arXiv:2601.13170v1 [q-bio.NC] 19 Jan 2026

主成分的赫布式/反赫布式网络的全局稳定性

子空间学习

David Lipshutz*和 Robert J. Lipshutz

2026年1月21日

摘要

生物神经网络会根据局部突触修饰进行自组织，从而产生稳定的计算。突触层面的修饰如何在网络层面产生此类计算，仍是一个悬而未决的问题。Pehlevan 等人 [1] 提出了一种自组织神经网络模型，该模型采用赫布式和反赫布式突触更新来实现主子空间分析算法；然而，非线性突触动力学的全局稳定性尚未得到证实。在此，我们针对前馈权重和循环权重以相同时间尺度演化的情形，证明突触动力学连续极限的全局稳定性，并表明该动力学分两个阶段演化。在第一阶段，突触权重收敛到一个不变流形，在该流形上，“神经滤波器”是正交归一的。在第二阶段，突触动力学遵循一个非凸势函数的梯度流，其极小值对应于张成输入数据主子空间的神经滤波器。

1 引言

生物神经网络根据局部突触相互作用进行自组织，从而在网络层面产生稳定的计算。理论神经科学面临的一项挑战，是将这些局部相互作用与稳定的网络级计算联系起来。

*系 贝勒医学院神经科学系及莱斯大学神经工程计划 y.

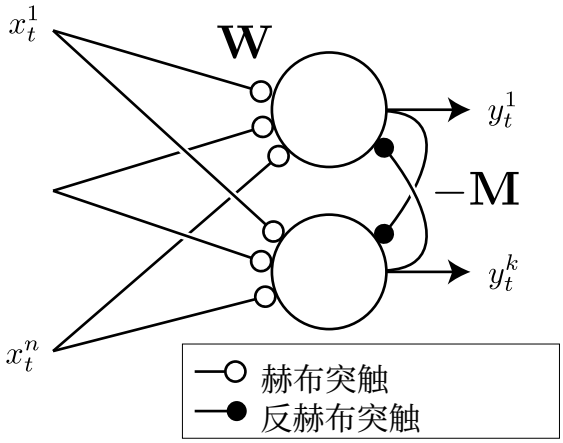

图 1：用于主子空间分析的赫布式/反赫布式网络。包含 k 个神经元的单层网络接收 n 个输入。前馈赫布突触$\mathbf{W}$ 将 n 个输入连接到 k 个神经元，循环反赫布突触 $-\mathbf{M}$则连接这些 k 个神经元。

在一项开创性研究中，Oja [2]提出了一种神经元计算模型：通过局部赫布式突触更新，实现在线算法，从输入数据中学习其最高主成分。因此，Oja算法建立了突触层面的局部赫布式相互作用与神经元层面的主成分分析（PCA）之间的联系。此外，这种在线（随机）算法表现出惊人的稳定性和数据效率——它具有全局稳定性，且其收敛速率达到信息论下界 [3]这表明，采用局部突触更新的神经网络也可能适用于机器学习和神经形态计算。

在多通道情形下，对于采用局部突触更新的神经网络，其稳定性所知甚少。继 Oja的工作 [2], 之后，研究者对主子空间分析（PSA）算法进行了若干扩展，使其能够在采用局部赫布和反赫布突触更新 [1, 4–8]的神经网络中实现；然而，这些多通道网络包含循环交互，使分析变得复杂，而且尚未确立网络动力学的全局稳定性。另一方向的研究提出并分析了采用非局部突触更新的主子空间分析网络 [9–15], ，但这些网络无法解释局部交互如何产生稳定的网络级计算。

本研究重点分析多通道网络的全局突触动力学，用于平台主管彩平台 马会.

主子空间分析的多通道网络，该网络采用 Pehlevan 等人提出的赫布式前馈突触和反赫布式循环突触。 [1, 8] ，如图1所示。具体而言，我们证明，与突触动力学相关的常微分方程（ODE）是全局稳定的，即对于几乎所有的初始化，突触权重都会收敛到与主子空间分析（PSA）相关的最优配置。我们还证明，动力学演化分为两个阶段。在第一阶段，突触权重收敛到一个不变流形，该流形以神经滤波器的正交归一性为特征。在第二阶段，突触动力学遵循非凸势函数的梯度流。

2 赫布式/反赫布式网络模型

本节首先回顾 Pehlevan 等人提出的用于 [1]主子空间分析的赫布／反赫布神经网络模型 [1, 8]。关于如何从所谓的“相似性匹配”目标函数推导该网络的详细过程，参见 [1, 8]。随后，我们介绍与突触动力学相关的常微分方程，这将是本文分析的重点。

考虑一个包含 k 个主神经元的网络，该网络接收 $n > k$ 个输入，如图1所示。前馈突触 $\mathbf{W}$ 将输入连接到主神经元，循环突触 $-\mathbf{M}$则连接各主神经元。在每个时间步 $t = 1,2,\ldots$，网络接收一个输入向量$\mathbf{x}_t = (x_t^1, \ldots, x_t^n)$。网络以两个时间尺度运行：快速的神经动力学时间尺度和缓慢的突触更新时间尺度。

在第一阶段， k 个输出神经元的神经活动（记为k维向量 $\mathbf{y}_t = (y_t^1, \ldots, y_t^k)$）按照快速线性神经动力学演化

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

它们收敛至 $\mathbf{y}_t = \mathbf{F}_t\mathbf{x}_t$，其中 $\mathbf{W}_t$ 和 $\mathbf{M}_t$ 分别表示时间 t时前馈突触权重和循环突触权重的状态，而 $k \times n$ 矩阵 $\mathbf{F}_t := \mathbf{M}_t^{-1}\mathbf{W}_t$的行向量称为“神经滤波器”。神经活动收敛后，突触权重将

根据缓慢的突触可塑性规则更新

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

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

其中， $\eta > 0$ 是前馈突触$\mathbf{W}$ 的学习率， τ 表示前馈突触$\mathbf{W}$ 与侧向突触 $-\mathbf{M}$的学习率之比。前馈突触$\mathbf{W}$ 的可塑性规则包含一项与突触前、突触后活动乘积成正比的项，即$\mathbf{y}_t\mathbf{x}_t^\top$，因此称为“赫布式”。侧向突触 $-\mathbf{M}$的可塑性规则包含一项与突触前、突触后活动乘积成反比的项，即$\mathbf{y}_t\mathbf{y}_t^\top$，因此称为“反赫布式”。

为了分析其算法的稳定性，Pehlevan 等人 [8] 考虑了更新的连续极限。形式上，当输入数据 $\{\mathbf{x}_t\}$ 是均值为零且具有固定 $n\times n$ 协方差矩阵$\mathbf{A}$ 的独立同分布样本，并且步长 $\eta > 0$ 无限小时，突触动力学可由以下常微分方程近似表示

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

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

在适当的正则性条件下（例如，$\mathbf{M}$ 的谱一致地远离零），可以精确描述在线算法与常微分方程之间的关系，前提是学习率随时间变化 [16, 17]。

Pehlevan 等人 [8] 证明，常微分方程的每个平衡点都对应于$\mathbf{A}$的一个特征子空间；并且，当 $\tau \leq \frac{1}{2}$ 时，所有线性稳定的平衡点都对应于 $\mathbf{A}$ 的主（特征）子空间。虽然他们的理论分析有助于了解平衡点附近的突触动力学，但对于平衡点之外的突触动力学（这对应于大多数随机初始化）却没有提供相关信息。Pehlevan 等人 [1, 8] 提供了实证证据，表明在线算法和常微分方程都是全局稳定的，并将二者与可比算法进行了基准测试；然而，对全局动力学的理论分析仍然

是一个尚未解决的问题。本研究旨在证明常微分方程的全局稳定性。

常微分方程（3）–（4）可自然地视为用于求解非凸—凹极小极大问题的梯度下降—上升流（具有时间尺度分离 τ）

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

其中，最小化是在 $k \times n$ 矩阵 $\mathbb{R}^{k\times n}$ 的集合上进行的，而最大化是在 $k \times k$ 正定矩阵 $\mathcal{S}^k_{++}$的集合上进行的。类似地，离散算法（1）–（2）可自然地解释为用于求解该极小极大问题的随机梯度下降—上升算法。一般而言，证明非凸—凹极小极大问题的梯度下降—上升算法的收敛性颇具挑战，现有结果 [16, 18, 19] 依赖于时间尺度的分离（即令 $\tau \to 0$）。[1] 然而，数值实验表明，最优收敛速率出现在 $\tau = \frac{1}{2}$ 附近，参见 [8, 图 4], ，这表明，证明全局收敛性并不需要进行时间尺度分离，并且 $\tau = \frac{1}{2}$ 可能是一个尤其值得关注的参数。

3 突触动力学的全局稳定性

对于 $\tau = \frac{1}{2}$ 的情况，我们证明常微分方程 (3)–(4) 全局收敛至所需的主子空间。我们假设 $\mathbf{A}$ 是一个正定 $n \times n$ 矩阵，其特征值为 $\lambda_1 \geq \cdots \geq \lambda_n > 0$ 且$k < n$ 固定。以下定理是我们的主要结果。

定理 1. 设 $A \in \mathcal{S}^n_{++}$。对于每个 $\tau > 0$ 以及 $(W_0,M_0) \in \mathcal{D} := \mathbb{R}^{k\times n} \times \mathcal{S}^k_{++}$，常微分方程 *(3)–(4)* 存在唯一解 $(W(t), M(t))$，其初始条件为 (W_0, M_0)，对所有 $t \geq 0$ 均成立。此外，假设 $\lambda_k > \lambda_{k+1}$ 且 $\tau = \frac{1}{2}$。那么存在一个集合 $\mathcal{Z}$，其勒贝格测度为零；若 $(W_0,M_0) \in \mathcal{D} \setminus \mathcal{Z}$，则解 $(W(t), M(t))$ 在 $t \to \infty$ 时收敛至常微分方程的平衡点集合 (W_*, M_*)，其中神经滤波器（即 $F_* := M_*^{-1}W_*$ 的行向量）正交归一，并张成 A 的主子空间。

接下来的两节和附录A 专用于证明定理1。在附录

[1] 在准备本手稿期间，我们了解到了一篇预印本 [19]，该预印本同样分析了常微分方程（3）–（4）的稳定性。然而，该分析考察的是 $\tau \to 0$ 这一情形，而我们处理的是 $\tau = \frac{1}{2}$ 这一情形。

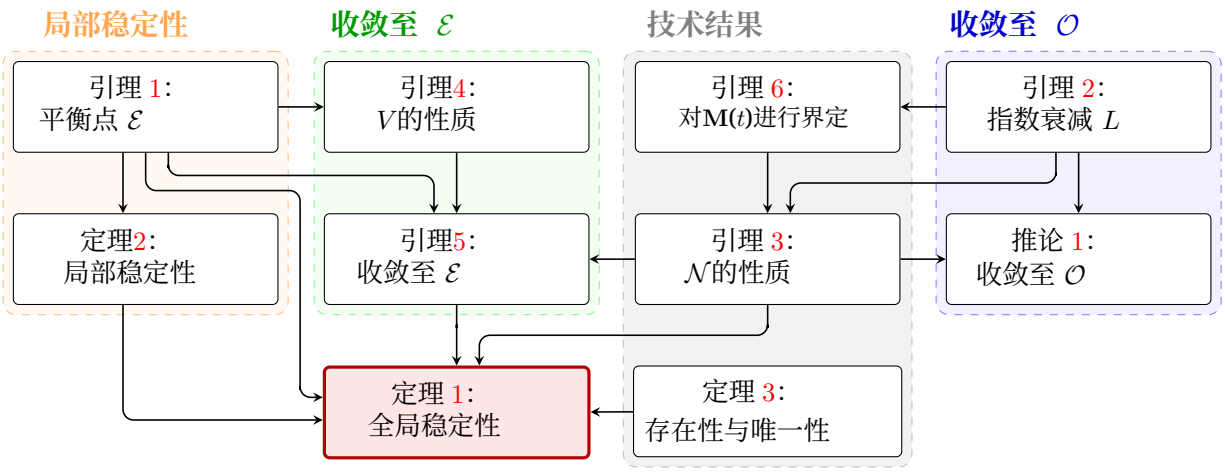

图 2：我们的结果依赖关系图。第 4 节证明了平衡点的局部稳定性（橙色区域）。第 $\mathcal{O}$ 5.1 节证明了解收敛至不变流形 (蓝色区域) $\mathcal{E}$ 。第 5.2 节证明了从不变流形上或其附近出发的解收敛至平衡点 （绿色区域）。最后，第 5.3 节证明了常微分方程的全局稳定性（红色方框）。技术结果（灰色区域）在附录 A 和 C 中证明。

A 中，我们证明解的存在唯一性，这保证了从 $\mathcal{D}$ 初始化的解在所有 $t \geq 0$ 内都保持在 $\mathcal{D}$ 中。在第 4 节中，我们回顾 [8] 关于平衡点局部线性稳定性的结果，平衡点记为 $\mathcal{E}$。全局收敛将在第 5 节中证明。我们表明，突触权重的演化分为两个阶段。在第一阶段，对于任何不属于零测集 $\mathcal{N}$ 的初始化，突触权重首先收敛到不变流形 $\mathcal{O}$ ，该流形对应于正交归一神经滤波器；也就是说，满足 $\mathrm{F} := \mathrm{M}^{-1}\mathrm{W}$ 具有正交归一行向量的 (W, M) 集合 $\in \mathcal{D}$ （第 5.1 节和图 3）。这一收敛过程由凸李雅普诺夫函数 $L(\mathrm{W}, \mathrm{M})$ 描述。随后，在第二阶段，从不变流形 $\mathcal{O}$ 上（或其附近）出发，突触动力学可近似为非凸势函数 $V(\mathrm{W})$ 的梯度流（第 5.2 节和图 4）。因此，对于几乎所有初始化，突触权重都会收敛到一个平衡点，使得神经滤波器对应于所需的主子空间投影。主要结果的依赖关系图如图 2 所示。

在第6节中，我们提供了实证证据，表明对于离散在线算法（1）–（2），突触权重也分两个阶段演化。在第7节中，我们推测定理1 可以

推广后适用于所有 $0 < \tau \leq \frac{1}{2}$。 –

4 平衡点的局部稳定性

接下来，我们刻画常微分方程（3）–（4）的平衡点，并回顾 Pehlevan 等人关于其线性稳定性的结果。为此，令 [8] 在其线性稳定性方面。为此，令

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

表示平衡点集合，其中 $\mathbf{G}\colon\ \mathcal{D} \mapsto \mathcal{D}$ 是由下式定义的向量场

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

4.1 平衡点的刻画

下面的引理（证明见附录B）根据奇异值分解（奇异值分解）刻画了平衡点。

引理 1。假设 $\tau > 0$ 且$\mathbf{A} \in \mathcal{S}^n_{++}$。当且仅当$(\mathbf{W}_*,\mathbf{M}_*) \in \mathcal{E}$ 满足$\mathbf{W}_* = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ 且$\mathbf{M}_* = \mathbf{U}\mathbf{S}\mathbf{U}^\top$时，其中$\mathbf{U}$ 是任意的 $k \times k$ 正交矩阵，$\mathbf{V}$ 是一个 $d \times k$ 矩阵，其列向量是协方差矩阵$\mathbf{A}$的正交归一特征向量，而$\mathbf{S}$ 是一个 $k \times k$ 对角矩阵，其对角元素是与$\mathbf{V}$的列向量相对应的$\mathbf{A}$ 的特征值。

根据引理 1，若 $(\mathrm{W}_*, \mathrm{M}_*) \in \mathcal{E}$，则 W_* 满秩（因为我们假设 A 的特征值均为正），且 $\mathrm{F}_* := \mathrm{M}_*^{-1}\mathrm{W}_* = \mathrm{U}\mathrm{V}^\top$，其中 U 是一个 $k \times k$ 正交矩阵，V 的列向量是 A 的正交归一特征向量。换言之，F_* 的行向量（即神经滤波器）是正交归一的，并张成 A 的一个特征子空间。此外，若 $(\mathrm{W}_*, \mathrm{M}_*) \in \mathcal{E}$，则对于每个 $k \times k$ 正交矩阵 Q，$(\mathrm{Q}\mathrm{W}_*, \mathrm{Q}\mathrm{M}_*\mathrm{Q}^\top) \in \mathcal{E}$。特别地，每个平衡点都是一个 ^{k}（$_{k-1}$）二维平衡点流形中的元素，该流形对应于 A 的一个特征子空间。

4.2 局部线性稳定性分析

对于我们的目的，若平衡点 $(\mathbf{W}_*,\mathbf{M}_*) \in \mathcal{E}$ 处的 $\mathbf{G}$ 的雅可比矩阵的所有特征值实部均非正，则称该平衡点为“线性稳定”；若其不线性稳定，则称其为“线性不稳定”；也就是说，在 $(\mathbf{W}_*,\mathbf{M}_*)$ 处计算的 $\mathbf{G}$ 的雅可比矩阵至少有一个特征值实部为正。令

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

下一个结果可由以下内容直接推出： [8, 定理 1]。

定理 2。 假设 $0 < \tau \leq \frac{1}{2}$ 且 $\mathbf{A} \in \mathcal{S}^n_{++}$。当且仅当 $(\mathbf{W}_*,\mathbf{M}_*) \in \mathcal{E}$ 时，平衡点 $(\mathbf{W}_*,\mathbf{M}_*) \in \mathcal{E}_0$是线性稳定的。

5 全局收敛分析

现在，我们证明关于常微分方程 (3)–(4) 在 $\tau = \frac{1}{2}$ 下的全局稳定性的主要结果，本节中 3 展示了标量情形 $\mathbf{G}(W,M)$ 的向量场 $k = n = 1$。在第 5.1 节中，我们证明神经滤波器是渐近正交归一的——在图 3 中，这对应于轨迹最初收敛到蓝线上。接着，在第 5.2 节中，我们证明，从正交归一初始化（或其附近）出发时，描述前馈权重 3 动力学的常微分方程 $\mathbf{W}$ 可以近似为某个势函数的梯度流，其极小值对应于主子空间——在图 3 中，这对应于初始条件位于蓝线上（或其附近）的轨迹收敛至红点。最后，在第 5.3 节中，我们结合这些结果证明定理 1。

5.1 神经滤波器的渐近正交归一性

在本节中，我们证明，对于几乎所有的初始化，常微分方程（$\mathbf{W}(t)$,$\mathbf{M}(t)$）（3）–（4）的解收敛到以下矩阵不变子集，该子集位于 $\mathcal{D}$ 中，对应于正交归一

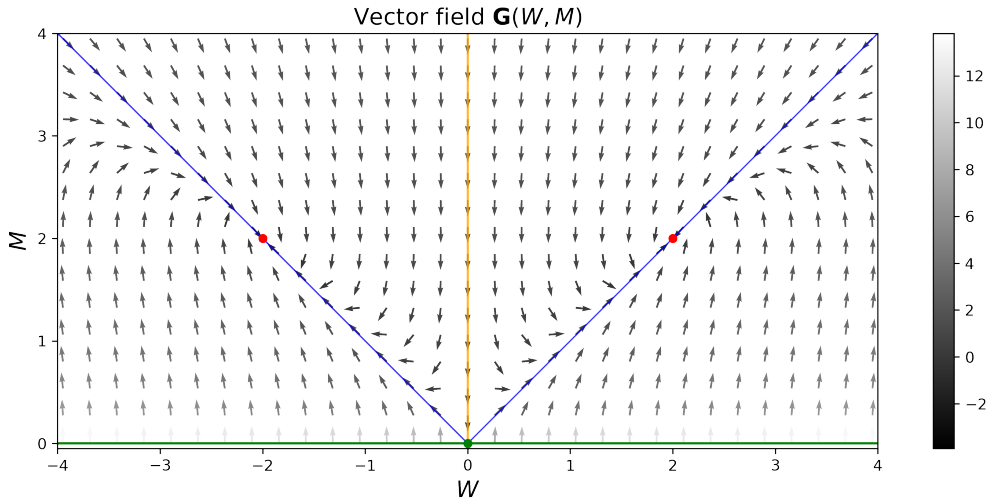

图 3：在 $k = n = 1$ 和 $\lambda_1 = 2$的情况下，向量场 $\mathbf{G}(W, M)$ 的图像。灰度表示向量模的对数。蓝色线条表示集合 $\mathcal{O}$，橙色竖线表示集合 $\mathcal{N}$，两个红点表示集合 $\mathcal{E}_0$（在此情况下，它等于 $\mathcal{E}$），绿色线条表示直线 $M = 0$ 不属于 $\mathcal{D} = \mathbb{R} \times (0, \infty)$。

神经滤波器 $\mathbf{F} = \mathbf{M}^{-1}\mathbf{W}$：

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

为证明这一收敛性，我们在 $\mathcal{D}$上定义如下凸李雅普诺夫函数：

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

注意 $L(\mathbf{W}, \mathbf{M})$处处非负，当且仅当$(\mathbf{W}, \mathbf{M}) \in \mathcal{O}$.

引理 2。 设 $\mathbf{A} \in \mathcal{S}^n_{++}$ 和 $(\mathbf{W}(t)、\mathbf{M}(t))$ 是常微分方程 (3)–(4) 的一个解，满足 $\tau = \frac{1}{2}$，且初始条件为 $(\mathbf{W}_0,\mathbf{M}_0) \in \mathcal{D}$。则

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

证明。 首先注意到，常微分方程（3）–（4）意味着

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

因此，根据链式法则、乘积法则、迹的循环性质以及上面的推导，

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

S通过变量分离法求解该微分方程，可得方程（8）。 □

为了证明神经滤波器渐近正交归一，我们需要一个技术引理，说明解（**W**(*t*)，**M**(*t*))不收敛到零的条件。为此，定义集合

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

那么， $\mathcal{N}$ 是满足以下条件的（W, M）对的集合：$\mathbf{W}\mathbf{W}^\top$ 是奇异的，且 M 存在一个属于 $\mathbf{W}\mathbf{W}^\top$零空间的特征向量。注意，如果 W 满秩且 M 正定，那么（W, M） $\notin \mathcal{N}$。集合 $\mathcal{N}$对应图 3 中的橙色竖线。下面的技术引理证明，对于任意不属于 $\mathcal{N}$ 的初始化，常微分方程（3）–（4）的解始终远离零和无穷大。证明见附录 C。

引理 *3*。集合 $\mathcal{N}$ 的勒贝格测度为零。设 $A \in \mathcal{S}^n_{++}$，且 *(W(t),M(t))* 是满足 $\tau = \frac{1}{2}$、初始于 *(*W_0, M_0*)* $\in \mathcal{D}$ 的常微分方程 *(3)–(4)* 的解。如果 *(*W_0, M_0*)* $\notin \mathcal{N}$，那么

$(\mathbf{W}(t),\mathbf{M}(t)) \notin \mathcal{N}$ 对所有 $t \geq 0$ 以及

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

另一方面，如果$(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{N}$，那么$(\mathbf{W}(t)，\mathbf{M}(t)) \in \mathcal{N}$对所有 $t \geq 0$均成立，且

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

引理 2 和 3 的以下推论指出，常微分方程的几乎所有解都以指数速度收敛到不变流形 $\mathcal{O}$。

推论 *1.* 假设 $A \in \mathcal{S}^n_{++}$，且 *(W(t), M(t))* 是常微分方程 *(3)–(4)* 的一个解，其中 $\tau = \frac{1}{2}$，初始条件为 $(W_0, M_0) \in \mathcal{D} \setminus \mathcal{N}$。那么，当 $t \to \infty$ 时，*(W(t), M(t))* 收敛到 $\mathcal{O}$。

证明。 由引理 3， $K :=$ 上确界$\{\|\mathbf{M}(t)^{-1}\| : \ t \geq 0\} < \infty$。因此，

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

其中第一个不等式由柯西-施瓦茨不等式得出，第二个不等式则由引理 2得出。

□

5.2 收敛到平衡点

既然已经证明 (W(*t*), M(*t*)) 在 $t \to \infty$ 时收敛到 $\mathcal{O}$，现在我们来分析 (W(*t*), M(*t*)) 接近集合 $\mathcal{O}$ 时的动力学。首先，考虑 (W$_0$,M$_0$) $\in \mathcal{O}$ 的情况，因此对于所有 $t \geq 0$，(W(*t*), M(*t*)) $\in \mathcal{O}$。于是，我们可以将常微分方程 (3) 的右端重写为仅关于 W(*t*) 的函数：

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

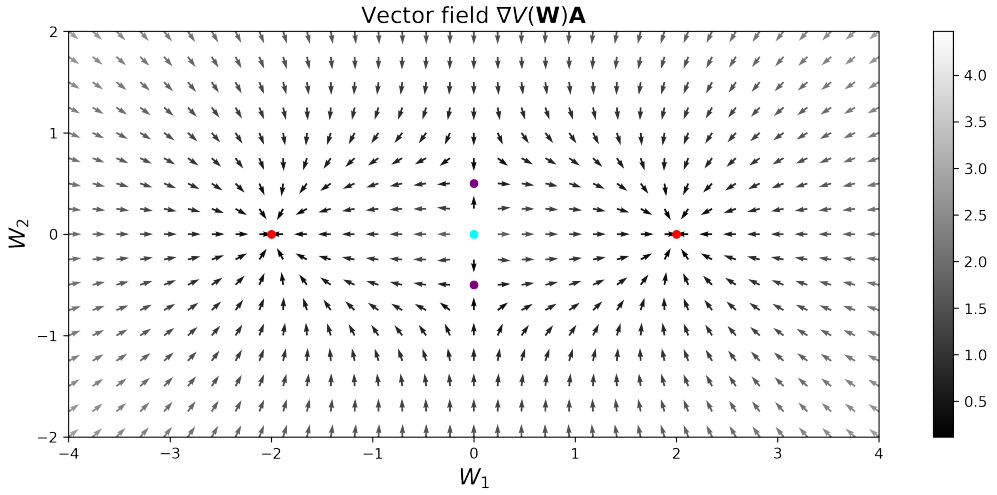

图 4：在 $n = 2$, $k = 1$ 且 $\mathbf{A} = \text{diag}(2, 1\,2)$ 的情况下，向量场 $-\nabla V(\mathbf{W})$ 的图。灰度表示向量模。红点表示 V 的全局极小值，紫点表示 V 的鞍点，原点处的青色点表示集合$\{\mathbf{W} : \det(\mathbf{W}\mathbf{W}^\top) = 0\}$。

其中 $V : \mathbb{R}^{k\times n} \mapsto \mathbb{R}$ 是非凸势函数

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

其中 $f(\mathbf{W}, \mathbf{M})$ 定义于（5）中。因此，我们可以将 $\mathbf{W}(t$ 在 $\mathcal{O}$ 上的动力学解释为势函数 V 的梯度流。（注意，V 仅在满秩矩阵 $\mathbf{W}$ 构成的子集 $\mathbb{R}^{k\times n}$上可微。）在图 4中，我们绘制了 $-\nabla V(\mathbf{W})$ 在 $d = 2$和 $k = 1$ 情况下的向量场，以展示 $\mathbf{W}(t)$ 在集合 $\mathcal{O}$上的动力学。

引理 4. 设 $\mathbf{A} \in \mathcal{S}^n_{++}$。函数 V有下界。此外，$\nabla V(\mathbf{W})$ 存在，且满足 $\nabla V(\mathbf{W}) = 0$ 当且仅当 $\mathbf{W}$ 为满秩矩阵，且 $(\mathbf{W}, (\mathbf{W}\mathbf{W})^{\frac{1}{2}}) \in \mathcal{E}$。

证明。令 $\mathbf{W} \in \mathbb{R}^{k\times n}$，并令 $\mathbf{W} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ 表示其奇异值分解。于是

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

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

其中 $\sigma_{\max}$(A) 表示 A 的谱范数。由于这对所有 W 都成立，函数 V 有下界。接下来，假设 W 满秩且 ∇V(W) $= 0$。则 2(WW$^\top$)$^{-\frac{1}{2}}$WAW$^\top$ =(WW$^\top$)$^{\frac{1}{2}}$，因此 V$^\top$AV $=$ S。由于 S 是对角矩阵，可知 V 的列向量是 A 的特征向量，而 S 的对角元素是 A 对应的特征值。因此，根据引理 1，(W, (WW$^\top$)$^{\frac{1}{2}}$) $\in \mathcal{E}$。通过代入即可轻易验证其逆命题。

□

根据 LaSalle 不变性原理 [20], ，常微分方程（10）和引理4 可知，当以（$\mathbf{W}_0, \mathbf{M}_0$） $\in \mathcal{O}$为初始条件时，解（$\mathbf{W}$（t），$\mathbf{M}$（t））收敛到 V的一个不动点，该不动点对应于 $\mathcal{E}$中的一个平衡点。接下来，对于一般情形（$\mathbf{W}_0, \mathbf{M}_0$） $\in \mathcal{D} \setminus \mathcal{N}$，我们可以将常微分方程（3）的右端改写如下:

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

其中回顾一下，$\mathbf{W}$（t）$\mathbf{W}$（t）$^\top$ 对所有 $t \geq 0$ 均非奇异，这是由引理3得出的。根据链式法则，

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

我们断言

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

假设该断言成立，则有

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

因此，根据 LaSalle 不变性原理，$\mathbf{W}(t)$ 收敛到 V的不动点集合，该集合对应于平衡点集合 $\mathcal{E}$。我们在以下引理中总结这一结果，并在附录D中给出详细证明。

引理 *5.* 假设 $A \in \mathcal{S}^n_{++}$ 且 *(W(t), M(t))* 是常微分方程 *(3)–(4)* 的一个解，其中 $\tau = \frac{1}{2}$ 且初始条件 $(W_0, M_0) \in \mathcal{D} \setminus \mathcal{N}$。则 *(W(t), M(t))* 收敛到集合 $\mathcal{E}$，当 $t \to \infty$。

5.3 定理1的证明

解的存在唯一性已在附录A的定理3中证明。现在我们结合引理 5 和定理 2 来证明全局收敛。定义子集

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

其元素对应的初始条件所产生的轨迹收敛到线性不稳定平衡点集合。如果 $(\mathrm{W}_*, \mathrm{M}_*) \in \mathcal{E} \setminus \mathcal{E}_0$，那么根据定理 2，G 在 $(\mathrm{W}_*, \mathrm{M}_*)$ 处的雅可比矩阵有一个实部为正的特征值。因此，根据 [21, 命题 3]，，集合 $\mathcal{U}$ 的勒贝格测度为零。结合引理 3，这意味着集合 $\mathcal{Z} := \mathcal{N} \cup \mathcal{U}$ 的勒贝格测度也为零。假设 $(\mathrm{W}(t), \mathrm{M}(t))$ 是常微分方程 (3)–(4) 的一个解，其初始条件为 $(\mathrm{W}_0, \mathrm{M}_0) \in \mathcal{D} \setminus \mathcal{Z}$。那么，根据引理 5 和 $\mathcal{U}$的定义，$(\mathrm{W}(t), \mathrm{M}(t))$ 随着 $t \to \infty$ 收敛到集合 $\mathcal{E}_0$。最后，根据引理 1 和 $\mathcal{E}_0$的定义，对于每个 $(\mathrm{W}_*, \mathrm{M}_*) \in \mathcal{E}_0$，$\mathrm{F}_* := \mathrm{M}_*^{-1}\mathrm{W}_*$ 的行向量正交归一，并张成 A 的主子空间。

6 比较常微分方程与在线算法

现在，我们通过数值模拟来考察在常微分方程（3）–（4）中观察到的两阶段收敛是否也会出现在在线算法（1）–（2）中。

设置。我们考虑一个包含 $d = 4$ 个输入和 $k = 2$ 个神经元的网络。输入 $\{\mathrm{x}_t\}$ 独立同分布地取自均值为零、协方差矩阵为 A $=$ diag(0.5, 0.25, 0.2, 0.05) 的正态分布。前馈权重矩阵 W 初始化为包含独立同分布标准正态条目的矩阵，循环权重矩阵 M 初始化为对角矩阵，其对角条目从 $[1, 2]$ 中均匀采样。使用 Mathematica 内置函数 NDSolve 求解常微分方程 (3)–(4)。在线算法 (1)–(2) 采用时变步长 $\eta_t = \frac{c_0}{c_1+t}$ 进行模拟，其中 c_0、$c_1 > 0$ 的选取满足

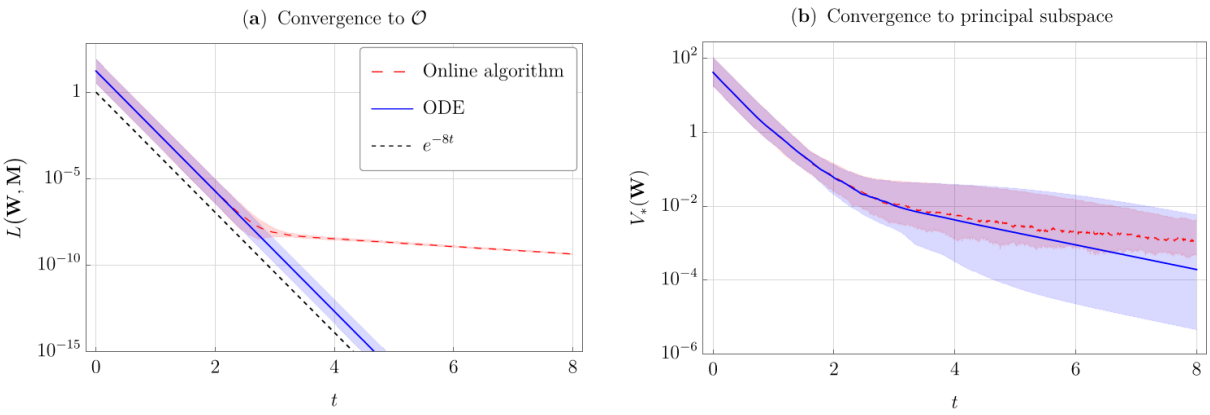

图5：在具有 $d = 4$ 个输入和$k = 2$个神经元的网络中，常微分方程与在线算法的收敛性。（**a**）使用李雅普诺夫函数$L(\mathbf{W},\mathbf{M})$测量向不变流形 $\mathcal{O}$的收敛性。（**b**）使用 $V_*(\mathbf{W}) = V(\mathbf{W}) - V(\mathbf{W}_*)$测量向主子空间的收敛性，其中$(\mathbf{W}_*,\mathbf{M}_*)$是常微分方程的任意稳定平衡点。在面板（**a**）和（**b**）中，阴影区域表示100次随机初始化结果的中间80%分位区间；实线/点线表示中位数值。仿真细节见正文。

$\eta_1 = 0.001$ 和 $\sum_{t=1}^{25,000} \eta_t = 8$。为跟踪收敛情况，我们计算李雅普诺夫函数 $L(\mathbf{W}, \mathbf{M})$ 和 $V_*(\mathbf{W}) = V(\mathbf{W}) - V(\mathbf{W}_*)$，其中 $(\mathbf{W}_*、\mathbf{M}_*)$ 是常微分方程的任意稳定临界点。常微分方程和在线算法均采用 100 次随机初始化进行模拟。

结果。图5 展示了常微分方程和在线算法的两阶段收敛。面板 (a) 显示，对于常微分方程（蓝色实线），$L(\mathrm{W}, \mathrm{M})$ 以衰减率 e^{-8t}（黑色虚线）呈指数收敛，这与引理 2 一致。对于在线算法（红色虚线），$L(\mathrm{W}, \mathrm{M})$ 最初以衰减率 e^{-8t} 呈指数收敛，随后逐渐趋于平稳（约在$t = 2.5$），这是随机波动所致。面板 (b) 显示，常微分方程和在线算法均通过 $V_*(\mathrm{W})$ 的衰减收敛到主子空间。当轨迹到达$\mathcal{O}$ 附近的 $t = 2$后，由于 W 的动力学主要受梯度 $\nabla V(\mathrm{W})$ 支配，常微分方程和在线算法中 $V_*(\mathrm{W})$ 的衰减率都会继续下降，但速度较为缓慢。总体而言，这些模拟结果表明，在线算法也遵循与常微分方程相同的两个阶段演化。

7 一般突触学习率下的全局稳定性

我们的分析聚焦于特殊情形 $\tau = \frac{1}{2}$，其中前馈突触和循环突触以相同速率演化。这一选择简化了不变流形 $\mathcal{O}$ 和李雅普诺夫函数 L的结构，从而能够清晰地证明全局收敛。然而，假设 $\tau = \frac{1}{2}$ 可能显得过于严格，也引出了一个问题：对于其他 $0 < \tau \leq \frac{1}{2}$，是否同样存在两阶段收敛?

Centorrino 等人 [19] 最近在极限 $\tau \to 0$ 下确立了全局稳定性；在该极限下，循环突触与前馈突触以不同的时间尺度演化。在此情形下，突触动力学也呈现两阶段结构，但两个阶段具有不同的时间尺度。在第一阶段，前馈权重 W 保持固定，而循环权重 M 演化并趋向不变流形 $\mathcal{O}_0 = \{(\mathrm{W}, \mathrm{M}) \in \mathcal{D} : \mathrm{M}^{-1}\mathrm{WAW}^\top\mathrm{M}^{-1} = \mathrm{M}\}$。如附录 E.1 所示，这种收敛可由凸李雅普诺夫函数 $L_0(\mathrm{W}, \mathrm{M}) = \|(\mathrm{WAW}^\top)^2 - \mathrm{M}^3\|^2$ 描述。在第二阶段，如 [19, 第 5 节所示，3], 突触矩阵 (W, M) 在不变流形 $\mathcal{O}_0$ 内演化，前馈突触 W 则遵循非凸势函数 $V_0(\mathrm{W}) = \frac{3}{2}\|(\mathrm{WAW}^\top)^{\frac{1}{3}}\|^2 - \|\mathrm{W}\|^2$ 的梯度流。值得注意的是，不变流形 $\mathcal{O}_0$以及函数 $L_0(\mathrm{W}, \mathrm{M})$ 和 $V_0(\mathrm{W})$ 均不同于本文对 $\tau = \frac{1}{2}$ 情形所分析的对象。

我们猜想，对于所有 $0 < \tau \leq \frac{1}{2}$，均存在全局收敛，且动力学呈现类似的两阶段结构：(i) 快速收敛至广义不变流形 $\mathcal{O}_\tau$，这一收敛过程由广义凸李雅普诺夫函数 $L_\tau(\mathbf{W}, \mathbf{M})$ 描述；(ii) 随后沿该流形较慢地收敛至主子空间，其过程遵循广义非凸势函数 $V_\tau(\mathbf{W})$ 的梯度流。此外，我们猜想，对于每个 $0 \leq \tau \leq \frac{1}{2}$，都存在一个函数 $\Phi_\tau : \mathbb{R}^{k\times n} \to \mathcal{S}_+^k$，它将前馈权重映射到循环权重，并且 $p_\tau > 0$ 满足

- 不变流形满足 $\mathcal{O}_\tau = \{(\mathbf{W}, \mathbf{M}) \in \mathcal{D} : \Phi_\tau(\mathbf{W}) = \mathbf{M}\}$,
- 凸李雅普诺夫函数满足 $L_\tau(\mathbf{W}, \mathbf{M}) = \|\Phi_\tau(\mathbf{W})^{p_\tau} - \mathbf{M}^{p_\tau}\|^2$,
- 且势函数为 $V_\tau(\mathbf{W}) = f(\mathbf{W}\ \Phi_\tau(\mathbf{W}))$，其中 $f(\mathbf{W}, \mathbf{M})$ 定义见（5）。

此外， Φ_τ和 p_τ满足$\Phi_{\frac{1}{2}}(\mathbf{W}) = (\mathbf{WW}^\top)^{\frac{1}{2}}$以及$p_{\frac{1}{2}} = 2$，且 $\lim_{\tau\to 0}\Phi_\tau(\mathbf{W}) = (\mathbf{WAW}^\top)^{\frac{1}{3}}$ 以及 $\lim_{\tau\to 0} p_\tau = 3$。我们预期可以在 τ的一个小邻域内证明该猜想

– 1 2 通过应用以下《动力系统》隐函数定理（扰动下正规双曲不变流形的持久性 [22]）来实现。挑战在于将证明扩展到 $\tau = \frac{1}{2}$ 的一个邻域之外。在附录E.2中，我们提供了一种用于估计不变流形的数值方法 $\mathcal{O}_\tau$，并在标量情形下展示这些流形 $n = k = 1$。

–

8 讨论

我们证明了当 $\tau = \frac{1}{2}$ 时，常微分方程（3）–（4）的解全局收敛。我们的分析揭示了收敛过程具有两阶段结构：（1）迅速收敛到一个不变集，其中神经滤波器正交归一；（2）沿着该流形较缓慢地演化，并遵循一个非凸势函数的梯度，其极小值对应于主子空间。这一结果为局部赫布/反赫布相互作用与稳定的网络级计算建立了严格联系。从实践角度看，结果表明，在不变流形上初始化突触权重（例如，设置 $\mathbf{M}_0 = \mathbf{I}_k$，并使 $\mathbf{W}_0$ 的行向量正交归一）可能会加快收敛。

这项工作是朝着证明在线算法的收敛速率保证迈出的重要一步，其对应结果已由 Chou 和 Wang [3] 针对单个神经元的 Oja 主成分分析模型确立。这项分析也可能有助于理解相关算法的全局动力学：这些算法可通过采用局部赫布突触学习规则的神经网络来求解非负矩阵分解问题 [23]，也可通过采用局部非赫布突触学习规则的网络来求解对称广义特征子空间问题，例如典型相关分析。

虽然假设 $\tau = \frac{1}{2}$ 简化了分析，并使得能够构造闭式李雅普诺夫函数，但这一假设也具有限制性，可能不符合生物学实际。未来的工作应将全局稳定性证明扩展到范围 $0 < \tau \leq \frac{1}{2}$，数值证据（第 7 节和附录E）以及针对 $\tau \to 0$ [19] 区间的解析结果均支持这一方向。

–

A 解的存在性与唯一性

本节中，我们证明对于任意 $\tau > 0$，常微分方程（3）–（4）的解的存在唯一性。常微分方程的一个解是一个连续可微函数 $t \mapsto (\mathbf{W}(t), \mathbf{M}(t))$，它从 $[0, \infty)$ 到 $\mathcal{D}$，其导数满足方程（3）–（4）。回顾一下，$\|\cdot\|$ 表示弗罗贝尼乌斯范数。

定理 3。设 $A \in \mathcal{S}^n_{++}$。对于任意 $\tau > 0$ 和 $(W_0, M_0) \in \mathcal{D}$，常微分方程 (3)–(4) 存在唯一解 $(W(t), M(t))$，其初始条件为 (W_0, M_0)。

证明。对于每个 $K < \infty$，定义集合

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

由于 G 在 $\mathcal{D}$ 上解析，并且对于 $K < \infty$，$\mathcal{D}_K$ 的闭包是紧致的，因此 G 在 $\mathcal{D}_K$ 上利普希茨连续且一致。令 $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}$。于是，对于每个足够大的 $K < \infty$，使得 $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}_K$，常微分方程 (3)–(4) 在区间 $[0, T_K)$ 上存在唯一解 $(\mathbf{W}(t), \mathbf{M}(t))$，其中

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

is 首次离开集合 $\mathcal{D}_K$ 的时刻。接下来只需证明 $T_\infty := \lim_{K\to\infty} T_K = \infty$。令 $\mathbf{v} \in \mathbb{R}^k$ 为任意单位向量，并定义对所有 $t \in [0, T_\infty)$，$a(t) := \mathbf{v}^\top \mathbf{M}(t)\mathbf{v}$。于是

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

因此，对于所有 $t \in [0, T_\infty)$，$a(t) \geq a(0)e^{-t/\tau}$。由于这对所有单位向量 $\mathbf{v} \in \mathbb{R}^k$ 都成立，因此对于所有 $t \in [0, T_\infty)$，我们有$\sigma_{\min}(\mathbf{M}(t)) \geq me^{-t/\tau} > 0$，其中 $m := \sigma_{\min}(\mathbf{M}_0)$，而 $\sigma_{\min}(\mathbf{M}) > 0$ 表示 $\mathbf{M}$ 的最小特征值。因此，

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

接下来，对于所有 $t \in [0, T_\infty)$，我们有：

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

根据格朗沃尔不等式，

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

令 $T < \infty$ 为任意值。设

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

由方程（12）–（13）可得 $T_K \geq T$。因此， $T_\infty = \text{极限}_{K\to\infty} T_K = \infty$。 □

B 刻画临界点

本节证明引理 1，该引理刻画了常微分方程（3）–（4）的临界点。

引理 *1* 的证明。 首先，假设（$\mathbf{W}_*$，$\mathbf{M}_*$） $\in \mathcal{E}$。令方程（4）中的导数为零，我们看到（$\mathbf{W}_*$，$\mathbf{M}_*$）满足

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

在等式两边分别左乘和右乘 $\mathbf{M}_*$后，我们得到关系式$\mathbf{M}_*^3 = \mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top$。取立方根可得 $\mathbf{M}_* = (\mathbf{W}_*\mathbf{A}\mathbf{W}_*^\top)^{\frac{1}{3}}$。接下来，将 (3) 中的导数设为零，我们得到

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

在等式两边右乘$\mathbf{W}_*^\top$，并代入$\mathbf{M}_*^{-1}$ =

$(W_*AW_*^\top)^{-\frac{1}{3}}$，得到 $(W_*AW_*^\top)^{\frac{2}{3}} = W_*W_*^\top$。代入奇异值分解 $W_* = USV^\top$，得到 $V^\top AV = S$。由于 S 是对角矩阵，因此 V 的列向量是 A 的特征向量，而 S 的对角元素是相应的特征值。最后，我们看到 $M_* = USU^\top$。另一方面，若 W_* 和 M_* 如引理陈述所示，则通过代入即可验证 $G(W_*, M_*) = \mathbf{0}$。

□

C 有界的突触权重

在本节中，我们证明引理 3，其指出集合 $\mathcal{N}$ 中具有发散轨迹的初始条件的勒贝格测度为零。在附录的其余部分，我们固定 $\tau = \frac{1}{2}$。首先需要以下引理。

引理 6。设 $A \in \mathcal{S}_{++}^n$ 且 $(W(t), M(t))$ 是常微分方程 (3)–(4) 的一个解，其初始条件为 $(W_0(t), M_0(t)) \in \mathcal{D} \setminus \mathcal{N}$。则 $\liminf_{t\to\infty} \det(M(t)) > 0$，且 $\limsup_{t\to\infty} \|M(t)\| < \infty$。证明。采用反证法，假设 $\liminf_{t\to\infty} \det(M(t)) = 0$。那么，结合引理 2，这意味着存在一个序列 $\{t_j\}$，满足当 $j \to \infty$ 时 $t_j \to \infty$，使得

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

并且对于每个 $j = 1, 2, \ldots\ldots$，

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

由于（$\mathbf{W}$，$\mathbf{M}$）$\notin \mathcal{N}$，我们可以选择足够大的 j，使得

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

然而，根据常微分方程（3），

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

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

这与方程（14）矛盾。因此，$\liminf_{t\to\infty}$ 行列式$(\mathbf{M}(t)) > 0$。

接下来只需证明 $\limsup_{t\to\infty}\|\mathbf{M}(t)\| < \infty$。再次使用反证法，假设 $\limsup_{t\to\infty}\|\mathbf{M}(t)\| = \infty$。根据前述结果，$K = \sup_{t\geq 0}\|\mathbf{M}(t)^{-1}\| < \infty$。因此，存在$t > 0$ 充分大的，使得

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

and

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

根据常微分方程（4）和引理2,

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

这与方程（15）矛盾。因此，$\limsup_{t\to\infty}\|\mathbf{M}(t)\| < \infty$。 □

引理3的证明。 事实是 $\mathcal{N}$ 的勒贝格测度为零，这显然是因为 $\mathcal{N} \subset\{(\mathbf{W},\mathbf{M}) \in \mathcal{D}:$ 行列式$(\mathbf{W}\mathbf{W}^\top) = 0\}$，行列式$(\cdot)$ 是一个非零多项式，而非零多项式的零点集的勒贝格测度为零 [26]。

设 $(\mathbf{W}(t), \mathbf{M}(t))$ 是常微分方程 (3)–(4) 的一个解，其初始条件为 $(\mathbf{W}_0,\mathbf{M}_0)$。在 $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{D}\setminus\mathcal{N}$ 的情况下，引理 2 和引理 6 表明，$\|\mathbf{M}(t)^{-1}\|$ 和 $\|\mathbf{W}(t)\|$ 在

t 上一致有界。另一方面，假设 $(\mathbf{W}_0, \mathbf{M}_0) \in \mathcal{N}$，且 $(\mathbf{W}(t), \mathbf{M}(t))$ 是方程 (3)–(4) 在初始条件 $(\mathbf{W}_0, \mathbf{M}_0)$ 下的解。那么存在 $t > 0$、$\mathbf{v} \in \mathbb{R}^k$ 和 $\lambda \in \mathbb{R}$，使得 $\mathbf{W}(t)^\top \mathbf{v} = 0$ 且 $\mathbf{M}(t)\mathbf{v} = \lambda \mathbf{v}$。于是

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

由此可知，对于 $t > 0$，$\mathbf{W}(t)^\top \mathbf{v} = 0$ 且 $\mathbf{M}(t)\mathbf{v} = \lambda e^{-2t}\mathbf{v}$，$(\mathbf{W}(t), \mathbf{M}(t)) \in \mathcal{N}$。因此，对于每个 $t > 0$，$\mathbf{v}$ 都是 $\mathbf{M}(t)$ 的特征向量，其特征值为 λe^{-2t}。由此可得 $\lim_{t\to\infty} \det(\mathbf{M}(t)) = 0$。 □

D 神经滤波器收敛到特征子空间

本节将证明引理 5。

引理 5 的证明。 设 $(\mathbf{W}(t), \mathbf{M}(t))$ 是常微分方程 (3)–(4) 的一个解。根据 $V(\mathbf{W})$ 在方程 (11) 中的定义以及常微分方程 (3)，有（省略对 t的依赖）

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

根据引理 3，$\mathbf{W}(t) - (\mathbf{W}(t)\mathbf{W}(t)^\top)^{-\frac{1}{2}}\mathbf{W}(t)\mathbf{A}$ 在 t中一致有界。由引理 2 和 3，

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

因此，结合 LaSalle 不变性原理和引理 4，这意味着 $(\mathbf{W}(t),\mathbf{M}(t))$

收敛到集合 $\mathcal{E}$，当 $t \to \infty$。证明至此完成。 □

E 不同时间尺度下常微分方程的稳定性

在第 7 节中，我们推测，正文中针对 $\tau = \frac{1}{2}$（定理 1）证明的全局收敛结果适用于所有 $0 < \tau \leq \frac{1}{2}$，且动力学保留相同的两阶段结构：(i) 快速收敛到依赖于 τ 的不变流形 $\mathcal{O}_\tau$，由凸李雅普诺夫函数 L_τ刻画；(ii) 沿 $\mathcal{O}_\tau$缓慢演化，该演化由依赖于 τ 的非凸势函数 V_τ的梯度流所支配，其极小值对应于主子空间。这里，我们提供了一些有限的实证证据，以支持这一猜想。

E.1 $\tau \to 0$ 情况下的凸李雅普诺夫函数

在 $\tau \to 0$ 区间内，循环权重首先收敛到不变流形 $\mathcal{O}_0 = \{(\mathrm{W}, \mathrm{M}) \in \mathcal{D} : \mathrm{M}^{-1}\mathrm{WAW}^\top\mathrm{M}^{-1} = \mathrm{M}\}$，而前馈权重保持不变；随后，两种权重都在该不变流形内演化。下面简要说明第一阶段指数收敛的论证，其中假设前馈权重 W 保持不变；也就是说，$\mathrm{dW}\,\mathrm{d}_t = 0$。回顾第7节中的凸李雅普诺夫函数 $L_0(\mathrm{W}, \mathrm{M}) = \|\mathrm{WAW}^\top - \mathrm{M}^3\|^2$。假设 W 固定，且 M($t$) 是常微分方程 (4) 的一个解。根据乘积法则和常微分方程 (4)，并省略对 t的依赖，我们有

—

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

接下来，根据链式 法则、前面的公式以及迹算子的循环性质，我们有 e

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

因此，在第一阶段，李雅普诺夫函数 L_0(**W**,**M**) 指数收敛至零。

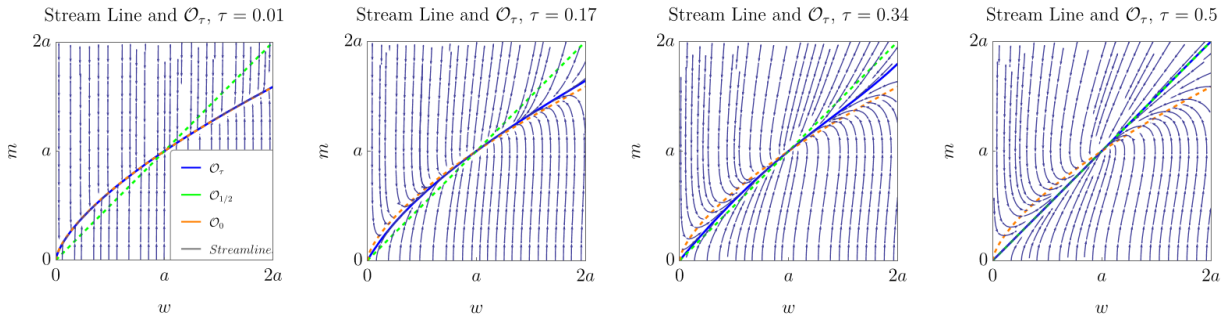

图6：对于每个 τ，箭头表示常微分方程（3）–（4）的流线。蓝色实线是不变集 $\mathcal{O}_\tau$ 的数值估计。橙色和绿色虚线分别是不变集 $\mathcal{O}_0$ 和 $\mathcal{O}_{\frac{1}{2}}$。

E.2 标量情形下不变流形的估计

考虑标量情形 $n = k = 1$，此时临界点为 $\pm(a, a)$。图 6 展示了常微分方程 (3)–(4) 在 $w > 0$ 下的流线，这些流线是使用 Mathematica 内置的 Streamline 函数计算得到的，并在不同的 $\tau \in (0, \frac{1}{2}]$ 下收敛。不变流形 $\mathcal{O}_0$ 和 $\mathcal{O}_{\frac{1}{2}}$ 分别以橙色和绿色显示。对于中间的 τ，我们使用下一段中描述的方法数值估计 $\mathcal{O}_\tau$。请注意，随着 τ 从 0 变化到 $\frac{1}{2}$，$\mathcal{O}_\tau$ 似乎在 $\mathcal{O}_0$ 和 $\mathcal{O}_{\frac{1}{2}}$ 之间插值；进一步的模拟（未展示）也支持这一点。

$\mathcal{O}_\tau$ 的数值估计（$w > 0$）。为估计 $a < w < 2a$时的 $\mathcal{O}_\tau$，我们定义函数h：$(0, \infty)^2 \to \mathbb{R}^2$ 如下。给定一点（w_0，m_0），令（$w(t)$ $m(t)$）表示初始条件为（w_0，m_0）的常微分方程的解，并设 $h(w_0, m_0) = \lim_{t\to\infty}(\dot{w}(t)\ \dot{m}(t))$。若点（$w_0$，$m_0$）处的 $\mathcal{O}_\tau$ 满足：$h(w_0, m_0)$与向量场 G 的雅可比矩阵中对应最小特征值的特征向量平行，则该点属于 $\mathcal{O}_\tau$。随后，我们进行数值搜索，以找到直线 $w = 2a$上的一点（w_0，m_0）$\in \mathcal{O}_\tau$，再计算从（w_0，m_0）出发的常微分方程的解，以估计 $a < w < 2a$时的 $\mathcal{O}_\tau$。我们借助 Mathematica 内置的 NDSolve 和 FindRoot 函数实现了这一数值过程。为估计 $0 < w < a$时的 $\mathcal{O}_\tau$，可以采用类似的论证，从原点附近的点（w_0，m_0）出发。不过，还有一种不需要数值搜索的方法。具体来说，我们假设原点属于不变

流形 $\mathcal{O}_\tau$ 的闭包，并且当 m 较小时，不变流形可以用线性近似；也就是说，原点附近的不变流形可以用 $\{(v_0 m,\ m):\ m \approx 0\}$ 很好地近似，其中 $v_0 > 0$为某个常数。在此假设下，当（w_0，m_0）较小时，常微分方程（3）–（4）的解在 $t = 0$ 处的导数近似为 $\dot{w}(0) \approx \frac{2aw_0}{m_0}$和 $\dot{m}(0) \approx \frac{aw_0^2}{\tau m_0^2}$。利用 $w_0 \approx v_0 m_0$以及 $\dot{w}(0) \approx v_0 \dot{m}(0)$，可知$2av_0 \approx \frac{a}{\tau} v_0^3$，因此 $v_0 = \sqrt{2\tau}$。于是，我们通过计算从（$\sqrt{2\tau\alpha}$，α）出发、 $\alpha > 0$ 较小时的解，来估计 $\mathcal{O}_\tau$。这一数值过程也可以推广到估计 d、$k > 1$时的不变流形；不过，流形的可视化会更加困难。

参考文献

[1] Cengiz Pehlevan、Tao Hu 和 Dmitri B Chklovskii。《用于线性子空间学习的赫布／反赫布神经网络：从流式数据的多维尺度分析推导而来》。《神经计算》，27(7):1461–1495，2015年。

[2] Erkki Oja。作为主成分分析器的简化神经元模型。《数学生物学杂志》，15(3):267–273，1982年。[3] Chi-Ning Chou 和 Mien Brabeeba Wang。用于解决流式主成分分析的生物学版本 Oja 规则的常微分方程启发式分析。载于《学习理论会议》，第 1339–1343 页。PMLR，2020年。[4] Peter F¨oldiak。用于最优线性特征提取的自适应网络。载于《IEEE/INNS 国际联合神经网络会议论文集》，第1卷，第401–405页，1989年。[5] Jeanne Rubner 和 Paul Tavan。用于主成分分析的自组织网络。《欧洲物理快报》，10(7):693，1989年。[6] Jeanette Rubner 和 Klaus Schulten。通过自组织开发特征检测器。《生物控制论》，62(3):193–199，1990年。[7] Todd K Leen。线性特征发现网络中的学习。载于《自适应信号处理》，第1565卷，第472–481页。国际光学与光子学学会，1991年。

[8] Cengiz Pehlevan、Anirvan M Sengupta 和 Dmitri B Chklovskii。为什么相似性匹配目标会导出赫布/反赫布网络？《神经计算》，30(1):84–124，2018年。[9] Erkki Oja 和 Juha Karhunen。对学习版本子空间方法收敛性的分析。《数学分析与应用杂志》，91(1):102–111，1983年。[10] Terence D Sanger。单层线性前馈神经网络中的最优无监督学习。《神经网络》，2(6):459–473，1989年。[11] Todd Leen、Mike Rudnick 和 Dan Hammerstrom。赫布特征发现提高分类器效率。载于《1990年IJCNN国际神经网络联合会议》，第51–56页。IEEE，1990年。

[12] Erkki Oja 和 Juha Karhunen。关于随机矩阵期望的特征向量与特征值的随机逼近。《数学分析与应用杂志》，106(1):69–84，1985年。

[13] Kurt Hornik 和 Chung-Ming Kuan。局部特征提取算法的收敛分析。神经网络，5(2):229–240，1992年。

[14] Erkki Oja。主成分、次成分与线性神经网络。神经网络，5(6):927–935，1992年。

[15] Mark D Plumbley. 主成分算法收敛的李雅普诺夫函数。《神经网络》，8(1):11–23，1995年。
[16] Vivek S Borkar. 双时间尺度下的随机逼近。《系统与控制快报》，29(5):291–294，1997年。
[17] Harold Joseph Kushner 和 Dean S Clark. 受约束和无约束系统的随机逼近方法，第26卷。Springer Science & Business Media，2012年。[18] Tianyi Lin、Chi Jin 和 Michael Jordan. 关于非凸-凹极小极大问题的梯度下降上升。载于国际机器学习会议，第6083–6093页。PMLR，2020年。

[19] Veronica Centorrino、Francesco Bullo 和 Giovanni Russo. 相似性匹配网络：赫布学习与跨多个时间尺度的收敛。*arXiv*预印本 *arXiv:2506.06134*，2025年。[20] Joseph P LaSalle. 常微分方程的稳定性理论。《微分方程杂志》，4(1):57–65，1968年。[21] Rafael Potrie 和 Pablo Monz´on. 几乎全局稳定性的局部推论。《动力系统》，24(1):109–115，2009年。

[22] Morris W. Hirsch、Charles C. Pugh 和 Michael Shub。《不变流形》，《数学讲义》第 583 卷。施普林格出版社，柏林、海德堡，1977年。ISBN 978-3-540- 08054-5。doi: 10.1007/BFb0087003。2006年再版。

[23] Cengiz Pehlevan 和 Dmitri B Chklovskii。受神经科学启发的在线无监督学习算法：人工神经网络。《IEEE信号处理杂志》，36(6):88–96，2019年。

[24]David Lipshutz、Yanis Bahroun、Siavash Golkar、Anirvan M Sengupta 和 Dmitri B Chklovskii。用于多通道典型相关分析的生物学上合理的神经网络。《神经计算》，33(9):2309–2352，2021年。

[25]David Lipshutz、Yanis Bahroun、Siavash Golkar、Anirvan M Sengupta 和 Dmitri B Chklovskii。用于推导具有多区室神经元和非赫布可塑性的神经网络的规范性框架。《PRX Life》，1:013008，2023年。

[26] Richard Caron 和 Tim Traynor。多项式的零点集。技术报告，第 1–2 页，2005年。