arXiv:2602.01883v1 [math.NA] 2 Feb 2026

临界流形上退化鞍点的高指标鞍点动力学收敛性*

Tao Luo†、Jianyuan Yin‡、Lei Zhang§、和Shixue Zhang¶

摘要。 高指标鞍点动力学（HiSD）方法为寻找鞍点和构建解景观提供了强有力的框架。尽管 HiSD 最初是针对非退化临界点推导的，但它在退化情形下也已展现出良好的实证效果，此时 Hessian矩阵具有零特征值。然而，针对退化鞍点的 HiSD 数学分析和数值分析仍未得到探索。本文利用 Morse–Bott 函数，对在临界流形上计算退化鞍点的 HiSD 进行了严格分析。我们证明了连续高指标鞍点动力学的局部收敛性，并确立了离散高指标鞍点动力学算法的线性收敛速率。此外，我们为梯度对齐趋势提供了理论解释，揭示梯度方向渐近地与某个特定 Hessian特征向量对齐。我们的分析还阐明了在退化鞍点情形下，HiSD 指标选择的灵活性。我们通过神经网络损失景观上的数值实验验证了分析结果，并证明 HiSD 的动量加速变体能够快速收敛至退化鞍点。

K关键词。高指标鞍点动力学、鞍点、退化鞍点、解景观， Morse–Bott 函数、线性收敛

MSC 分类号。37M05, 37N30, 65L20

1. 引言。鞍点的计算长期以来一直是各个科学领域的重要课题。光滑能量景观上的鞍点在相变 [5, 36, 33, 47], 化学反应 [17, 19], 以及机器学习中神经网络的训练 [7, 6, 50]方面发挥着关键作用。然而，由于鞍点具有不稳定性，定位鞍点是一项极具挑战性的问题，尤其是在转变机制未知的高维系统中。

人们开发了多种计算鞍点的数值方法，大致可分为两类。路径搜索方法，例如弦方法 [8] 和微扰弹性带方法 [15], ，旨在寻找初态与终态之间的最小能量路径和过渡态。相比之下，曲面行走方法旨在从初始猜测出发定位鞍点。用于指标-1鞍点的曲面行走方法包括最温和上升动力学[9], 、二聚体类方法 [16, 45], 以及迭代最小化表述 [11]。近年来，高指标鞍点动力学（HiSD）方法 [43] 提供了一个统一的曲面行走框架，用于搜索指标-k 鞍点，为

*投稿至。**经费支持：** L.Z. 得到国家自然科学基金委员会（编号：12225102、T2321001、12288101）和国家重点研发计划（编号：2024YFA0919500）的资助。L.T. 得到国家重点研发计划（编号：2022YFA1008200）及上海数学与交叉学科研究院（SIMIS）SIMIS-ID-2025-ST 项目的资助。J.Y. 得到“中央高校基本科研业务费专项资金”（编号：2253100016）的资助。

†上海交通大学数学科学学院、自然科学研究院、教育部数学与复杂系统重点实验室、上海市计算数学学会，中国上海 200240（luotao41@sjtu.edu.cn）。

‡北京师范大学数学科学学院、数学与复杂系统实验室、教育部，中国北京 100875（jyyin@bnu.edu.cn）。§北京大学数学科学学院、北京国际数学研究中心、定量生物学中心、机器学习研究中心，中国北京 100871（zhangl@math.pku.edu.cn）。

¶北京大学数学科学学院，中国北京 100871（zhang-shixue@stu.pku.edu.cn）。

各种扩展和变体 [42, 25, 35]。解景观由所有驻点及其连接构成，可通过将 HiSD 方法与向上/向下搜索算法结合来构建 [41, 42, 46]。解景观方法已在许多科学应用中展现出其有效性，包括液晶 [41, 14], 准晶体成核 [40], 反应扩散系统 [37],和玻色–爱因斯坦凝聚体 [38]。

对于 HiSD 方法，在推导与理论分析中，通常假定目标鞍点是非退化的。若能量函数在该点处的海森矩阵（二阶导数）具有有界逆，即海森矩阵的所有特征值均非零，则称该鞍点为非退化鞍点。根据经典 Morse理论 [27], ，非退化鞍点本质上是孤立的，且莫尔斯引理保证在此类点附近，能量函数局部表现为非退化二次型。在非退化性假设下，鞍点可刻画为极小极大问题的解。对 HiSD 的数学与数值分析已确立其局部稳定性与收敛性 [43, 26, 48, 35]。

然而，在许多实际应用的能量函数中，驻点（包括鞍点和极小值点）并非孤立，而是构成一个连续流形，通常称为临界流形，从而导致退化。例如，在物理系统中，当连续对称性破缺为离散对称性时，会出现戈德斯通模 [12], ，其数目等于零特征值的重数。值得注意的是，在过参数化神经网络中，经验研究表明，训练过程中很大一部分海森矩阵特征值聚集在零附近 [32]；因此，损失景观呈现出显著的退化程度。此外，如训练过程中的参数凝聚现象所示[24], ，神经网络的内在对称性导致退化鞍点的存在 [50]。

为了从数学上描述这类结构，Morse–Bott 函数理论将经典Morse理论推广到这些退化情形，使临界点能够形成光滑流形而非孤立点 [1]。与经典Morse理论一致 [27], ，临界点的指标定义为海森矩阵的负惯性指数，而零化度对应于零特征值的重数。该框架为分析复杂系统中的能量景观提供了一种自然的语言，并已被应用于神经网络的损失景观 [49]。

尽管HiSD 方法的推导与分析基于非退化性假设，该方法已在各种科学问题中成功识别出退化鞍点 [40, 44, 39, 38]。通过将这些零特征向量视为不稳定方向，HiSD 方法能够成功爬出由极小值点构成的临界流形并定位一个退化鞍点，而许多针对指标-1鞍点的曲面行走方法因无法容纳零特征向量而无法做到这一点。具体而言，为搜索零化度 m的指标-s鞍点，数值证据表明可以采用指标从 s 到（$s+m$）的HiSD。然而，关于为何这些指标下的HiSD会收敛到退化鞍点的严格理论解释，包括稳定性分析与收敛速率，在以往文献中仍然缺失。

本文旨在将 Morse–Bott 函数视为 Morse 函数的推广，从而建立将 HiSD 应用于退化鞍点的理论框架。具体而言，对于位于临界流形上且在法空间中非退化的鞍点，在若干假设下，我们得出以下结果：

- 我们证明了连续高指标鞍点动力学下临界流形的局部渐近稳定性，并证明其收敛到流形上的一个点。
- 我们证明，采用精确特征向量的离散 HiSD 算法以线性收敛速率收敛到临界流形。
- 我们对梯度对齐趋势给出了严格论证：在迭代过程中，梯度方向渐近地与某个特定的海森矩阵特征向量对齐。
- 我们展示了 HiSD 指标的灵活性，表明理论结果适用于从 s 到 $s+m$的指标范围，其中 s 为指标， m为零化度。
- 我们通过数值实验验证了这些发现，并表明 HiSD 的动量加速变体能够快速收敛到退化鞍点。

本文其余部分安排如下。在第 2 节中，我们给出 Morse–Bott 函数的定义，并回顾 HiSD 的连续动力学和离散算法。在第 3 节中，我们证明渐近稳定性和收敛结果，并解释梯度对齐趋势。在第 4 节中，我们通过数值实验验证理论发现，并展示 HiSD 的动量加速变体在退化情形下的高效性。最后，我们在第 5 节中给出结论，并在附录A中提供引理的证明。

2. 预备知识。

2.1. Morse–Bott 函数。 对于光滑函数 $E\colon \mathbb{R}^d \to \mathbb{R}$，若点 $\theta \in \mathbb{R}^d$处的梯度为零，即 $\nabla E(\theta) = 0$，则称其为临界点。临界点的（*Morse*）指标定义为其海森矩阵 $\nabla^2 E$ 的负特征值个数，它刻画了临界点的局部性质。临界点的零化度定义为其海森矩阵 $\nabla^2 E$的零特征值重数，即海森矩阵的核的维数。如果某点处的海森矩阵存在有界逆，则称该临界点为非退化，也就是说，海森矩阵没有零特征值。由此得到 Morse 函数 [27]，的以下定义；Morse 函数构成了 Morse理论的基础，并在梯度流分析中发挥核心作用。

Definition 2.1（Morse 函数）。若光滑函数 E 的所有临界点均为非退化点，则称其为Morse 函数。

为处理退化情形，Morse–Bott理论通过以临界子流形代替孤立临界点来推广经典Morse理论 [1, 4]。记 $\mathcal{M}_E$ 为光滑函数 E： $\mathbb{R}^d \to \mathbb{R}$的全部临界点所成之集，并令 $\mathcal{M}$ 为 $\mathcal{M}_E$的一个连通分支。Morse–Bott理论的关键洞见在于考察海森矩阵在临界子流形 $\mathcal{M}$上的行为。基于 $\mathbb{R}^d$中的欧氏度量，每个点 $\theta \in \mathcal{M}$ 处的整个空间容许正交分解 $\mathbb{R}^d = T_\theta\mathcal{M} \oplus N_\theta\mathcal{M}$，其中 $N_\theta\mathcal{M}$ 是 $\mathcal{M}$ 在 θ处的法空间。

一个重要事实是，临界点 θ 处的切空间 $T_\theta\mathcal{M}$完全位于海森矩阵的核中，这可以直接推导得出。对任意$w \in T_\theta\mathcal{M}$，存在光滑曲线 γ： $(-\varepsilon,\varepsilon) \to \mathcal{M}$ 满足 $\gamma(0) = \theta$ 与$\dot\gamma(0) = w$。由于对所有 $t \in (-\varepsilon,\varepsilon)$都有 $\nabla E(\gamma(t)) = 0$，对 t求导可得 $\nabla^2 E(\gamma(t))\dot\gamma(t) = 0$。因此，在 $t = 0$处，我们有 $\nabla^2 E(\theta)w = 0$，这意味着 $T_\theta\mathcal{M} \subseteq \mathrm{Ker}(\nabla^2 E(\theta))$。这一观察表明，海森矩阵的非平凡行为被限制在法空间 $N_\theta\mathcal{M}$上。因此，非退化性的概念可以自然地推广到临界子流形 $\mathcal{M}$。

Definition 2.2（非退化临界子流形）。光滑子流形 $\mathcal{M}$

$\hookrightarrow \mathbb{R}^d$ 称为 非退化临界子流形 E，如果 [29]：
(1) $\mathcal{M} \subset \mathcal{M}_E$; *(2)* $\mathcal{M}$ 是连通的；*(3)* $\forall \theta \in \mathcal{M}$, $\nabla^2 E(\theta)$ 在 $N_\theta \mathcal{M}$ 上是非退化的。

注意，条件 (3) 等价于 $T_\theta \mathcal{M} = \mathrm{Ker}(\nabla^2 E(\theta))$。最后，我们介绍 Morse–Bott 函数的定义。

DEFINITION 2.3（Morse–Bott 函数）。光滑函数 $E : \mathbb{R}^d \to \mathbb{R}$ 称为 Morse–Bott 函数，如果 $\mathcal{M}_E$ 的每个连通分支都是非退化临界子流形。

对于 Morse–Bott 函数，以下引理给出了一个基本结果，它在临界子流形 [2] 上的临界点附近提供了局部标准形。

LEMMA 2.4（*Morse–Bott* 引理）。设 $E: \mathbb{R}^d \to \mathbb{R}$ 是一个 *Morse–Bott* 函数，且 $\mathcal{M}$ 是 $\mathcal{M}_E$ 的一个维数为 m 的连通分支。对于任意指标为 s 的鞍点 $\theta^* \in \mathcal{M}$，存在 θ^* 的一个邻域 $U \subset \mathbb{R}^d$，以及一个局部微分同胚 $\phi: U \to V \subset \mathbb{R}^m \times \mathbb{R}^{d-m}$，使得 $\phi(\theta^*) = 0$。若将局部坐标记为 $(x, y) = (x_1, \ldots, x_m, y_1, \ldots, y_{d-m})$，则 ϕ 在局部将流形展平，使得 $\phi(U \cap \mathcal{M}) = \{(x, y) \in V: y = 0\}$。此外，对所有 $\theta \in U$，函数 E 可在这些坐标中表示为 $\phi(\theta) = (x, y)$：

$$E(\theta) = E(\mathcal{M}) - y_1^2 - \cdots - y_s^2 + y_{s+1}^2 + \cdots + y_{d-m}^2. \tag{2.1}$$

该引理表明，在任意临界子流形附近，Morse–Bott 函数沿临界子流形保持常数，并且在法向方向上表现为非退化二次型。

2.2. HiSD 方法回顾。近年来，HiSD 方法已成为搜索任意指标鞍点的有力工具 [43]。假设能量函数为 $E \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$，则能量函数 E 的指标为 k 的鞍点（k-HiSD）所对应的 HiSD 形式如下：

$$\begin{cases} \dot{\theta} = -\beta \left(I - \sum_{p=1}^{k} 2 v_p v_p^\top \right) \nabla E(\theta), \\ \dot{v}_p = -\zeta \left(I - v_p v_p^\top - \sum_{q=1}^{p-1} 2 v_q v_q^\top \right) \nabla^2 E(\theta) v_p, \quad p = 1, \cdots, k, \end{cases} \tag{2.2}$$

其中，$\{v_p\}_{p=1}^k$ 为正交归一向量，用于近似海森矩阵 $\nabla^2 E(\theta)$ 最小的 k 个特征值所对应的特征向量。β 和 ζ 为正松弛参数。$v^\top w$ 表示欧几里得空间 $\mathbb{R}^d$ 中 v 与 w 的内积。

该动力系统（2.2）源自非退化指标为 k 的鞍点的极小极大变分刻画，以及与其相应的 k 维不稳定子空间的构造。该动力学在由 $\{v_p\}_{p=1}^k$ 张成的子空间上采用梯度上升流，这些向量近似 k 的最小 $\nabla^2 E(\theta)$ 个特征向量；同时在其正交补上采用梯度下降流，从而引导搜索指标为 k 的鞍点。若初始条件满足 $\{v_p\}_{p=1}^k$ 为正交归一向量，则方程（2.2）的解也满足这一正交归一条件。

2.3. HiSD 的数值算法。在离散数值算法中，我们用 $\theta^{(t)} \in \mathbb{R}^d$ 表示经过 t 个迭代步后的参数值，并用 $v_p^{(t)}$ 表示方向。HiSD 方法的显式欧拉格式见算法 2.1。

这里，“EigenSol”表示一个特征向量求解器，在给定初值 $\{v_p^{(t)}\}_{p=1}^k$ 后，计算对应于最小的 k 个 $\nabla^2 E(\theta^{(t+1)})$ 的特征值的特征向量，既可以精确计算，也可以近似计算。典型选择包括同步瑞利商迭代最小化方法 [23], 以及局部最优块预条件共轭梯度法（LOBPCG） [21]。

算法 2.1 k-HiSD 数值算法

Input: $k \in \mathbb{N}$, $\theta^{(0)} \in \mathbb{R}^d$, $\left\{v_p^{(0)}\right\}_{p=1}^k \subset \mathbb{R}^d$ satisfying $v_p^{(0)\top} v_q{}^{(0)} = \delta_{pq}$.

for $t = 0, 1, \cdots, T-1$ **do**

$$\theta^{(t+1)} = \theta^{(t)} - \beta_t \left(I - \textstyle\sum_{p=1}^k 2 v_p^{(t)} v_p^{(t)\top}\right) \nabla E(\theta^{(t)}). \qquad \text{(main iteration)}$$

$$\left\{v_p^{(t+1)}\right\}_{p=1}^k = \text{EigenSol}\left(\left\{v_p^{(t)}\right\}_{p=1}^k, \nabla^2 E(\theta^{(t+1)})\right).$$

Output: $\theta^{(T)}$

在实际计算中，HiSD 的显式欧拉格式通常会遭遇收敛缓慢的问题，因此需要进行大量迭代才能计算出鞍点。为了解决这一问题，人们开发了 HiSD 的动量加速变体。受 Heavy-Ball 技术的简洁性和加速效率启发， [31], 提出了采用 Heavy-Ball 加速的 HiSD 方法[25]。令 $\theta^{(-1)} = \theta^{(0)}$，则算法 2.1 中的主要迭代步骤为：

$$\theta^{(t+1)} = \theta^{(t)} - \beta_t \left(I - \sum_{p=1}^k 2 v_p^{(t)} v_p^{(t)\top}\right) \nabla E(\theta^{(t)}) + \gamma(\theta^{(t)} - \theta^{(t-1)}), \tag{2.3}$$

其中， $\gamma \in [0, 1)$ 为动量参数。

类似地，Nesterov 加速 [28] 也被应用于改进 HiSD 方法的性能。具体而言，已在 [22], 中实现了结合该技术的变体，其算法 2.1 中的主要迭代步骤如下：

$$\xi^{(t)} = \theta^{(t)} + \gamma_t \left(\theta^{(t)} - \theta^{(t-1)}\right), \quad \theta^{(t+1)} = \xi^{(t)} - \beta_t \left(I - \sum_{p=1}^k 2 v_p^{(t)} v_p^{(t)\top}\right) \nabla E(\xi^{(t)}),$$

其中 γ_t 是动量参数，通常设为 $\gamma_t = \frac{t}{t+3}$。在实际计算中，常采用重启策略来进一步提升性能。

3. 退化情形下 HiSD 的分析。 本节将对临界流形上退化鞍点的 HiSD 进行理论分析。

首先，基于 引理 2.4中的局部标准形，我们通过一个简单的例子来阐述我们的思路。考虑 $\theta = (x, y) = (x_1, \cdots, x_m, y_1, \cdots, y_{d-m}) \in \mathbb{R}^d$ 以及能量函数的标准形式，

$$E(\theta) = -y_1^2 - \cdots - y_s^2 + y_{s+1}^2 + \cdots + y_{d-m}^2.$$

在此示例中， $\mathcal{M} = \{(x, 0) \in \mathbb{R}^d\}$ 构成一个指标为 s 的 m 维非退化临界流形，其上的点均为鞍点；而且，对于任意一点， $U = \{(x, 0) \in \mathbb{R}^d\}$ 也构成海森矩阵的零空间。对于带有 $k = s, \cdots, s+m$ 的 k-HiSD，我们取前 s 个不稳定方向作为特征向量 $v_p = (0, e_p)$、 $p = 1, \cdots, s$，其中 e_p 是

标准基向量，其余的（$k-s$）个不稳定方向 v_p、 $p=s+1,\cdots,k$取为 U 中的任意标准正交集。于是，采用固定步长 $\beta_t\equiv\beta\in$ (0,1) 的显式欧拉格式，

$$\theta^{(t+1)}=\theta^{(t)}-\beta\left(0,2y^{(t)}\right)=\left(x^{(t)},(1-2\beta)y^{(t)}\right),$$

以线性速率收敛到鞍点 $\left(x^{(0)},\ 0\right)\in\mathcal{M}$ 。此例直观地说明了 HiSD 的指标 k 以及退化情形下 $(k-s)$ 个不稳定方向的灵活性。

现在我们继续进行严格的理论分析。假设 $\mathcal{M}\subset\mathbb{R}^d$ 是一个闭的、 $\mathcal{C}^2$嵌入的、连通的、维数为 m 的非退化临界流形，由指标为 s 的鞍点组成；函数 E 是一个海森矩阵有界的 $\mathcal{C}^2$ 函数。在此设定下，我们研究 k-HiSD 向鞍流形 $\mathcal{M}$的收敛。具体而言，我们证明 k-HiSD 在 $k=s$、.....、 $s+m$ 条件下的局部收敛，并推导相应的收敛速率。

3.1. 记号。对于子空间 $V\subset\mathbb{R}^d$，符号 $V^\perp$ 表示其正交补，而 $\mathcal{P}_V\colon\mathbb{R}^d\to\mathbb{R}^d$ 表示到 V上的正交投影算子。类似地，我们定义最近点投影算子 $\mathcal{P}_\mathcal{M}$,

$$\mathcal{P}_\mathcal{M}(\theta)=\operatorname*{arg\,min}_{\theta'\in\mathcal{M}}\|\theta-\theta'\|_2, \tag{3.1}$$

将 $\theta\in\mathbb{R}^d$ 投影到鞍流形 $\mathcal{M}$ 上，为简洁起见，我们用 $\hat{}\theta$ 表示 $\mathcal{P}_\mathcal{M}(\theta)$。最近点投影满足如下重要性质:

$$\theta-\hat{\theta}\in\operatorname{Im}(\nabla^2E(\hat{\theta}))=\left(\operatorname{Ker}(\nabla^2E(\hat{\theta}))\right)^\perp. \tag{3.2}$$

注 3.1. 最近点投影算子 $\mathcal{P}_\mathcal{M}$ 至少在 $\mathcal{M}$的一个邻域内有定义。

1. 存在性。 由于 $\mathcal{M}$ 在 $\mathbb{R}^d$中是闭的，定义中的下确界

$$\operatorname{dist}(\theta,\mathcal{M})=\inf_{\theta'\in\mathcal{M}}\|\theta-\theta'\|_2, \tag{3.3}$$

对每个 $\theta\in\mathbb{R}^d$都可以取到。

2. 唯一性。 根据管状邻域定理 [18], 对每个 $\vartheta\in\mathcal{M}$，存在 $r_\vartheta>0$ 使得法丛映射

$$(\theta',v)\mapsto\theta'+v,\quad \theta'\in\mathcal{M},\ v\in N_{\theta'}\mathcal{M},\ \|v\|_2<r_\vartheta,$$

是到开管上的微分同胚

$$\mathcal{U}_\vartheta=\left\{\theta\in\mathbb{R}^d:\operatorname{dist}(\theta,\mathcal{M})<r_\vartheta,\|\theta-\vartheta\|_2<r_\vartheta\right\}.$$

因此，对于每个 $\theta\in\mathcal{U}_\vartheta$，最近点投影 $\mathcal{P}_\mathcal{M}(\theta)$ 是唯一的 .

注3.2. $\hat{}\theta$ 是约束问题 $\min_{\theta'\in\mathcal{M}}\ \frac{1}{2}\|\theta-\theta'\|_2^2$的局部极小值点，且 KKT条件 $\theta-\hat{}\theta+\nabla^2E(\hat{}\theta)\lambda=0$ 给出 $\theta-\hat{}\theta\in\operatorname{Im}(\nabla^2E(\hat{}\theta))$。式（3.2）中的等式源于 $\nabla^2E(\hat{}\theta)$ 的对称性。

以下假设在本章节中始终成立，并在鞍流形 $\mathcal{M}$的一个适当管状邻域内提出。将$\nabla^2E(\theta)$的特征值记为 $\{\lambda_p\}_{p=1}^d$ ，并按升序排列；对应的标准正交特征向量记为$\{u_p\}_{p=1}^d$。将 $\nabla^2E(\hat{}\theta)$的特征值和对应的标准正交特征向量分别记为 $\{\hat{}\lambda_p\}_{p=1}^d$ 和 $\{\hat{}u_p\}_{p=1}^d$。

ASSUMPTION 3.3. 存在一个 δ-管状邻域 $\mathcal{U}(\mathcal{M},\delta) = \{\theta \in \mathbb{R}^d : \hat{}\|\theta-\theta\|_2 < \delta\}$ 以及满足 $M\delta \leqslant \mu/4$的正常数 M、 L、 $\mu>0$，使得：

(i) 最近点投影算子 $\mathcal{P}_{\mathcal{M}}$ 在 $\mathcal{U}(\mathcal{M},\delta)$中定义良好。*(ii)* 对所有 $\theta,\theta' \in \mathcal{U}(\mathcal{M},\delta)$, $\|\nabla^2 E(\theta) - \nabla^2 E(\theta')\|_2 \leqslant M\|\theta-\theta'\|_2$。*(iii)* 对所有 $\theta \in \mathcal{U}(\mathcal{M},\delta)$， $\lambda_1,\cdots,\lambda_s \in [-L,-\mu]$, $\lambda_{s+m+1},\cdots,\lambda_d \in [\mu,L]$。

注3.4. 对所有 $\theta \in \mathcal{U}(\mathcal{M},\delta)$，假设 3.3 给出了 $\hat{}\,|\lambda_p|$ 的上界，其中 $p = s+1, \cdots, s+m$。由于 $\theta \in \mathcal{M}$，我们可以直接得到 $\lambda_p = 0$，因此根据 [13,定理 8.1。5], 我们有，

$$|\lambda_p| = |\lambda_p - \hat{\lambda}_p| \leqslant \|\nabla^2 E(\theta) - \nabla^2 E(\hat{\theta})\|_2 \leqslant M\|\theta-\hat{\theta}\|_2 \leqslant M\delta \leqslant \mu/4.$$

因此，假设3.3 确保在 $\mathcal{M}$ 的邻域内，靠近零的特征值与其他特征值之间有良好的分离。

3.2. 连续 HiSD 的局部收敛。对于非退化的指标为 k 的鞍点，可以通过线性化直接确定 k-HiSD 的稳定性[43],，但这种方法不再适用于鞍流形。在退化情形下，我们旨在证明，在鞍流形 $\mathcal{M}$ 附近， $G(\theta) = \frac{1}{2}\|\nabla E(\theta)\|_2^2$ 是 k-HiSD 的局部李雅普诺夫函数。

根据假设3.3，我们定义 m 维子空间 $U = \operatorname{span}\{u_p\}_{p=s+1}^{s+m}$和 $\widehat{U} = \operatorname{span}\{\hat{} u_p\}_{p=s+1}^{s+m}$。与 $\widehat{U}$之间的距离定义为 [13]:

$$\operatorname{dist}(U,\widehat{U}) := \left\|\mathcal{P}_U - \mathcal{P}_{\widehat{U}}\right\|_2 = \left\|\mathcal{P}_{U^\perp} - \mathcal{P}_{\widehat{U}^\perp}\right\|_2, \tag{3.4}$$

可利用以下引理并结合 $C = 2M/\mu$ 进行界定。

LEMMA 3.5（子空间间距离）。对于任意 $\theta \in \mathcal{U}(\mathcal{M}, \delta)$，我们有^ ^ 如下估计： $\operatorname{dist}(U, U) \leqslant C\|\theta-\theta\|_2$。
梯度也可以用以下引理来刻画。

LEMMA 3.6。 对于任意 θ、 $\theta' \in \mathbb{R}^d$，以下恒等式成立：

$$\nabla E(\theta') = \nabla E(\theta) + \nabla^2 E(\theta)(\theta'-\theta) + J(\theta'-\theta),$$
$$where\ J = \int_0^1 \nabla^2 E\left(\theta + t(\theta'-\theta)\right) \mathrm{d}t - \nabla^2 E(\theta) \in \mathbb{R}^{d\times d}.$$

此外，如果 θ、 $\theta' \in \mathcal{U}$（$\mathcal{M}$、 δ），则有 $\|J\|_2 \leqslant \frac{M}{2}\|\theta'-\theta\|_2$。
此外，梯度 $\nabla E(\theta)$ 及其在 U 和 $U^\perp$上的正交投影满足尖锐估计：$\mathcal{P}_{U^\perp}\nabla E(\theta)$ 占主导，并随到流形的距离线性增长，而 $\mathcal{P}_U \nabla E(\theta)$ 是高阶项。

LEMMA 3.7（阶次分离的梯度分解）。对于 $\theta \in \mathcal{U}(\mathcal{M}, \delta)$，我们^有 $\|\nabla E(\theta)\|_2 \leqslant L\|\theta-\theta\|_2$ 以及

$$\|\mathcal{P}_{U^\perp}\nabla E(\theta)\|_2 \geqslant \left(\mu - \tfrac{5}{2}M\|\theta-\hat{\theta}\|_2\right)\|\theta-\hat{\theta}\|_2, \quad \|\mathcal{P}_U \nabla E(\theta)\|_2 \leqslant M\|\theta-\hat{\theta}\|_2^2. \tag{3.5}$$

注3.8. 由（3.5），可得 $\|\nabla E(\theta)\|_2 \geqslant \left(\mu - \frac{5}{2}M\|\theta - \hat{}\theta\|_2\right)\|\theta - \hat{}\theta\|_2 \geqslant \hat{}$ $3\mu\|\theta-\theta\|_2/8$，由于 $M\delta \leqslant \mu/4$，因此管状邻域中的任意临界点$\mathcal{U}(\mathcal{M},\delta)$ 都位于临界流形 $\mathcal{M}$。

我们的第一个结果证明了逐点下降性质：配备某些精确特征向量的 k-HiSD 会使 Lyapunov 函数 G 减小。

THEOREM 3.9（逐点下降）。在假设3.3下，对于 $s \leqslant k \leqslant s+m$，若 $\theta \in \mathcal{U}$（$\mathcal{M}$，δ）$\backslash \mathcal{M}$，那么对于任意正交归一向量 $\{w_p\}_{p=s+1}^{k}$ 属于 U，方向

$$-\left(I-\sum_{p=1}^{s} 2u_p u_p^{\top}-\sum_{p=s+1}^{k} 2w_p w_p^{\top}\right)\nabla E(\theta), \tag{3.6}$$

是 $G(\theta) = \frac{1}{2}\|\nabla E(\theta)\|_2^2$ 的严格下降方向。

证明。记 $g=\nabla E(\theta)$，只需证明

$$K(\theta) := -g^{\top}\nabla^2 E(\theta)\left(I-\sum_{p=1}^{s} 2u_p u_p^{\top}-\sum_{p=s+1}^{k} 2w_p w_p^{\top}\right) g < 0. \tag{3.7}$$

记 $W =$ 张成空间$\{w_p\}_{p=s+1}^{k} \subseteq U$ 且 $\mathcal{P}_W = \sum_{p=s+1}^{k} w_p w_p^{\top}$是到 W 的正交投影。由特征值分解 $\nabla^2 E(\theta) = \sum_{p=1}^{d} \lambda_p u_p u_p^{\top}$以及假设 3.3中的特征值界可得

$$\begin{aligned}
K(\theta) &= \sum_{p=1}^{s} \lambda_p (u_p^{\top} g)^2 - \sum_{p=s+1}^{d} \lambda_p (u_p^{\top} g)^2 + 2g^{\top}\nabla^2 E(\theta)\mathcal{P}_W g \\
&\leqslant -\mu\left(\sum_{p=1}^{s}(v_p^{\top} g)^2 + \sum_{p=s+m+1}^{d}(v_p^{\top} g)^2\right) + \frac{\mu}{4}\left(\sum_{p=s+1}^{s+m}(v_p^{\top} g)^2 + 2\|g\|_2\|\mathcal{P}_W g\|_2\right) \\
&\leqslant -\mu\|\mathcal{P}_{U^{\perp}} g\|_2^2 + \frac{\mu}{4}\left(\|\mathcal{P}_U g\|_2^2 + \|g\|_2^2 + \|\mathcal{P}_U g\|_2^2\right) = -\frac{3\mu}{4}\left(\|\mathcal{P}_{U^{\perp}} g\|_2^2 - \|\mathcal{P}_U g\|_2^2\right).
\end{aligned}$$

应用 引理 3.7，可得

$$\|\mathcal{P}_{U^{\perp}} g\|_2 \geqslant \left(\mu - \tfrac{5}{2}M\|\theta-\hat{\theta}\|_2\right)\|\theta-\hat{\theta}\|_2, \quad \|\mathcal{P}_U g\|_2 \leqslant M\|\theta-\hat{\theta}\|_2^2. \tag{3.8}$$

ˆ 因此，由 $M\|\theta-\theta\|_2 < M\delta \leqslant \mu/4$ 可得

$$\begin{aligned}
K(\theta) &= -\frac{3\mu}{4}\left(\|\mathcal{P}_{U^{\perp}} g\|_2 + \|\mathcal{P}_U g\|_2\right)\left(\|\mathcal{P}_{U^{\perp}} g\|_2 - \|\mathcal{P}_U g\|_2\right) \\
&\leqslant -\frac{3\mu}{4}\|g\|_2\left(\mu - \frac{7}{2}M\|\theta-\hat{\theta}\|_2\right)\|\theta-\hat{\theta}\|_2 \leqslant -\frac{3\mu^2}{32}\|g\|_2\|\theta-\hat{\theta}\|_2 < 0,
\end{aligned} \tag{3.9}$$

至此，证明完成。 □

注 3.10. 由于 $\nabla^2 E(\theta)$ 可能存在多个特征值，特征向量方向的选择可能并不唯一，因此，采用精确特征向量的 k-HiSD 的轨迹可能无法明确定义。对于 $\theta \in \mathcal{U}(\mathcal{M}, \delta)$，注 3.4表明特征值 $\{\lambda_p\}_{p=s+1}^{s+m}$ 接近于零，且引理 3.7 指出，$\nabla E(\theta)$ 到 U 的正交投影足够小。因此，在定理 3.9中，我们不明确要求向量 w_p 是特征向量，而是假设它们是子空间 U内的正交归一向量。

以下结果证明了k-HiSD (2.2) 的局部渐近稳定性和收敛，前提是前 s 个不稳定方向足够精确，而其余（$k-s$）个不稳定方向允许在子空间$U(t)$ 附近变化。这里，我们考虑依赖于 θ 的轨迹 t，并沿用先前关于 λ_p、u_p、U的记号，它们也依赖于 t。

THEOREM 3.11（局部渐近稳定性与轨迹收敛）。在假设3.3下，对于 $s \leqslant k \leqslant s+m$，假设由 $\theta(t)$ 生成的轨迹 k-*HiSD* (2.2) 对所有 $\mathcal{U}(\mathcal{M},\delta)$ 均保持在 $t \geqslant 0$内。若存在常数 $\alpha < \frac{3\mu^2}{64L^2}$，使得对于任意 $t \geqslant 0$，$v_p(t)$ 中的不稳定方向 k满足 -*HiSD*

$$\left\| \sum_{p=1}^{k} v_p(t)v_p(t)^\top - \sum_{p=1}^{s} u_p(t)u_p(t)^\top - \sum_{p=s+1}^{k} w_p(t)w_p(t)^\top \right\|_2 < \alpha, \tag{3.10}$$

其中 $\{w_p(t)\}_{p=s+1}^{k}$ 是 $U(t)$中的一些正交归一向量，则 $\mathcal{M}$ 是渐近稳定的，并且轨迹 $\theta(t)$收敛于 $\mathcal{M}$中的一个指标为 s 的鞍点。

证明。根据 [3, 定理 4，18], 要证明 $\mathcal{M}$ 是渐近稳定的，只需证明 $G(\theta) = \frac{1}{2}\|\nabla E(\theta)\|_2^2$ 是k-HiSD（2.2）的严格李雅普诺夫函数：

(i) 对于 $\vartheta \in \mathcal{M}$，$G(\vartheta) = 0$。对于 $\vartheta \in \mathcal{U}(\mathcal{M},\delta)\backslash\mathcal{M}$，$G(\vartheta) > 0$。(ii) 存在严格递增函数 $a(x)$ $b(x)$ $a(0) = b(0) = 0$，定义在 $[0,+\infty)$上，使得 $a(\|\vartheta-\vartheta\|_2) \leqslant G(\vartheta) \leqslant b(\|\vartheta-\vartheta\|_2)$ 对于 $\vartheta \in \mathcal{U}(\mathcal{M},\delta)$成立。(iii) 如果 $\theta(0) \notin \mathcal{M}$，则沿着 k-HiSD轨迹，我们有 $\mathrm{d}\, \mathrm{d}_t G(\theta(t)) < 0$，且 $\lim_{t\to\infty} G(\theta(t)) = 0$。

由 $\|\nabla E(\vartheta)\|_2 \geqslant 3\mu\|\vartheta-\vartheta\|_2/8$ 可知，这由备注 3.8给出；由 $\|\nabla E(\vartheta)\|_2 \leqslant L\|\vartheta-\vartheta\|_2$ 可知，这由引理 3.7给出，因此条件 (i) 和 (ii) 对 $a(x) = 9\mu^2x^2/128$ 以及 $b(x) = L^2x^2/2$成立。

接下来，为了证明 (iii)，在以下推导中，为简洁起见，我们省略对 t 的显式依赖。记 $g = \nabla E(\theta)$、$K(\theta)$ 为 (3.7)，并且

$$R = \sum_{p=1}^{k} v_p v_p^\top - \sum_{p=1}^{s} u_p u_p^\top - \sum_{p=s+1}^{k} w_p w_p^\top. \tag{3.11}$$

根据 k-HiSD 中的 θ-动力学，

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\left(I - \sum_{p=1}^{k} 2v_p v_p^\top\right) g = -\left(I - \sum_{p=1}^{s} 2u_p u_p^\top - \sum_{p=s+1}^{k} 2w_p w_p^\top - 2R\right) g, \tag{3.12}$$

我们定义一个常数 $\eta = \frac{3\mu^2}{32L} - 2L\alpha > 0$，并得到

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}G(\theta) &= g^\top \nabla^2 E(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} = K(\theta) + 2g^\top\nabla^2 E(\theta)Rg \leqslant K(\theta) + 2\|g\|_2^2 \left\|\nabla^2 E(\theta)\right\|_2 \|R\|_2 \\ &\leqslant -\left(\frac{3\mu^2}{32} - 2L^2\alpha\right)\|g\|_2\|\theta-\hat{\theta}\|_2 \leqslant -2\eta G(\theta),\end{aligned}$$

其中我们应用了 (3.9)，$\|\nabla^2 E(\theta)\|_2 \leqslant L$，以及 $\|g\|_2 \leqslant L\|\theta-\hat{\theta}\|_2$，由引理 3.7得到。

由格朗沃尔不等式，直接得到 $G(\theta(t)) \leqslant G(\theta(0))\mathrm{e}^{-2\eta t}$，即 $\|\nabla E(\theta(t))\|_2 \leqslant C_0\mathrm{e}^{-\eta t}$，其中 $C_0 = \|\nabla E(\theta(0))\|_2 > 0$。因此，我们有 $\lim_{t\to\infty} G(\theta(t)) = 0$，这就完成了 (iii) 的证明。

最后，由 $\left\|\frac{\mathrm{d}\theta}{t\,\mathrm{d}(t)}\right\|_2 = \|\nabla E(\theta(t))\|_2$，对于任意 $t \geqslant 0, \tau > 0$，我们有

$$\|\theta(t+\tau)-\theta(t)\|_2 = \left\|\int_t^{t+\tau}\frac{\mathrm{d}\theta}{\mathrm{d}t}(s)\mathrm{d}s\right\|_2 \leqslant \int_t^{+\infty} C_0\mathrm{e}^{-\eta s}\mathrm{d}s = \frac{C_0}{\eta}\mathrm{e}^{-\eta t},$$

这意味着 $\{\theta(t)\}_{t\geqslant 0}$ 是一个柯西函数，因为 $t \to +\infty$，所以 $\lim_{t\to\infty}\theta(t)$ 存在。由 $\lim_{t\to\infty}\|\nabla E(\theta(t))\|_2 = 0$ 以及注 3.8，可知 $\lim_{t\to\infty}\theta(t)$ 是 $\mathcal{M}$中的一个临界点。 □

3.3. 离散 HiSD 算法的局部收敛。 在非退化性假设下，文献 [26]从理论上证明，采用固定步长的 HiSD 方法的显式欧拉格式在局部区域内具有线性收敛速率。对于退化鞍点，建立局部收敛性的主要挑战在于其固有的退化性，这使得无法直接应用文献 [26] 中提出的技术，因此需要进行专门分析。为解决这一问题，我们扩展文献 [26] 中的数值分析，证明在适当假设下，该算法能够以线性收敛速率收敛到鞍流形。

首先，我们介绍 $\theta^{(t)} - \hat{}\theta^{(t)}$ 的一个刻画，其证明与 [26,Lemma 3.2] 非常相似。

LEMMA 3.12（单步迭代的误差）。对于满足 $A^{(t)} \in \mathbb{R}^{n\times n}$ 的迭代格式 $\theta^{(t+1)} = \theta^{(t)} - \beta_t A^{(t)} \nabla E(\theta^{(t)})$，有如下恒等式，

$$\begin{aligned} &\theta^{(t+1)} - \hat{\theta}^{(t)} = \left[Q^{(t)} + B^{(t)}\right](\theta^{(t)} - \hat{\theta}^{(t)}) \\ &where\ Q^{(t)} = I - \beta_t A^{(t)} \nabla^2 E(\theta^{(t)}), \\ &B^{(t)} = \beta_t A^{(t)} \left[\nabla^2 E(\theta^{(t)}) - \int_0^1 \nabla^2 E\left(\hat{\theta}^{(t)} + t(\theta^{(t)} - \hat{\theta}^{(t)})\right) \mathrm{d}t\right]. \end{aligned} \tag{3.13}$$

此外，若 $\theta^{(t)} \in \mathcal{U}(\mathcal{M}, \delta)$，则有 $\|B^{(t)}\|_2 \leqslant \frac{1}{2}\beta_t M \|A^{(t)}\|_2 \|\theta^{(t)} - \hat{}\theta^{(t)}\|_2$。

特别地，对于以下 k-HiSD 的显式欧拉格式：

$$\theta^{(t+1)} = \theta^{(t)} - \beta_t \left(I - 2\sum_{p=1}^{k} v_p^{(t)} v_p^{(t)\top}\right) \nabla E(\theta^{(t)}), \tag{3.14}$$

在 $v_p^{(t)\top} v_q^{(t)} = \delta_{pq}$ 的条件下，分别将 $\nabla^2 E(\theta^{(t)})$ 和 $\nabla^2 E(\hat{}\theta^{(t)})$ 的特征值及对应的标准正交特征向量按升序记为 $\left(\lambda_p^{(t)}, u_p^{(t)}\right)_{p=1}^{d}$ 和 $\left(\hat{}\lambda_p^{(t)}, \hat{}u_p^{(t)}\right)_{p=1}^{d}$。对于 $\theta^{(t)} \in \mathcal{U}(\mathcal{M}, \delta)$，我们还可以进一步定义 $U_{(t)} = \mathrm{span}\{u_p^{(t)}\}_{p=s+1}^{s+m}$ 和 $\widehat{U}^{(t)} = \mathrm{span}\{\hat{}u_p^{(t)}\}_{p=s+1}^{s+m}$。现在，我们给出基于精确特征向量的单步数值分析。

THEOREM 3.13（单步迭代的误差估计）。在假设*3.3*下，假设 *k-HiSD* 的指标选为 $s \leqslant k \leqslant s+m$，$\theta^{(t)} \in \mathcal{U}(\mathcal{M}, \delta)$，且 $\beta_t \leqslant 2/(L+\mu)$。若 $v_p^{(t)} = u_p^{(t)}$ 对 $1 \leqslant p \leqslant s$ 成立，且 $v_p^{(t)} \in U^{(t)}$ 对 $s+1 \leqslant p \leqslant k$ 成立，则对于 $r_t = \|\theta^{(t)} - \hat{}\theta^{(t)}\|_2$，下列估计成立：

$$r_{t+1} \leqslant (1 - \mu\beta_t)\, r_t + \left(C + \frac{5M\beta_t}{2}\right) r_t^2. \tag{3.15}$$

证明。由（3.14），取 $A^{(t)} = I - 2\sum_{p=1}^{k} v_p^{(t)} v_p^{(t)\top}$ 并应用引理 3.12 及 $\|A\|_2 = 1$ 可得

$$\theta^{(t+1)} - \hat{\theta}^{(t)} = \left[Q^{(t)} + B^{(t)}\right]\left(\theta^{(t)} - \hat{\theta}^{(t)}\right), \quad \left\|B^{(t)}\right\|_2 \leqslant \frac{1}{2} M \beta_t r_t. \tag{3.16}$$

根据三角不等式，我们可以确定

$$\left\|\theta^{(t+1)} - \hat{\theta}^{(t)}\right\|_2 \leqslant \left\|Q^{(t)}\left(\theta^{(t)} - \hat{\theta}^{(t)}\right)\right\|_2 + \frac{M}{2}\beta_t r_t^2, \tag{3.17}$$

接下来，我们需要估计 $\left\|Q^{(t)}\left(\theta^{(t)}-\hat{\theta}^{(t)}\right)\right\|_2$。

下文中，为简洁起见，我们省略所有下标（t），并将 $\theta-\theta$ 分解为 $\theta-\theta=g_1+g_2$，其中 $g_1=\mathcal{P}_{U^\perp}(\theta-\theta)$ 且 $g_2=\mathcal{P}_U(\theta-\theta)$。由（3.2）以及临界流形 $\mathcal{M}$的非退化性可得 $\theta-\hat{\theta}\in\widehat{U}^\perp$，因此 $\|g_2\|_2$ 可由引理 3.5 估计如下：

$$\|g_2\|_2=\|\mathcal{P}_U(\theta-\hat{\theta})-\mathcal{P}_{\widehat{U}}(\theta-\hat{\theta})\|_2\leqslant\|\theta-\hat{\theta}\|_2\cdot\operatorname{dist}(U,\widehat{U})\leqslant C\|\theta-\hat{\theta}\|_2^2. \tag{3.18}$$

根据 $\nabla^2E(\theta)=\sum_{p=1}^d\lambda_pu_pu_p^\top$的特征值分解，$Q$ 在（3.16）中可表示为：

$$\begin{aligned}Q&=I-\beta_t\left(I-2\sum_{p=1}^k v_pv_p^\top\right)\left(\sum_{p=1}^d\lambda_pu_pu_p^\top\right)\\&=\underbrace{\sum_{p=1}^s(1+\beta_t\lambda_p)u_pu_p^\top+\sum_{p=s+m+1}^d(1-\beta_t\lambda_p)u_pu_p^\top}_{Q_1}+\underbrace{\sum_{p=s+1}^{s+m}(1-\beta_t\lambda_p)u_pu_p^\top}_{Q_2}\\&\quad\underbrace{+2\beta_t\left(\sum_{p=s+1}^k v_pv_p^\top\right)\left(\sum_{p=s+1}^{s+m}\lambda_pu_pu_p^\top\right)}_{Q_3}.\end{aligned} \tag{3.19}$$

由于 $\theta\in\mathcal{U}(\mathcal{M},\delta)$ 以及假设3.3，可得

$$-\lambda_p\in[\mu,L],\ \text{for } 1\leqslant p\leqslant s;\quad \lambda_p\in[\mu,L],\ \text{for } s+m+1\leqslant p\leqslant d,$$

并且由备注3.4可得 $|\lambda_p|\leqslant\mu/4$ 对于 $s+1\leqslant p\leqslant s+m$。因此，根据 Q 在（3.19）中的表达式可得

$$\begin{aligned}&\|Q\mathcal{P}_{U^\perp}\|_2=\|Q_1\|_2\leqslant1-\mu\beta_t,\\&\|Q\|_2\leqslant\|Q_1+Q_2\|_2+\|Q_3\|_2\leqslant1+\frac{\beta_t\mu}{4}+2\beta_t\cdot\frac{\mu}{4}<1+\mu\beta_t.\end{aligned} \tag{3.20}$$

基于上述观察，我们现在可以给出 $\|Q(\theta-\theta)\|_2$ 的最终估计：

$$\begin{aligned}\|Q(\theta-\hat{\theta})\|_2&\leqslant\|Qg_1\|_2+\|Q\|_2\|g_2\|_2=\|Q\mathcal{P}_{U^\perp}(\theta-\hat{\theta})\|_2+\|Q\|_2\|g_2\|_2\\&\leqslant\|Q\mathcal{P}_{U^\perp}\|_2\|\theta-\hat{\theta}\|_2+\|Q\|_2\|g_2\|_2\leqslant(1-\mu\beta_t)\,r_t+(1+\mu\beta_t)\,Cr_t^2.\end{aligned} \tag{3.21}$$

结合（3.17）和（3.21），最终得到

$$r_{t+1}\leqslant\left\|\theta^{(t+1)}-\hat{\theta}^{(t)}\right\|_2\leqslant(1-\mu\beta_t)\,r_t+(1+\mu\beta_t)\,Cr_t^2+\frac{M}{2}\beta_tr_t^2, \tag{3.22}$$

由此，结合 $C=2M/\mu$，证明得证。 □

最后，为了根据定理 3.13建立线性收敛速率，我们引入 [26,引理 3.2]。

LEMMA 3.14. 设 $\{r_t\}_{t\geqslant0}$ 是一个满足以下条件的非负级数

$$r_{t+1}\leqslant(1-q)r_t+cr_t^2,\quad t\geq0,\quad q\in(0,1),\quad c>0.$$

(a) If $r_t < \frac{q}{c}$ *for some* $t \geqslant 0$, *then* $r_{t+1} < r_t < \frac{q}{c}$;

(b) If $r_0 < \frac{q}{c}$, *then* $r_{t+1} \leqslant \left(\frac{1}{1+q}\right)^{t+1} \frac{qr_0}{q - cr_0}$ *for all* $t \geqslant 0$.

基于 定理 3.13 和 引理 3.14，我们随后给出收敛速率的主要定理。

THEOREM 3.15（线性收敛速率）。在假设 *3.3* 下，假设 *k-HiSD* 的指标选为 $s \leqslant k \leqslant s+m$，其中 $v_p^{(t)} = u_p^{(t)}$对于 $1 \leqslant p \leqslant s$ 且$v_p^{(t)} \in U^{(t)}$ 对于 $s+1 \leqslant p \leqslant k$。若固定步长选为 $\beta_t \equiv \beta \leqslant 2/(L+\mu)$ 对任意 $t \geqslant 0$，且初始点 $\theta^{(0)} \in \mathcal{U}(\mathcal{M}, \delta)$ 满足

$$r_0 = \left\|\theta^{(0)} - \hat{\theta}^{(0)}\right\|_2 < \hat{r} = \frac{2\mu^2\beta}{(5\mu\beta+4)M}, \tag{3.23}$$

则 $r_t = \|\theta^{(t)} - \hat{\theta}^{(t)}\|_2$ 单调递减，并在 $t \to \infty$ 时收敛至 0，且至少具有线性收敛速率：

$$r_t \leqslant (1+\mu\beta)^{-t} \frac{\hat{r} r_0}{\hat{r} - r_0}. \tag{3.24}$$

证明。由 定理 3.13 可知，$t = 0$，由于 $r_0 < \hat{r}$，有 $r_1 < r_0 < \hat{r}$，应用 引理 3.14 (a)，取 $q = \mu\beta \in (0, 1)$，且 $c = q/\hat{r} = C + \frac{5M\beta}{2} > 0$。归纳可得 $r_{t+1} < r_t < \hat{r}$对任意 $t \geqslant 0$ 均成立。然后结合 定理 3.13 对所有 $t \geqslant 0$ 以及引理 3.14 (b) 即可完成证明。 □

注记 3.16. 通过应用 $\{v_p^{(t)}\}_{p=1}^k$ 中的技术，可以将这一结果推广到满足特定精度界的不精确特征向量情形Theo-rem3.11 和 [26]。

3.4. HiSD 的梯度对齐。 与梯度下降类似，[10], HiSD 算法也表现出梯度 $\nabla E(\theta)$ 与海森矩阵 $\nabla^2 E(\theta)$ 的某个特定特征向量对齐的倾向。这种现象称为“梯度对齐”，它阐明了梯度的渐近方向，并使我们能够更精确地刻画算法的收敛速率。自然地，这一性质对非退化鞍点也成立。然而，对于鞍流形，零特征值的存在给严格证明梯度对齐带来了重大挑战。为应对这些挑战，我们现在引入额外的假设和记号，以便对退化情形下的梯度对齐进行精确分析。

ASSUMPTION 3.17. 存在正常数 $\mu_1, \mu_2 > 0$，使得对于所有充分大的 t

ffi y g ,

(i) $\mu < \min\left(|\lambda_s^{(t)}|, |\lambda_{s+m+1}^{(t)}|\right) < \mu_1 < \mu_2 < \max\left(|\lambda_s^{(t)}|, |\lambda_{s+m+1}^{(t)}|\right) < L.$

(ii) $\mu < \min\left(|\hat{\lambda}_s^{(t)}|, |\hat{\lambda}_{s+m+1}^{(t)}|\right) < \mu_1 < \mu_2 < \max\left(|\hat{\lambda}_s^{(t)}|, |\hat{\lambda}_{s+m+1}^{(t)}|\right) < L.$

令 $\lambda^{(t)}{}_{\min}$ 表示 $\lambda_s^{(t)}$和 $\lambda_{s+m+1}^{(t)}$中绝对值较小的特征值，并令 $u^{(t)}$ 表示 $\nabla^2 E(\theta^{(t)})$对应的单位特征向量。事实上，$u^{(t)}$ 正是在迭代过程中梯度方向趋于对齐的特征向量。类似地，对于海森矩阵 $\nabla^2 E(\hat{\theta}^{(t)})$，我们定义$\hat{\lambda}^{(t)}{}_{\min}$ 和$\hat{u}^{(t)}$。为方便起见，我们用 $\mathcal{P}_u$ 和 $\mathcal{P}_{u^\perp}$ 分别表示投影到由向量 u 张成的子空间及其正交补上的投影算子。

为刻画梯度对齐，我们将 $y_t = \left\|\mathcal{P}_{u^{(t)}} \nabla E(\theta^{(t)})\right\|_2$ 定义为 $\nabla E(\theta^{(t)})$在所选特征向量方向上的分量大小，并

$z_t = \left\|\mathcal{P}_{u^{(t)\perp}}\nabla E(\theta^{(t)})\right\|_2 = \sqrt{\|\nabla E(\theta^{(t)})\|_2^2 - y_t^2}$。为避免边界情形中过多的技术细节，我们作出如下假设。

ASSUMPTION 3.18. 存在一个常数 $\mu_0 > 0$，使得对于所有充分大的 t，$y_t \geqslant \mu_0 r_t$，这里 $r_t = \|\theta^{(t)} - \hat{}\theta^{(t)}\|_2$。

梯度对齐的数学描述为 $y_t/\|\nabla E(\theta^{(t)})\|_2 \to 1$，或者等价地，$z_t/y_t \to 0$，下面给出定理，以说明 $\nabla E(\theta^{(t)})$ 渐近地与方向 $u^{(t)}$ 对齐:

THEOREM 3.19（梯度对齐）。在假设 *3.3*、*3.17* 和 *3.18*下，假设 k-*HiSD* 的指标选为 $s \leqslant k \leqslant s+m$，其中 $v_p^{(t)} = u_p^{(t)}$对于$1 \leqslant p \leqslant s$，且 $v_p^{(t)} \in U^{(t)}$ 对于 $s+1 \leqslant p \leqslant k$。若固定步长选为$\beta_t \equiv \beta \leqslant 1/L$ 对于任意 $t \geqslant 0$，且初始点 $\theta^{(0)} \in \mathcal{U}(\mathcal{M}, \delta)$ 满足

$$r_0 < \hat{r} = \frac{2\mu^2\beta}{(5\mu\beta+4)M}, \tag{3.25}$$

则 $z_t/y_t \to 0$ *as* $t \to \infty$。

证明。根据假设 3.18，对于所有充分大的 t，$u^{(t)}$ 和 $\hat{}u^{(t)}$ 均为海森矩阵 p 的第 $\nabla^2 E(\theta^{(t)})$ 个最小特征值和 $\nabla^2 E(\hat{}\theta^{(t)})$ 对应的特征向量，且具有相同的 $p = s$ 或 $s+m+1$。采用与引理 3.5 的证明类似的论证，可得

$$\|\mathcal{P}_{u^{(t)}} - \mathcal{P}_{\hat{u}^{(t)}}\|_2 \leqslant C\left\|\theta^{(t)} - \hat{\theta}^{(t)}\right\|_2, \quad \|\mathcal{P}_{u^{(t+1)}} - \mathcal{P}_{\hat{u}^{(t)}}\|_2 \leqslant C\left\|\theta^{(t+1)} - \hat{\theta}^{(t)}\right\|_2.$$

由于定理 3.15 的假设成立，令 $r_t = \|\theta^{(t)} - \hat{}\theta^{(t)}\|_2$ 并应用 (3.22)，得到 $\|\theta^{(t+1)} - \hat{}\theta^{(t)}\|_2 < r_t$。因此，

$$\|\mathcal{P}_{u^{(t+1)}} - \mathcal{P}_{u^{(t)}}\|_2 \leqslant \|\mathcal{P}_{u^{(t+1)}} - \mathcal{P}_{\hat{u}^{(t)}}\|_2 + \|\mathcal{P}_{u^{(t)}} - \mathcal{P}_{\hat{u}^{(t)}}\|_2 \leqslant 2Cr_t. \tag{3.26}$$

在 引理 3.6 处应用 $\theta^{(t)}$ 和 $\theta^{(t+1)}$，可得

$$\begin{aligned} &\nabla E(\theta^{(t+1)}) = \nabla E(\theta^{(t)}) + \left(\nabla^2 E(\theta^{(t)}) + D^{(t)}\right)\left(\theta^{(t+1)} - \theta^{(t)}\right), \\ &\left\|D^{(t)}\right\|_2 \leqslant \frac{M}{2}\left\|\theta^{(t+1)} - \theta^{(t)}\right\|_2 \leqslant Mr_t, \quad \left\|\theta^{(t+1)} - \theta^{(t)}\right\|_2 \leqslant 2r_t. \end{aligned} \tag{3.27}$$

使用 k-HiSD 的显式欧拉格式:

$$\theta^{(t+1)} = \theta^{(t)} - \beta R^{(t)}\nabla E(\theta^{(t)}), \quad \text{where } R^{(t)} = I - 2\sum_{p=1}^{k} v_p^{(t)} v_p^{(t)\top}, \tag{3.28}$$

我们有

$$\nabla E(\theta^{(t+1)}) = \left(I - \beta\nabla^2 E(\theta^{(t)})R^{(t)}\right)\nabla E(\theta^{(t)}) + D^{(t)}\left(\theta^{(t+1)} - \theta^{(t)}\right). \tag{3.29}$$

通过利用 $\nabla^2 E(\theta^{(t)})$的特征分解直接计算，我们得到

$$\begin{aligned} F^{(t)} &:= I - \beta\nabla^2 E(\theta^{(t)})R^{(t)} \\ &= \sum_{p=1}^{s}\left(1 + \beta\lambda_p^{(t)}\right)u_p^{(t)}u_p^{(t)\top} + \sum_{p=s+1}^{d}\left(1 - \beta\lambda_p^{(t)}\right)u_p^{(t)}u_p^{(t)\top} \\ &\quad + 2\beta\left(\sum_{p=s+1}^{s+m}\lambda_p u_p^{(t)}u_p^{(t)\top}\right)\left(\sum_{p=s+1}^{k} v_p^{(t)}v_p^{(t)\top}\right). \end{aligned} \tag{3.30}$$

以及 $\|F^{(t)}\|_2 \leqslant 1+\beta \max_p\{|\lambda_p^{(t)}|\} \leqslant 2$。

将（3.29）代入 z_{t+1}的定义，得到

$$\begin{aligned} z_{t+1} &= \|\mathcal{P}_{u^{(t+1)\perp}} \nabla E(\theta^{(t+1)})\|_2 \\ &\leqslant \|\mathcal{P}_{u^{(t)\perp}} \nabla E(\theta^{(t+1)})\|_2 + \|\mathcal{P}_{u^{(t+1)\perp}} - \mathcal{P}_{u^{(t)\perp}}\|_2 \|\nabla E(\theta^{(t+1)})\|_2 \\ &\leqslant \|\mathcal{P}_{u^{(t)\perp}} F^{(t)} \nabla E(\theta^{(t)})\|_2 + \|D^{(t)}(\theta^{(t+1)} - \theta^{(t)})\|_2 + 2CLr_t^2 \\ &\leqslant \|\mathcal{P}_{u^{(t)\perp}} F^{(t)} \nabla E(\theta^{(t)})\|_2 + (2M + 2CL) r_t^2. \end{aligned} \tag{3.31}$$

这里用到（3.26）、（3.27）以及 $\|\nabla E(\theta^{(t+1)})\|_2 \leqslant Lr_{t+1} \leqslant Lr_t$，它们来自引理 3.7。根据引理 3.7，我们将 $\nabla E(\theta^{(t)}) = x^{(t)} + w^{(t)}$ 写成 $x^{(t)} = \mathcal{P}_{U^{(t)\perp}} \nabla E(\theta^{(t)})$ 和 $w^{(t)} = \mathcal{P}_{U^{(t)}} \nabla E(\theta^{(t)})$，使其满足 $\|w^{(t)}\|_2 \leqslant Mr_t^2$。此外，我们注意到 $\mathcal{P}_{u^{(t)\perp}}$与 $F^{(t)}$ 和 $\mathcal{P}_{U^{(t)\perp}}$ 可交换。因此，我们得到以下估计：

$$\begin{aligned} \|\mathcal{P}_{u^{(t)\perp}} F^{(t)} \nabla E(\theta^{(t)})\|_2 &\leqslant \|F^{(t)} \mathcal{P}_{u^{(t)\perp}} x^{(t)}\|_2 + \|F^{(t)}\|_2 \|w^{(t)}\|_2 \\ &\leqslant \|F^{(t)} \mathcal{P}_{U^{(t)\perp}} \mathcal{P}_{u^{(t)\perp}}\|_2 \cdot z_t + 2\|w^{(t)}\|_2 \leqslant (1-\beta\mu_2) z_t + 2Mr_t^2. \end{aligned} \tag{3.32}$$

最后一个不等式由（3.30）和 $\max\left(|\lambda_s^{(t)}|, |\lambda_{s+m}^{(t)}|\right) > \mu_2$推出，它们蕴含

$$\|F^{(t)} \mathcal{P}_{U^{(t)\perp}} \mathcal{P}_{u^{(t)\perp}}\|_2 = \|F^{(t)}(I - \mathcal{P}_{U^{(t)}} - \mathcal{P}_{u^{(t)}})\|_2 \leqslant 1 - \beta\mu_2. \tag{3.33}$$

结合（3.31）和（3.32），得到

$$z_{t+1} \leqslant (1-\beta\mu_2) z_t + \overline{C} r_t^2, \quad \text{where } \overline{C} = 4M + 2CL. \tag{3.34}$$

同理，对于 y_{t+1}，得到

$$\begin{aligned} y_{t+1} &= \|\mathcal{P}_{u^{(t+1)}} \nabla E(\theta^{(t+1)})\|_2 \\ &\geqslant \|\mathcal{P}_{u^{(t)}} \nabla E(\theta^{(t+1)})\|_2 - \|\mathcal{P}_{u^{(t+1)}} - \mathcal{P}_{u^{(t)}}\|_2 \|\nabla E(\theta^{(t+1)})\|_2 \\ &\geqslant \|\mathcal{P}_{u^{(t)}} F^{(t)} \nabla E(\theta^{(t)})\|_2 - \|D^{(t)}(\theta^{(t+1)} - \theta^{(t)})\|_2 - 2CLr_t^2 \\ &\geqslant \|F^{(t)} \mathcal{P}_{u^{(t)}} \mathcal{P}_{u^{(t)}} \nabla E(\theta^{(t)})\|_2 - (2M + 2CL) r_t^2 \\ &\geqslant \|(1 - \beta|\lambda_{\min}^{(t)}|) \mathcal{P}_{u^{(t)}} \nabla E(\theta^{(t)})\|_2 - \overline{C} r_t^2 \geqslant (1-\beta\mu_1) y_t - \overline{C} r_t^2. \end{aligned} \tag{3.35}$$

由 $r_t \leqslant y_t/\mu_0$ 假设3.18，可得 $y_{t+1} \geqslant (1-\beta\mu_1)\, y_t - Cr_t y_t/\mu_0$。定理 3.15 意味着 $r_t \leqslant C_1 \xi^t$，其中某个常数为 $C_1 > 0$，其中 $\xi = (1+\mu\beta)^{-1}$。因此，存在一个整数 $T > 0$，使得对于所有 $t \geqslant T$，假设3.17和3.18 均成立，且条件 $Cr_t/\mu_0 < (1-\beta\mu_1)/2$ 成立。因此，

$$y_{t+1} \geqslant (1-\beta\mu_1) y_t - \overline{C} r_t y_t/\mu_0 > (1-\beta\mu_1) y_t/2 > 0. \tag{3.36}$$

结合（3.34）和（3.35），得到

$$\frac{z_{t+1}}{y_{t+1}} \leqslant \frac{(1-\beta\mu_2) z_t + \overline{C} r_t^2}{(1-\beta\mu_1) y_t - \overline{C} r_t^2} = \rho \frac{z_t}{y_t} + \frac{(1-\beta\mu_2) z_t + (1-\beta\mu_1) y_t}{(1-\beta\mu_1) y_t ((1-\beta\mu_1) y_t - \overline{C} r_t^2)} \overline{C} r_t^2, \tag{3.37}$$

其中 $0 < \beta\mu_1 < \beta\mu_2 = \mu_2/L < 1$，且 $\rho = (1-\beta\mu_2)/(1-\beta\mu_1) \in (0,1)$。利用 $r_t \leqslant y_t/\mu_0$，$z_t \leqslant Lr_t \leqslant Ly_t/\mu_0$ 根据 引理 3.7和（3.36），可得

$$\frac{z_{t+1}}{y_{t+1}} - \rho \frac{z_t}{y_t} \leqslant \frac{(1-\beta\mu_2) L y_t/\mu_0 + (1-\beta\mu_1) y_t}{(1-\beta\mu_1)^2 y_t^2/2} \overline{C} r_t y_t/\mu_0 := \tilde{C} r_t, \tag{3.38}$$

这里 $\tilde{C} = 2\overline{C}\mu_0^{-2}(1-\beta\mu_1)^{-2}[(1-\beta\mu_2)L + (1-\beta\mu_1)\mu_0]$。

最后，对于 $t \geqslant T$，利用 $r_t \leqslant C_1\xi^t$，我们有

$$\frac{z_t}{y_t} \leqslant \rho^{t-T}\frac{z_T}{y_T} + \tilde{C}C_1\sum_{\tau=T}^{t-1}\rho^{t-1-\tau}\xi^\tau \leqslant \rho^{t-T}\frac{z_T}{y_T} + \tilde{C}C_1 t\left(\max\{\rho,\xi\}\right)^{t-1} \to 0, \tag{3.39}$$

当 $t \to \infty$ 时，结论成立，从而完成证明。 □

注3.20. 在梯度对齐条件 $\mathcal{P}_{u^\perp}\nabla E(\theta) \approx 0$ 下，$\|\nabla E(\theta)\|_2$ 的后期衰减率可渐近估计如下：

$$\frac{\mathrm{d}}{\mathrm{d}t}G(\theta) = -(\nabla E(\theta))^\top \nabla^2 E(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} \approx -|\lambda_{\min}| \cdot \|\nabla E(\theta)\|_2^2,$$

或者等价地，$^{d}{}_{\mathrm{d}t}\|\nabla E(\theta)\|_2 \approx -|\lambda_{\min}| \cdot \|\nabla E(\theta)\|_2$。因此，$\|\nabla E(\theta)\|_2$ 以速率 $-|\lambda_{\min}|$近似指数衰减，其中 $\lambda_{\min}$ 是 $\nabla^2 E(\theta)$ 的最小非零特征值，排除第 $(s+1)$ 至第 $(s+m)$ 个最小特征值。

在离散迭代中，通过将引理 3.6 应用于 $\hat{}\theta^{(t)}$ 和 $\theta^{(t)}$，我们得到

$$\nabla E(\theta^{(t)}) = \nabla^2 E(\theta^{(t)})(\theta^{(t)} - \hat{\theta}^{(t)}) + O(r_t^2). \tag{3.40}$$

由于 $\mathcal{P}_{U^{(t)}}(\theta^{(t)} - \hat{}\theta^{(t)}) = O(r_t^2)$，因此有

$$\theta^{(t)} - \hat{\theta}^{(t)} = \left(\lambda_{\min}^{(t)}\right)^{-1}\nabla E(\theta^{(t)}) + o(r_t), \quad \text{as } t \to \infty, \tag{3.41}$$

其中我们用到了梯度对齐结果（3.39）。实际中，可以将$\|\nabla E(\theta)\|_2$ 作为终止条件，用于监测数值算法的收敛情况。误差 $\|\theta^{(t)} - \hat{}\theta^{(t)}\|$ 可在迭代结束时利用（3.41）进行估计。

4. 数值实验。对于具有退化鞍点的能量景观，过参数化神经网络的损失景观呈现出极端退化性和非凸性的显著特征，使其有别于具有明确对称性的简单模型。损失函数的大多数临界点都是退化的，且该景观中既包含多个局部极小值，也包含各种各样的鞍点。这种复杂性给神经网络优化带来了重大挑战 [7]。此外，近期研究表明，采用较小初始值的基于梯度的优化方法可能会引发“鞍点到鞍点”的训练机制[30]。因此，理解鞍点在深度学习中的影响至关重要。对鞍点进行计算分析，有助于深入了解神经网络的优化动力学和训练行为。

在本章节中，我们给出神经网络的一些数值结果，以展示 HiSD 在计算退化鞍点方面的性能。我们采用一个包含六个神经元的两层神经网络实验设置：

$$f_\theta : \mathbb{R} \to \mathbb{R}, \quad f_\theta(x) = \sum_{i=1}^{6} a_i \tanh(w_i x + b_i), \tag{4.1}$$

作为假设空间。真实函数 $f^* = \sum_{i=1,\ 2} a_i^* \tanh(w_i^* x + b_i^*)$ 有两个神经元。在我们的实验中，训练数据集为 $S = \{(x_i,\ y_i)\}_{i=1}^n$，其中 $n = 15$，且 $x_i \sim \mathcal{N}(0,\ 1)$ 独立同分布，$y_i = f^*(x_i)$。作为能量函数的经验风险定义为

$$E(\theta) = \frac{1}{2n}\sum_{i=1}^{n}\left(f_\theta(x_i) - y_i\right)^2. \tag{4.2}$$

文献 [50] 指出，在神经网络训练过程中，参数可能趋近于某些鞍流形，这些鞍流形对于理解训练动力学和损失景观都发挥着重要作用。这类鞍流形可以通过神经元拆分构造。为确保数值实验的精确性，我们采用这种方法构造鞍流形 $\mathcal{M}$。在我们的数值实验中，所研究的鞍流形 $\mathcal{M}$是一个五维线性流形，其海森矩阵具有五个负特征值、五个零特征值和八个正特征值，即，指标为 $s = 5$，零化度为 $m = 5$（采用本文的记号）。初始点 $\theta^{(0)}$ 选取为固定鞍点 $\theta^* \in \mathcal{M}$的随机扰动，并且在所有实验中始终使用同一个初始点。

4.1. 不同指标下的 HiSD。我们的理论结果表明，k- HiSD 方法在 $s \leqslant k \leqslant s+m$下表现出收敛性和类似的定性行为。因此，我们采用 k-HiSD 来处理 $5 \leqslant k \leqslant 10$，使用固定步长 $\beta = 0.1$，并采用 LOBPCG 作为特征值求解器。图 4.1展示了不同 k下迭代过程中梯度的 2-范数、到鞍流形的距离以及梯度对齐程度。值得注意的是，由于投影 $\hat{\theta}^{(t)} = \mathcal{P}_{\mathcal{M}}\theta^{(t)}$ 到线性流形 $\mathcal{M}$ 可以显式计算，因此我们可以直接计算 $\|\theta^{(t)} - \hat{\theta}^{(t)}\|_2$。

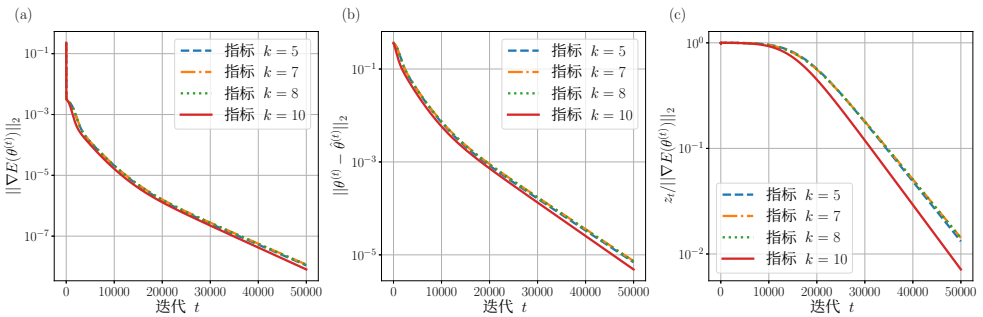


图 4.1：不同指标下 HiSD 的 (a) 梯度范数 $\|\nabla E(\theta^{(t)})\|_2$、(b) 到鞍流形的距离$\|\theta^{(t)} - \hat{\theta}^{(t)}\|_2$ 以及 (c) 梯度对齐程度 $z_t/\|\nabla E(\theta^{(t)})\|_2$ 相对于迭代次数 t的变化。

图 4.1 说明，对于不同的指标，$\|\nabla E(\theta^{(t)})\|_2$ 和 $\|\theta^{(t)} - \hat{\theta}^{(t)}\|_2$ 均以相近的速率收敛至零 k。此外，比值 $z_t/\|\nabla E(\theta^{(t)})\|_2$表示梯度中与对应于 $\lambda_{\min}$的特征向量正交的分量，该比值也收敛至零，这与梯度对齐趋势一致。

4.2. HiSD 的不同加速方法。正如前面的数值结果所示，鞍点退化给搜索算法带来了重大挑战，往往导致收敛速率缓慢，并需要大量迭代。为应对这一问题，人们开发了 HiSD 的动量加速变体，例如采用 Heavy-Ball 加速 [25] 和 Nesterov 加速 [22]的 HiSD，有效提升了收敛速度。在数值实验中，我们比较了显式欧拉格式、动量参数为 $\gamma = 0.6$ 和 0.9 的 Heavy-Ball 加速，以及 Nesterov 加速的性能。所有方法均采用固定步长 $\beta = 0.1$，并使用 LOBPCG 作为特征值求解器。对于 Nesterov 加速，每 500 次迭代实施一次重启策略。

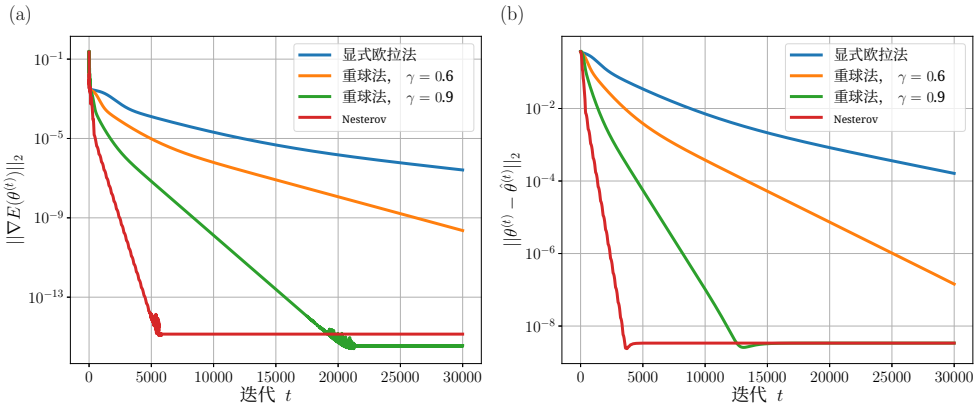


图 4.2：不同加速方法下，(a) 梯度范数 $\|\nabla E(\theta^{(t)})\|_2$ 和 (b) 到鞍流形 $\|\theta^{(t)} - \hat{\theta}^{(t)}\|_2$ 的距离随迭代次数 t 的变化图。

如图 4.2所示，与显式欧拉格式相比，这些加速方法显著提高了收敛速度。Nesterov 加速表现最佳，$\|\nabla E(\theta^{(t)})\|_2$ 在约 5,000 步内降至 10^{-15} 。同样，动量参数为 $\gamma = 0.9$ 的 Heavy-Ball 方法在约 20,000 步内将梯度范数降至 10^{-15} ，而动量参数为 $\gamma = 0.6$ 的方法收敛速度较慢。到鞍流形的距离 $\|\theta^{(t)} - \hat{\theta}^{(t)}\|_2$表现出类似的收敛行为。

5. 结语。本文利用 Morse–Bott 函数，对临界流形上退化鞍点的 HiSD 进行了严格分析。针对海森矩阵中零特征值带来的挑战，我们证明了连续高指标鞍点动力学的局部收敛性，并且尤为重要的是，确立了离散高指标鞍点动力学算法在临界流形附近的线性收敛速率。此外，我们为梯度对齐趋势提供了理论依据，并阐明了实际计算中指标选择的灵活性。神经网络上的数值实验验证了我们的分析，表明即使在高度退化的景观中，HiSD 仍能保持快速收敛。这些结果弥合了 HiSD 在退化情形下的经验成功与其数学理解之间的差距。

我们的研究结果证实，HiSD 可用于搜索退化鞍点，为其今后在退化问题中的应用提供了坚实的理论基础，并进一步彰显了 HiSD 作为一种高效鞍点搜索方法的广泛适用性。鉴于实际问题中的能量函数往往存在退化，开发更高效、更具针对性的算法仍是当务之急。例如，近期研究通过筛除零特征向量改进了 HiSD，从而加快收敛速度 [20]。展望未来，仍需持续研究如何进一步改进算法，并设计稳健的算法，以便在科学应用中高效搜索退化鞍点。

[1] A. Banyaga and D. Hurtubise, *Morse-Bott homology*, 美国数学会汇刊, 362 (2010), 第 3997–4043 页, https://doi.org/10.1090/S0002-9947-10-05073-7. [2] A. Banyaga and D. E. Hurtubise, *A proof of the Morse-Bott Lemma*, 数学展望, 22 (2004), 第 365–373 页, https://doi.org/10.1016/S0723-0869(04)80014-8. [3] N. P. Bhatia and G. P. Szegö, *Stability Theory of Dynamical Systems*, 施普林格出版社, 柏林、海德堡, 2002, https://link.springer.com/book/9783540427483. [4] R. Bott, *Nondegenerate critical manifolds*, 数学年刊, 60 (1954), 第 248–261 页, https://doi. org/10.2307/1969631. [5] X. Cheng, L. Lin, W. E, P. Zhang, and A.-C. Shi, *Nucleation of ordered phases in block copolymers* 104 2010 148301, 10 1103 , 《物理评论快报》, ( ), 第 https://doi.org/ 页。/PhysRevLett. 104.148301.[6] H. Daneshmand, J. Kohler, A. Lucchi, and T. Hofmann, *Escaping saddles with stochastic gradients* 80 , 载于机器学习国际会议论文集, J. Dy 和 A. Krause 编, 第 卷《机器学习研究论文集》, PMLR, 2018, 第 1155–1164 页, https://proceedings.mlr. press/v80/daneshmand18a.html. [7] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, *Identifying and attacking the saddle point problem in high-dimensional non-convex optimization*, 载于神经信息处理系统进展, 第 27 卷, Curran Associates, Inc., 2014, 第 2933–2941 页, https://dl.acm.org/doi/10.5555/2969033.2969154. [8] W. E, W. Ren, and E. Vanden-Eijnden, *String method for the study of rare events*, 物理评论 B, 66 (2002), 第 052301 页, https://doi.org/10.1103/PhysRevB.66.052301. [9] W. E and X. Zhou, *The gentlest ascent dynamics*, 非线性, 24 (2011), 第 1831 页, https: //doi.org/10.1088/0951-7715/24/6/008. [10] G. E. Forsythe, *On the asymptotic directions of the s-dimensional optimum gradient method*, 数值数学, 11 (1968), 第 57–76 页, https://doi.org/10.1007/BF02165472. [11] W. Gao, J. Leng, and X. Zhou, *An iterative minimization formulation for saddle point search*, SIAM 数值分析杂志, 53 (2015), 第 1786–1805 页, https://doi.org/10.1137/130930339. [12] J. Goldstone, A. Salam, and S. Weinberg, *Broken symmetries*, 物理评论, 127 (1962), 第 965–970 页, https://doi.org/10.1103/PhysRev.127.965. [13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 约翰斯·霍普金斯大学出版社, 宾夕法尼亚州费城, 第 4 版, 2013, https://epubs.siam.org/doi/abs/10.1137/1. 9781421407944.[14] Y. Han, J. Yin, P. Zhang, A. Majumdar, and L. Zhang, *Solution landscape of a reduced Landau–de Gennes model on a hexagon* 34 2021 2048, , 非线性, ( ), 第 https://doi.org/ 10.1088/1361-6544/abc5d4. [15] G. Henkelman and H. Jónsson, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points* 113 2000 9978 , 化学物理学杂志, ( ), 第 – 9985 页, https://doi.org/10.1063/1.1323224. [16] G. Henkelman and H. Jónsson, *A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives* 111 1999 7010 7022, , 化学物理学杂志, ( ), 第 – 页, https://doi.org/10.1063/1.480097. [17] A. Heyden, A. T. Bell, and F. J. Keil, *Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method* 123 2005 224101, 10 1063 , 化学物理学杂志, ( ), 第 https://doi.org/ 页。/ 1.2104507.[18] M. W. Hirsch, *Differential Topology*, 第 33 卷, Springer Science & Business Media, 纽约, 2012, https://doi.org/10.1007/978-1-4684-9449-5. [19] L. D. Jacobson, A. D. Bochevarov, M. A. Watson, T. F. Hughes, D. Rinaldo, S. Ehrlich, T. B. Steinbrecher S. Vaitheeswaran D. M. Philipp M. D. Halls et al. *Automated transition state search and its application to diverse types of organic reactions* , , , , , , 化学理论与计算杂志, 13 (2017), 第 5780–5797 页, https://doi.org/10.1021/acs.jctc.7b00764.[20] K. Jiang, L. Zhang, X. Zheng, and T. Zhou, *Natispace-preserving high-index saddle dynamics method for degenerate multiple solution problems* 2025 2510 , , https://arxiv.org/abs/ . 24292.[21] A. V. Knyazev, *Toward the preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method* 23 2001 517 541, , SIAM 科学计算杂志, ( ), 第 – 页 https://doi. org/10.1137/S1064827500366124.[22] Y. Liu, L. Zhang, and J. Zhao, *Neural-network-based high-index saddle dynamics method for searching saddle points and solution landscape* 2024 2411 16200 , , https://arxiv.org/abs/ . .

[23] D. Longsine and S. McCormick, *Simultaneous Rayleigh-quotient minimization methods for $Ax = \lambda Bx$* 34 1980 195 234, 10 1016 , 线性代数及其应用, ( ), 第 – 页 https://doi.org/ . / 0024-3795(80)90166-4.

[24] T. Luo, Z.-Q. J. Xu, Z. Ma, and Y. Zhang, *Phase diagram for two-layer ReLU neural networks at infinite-width limit* 22 2021 1 47, , 《机器学习研究期刊》, (), 第–页, http://jmlr. org/papers/v22/20-1123.html。[25] Y. Luo, L. Zhang, and X. Zheng, *Accelerated high-index saddle dynamics method for searching high-index saddle points* 102 2025 31, 10 1007 , 《科学计算杂志》, (), 第页, https://doi.org/ . / s10915-024-02760-6。[26] Y. Luo, X. Zheng, X. Cheng, and L. Zhang, *Convergence analysis of discrete high-index saddle dynamics* 60 2022 2731 2750, 10 , SIAM 数值分析杂志, (), 第–页, https://doi.org/ . 1137/22M1487965。[27] J. Milnor, *Morse Theory*, 普林斯顿大学出版社, 普林斯顿, 1963, https://doi.org/10. 1515/9781400881802。[28] Y. Nesterov, *A method for solving the convex programming problem with convergence rate* $O(1/k^2)$, 苏联科学院院刊, 269（1983）, 第543–547页, https://api.semanticscholar. org/CorpusID:145918791。[29] L. Nicolaescu, *An Invitation to Morse Theory*, 施普林格, 纽约, 2007, https://doi. org/10.1007/978-0-387-49510-1。[30] S. Pesme and N. Flammarion, *Saddle-to-saddle dynamics in diagonal linear networks*, 统计力学杂志, 2024（2024）, 第104016页, https://doi.org/10.1088/1742-5468/ad65e3。[31] B. Polyak, *Some methods of speeding up the convergence of iteration methods*, 苏联计算数学与数学物理, 4（1964）, 第1–17页, https://doi.org/10.1016/0041-5553(64)90137-5。[32] L. Sagun, U. Evci, V. U. Guney, Y. Dauphin, and L. Bottou, *Empirical analysis of the hessian of over-parametrized neural networks* 2018 1706 04454 , , https://arxiv.org/abs/。。[33] A. Samanta, M. E. Tuckerman, T.-Q. Yu, and W. E, *Microscopic mechanisms of equilibrium melting of a solid* 346 2014 729 732, 10 1126 , 《科学》, (), 第–页, https://doi.org/ . /science. 1253810。[34] G. Stewart and J. Sun, *Matrix Perturbation Theory*, 爱思唯尔科学出版社, 阿姆斯特丹, 1990, https://api.semanticscholar.org/CorpusID:265829890。[35] H. Su, H. Wang, L. Zhang, J. Zhao, and X. Zheng, *Improved high-index saddle dynamics for finding saddle points and solution landscape* 63 2025 1757 , SIAM 数值分析杂志, (), 第–1775页, https://doi.org/doi.org/10.1137/25M173212X。[36] Y. Wang and J. Li, *Phase field modeling of defects and deformation*, 《材料学报》, 58（2010）, 第1212–1235页, https://doi.org/10.1016/j.actamat.2009.10.041。[37] S. Wu, B. Yu, Y. Tu, and L. Zhang, *Solution landscape of reaction-diffusion systems reveals a nonlinear mechanism and spatial robustness of pattern formation* 2025 , 《基础研究》, (), https://doi.org/10.1016/j.fmre.2025.10.006。[38] J. Yin, Z. Huang, X. Cai, Q. Du, and L. Zhang, *Revealing excited states of rotational Bose-Einstein condensates* 5 2024 100546, 10 1016 , 《创新》, (), 第页, https://doi.org/ . /j.xinn. 2023.100546。[39] J. Yin, Z. Huang, and L. Zhang, *Constrained high-index saddle dynamics for the solution landscape with equality constraints* 91 2022 62, 10 , 《科学计算杂志》, (), 第页, https://doi.org/ . 1007/s10915-022-01838-3。[40] J. Yin, K. Jiang, A. C. Shi, P. Zhang, and L. Zhang, *Transition pathways connecting crystals and quasicrystals* 118 2021 2106230118, , 美国国家科学院院刊, (), 第e页, https: //doi.org/10.1073/pnas.2106230118。[41] J. Yin, Y. Wang, J. Z. Y. Chen, P. Zhang, and L. Zhang, *Construction of a pathway map on a complicated energy landscape* 124 2020 090601, , 《物理评论快报》, (), 第页, https: //doi.org/10.1103/PhysRevLett.124.090601。[42] J. Yin, B. Yu, and L. Zhang, *Searching the solution landscape by generalized high-index saddle dynamics* 64 2021 1801 1816, 10 1007 , 《中国科学: 数学》, (), 第–页, https://doi.org/ . / s11425-020-1737-1。[43] J. Yin, L. Zhang, and P. Zhang, *High-index optimization-based shrinking dimer method for finding high-index saddle points* 41 2019 3576 3595, , SIAM 科学计算杂志, (), 第A–A页, https://doi.org/10.1137/19M1253356。[44] J. Yin, L. Zhang, and P. Zhang, *Solution landscape of the Onsager model identifies non-axisymmetric critical points* 430 2022 133081, 10 1016 , 《物理学D辑》, (), 第页, https://doi.org/ . /j. physd.2021.133081。[45] J. Zhang and Q. Du, *Shrinking dimer dynamics and its applications to saddle point search*, SIAM 数值分析杂志, 50（2012）, 第1899–1921页, https://doi.org/10.1137/110843149。[46] L. Zhang, *Construction of solution landscapes for complex systems*, 《数学物理学报》, 45（2023）, 第267–283页, https://doi.org/10.12286/jssx.j2023-1121。[47] L. Zhang, W. Ren, A. Samanta, and Q. Du, *Recent developments in computational modelling of nucleation in phase transformations* 2 2016 16003, , 《npj计算材料学》, (), 第页, https:

//doi.org/10.1038/npjcompumats.2016.3。 [48] L. Zhang, P. Zhang, and X. Zheng, *Error estimates for euler discretization of high-index saddle dynamics* 60 2022 2925 2944, 10，SIAM 数值分析杂志，（），第–页https://doi.org/. 1137/21M1458314。 [49] L. Zhang, Y. Zhang, and T. Luo, *Geometry and local recovery of global minima of two-layer neural networks at overparameterization* 2025 2309 00508，，https://arxiv.org/abs/。。 [50] Y. Zhang, Z. Zhang, T. Luo, and Z. J. Xu, *Embedding principle of loss landscape of deep neural networks* 34 2021，神经信息处理系统进展，（），第14848–14859页，https://proceedings.neurips.cc/paper files/paper/2021/file/7cc532d783a7461f227a5da8ea80bfe1-Paper.pdf。

附录 A. 引理的证明。

证明引理 3.5。 [13,定理 8.1。5] 可得

$$|\lambda_p - \hat{\lambda}_p| \leqslant \|\nabla^2 E(\hat{\theta}) - \nabla^2 E(\theta)\|_2 \leqslant M\|\theta - \hat{\theta}\|_2 \leqslant M\delta \leqslant \mu/2. \tag{A.1}$$

由 [34, 定理 3.6],可得

$$\operatorname{dist}(U, \widehat{U}) \leqslant \frac{1}{\mu - \mu/2}\|\nabla^2 E(\hat{\theta}) - \nabla^2 E(\theta)\|_2 = \frac{2}{\mu}\|\nabla^2 E(\hat{\theta}) - \nabla^2 E(\theta)\|_2. \tag{A.2}$$

该结果可通过联立（A.1）与（A.2）得到。 □

证明引理 3.6。 第一部分直接由恒等式得到

$$\nabla E(\theta') - \nabla E(\theta) = \left(\int_0^1 \nabla^2 E\left(\theta + t(\theta' - \theta)\right) \mathrm{d}t\right)(\theta' - \theta), \tag{A.3}$$

若 θ, $\theta' \in \mathcal{U}$（$\mathcal{M}$, δ），则根据假设 3.3（i），可得

$$\|J\|_2 \leqslant \int_0^1 \left\|\nabla^2 E(\theta') - \nabla^2 E(\theta + t(\theta' - \theta))\right\|_2 \mathrm{d}t \leqslant \frac{1}{2}M\|\theta' - \theta\|_2.$$

□

ˆ引理 3.7 的证明。 利用 （A.3）于 θ 和 θ处，可得

$$\|\nabla E(\theta)\|_2 \leqslant \int_0^1 \left\|\nabla^2 E(\hat{\theta} + t(\theta - \hat{\theta}))\right\|_2 \mathrm{d}t \cdot \|\theta - \hat{\theta}\|_2 \leqslant L\|\theta - \hat{\theta}\|_2. \tag{A.4}$$

ˆ 利用 引理 3.6 在基点 θ 和 θ处，可得

$$\nabla E(\theta) = \nabla^2 E(\theta)\mathcal{P}_{U^\perp}(\theta - \hat{\theta}) + \nabla^2 E(\theta)\mathcal{P}_U(\theta - \hat{\theta}) + J(\theta - \hat{\theta}), \tag{A.5}$$

其中第一项属于 $U^\perp$，第二项属于 U，且 $\|J\|_2 \leqslant \frac{M}{2}\|\theta - \hat{\theta}\|_2$。 由ˆ ˆ dist($U$, U) $\leqslant C\|\theta - \theta\|_2$ 以及引理 3.5，可知

$$\|\mathcal{P}_U(\theta - \hat{\theta})\|_2 = \|\mathcal{P}_U(\theta - \hat{\theta}) - \mathcal{P}_{\widehat{U}}(\theta - \hat{\theta})\|_2 \leqslant \operatorname{dist}(\widehat{U}, U)\|\theta - \hat{\theta}\|_2 \leqslant C\|\theta - \hat{\theta}\|_2^2.$$

现在我们估计 $\|\mathcal{P}_{U^\perp}\nabla E(\theta)\|_2$ 和 $\|\mathcal{P}_U\nabla E(\theta)\|_2$。考虑到将 $\nabla^2 E(\theta)$限制在 $U^\perp$ 上时，其所有特征值均位于（∞, $-\mu$] $\cup$ [μ, $+\infty$）内，我们有

$$\begin{aligned}\|\mathcal{P}_{U^\perp}\nabla E(\theta)\|_2 &\geqslant \|\nabla^2 E(\theta)\mathcal{P}_{U^\perp}(\theta - \hat{\theta})\|_2 - \|J(\theta - \hat{\theta})\|_2 \\ &\geqslant \mu\left(\|\theta - \hat{\theta}\|_2 - \|\mathcal{P}_U(\theta - \hat{\theta})\|_2\right) - \|J\|_2\|\theta - \hat{\theta}\|_2 \geqslant \|\theta - \hat{\theta}\|_2\left(\mu - \tfrac{5}{2}M\|\theta - \hat{\theta}\|_2\right),\end{aligned}$$

由 $\|\nabla^2 E$（θ）$\mathcal{P}_U\|_2 \leqslant \frac{\mu}{4}$ 在注记3.4中，我们可以推导出

$$\begin{aligned}\|\mathcal{P}_U(\nabla E(\theta))\|_2 &\leqslant \|\nabla^2 E(\theta)\mathcal{P}_U\|_2\|\mathcal{P}_U(\theta - \hat{\theta})\|_2 + \|J\|_2\|\theta - \hat{\theta}\|_2 \\ &\leqslant \left(\frac{\mu}{4}C + \frac{M}{2}\right)\|\theta - \hat{\theta}\|_2^2 = M\|\theta - \hat{\theta}\|_2^2,\end{aligned} \tag{A.6}$$

□