figure 4, we plot the vector field $-\nabla V(\mathbf{W})$ in the case $d = 2$ and $k = 1$ to illustrate the dynamics of $\mathbf{W}(t)$ on the set $\mathcal{O}$.

Lemma 4. *Suppose $\mathbf{A} \in \mathcal{S}^n_{++}$. The function V is bounded below. Furthermore, $\nabla V(\mathbf{W})$ exists and satisfies $\nabla V(\mathbf{W}) = 0$ if and only if $\mathbf{W}$ is full rank and $(\mathbf{W}, (\mathbf{W}\mathbf{W})^{\frac{1}{2}}) \in \mathcal{E}$.*

Proof. Let $\mathbf{W} \in \mathbb{R}^{k\times n}$ and let $\mathbf{W} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ denote its SVD. Then

$$\begin{aligned} V(\mathbf{W}) &\geq \mathrm{Tr}\left[-\sigma_{\max}(\mathbf{A})(\mathbf{W}\mathbf{W}^\top)^{\frac{1}{2}} + \frac{1}{2}\mathbf{W}\mathbf{W}^\top\right] \\ &\geq -\sigma_{\max}(\mathbf{A})\,\mathrm{Tr}(\mathbf{S}) + \frac{1}{2}\,\mathrm{Tr}(\mathbf{S}^2) \end{aligned}$$

$$\geq -\frac{k}{2}\sigma_{\max}^2(\mathbf{A}),$$

where $\sigma_{\max}(\mathbf{A})$ denotes the spectral norm of $\mathbf{A}$. Since this holds for all $\mathbf{W}$, the function V is bounded below. Next, suppose $\mathbf{W}$ is full rank and $\nabla V(\mathbf{W}) = 0$. Then $2(\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\mathbf{W}\mathbf{A}\mathbf{W}^\top = (\mathbf{W}\mathbf{W}^\top)^{\frac{1}{2}}$ and so $\mathbf{V}^\top\mathbf{A}\mathbf{V} = \mathbf{S}$. Since $\mathbf{S}$ is diagonal, it follows that the column vectors of $\mathbf{V}$ are eigenvectors of $\mathbf{A}$ and the diagonal entries of $\mathbf{S}$ are the corresponding eigenvalues of $\mathbf{A}$. Thus, by Lemma 1, $(\mathbf{W}, (\mathbf{W}\mathbf{W}^\top)^{\frac{1}{2}}) \in \mathcal{E}$. The converse is readily verified by substitution. □

It follows from LaSalle's invariance principle [20], the ODE (10) and Lemma 4 that when initialized with $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{O}$, the solution $(\mathbf{W}(t), \mathbf{M}(t))$ converges to a fixed point of V, which corresponds to an equilibrium point in $\mathcal{E}$. Next, for the general case $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} \setminus \mathcal{N}$, we can rewrite the right-hand side of the ODE (3) as follows:

$$\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t} = -2\nabla V(\mathbf{W}(t)) + \left[\mathbf{M}(t)^{-1} - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\right]\mathbf{W}(t)\mathbf{A},$$

where we recall that $\mathbf{W}(t)\mathbf{W}(t)^\top$ is non-singular for all $t \geq 0$ by Lemma 3. By the chain rule,

$$\frac{\mathrm{d}V(\mathbf{W}(t))}{\mathrm{d}t} \leq -2\|\nabla V(\mathbf{W}(t))\|^2 + \|\nabla V(\mathbf{W}(t))\mathbf{A}\mathbf{W}(t)^\top\|\|\mathbf{M}(t)^{-1} - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\|.$$

We claim that

$$\limsup_{t\to\infty} \|\nabla V(\mathbf{W}(t))\mathbf{A}\mathbf{W}(t)^\top\|\|\mathbf{M}(t)^{-1} - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\| = 0.$$

Assuming the claim holds, we have

$$\limsup_{t\to\infty} \frac{\mathrm{d}V(\mathbf{W}(t))}{\mathrm{d}t} \leq 0.$$

Therefore, by LaSalle's invariance principle, $\mathbf{W}(t)$ converges to the set of fixed points of V, which correspond to the set of equilibrium points $\mathcal{E}$. We summarize this result in the following lemma, and provide a detailed proof in appendix D.

Lemma 5. *Suppose* $\mathbf{A} \in \mathcal{S}^n_{++}$ *and* $(\mathbf{W}(t), \mathbf{M}(t))$ *is a solution to ODE* (3)–(4) *with* $\tau = \frac{1}{2}$ *and initial condition* $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} \setminus \mathcal{N}$. *Then* $(\mathbf{W}(t), \mathbf{M}(t))$ *converges to the set* $\mathcal{E}$ *as* $t \to \infty$.

5.3 Proof of Theorem 1

Existence and uniqueness of solutions is shown in Theorem 3 of appendix A. We now combine Lemma 5 and Theorem 2 to prove global convergence. Define the subset

$$\mathcal{U} := \{(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} : (\mathbf{W}(t), \mathbf{M}(t)) \text{ converges to } \mathcal{E} \setminus \mathcal{E}_0 \text{ as } t \to \infty\}$$

of initializations whose trajectories converge to the set of linearly unstable equilibrium points. If $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E} \setminus \mathcal{E}_0$, then by Theorem 2, the Jacobian of $\mathbf{G}$ evaluated at $(\mathbf{W}_*, \mathbf{M}_*)$ has an eigenvalue with positive real part. Therefore, by [21, Proposition 3], the set $\mathcal{U}$ has Lebesgue measure zero. Along with Lemma 3, this implies that the set $\mathcal{Z} := \mathcal{N} \cup \mathcal{U}$ also has Lebesgue measure zero. Suppose $(\mathbf{W}(t), \mathbf{M}(t))$ is a solution of the ODE (3)–(4) with initial condition $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D} \setminus \mathcal{Z}$. Then by Lemma 5 and the definition of $\mathcal{U}$, $(\mathbf{W}(t), \mathbf{M}(t))$ converges to the set $\mathcal{E}_0$ as $t \to \infty$. Finally, by Lemma 1 and the definition of $\mathcal{E}_0$, for every $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}_0$, the row vectors of $\mathbf{F}_* := \mathbf{M}_*^{-1}\mathbf{W}_*$ are orthonormal and span the principal subspace of $\mathbf{A}$.

6 Comparing the ODE and the online algorithm

We now use numerical simulations to examine whether the two-phase convergence observed in the ODE (3)–(4) also appears in the online algorithm (1)–(2).

Setup. We consider a network with $d = 4$ inputs and $k = 2$ neurons. The inputs $\{\mathbf{x}_t\}$ are sampled i.i.d. from a mean zero normal distribution with covariance matrix $\mathbf{A} = \mathrm{diag}(0.5, 0.25, 0.2, 0.05)$. The feedforward weight matrix $\mathbf{W}$ is initialized to have i.i.d. standard normal entries and the recurrent weight matrix $\mathbf{M}$ is initialized to be diagonal with entries sampled uniformly from $[1, 2]$. The ODE (3)–(4) is solved using the built-in Mathematica function `NDSolve`. The online algorithm (1)–(2) is simulated with time-dependent step size $\eta_t = \frac{c_0}{c_1 + t}$, where $c_0, c_1 > 0$ are chosen so that

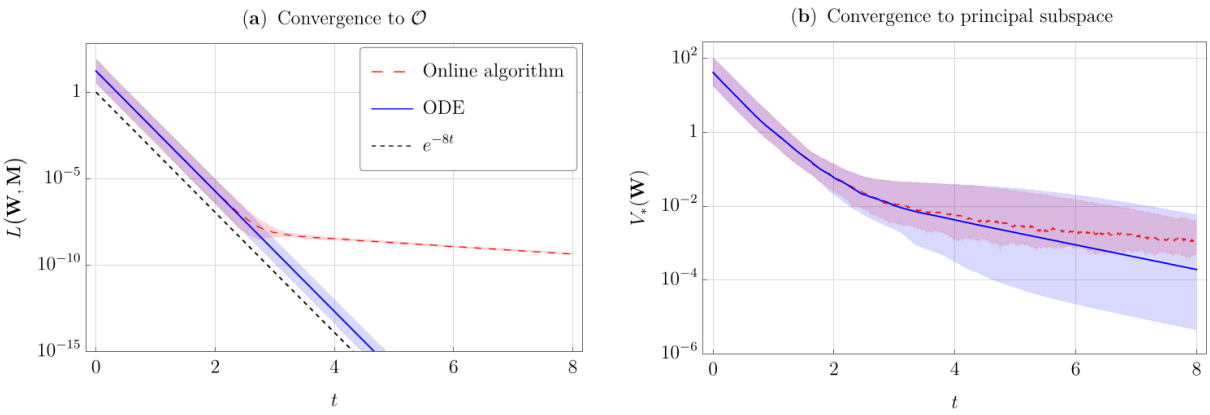

Figure 5: Convergence of the ODE and the online algorithm in a network with $d = 4$ inputs and $k = 2$ neurons. **(a)** Convergence to the invariant manifold $\mathcal{O}$, measured using the Lyapunov function $L(\mathbf{W}, \mathbf{M})$. **(b)** Convergence to the principal subspace, measured using $V_*(\mathbf{W}) = V(\mathbf{W}) - V(\mathbf{W}_*)$, where $(\mathbf{W}_*, \mathbf{M}_*)$ is any stable equilibrium point of the ODE. In panels **(a)** and **(b)**, shaded regions indicate the middle 80th percentile over 100 random initializations; solid/dotted lines denote the median values. Simulation details are in the main text.

$\eta_1 = 0.001$ and $\sum_{t=1}^{25{,}000} \eta_t = 8$. To track convergence, we evaluate the Lyapunov functions $L(\mathbf{W}, \mathbf{M})$ and $V_*(\mathbf{W}) = V(\mathbf{W}) - V(\mathbf{W}_*)$, where $(\mathbf{W}_*, \mathbf{M}_*)$ is any stable critical point of the ODE. The ODE and online algorithm are simulated with 100 random initializations.

Results. Figure 5 illustrates the two-phase convergence for both the ODE and the online algorithm. Panel **(a)** shows that $L(\mathbf{W}, \mathbf{M})$ converges exponentially with decay rate e^{-8t} (dotted black line) for the ODE (solid blue line), consistent with Lemma 2. For the online algorithm (dashed red line), $L(\mathbf{W}, \mathbf{M})$ initially converges exponentially with decay rate e^{-8t} before leveling off (around $t = 2.5$) due to stochastic fluctuations. Panel **(b)** shows convergence to the principal subspace via decay of $V_*(\mathbf{W})$ for the ODE and online algorithm. After the trajectories reach a neighborhood of $\mathcal{O}$ around $t = 2$, the decay rate of $V_*(\mathbf{W})$ for both the ODE and online algorithm continues more gradually as the dynamics of $\mathbf{W}$ are dominated by the gradient $\nabla V(\mathbf{W})$. Overall, these simulations suggest that the online algorithm also evolves according to the same two phases as the ODE.

7 Global stability for general synaptic learning rates

Our analysis focused on the special case $\tau = \frac{1}{2}$, where feedforward and recurrent synapses evolve at equal rates. The choice simplified the structure of the invariant manifold $\mathcal{O}$ and Lyapunov function L, enabling a transparent proof of global convergence. However, the assumption $\tau = \frac{1}{2}$ may appear restrictive and raises the question of whether the same two-phase convergence holds for other $0 < \tau \leq \frac{1}{2}$.

Centorrino et al. [19] recently established global stability in the limit $\tau \to 0$, where the recurrent synapses evolve on a separate timescale than the feedforward synapses. In this regime, the synaptic dynamics also exhibit a two-phase structure, but with separate timescales. In the first phase, the feedforward weights $\mathbf{W}$ are fixed while the recurrent weights $\mathbf{M}$ evolve towards the invariant manifold $\mathcal{O}_0 = \{(\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \mathbf{M}^{-1}\mathbf{W}\mathbf{A}\mathbf{W}^\top\mathbf{M}^{-1} = \mathbf{M}\}$. As shown in appendix E.1, this convergence can be described by the convex Lyapunov function $L_0(\mathbf{W}, \mathbf{M}) = \|(\mathbf{W}\mathbf{A}\mathbf{W}^\top)^2 - \mathbf{M}^3\|^2$. In the second phase, as shown in [19, section 5.3], the synaptic matrices $(\mathbf{W}, \mathbf{M})$ evolve within the invariant manifold $\mathcal{O}_0$ and the feedforward synapses $\mathbf{W}$ follow the gradient flow of the non-convex potential function $V_0(\mathbf{W}) = \frac{3}{2}\|(\mathbf{W}\mathbf{A}\mathbf{W}^\top)^{\frac{1}{3}}\|^2 - \|\mathbf{W}\|^2$. Notably, both the invariant manifold $\mathcal{O}_0$ and functions $L_0(\mathbf{W}, \mathbf{M})$ and $V_0(\mathbf{W})$ are distinct from those analyzed here for the case that $\tau = \frac{1}{2}$.

We conjecture that global convergence holds for all $0 < \tau \leq \frac{1}{2}$, with dynamics exhibiting a similar two-phase structure: (i) rapid convergence to a generalized invariant manifold $\mathcal{O}_\tau$ with convergence described by a generalized convex Lyapunov function $L_\tau(\mathbf{W}, \mathbf{M})$, and (ii) slower convergence along this manifold to the principal subspace that follows the gradient flow of a generalized non-convex potential function $V_\tau(\mathbf{W})$. Moreover, we conjecture that for each $0 \leq \tau \leq \frac{1}{2}$, there is a function $\Phi_\tau : \mathbb{R}^{k\times n} \to \mathcal{S}^k_+$ mapping feedforward weights to recurrent weights and $p_\tau > 0$ such that

- the invariant manifold satisfies $\mathcal{O}_\tau = \{(\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \Phi_\tau(\mathbf{W}) = \mathbf{M}\}$,
- the convex Lyapunov function satisfies $L_\tau(\mathbf{W}, \mathbf{M}) = \|\Phi_\tau(\mathbf{W})^{p_\tau} - \mathbf{M}^{p_\tau}\|^2$,
- and the potential function is $V_\tau(\mathbf{W}) = f(\mathbf{W}, \Phi_\tau(\mathbf{W}))$, where $f(\mathbf{W}, \mathbf{M})$ is defined in (5).

Furthermore, Φ_τ and p_τ satisfy $\Phi_{\frac{1}{2}}(\mathbf{W}) = (\mathbf{W}\mathbf{W}^\top)^{\frac{1}{2}}$ and $p_{\frac{1}{2}} = 2$, and $\lim_{\tau\to 0} \Phi_\tau(\mathbf{W}) = (\mathbf{W}\mathbf{A}\mathbf{W}^\top)^{\frac{1}{3}}$ and $\lim_{\tau\to 0} p_\tau = 3$. We anticipate that the conjecture can be proved for τ in a small neighborhood of

$\frac{1}{2}$ using an application of the following implicit function theorem for dynamical systems (Persistence of Normally Hyperbolic Invariant Manifolds under Perturbations [22]). The challenge is extending the proof beyond a neighborhood of $\tau = \frac{1}{2}$. In appendix E.2, we provide a numerical method for estimating the invariant manifolds $\mathcal{O}_\tau$ and illustrate the manifolds in the scalar setting $n = k = 1$.

8 Discussion

We proved global convergence of solutions to the ODE (3)–(4) when $\tau = \frac{1}{2}$. Our analysis revealed a two-phase structure to the convergence: (1) rapid convergence to an invariant set where the neural filters are orthonormal, and (2) slower evolution along this manifold following the gradient of a non-convex potential function whose minima correspond to the principal subspace. This result provides a rigorous link between local Hebbian/anti-Hebbian interactions and stable network-level computations. Practically, the results suggest that initializing synaptic weights on the invariant manifold (e.g., setting $\mathbf{M}_0 = \mathbf{I}_k$ and $\mathbf{W}_0$ to have orthonormal row vectors) could accelerate convergence.

This work is an important step towards proving convergence rate guarantees for the online algorithm analogous to the results established by Chou and Wang [3] for Oja's PCA model of a neuron. The analysis may also provide insight into the global dynamics of related algorithms that can be implemented in neural networks with local Hebbian synaptic learning rules for solving non-negative matrix factorization problem [23] and networks with local non-Hebbian synaptic learning rules for solving symmetric generalized eigen-subspace problems such as canonical correlation analysis [24, 25].

While the assumption $\tau = \frac{1}{2}$ simplifies the analysis and enables closed-form Lyapunov functions, it is also restrictive and may not be biologically realistic. Future work should extend the proof of global stability to the range $0 < \tau \leq \frac{1}{2}$, as supported by numerical evidence (section 7 and appendix E) and analytical results for the regime $\tau \to 0$ [19].

A Existence and uniqueness of solutions

In this section, we prove existence and uniqueness of solutions to the ODE (3)–(4) for any $\tau > 0$. A solution of the ODE is a continuously differentiable function $t \mapsto (\mathbf{W}(t), \mathbf{M}(t))$ from $[0, \infty)$ to $\mathcal{D}$ whose derivative satisfies equations (3)–(4). Recall that $\|\cdot\|$ denotes the Frobenius norm.

Theorem 3. *Suppose $\mathbf{A} \in \mathcal{S}^n_{++}$. For any $\tau > 0$ and $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}$, there exists a unique solution $(\mathbf{W}(t), \mathbf{M}(t))$ to the ODE (3)–(4) with initial condition $(\mathbf{W}_0, \mathbf{M}_0)$.*

Proof. For each $K < \infty$, define the set

$$\mathcal{D}_K := \left\{(\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \|\mathbf{M}^{-1}\| + \|\mathbf{W}\| < K\right\}.$$

Since $\mathbf{G}$ is analytic on $\mathcal{D}$ and, for $K < \infty$, the closure of $\mathcal{D}_K$ is compact, it follows that $\mathbf{G}$ is uniformly Lipschitz continuous on $\mathcal{D}_K$. Let $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}$. Then for each $K < \infty$ sufficiently large such that $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}_K$, there exists a unique solution $(\mathbf{W}(t), \mathbf{M}(t))$ of the ODE (3)–(4) on the interval $[0, T_K)$, where

$$T_K := \inf\left\{t \geq 0 : \|\mathbf{M}(t)^{-1}\| + \|\mathbf{W}(t)\| \geq K\right\}$$

is the first time $(\mathbf{W}(t), \mathbf{M}(t))$ exits the set $\mathcal{D}_K$. We are left to show that $T_\infty := \lim_{K\to\infty} T_K = \infty$.

Let $\mathbf{v} \in \mathbb{R}^k$ be an arbitrary unit vector and define $a(t) := \mathbf{v}^\top \mathbf{M}(t)\mathbf{v}$ for all $t \in [0, T_\infty)$. Then

$$\tau \frac{\mathrm{d}a(t)}{\mathrm{d}t} = \mathbf{v}^\top \mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A}\mathbf{W}(t)^\top \mathbf{M}(t)^{-1}\mathbf{v} - a(t) \geq -a(t),$$

and so $a(t) \geq a(0)e^{-t/\tau}$ for all $t \in [0, T_\infty)$. Since this holds for all unit vectors $\mathbf{v} \in \mathbb{R}^k$, we have $\sigma_{\min}(\mathbf{M}(t)) \geq me^{-t/\tau} > 0$ for all $t \in [0, T_\infty)$, where $m := \sigma_{\min}(\mathbf{M}_0)$ and $\sigma_{\min}(\mathbf{M}) > 0$ denotes the smallest eigenvalue of $\mathbf{M}$. Therefore,

$$\|\mathbf{M}(t)^{-1}\| = \sqrt{\mathrm{Tr}(\mathbf{M}(t)^{-2})} \leq \frac{\sqrt{k}}{\sigma_{\min}(\mathbf{M}(t))} \leq \frac{\sqrt{k}}{m} e^{t/\tau}, \qquad t \in [0, T_\infty). \tag{12}$$

Next, we have, for all $t \in [0, T_\infty)$,

$$\begin{aligned}\left\| \frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t} \right\| &\le 2\|\mathbf{M}(t)^{-1}\|\|\mathbf{W}(t)\|\|\mathbf{A}\| + 2\|\mathbf{W}(t)\| \\ &\le 2\left(\frac{\sqrt{k}}{m}\|\mathbf{A}\|e^{t/\tau} + 1\right)\|\mathbf{W}(t)\|.\end{aligned}$$

By Gronwall's inequality,

$$\|\mathbf{W}(t)\| \le \|\mathbf{W}_0\| \exp\left[2\left(\frac{\sqrt{k}}{m}\|\mathbf{A}\|e^{t/\tau} + 1\right)t\right], \qquad t \in [0, T_\infty). \tag{13}$$

Let $T < \infty$ be arbitrary. Setting

$$K := \max\left\{\frac{2\sqrt{k}}{m}e^{T/\tau}, 2\|\mathbf{W}_0\| \exp\left[2\left(\frac{\sqrt{k}}{m}\|\mathbf{A}\|e^{T/\tau} + 1\right)T\right]\right\} < \infty,$$

it follows from equations (12)–(13) that $T_K \ge T$. Therefore, $T_\infty = \lim_{K\to\infty} T_K = \infty$. □

B Characterizing the critical points

In this section, we prove Lemma 1, which characterizes the critical points of the ODE (3)–(4).

Proof of Lemma 1. First suppose $(\mathbf{W}_*, \mathbf{M}_*) \in \mathcal{E}$. Setting the derivative in (4) to zero, we see that $(\mathbf{W}_*, \mathbf{M}_*)$ satisfy

$$\mathbf{M}_*^{-1}\mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top\mathbf{M}_*^{-1} = \mathbf{M}_*.$$

After left- and right-multiplying on both sides of the equality by $\mathbf{M}_*$, we obtain the relation $\mathbf{M}_*^3 = \mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top$. Taking cube roots yields $\mathbf{M}_* = (\mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top)^{\frac{1}{3}}$. Next, setting the derivative in (3) to zero, we obtain

$$\mathbf{M}_*^{-1}\mathbf{W}_*\mathbf{A} = \mathbf{W}_*.$$

After right-multiplying by $\mathbf{W}_*^\top$ on both sides of the equality and substituting in with $\mathbf{M}_*^{-1} =$

$(\mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top)^{-\frac{1}{3}}$, we obtain $(\mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top)^{\frac{2}{3}} = \mathbf{W}_*\mathbf{W}_*^\top$. Substituting in with the SVD $\mathbf{W}_* = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ yields $\mathbf{V}^\top\mathbf{A}\mathbf{V} = \mathbf{S}$. Since $\mathbf{S}$ is diagonal, it follows that the column vectors of $\mathbf{V}$ are eigenvectors of $\mathbf{A}$ and the diagonal elements of $\mathbf{S}$ are the corresponding eigenvalues. Finally, we see that $\mathbf{M}_* = \mathbf{U}\mathbf{S}\mathbf{U}^\top$. On the other hand, if $\mathbf{W}_*$ and $\mathbf{M}_*$ are as in the statement of the lemma, then it is readily verified by substitution that $\mathbf{G}(\mathbf{W}_*, \mathbf{M}_*) = \mathbf{0}$. □

C Bounding the synaptic weights

In this section, we prove Lemma 3, which states that the set $\mathcal{N}$ of initial conditions with divergent trajectories, has Lebesgue measure zero. Through the remainder of the Appendices we fix $\tau = \frac{1}{2}$. We first need the following lemma.

Lemma 6. *Suppose* $\mathbf{A} \in \mathcal{S}^n_{++}$ *and* $(\mathbf{W}(t), \mathbf{M}(t))$ *is a solution of the ODE* (3)–(4) *with initial condition* $(\mathbf{W}_0(t), \mathbf{M}_0(t)) \in \mathcal{D} \setminus \mathcal{N}$. *Then* $\liminf_{t\to\infty} \det(\mathbf{M}(t)) > 0$ *and* $\limsup_{t\to\infty} \|\mathbf{M}(t)\| < \infty$.

Proof. For a proof by contradiction, suppose $\liminf_{t\to\infty} \det(\mathbf{M}(t)) = 0$. Then, along with Lemma 2, this implies there is a sequence $\{t_j\}$ with $t_j \to \infty$ as $j \to \infty$ such that

$$\lim_{j\to\infty} \det(\mathbf{M}(t_j)^2) = \lim_{j\to\infty} \det(\mathbf{W}(t_j)\mathbf{W}(t_j)^\top) = 0$$

and for each $j = 1, 2, \ldots,$

$$-\infty < \frac{\mathrm{d}}{\mathrm{d}t} \log\det(\mathbf{W}(t_j)\mathbf{W}(t_j)^\top)) < 0.$$

Since $(\mathbf{W}, \mathbf{M}) \notin \mathcal{N}$, we can choose j sufficiently large such that

$$\frac{k}{\sigma_{\min}(\mathbf{A})} < \mathrm{Tr}\left[\mathbf{M}(t_j)^{-1}\right] < \infty. \tag{14}$$

However, by the ODE (3),

$$\frac{\mathrm{d}}{\mathrm{d}t} \log(\det(\mathbf{W}(t_j)\mathbf{W}(t_j)^\top)) = \mathrm{Tr}\left[\left(\mathbf{W}(t_j)\mathbf{W}(t_j)^\top\right)^{-1} \frac{\mathrm{d}\mathbf{W}(t_j)\mathbf{W}(t_j)^\top}{\mathrm{d}t}\right]$$

$$
\begin{aligned}
&= 2\,\mathrm{Tr}\left[\left(\mathbf{W}(t_j)\mathbf{W}(t_j)^\top\right)^{-1}\mathbf{M}(t_j)^{-1}\mathbf{W}(t_j)\mathbf{A}\mathbf{W}(t_j)^\top - \mathbf{I}_k\right] \\
&\geq 2\,\mathrm{Tr}\left[\sigma_{\min}(\mathbf{A})\left(\mathbf{W}(t_j)\mathbf{W}(t_j)^\top\right)^{-1}\mathbf{M}(t_j)^{-1}\mathbf{W}(t_j)\mathbf{W}(t_j)^\top - \mathbf{I}_k\right] \\
&> 0,
\end{aligned}
$$

which contradicts equation (14). Therefore, $\liminf_{t\to\infty}\det(\mathbf{M}(t)) > 0$.

We are left to show that $\limsup_{t\to\infty}\|\mathbf{M}(t)\| < \infty$. Again, for a proof by contradiction, suppose $\limsup_{t\to\infty}\|\mathbf{M}(t)\| = \infty$. By the previous result, $K = \sup_{t\geq 0}\|\mathbf{M}(t)^{-1}\| < \infty$. Thus, there exists $t > 0$ sufficiently large such that

$$
\frac{\mathrm{d}\|\mathbf{M}(t)\|^2}{\mathrm{d}t} > 0 \tag{15}
$$

and

$$
-\infty < \mathrm{Tr}\left[\sigma_{\max}(\mathbf{A})\mathbf{M}(t) - \mathbf{M}(t)^2\right] + 4KL(\mathbf{W}(0),\mathbf{M}(0))e^{-4t}\sigma_{\max}(\mathbf{A}) < 0.
$$

By the ODE (4) and Lemma 2,

$$
\begin{aligned}
\frac{\mathrm{d}\|\mathbf{M}(t)\|^2}{\mathrm{d}t} &= 4\,\mathrm{Tr}\left[\mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A}\mathbf{W}(t)^\top - \mathbf{M}(t)^2\right] \\
&\leq 4\,\mathrm{Tr}\left[\sigma_{\max}(\mathbf{A})\mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{W}(t)^\top - \mathbf{M}(t)^2\right] \\
&\leq 4\,\mathrm{Tr}\left[\sigma_{\max}(\mathbf{A})\mathbf{M}(t) - \mathbf{M}(t)^2\right] + 4KL(\mathbf{W}(0),\mathbf{M}(0))e^{-4t}\sigma_{\max}(\mathbf{A}) \\
&< 0.
\end{aligned}
$$

This contradict equation (15). Therefore, $\limsup_{t\to\infty}\|\mathbf{M}(t)\| < \infty$. □

Proof of Lemma 3. The fact that $\mathcal{N}$ has Lebesgue measure zero follows immediately because $\mathcal{N} \subset \{(\mathbf{W},\mathbf{M}) \in \mathcal{D} : \det(\mathbf{W}\mathbf{W}^\top) = 0\}$, $\det(\cdot)$ is a nonzero polynomial, and the vanishing set of a nonzero polynomial has Lebesgue measure zero [26].

Let $(\mathbf{W}(t),\mathbf{M}(t))$ be a solution of the ODE (3)–(4) with initial condition $(\mathbf{W}_0,\mathbf{M}_0)$. In the case $(\mathbf{W}_0,\mathbf{M}_0) \in \mathcal{D}\setminus\mathcal{N}$, Lemmas 2 and 6 imply that $\|\mathbf{M}(t)^{-1}\|$ and $\|\mathbf{W}(t)\|$ are uniformly bounded in

t. On the other hand, suppose $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{N}$, and $(\mathbf{W}(t), \mathbf{M}(t))$ is the solution to (3)–(4) with initial condition $(\mathbf{W}_0, \mathbf{M}_0)$. Then there exists $t > 0$, $\mathbf{v} \in \mathbb{R}^k$ and $\lambda \in \mathbb{R}$ such that $\mathbf{W}(t)^\top \mathbf{v} = 0$ and $\mathbf{M}(t)\mathbf{v} = \lambda \mathbf{v}$. Then

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{v}^\top \mathbf{W}(t)}{\mathrm{d}t} &= 2\left(\mathbf{v}^\top \mathbf{W}(t)\mathbf{M}(t)^{-1}\mathbf{A} - \mathbf{v}^\top \mathbf{W}(t)\right) = 0 \\
\frac{\mathrm{d}\mathbf{M}(t)\mathbf{v}}{\mathrm{d}t} &= 2\left(\mathbf{M}(t)^{-1}\mathbf{W}(t)\mathbf{A}\mathbf{W}(t)^\top \mathbf{M}(t)^{-1}\mathbf{v} - \mathbf{M}(t)\mathbf{v}(t)\right) = -2\lambda \mathbf{v}.
\end{aligned}$$

It follows that for $t > 0$, $\mathbf{W}(t)^\top \mathbf{v} = 0$ and $\mathbf{M}(t)\mathbf{v} = \lambda e^{-2t}\mathbf{v}$, $(\mathbf{W}(t), \mathbf{M}(t)) \in \mathcal{N}$. Therefore, for each $t > 0$, $\mathbf{v}$ is an eigenvector of $\mathbf{M}(t)$ with eigenvalue λe^{-2t}. It follows that $\lim_{t\to\infty} \det(\mathbf{M}(t)) = 0$. □

D Convergence of neural filters to an eigen-subspace

In this section, we prove Lemma 5.

Proof of Lemma 5. Suppose $(\mathbf{W}(t), \mathbf{M}(t))$ is a solution to ODE (3)–(4). By the definition of $V(\mathbf{W})$ in equation (11) and the ODE (3), we have (suppressing the dependence on t)

$$\begin{aligned}
\operatorname{Tr}\left[\nabla_{\mathbf{W}}(V(\mathbf{W}))\frac{\mathrm{d}\mathbf{W}^\top}{\mathrm{d}t}\right] &= 2\operatorname{Tr}\left[\left(\mathbf{W} - (\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\mathbf{W}\mathbf{A}\right)\left(\mathbf{A}\mathbf{W}^\top\mathbf{M}^{-1} - \mathbf{W}\right)\right] \\
&= 2\operatorname{Tr}\left[\left(\mathbf{W} - (\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\mathbf{W}\mathbf{A}\right)\left(\mathbf{A}\mathbf{W}^\top(\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}} - \mathbf{W}\right)\right] \\
&\quad + \operatorname{Tr}\left[\left(\mathbf{W} - (\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\mathbf{W}\mathbf{A}\right)\mathbf{A}\mathbf{W}^\top\left(\mathbf{M}^{-1} - (\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\right)\right] \\
&= -2\|\nabla_{\mathbf{W}}(V(\mathbf{W}^\top))\|^2 \\
&\quad + \operatorname{Tr}\left[\left(\mathbf{W} - (\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\mathbf{W}\mathbf{A}\right)\mathbf{A}\mathbf{W}^\top\left(\mathbf{M}^{-1} - (\mathbf{W}\mathbf{W}^\top)^{-\frac{1}{2}}\right)\right].
\end{aligned}$$

By Lemma 3, $\mathbf{W}(t) - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\mathbf{W}(t)\mathbf{A}$ is uniformly bounded in t. By Lemmas 2 and 3,

$$\begin{aligned}
&\limsup_{t\to\infty} \left\|\mathbf{M}(t)^{-1} - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\right\|^2 \\
&\qquad \le \limsup_{t\to\infty} \left\|\mathbf{M}(t)^{-1}\right\|^2 \left\|(\mathbf{W}(t)\mathbf{W}(t)^\top)^{\frac{1}{2}}\right\|^2 \left\|\mathbf{M}(t) - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{\frac{1}{2}}\right\|^2 = 0.
\end{aligned}$$

Therefore, along with LaSalle's invariance principle and Lemma 4, this implies that $(\mathbf{W}(t), \mathbf{M}(t))$

converges to the set $\mathcal{E}$ as $t \to \infty$. This completes the proof. $\square$

E Stability of the ODE for general timescales

In section 7 we conjectured that the global convergence result proved in the main text for $\tau = \frac{1}{2}$ (Theorem 1) extends to all $0 < \tau \leq \frac{1}{2}$, and that the dynamics retain the same two-phase structure: (i) rapid convergence towards a τ-dependent invariant manifold $\mathcal{O}_\tau$, captured by a convex Lyapunov function L_τ, and (ii) slow evolution along $\mathcal{O}_\tau$ governed by the gradient flow of a τ-dependent nonconvex potential V_τ, whose minima correspond to the principal subspace. Here we provide some limited empirical evidence in support of our conjecture.

E.1 Convex Lyapunov function for the case $\tau \to 0$

In the $\tau \to 0$ regime, the recurrent weights first converge to the invariant manifold $\mathcal{O}_0 = \{(\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \mathbf{M}^{-1}\mathbf{W}\mathbf{A}\mathbf{W}^\top\mathbf{M}^{-1} = \mathbf{M}\}$ while the feedforward weights remain fixed, after which both weights evolve within the invariant manifold. Here we sketch out an argument showing exponential convergence in the first phase, where we assume that the feedforward weights $\mathbf{W}$ remain fixed; that is, $\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t} = 0$. Recall the convex Lyapunov function $L_0(\mathbf{W}, \mathbf{M}) = \|\mathbf{W}\mathbf{A}\mathbf{W}^\top - \mathbf{M}^3\|^2$ from section 7. Suppose $\mathbf{W}$ is fixed and $\mathbf{M}(t)$ is a solution of the ODE (4). By the product rule and the ODE (4), and suppressing the dependence on t, we have

$$\tau \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}^3 = \mathbf{M}\mathbf{W}\mathbf{A}\mathbf{W}^\top\mathbf{M}^{-1} + \mathbf{W}\mathbf{A}\mathbf{W}^\top + \mathbf{M}^{-1}\mathbf{W}\mathbf{A}\mathbf{W}^\top\mathbf{M} - 3\mathbf{M}^3.$$

Next, by the chain rule, the previous display and the cyclic property of the trace operator, we have

$$\begin{aligned}
\tau \frac{\mathrm{d}}{\mathrm{d}t} L_0(\mathbf{W}, \mathbf{M}) &= -2 \operatorname{Tr}\left[\left(\mathbf{W}\mathbf{A}\mathbf{W}^\top - \mathbf{M}^3\right)\left(2\mathbf{M}\mathbf{W}\mathbf{A}\mathbf{W}^\top\mathbf{M}^{-1} + \mathbf{W}\mathbf{A}\mathbf{W}^\top - 3\mathbf{M}^3\right)\right] \\
&= -4\left\|\mathbf{M}^{\frac{1}{2}}\left(\mathbf{W}\mathbf{A}\mathbf{W}^\top - \mathbf{M}^3\right)\mathbf{M}^{-\frac{1}{2}}\right\|^2 - 2L_0(\mathbf{W}, \mathbf{M}) \\
&\leq -2L_0(\mathbf{W}, \mathbf{M}).
\end{aligned}$$

Therefore, in the first phase, the Lyapunov function $L_0(\mathbf{W}, \mathbf{M})$ converges exponentially to zero.

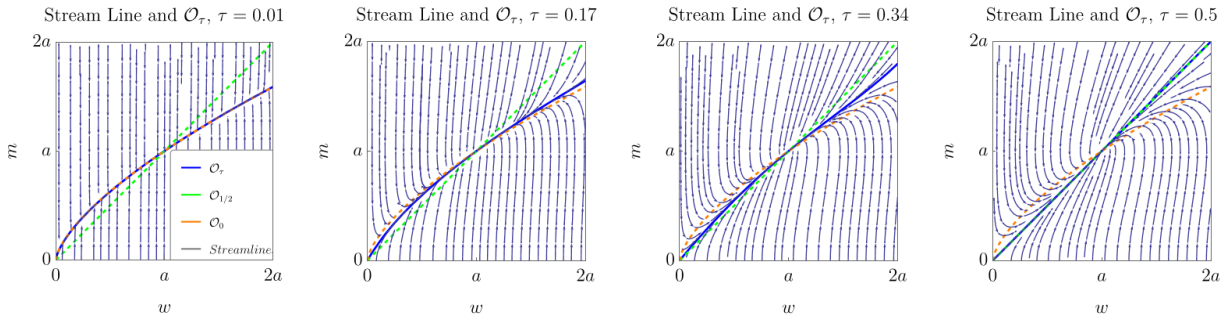

Figure 6: For each τ, the arrows show the streamlines for the ODE (3)–(4). The solid blue line is a numerical estimate of the invariant set $\mathcal{O}_\tau$. The orange and green dashed lines are the invariant sets $\mathcal{O}_0$ and $\mathcal{O}_{\frac{1}{2}}$, respectively.

E.2 Estimation of invariant manifolds in the scalar setting

Consider the scalar setting $n = k = 1$, in which case the critical points are $\pm(a, a)$. Figure 6 shows the streamlines of the ODE (3)–(4) for $w > 0$, computed using Mathematica's built-in function `Streamline` function, converging for different $\tau \in (0, \frac{1}{2}]$. The invariant manifolds $\mathcal{O}_0$ and $\mathcal{O}_{\frac{1}{2}}$ are shown in orange and green, respectively. For intermediate τ, we numerically estimate $\mathcal{O}_\tau$ using the method described in the next paragraph. Note that $\mathcal{O}_\tau$ appears to interpolate between $\mathcal{O}_0$ and $\mathcal{O}_{\frac{1}{2}}$ as τ varies from 0 to $\frac{1}{2}$, and this is supported by further simulations (not shown).

Numerical estimation of $\mathcal{O}_\tau$ for $w > 0$. To estimate $\mathcal{O}_\tau$ when $a < w < 2a$, we define a function $h : (0, \infty)^2 \to \mathbb{R}^2$ as follows. Given a point (w_0, m_0), let $(w(t), m(t))$ denote the solution of the ODE with initial condition (w_0, m_0) and set $h(w_0, m_0) = \lim_{t\to\infty}(\dot{w}(t), \dot{m}(t))$. A point (w_0, m_0) is in $\mathcal{O}_\tau$ if $h(w_0, m_0)$ is parallel to the eigenvector of the Jacobian of the vector field $\mathbf{G}$ with smallest associated eigenvalue. We then perform a numerical search to find a point $(w_0, m_0) \in \mathcal{O}_\tau$ on the line $w = 2a$, and then compute the solution of the ODE starting at (w_0, m_0) to estimate $\mathcal{O}_\tau$ for $a < w < 2a$. We implemented this procedure with the help of Mathematica's built-in functions `NDSolve` and `FindRoot`. To estimate $\mathcal{O}_\tau$ for $0 < w < a$, we could use a similar argument starting with points (w_0, m_0) near the origin. However, there is an alternative approach that does not require a numerical search. Rather, we assume that the origin is in the closure of the invariant

manifold $\mathcal{O}_\tau$ and the invariant manifold can be linearly approximated when m is small; that is the invariant manifold near the origin is well-approximated by $\{(v_0 m, m) : m \approx 0\}$ for some $v_0 > 0$. Under this assumption, when (w_0, m_0) are small, the derivatives of the solutions to the ODE (3)–(4) at $t = 0$ are approximately $\dot{w}(0) \approx \frac{2aw_0}{m_0}$ and $\dot{m}(0) \approx \frac{aw_0^2}{\tau m_0^2}$. Using the fact that $w_0 \approx v_0 m_0$ and $\dot{w}(0) \approx v_0 \dot{m}(0)$, we see that $2av_0 \approx \frac{a}{\tau} v_0^3$ and so $v_0 = \sqrt{2\tau}$. Therefore, we estimate $\mathcal{O}_\tau$ by computing a solution starting at $(\sqrt{2\tau}\alpha, \alpha)$ for $\alpha > 0$ small. This numerical procedure can also be generalized to estimate the invariant manifolds when $d, k > 1$; however, visualizing the manifolds is more difficult.

References

[1] Cengiz Pehlevan, Tao Hu, and Dmitri B Chklovskii. A Hebbian/anti-Hebbian neural network for linear subspace learning: A derivation from multidimensional scaling of streaming data. *Neural Computation*, 27(7):1461–1495, 2015.

[2] Erkki Oja. Simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15(3):267–273, 1982.

[3] Chi-Ning Chou and Mien Brabeeba Wang. ODE-inspired analysis for the biological version of Oja's rule in solving streaming PCA. In *Conference on Learning Theory*, pages 1339–1343. PMLR, 2020.

[4] Peter Földiak. Adaptive network for optimal linear feature extraction. In *Proceedings of IEEE/INNS Int. Joint. Conf. Neural Networks*, volume 1, pages 401–405, 1989.

[5] Jeanne Rubner and Paul Tavan. A self-organizing network for principal-component analysis. *EPL (Europhysics Letters)*, 10(7):693, 1989.

[6] Jeanette Rubner and Klaus Schulten. Development of feature detectors by self-organization. *Biological Cybernetics*, 62(3):193–199, 1990.

[7] Todd K Leen. Learning in linear feature-discovery networks. In *Adaptive Signal Processing*, volume 1565, pages 472–481. International Society for Optics and Photonics, 1991.

[8] Cengiz Pehlevan, Anirvan M Sengupta, and Dmitri B Chklovskii. Why do similarity matching objectives lead to Hebbian/anti-Hebbian networks? *Neural Computation*, 30(1):84–124, 2018.

[9] Erkki Oja and Juha Karhunen. An analysis of convergence for a learning version of the subspace method. *Journal of Mathematical Analysis and Applications*, 91(1):102–111, 1983.

[10] Terence D Sanger. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2(6):459–473, 1989.

[11] Todd Leen, Mike Rudnick, and Dan Hammerstrom. Hebbian feature discovery improves classifier efficiency. In *1990 IJCNN International Joint Conference on Neural Networks*, pages 51–56. IEEE, 1990.

[12] Erkki Oja and Juha Karhunen. On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix. *Journal of Mathematical Analysis and Applications*, 106(1):69–84, 1985.

[13] Kurt Hornik and Chung-Ming Kuan. Convergence analysis of local feature extraction algorithms. *Neural Networks*, 5(2):229–240, 1992.

[14] Erkki Oja. Principal components, minor components, and linear neural networks. *Neural Networks*, 5(6):927–935, 1992.

[15] Mark D Plumbley. Lyapunov functions for convergence of principal component algorithms. *Neural Networks*, 8(1):11–23, 1995.

[16] Vivek S Borkar. Stochastic approximation with two time scales. *Systems & Control Letters*, 29(5):291–294, 1997.

[17] Harold Joseph Kushner and Dean S Clark. *Stochastic approximation methods for constrained and unconstrained systems*, volume 26. Springer Science & Business Media, 2012.

[18] Tianyi Lin, Chi Jin, and Michael Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In *International Conference on Machine Learning*, pages 6083–6093. PMLR, 2020.

[19] Veronica Centorrino, Francesco Bullo, and Giovanni Russo. Similarity matching networks: Hebbian learning and convergence over multiple time scales. *arXiv preprint arXiv:2506.06134*, 2025.

[20] Joseph P LaSalle. Stability theory for ordinary differential equations. *Journal of Differential Equations*, 4(1):57–65, 1968.

[21] Rafael Potrie and Pablo Monzón. Local implications of almost global stability. *Dynamical Systems*, 24(1):109–115, 2009.

[22] Morris W. Hirsch, Charles C. Pugh, and Michael Shub. *Invariant Manifolds*, volume 583 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, Heidelberg, 1977. ISBN 978-3-540-08054-5. doi: 10.1007/BFb0087003. Reprinted 2006.

[23] Cengiz Pehlevan and Dmitri B Chklovskii. Neuroscience-inspired online unsupervised learning algorithms: Artificial neural networks. *IEEE Signal Processing Magazine*, 36(6):88–96, 2019.

[24] David Lipshutz, Yanis Bahroun, Siavash Golkar, Anirvan M Sengupta, and Dmitri B Chklovskii. A biologically plausible neural network for multichannel canonical correlation analysis. *Neural Computation*, 33(9):2309–2352, 2021.

[25] David Lipshutz, Yanis Bahroun, Siavash Golkar, Anirvan M Sengupta, and Dmitri B Chklovskii. A normative framework for deriving neural networks with multicompartmental neurons and non-Hebbian plasticity. *PRX Life*, 1:013008, 2023.

[26] Richard Caron and Tim Traynor. The zero set of a polynomial. *Technical Report*, pages 1–2, 2005.