arXiv:2603.05703v1 [stat.ME] 5 Mar 2026

随机点积图作为动力系统：局限与机遇

Giulio Valentino Dalla Riva Baffelan.comme@gvdallariva.net

2026年3月9日

摘要

在生态学、历史学、经济学、社会行为、文化动态等领域，许多现象都可以描述为实体之间建立和解除相互作用。这类情形通常以时序网络的形式进行数学建模，并将这些对象随时间的演化视为时间序列加以研究，目的是预测未来的网络状态。与此不同，我们在此采用动力系统的视角：如果目标是理解网络为何 会以这样的方式演化，我们能否学得支配其演化的微分方程？我们在随机点积图（RDPG）框架下研究这一问题，其中每个网络快照都由遵循未知动力学演化的潜在位置生成。我们指出，恢复这些动力学存在三个根本障碍：潜在位置的旋转歧义带来的规范自由度、概率矩阵流形结构导致的可实现性约束，以及谱嵌入伪迹引发的轨迹恢复问题。我们基于主纤维丛建立了一个几何框架，将这些障碍形式化并揭示它们之间的相互作用。对于水平族，多项式动力学具有平凡和乐，而拉普拉斯动力学满足一个明确的非交换性判据，从而具有非平凡整体单值性；在 $d = 2$这一情形下，我们得到完整的受限和乐 SO(2)；一般而言， d 我们给出一个有条件的全整体判据，而对于 $d \geq 3$ 一般情形下完整 SO(d) 的结论仍属猜想。Cramér–Rao下界表明，控制几何难度的谱隙同时也控制统计难度，二者密不可分、相互对偶。我们确立了一项可辨识性原理：对称动力学无法吸收反对称规范污染。然而，我们也表明，在有限样本下仍存在显著的实际障碍。我们将可辨识性与实际恢复之间的差距视为一项开放挑战，并讨论未来的研究方向。

关键词 随机点积图 · 谱嵌入 · 动力系统 ·主丛 ·gauge 自由度

1 引言

时序网络是其边和节点随时间变化的网络。它们遍布整个科学领域，从随空间和时间变化的生态食物网 [Poisot 等人， 2015]到在发育过程中重新连接的神经连接组。随机点积图已被证明对此类网络特别有用：Dalla Riva 和 Stouffer [2016] 首次将随机点积图应用于生态网络，表明潜在位置捕捉了食物网中的演化特征；后续工作利用随机点积图嵌入来预测营养级相互作用 [Strydom 等人， 2022], 而Athreya 等人 [2018] 的综述展示了其在连接组分析和社会网络推断中的应用。一个基本问题是：**哪些微分方程支配网络结构的演化？**

本文在随机点积图（随机点积图）的框架内研究该问题 [Athreya 等人，2018年，Young 和 Scheinerman， 2007]。在随机点积图中，每个节点 i 具有潜在位置 $x_i \in \mathbb{R}^d$，节点 i 与 j 之间存在边的概率为 $P_{ij} = x_i^\top x_j$。当这些位置按照某种动力学 $\dot{X} = f$（X）演化时，网络结构随之改变。我们的目标是恢复 f（从观测值中）。

这种方法之所以吸引人，有几个原因。首先，潜在空间提供了一种连续表示，使动力学可以用动力系统理论中的标准工具来建模。其次，X-空间中可解释的动力学（例如吸引、排斥、扩散）会转化为网络结构中可解释的变化。第三，随机点积图框架拥有成熟的统计理论，将潜在位置与所观测邻接矩阵的谱性质联系起来。
然而，我们指出从RDPG观测进行动力学学习存在三**个根本障碍**:

1. 规范自由度（第3.1节）：潜在位置仅在正交变换意义下确定（一种正交不可识别性）。事实上，对于任意 $Q \in O(d)$， X 与 XQ 产生相同的网络。这意味着某些动力学是不可见的：它们改变 X 而不改变可观测网络结构。2. 可实现性约束（第3.2节）：概率矩阵 $P = XX^\top$ 位于低维流形上。并非每一个对称扰动 P 都可实现；那些会提高 P 的秩的扰动是被禁止的。3. 从嵌入恢复轨迹（第4节）：在RDPG流程中，我们并不观测 $X(t)$ 或$P(t)$。我们观测邻接矩阵，然后通过邻接谱嵌入估计 $X(t)$。这一估计步骤必然引入随时间变化的规范选择。即使底层的 $X(t)$ 是光滑的，估计得到的嵌入也可能不规则地跳跃，因为特征分解在每个时间步都会选取 $O(d)$ 纤维的一个代表元。这并非通过取更小的时间步就能解决的问题。它是特征向量不确定性与从含噪图进行估计相结合所产生的结构性伪影。

我们表明，现有方法未能解决这些障碍。诸如UASE [Gallagher等人， 2021] 以及Omnibus [Levin等人， 2017] 等联合嵌入方法所假设的生成模型，与潜在位置上的常微分方程动力学并不相容。事实上，它们建模的是相对于固定“身份”而随时间变化的“活动”，而不是真正演化的位置。带有层次平滑先验的贝叶斯方法 [Loyal， 2025] 能产生光滑的轨迹，但缺乏动力学一致性：它们插值效果良好，但并未强制速度依照 $X = f(X)$ 依赖于状态。成对普罗克鲁斯对齐是局部的，无法确保全局一致性。
一种相关策略由Athreya 等人 [2025], 提出，它通过定义普罗克鲁斯对齐距离 d_{MV} （即不同时间的网络状态之间的距离），并利用经典多维尺度分析提取一条概括网络演化的低维曲线（“镜像”），从而避免规范自由度。他们的 d_{MV} 是规范不变的，并且可以由邻接谱嵌入一致地估计。它非常适合变点检测和可视化，但捕捉的是距离（网络状态之间的距离），而非支配其演化的微分方程，而这正是本文的目标。

我们的主要贡献是理论性的。我们建立了一个基于主纤维丛的几何框架，将规范自由度、可实现性和轨迹恢复形式化为彼此独立但相互关联的障碍。我们证明，随机规范伪影会引入反对称污染，而这种污染无法被对称动力学吸收（定理 4），从而确立了可辨识性原理。我们阐明了水平动力学族之间的具体差异：多项式动力学具有可交换生成元和平凡和乐，因此规范对齐纯粹是一个统计问题；对于拉普拉斯动力学，我们证明了非平凡整体单值性的局部交换子准则，以及在维度 $d = 2$下的完整受限 d，并将一般情况下的全 $\mathrm{SO}(d)$留作 $d \geq 3$ 猜想。我们推导出克拉美–罗下界，表明控制曲率的谱隙同时控制费舍尔信息，因此几何难度与统计难度密不可分。然而，我们也表明，在实践中利用这些结构面临根本性困难：有限样本导致的偏差，以及自然动力学族的表达能力，使构造性问题远比可辨识性理论所暗示的更加困难。我们将可辨识性与实际恢复之间的差距界定为一个开放问题，并讨论取得进展的方向。

论文结构。 第2 节回顾随机点积图的基础知识。第3 节阐述几何框架：规范自由度、可实现性，以及包含联络、曲率和和乐的纤维丛视角。第4 节探讨从谱嵌入中恢复轨迹这一实际问题，说明现有方法为何失效，并阐明为何需要动力学感知方法。第5 节整理具体的动力学族，通过李雅普诺夫方程分析其可观测后果，并明确形式化阐述多项式与拉普拉斯算子和乐之间的对比，同时标明结论属于已证明、有条件成立或猜想。第6 节推导信息论下界，揭示统计-几何对偶。第7节探讨构造性问题：可辨识性原理及其算法意义、尚存的实际障碍，以及一个利用锚点节点的可处理特殊情形，并通过数值实验加以说明。第8 节总结研究成果并讨论开放问题。

记号。 $X \in \mathbb{R}^{n\times d}$ 表示以 $x_i^\top$为行的潜在位置矩阵。$\mathbb{R}_*^{n\times d} = \{X \in \mathbb{R}^{n\times d} : \text{秩}(X) = d\}$表示满秩 $n \times d$ 矩阵的集合。 $P = XX^\top$ 是概率矩阵。 $O(d)$是正交群；

$\mathfrak{so}(d)$是其斜对称矩阵的李代数。$\hat{X}$ 表示一个估计（例如来自谱嵌入）。 $\|\cdot\|_F$是Frobenius范数； $\|M\|_{2\to\infty} = \max_i \|e_i^\top M\|_2$是二到无穷范数，用于度量$M$的最大行范数。 $O_p(\cdot)$表示随机阶： $Y_n = O_p(a_n)$意味着 Y_n/a_n 依概率有界。Sym(d)表示 $d \times d$对称矩阵。下标 i, j指节点；下标 ι, γ在谱基下工作时指特征向量方向。

基本假设（全文默认）。除非结果另有说明，我们均采用以下假设：

（S1）维度固定，且 $n \geq d \geq 2$； **（S2）**轨迹位于某个时间区间内，其中 $X(t)$始终处于 $\mathcal{E}$ 的内部（因此秩($X(t) = d$ 且 $0 < P_{ij}(t) < 1$）； **（S3）**轨迹/场具有足够的光滑性，以满足所述导数和李括号的要求； **（S4）**对于统计界，正则性假设采用命题14中的假设。

结果	状态	范围
定理 1	已证明	满秩潜在状态
命题 9	已证明	水平对称族
命题 11	已证明	局部非零对易子判据
推论 3	已证明	线性动力学
命题 14	已证明	多项式动力学 + 正则性假设
定理 4	已证明	对称生成元可辨识性
通过曲率张成空间得到完整限制全纯群	条件定理	要求满足曲率张成假设
基点处的有限时间秩判据	条件命题	完整限制全纯群的充分条件
猜想 1	猜想	$d \geq 3$一般整体群

表 1：结果状态图：已证明的结果、条件性判据与猜想。

我们以表1 作为整篇手稿中定理主张的权威状态图例 .

2 随机点积图

我们从RD开始 无向网络的PG框架。向有向图的扩展见第D .

2.1 潜在位置与联络概率

在随机点积图中，每个节点 $i \in \{1, \ldots, n\}$ 都关联一个潜在位置 $x_i \in \mathbb{R}^d$。节点 i与 j 之间存在一条边的概率为：

$$P_{ij} = x_i^\top x_j \tag{1}$$

要使该概率有效，我们需要对所有节点对满足 $P_{ij} \in [0, 1]$。一个充分条件是，所有位置都位于单位球的正正交象限内：

$$B_+^d = \{x \in \mathbb{R}^d : x_k \geq 0 \text{ for all } k, \quad \|x\| \leq 1\} \tag{2}$$

非负性确保了 $P_{ij} \geq 0$，而柯西–施瓦茨不等式给出 $P_{ij} \leq \|x_i\|\|x_j\| \leq 1$。

将位置收集到一个矩阵 $X \in \mathbb{R}^{n\times d}$ 中，其各行为 $x_i^\top$，概率矩阵为：

$$P = XX^\top \tag{3}$$

该矩阵是对称的、半正定的，且秩至多为 d。本文始终仅考虑所有边概率均位于开区间内的构形：$0 < P_{ij} < 1$ 对所有 i， j。这排除了会导致费舍尔信息出现奇异性以及微分几何处理复杂化的边界情形（确定性边或必然不存在的边）；详情请参见第 3.3 节。

给定潜在位置后，各条边独立生成：

$$A_{ij} \sim \text{Bernoulli}(P_{ij}) \quad \text{for } i \leq j \tag{4}$$

其中 $A_{ji} = A_{ij}$（无向），且 $A_{ii} \sim$ 伯努利(P_{ii})，因此 $\mathbb{E}[A] = P$ 恰好成立。在某些应用中，通常会通过设置 $A_{ii} = 0$ 来忽略自环，从而得到 $\mathbb{E}[A] = P - \text{diag}(P)$。这会使对角线变为空，并引入系统性的流形外偏差，使谱理论更加复杂（扰动 diag(P) 的谱范数至多为 1；对于适中的 n，这一量级与抽样噪声相当）。我们的几何和统计论证可通过将缺失或置零的对角线视为额外的结构化扰动项，扩展到这种情形；正文中我们暂不考虑这一情形，以便聚焦于主题，并使 $\mathbb{E}[A]$严格位于秩为 d 的流形 $\{XX^\top\}$ 上。（如有需要，可以在附录中通过明确追踪额外对角扰动在嵌入和对齐步骤中的影响，推导空对角线扩展。）

2.2 邻接谱嵌入

给定观测到的邻接矩阵 A，我们通过**邻接谱嵌入**（ASE）[Athreya等人， 2018]。

由于 A是对称的，因此它具有特征分解 $A = U\Lambda U^\top$，其特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$均为实数，且对应正交归一特征向量。秩为$d$的邻接谱嵌入为:

$$\hat{X} = U_d|\Lambda_d|^{1/2} \tag{5}$$

其中， $U_d \in \mathbb{R}^{n\times d}$ 包含按特征值绝对值排序的前 d 个特征向量， $\Lambda_d = \text{diag}(\lambda_1, \ldots\ldots, \lambda_d)$。我们取绝对值，是因为 A 可能因抽样噪声而具有负特征值，尽管真实的 P是半正定的。在稠密区间中，当 $\lambda_d(P) = \Omega(\sqrt{n})$时，韦尔不等式保证$A$ 的前 d 个特征值以高概率为正，因此渐近而言，绝对值不会产生影响。在有限的 n下，为明确起见，我们仍保留绝对值。

注。 对于对称矩阵，特征分解与奇异值分解密切相关：若 $A = U\Lambda U^\top$(特征) 且$A = \tilde{U}\Sigma\tilde{V}^\top$(奇异值分解)，则 $\tilde{U} = \tilde{V} = U$(至多相差符号) 且 $\sigma_i = |\lambda_i|$。我们对邻接谱嵌入采用特征分解（这是随机点积图文献中的标准做法），而对 Procrustes 对齐采用奇异值分解（这是该问题中的标准做法）。

统计性质。随机点积图的统计性质已得到充分研究，这里我们回顾文献中的一些基本^结果。在温和条件下，邻接谱嵌入是一致的：当 $n \to \infty$ 时， X 的各行以 $O(1/\sqrt{n})$的速率收敛到真实的潜在位置（至多相差正交变换）。 [Athreya 等人， 2016, 2018]。具体而言，存在 $Q \in O(d)$，使得 $\max_i \|\hat{}x_i - x_iQ\| = O_p(1/\sqrt{n})$；并且在给定每个潜在位置的条件下， $\sqrt{n}(\hat{}x_i - Qx_i)$收敛到一个高斯分布，其协方差取决于边方差剖面 [Athreya 等人，2016年，Rubin-Delanchy 等人， 2022]。扰动展开^ $X = XW + (A-P)X(X^\top X)^{-1}W + R$，其中 $W \in O(d)$且 $\|R\|_{2\to\infty} = o(n^{-1/2})$，表明主要误差项关于 $A - P$是线性的，且其条件均值为零 [Cape 等人， 2019]。邻接谱嵌入渐近无偏，但对单个潜在位置而言并非完全有效；Xie 和 Xu 的一步法[2023] 实现了他们所称的局部效率：对于每个顶点，估计量的渐近协方差与已知其他所有潜在位置的预言机最大似然估计的渐近协方差相匹配（至多相差正交变换）。对于随机块模型参数，谱估计量本身渐近有效 [Tang 等人， 2022]。 A 的特征值也渐近正态地集中在 P的特征值附近，其偏差为 $O(1)$，源于伯努利方差 [Tang 等人， 2018]。相对于主导特征值而言，这种偏差可以忽略（$\lambda_1 \sim n$），但当谱隙较小时，它可能与谱隙相当 $\delta = \lambda_d$；此外，其量级与特征值波动相同（$O(\sqrt{n})$）。对于动力学估计，重要的是偏差相对于各时间步之间特征值变化的大小；我们将在第6节重新讨论这一点。

正交矩阵 Q 反映了根本性的规范歧义：特征向量的符号无法唯一确定，而重特征子空间内的旋转则是任意的。这种歧义无法避免，也是第4节讨论的轨迹估计问题的根源。

3 规范障碍

我们分析从 RDPG观测中学习动力学所面临的根本障碍，从理论角度刻画哪些内容可以学习，哪些内容无法学习。

3.1 规范自由度与可观测性

潜在位置 X 无法由概率矩阵 P 唯一确定。对于任意正交矩阵 $Q \in O(d)$:

$$(XQ)(XQ)^\top = XQQ^\top X^\top = XX^\top = P \tag{6}$$

因此， X 和 XQ会产生完全相同的连接概率。这种 $O(d)$对称性就是随机点积图的**规范自由度**： 等价类 $[X] = \{XQ : Q \in O(d)\}$ 对应于单一的可观测网络结构。

从几何视角来看，对潜在空间中所有位置进行全局旋转，不会改变两两夹角和大小，因此所有点积都保持不变。

这种规范自由度对动力学学习有深远影响。如果 X 和 XQ 在观测上无法区分，那么任何沿等价类移动的动力学，也就是通过随时间变化的共同正交变换旋转所有位置的动力学，都会**不会在网络中产生可观测的变化**。

考虑潜在位置上的动力学 $\dot{X} = f(X)$。概率矩阵 $P = XX^\top$上的诱导动力学可由乘积法则得出：

$$\dot{P} = \dot{X}X^\top + X\dot{X}^\top = f(X)X^\top + Xf(X)^\top \tag{7}$$

定义 1（可观测与不可观测动力学）。 若向量场 $f\colon \mathbb{R}^{n\times d} \to \mathbb{R}^{n\times d}$ 产生**可观测动力学**，则称其**可观测动力学**，当且仅当 $P \neq 0$。否则，该动力学为**不可见**：潜在位置发生变化，但网络结构保持静止。

该定义自然引出一个问题：哪些动力学是不可见的?

定理 1（不可观测动力学的刻画）。 对于 $X \in \mathbb{R}^{n\times d}_*$（即秩$(X) = d$），向量场 f 产生不可见动力学（$P = 0$），当且仅当 $f(X) = XA$，其中某个矩阵 $A \in \mathfrak{so}(d)$ 为反对称矩阵。

证明。 ($\Leftarrow$) 若 $f(X) = XA$ 满足 $A^\top = -A$，则：

$$\dot{P} = f(X)X^\top + Xf(X)^\top = XAX^\top + XA^\top X^\top = XAX^\top - XAX^\top = 0 \tag{8}$$

($\Rightarrow$) 假设 $\dot{P} = f(X)X^\top + Xf(X)^\top = 0$。令 $G = X^\top X$，由于 $\mathrm{rank}(X) = d$，它是可逆的。将恒等式左乘 $(I - XG^{-1}X^\top)$，以投影到 $\mathrm{col}(X)^\perp$：

$$0 = (I - XG^{-1}X^\top)f(X)X^\top \tag{9}$$

由于 $X^\top$ 具有满行秩，这意味着

$$(I - XG^{-1}X^\top)f(X) = 0 \tag{10}$$

因此 $f(X) \in \mathrm{col}(X)$，且存在 $B \in \mathbb{R}^{d\times d}$使得 $f(X) = XB$。

代回 $P = 0$ 可得

$$0 = X(B + B^\top)X^\top \tag{11}$$

并且由于 X 具有满列秩，这迫使 $B + B^\top = 0$成立，即 $B \in \mathfrak{so}(d)$。取 $A = B$ 即完成证明。 □

我们可以很容易地解释这一结果：不可见动力学恰好是潜在空间中围绕原点的均匀旋转；所有其他动力学（吸引、排斥、非均匀旋转、漂移等）都会导致网络结构出现可观测的变化。

这意味着不可见动力学的类别很小（维数为 $d(d-1)/2$，即 $\mathfrak{so}(d)$ 的维数），而可观测动力学则遍布一个大得多的空间。

3.2 可实现动力学

除了规范自由度之外，随机点积图动力学还受到几何约束：概率矩阵 $P = XX^\top$ 位于一个低维流形上，因此大多数（对称）扰动 P 都无法通过 X 的动力学实现。

命题 1 (切空间约束)。 设 $V \in \mathbb{R}^{n\times d}$ 是$\mathrm{col}(P)$的一组标准正交基，且 $V_\perp \in \mathbb{R}^{n\times(n-d)}$张成其正交补。任何可实现的 $\dot{P}$（即，对于某个 $\dot{X}$，有 $\dot{P} = \dot{X}X^\top + X\dot{X}^\top$）都满足：

$$V_\perp^\top \dot{P} V_\perp = 0 \tag{12}$$

可实现切空间的维数为 $nd - d(d-1)/2$。

证明。 任意可实现的 $\dot{P} = FX^\top + XF^\top$ 都存在某个 $F \in \mathbb{R}^{n\times d}$。由于 $\mathrm{col}(X) = \mathrm{col}(V)$，因此对于可逆的 R，有 $X = VR$。于是：

$$V_\perp^\top \dot{P} V_\perp = V_\perp^\top FR^\top V^\top V_\perp + V_\perp^\top VRF^\top V_\perp = 0 \tag{13}$$

使用 $V^\top V_\perp = 0$。 □

通过考虑生成矩阵的**分块分解**，我们可以更好地理解其动力学。注意，任意对称矩阵 M 都可以相对于 $(V, V_\perp)$ 分解为若干块：

$$M = \underbrace{VAV^\top}_{\text{range-range}} + \underbrace{VBV_\perp^\top + V_\perp B^\top V^\top}_{\text{cross terms}} + \underbrace{V_\perp C V_\perp^\top}_{\text{null-null}} \tag{14}$$

对于可实现的 P：A 和 B 块可以是任意的，但 $C = 0$ 始终如此。零-零块表示完全支撑在 $\text{col}(P)^\perp$ 上的方向，因此这些方向不能由 $\dot{P} = FX^\top + XF^\top$ 产生，并且通常会增加 P的秩。对于固定潜在维度 d，沿这些方向的移动是被禁止的（但对于起始时秩亏且维度会增长的网络，它们或许可以实现）。

推论 1（维数计数）。对于底空间 $\mathcal{B}$中的每个概率矩阵 P，无穷小可实现扰动 $\dot{P}$ 的集合构成 P处的切空间，其维数为：

$$\dim(T_P\mathcal{B}) = nd - d(d-1)/2 \tag{15}$$

这等于 X空间的维数（nd）减去规范自由度（$d(d-1)/2$）。

分块分解为该模型提供了一种直接诊断：如果观测到的动力学在零-零块 $V_\perp^\top \dot{P} V_\perp$中具有非零结构，则这表明以下二者之一：

1. **模型误设**：真实动力学不能保持低秩结构 2. **维度涌现**：潜在维度 d 正在增加，这表明新的交互维度正在涌现

3.3 纤维丛视角

随机点积图中的规范自由度具有自然的几何结构，有助于我们理解从网络观测值中可以以及不可以学习到什么。$\mathbb{R}_*^{n\times d}/O(d)$的商几何已在优化文献中得到广泛发展，尤其见 Absil 等人 [2008], Journée 等人 [2010], 以及 Massart 和 Absil [2020], Massart 等人 [2019]。我们在此回顾这一几何结构，并将其应用于随机点积图设定；在该设定下，联络、曲率和和乐结构具有具体的统计学含义，它们构成了动力学估计的障碍。

注。**记号（切空间）**。对于流形 $\mathcal{M}$及其上的点 $p \in \mathcal{M}$，我们用 $T_p\mathcal{M}$ 表示点p处的切空间，用 $T\mathcal{M}$ 表示切丛（所有切空间的并集）。本文中，对于 $\mathcal{E} \subset \mathbb{R}^{n\times d}$的内部点，我们将 $T_X\mathcal{E}$与 $\mathbb{R}^{n\times d}$等同（即 X的可允许无穷小扰动）。同样，$T_P\mathcal{B}$表示在 P处可实现的无穷小扰动 $\dot{P}$所构成的空间。

3.3.1 为什么需要纤维丛?

不妨想象这样一个场景。撰写本文时，我常常面朝海滩工作，羡慕着海里游泳的人。时不时会有一条好奇的鲨鱼游向岸边，被人发现。随后，当地警方会乘小艇出海搜寻。在我们此处用来帮助理解的那个（假想）场景中，他们的雷达上出现了一个移动的点：鲨鱼的位置被投影到以船为中心的平面坐标空间中。这条好奇的鲨鱼在三维空间中非常平滑地移动：它可以游得离船更近或更远，改变接近的角度，也可以潜入更深的海中。然而，单凭曲面轨迹，你无法判断鲨鱼是在浮出水面，还是在向深处下潜。因此，只观测到这个表面投影时，许多不同的三维轨迹都可能产生相同的二维轨迹。

类似地，在动态随机点积图的设定中，我们有一个潜在结构 X，它按照某种机制随时间演化。然而，我们只能观测到 X 的一个函数，其中包含部分信息：我们从交互概率矩阵 $P(t) = X(t)X(t)^\top$ 中抽样得到图，并且知道 $P(t)$ 在 $O(d)$ 旋转下保持不变。

在我们的情形中，可观测量是 $P = XX^\top$（或者在实际应用中，是从中抽取的含噪图），而潜在构型就是 X 本身。潜在构型包含了我们讨论潜在空间中的速度和加速度所需的信息，但无法直接测量。将 X 替换为旋转后的版本 XQ，不会改变 P。它是同一“曲面轨迹”的另一个代表元。（这只是一个类比。规范方向并非字面意义上的物理深度，但要点相同：存在一个无法从底空间观察到的隐藏方向。）

纤维丛将这一结构形式化，使我们能够研究底空间与全（潜在）空间之间的关系。我们将所有可能的可观测量组成的空间（例如 $P(t)$）称为底空间 $\mathcal{B}$。每个可观测量 P上方都有一组等价的潜在构型 $\{XQ$，即：$Q \in O(d)\}$，它们都会产生相同的 P。我们将完整的

潜在构型空间 X 称为全空间 $\mathcal{E}$。我们还配备了一个投影 $\pi(X) = XX^\top$，它将全空间映射到底空间。

在纤维丛中，我们可以考虑**提升**，将轨迹从底空间映射到全空间：给定底空间中可观测量随时间变化的路径 $P(t)$，我们将其"提升"为全空间中的路径 $X(t)$。然而，对于任意一个 $P(t)$，都有无穷多种可能的 $X(t)$ 选择，它们之间相差时变规范选择 $Q(t)$。丛的联络 和曲率 告诉我们哪些提升是"自然的"（没有伪旋转），以及是否根本可能实现一致的提升。如果不可能，而且丛本身具有曲率，我们就会得到**和乐**。这是沿着最谨慎的轨迹也会不断累积的规范漂移。

现在我们来精确说明这一点。

3.3.2 概率约束

在定义丛之前，我们必须先处理此前一笔带过的一个约束：要让 P 表示连接概率，对于所有成对的 i, j，我们需要 $P_{ij} \in [0, 1]$。

并非每个 $X \in \mathbb{R}^{n\times d}$ 都满足这一条件。约束 $(XX^\top)_{ij} = x_i^\top x_j \in [0, 1]$ 定义了$\mathbb{R}^{n\times d}$的一个半代数子集。

定义 2（有效潜在位置）。有效配置空间为：

$$\mathcal{E} = \{X \in \mathbb{R}^{n\times d} : \operatorname{rank}(X) = d, \quad 0 \le x_i^\top x_j \le 1 \text{ for all } i, j\} \tag{16}$$

有效概率空间为：

$$\mathcal{B} = \{P \in \mathbb{R}^{n\times n} : P = XX^\top \text{ for some } X \in \mathcal{E}\} \tag{17}$$

满足 $X \in \mathcal{E}$ 的一个充分条件是所有行都位于**单位球的正正交象限**内 ：

$$x_i \in B_+^d = \{x \in \mathbb{R}^d : x_k \ge 0 \text{ for all } k, \quad \|x\| \le 1\} \tag{18}$$

这确保了 $x_i^\top x_j \ge 0$（非负条目）和 $x_i^\top x_j \le \|x_i\|\|x_j\| \le 1$（柯西–施瓦茨不等式）。然而， $X \in B_+^d$ 是充分但非必要条件：许多有效构形的部分条目可以位于正正交象限之外，只要所有成对点积仍属于 $[0, 1]$。

内部与边界。$\mathcal{E}$ 的内部由所有约束均为严格不等式的构形组成： $0 < x_i^\top x_j < 1$。边界包括某些 $P_{ij} = 0$的构形（节点的位置正交，因此联络概率为零），或某些 $P_{ij} = 1$ 的构形（节点的位置方向一致且长度为单位长度，因此必然发生联络）。

纤维丛结构要求全空间光滑，这一条件在内部成立：在那里， $\mathcal{E}$ 是光滑流形 $\{X \in \mathbb{R}^{n\times d}$ 的开子集：$\operatorname{rank}(X) = d\}$，因此联络、曲率和和乐的所有理论工具都可直接应用，无需额外说明。

在边界处，会出现三个问题。首先，可行集 $\{X : \ 0 \le x_i^\top x_j \le 1$ 对所有 $i, \ j\}$ 由多项式不等式定义，并且在多个约束同时激活的位置具有角。因此，它是带角流形，而不是光滑流形。其次，在边界点，切空间被切锥取代：并非所有方向都可行，因为某些方向会将 P_{ij} 推到 $[0, 1]$之外。命题2 中的水平-垂直分解可能无法保持可行性：可行速度的水平分量可能指向约束集之外。第三，如果轨迹到达 $P_{ij} = 0$ 或 $P_{ij} = 1$，无约束常微分方程 $\dot{X} = f(X)$ 可能会试图离开可行区域，因此需要进行约束处理（投影、反射或障碍项，详见第 7 节）。

在实际应用中，对于关注的时序网络而言，边界问题很少构成限制：边概率恰好为 0 或 1，分别对应于从不发生交互的节点（这类节点通常因"无法检测"而被排除在时序网络分析之外），或总是发生交互的节点（这意味着网络结构不具有随机变异性）。本文余下部分将在 $\mathcal{E}$ 内部的一个连通分支上开展研究，此处的几何结构平滑，丛结构也很规整。这确保轨迹 $X(t)$ 不会穿过纤维丛结构退化的秩亏分层。

3.3.3 主丛结构

定义 3（随机点积图主丛）。 在有效空间的内部，$(\mathcal{E}, \mathcal{B}, \pi, O(d))$ 构成一个主纤维丛：

- **全空间** $\mathcal{E}$：有效潜在构型（内部）

- **底空间** $\mathcal{B}$：有效概率矩阵（内部）
- **投影** π： $\mathcal{E} \to \mathcal{B}$：映射 $X \mapsto XX^\top$
- **结构群** $O(d)$：作用于 $\mathcal{E}$ 通过 $X \cdot Q = XQ$

P 上方的**纤维** 是 $\pi^{-1}(P) = \{XQ : \ Q \in O(d)\} \simeq O(d)$。

这是一个**主丛**，因为 $O(d)$ 的作用是自由的（没有任何 X 会被非恒等的 Q固定），并且在每个纤维上都是传递的（ P 的任意两个提升只相差一个 Q）。

3.3.4 分解运动：垂直与水平

在每个 $X \in \mathcal{E}$，我们都可以问：哪些运动方向会改变 P，哪些不会？

垂直子空间 $\mathcal{V}_X$ X $P =$ 由沿着纤维的方向组成。沿这些方向运动会发生变化，但不会改变：$XX^\top$

$$\mathcal{V}_X = \ker(d\pi_X) = \{X\Omega : \Omega \in \mathfrak{so}(d)\} \tag{19}$$

这些正是**不可见动力学**，来自定理 1：无穷小旋转 $X = X\Omega$，其中 Ω 是反对称的。

水平子空间 $\mathcal{H}_X$ 由与纤维横向的方向组成。沿这些方向运动确实会改变 P（以及 X）：

$$\mathcal{H}_X = \{Z \in T_X\mathcal{E} : X^\top Z \in \mathrm{Sym}(d)\} \tag{20}$$

每个切向量都可以唯一地分解为垂直部分和水平部分：

$$T_X\mathcal{E} = \mathcal{V}_X \oplus \mathcal{H}_X \tag{21}$$

命题 2（水平性刻画）。切向量 $\dot{X} \in T_X\mathcal{E}$ 是水平的，当且仅当 $X^\top\dot{X}$ 是对称的。

证明。 将 $\dot{X} = X\Omega + H$ 写成 $\Omega \in \mathfrak{so}(d)$（垂直部分）和 $H \in \mathcal{H}_X$（水平部分）。那么 $X^\top\dot{X} = X^\top X\Omega + X^\top H$。根据定义，项 $X^\top H$ 是对称的。项 $(X^\top X)\Omega$ 只有在 $\Omega = 0$时才是对称的，因为正定对称矩阵与非零反对称矩阵的乘积绝不可能是对称的。因此，$X^\top\dot{X}$ 是对称的，当且仅当 $\Omega = 0$，也当且仅当 $\dot{X}$ 完全水平。 □

水平条件 $X^\top\dot{X} \in \mathrm{Sym}(d)$ 是一种**规范固定条件**：它从所有在 $\mathcal{E}$中移动的方式里，选出那些不会沿纤维“白费力气”的方式。

3.3.5 联络 1-形式：提取规范分量

水平子空间 $\mathcal{H}_X$ 在 $\mathcal{E}$ 上平滑变化，这种选择称为 **Ehresmann 联络**。它提供了一种一致的方法，在整个丛中区分“可观测”方向和“规范”方向。有了联络之后，我们就可以定义平行移动（沿底空间移动，同时保持水平）和曲率（衡量联络的扭曲程度）。

Ehresmann 联络可以用一个**联络 1-形式** ω 来表示，它用于提取任意运动的垂直（规范）分量：

定义 4（联络 1-形式）。对于每个构形 $X \in \mathcal{E}$，切空间 $T_X\mathcal{E}$ 是允许的无穷小扰动 Z 的向量空间，其中 Z 是对 X 的无穷小扰动（在内部点处，它与 $\mathbb{R}^{n\times d}$重合）。联络 1-形式是将每个切向量 $Z \in T_X\mathcal{E}$ 映射到其垂直（规范）分量的映射：

$$\omega_X : T_X\mathcal{E} \to \mathfrak{so}(d) \tag{22}$$

定义为

$$\omega_X(Z) = \Omega \quad \text{where} \quad (X^\top X)\Omega + \Omega(X^\top X) = X^\top Z - Z^\top X \tag{23}$$

也就是说，$\omega_X(Z)$是求解该李雅普诺夫方程的唯一反对称矩阵 Ω。

水平空间 $\mathcal{H}_X = \{Z : \ X^\top Z$ 对称$\}$ 并非任意对象：它是度量联络，由 Frobenius 内积诱导。等价地，$\omega_X(Z)$ 最小化 $\|Z - X\Omega\|_F^2$ 在 $\Omega \in \mathfrak{so}(d)$上：它提取动能最小的规范分量。这是商流形上的标准黎曼联络 [Absil 等人，2008 年，Massart 和 Absil，2020],从而确保水平提升在适当投影后为测地线。

为推导其表达式，请记住任意切向量可分解为 $Z = X\Omega + H$，其中 $\Omega \in \mathfrak{so}(d)$ 且 $X^\top H$对称。因此，计算 $X^\top Z - Z^\top X$：由于 $\Omega^\top = -\Omega$，我们有 $Z^\top X = -\Omega(X^\top X) + H^\top X$。于是$X^\top Z - Z^\top X = (X^\top X)\Omega + \Omega(X^\top X) + (X^\top H - H^\top X)$。由于 $X^\top H$ 是对称的，$(X^\top H - H^\top X) = 0$，从而得到李雅普诺夫方程。唯一性成立，因为当 $G = X^\top X$ 正定时，李雅普诺夫算子 $\Omega \mapsto G\Omega + \Omega G$ 可逆：在 G 的特征基中（特征值为 λ_1，……，λ_d），该算子将每一项 $\Omega_{\iota\gamma}$独立地映射为 $(\lambda_\iota + \lambda_\gamma)\Omega_{\iota\gamma}$，因此它在斜对称矩阵的向量空间上充当对角算子，且所有特征值为 $\lambda_\iota + \lambda_\gamma > 0$。在秩亏点（$\mathrm{rank}(X) < d$）处，$G$ 具有零特征值，李雅普诺夫算子获得一个核，水平-垂直分解不再唯一。这是纤维丛结构在 $\mathcal{E}$的边界处退化这一几何事实的代数对应，并预示了谱隙在命题4中的作用。

联络 1-形式 $\omega_X(X)$ 表示运动 X的“瞬时旋转速率”：如果 $\omega_X(X) = 0$，则运动完全是水平的（没有规范分量）；如果 $\omega_X(X) = \Omega$，则 X 包含速率为 Ω 的旋转。

联络满足三个关键性质：

1. $\ker(\omega_X) = \mathcal{H}_X$：水平向量的规范分量为零 2. $\omega_X(X\Omega) = \Omega$：垂直向量与其旋转速率相对应 3. 等变性：$\omega$ 在规范变换下按相应方式变换

3.3.6 水平提升与平行移动

假设我们观察到一条底空间中的轨迹 $P(t)$（可观测动力学）。我们希望将其“提升”为全空间中的一条轨迹$X(t)$。然而，提升有无穷多个，那么我们应该选择哪一个？

水平提升是没有规范漂移的唯一提升：

$$\pi(X(t)) = P(t) \quad \text{and} \quad \omega_{X(t)}(\dot{X}(t)) = 0 \tag{24}$$

第二个条件表示，速度 $X(t)$ 在任意时间都没有垂直分量：轨迹完全沿水平方向移动。

存在唯一性： 给定 $P(t)$ 以及初始提升 $X(0)$，满足 $\pi(X(0)) = P(0)$，水平提升存在且唯一 [Kobayashi 和 Nomizu。 1963] 水平条件 $\omega = 0$ 定义了 $\mathcal{E}$ 上的一个常微分方程，其解投影到 $P(t)$。

水平提升为何重要： 在所有 $X(t)$ 投影到 $P(t)$的轨迹中，水平提升是“规范意义上的规范”轨迹，能够追踪可观测动力学，同时不会引入伪旋转。如果我们能够从数据中计算水平提升，就能解决规范问题。困难在于，实践中我们观测到的是含噪的邻接矩阵，而不是直接观测到 $P(t)$（关于这一估计问题的详细讨论，见第 5.2 节）。

计算水平提升。 水平提升公式与命题 1的分块分解直接相关。设 $P = V\Lambda V^\top$，其中 $V \in \mathbb{R}^{n\times d}$具有正交归一列，且 $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$，其中 $\lambda_\iota > 0$。自然提升为$X = V\Lambda^{1/2}$，从而有 $X^\top X = \Lambda$。

给定可实现的 P，分块分解得：

$$\dot{P} = VAV^\top + VBV_\perp^\top + V_\perp B^\top V^\top \tag{25}$$

其中 $A = V^\top \dot{P} V$（对称的，$d \times d$）和 $B = V^\top \dot{P} V_\perp$（交叉项）。

我们寻求满足以下条件的水平提升 $X = VS + V_\perp T$：

1. *Projection*: $\dot{P} = \dot{X}X^\top + X\dot{X}^\top$
2. *Horizontality*: $X^\top \dot{X} = \Lambda^{1/2} S$ is symmetric

根据投影条件，匹配各块：

- Range-range: $A = S\Lambda^{1/2} + \Lambda^{1/2} S^\top$
- Cross terms: $B = T^\top \Lambda^{1/2}$, giving $T = B^\top \Lambda^{-1/2}$

水平性条件要求 $S = \Lambda^{-1/2}\Sigma$，其中存在某个对称的 Σ。代入值域-值域方程并令 $\tilde{\Sigma} = \Lambda^{-1/2}\Sigma\Lambda^{-1/2}$:

$$\Lambda\tilde{\Sigma} + \tilde{\Sigma}\Lambda = A \tag{26}$$

一个小小的符号说明： $A = V^\top \dot{P} V$已经是用 $P = V\Lambda V^\top$ 的特征基表示的值域-值域块。因此，等式右侧就是 A，无需再用 Λ 的幂进行共轭。
这与上文联络 1-形式中的李雅普诺夫结构相同。这并非巧合：二者都涉及将规范分量与可观测分量分离。下面的公式本质上是Absil 等人 [2008] 和Massart 和 Absil [2020], 的水平投影，专门应用于 P的谱分解。

命题 3（水平提升公式）。给定 $P = V\Lambda V^\top$ 和可实现的 $\dot{P}$，其块为 $A = V^\top \dot{P} V$ 和 $B = V^\top \dot{P} V_\perp$，则在 $X = V\Lambda^{1/2}$ 处的水平提升为：

$$\dot{X} = V\tilde{\Sigma}\Lambda^{1/2} + V_\perp B^\top \Lambda^{-1/2} \tag{27}$$

˜其中 Σ 是方程（*26*）的对称解，其逐元素定义如下：

$$\tilde{\Sigma}_{\iota\gamma} = \frac{A_{\iota\gamma}}{\lambda_\iota + \lambda_\gamma} \tag{28}$$

˙证明。上述推导（方程（26）及此前的分块代数）构造了 X，使其同时满足投影条件 $\dot{P} = \dot{X}X^\top + X\dot{X}^\top$ 和水平性条件 $X^\top \dot{X}$，且为对称矩阵。由于对所有 $\lambda_\iota + \lambda_\gamma > 0$， ι，李雅普诺夫方程都有唯一的对称解。γ。 □

该公式揭示了水平运动的两个组成部分：

- 第一项 $V\tilde{\Sigma}\Lambda^{1/2}$ 处理当前列空间内部的运动 X，该运动由值域-值域块 A 通过李雅普诺夫方程确定
- 第二项 $V_\perp B^\top \Lambda^{-1/2}$ 处理向新方向扩展的运动，该运动由交叉块 B直接确定

在实践中，根据离散且含噪的观测值计算水平提升，需要进行插值和常微分方程积分。我们将在第 7 节讨论离散 Procrustes 对齐如何近似水平输运。

3.3.7 曲率与和乐

一个微妙的问题是：即使是水平提升，沿闭合回路也可能累积规范漂移。
联络的**曲率** 衡量了水平向量方向不满足对易关系的程度：
定义 5（曲率二形式）。 对于水平向量场 H_1， H_2 在 $\mathcal{E}$上：

$$\Omega(H_1, H_2) = -\omega([H_1, H_2]) \tag{29}$$

其中 $[\cdot,\cdot]$ 是李括号。

当曲率不为零时，先沿方向 H_1 再沿方向 H_2 水平移动，最终到达的位置与先沿方向 H_2 再沿方向H_1不同。因此，会产生垂直（规范）偏差。
这对闭合回路有一个惊人的推论：
定义 6（和乐）。 设 γ: $[0,1] \to \mathcal{B}$ 是一条闭曲线，且满足 $\gamma(0) = \gamma(1) = P$。从 X 开始的水平提升最终到达 XQ_γ，其中某个 $Q_\gamma \in O(d)$。 γ 的**和乐** 就是这一累积旋转 Q_γ。
一个经典结果对曲率与和乐之间的联系给出了精确的定性刻画。具体而言，Ambrose–Singer 定理 [Kobayashi 和 Nomizu， 1963]将受限全纯群（可缩环路的和乐）的李代数刻画为由水平输运过程中遇到的曲率值生成的李代数。（稍后我们将回到特定随机点积图动力学族的具体和乐行为。）
定理 2（全纯障碍（安布罗斯–辛格））。 如果曲率非零，那么受限全纯群就是非平凡的。特别地，在 $\mathcal{B}$（可缩环）中存在闭路径，使得不存在全局一致的规范：任意提升都满足 X（1） $= X$（0）Q， 其中 $Q \neq I$。

结果表明，如果真实动力学 $P(t)$ 描出一个闭环（周期性网络行为），那么底层的 $X(t)$ 可能无法自行闭合。相反，它会在返回时发生和乐旋转。这是一个根本性的障碍：即使局部对齐完美无缺，经过多个环后仍会累积全局规范漂移。

我们可以将全纯障碍与（观测到的）图的谱性质联系起来： $\mathcal{B}$ 的曲率取决于 $P = XX^\top$的特征值。由于 P的非零特征值恰好等于 $X^\top X$ （潜在位置的格拉姆矩阵）的特征值，因此在若干常见情形下会出现较小的 λ_d 。

注。 **谱隙较小时的情形（示例）**。 较小的 λ_d 例如会在以下情形中出现： ：

- 潜在位置聚集在低维子空间中： 如果节点在 $\mathbb{R}^d$中的位置近乎共面， X 的列就会近乎线性相关。
- 社区结构较弱的随机块模型： 对于具有 $X = ZB^{1/2}$ （成员矩阵 Z、块矩阵 B）的随机块模型， P的谱隙反映了 B的谱隙。当社群难以区分时， λ_d 较小。
- 稀疏网络： 如果 X 的所有条目都按 $\rho_n \to 0$ （稀疏度参数）的量级缩放，那么 $\lambda_d \sim \rho_n^2 \to 0$。

这点尤为重要：几何障碍不仅会随着 $\lambda_d \to 0$ （单射半径会随着$\sqrt{\lambda_d}$趋于零，且当 λ_{d-1} 也很小时，曲率会急剧增大）而加剧，ASE 的收敛速率也会变差（其变化尺度为$O(\sqrt{\log n/\lambda_d})$ [Cape 等人， 2019]）。因此，谱隙较小的构形在统计上（更难估计）和几何上（更难追踪规范）都会带来问题。我们在此强调这一预期提示，因为之后还会再次提到：在第 5 节和第 6节中，控制曲率和单射性的相同谱量，也会控制估计误差和费舍尔信息。

命题4 （曲率与谱隙）。配备普罗克鲁斯特斯度量的商流形 $\mathcal{B} = \mathbb{R}_*^{n\times d}/O(d)$的截面曲率 K （一种标量，用于度量二维截面的高斯曲率；不同于上文李代数值曲率二形式Ω ） 由黎曼浸没的 ***O'Neill***公式给出 *[O' Neill， 1966]*： ¯对于正交归一的水平向量 ξ， $\bar{\eta} \in \mathcal{H}_X$，

$$K(\xi,\eta) = 3\|\mathcal{A}_{\bar{\xi}}\bar{\eta}\|^2 = \frac{3}{4}\|[\bar{\xi},\bar{\eta}]^{\mathcal{V}}\|^2 \tag{30}$$

其中， $\mathcal{A}_{\bar{\xi}}\bar{\eta} = \frac{1}{2}[\bar{\xi},\bar{\eta}]^{\mathcal{V}}$是 ***O'Neill*** A张量， $[\bar{\xi},\bar{\eta}]^{\mathcal{V}}$表示李括号的垂直分量。由于全空间 $\mathbb{R}_*^{n\times d}$是平坦的，所有底空间曲率都源自 A张量。

Massart 等人 [2019] 明确计算了该商空间的截面曲率：

- 如果 $d = 1$，则截面曲率恒为零（底空间是平坦的）。
- 如果 $d \geq 2$，则最小截面曲率为零，在水平向量场彼此对易时达到。
- 在 $[X]$ 处的最大截面曲率会在 $P = XX^\top$的两个最小特征值 λ_{d-1}、 λ_d 同时趋近于零时发散。

谱隙会引出两个相关但不同的几何障碍。曲率取决于 A<style id='8}张量，该张量涉及联络系数 $1/(\lambda_\iota + \lambda_\gamma)$ 对于 $\iota \neq \gamma$。由于联络是反对称的，最坏情况是 $1/(\lambda_{d-1} + \lambda_d)$：只有当两个最小特征值都趋近于零时，曲率才会发散。如果 $\lambda_d \to 0$但 λ_{d-1} 仍然有界，则曲率保持有限（流形在塌缩方向上局部看起来像圆柱面）。相比之下，在 $[X]$ 处的单射半径等于 $\sqrt{\lambda_d}$ [Massart 和 Absil， 2020] 并且只要$\lambda_d \to 0$ 单独满足条件，就会消失：超过这个半径后，测地线不再是长度最短路径，而且同一对点之间可能存在多条路径。

对于动力学估计，这两种现象都很重要。曲率大意味着即使是很短的路径也会产生显著的和乐，使得局部 Procrustes 对齐在中等距离范围内不一致。单射半径小意味着时间步之间的插值并不唯一，即使不存在曲率也是如此。 $d = 1$情形以一种有用的方式退化： $O(1) = \{\pm 1\}$ 是离散的，因此唯一的规范歧义是整体符号翻转；由于没有连续旋转可以累积，曲率为零。

3.3.8 黎曼结构

商空间 $\mathcal{B} \simeq \mathcal{E}/O(d)$继承了一个自然度量：

定义 7 （商度量）。 $[X],[Y] \in \mathcal{B}$ 之间的黎曼距离就是**Procrustes距离** ：

$$d_{\mathcal{B}}([X],[Y]) = \min_{Q\in O(d)} \|X - YQ\|_F \tag{31}$$

这将几何与计算联系起来：普罗克鲁斯对齐可计算底空间上的测地距离 [Mas- sart 和 Absil，2020]。全局单射半径 $\mathcal{B}$ 为零（因为 λ_d 可以任意小），这反映了商空间非平凡的拓扑结构。

4 从谱嵌入中恢复轨迹

上述障碍涉及动力学在理论上可能实现的情况。现在我们转向实际问题：即使动力学是可观测且可实现的，从数据中恢复轨迹仍然很困难，因为谱嵌入会在每个时间步引入任意的规范变换。

4.1 邻接谱嵌入会引入任意的规范变换

在每个时间步 t，邻接谱嵌入都会计算 $A^{(t)}$ 的特征分解。特征向量仅在符号上（以及在重特征子空间内的旋转上）确定。因此：

$$\hat{X}^{(t)} = X^{(t)} R^{(t)} + E^{(t)} \tag{32}$$

其中，$E^{(t)}$ 是统计噪声，$R^{(t)} \in O(d)$ 是规范因子。尽管 $R^{(t)}$ 是确定的（由特征值求解器的约定固定，例如第一个非零项的符号），但它在 t 中不连续：标准特征值求解器所采用的约定会在某个项穿过零或特征值顺序改变时发生跳变。这些跳变与动力学无关，因此 $R^{(t)}$ 的表现实际上就像在每个时间步从纤维 $O(d)$ 中任意选取一样。

关键在于：即使真实位置 $X^{(t)}$ 平滑演变，估计值 $\hat{X}^{(t)}$ 也会不规则地跳变，因为 $R^{(t)}$在不同时间之间彼此无关。原则上，当特征值是单一的且解析变化时，存在一种平滑的特征向量选取方式，但标准数值特征值求解器不会强制采用这种选取方式。

4.2 有限差分失效

考虑通过有限差分来估计速度：

$$\hat{\dot{X}}^{(t)} = \frac{\hat{X}^{(t+\delta t)} - \hat{X}^{(t)}}{\delta t} \tag{33}$$

代入受规范影响的估计值：

$$\hat{\dot{X}}^{(t)} = \frac{X^{(t+\delta t)} R^{(t+\delta t)} - X^{(t)} R^{(t)}}{\delta t} + O(E) \tag{34}$$

即使忽略噪声 E，主导这一结果的仍是规范跳变 $R^{(t+\delta t)} - R^{(t)}$，其大小为 $O(1)$，与 δt无关。随着“速度”按 $\delta t \to 0$发散，我们测量到的是规范跳变，而非动力学。

4.3 成对对齐与误差累积

可以尝试通过 Procrustes 对齐连续的嵌入：

$$Q^{(t)} = \arg\min_{Q \in O(d)} \|\hat{X}^{(t+1)} - \hat{X}^{(t)} Q\|_F \tag{35}$$

这个问题可以通过奇异值分解得到闭式解，从而找到使相邻帧匹配的最佳旋转。然而，顺序成对对齐存在根本性的局限：

1. 局部而非全局：每次对齐都会最小化相邻帧之间的误差，但并不能确保整条轨迹的一致性。2. 误差累积：ASE规范 $R^{(t)}$ 在每个时间步的行为类似于任意的 $O(1)$旋转，且在时间上互不相关。成对Procrustes针对连续帧之间的相对规范（$R^{(t)}$）$^{-1}R^{(t+1)}$，其误差由ASE噪声 $E^{(t)}$控制。当这些噪声引起的误差在各步之间大致独立时，对齐 T帧后的累积旋转误差会扩散式地变化，尺度类似于 $O(\sqrt{T}\sigma)$，其中σ 为每步噪声水平。 —

3. 和乐是一种独立的障碍：该 $O(\sqrt{T}\sigma)$ 扩散式误差是统计对齐难度。全纯障碍（第3.3节）是几何/拓扑性的，并且即使在完美数据（$\sigma = 0$）下也存在：对于具有非平凡整体单值性的动力学，无论每次成对对齐执行得多么精确，在闭合回路上都不存在全局一致的规范。这两种障碍会相互作用，但不应混为一谈。

4. 没有动力学约束： 即使序列完美对齐，也未必对应于具有动力学意义的提升：普罗克鲁斯对齐纯粹是几何操作，并不保证对任何结构化族都有 $X = f(X)$。

谱嵌入中的噪声： 随机点积图的谱嵌入存在估计误差 $\| \hat{X} - XQ\|_F = O_p(\sqrt{n})$ 其中存在某个$Q \in O(d)$，由此产生的每个节点误差为 $O_p(1/\sqrt{n})$ [Athreya 等人，2018年，Cape 等人， 2019]。对于相邻帧之间的成对普罗克鲁斯对齐，这意味着旋转估计误差取决于 n 以及帧之间的间隔。

当真实轨迹变化缓慢（ $\|X^{(t+1)} - X^{(t)}\|$较小）时，对齐的信噪比会下降。在这种情况下，我们试图从估计噪声的背景中检测微小的真实旋转。

4.4 为什么现有的联合嵌入方法无济于事

像 UASE 这样的联合嵌入方法 [Gallagher等人， 2021] 会同时嵌入所有时间点，生成规范一致的估计值。然而，它们假设的是不同的生成模型。

UASE 模型（多层随机点积图）：

$$P_{ij}^{(t)} = x_i \Lambda^{(t)} y_j^{(t)\top} \tag{36}$$

其中， x_i 是一个固定的“身份”，而 $y_j^{(t)}$ 是一个随时间变化的“活动”。

我们的模型（具有常微分方程动力学的时序随机点积图）：

$$P_{ij}^{(t)} = x_i^{(t)\top} x_j^{(t)} \tag{37}$$

其中，同一位置 $x_i^{(t)}$ 根据 $\dot{x}_i = f(x_i,$ .....)演化。

我们可以通过共享子空间假设来表达这种不匹配。UASE 假设拼接矩阵$[A^{(1)};\ A^{(2)};\ \ldots;\ A^{(T)}]$ 具有维度为 d的<样式 id='6'>公共不变列空间：所有时间切片共享相同的左奇异向量。在潜在位置的常微分方程动力学下，$P^{(t)} = X^{(t)} X^{(t)\top}$ 的列空间可能随时间演化；对于不可交换的矩阵族（尤其是拉普拉斯动力学），特征向量会发生旋转，通常不存在单一的不变子空间。因此，UASE 将静态平均子空间拟合到本质上不断演化的子空间上，把动力学信号压缩为固定基底中“活动分数”的变形。简而言之：联合嵌入强制要求共享子空间；由常微分方程驱动的时序随机点积图不一定具有这样的子空间（可交换/对称情形等特殊情况除外）。将 UASE 应用于我们模型生成的数据会扭曲所恢复的轨迹，把真实的位置演化归因于固定身份下的时变活动。

类似的问题也会影响 Omnibus 嵌入 [Levin等人、 2017]以及 COSIE [Arroyo等人： 2021]它们都假设存在一个共享子空间结构，而常微分方程动力学并不能保持这一结构。

4.5 为什么贝叶斯平滑方法并不充分

近期研究 [Loyal， 2025] 提出了使用潜在位置的层次先验对动态随机点积图进行贝叶斯推断。通过对连续差分施加先验（或者等价地，使用具有平滑核的高斯过程），得到的轨迹是平滑的：速度和加速度都有明确且连续的定义。

本节提出一个概念性观点：光滑性是动力学一致性的必要条件，但并非充分条件。随后，我们将形式化地说明这一间隙（通过一个零测集结果），并以实用的诊断方法作结。

然而，光滑性是动力学一致性的必要条件，但并非充分条件。常微分方程 $\dot{X} = f(X)$ 将速度约束为当前状态的函数；平滑先验约束速度连续变化，但并不要求 $\dot{X}(t)$ 由 $X(t)$ 决定。

为明确讨论范围：本节的批评针对的是仅平滑 模型。这类模型直接对轨迹（或其增量）设置先验，却未将 f 和常微分方程约束纳入生成模型。具有明确常微分方程／随机微分方程结构的贝叶斯推断则完全是另一回事（见下方对比表）。

定义 8（动力学一致性）。若存在 $f \in \mathcal{F}$，使得对所有 t，$X(t)$ 满足 $\dot{X}(t) = f(X(t))$，则称轨迹 $X(t)$ 关于函数类$\mathcal{F}$是**动力学一致的**。

层次贝叶斯先验确保 $X(t)$、$\dot{X}(t)$ 以及高阶导数都具有光滑性，但它们并不强制满足状态依赖约束 $\dot{X}(t) = f(X(t))$，其中任意 f 均如此。这一区别正是运动学正则性与动力学结构之间的间隙。

先验支撑与常微分方程解流形。对于给定的初始条件 $X(0) = X_0$ 和动力学 $\dot{X} = f(X)$，解 $X(t)$ 描出一条唯一的曲线：给定 (f, X_0)，轨迹不存在不确定性。所有常微分方程解（涵盖所有 $f \in \mathcal{F}$ 和 X_0）构成了无限维光滑路径空间中的一个有限维流形。层次光滑先验为所有光滑路径赋予正概率，而这类路径远多于仅满足某个常微分方程的路径。因此，在此类先验下，常微分方程解流形的测度为零。

命题5（有限维常微分方程族的测度为零）。设 $\mathcal{F} = \{f_\theta : \theta \in \Theta \subset \mathbb{R}^p\}$ 是 $\mathbb{R}^{n\times d}$ 上一个有限维向量场 C^1 族，并假设每个常微分方程 $\dot{X} = f_\theta(X)$ 对于每个初始条件，在 $[0,T]$上都有唯一解。设 μ 是 $C^k([0,T], \mathbb{R}^{n\times d})$ 上的一个具有全支撑的非退化高斯测度。定义

$$\mathcal{M}_{\mathcal{F}} = \{X(\cdot) : \dot{X} = f_\theta(X) \text{ for some } \theta \in \Theta\}. \tag{38}$$

那么 $\mu(\mathcal{M}_{\mathcal{F}}) = 0$。

证明。参见第 A节。 □

命题6（没有常微分方程约束先验时后验概率为零）。设 $\Pi_0 = \mu$ 是路径空间上的一个先验，且满足$\mu(\mathcal{M}_{\mathcal{F}}) = 0$。对于观测数据 y，定义后验分布如下

$$\Pi(A \mid y) = \frac{\int_A L(y \mid X)\, d\mu(X)}{\int L(y \mid X)\, d\mu(X)} \tag{39}$$

其中似然 L 可测，且归一化常数有限且为正。那么，对于每个这样的数据集 y，都有 $\Pi(\mathcal{M}_{\mathcal{F}} \mid y) = 0$。

证明。后验分布相对于 μ 绝对连续，其拉东–尼科迪姆导数正比于$L(y \mid X)$。因此，每个 μ 零测集都是后验零测集，尤其是 $\mathcal{M}_{\mathcal{F}}$。 □

因此，除非将常微分方程约束纳入先验/模型，否则贝叶斯更新无法为动力学一致轨迹赋予正的后验概率。对 $\mathcal{M}_{\mathcal{F}}$附近（而非其上）的渐近集中取决于模型误设，这是一个独立的问题。这就是插值与动力学学习之间的区别：两者都能生成穿过数据的平滑曲线，但只有后者遵循速度取决于状态这一约束。

用于检测动力学不一致的诊断方法。有几种检验可以区分平滑但动力学不一致的轨迹与真正的常微分方程解。首先是状态依赖性检验：对于自治动力学$\dot{X} = f(X)$，如果轨迹在两个不同时间经过相似的位置（$X(t_1) \approx X(t_2)$），那么推断速度也必须相似（$\dot{X}(t_1) \approx \dot{X}(t_2)$；平滑先验并不强制满足这一点，而在再次经过的状态处出现较大的速度差异，则表明存在动力学不一致。其次是常微分方程残差检验：拟合候选动力学 $\hat{f}$ 到平滑轨迹，并计算残差 $r(t) = \|\dot{\hat{X}}(t) - \hat{f}(\hat{X}(t))\|$。对于动力学一致轨迹，$r(t)$ 应当较小；而对于平滑但错误的轨迹，残差会系统性地偏大。第三，流一致性：从轨迹上的不同初始条件出发，对拟合得到的 $\hat{f}$ 进行向前积分；动力学一致轨迹会紧随积分流，而插值轨迹则会逐渐偏离。最后，中长期预测 可以区分这两种方法：运动学外推假设速度和加速度保持不变，而动力学外推则通过 $\hat{f}(X)$根据当前状态调整速度；只有后者在轨迹进入状态空间的新区域时仍能保持准确。

具有动力学一致性的替代方法。

随机微分方程表述 $dX = f(X)\, dt + \sigma\, dW$提供了一条原则性桥梁：Freidlin–Wentzell理论 [Freidlin与Wentzell, 1998] 表明，当 $\sigma \to 0$时，解集中在常微分方程解附近，速率函数为 $J_T(\phi) = \frac{1}{2}\int_0^T \|\dot{\phi}(t) - f(\phi(t))\|^2\, dt$。该速率函数正是动力学一致性惩罚，衡量一条路径偏离常微分方程解的程度。

注。**局限性（轨迹恢复）**。对齐谱嵌入以恢复连续时间轨迹仍然是一个困难的开放问题。上文讨论的方法针对相关任务，但它们并未解决

方法	光滑性	动力学一致性
层次高斯过程先验	是（C^k）	否：仅插值
随机微分方程 $dX = f(X)\,dt + \sigma\,dW$	否（Hölder $< 1/2$）	是，作为 $\sigma \to 0$
神经常微分方程	是（C^k）	是（按构造）
GP-ODE（NPODE）[Heinonen等人， 2018]	是（C^1）	是：将 f 学习为高斯过程
结构约束（我们的方法）	Yes	是：族 $\mathcal{F}$ 已强制施加

表 2：方法比较：光滑性与动力学一致性。

轨迹恢复的 r随机点积图中ODE驱动的潜在位置，且以可证明的方式。据我们所知，存在n o现有方法，可证 明能够从随机点积图潜在位置上的常微分方程动力学中恢复轨迹。

特别是，序贯对齐会在长轨迹上遭受误差累积（第 4.3 节）。此外，仅凭光滑性并不能保证动力学一致性，因此插值可能被误认为是真正的动力学学习（第 4.5 节）。最后，和乐意味着，即使不存在统计噪声，在底空间的回路上，全局一致的规范选择也可能不存在（第 3.3 节）。

这些考虑促使我们将重点放在动力学结构上，将其作为对齐的额外约束来源；在第7 节分析具体的动力学族之后，我们将在第 5 节回到这一构造性问题。

5 随机点积图上的动力学

第 3 节的几何框架刻画了规范自由度、曲率和和乐的抽象结构，而第 4 节展示了从谱嵌入中恢复轨迹的实际困难。我们分析具体的动力学族、它们可观测的后果，以及几何与统计方面的估计障碍之间的相互作用。

5.1 随机点积图动力学族

我们梳理了作用于随机点积图潜在位置的具体动力学族。这些族之后将作为对齐问题（第 7节）的归纳偏置；后续的和乐分析和统计分析取决于它们的具体结构。

线性动力学。 最简单的族是 $\dot{X} = NX$，其中 $N \in \mathbb{R}^{n\times n}$，其中 N_{ij} 决定了节点 j的位置如何影响节点 i的速度。根据命题 2，当且仅当 $X^\top NX$ 对称时，该动力学是水平的；一个充分条件是 $N = N^\top$，这使得 $X^\top NX$对所有 X都对称。诱导的 P-动力学为 $\dot{P} = NP+PN^\top$，当 $\dot{}$对称时，它可化为李雅普诺夫方程 $\dot{P} = NP + PN$。对称线性动力学有 N个自由参数（$n\,n+1$）$/2$。

多项式动力学于 P。 一个简约得多的族将常数 N 替换为概率矩阵上的多项式：

$$\dot{X} = N(P)X, \quad N(P) = \sum_{k=0}^{K} \alpha_k P^k \tag{40}$$

由于 P是对称的， $N(P)$自动对称，且无论系数 α_k如何，这些动力学都始终水平。各项具有清晰的解释：$P^0 = I$给出自作用动力学（衰减或增长）， $P^1 = P$编码直接邻居影响，而 P^k通过加权路径计数$(P^k)_{ij}$刻画 k跳效应。关键在于，$P = XX^\top$是规范不变的，因此 $N(P)$可由观测值计算，而无需求解对齐问题。该族仅有 $K+1$ 个独立于 n的参数，相较一般对称N 的 $n(n+1)/2$ 大幅减少，这对可辨识性至关重要。

图拉普拉斯动力学。 令 $\dot{X} = -LX$取 $L = D - P$且 $D = \mathrm{diag}(P\mathbf{1})$给出图上的扩散：逐分量地，$\dot{x}_i = \sum_j P_{ij}(x_j - x_i)$，因此每个节点趋向其邻居的加权平均。由于 L 是对称的，故该动力学是水平的。与多项式动力学不同，拉普拉斯算子 并非 关于 P 的多项式：度矩阵D 依赖于行和，而行和混合了多项式在 P 中无法获取的特征向量信息。这一区别将对和乐至关重要（第5.3节）。

消息传递动力学。更一般地，对于某个交互函数 g： $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$， $\dot{x}_i = \sum_j P_{ij} g(x_i, x_j)$。特殊情况包括邻居吸引（$g(x_i, x_j) = x_j$，得到 $\dot{X} = PX$）和拉普拉斯扩散（$g(x_i, x_j) =$

$x_j - x_i$）。水平性取决于 g；局部性约束（速度仅取决于由 P 加权的邻居）本身就是一种有用的结构性假设。

可观测的非水平动力学。并非所有可观测动力学都是水平的。质心环流 $\dot{x}_i = (x_i - \bar{x})A$ with $A \in \mathfrak{so}(d)$ 可分解为 $\dot{x}_i = x_i A - \bar{x}A$：第一项是纯规范项（定理 1），而共享漂移$-\bar{x}A$ 通常会改变点积，因此 P 通常不为零，除非处于特殊的对称构型中。然而，$X^\top \dot{X} = CA$ 其中 C 是位置的样本协方差；只有在额外的代数约束下（例如，与交换相容的结构），它才是对称的，因此该动力学通常是非水平的，并且轨迹沿纤维螺旋前进。类似地，微分旋转 $\dot{x}_i = x_i A_i$ with node-specific $A_i \in \mathfrak{so}(d)$ 给出 $P_{ij} = x_i(A_i - A_j)x_j^\top$，当旋转速率不同时，这是可观测的。

总结。

族	Form	水平?	参数
线性对称	$\dot{X} = NX, N = N^\top$	始终	$n(n+1)/2$
多项式	$\dot{X} = (\sum_k \alpha_k P^k)X$	始终	$K+1$
拉普拉斯算子	$\dot{X} = -LX$	始终	0（固定）
消息传递	$\dot{x}_i = \sum_j P_{ij} g(x_i, x_j)$	如果 g 对称	$\lvert\theta\rvert$
质心环流	$\dot{x}_i = (x_i - \bar{x})A$	No	$d(d-1)/2$
一般线性	$\dot{X} = NX$	如果 $X^\top NX$ 对称。	n^2

表 3：随机点积图动力学的各个族。多项式族始终是水平的，参数较少（$K+1$独立于 n），且规范不变（$N(P)$无需对齐即可计算）。质心环流是可观测量，但不是水平的。

5.2 可观测量动力学与李雅普诺夫方程

定理 1 刻画了规范自由度会导致什么内容丢失。我们刻画了什么得以保留：潜在空间动力学的规范不变特征，即投影到底空间 $\mathcal{B}$ 概率矩阵后仍然保留的内容。

关键观察结合了两个要素。首先，随机点积图模型给出了因式分解 $P = XX^\top$，因此P 继承了 X 的光滑性，并且乘积法则适用。其次，动力学模型（先从简单情形.

$X = NX$开始）规定了位置如何演化。两者结合，便可得到关于 P的封闭方程：

$$\dot{P} = \dot{X}X^\top + X\dot{X}^\top = NXX^\top + XX^\top N^\top = NP + PN^\top \tag{41}$$

当 N为对称矩阵时，这就成为李雅普诺夫方程 $P = NP + PN$。

这并非关于时序网络的一般性陈述。若没有低秩因式分解， $P(t)$ 就只是一个任意随时间变化的概率矩阵，其元素可能各自独立演化、遵循不同的矩阵常微分方程，或者根本不可微。（41）则是随机点积图作为动力系统这一视角所特有的：它展现了舍弃规范自由度后潜在常微分方程的形式。

李雅普诺夫算子 $\mathcal{L}_P$: Sym(n) $\to$Sym(n) 定义为 $\mathcal{L}_P(N) = NP + PN$，是作用于对称矩阵上的线性映射。为了分析它，我们转到 P 的特征基。设 $P = U\Lambda U^\top$为特征分解，并将 $\tilde{N} = U^\top NU$和 $\dot{P} = U^\top \dot{P} U$表示为该基下的表示，其中索引 ι， $\gamma \in \{1, \ldots, n\}$标记特征向量方向（而非节点）。在此基下， $\mathcal{L}_P$是对角作用的：$(\mathcal{L}_P(N))_{\iota\gamma} = (\lambda_\iota + \lambda_\gamma)N_{\iota\gamma}$。

对于随机点积图的情形，其中 $P = XX^\top$且 $X \in \mathbb{R}^{n\times d}$，矩阵 P的秩为 d，其特征值为 $\lambda_1 \geq \cdots \geq \lambda_d > 0$和 $\lambda_{d+1} = \cdots = \lambda_n = 0$。因此，李雅普诺夫算子在整个 Sym($n$) 上不可逆：当 ι、 $\gamma > d$、 时，方程 $(\lambda_\iota + \lambda_\gamma)N_{\iota\gamma} = P_{\iota\gamma}$化简为 $0 \cdot N_{\iota\gamma} = P_{\iota\gamma}$。可实现性约束（命题 1）强制要求该块满足 $P_{\iota\gamma} = 0$，因此该方程对 $N_{\iota\gamma}$没有提供任何信息。对于 $n = 100$和 $d = 3$，这意味着 $(n-d)(n-d+1)/2 = 4753$个 $n(n+1)/2 = 5050$ 项完全不受数据约束。

然而，结论很明确： N 中无法确定的部分，恰好是不影响给定 X下 $X = NX$ 的部分。

命题7 (部分李雅普诺夫可辨识性)。 设 $P = XX^\top$ 满足秩(P) $= d < n$，且 $V \in \mathbb{R}^{n\times d}$ 和$V_\perp \in \mathbb{R}^{n\times(n-d)}$ 分别张成 P 的值域和零空间。在 P 的特征基下（指标为 ι, γ），李雅普诺夫方程 $\dot{P} = NP + PN$唯一确定 $\tilde{N} = U^\top NU$的以下各块：

- Range-range block *($\iota, \gamma \le d$): $\tilde{N}_{\iota\gamma} = \tilde{\dot{P}}_{\iota\gamma}/(\lambda_\iota + \lambda_\gamma)$, giving $d(d+1)/2$ entries.*
- Cross block *($\iota \le d, \gamma > d$ or vice versa): $\tilde{N}_{\iota\gamma} = \tilde{\dot{P}}_{\iota\gamma}/\lambda_\iota$, giving $d(n-d)$ entries.*

零-零块 *($\iota, \gamma > d$)*：$(n-d)(n-d+1)/2$ 个元素仍不受约束。
已确定条目的总数为 $nd - d(d-1)/2$，与可实现切空间的维数相符。

证明。在特征基下，李雅普诺夫方程逐项给出 $(\lambda_\iota + \lambda_\gamma)N_{\iota\gamma} = P_{\iota\gamma}$。对于 $\iota, \gamma \le d$：$\lambda_\iota + \lambda_\gamma > 0$，因此 $N_{\iota\gamma}$ 唯一确定。对于 $\iota \le d, \gamma > d$：$\lambda_\iota + \lambda_\gamma = \lambda_\iota > 0$，因此 $N_{\iota\gamma}$ 唯一确定。对于 $\iota, \gamma > d$：$\lambda_\iota + \lambda_\gamma = 0$，而可实现性要求 $P_{\iota\gamma} = 0$，因此 $N_{\iota\gamma}$ 是自由的。计数如下：$d(d+1)/2 + d(n-d) = nd - d(d-1)/2$。 □

未确定的零-零块决定了 N 如何作用于与 col(X) 正交的向量。由于 X 在这个子空间中没有分量，NX 不依赖于该块：动力学 $X = NX$ 对其不敏感。李雅普诺夫方程恰好在对动力学重要的子空间上确定 N，并恰好在动力学无关的子空间上留下未受约束的部分。

总之，对称线性动力学可以从（P, P）中辨识出来，且仅差一个在动力学上无关的零-零块，整个过程中无需选择任何规范。该视角自然延伸到多项式动力学：如果 $X = N(P)X$ 其中 $N(P) = \sum_k \alpha_k P^k$，那么

$$\dot{P} = N(P)P + PN(P) \tag{42}$$

并且原则上可以通过将 α_k 拟合为对称矩阵 $\{P^k P + PP^k\}_{k=0}^{K}$ 张成空间中的元素来恢复系数 $\dot{P}$。这里，多项式结构发挥了实质性作用：它确定了 N 在 col(P) 之外的、原本欠确定的作用，从结构上消除了自由的零-零块。

然而，有两个根本性障碍阻碍了直接应用。

我们无法观测 $P(t)$。在实践中，我们观测到的是邻接矩阵 $A^{(t)}$，其服从 $A_{ij}^{(t)} \sim$ 伯努利($P_{ij}(t)$) 分布。逐条来看，$A_{ij}^{(t)}$ 是 $P_{ij}(t)$ 的无偏估计，其方差为 $P_{ij}(1 - P_{ij})$。如果每个时间点有 m 个独立图，则对 $\bar{A}^{(t)} = \frac{1}{m}\sum_{\ell=1}^{m} A_\ell^{(t)}$ 取平均会将逐条目方差降低 m 倍。

但李雅普诺夫反演（以及大多数几何性质）取决于谱，而不仅仅是其元素；谱精度 P取决于$\bar{}$和 n两者 m。由于特征分解是非线性的，A 的特征值/特征向量在有限样本量下会产生偏差。从渐近角度看，主导扰动的条件均值为零，残差偏差也很小 [Athreya 等人，2016年，Cape 等人，2019]，但命题7 中的逆映射要除以$(\lambda_\iota + \lambda_\gamma)$。这种除法恰好会放大谱隙较小时本就病态的方向。这与第 $1/\lambda_d$3.3节从几何角度揭示的灵敏度相同，只不过现在披上了代数的外衣。

在相关子空间上的参数数量 N。即使丢弃了动态上无关的零-零块，剩余的自由度仍可能很大：值域-值域块与交叉块合计有 $nd - d(d-1)/2$ 个元素。

如果 N 是常数，并且能够以某种方式获得一对估计良好的（P, P），那么约束数量就与未知量数量相匹配。但只要 N 随时间/状态变化，就需要轨迹中有足够的激励来区分各个参数。即使在多项式动力学中，系数 α_k 是常数，要识别这些系数也要求 $P(t)$在底空间中充分移动，从而区分不同的幂次 P^k。这正是多项式族具有吸引力的原因：$K+1$ 个参数（与 n无关）不仅简约，而且在结构上与数据有望约束的内容相契合。如果没有强结构假设或正则化，更丰富的族很快就会变得欠定。

注。 P-动力学视角与纤维丛视角分别处理同一问题的不同方面。命题 7 指出，从规范不变可观测量 P出发，原则上可以识别动力学，但动态无关的零-零块除外。纤维丛理论（第3.3节）探讨了更难的问题：当我们 必须 使用 X（例如，因为我们需要潜在位置用于解释或用于通用微分方程流程）时会发生什么，以及商空间的几何结构如何决定这样做的难度。这两种视角互为补充：代数结果告诉我们有哪些信息可用；几何框架则告诉我们提取这些信息需要付出什么代价。

欧几里得镜像 Athreya 等人的 [2025] 提供了一个具体的、可估计的规范不变轨迹摘要 . 它们之间的时间距离 ce

$$\hat{d}_{MV}(\hat{X}_t, \hat{X}_{t'}) = \min_{W \in O(d)} n^{-1/2} \|\hat{X}_t - \hat{X}_{t'} W\|_2 \tag{43}$$

是谱嵌入上的Procrustes距离。与我们的黎曼距离 $d_{\mathcal{B}}([X],\ [Y]) = \min_{Q\in O(d)} \|X - YQ\|_F$一样，它商掉了 $O(d)$规范，并在底空间 $\mathcal{B}$上给出一个度量。二者的区别在于范数的选择： $d_{\mathcal{B}}$ 使用Frobenius范数（从而得到我们的联络与曲率所源自的黎曼下沉度量），而 d_{MV} 使用谱范数（在 $1/\sqrt{n}$缩放之后，度量最坏情形 $-$的方向变化）。对于 $d = 1$，二者重合；对于 $d \geq 2$，最小化旋转以及由此得到的距离通常并不相同。

尽管存在这种差异，但两者共享同一个概念要点：它们都从全空间投影到 $\mathcal{B}$，从而舍弃ˆ规范。 Athreya 等人 [2025] 证明了 d_{MV} 能够一致地估计 d_{MV}，其依据是邻接谱嵌入，收敛速率为 $O(\log(n)/\sqrt{n})$（其定理6）；随后，他们又证明了根据估计距离矩阵得到的 CMDS 镜像估计具有一致性（其定理7），为规范不变量提供了具体的估计保证。然而，从动力学学习的角度来看，任何标量时间间距离（无论基于 Frobenius 范数还是谱范数）只能反映移动了多远， P 却无法反映如何移动：它无法恢复生成元 N 或李雅普诺夫结构

. $P = NP + PN$。将镜像的估计保证从标量距离推广到满矩阵 P，是一个有趣的开放问题；这将把他们的框架与命题 7中的李雅普诺夫反演联系起来。

5.3 水平动力学族的和乐（已证明/条件性/猜想性）

全纯障碍（第3.3节）提出了一个具体问题：对于哪些水平动力学族，和乐为平凡和乐，而对于哪些则为非平凡整体单值性？这一区分具有实际意义：平凡和乐意味着全局一致的规范可以原则上实现（换言之：对齐“仅仅”是统计性的）；非平凡整体单值性意味着，即便完美的局部对齐，也可能在环上累积全局漂移。

如表1所总结，本节包含：

1. 已证明的局部判据以及 $d = 2$完整限制全纯群推论，2. 针对一般 d的
条件定理/命题，3. 以及针对 $d \geq 3$的猜想性一般完整和乐推广。

在本节中，我们针对两个动力学族明确阐述该机制。要点在于交换/非交换的对比：多项式动力学使用沿轨迹彼此交换的生成元（因此沿该路径没有曲率，而且一开始就不存在回路），而拉普拉斯型生成元可能无法交换；每当投影交换子判据非零时，曲率为正，非平凡整体单值性便成为可能。

来自不可交换生成元的曲率。考虑水平动力学 $\dot{X} = M(X)X$，其中 $M(X)$ 对称，因此 $X^\top \dot{X} = X^\top M(X)X$ 对称，且该动力学为水平动力学。在两个不同的时间 t_1，t_2，生成元$M_1 = M(X(t_1))$ 和 $M_2 = M(X(t_2))$ 均为对称矩阵。交换子 $[M_1, M_2] = M_1M_2 - M_2M_1$ 是反对称矩阵（因为转置会颠倒顺序）。

相应水平向量场 $\bar{\xi}_i(X) = M_iX$ （常系数延拓，由于每个 M_i 均为对称矩阵，故为水平）的李括号为：

$$[\bar{\xi}_1, \bar{\xi}_2](X) = (M_2M_1 - M_1M_2)X = -[M_1, M_2]X \tag{44}$$

令 $S = -[M_1, M_2]$（反对称），则李括号为 $[\xi_1, \xi_2](X) = SX$。这并非自动属于垂直方向：垂直方向是右乘 $X\Omega$，而 SX 是左乘。只有在特殊情况下二者才会重合（大致而言，即当 S 保持 col(X) 不变时）。

但对于曲率，我们只需要垂直分量 $[\bar{\xi}_1, \bar{\xi}_2]^{\mathcal{V}}$。将 $Z = SX$投影到 $\mathcal{V}_X = \{X\Omega : \Omega \in \mathfrak{so}(d)\}$，得到 $X\Omega^*$，其中Ω^* 满足李雅普诺夫方程

$$G\Omega^* + \Omega^*G = 2\,\mathrm{skew}(X^\top Z) = 2X^\top SX \tag{45}$$

其中 $G = X^\top X$ （见第B节）。由于 G 是正定的，因此该方程有唯一解。在 G的特征基中：

$$\Omega^*_{\iota\gamma} = \frac{2(X^\top SX)_{\iota\gamma}}{\lambda_\iota + \lambda_\gamma} \tag{46}$$

因此，垂直分量为零，当且仅当投影交换子为零： $X^\top[M_1, M_2]X = 0$.

命题8（投影反对称作用的一般非消失性）。设 $n \geq d \geq 2$ 且 $S \in \mathfrak{so}(n)$ 非零。定义

$$\mathcal{N}_S := \{X \in \mathbb{R}^{n\times d}_* : X^\top SX = 0\}. \tag{47}$$

那么 $\mathcal{N}_S$ 是一个恰当的 $\mathbb{R}^{n\times d}_*$的r实代数子集。特别地，它具有空内部，且勒贝格测度ze ro.

证明。 $X^\top SX$ 的每个元素都是 X的元素的多项式，因此 $\mathcal{N}_S$ 是代数的。它是真的，因为 $S \neq 0$ 蕴含存在指标 $i \neq j$ 满足 $S_{ij} \neq 0$。选取 $x_1 = e_i$ 和 $x_2 = e_j$；则$x_1^\top S x_2 = S_{ij} \neq 0$，因此对于 $X_0 = [x_1, x_2, x_3, \ldots, x_d]$，将其余各列选取为使其具有满列秩，$(X_0^\top S X_0)_{12} \neq 0$。故 X_0位于 $\mathcal{N}_S$之外，从而 $\mathcal{N}_S$ 是真的。欧几里得空间的真实代数子集具有空内部且测度为零。 □

因此， $X^\top[M_1, M_2]X = 0$严格弱于 $[M_1, M_2] = 0$；对于固定的非零交换子，它仅在 X中的一个代数例外集上失效。将该例外集与有效域 $\mathcal{E}$ 的内部相交，不会改变这一解释：在模型定义域内部，失效集仍是一个真子集。

回顾命题 4，根据 O'Neill公式 [O'Neill, 1966]由 $\bar{\xi}_1$和 $\bar{\xi}_2$在 $\mathcal{B}$中的投影所张成的二维平面的截面曲率为：$K(\xi_1, \xi_2) = \frac{3}{4}\|[\bar{\xi}_1, \bar{\xi}_2]^{\mathcal{V}}\|^2 / (\|\bar{\xi}_1\|^2\|\bar{\xi}_2\|^2 - \langle\bar{\xi}_1, \bar{\xi}_2\rangle^2)$。 $1/(\lambda_\iota + \lambda_\gamma)$中 Ω^*的因子将曲率直接关联到纤维丛的联络系数：放大规范敏感性的相同分母也会放大垂直括号。

由于 $\mathbb{R}_*^{n\times d}$是平坦的（它是欧几里得空间的一个开子集），基空间曲率完全来自垂直括号。

命题 9 （水平动力学的曲率判据）。 对于水平动力学 $\dot{X} = M(X)X$，其中 M对称，沿 $\mathcal{B}$ 中轨迹的截面曲率消失，当且仅当投影交换子消失：$X^\top[M(X(t_1))\ M(X(t_2))]X = 0$对沿轨迹的所有 t_1、t_2成立。一个充分条件是完整生成元可交换：$[M(X(t_1))\ M(X(t_2))] = 0$。当投影交换子非零时，曲率严格为正，且有：

$$\|[\bar{\xi}_1, \bar{\xi}_2]^{\mathcal{V}}\|^2 = 4\sum_{\iota<\gamma} \frac{\lambda_\iota\lambda_\gamma}{\lambda_\iota + \lambda_\gamma}[(U^\top[M_1, M_2]U)_{\iota\gamma}]^2 \tag{48}$$

其中 $P = U\Lambda U^\top$是特征分解。$1/(\lambda_\iota + \lambda_\gamma)$因子将曲率直接与纤维丛的联络系数联系起来。

证明。常系数水平延拓 $\xi_i(Y) = M_i Y$的李括号为 $[\xi_1, \xi_2](X) = -[M_1, M_2]X$（线性向量场的标准交换子）。$SX$（其中 $S = -[M_1, M_2]$ 反对称）的垂直投影为$X\Omega^*$，其中Ω^*满足 $G\Omega^* + \Omega^* G = 2X^\top SX$（B节）。当且仅当 $X^\top[M_1, M_2]X = 0$时，该投影消失。范数公式由在 G的特征基中计算的 $\|X\Omega^*\|^2 = \mathrm{tr}(\Omega^{*\top} G\Omega^*)$得出（计算见C节）。 □

多项式动力学：平凡和乐。 对于多项式动力学 $\dot{X} = N(P)X$，其中 $N(P) = \sum_{k=0}^K \alpha_k P^k$，诱导的 P-动力学为：

$$\dot{P} = N(P)P + PN(P) = 2\sum_{k=0}^K \alpha_k P^{k+1} \tag{49}$$

这是仅关于 P 的一个多项式。将 $P = U\Lambda U^\top$ 写成其特征分解后，每一幂 $P^k = U\Lambda^k U^\top$ 都共享相同的特征向量 U。因此， $\dot{P} = U(2\sum \alpha_k \Lambda^{k+1})U^\top$ 同样具有特征向量 U：只要 $P(0)$ 的谱是单谱（所有特征值互不相同）， P 的特征向量在多项式动力学下便保持静止。

轨迹 $P(t)$ 位于 $\mathcal{B}$ 的一个 d 维子流形上，该子流形仅由特征值 $(\lambda_1(t)\ldots, \lambda_d(t))$ 参数化，且每个特征值独立演化：

$$\dot{\lambda}_\iota = 2\sum_{k=0}^K \alpha_k \lambda_\iota^{k+1} \tag{50}$$

由此立即得出两个推论。

命题 10 （多项式动力学：平凡和乐）.多项式动力学 $\dot{X} = N(P)X$，其中 $N(P) = \sum \alpha_k P^k$，且初始谱为单谱（$\lambda_1(0) > \lambda_2(0) > \cdots > \lambda_d(0) > 0$），具有平凡和乐，其具体含义如下：

1. 生成元 $N(P(t_1))$与 $N(P(t_2))$对所有 t_1，t_2（它们共享特征向量）均可交换，因此曲率沿该轨迹由命题 9 可知为零。*2.* 每个特征值满足一维自治常微分方程，因此非常值特征值轨迹是单调的。特别地，所诱导的 $P(t)$ 不能在 $\mathcal{B}$ 中描绘非平凡闭环。

证明。 对于（1）：由于 P 的特征向量是静止的， $N(P(t)) = U(\sum \alpha_k \Lambda(t)^k)U^\top$ 对所有 t 成立，且具有相同的 U。任意两个对角矩阵均可交换，因此 $[N(P(t_1))\ N(P(t_2))] = U[\sum \alpha_k \Lambda(t_1)^k,\ \sum \alpha_k \Lambda(t_2)^k]U^\top = 0$。

对于 (2)：每个 $\lambda_\iota(t)$ 都满足自治常微分方程 $\dot{\lambda}_\iota = f(\lambda_\iota)$，其中 $f(x) = 2\sum \alpha_k x^{k+1}$。根据常微分方程解的唯一性，如果对于 $t_0 \neq t_1$， $\lambda_\iota(t_0) = \lambda_\iota(t_1)$，那么 λ_ι 必须为常数。因此，非恒定解不可能是周期解：它们无法回到先前的取值。因此，特征值空间中不可能存在非平凡闭环。由此可知， $\mathcal{B}$ 中也不可能存在非平凡闭环。

□

节点的多项式动力学表现如何。关于 P 的特征向量处于静止状态这一论述，既需要严谨的论证，也需要具体的解释。

特征向量为何固定不变。设 $P(t) = U(t)\Lambda(t)U(t)^\top$，并令 $\Omega = U^\top \dot{U} \in \mathfrak{so}(d)$ 为特征向量旋转速率。对 $P = U\Lambda U^\top$ 求导并投影到特征基上，得到对于 $\iota \neq \gamma$，$(U^\top \dot{P} U)\iota\gamma = \Omega\iota\gamma(\lambda_\gamma - \lambda_\iota)$。由于 $\dot{P} = 2\sum \alpha_k P^{k+1}$ 在特征基下为对角矩阵，左侧为零，因此对所有 $\iota \neq \gamma$，$\Omega\iota\gamma(\lambda_\gamma - \lambda_\iota) = 0$。接下来需要证明特征值始终互不相同。所有 d 特征值都满足相同的自治常微分方程 $\lambda = f(\lambda)$，其中$f(x) = 2\sum \alpha_k x^{k+1}$，彼此仅初始条件不同。如果 $\lambda_\iota(0) \neq \lambda_\gamma(0)$，那么根据常微分方程解的唯一性，对于所有 t，$\lambda_\iota(t) \neq \lambda_\gamma(t)$：如果它们曾经相等，就会是同一个解，这与它们不同的初始条件相矛盾。因此 $\lambda_\gamma - \lambda_\iota \neq 0$，这迫使 $\Omega = 0$： P 的特征向量确实保持静止，而不仅仅是一阶近似下静止。

简单谱假设至关重要：如果对于某个 $\iota \neq \gamma$， $\lambda_\iota(0) = \lambda_\gamma(0)$，那么重特征子空间中的特征基并不唯一，而多项式动力学会保留这种简并（重复特征值以相同方式演化，使简并在所有时间内持续存在）。在这种情况下，特征向量的“平稳性”定义不明确，因为基本身并不唯一。例如，对于块大小相同的块模型，重复特征值源于模型对称性。这是初始条件空间中的一个零测度条件。

这对节点意味着什么。在典范规范下，节点 i 的潜在位置为 $x_i(t) = (U_{i1}\sqrt{\lambda_1(t)}, \ldots, U_{id}\sqrt{\lambda_d(t)})$。载荷 $U_{i\iota}$（可理解为节点 i 在特征向量方向 ι 上的“成员权重”）保持不变；发生变化的是每个方向的尺度 $\sqrt{\lambda_\iota(t)}$。

由于当 $K \geq 1$ 时，特征值常微分方程 $\dot{\lambda} = f(\lambda)$ 是非线性的，不同模长的特征值会以不同速率演化，因此比值 $\lambda_\iota(t)/\lambda_\gamma(t)$ 会随时间变化。其对节点位置的影响，是沿固定轴对潜在空间进行各向异性缩放：在 $d = 2$ 中，点云发生形变，仿佛内接于一个椭圆之中；椭圆的轴不发生旋转，但离心率会变化。$\mathbb{R}^d$ 中的节点方向确实会改变（并非仅仅是模长被缩放），但这种变化受到严格约束：所有形变都由 d 标量函数 $\lambda_\iota(t)$刻画。

例外情况是线性动力学（$K = 0$， $f(\lambda) = 2\alpha_0\lambda$），此时 $\lambda_\iota(t) = \lambda_\iota(0)e^{2\alpha_0 t}$，所有特征值都以相同的指数速率增长或衰减。在这种情况下，比值保持不变，点云进行各向同性缩放，且 $\mathbb{R}^d$ 中的节点角位置真正保持固定。对于二次或更高阶动力学（$K \geq 1$）， f 的非线性会使较大的特征值与较小的特征值以不同方式演化，从而真实地改变点云的几何形状，同时保留定义社区对齐的特征向量轴。不过需要注意，保留这些轴并不保证保留社区可分性：如果特征值比值发散（例如，当 λ_1 增长时， $\lambda_d(t) \to 0$），点云会塌缩到一个低维子空间中，而原本沿 λ_d 方向可区分的社群可能变得无法区分。

拉普拉斯动力学：非平凡整体单值性。图拉普拉斯动力学 $\dot{X} = -LX$，其中 $L = D - P$，且 $D = \mathrm{diag}(P\mathbf{1})$，提供了一个具有非平凡整体单值性的水平动力学的自然例子。生成元 $-L = P - D$是对称的，因此该动力学是水平的。但 L 不是 P的多项式：度矩阵 D 取决于 P的行和，而这些行和混合了多项式作用于 P时无法获取的特征向量信息。

有一个值得注意的例外：若 D 是单位矩阵的标量倍数（$D = cI$，即图是所有节点度都相等的正则图），那么 $L = cI - P$ 是 P的仿射函数，拉普拉斯动力学便约化为多项式族。因而，正则图（更一般地说， P的行和为常数的构形，即 $P\mathbf{1} = c\mathbf{1}$）在拉普拉斯动力学下具有平凡和乐。下文关于非平凡整体单值性的结果要求 D 不是标量：图的结构异质性正是整体单值性的来源。

拉普拉斯流下的 P-动力学为：

$$\dot{P} = -LP - PL = -(D - P)P - P(D - P) = 2P^2 - DP - PD \tag{51}$$

项 $2P^2$保持 P的特征向量（它是关于 P的多项式），但 $DP + PD$并非如此：在特征基$P = U\Lambda U^\top$下，度矩阵 $D = \mathrm{diag}(U\Lambda U^\top\mathbf{1})$在节点基中是对角的，而不是在特征基中。因此 $U^\top DU$通常为满矩阵，且 P具有横向于 P的特征向量分量。

推论： P的特征向量在拉普拉斯动力学下旋转，轨迹在 $\mathcal{B}$ 中扫出横向于特征值参数化子流形的面积，且命题9 保证在轨迹的任意部分上具有正曲率，只要 $[L(P(t_1))\ L(P(t_2))] \neq 0$。

节点的拉普拉斯动力学表现。 在节点层面，拉普拉斯动力学与多项式动力学形成鲜明对比。在拉普拉斯扩散下，每个节点的速度 $\dot{x}_i = -\sum_j L_{ij}x_j$ 是其邻居位置的加权平均（权重来自 P），再减去由该节点自身的度产生的回复项。这就像在潜在位置上进行热扩散：高阶节点受到更强的均值回复，而扩散则通过网络结构将所有位置耦合起来。由于这种耦合取决于 P的行和，会混合来自所有特征向量方向的贡献，因此 P 的特征向量会随时间旋转。对于 $d = 2$的椭圆类比而言，不仅离心率会改变，椭圆的轴本身也会旋转。社区结构不再是静态的；节点在不同潜在维度上的隶属权重确实会演化，从而产生更丰富的动力学，并导致非平凡整体单值性。

命题 11（拉普拉斯动力学：非平凡整体单值性的局部对易子准则）。 设 $X(t)$ 满足拉普拉斯动力学 $\dot{X} = -L(P)X$，其中 $P = XX^\top$ 且 $L = D - P$。假设存在 $t_* \in [0, T)$ 使得

$$C_* := X(t_*)^\top [L(t_*), \dot{L}(t_*)]X(t_*) \neq 0. \tag{52}$$

那么存在 $\epsilon_0 > 0$，使得对于所有 $0 < |h| < \epsilon_0$，

$$X(t_*)^\top [L(t_*), L(t_* + h)]X(t_*) \neq 0. \tag{53}$$

因此，根据命题9，沿该轨迹的截面曲率在该局部双时间跨度上严格为正。特别地，在该区域的一个邻域内，存在受限和乐非平凡的可缩环路。

证明。 由于 L 关于 t可微，根据泰勒展开可得 $L(t_* + h) = L(t_*) + h\dot{L}(t_*) + O(h^2)$。因此$[L(t_*)\ L(t_* + h)] = h[L(t_*)\ \dot{L}(t_*)] + O(h^2)$。分别左乘和右乘固定矩阵 $X(t_*)^\top$ 和 $X(t_*)$，得到

$$X(t_*)^\top [L(t_*), L(t_* + h)]X(t_*) = hC_* + O(h^2). \tag{54}$$

因为 $C_* \neq 0$ ，由连续性可知，存在 $\epsilon_0 > 0$，使得对所有 $0 < |h| < \epsilon$，该表达式均非零 $_0$.

应用命题 9 于 $M_1 = -L(t_*)$ 和 $M_2 = -L(t_* + h)$（二者均为对称矩阵）：非零投影交换子意味着相应的双时间张成空间具有严格正的截面曲率。根据安布罗斯–辛格定理，最终的和乐结论随之成立：非零曲率意味着受限和乐非平凡，因此存在具有非恒等和乐的可缩环路。 □

上述判据可沿轨迹逐点检验。在初始化时，该条件也具有一般性 .

命题 12（初始化时局部对易子条件的一般性）。 固定 $n \geq 3$ 和 $2 \leq d \leq n$，并定义

$$\Psi(X) := X^\top [L(X), \dot{L}(X)]X \tag{55}$$

其中 $P = XX^\top$, $D = \mathrm{diag}(P\mathbf{1})$, $L = D - P$, $\dot{P} = 2P^2 - DP - PD$, *and* $\dot{L} = \mathrm{diag}(\dot{P}\mathbf{1}) - \dot{P}$. *Then*

$$\mathcal{Z} := \{X \in \mathbb{R}_*^{n\times d} : \Psi(X) = 0\} \tag{56}$$

是 $\mathbb{R}_*^{n\times d}$的真实代数子集。因此， $\mathcal{Z}$ 空内部且勒贝格测度为零；换言之，对于一般的初始条件，命题 *11*的假设在 $t_* = 0$处成立。限制到有效的随机点积图状态后，与 $\mathcal{E}$取交集仍得到模型定义域内部中的真例外集。

证明。 $\Psi(X)$的每个元素都是 X 各元素的多项式（所有运算均为求和与乘积），因此 $\mathcal{Z}$ 是代数的。为证明其为恰当的，只需给出一个满秩的 X，使得 $\Psi(X) \neq 0$。

对于 $n = 3$、 $d = 2$，取

$$X_0 = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 2 \end{pmatrix}. \tag{57}$$

将上述公式直接代入可得$(\Psi(X_0))_{12} = 2/3^7 \neq 0$，因此 $\Psi(X_0) \neq 0$.

对于一般的 $n \geq 3$ 和 $2 \leq d \leq n$，将这个二维见证嵌入前两列和前三行，其余元素设为零，然后在第 3、.....、 d列中加入任意小的扰动，使矩阵

达到满秩。由于 Ψ 在嵌入见证处连续且非零，扰动足够小时，该值仍然非零。因此，对于每个 $n \geq 3$和 $2 \leq d \leq n$，都存在一个满秩见证，所以 $\mathcal{Z}$ 是真代数子集。

欧几里得空间中的真代数子集内部为空，且勒贝格测度为零。 □

对于 $d = 2$，这个结果尤其强。

推论 2 (二维：由一个非零曲率值得到完整受限全纯群)。 设 $d = 2$，并假设命题 *11* 的假设在某个时间 t_*成立。那么，该丛的受限全纯群为 $\mathrm{SO}(2)$。等价地，对于每个角度 $\phi \in [0, 2\pi)$，存在一个位于 $\mathcal{B}$ 中的可缩环路，其水平输运产生的和乐旋转角为 ϕ。

证明。 李代数 $\mathfrak{so}(2)$ 是一维的，由单个生成元 $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$张成。根据命题 11，曲率二形式产生了 $\mathfrak{so}(2)$中的一个非零元素。单个非零元素即可张成这一维李代数，因此根据 Ambrose–Singer 定理，holonomy 代数就是整个 $\mathfrak{so}(2)$，而受限全纯群（可缩环路的 holonomy）为$\mathrm{SO}(2)$。 □

这意味着，对于拉普拉斯动力学下的秩二潜空间，障碍并不仅仅是离散的符号翻转（如最简单情形中那样），而是任意角度的连续旋转：根据环路的不同，可以得到任意程度的规范漂移。从对齐的角度来看，这几乎已是最坏的情形。

注。 **启发式定量估计（二维）**。 对于 $d = 2$，环路 γ 在 $\mathcal{B}$ 中的和乐通常由截面曲率的面积积分近似： $\phi(\gamma) \approx \iint_\Sigma K\, dA$ 其中 Σ 是以 γ为边界的曲面， K 是截面曲率，且 dA 是面积元。这是一种直觉辅助，而非手稿其他部分所使用的定理。结合 Massart等人 [2019], ，这表明经过秩亏状态（尾部特征值较小）附近的轨迹可能累积更大的和乐。

高维情形（$d \geq 2$）。 对于一般的 d，李代数 $\mathfrak{so}(d)$ 的维数为 $d(d-1)/2$。曲率二形式产生和乐代数元素 $\Omega^* \in \mathfrak{so}(d)$，满足 $\Omega^*_{\iota\gamma} = 2(X^\top [M_1, M_2] X)_{\iota\gamma}/(\lambda_\iota + \lambda_\gamma)$。当轨迹经过不同状态时，这些元素在 $\mathfrak{so}(d)$ 中的方向可能改变，其张成也可以增大。

定理 3 (由曲率张成条件推出完整受限全纯群)。 设 $d \geq 2$并固定一个基点 $p \in \mathcal{B}$。假设从 p 可达的曲率值所生成的李代数张成整个 $\mathfrak{so}(d)$，即

$$\mathrm{span}\{\mathrm{Ad}_{\tau_{q\to p}}(\Omega_q(u,v)) : q \textit{ horizontally reachable from } p,\ u, v \in \mathcal{H}_q\} = \mathfrak{so}(d). \tag{58}$$

那么， p处的受限全纯群为： $\mathrm{Hol}^0_p = \mathrm{SO}(d)$。

证明。 根据安布罗斯–辛格定理，Hol^0_p 的李代数恰好是上述张成空间。在该假设下，$\mathrm{Lie}(\mathrm{Hol}^0_p) = \mathfrak{so}(d)$。群 Hol^0_p 是连通的，且包含于 $O(d)$；因此它包含于 $\mathrm{SO}(d)$ 的单位连通分支中。$\mathrm{SO}(d)$ 的连通李子群若其李代数为 $\mathfrak{so}(d)$，则必为 $\mathrm{SO}(d)$ 本身。 □

该定理is lates 给出了 $d \geq 3$所缺的确切要素：对曲率生成方向的秩条件 s.

命题 13 (固定基点处的有限时间秩判据)。 设 $X(t)$ 满足 $\dot{X} = -L(P)X$，并固定 t_0，使得 $P(t_0)$ 的秩为 d。假设存在时间 s_1、……、s_m，满足 $m = d(d-1)/2$，使得这些反对称矩阵

$$B_a := X(t_0)^\top [L(t_0), L(s_a)] X(t_0) \in \mathfrak{so}(d), \quad a = 1, \ldots, m, \tag{59}$$

线性无关。那么，在基点 $p = [X(t_0)]$，，受限全纯群为 $\mathrm{SO}(d)$。

证明。根据命题 9 和 B节，每个 B_a 都通过李雅普诺夫映射 $B_a \mapsto \Omega^*_a$，在同一基点 p 处确定一个曲率值，并通过正因子 $1/(\lambda_\iota + \lambda_\gamma)$ 进行逐分量缩放（在 t_0 处计算）。该作用于 $\mathfrak{so}(d)$ 上的线性映射是可逆的，因此 (B_a) 的线性无关性蕴含了 (Ω^*_a) 的线性无关性。由于 m =dim($\mathfrak{so}(d)$)，这些曲率值在 p处张成 $\mathfrak{so}(d)$。因此，定理 3 的曲率张成假设成立（取 $q = p$，所以不存在传输问题），且 $\mathrm{Hol}^0_p = \mathrm{SO}(d)$。 □

注。 **关于一般性的启发式说明**。 对于具有 $n \gg d$ 的拉普拉斯动力学，人们预期，不同时间的度矩阵会使有限时间秩判据在大多数轨迹上成立。要将其推广为适用于所有 $d \geq 3$ 的定理，需要通过横截性论证来控制对称性分层。因此，我们将完全一般性的陈述保留为猜想。

猜想 1（高维完全整体性）。 对于 $d \geq 3$、 $n > d$，以及满足 $D(P(0))$ 非标量（非正则图）的初始条件，拉普拉斯动力学的受限全纯群 $X = -L(P)$ X 为 SO(d)。

对猜想 1 的证明将意味着在一般情形下证明上述秩判据。等价地，必须证明投影交换子 $X^\top[L(P(t_1))\ L(P(t_2))]X$ 张成 $\mathfrak{so}(d)$ 对一般轨迹成立。主要障碍是结构对称性：如果度算子 $P \mapsto D(P)$ 在 $\mathfrak{so}(d)$中保持一个真不变子空间，则生成的 holonomy 代数可能退化为一个真子代数。正则/可平衡族是这种简并的典型例子。

同样的论证也适用于消息传递动力学 $\dot{x}_i = \sum_j P_{ij} g(x_i, x_j)$，其中的对称生成元并非 P 的多项式。

总结： 和乐障碍清晰地揭示了水平动力学族之间的差异。多项式动力学作用于 P的谱结构，并通过可交换生成元运行，因此保持特征向量不变并具有平凡和乐：无需拓扑障碍即可实现全局规范一致性。拉普拉斯算子动力学和消息传递动力学通过不可交换生成元混合谱结构与空间结构；当命题 9 的投影交换子判据被触发时，便可能产生非平凡整体单值性：即使沿水平轨迹，规范漂移也会不断累积，并且任何局部对齐过程都无法在环上实现全局一致性。
这种区别具有实际意义：对于多项式动力学，构造性对齐问题（第 7 节）纯粹是一个统计学挑战（克服伯努利噪声）；对于拉普拉斯动力学，它同时也是一个统计学和 拓扑学挑战。

注。 **安布罗斯–辛格视角（快速直觉构建）**。 局部曲率与整体和乐之间的关系由 Ambrose–Singer 定理 [Kobayashi 和 Nomizu， 1963]形式化如下：受限全纯群在 $\mathfrak{hol}_p$ 处的李代数 p 由曲率端omorphism $\Omega_q(u, v)$ 生成；这些曲率端omorphism取值于从 q出发、可通过水平曲线到达的点，随后将其平行输运回 p。

对于 $d = 2$， $\mathfrak{so}(2)$没有真非零子代数，因此单个非零曲率元素就会强制产生完整限制全纯群$\mathrm{Hol}_p^0 = \mathrm{SO}(2)$（如上所用）。
对于 $d \geq 3$，一个曲率值只能给出一个方向，而处于规模大得多的代数之中。整体群要求出现括号生成现象：沿轨迹遇到的曲率方向不能全部位于某个可交换子代数内。由于度矩阵 $D(X)$依赖于潜在位置的非线性函数，诱导曲率元素的方向通常会随着轨迹演化而旋转，这正是猜想 1 背后的混合机制。

最后，谱隙再次出现：正如命题 9 所示，曲率幅度按 $1/(\lambda_\iota + \lambda_\gamma)$ 加权。因此，较小的谱隙不仅会放大统计噪声（通过 ASE误差界），还会使和乐效应变得更大，因为曲率生成的规范旋转会在病态方向上发散。

6 信息论极限

第 3 节中的几何障碍（规范自由度、曲率、和乐）、第 4 节中的轨迹恢复难题，以及第 5 节中的动力学分析，共同限制了原则上可学习的内容。除此之外，我们还考虑了统计 障碍：给定 T 含噪邻接矩阵，无论采用何种估计方法，我们能够以多高的精度估计控制动力学的参数?

答案揭示了一种对偶：控制几何难度的同一个谱隙 λ_d 也控制着统计难度，因此这两种障碍会相互强化。

6.1 动力学参数的费舍尔信息

考虑一个参数为 $\theta \in \mathbb{R}^k$、生成轨迹 $P(t;\ \theta)$ 的参数动力学族。我们观测到邻接矩阵 $\{A^{(t)}\}_{t=0}^T$，其元素服从伯努利($P(t;\ \theta)$) 分布；给定 $P(t;\ \theta)$ 时，所有元素均条件独立。对数似然可分解为：

$$\ell(\theta) = \sum_{t=0}^{T} \sum_{i<j} \left[A_{ij}^{(t)} \log P_{ij}(t;\theta) + (1 - A_{ij}^{(t)}) \log(1 - P_{ij}(t;\theta)) \right] \tag{60}$$

费舍尔信息矩阵 $\mathcal{I}(\theta) \in \mathbb{R}^{k\times k}$ 的元素为：

$$\mathcal{I}(\theta)_{ab} = \sum_{t=0}^{T}\sum_{i<j} \frac{\partial P_{ij}}{\partial \theta_a}\frac{\partial P_{ij}}{\partial \theta_b}\cdot\frac{1}{P_{ij}(1-P_{ij})} \tag{61}$$

其中 $P_{ij} = P_{ij}(t;\theta)$以及导数 $\partial P_{ij}/\partial\theta_a$ 刻画了动力学参数中的扰动如何传播到观测值。

6.2 灵敏度通过李雅普诺夫方程的传播

（61）中的关键量是灵敏度 $\partial P_{ij}/\partial\theta_a$。对于多项式动力学 $N(\theta) = \sum_{k=0}^{K}\theta_k P^k$， P的动力学为：$\dot{P} = 2\sum_{k=0}^{K}\theta_k P^{k+1}$。一个关键的简化来自命题 10： P的特征向量在多项式动力学下是静止的。记 $P = U\Lambda U^\top$，其中 $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_d)$且 $U\in\mathbb{R}^{n\times d}$固定，灵敏度$\partial P/\partial\theta_a = U(\partial\Lambda/\partial\theta_a)U^\top$在特征基下是对角的。

每个特征值满足 $\dot{\lambda}_\iota = 2\sum_{k=0}^{K}\theta_k\lambda_\iota^{k+1}$，因此：

$$\frac{\partial}{\partial t}\frac{\partial\lambda_\iota}{\partial\theta_a} = 2\lambda_\iota^{a+1} + \frac{\partial\dot{\lambda}_\iota}{\partial\lambda_\iota}\cdot\frac{\partial\lambda_\iota}{\partial\theta_a} \tag{62}$$

其中 $\partial\dot{\lambda}_\iota/\partial\lambda_\iota = 2\sum_{k=0}^{K}\theta_k(k+1)\lambda_\iota^k$是线性化特征值动力学。对于每一对 $(\iota,\ a)$，这是一个标量线性常微分方程，带有驱动项 $2\lambda_\iota^{a+1}$以及通过线性化动力学的反馈。
由此得到的观测概率灵敏度为：

$$\frac{\partial P_{ij}}{\partial\theta_a} = \sum_{\iota=1}^{d} U_{i\iota}U_{j\iota}\frac{\partial\lambda_\iota}{\partial\theta_a} \tag{63}$$

每个节点对 $(i,\ j)$ 接收来自全部 d 个特征值敏感度的信号，并由特征向量载荷 $U_{i\iota}U_{j\iota}$加权。

示例：线性动力学（$K=0$）。对于 $X = \alpha_0 X$（单参数），特征值动力学为 $\lambda_\iota = 2\alpha_0\lambda_\iota$，从而得到 $\lambda_\iota(t) = \lambda_\iota(0)e^{2\alpha_0 t}$。灵敏度为 $\partial\lambda_\iota/\partial\alpha_0 = 2t\lambda_\iota(t)$，因此：

$$\frac{\partial P_{ij}}{\partial\alpha_0} = 2t\sum_{\iota} U_{i\iota}\lambda_\iota(t)U_{j\iota} = 2tP_{ij}(t) \tag{64}$$

单个快照下的费舍尔信息 t为：

$$\mathcal{I}_t(\alpha_0) = 4t^2\sum_{i<j}\frac{P_{ij}(t)^2}{P_{ij}(t)(1-P_{ij}(t))} = 4t^2\sum_{i<j}\frac{P_{ij}(t)}{1-P_{ij}(t)} \tag{65}$$

对 T 个等间隔快照求和，得到 $\mathcal{I}(\alpha_0) = 4\sum_{t=0}^{T}t^2\sum_{i<j}P_{ij}(t)/(1-P_{ij}(t))$；当边概率远离 0 和 1 时，其尺度为$O(T^3\cdot n^2)$。

高阶项与谱隙。对于 $a\geq 1$，靠近谱隙的特征值对应的驱动项 $2\lambda_\iota^{a+1}$较小： $2\delta^{a+1}$ 与 $2\lambda_1^{a+1}$。这意味着灵敏度 $\partial\lambda_d/\partial\theta_a$ 相对于 $\partial\lambda_1/\partial\theta_a$较小，因此高阶参数 θ_a 的费舍尔信息从小特征值方向获得的信号相对较少。具体而言， λ_d 方向对 $\mathcal{I}(\theta)_{ab}$ 的贡献尺度为 δ^{a+b+2}，而 λ_1 方向的贡献为 λ_1^{a+b+2}。

6.3 一般动力学与李雅普诺夫放大

多项式情形之所以清晰，是因为特征向量是固定的，从而将全部灵敏度限制在特征基的对角线上。对于一般对称动力学 $\dot{X} = M(P)X$，若 M 对称但并非 P 的多项式，则特征向量会发生旋转，且灵敏度 $\partial P/\partial\theta$ 在特征基中同时具有对角分量与非对角分量。

在此设定下，李雅普诺夫方程 $\dot{P} = MP+PM$必须求逆，才能恢复 M，其依据为 P。在 P 的特征基中，该求逆按分量作用： $M_{\iota\gamma} = P_{\iota\gamma}/(\lambda_\iota+\lambda_\gamma)$。因子 $1/(\lambda_\iota+\lambda_\gamma)$ 会放大非对角分量中的噪声，且与出现在联络 1-形式（第 3.3 节）中并支配曲率的因子相同。由此得到直接对应：参数空间中在几何上困难的方向（较小的 $\lambda_\iota+\lambda_\gamma$ 会产生较大曲率与较大联络系数）在统计上也困难（较小的 $\lambda_\iota+\lambda_\gamma$ 会放大估计噪声）。

对于拉普拉斯动力学而言，和乐使情况更加复杂。特征向量的旋转为灵敏度带来了额外的自由度，但这些信息与规范自由度交织在一起；当命题 11 的非零曲率判据成立时，这些规范自由度可能无法在局部解析。

6.4 统计—几何对偶

结合几何与统计图景，可以得到一个统一的障碍。我们先陈述多项式动力学的结果，再解释更广泛的对偶关系。

命题 14 （多项式动力学的克拉美-罗界）。 对于度为 K 的多项式动力学，参数为 $\theta = (\theta_0, \dots, \theta_K) \in \mathbb{R}^k (k = K+1)$，假设：

1. θ 属于开参数集 $\Theta \subset \mathbb{R}^k$，且真实参数位于内部。*2.* 对于每个 t、i、j，映射 $\theta \mapsto P_{ij}(t;\theta)$ 是 C^1，且在真实参数的某个邻域内，$0 < P_{ij}(t;\theta) < 1$。

3. 得分函数可积，且微分与期望运算可以交换（标准的费舍尔正则性），因此信息恒等式成立。

4. $\mathcal{I}(\theta)$ 在可识别参数子空间上是正定的。

ˆ那么，任何无偏估计量 θ 都满足：

$$\mathbb{E}[\|\hat{\theta} - \theta\|^2] \geq \operatorname{tr}(\mathcal{I}(\theta)^{-1}) \tag{66}$$

费舍尔信息矩阵 $\mathcal{I}(\theta)$的元素为：

$$\mathcal{I}(\theta)_{ab} = \sum_{t=0}^{T} \sum_{\iota,\gamma=1}^{d} \frac{\partial \lambda_\iota}{\partial \theta_a} \frac{\partial \lambda_\gamma}{\partial \theta_b} \cdot C_{\iota\gamma}(t) \tag{67}$$

其中 $C_{\iota\gamma}(t) = \sum_{i<j} U_{i\iota} U_{j\iota} U_{i\gamma} U_{j\gamma} / [P_{ij}(t(1-P_{ij}(t))]$。对角元素 $C_{\iota\iota}(t) = \mathcal{W}_\iota(t)$ 衡量特征值方向 ι 在时间 t上的信息量；对于 $\iota \neq \gamma$，非对角元素 $C_{\iota\gamma}(t)$ 则捕捉特征值间的相关性。根据矩阵不等式$(\mathcal{I}^{-1})_{aa} \geq 1/\mathcal{I}_{aa}$，可得各参数方差的基线下界：

$$\operatorname{Var}(\hat{\theta}_a) \geq \frac{1}{\mathcal{I}_{aa}} = \frac{1}{\sum_t \sum_\iota \left(\frac{\partial \lambda_\iota}{\partial \theta_a}\right)^2 \mathcal{W}_\iota(t) + \textit{cross terms}} \tag{68}$$

非零的特征值间交叉项源于结构化图（其中 *Fisher* 权重与特征向量 r 分量相关）只能 提高真实的克拉美-罗界，使其高于这一基准。

证明。由（63）， P_{ij} 的灵敏度可分解为$\partial P_{ij}/\partial \theta_a = \sum_\iota U_{i\iota} U_{j\iota} (\partial \lambda_\iota / \partial \theta_a)$。代入（61）：

$$\mathcal{I}(\theta)_{ab} = \sum_t \sum_{i<j} \left[\sum_\iota U_{i\iota} U_{j\iota} \frac{\partial \lambda_\iota}{\partial \theta_a}\right] \left[\sum_\gamma U_{i\gamma} U_{j\gamma} \frac{\partial \lambda_\gamma}{\partial \theta_b}\right] \cdot \frac{1}{P_{ij}(1-P_{ij})} \tag{69}$$

对 ι， γ 上的乘积展开，得到：

$$\mathcal{I}(\theta)_{ab} = \sum_t \sum_{\iota,\gamma} \frac{\partial \lambda_\iota}{\partial \theta_a} \frac{\partial \lambda_\gamma}{\partial \theta_b} \underbrace{\sum_{i<j} \frac{U_{i\iota} U_{j\iota} U_{i\gamma} U_{j\gamma}}{P_{ij}(1-P_{ij})}}_{C_{\iota\gamma}(t)} \tag{70}$$

关于交叉特征值项。 U $\sum_i U_{i\iota} U_{i\gamma} = \delta_{\iota\gamma}$ $1/[P_{ij}$ $1-$的列 P_{ij}] 正交性给出，但费舍尔权重（ ）并不均匀。在结构化随机点积图中（例如具有相同大小块的随机块模型），这些权重在各块内为常数，而块由特征向量本身定义，因此权重与特征向量分量相关，交叉项 $C_{\iota\gamma}$ 对于 $\iota \neq \gamma$ 可能非零（在明确的块模型配置中确实非零）。因此，对于结构化图，费舍尔信息矩阵在特征值基底下并非对角矩阵。

然而，谱隙标度不受影响：特征值灵敏度 $\partial \lambda_\iota / \partial \theta_a$ 由标量常微分方程$\dot{\lambda}_\iota = 2\sum_k \theta_k \lambda_\iota^{k+1}$决定。驱动项 $2\lambda_\iota^{a+1}$ 在 $\iota = d$时按 δ^{a+1} 缩放，因此无论这一贡献来自对角项还是交叉项， λ_d 方向对$\mathcal{I}_{ab}$ 的贡献都是 $O(\delta^{a+b+2})$。当 $\delta \to 0$时，费舍尔信息矩阵会变得病态：涉及 δ 高次幂的行和列会缩小，导致 $\operatorname{tr}(\mathcal{I}^{-1})$ 发散。

基线界限 $(\mathcal{I}^{-1})_{aa} \geq 1/\mathcal{I}_{aa}$ 可直接由柯西–施瓦茨不等式得出： $1 = (e_a^\top e_a)^2 \leq (e_a^\top \mathcal{I} e_a)(e_a^\top \mathcal{I}^{-1} e_a) = \mathcal{I}_{aa} \mathcal{I}^{-1}_{aa}$ $\mathcal{I}$ （ ）。当为对角矩阵时，该界限是紧的；当非对角项较大时，该界限则较松（即结构化图中的参数相关性会使估计比对角分析所暗示的更加困难）。 □

如果 $\mathcal{I}(\theta)$ 是奇异的，则可在可辨识子空间上使用Moore–Penrose伪逆$\mathcal{I}^{+}$进行相同的推导。

该界揭示了三种标度机制。每个快照都提供 $O(n^2)$ 个条件独立的伯努利观测，因此费舍尔信息随 n^2 增长。假设动力学随时间在 P 上产生足够的变化，则信息会随 T 线性累积。谱隙 δ 通过特征值敏感度控制估计难度：参数 θ_a 从最小特征值方向获得的信号与 δ^{a+1} 成正比。对于常数项 θ_0（线性动力学），δ 方向对费舍尔信息的贡献为 $O(\delta^2)$；对于最高次项 θ_K，其贡献则减弱为 $O(\delta^{2K+2})$。这些标度关系对角元素和满矩阵都成立；交叉特征值项具有相同的 δ 依赖关系。

对于线性动力学（单个标量参数 α_0），费舍尔信息是精确的，不存在交叉项歧义：

推论 3 (线性动力学：费舍尔信息的显式表达)。 对于线性动力学 $X = \alpha_0 X$，快照的索引为 $t = 0$、.....、T，费舍尔信息为：

$$\mathcal{I}(\alpha_0) = 4 \sum_{t=0}^{T} t^2 \sum_{i<j} \frac{P_{ij}(t)}{1 - P_{ij}(t)} \tag{71}$$

其规模为 $\Theta(T^3 \cdot n^2)$ 当边概率在整个轨迹中始终远离 *0* 和 *1* 时，得到一个 *Cramér–Rao*下界 为 $\Omega(1/(T^3 n^2))$。

证明。 参见第A节。 □

关于 T 的三次依赖（而非线性依赖）反映了这样一个事实：灵敏度 $\partial P_{ij}/\partial \alpha_0 = 2tP_{ij}$ 会随时间增长；后续快照包含更多信息，因为扰动有更长时间传播。这是动力学估计的普遍特征，与独立同分布情形形成对比；在后者中，信息会线性累积。

标度 $\Theta(T^3 \cdot n^2)$仅在轨迹仍处于概率空间内部时成立。线性动力学会使特征值发生指数增长或衰减：$\lambda_\iota(t) = \lambda_\iota(0) e^{2\alpha_0 t}$。对于 $\alpha_0 > 0$，概率 P_{ij}向1增长，Fisher权重 $P_{ij}/(1 - P_{ij})$发散；但当概率饱和时模型失效，因此 $\mathcal{I}$ 的增长最终停止。对于 $\alpha_0 < 0$，概率衰减至0，灵敏度 $\partial P_{ij}/\partial \alpha_0 = 2tP_{ij}(t)$呈指数收缩，每个快照的费舍尔信息趋于饱和。在这两种情况下，T^3 标度描述的是边界效应占主导之前的短时间区间；有效时间视界为 $T_{\text{eff}} \sim 1/|\alpha_0|$。

对偶。 统计—几何对应关系在一般（非多项式）对称动力学中最为清晰，此时完整的李雅普诺夫结构显现。联络 1-形式涉及因子 $1/(\lambda_\iota + \lambda_\gamma)$ 控制规范敏感性（第3.3节）；截面曲率中的垂直括号范数包含相同的分母（加权方式如命题9），因此弱特征方向在交换子能量位于这些方向且 λ_{d-1}，λ_d 较小时在几何上变得脆弱；而估计 $M_{\iota\gamma}$ 所依据的李雅普诺夫方程 $P_{\iota\gamma} = (\lambda_\iota + \lambda_\gamma) M_{\iota\gamma}$ 中的费舍尔信息涉及同一因子 $(\lambda_\iota + \lambda_\gamma)^2$。

一个重要的限定条件：这种对偶适用于完整算子恢复（即估计 M 在所有特征向量方向上的作用，包括 λ_d 附近的弱子空间）。对于简约参数化（例如，所有特征值方向共用一个标量参数 α_0），估计主要可以依靠主导模态进行：仅凭 λ_1 的贡献就可能足以识别 α_0，完全绕过病态的 λ_d 方向。当感兴趣的参数专门控制弱子空间，或必须恢复完整算子 M 时，几何与统计的耦合才会产生影响。

在这一限定条件下，两者的对应关系仍然紧密：对于动力学中任何依赖谱隙的组成部分，同一个 δ 同时决定曲率和费舍尔信息。接近秩亏的网络（$\delta \to 0$，且λ_{d-1} 也很小）既更难对齐（曲率 $\sim 1/(\lambda_{d-1} + \delta)$），更难插值（单射半径 $\sim \sqrt{\delta}$），统计估计也更困难（最弱方向上的费舍尔信息 $\sim \delta^2$）。

对于非多项式动力学（例如拉普拉斯算子），敏感性传播会更加复杂，因为P的特征向量也会发生演化。额外的特征向量敏感性会对费舍尔信息产生正向贡献（观测值携带信号的方向更多），但也会引入和乐：在命题11适用的区域，额外信息可能与规范自由度纠缠在一起。在存在和乐的情况下，针对动力学估计的完整极小极大理论仍是一个重要的开放问题。

注。 **与 ASE 估计理论的关系**。 命题 14 由观测到的邻接矩阵的伯努利似然推导得出，适用于 任何 对 θ的估计量，而非专门针对谱方法。因此，该界在概念上不同于邻接谱嵌入的中心极限定理 [Athreya 等人、2016, 2018], 二范数到无穷范数扰动理论 [Cape 等人以及静态网络谱估计量的效率结果 2019], 。

参数 [Tang 等人，2022，Xie 和 Xu， 2023]。这些结果刻画了从单个图（或一组共享相同 P 的固定图集合）中估计潜在位置 X 或块概率矩阵 B 的准确性。ASE 渐近无偏，且逐顶点误差为 $O(1/\sqrt{n})$ [Athreya 等人， 2016]，；随机块模型块概率的谱估计量可达到渐近有效性 [Tang 等人， 2022]；对于单个潜在位置，一步法可达到局部效率（在正交变换意义下与预言机最大似然估计协方差相匹配） [Xie 和 Xu， 2023]。基于弗罗贝尼乌斯损失的潜在位置估计极小极大速率为 $\Theta(d/n)$ 每个顶点 [Xie 和 Xu，2020]，；二范数到无穷范数速率则取决于谱隙 [Agterberg 等人， 2023]。

相比之下，命题14 讨论的是对动力学参数 θ 的估计，其数据来自图的时间序列，而此前并无适用于此情形的有效性理论。其估计目标在性质上截然不同：目标不是 $O(nd)$ 潜在位置坐标，而是支配其时间演化的 $O(1)$ 参数。命题14中的费舍尔信息特征值方向分解，带有涉及特征向量载荷的权重 $\mathcal{W}_\iota(t)$，在静态理论中没有对应概念。将现有的极小极大框架推广到参数化时间模型，仍是一个开放问题；我们的费舍尔信息结构或可助力解决这一问题。

关于“任意无偏估计量”这一限定，值得作些说明。克拉美-罗下界是针对伯努利似然的界限，对任意 n 都有明确定义，因此该界限本身并不依赖渐近论证。然而，对 θ 的有用估计量在有限 n 下是否无偏，则是另一个问题。任何先通过谱方法恢复 P，再拟合 θ 的估计量，都会受到特征分解的有限样本偏差影响：谱映射的非线性会引入偏差，但这属于 $o(1/\sqrt{n})$ [Cape 等人， 2019]，不会影响渐近界限。接近秩亏（$\delta \to 0$）时，情况会更加严重：邻接谱嵌入会压缩弱特征值方向，由此在 λ_d 中产生的偏差可能超过信号λ_d 本身。在这种机制下，依赖于 λ_d 方向的参数，其均方误差很可能主要由偏差而非方差所主导，而克拉美-罗下界（仅约束方差）会低估实际难度。对于有限样本推断，带偏的克拉美-罗变体 $\mathrm{Var}(\hat{\theta}) \geq (1 + b'(\theta))^2/\mathcal{I}(\theta)$ 提供了适当的推广形式，但偏差 $b(\theta)$作为 δ 的函数，其刻画仍是一个开放问题。

7 构造性问题：可辨识性及其局限

前几节中的障碍刻画了原则上哪些内容可观测、哪些不可观测。我们探讨构造性问题：既然可观测动力学存在，能否从谱嵌入中恢复它们？我们建立了一项理论上的可辨识性结果，表明原则上动力学结构可以消除规范歧义。随后，我们讨论为何这种可辨识性并不能直接转化为实用算法。

7.1 结构约束对齐

对齐问题是不适定的：缺少额外信息时，许多规范选择都能产生看似合理的轨迹。一个自然的想法是利用关于动力学结构的归纳偏置，以约束哪些轨迹是可接受的。

假设动力学属于第 $\mathcal{F}$5.1 节中介绍的一个族。例如，多项式动力学 $\dot{X} = N(P)X$，其中$N(P) = \sum_k \alpha_k P^k$。这些族具有两项特性，表明它们或可用于正则化对齐：它们是水平的（对称的 N 可确保不会发生规范漂移），且具有受限结构，随机规范误差会违反这种结构。

如果真实动力学属于 $\mathcal{F}$，那么正确的规范会产生可由某个 $f \in \mathcal{F}$ 拟合的轨迹，而错误的规范会产生需要 $\mathcal{F}$ 之外的动力学才能拟合的轨迹。动力学族对规范选择起到正则化作用。
我们将此表述为一个联合优化问题：

定义 9（联合对齐与学习问题）。 给定 ASE 嵌入 $\{\hat{X}^{(t)}\}_{t=0}^{T}$ 和一个动力学族 $\mathcal{F}$，求规范校正 $\{Q_t \in O(d)\}$ 和 $f \in \mathcal{F}$，使以下目标最小化：

$$\mathcal{L}(\{Q_t\}, f) = \sum_{t=0}^{T-1} \|\hat{X}^{(t+1)}Q_{t+1} - \hat{X}^{(t)}Q_t - \delta t\, f(\hat{X}^{(t)}Q_t)\|_F^2 \tag{72}$$

该目标衡量对齐后的离散轨迹 $\{\hat{X}^{(t)}Q_t\}$ 在多大程度上可由动力学 f解释：如果规范正确，且 f 捕捉到了真实动力学，那么每一步的位移都应与预测速度相匹配。

7.2 可辨识性原理

结构约束对齐的理论依据，在于将规范伪影与水平动力学清晰地区分开来。

定理 4 (规范速度污染**)**。 设 $X(t)$ 遵循真实动力学 $\dot{X} = NX$，其中 $N = N^\top$。设 $S(t) \in O(d)$ 为随时间变化的（错误）规范变换，并定义规范变换后的轨迹 $\tilde{X}(t) = X(t)S(t)$。那么

$$\dot{\tilde{X}} = N\tilde{X} + \tilde{X}\Omega \tag{73}$$

其中$\Omega(t) = S(t)^\top \dot{S}(t) \in \mathfrak{so}(d)$是反对称的。

此外，如果 $\tilde{X}(t)$ 在某个区间内对所有 t 均满列秩，那么在该区间上，以下条件等价 *l:*

1. 存在一个对称矩阵函数 $\tilde{N}(t) = \tilde{N}(t)^\top$，使得 $\dot{\tilde{X}} = \tilde{N}(t)\,\tilde{X}$；*2.* $\Omega(t) = 0$ （即，$S(t)$是时间常数）。

. .证明。 根据乘积法则，$\dot{\tilde{X}} = \dot{X}S + X\dot{S} = NXS + X\dot{S}$。利用 $X = \tilde{X}S^\top$ （因为 $S \in O(d)$），可得 $\dot{\tilde{X}} = N\tilde{X} + \tilde{X}(S^\top\dot{S}) = N\tilde{X} + \tilde{X}\Omega$，其中 $\Omega = S^\top\dot{S} \in \mathfrak{so}(d)$，因为对 $S^\top S = I$ 求导得到 $\dot{S}^\top S + S^\top\dot{S} = 0$。

˜ ˜ ˜ ˜现在假设对于某个对称（可能随时间变化的） N，有 $\dot{X} = NX$，且 X 满列秩，因此 $G = \tilde{X}^\top\tilde{X}$ 正定。左乘 $\tilde{X}^\top$ 得到 $\tilde{X}^\top\dot{\tilde{X}} = \tilde{X}^\top\tilde{N}\tilde{X}$，右侧是对称矩阵˜，因为 N 是对称的。

.另一方面，根据上述分解，$\tilde{X}^\top\dot{\tilde{X}} = \tilde{X}^\top N\tilde{X} + G\Omega$。项 $\tilde{X}^\top N\tilde{X}$ 是对称的，因此 $G\Omega$也必须是对称的。但如果 G 是对称正定的，且 Ω 是反对称的，那么 $G\Omega$ 是对称的，当且仅当 $\Omega = 0$ （等价地，$G\Omega + \Omega G = 0$ 意味着 $\Omega = 0$）。因此 $\Omega = 0$，这意味着 S 在时间上是常数。˜ ˜反向结论显而易见：如果 $\Omega = 0$，那么 $\dot{X} = NX$ 已经具有所需的形式，其生成元N是对称的。 □

其机制十分巧妙：随机的邻接谱嵌入规范误差 $R^{(t)}$ 会引入反对称污染，而这种污染无法被对称动力学吸收。要求动力学为对称的，实际上隐式地选取了污染消失的规范。

读者或许会注意到，这与命题2有相似之处；该命题指出，$\dot{X}$ 是水平的，当且仅当 $X^\top\dot{X}$ 是对称的。这两个结果基于同一个代数事实（正定矩阵与非零反对称矩阵的乘积绝不会是对称矩阵），但回答的是不同的问题。命题2 刻画单点处的方向：给定一个速度 X，它是否含有规范分量？定理4 刻画随时间演化的轨迹：给定一条完整轨迹˜X（t），它是在错误的规范下观测到的，那么能否用任何对称动力学来解释？前者是逐点分解；后者则是关于沿路径的动力学—规范耦合的可辨识性陈述。

对于特定的模型族，可以据此得到具体算法。对于线性对称动力学，可采用交替方案：固定规范，并通过带有对称性约束的最小二乘问题求解动力学矩阵 M；固定 M，并针对每个规范求解正交Procrustes问题。对于多项式动力学 $\dot{X} = (\sum_k \alpha_k P^k)X$，$P = \hat{X}\hat{X}^\top$ 的规范不变性可将动力学步骤简化为对 α_k 系数进行线性回归。两种方法的每一步都可以闭式更新。

7.3 下游流程：从轨迹到方程

乐观地假设轨迹恢复问题已经解决，也就是说，某种方法（结构约束对齐、未来某种 P 级别的估计量，或领域特定先验信息）得到了潜在位置 $\tilde{X}^{(t)}$ 的规范一致估计值，但其中仍含有噪声。

注。 **采样平均启发式方法**。 当每个时间点 t都有多个独立的网络样本 $A_1^{(t)}, \ldots, A_m^{(t)}$时，取平均值 $\bar{A}^{(t)} = \frac{1}{m}\sum_i A_i^{(t)}$ 可将逐条目方差降低至原来的 m。因此，在标准的稠密摄动状态下，预期逐顶点嵌入误差大致会改善 $\sqrt{m}$。我们将此作为实用的尺度启发式方法，而非本手稿中经证明的收敛速率。

在获得这样的轨迹后，其余步骤都已相当成熟。 通用微分方程(通用微分方程) [Rackauckas 等人，2020] 提供了一个框架，用于学习将已知机理结构与神经网络灵活性相结合的动力学：

$$\dot{X} = g(f_{\text{known}}(X, \phi), f_{\text{NN}}(X, \theta)) \tag{74}$$

对于随机点积图动力学，其已知结构来自第5.1节中的多项式族。例如，多项式架构 $\dot{X} = (\sum_k \alpha_k(\theta, X)P^k)X$ 通过构造确保了水平性，同时允许学习系数。梯度通过伴随灵敏度方法经由常微分方程求解器传播，而规范一致架构则对假设空间加以约束。

通用微分方程训练完成后，符号回归 [Cranmer， 2023] 可以提取出易于解释的闭式方程。通过从训练好的模型中采样状态—速度对$(X, f_\theta(X))$ 并搜索形式为$N(P)X$的符号表达式，可以恢复显式动力学方程，从而完成从观测到的邻接矩阵到微分方程的整个过程。需要注意的是：对潜变量进行符号回归会依赖规范。除非将通用微分方程的输出预先对齐到规范标架（例如 P的特征基），否则回归可能无法找到稀疏系数，因为X 在通用微分方程所用规范下，是“自然”坐标经过旋转后的结果。使用规范不变特征(特征值 P，或直接使用 P动力学 $P = NP + PN$）可以避免这一问题；对于动力学按特征值分离的多项式族而言，这种做法也很自然。

瓶颈在于获得 $\tilde{X}^{(t)}$。通用微分方程和符号回归的工具已相当成熟，但轨迹恢复问题尚未解决。

7.4 为什么构造性问题仍未解决

尽管可辨识性理论简洁明了，但要将定理 4 转化为可靠的恢复方法，仍面临三个相互作用的难题。

有限样本偏差。 每个谱嵌入 $\hat{X}^{(t)}$ 对真实位置的估计误差为 $O_p(1/\sqrt{n})$。当真实动力学较慢（$\|X^{(t+1)} - X^{(t)}\|$较小）时，对齐的信噪比会下降：我们试图从与之大小相近的估计噪声中辨别微小的真实位移。交替优化也会受到影响：动力学步骤拟合的是真实信号与噪声的组合，而规范步骤则对齐到一个有噪声的目标。两种误差来源相互叠加，而非彼此抵消。

动力学族的表达能力。 可辨识性论证要求 $\mathcal{F}$ 具有足够的限制性，使错误的规范必然导致结果落在该族之外。但在有限样本和离散时间下，出人意料地灵活的行为可能悄然出现。

一种机制源于数值因素，而非哲学层面：即使对于多项式族，针对 $(\alpha_0$、.....、$\alpha_K)$ 的回归问题也可能变得病态。矩阵 I、P、P^2、.....会逐渐与P 的最高特征子空间对齐（类似幂迭代），因此设计矩阵在特征值上呈类范德蒙结构，对于 $K \geq 2$ 可能会严重病态。在这种机制下，适度的规范诱导伪影（或普通的 ASE 噪声）可能会被高阶项中幅度很大、彼此抵消的系数拟合，从而难以区分“真实动力学”和“能够解释规范失配数据的动力学”。

第二种机制源于几何因素：当 d 较小时，规范自由度是低维的（例如， $\mathfrak{so}(2)$ 是一维的），因此缓慢变化的规范失配可能会造成一种连贯的漂移，在离散时间分辨率下看起来并非明显荒谬。道理相同：让 $\mathcal{F}$ 更加严格，有排除真实情况的风险；让它的表达能力更强，则可能放任规范伪影（以及噪声）混入。

和乐与全局一致性。 即使局部对齐成功（每一对相邻帧都对齐良好），和乐（第 3.3 节）也意味着规范漂移可能沿长轨迹不断累积。对于其基空间 路径 $P(t)$ 在 $\mathcal{B}$ 中形成（或近乎形成）回路的动力学，沿整个回路全局一致地选择规范可能是不可能的：闭曲线的水平提升不一定会闭合。

这不是增加数据就能解决的问题；它是模型类的几何障碍。增加数据可以 做的是减少叠加在其上的统计波动。这一点很重要，但这是另一个问题。

基本权衡。 核心难点在于两个要求之间的张力。动力学族必须足够受限，才能排除受规范污染的轨迹（可辨识性）；同时又必须足够灵活，才能捕捉真实动力学（模型充分性）。要使问题适定， $\mathcal{F}$ 能够拟合的内容与规范污染所产生的内容之间的“间隙”必须超过噪声水平。刻画这一间隙何时足够大（取决于 n、 d、T、 P 的谱隙以及动力学族），是核心开放问题。

注。 困难并不在于可辨识性原理在理论上失效。定理 4是尖锐的：在连续时间、无限数据极限下，对称动力学 确实能够 辨识规范。困难在于，有限样本、离散时间

设置会引入误差，而这些误差会以渐近理论无法捕捉的方式与规范-动力学耦合相互作用。要取得进展，可能需要更强的结构假设（例如直接研究 P-动力学，从而完全绕过规范问题），或从根本上采用不同的估计策略（例如能够刻画动力学恢复极小极大速率的信息论方法）。

7.5 基于锚点的对齐：可吸引的特殊情形

上述困难在一般情形下确实存在。不过，有一种自然的特殊情形可以干净地解决规范问题：已知（或假设）一部分节点是静止的。

锚点原理。假设一部分节点 $S \subset \{1, \dots, n\}$ 满足 $\dot{x}_i = 0$，对于 $i \in S$而言。这些是"锚点"，其潜在位置不会移动。在每个时间 t，邻接谱嵌入（ASE）都会产生 $\hat{X}^{(t)} = X^{(t)}R^{(t)} + E^{(t)}$，并带有一个规范因子$R^{(t)} \in O(d)$（源于特征分解的选择）以及估计噪声 $E^{(t)}$。对于锚点节点 $i \in S$，对所有 t，都有 $x_i^{(t)} = x_i^{(0)}$，因此锚点行满足 $\hat{X}_S^{(t)} = X_S R^{(t)} + E_S^{(t)}$，其中 $X_S \in \mathbb{R}^{|S| \times d}$ 为常数。

通过 Procrustes 将锚点行对齐到某一参考系（例如 $t = 0$）：

$$\hat{Q}^{(t)} = \arg\min_{Q \in O(d)} \|\hat{X}_S^{(t)} Q - \hat{X}_S^{(0)}\|_F. \tag{75}$$

在无噪声极限下，这直接恢复相对规范 $(R^{(t)})^{-1}R^{(0)}$，无需使用任何动力学模型，也从不按顺序传播规范。将 $\hat{Q}^{(t)}$ 应用于完整嵌入 $\hat{X}^{(t)}$，即可将所有节点对齐到$t = 0$规范。

为何有效。 锚点节点提供一个固定参考系，在每个时间步"钉住"规范。对齐是全局 （所有帧都对齐到 $t = 0$，而非顺序对齐），因此不存在误差累积。该方法无需了解动力学族 $\mathcal{F}$：它对多项式、拉普拉斯算子或任何其他动力学都有效，只要锚点真正静止。和乐无关紧要，因为我们从不试图沿轨迹传播规范；每一帧都独立对齐到固定参考。

条件与局限。 锚点原理要求：

1. 条件良好的锚点： 锚点位置矩阵 $X_S \in \mathbb{R}^{|S| \times d}$ 必须满列秩，Procrustes 问题才有确定解；同时矩阵还应具有良好的条件数（$\sigma_d(X_S)$不能太小），以确保抗噪性。仅仅计算锚点数量并不足够： $|S| \gg d$ 如果锚点近乎共线，或聚集在单个点附近，旋转在正交方向上就会受到较弱的约束。这与经典正交Procrustes问题中的稳定性情况相同：当参考构形接近秩亏时，旋转估计会变得敏感。（这里不再给出完整的扰动界；就我们的目的而言，关键在于对条件数的依赖。相关内容可参见标准的普罗克鲁斯扰动理论，或者直接从Procrustes的奇异值分解公式以及韦尔/韦丁型特征向量与奇异向量扰动界来理解。）2. 已知锚点身份： 我们必须知道哪些节点是静止的。在实践中，这些信息可能来自领域知识（例如生态网络中已确立的物种，或社会网络中的机构节点），也可能来自初步分析，即识别出时间方差较低的节点。3. 近似静止的锚点： 如果锚点节点以速度 $\|\dot{x}_i\| = \epsilon$ 缓慢漂移，持续时间为 $i \in S$，那么对齐就会产生随时间范围增长的系统偏差。一个粗略但实用的经验法则是：锚点漂移应远小于所关注时间窗口内累积的嵌入噪声；否则，"参考系"本身也会移动。

注。 第三个条件解释了我们在初步实验中观察到的一种现象：包含一大批缓慢移动节点的网络，可以通过朴素的 Procrustes 方法成功对齐；而所有节点移动速率相近的网络则无法对齐。这些缓慢移动的节点实际上充当了事实上的 锚点，在我们尚未明确意识到这一点时，就稳定了规范。

锚点符合实际的情形。 有几个应用领域天然存在动力学速率各异的节点。在生态食物网中，基底物种（初级生产者）的营养级位置通常较为稳定，而高营养级消费者则会迅速变化。在社会网络中，机构行为体（组织、永久职位）可能长期存在，而个体则不断变动。在神经连接组中，结构枢纽区域在外围联络重新布线的时间尺度内可能保持稳定。更一般地说，任何具有时间尺度分离的系统都是天然的候选对象：它拥有一个缓慢演化的"骨干"和一个快速演化的外围。

锚点方法并未解决一般性的规范问题，而是将一个棘手的几何统计问题转化为对领域知识的要求：识别静止节点。但在具备此类知识时，它能提供

一条简洁且计算上极为简单的路径，得到规范一致轨迹，从而支持第 7 节所述的后续 UDE 流程。

7.6 数值示例

我们通过在合成随机内积图数据上进行的两项受控实验来说明该理论。第一项实验展示规范等变多项式动力学上的基于锚点的对齐，其中对齐质量可以独立于动力学恢复进行评估。第二项实验展示完整的下游流程：UDE 训练与符号回归；为了更清晰地展示这一点，我们转向非规范等变动力学，此时对齐质量会直接影响动力学恢复。

7.6.1 实验1：基于锚点的对齐

设置。 我们生成一个具有 $n = 200$ 个节点、 $d = 2$ 个潜在维度的随机点积图。初始位置 $X(0)$从 $B_+^2 = \{x \in \mathbb{R}^2$ 中均匀抽取： $x_1, x_2 \geq 0, \|x\| \leq 1\}$，并指定大小为 n_a的锚点集 S。非锚点节点在多项式动力学 $\dot{X} = (\alpha_0 I + \alpha_1 P)X$ 下演化，参数为$(\alpha_0, \alpha_1) = (-0.3, 0.003)$，而锚点节点保持固定： $\dot{x}_i = 0$ 对于 $i \in S$。我们对常微分方程进行积分，在 T 个等间隔时间上生成轨迹 $X(t)$，每个时间步生成 $m = 3$ 个独立的邻接矩阵 $A_k^{(t)} \sim$ 伯努利$(X(t)X(t)^\top)$（并对其取平均以降低噪声），计算ASE 嵌入$\hat{X}^{(t)}$，并将基于锚点的对齐与序贯普罗克鲁斯对齐进行比较。

除非另有说明，我们使用 $T = 50$步，步长为 $\delta t = 0.05$ （总时间为 2.45），并报告 20 次蒙特卡洛重复试验的均值和标准差。

度量指标。 我们用平均普罗克鲁斯误差衡量对齐质量：

$$\mathrm{err}(t) = \frac{1}{n}\|\hat{X}^{(t)}\hat{Q}^{(t)} - X^{(t)}Q^*\|_F \tag{76}$$

其中， Q^* 表示与真实轨迹进行全局对齐时的最优结果（计入 $t = 0$处残余的规范自由度）。我们考察四种实验条件：锚点数量 n_a、轨迹长度 T、锚点漂移率 ϵ，以及初始嵌入范数尺度（用于控制信噪比）。

结果。实验　结果与基本理论预期相符，也揭示了“简单”直觉何时会失效　.

- 锚点数量。 锚点过少时，普罗克鲁斯问题无法确定：对于 $d = 2$， $n_a = 1$ 是最坏情形，对齐会失败；而 $n_a = 2$ 则处于临界状态，方差很高（有时有效，有时无效，取决于锚点几何分布）。当 n_a 明显增大到适中水平（在此处的实验中为 $n_a \geq 5$）后，基于锚点的误差便稳定在邻接谱嵌入噪声底限附近（在我们的设置中约为 0.005），这与锚点此时能够稳健地张成潜在平面的观点一致。
- 轨迹长度与漂移累积。对于较短的轨迹（$T \leq 50$），序贯普罗克鲁斯对齐略胜一筹（在我们的扫描中约有 2% 的优势），因为它在每一步都使用全部 $n = 200$ 个节点，而基于锚点的 Procrustes 仅使用 $n_a = 15$ 个锚点。但序贯对齐会累积漂移：两者的表现大约在$T \approx 100$处出现转折；到了 $T = 200$ ，序贯普罗克鲁斯对齐的表现约差 13%（比率为 ≈ 1.13），这与累积旋转误差预期的 $O(\sqrt{T})$ 增长相符。
- 漂移锚点。经过 $T = 50$步（总时间为 2.45）后，适度的锚点漂移率对平均对齐误差影响不大（因为该指标是对全部 200 个节点取平均值，而锚点只有 15 个）。其对下游变异性的影响则更明显：在 $\epsilon = 0.1$ 处，恢复参数的变异性显著增加，这与缓慢移动的参考系引入系统偏差的情况相符。

动力学恢复在此（大体上）是规范自由的。 一个有用的合理性检查是：对于多项式动力学，系数 (α_0, α_1) 可以直接由规范不变轨迹 $P(t) = X(t)X(t)^\top$估计得到。在我们的扫描中，$\hat{\alpha}_0$ 的标准误在不同锚点数量下基本保持不变（约为0.008），从而证实这一特定的参数恢复问题从根本上并不依赖于对 X的对齐。

然而，点估计 确实会随锚点数量而变化（例如，$\hat{\alpha}_0$ 从 $n_a = 0$ 时约为 -0.282 ，变为 $n_a = 15$ 时约为 -0.250 ），因为冻结大量节点会改变实际的 $P(t)$ 轨迹。这是数据生成过程的变化，而非规范伪影。这引出了下一个问题：当动力学依赖于 X-空间坐标（即不具备规范等变性）时，对齐质量是否会直接影响动力学恢复？实验2旨在让这种依赖性无法回避。

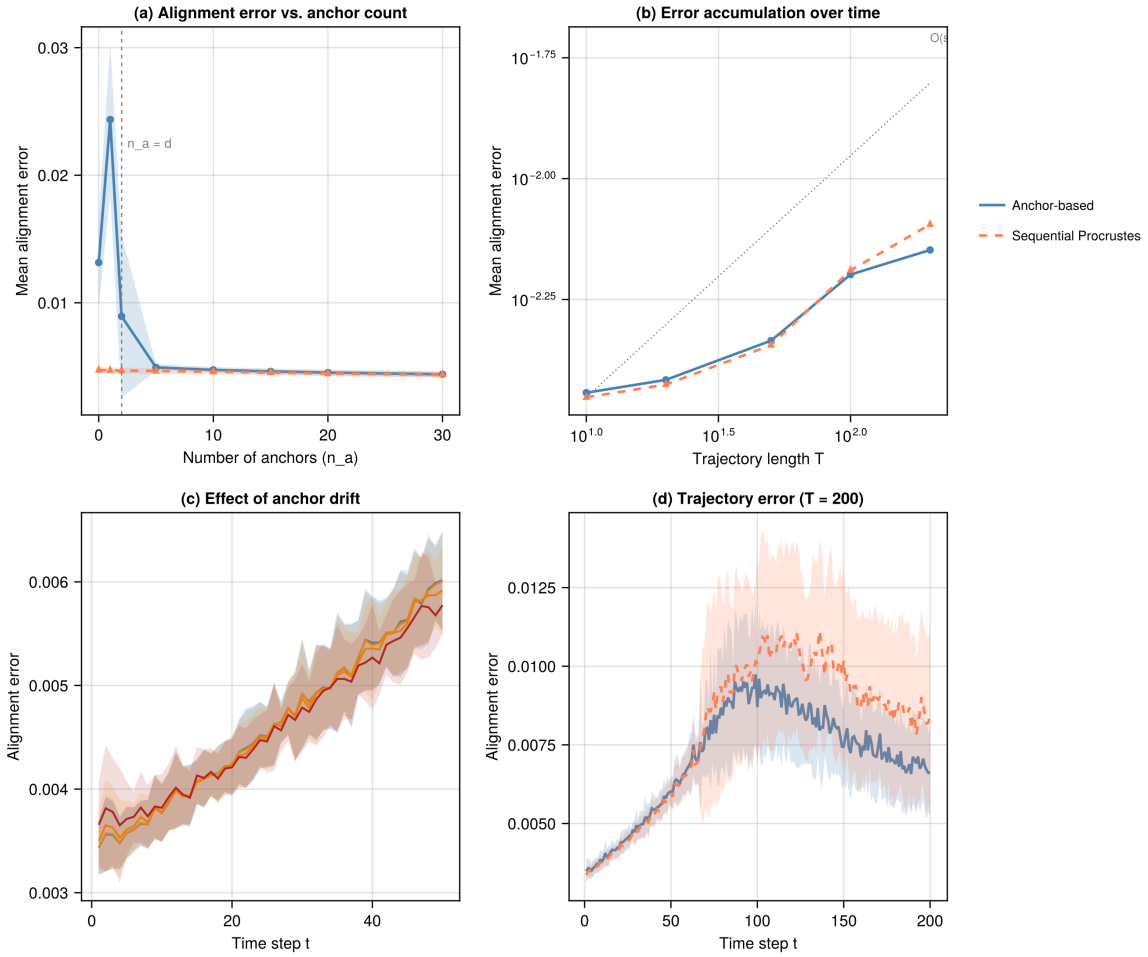

图1：基于锚点的对齐实验（$n = 200, d = 2$，多项式动力学 $X = (\alpha_0 I + \alpha_1 P)X$，每个时间步包含 $m = 3$ 个伯努利样本）。**(a)**对齐误差随锚点节点数量的变化：低于 $n_a = d = 2$（垂直虚线）时，Procrustes 问题欠定，对齐失败；锚点数量充足时，误差稳定在 ASE 噪声底限。**(b)** 轨迹长度与误差累积：基于锚点的对齐（蓝色）保持有界，而序贯普罗克鲁斯对齐（珊瑚色虚线）随 T增长，这与 $O(\sqrt{T})$ 漂移累积相符。**(c)** 锚点漂移率 ϵ的影响：随着 ϵ增大，漂移锚点导致的系统偏差变得明显。**(d)** $T = 200$ 轨迹每个时间步的对齐误差：基于锚点的对齐误差（蓝色）基本保持平稳，而序贯普罗克鲁斯对齐（珊瑚色）则不断累积漂移，且间隙在较晚的时间步逐渐扩大。阴影带表示20次蒙特卡洛重复试验的 ± 1 标准差。

7.6.2 实验 2：采用非规范等变动力学的 UDE 流程

动机。 实验1中的多项式动力学具有规范等变性：$P = XX^\top$ 具有旋转不变性，因此动力学参数位于 P-空间，无需对齐即可恢复。为了完整检验轨迹恢复→ UDE 流程，我们现在设计一种直接依赖于 X-空间坐标的动力学，使对齐质量成为真正的瓶颈。

设置。 我们生成一个包含 $n = 200$ 个节点、处于 $d = 3$ 维潜在空间中的随机点积图，其中 $n_a = 100$ 个锚点节点分属 $K_c = 3$ 个社群，分布在B_+^3 的顶点附近（位置为$(0.7, 0.2, 0.2)$、$(0.2, 0.7, 0.2)$、$(0.2,0.2, 0.7)$，并叠加高斯分布噪声）。100个非锚点节点围绕各自的社群质心 μ_k进行阻尼螺旋动力学演化：

$$\dot{x}_i = (-\gamma + \beta\|x_i - \mu_k\|^2)(x_i - \mu_k) + \omega J(x_i - \mu_k) \tag{77}$$

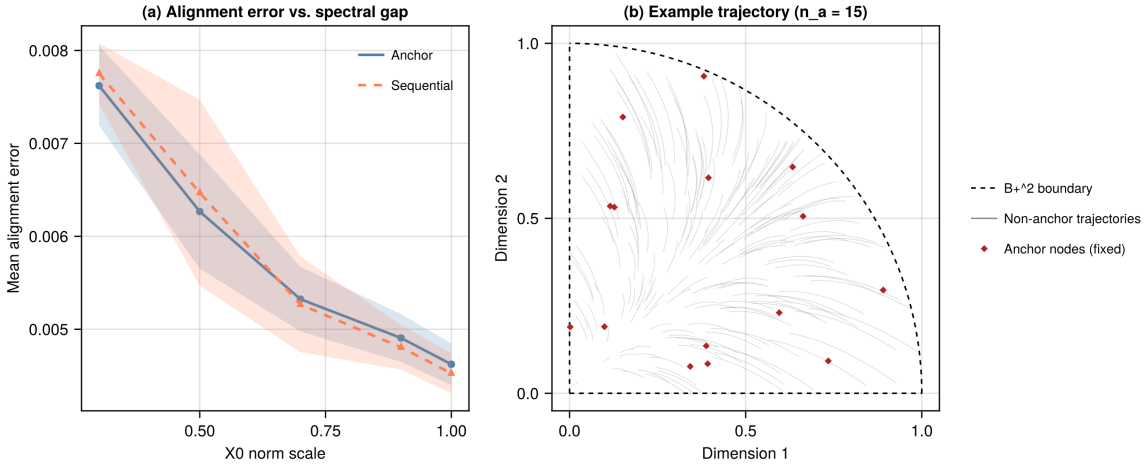

图 2：（**a**）对齐误差随初始嵌入范数尺度（信号强度的代理指标）的变化：范数越小，信号越弱，伯努利观测提供的信息越少，导致 ASE噪声增大，两种方法的对齐效果都变差。注意，均匀缩放会保持条件数 σ_1/σ_2 （恒定为 ≈ 2.2）；本实验改变的是信号幅度，而非谱间隙比。（**b**）在 B_+^2 中，包含 $n_a = 15$ 个锚点节点（红色菱形，静止）和非锚点节点（灰色轨迹，遵循多项式动力学演化）的相图。

其中，$J = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$ 是绕$(1,1,1)/\sqrt{3}$ 轴旋转的生成元，而$(\gamma, \beta, \omega) = (0.3, -0.5, 1.0)$。这些动力学并非 规范等变：质心 μ_k 和旋转轴都以 X空间坐标表示，因此，未对齐的轨迹 $X(t)$会产生不同的偏移量 $\tilde{x}_i - \mu_k \neq x_i - \mu_k$，破坏动力学结构。

我们在每个时间步于 $\delta t = 0.1$ 观测 $m = 10$ 个独立邻接矩阵，持续 $T = 50$ 个时间步，计算 ASE 嵌入，并采用三种对齐条件：基于锚点的 Procrustes、序贯普罗克鲁斯对齐，以及不进行对齐（原始 ASE 嵌入带有随机的逐帧旋转）。

通用微分方程架构。遵循第 7 节的框架，我们将动力学分解为已知和未知部分：

$$\dot{x}_i = \underbrace{-\hat{\gamma}(x_i - \mu_k)}_{f_{\text{known}}} + \underbrace{f_\theta(x_i - \mu_k)}_{f_{\text{NN}}} \tag{78}$$

其中，$\hat{\gamma}$ 是一个可学习的标量参数，而 $f_\theta : \mathbb{R}^3 \to \mathbb{R}^3$ 是一个小型神经网络（架构： $3 \to 16 \to 16 \to 3$，采用双曲正切激活函数，共有 ≈ 390 个参数）。已知部分编码了这样一种结构性假设：非锚点节点会被吸引到其所属社群的质心；未知部分则捕获其他可能存在的动力学。对网络权重进行 L2正则化（$\lambda = 10^{-3}$）可防止网络吸收线性阻尼项，从而提高已知部分与未知部分分解的可辨识性。通过从初始条件出发正向求解神经常微分方程，并利用伴随灵敏度方法最小化轨迹均方误差，来训练通用微分方程。

主要指标：总动力学均方误差（核心关注点）。由于神经网络可以吸收部分线性阻尼项，加性分解 $f_{\text{learned}}(\delta) = -\hat{\gamma}\delta + f_\theta(\delta)$存在固有的可辨识性注意事项：不同的（$\hat{\gamma}$， f_θ）组合可能产生基本相同的总向量场。因此，我们评估真正关心的内容，即学习到的总动力学，并将其与留出样本集中的随机测试输入 $\delta \in [-0.3, 0.3]^3$ （在我们的实现中为 2000 个样本）上的真实动力学 $f_{\text{true}}(\delta) = (-\gamma + \beta\|\delta\|^2)\delta + \omega J\delta$ 进行比较。该指标特意设计为不受模型如何在$\hat{\gamma}$ 与神经网络之间分摊责任的影响。

结果。 总动态 MSE 非常清晰地区分了三种对齐条件（图 3d），在 5 次蒙特卡洛重复实验ns:

- 锚点对齐： 平均 MSE $\approx 6 \times 10^{-4}$，标准差为 $\approx 4 \times 10^{-4}$，即对完整向量场的恢复效果极佳。
- 序贯普罗克鲁斯对齐： 平均 MSE $\approx 7.6 \times 10^{-3}$，标准差为 $\approx 3 \times 10^{-4}$，约差 $13\times$ 。
- 未对齐： 平均 MSE 为 ≈ 0.44，标准差为 ≈ 0.35，表现约差 $700\times$，且波动很大。

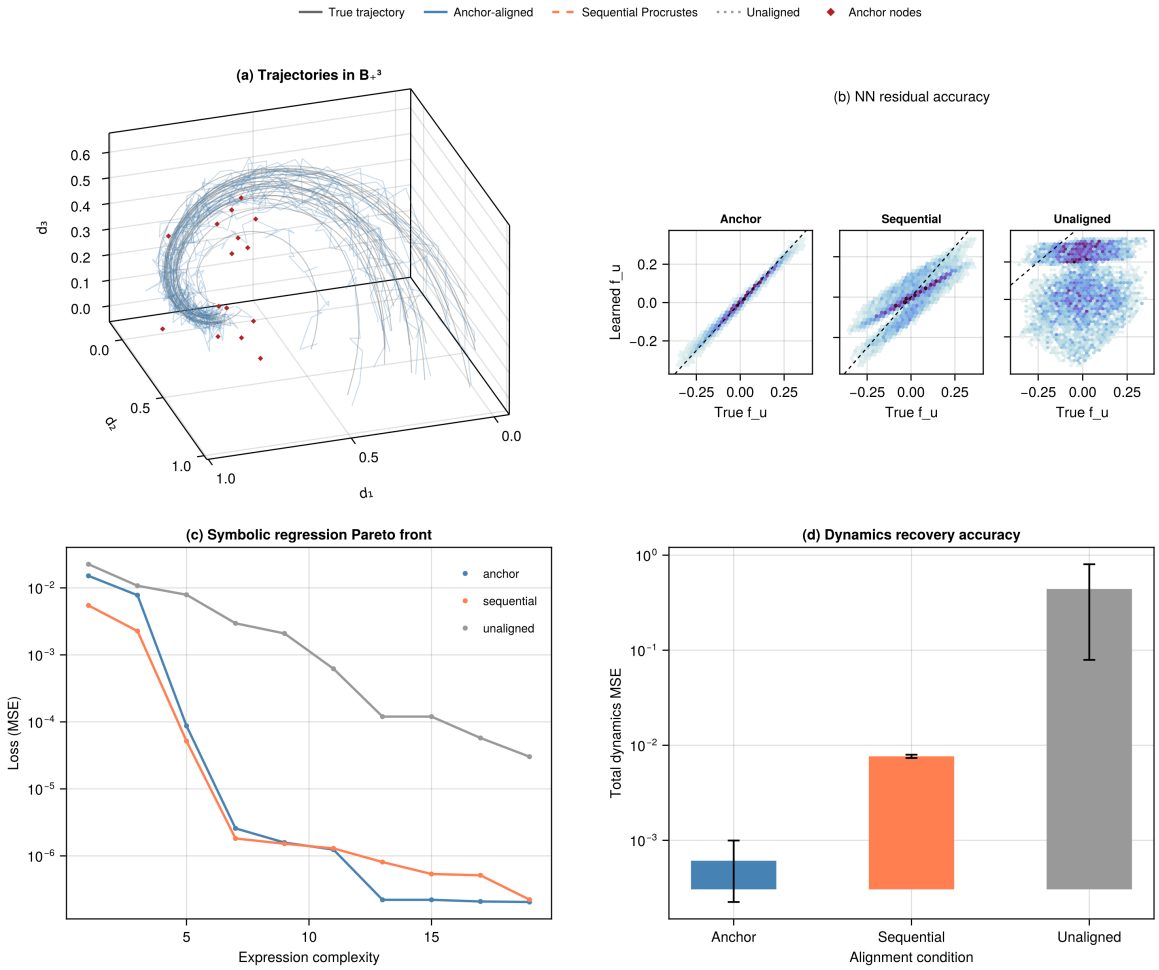

图 3：UDE 流程实验（$n = 200$、 $d = 3$，以 (1,1,1) 为旋转中心的阻尼螺旋动力学，每帧 $/\sqrt{3}$ $m = 10$ 个伯努利样本）。**(a)** B_+^3 中的真实轨迹（灰色）和锚点对齐的邻接谱嵌入（蓝色）；锚点节点以红色菱形标示。**(b)** 学习得到的神经网络残差 $f_\theta(\delta)$ 与真实残差 $f_u(\delta)$：锚点对齐的数据（左）沿对角线高度吻合；序贯普罗克鲁斯对齐（中）和未对齐数据（右）的效果则依次变差。**(c)** 符号回归帕累托前沿（损失与表达式复杂度）：在每个复杂度水平上，锚点对齐的数据都能将损失降低 1–2 个数量级。**(d)** 按对齐条件划分的总动力学均方误差（对数刻度）：锚点对齐的均方误差为 $\approx 6 \times 10^{-4}$，序贯普罗克鲁斯对齐为 $\approx 8 \times 10^{-3}$（13× 更差），未对齐为 ≈ 0.44（$\approx 700\times$ 更差）。误差线表示 5 次重复实验的 ± 1标准差。

神经网络残差 MSE 显示出同样的排序，而且更加明显：锚点对齐时，神经网络残差 MSE 约为9×10^{-4}；顺序对齐时约为 7.3×10^{-3}；未对齐时约为 0。44。这一机制显而易见：当帧得到连贯对齐时，神经网络会学习预期的非线性+旋转残差；当帧以不连贯的方式旋转时，它学到的大多是噪声。

符号回归。 对训练后网络的输入输出对进行符号回归（图3c）可得到帕累托最优表达式，其损失比其他条件低1–2 个数量级（在每个复杂度水平上，锚点对齐条件下均如此）。锚点对齐表达式包含预期的结构项（坐标乘积以及与旋转生成元 J一致的交叉项）。如第7节所述，符号回归需要使用规范不变特征或典范对齐标架；锚点对齐提供了后者。

可辨识性注意事项（以及为何它并不削弱结果）。学习得到的阻尼参数 $\hat{}\gamma$ 在多次重复实验中并不特别稳定（锚点对齐后的均值为 ≈ 0.255，标准差为 ≈ 0.081，而真实值为 $\gamma = 0.3$），因为即使采用 L2 正则化，神经网络仍然可以吸收部分线性项。这是加性通用微分方程分解的一个真实局限，并非对齐问题。

尽管如此，这并不会削弱主要结论，因为在锚点对齐下，总动力学仍得到了准确恢复，而这正是总动力学均方误差所衡量的内容。

要点。 当动力学具有规范等变性时（$X = N(P)X$），对齐质量与参数恢复无关（实验1）。当动力学取决于 X-空间坐标时，对齐质量会直接决定学习到的动力学的保真度（实验2）。基于锚点的对齐能够提供下游 UDE 流程所需的规范一致轨迹。

8 讨论与结论

我们希望已经让读者相信，将随机点积图视为动力系统颇具吸引力，尽管从几何和统计角度来看，这确实面临诸多困难。将时序随机点积图作为动力系统来刻画，主要优势在于能够获得严谨的理论理解，并明确一系列具体的失效模式，从而阐明困难的根源及其表现形式。

我们指出了三个核心障碍。第一，规范自由度会使整个潜在运动子空间不可见（定理 1），并迫使任何潜在空间方法都必须处理对齐问题。第二，可实现性限制了在固定潜在维度下哪些概率矩阵动力学是可能的。第三，轨迹恢复问题表明，在标准 RDPG 流程中——潜在位置通过邻接谱嵌入（ASE）估计——朴素的时间差分主要测量的是规范抖动，而非动力学。

我们采用并发展的微分几何框架（第 3.3 节）将规范自由度形式化为主丛，并明确了随机点积图框架中的联络／曲率／和乐机制。联络 1-形式将“规范速度”形式化，而曲率／和乐则解释了局部对齐为何可能无法推广到全局。对具体动力学的分析使我们得以识别出一种稳健的和乐对比（第 5.3 节）：多项式动力学通过可交换生成元发挥作用（沿轨迹的和乐平凡），而拉普拉斯动力学满足经证明的局部非平凡整体量判据，并在 $d = 2$ 中得出完整受限和乐$\mathrm{SO}(2)$ 这一结果；对于 $d \geq 3$，我们给出了有条件的全整体判据，并将一般情形下的全整体陈述保留为猜想。信息论分析（第 6 节）揭示了深刻的统计学—几何对偶：控制曲率和单射性的同一组谱量，也控制着费舍尔信息和病态条件。

我们所揭示的困难并非让我们毫无希望。例如，可辨识性原理（定理 4）表明，结构能够消解规范歧义，因为对称动力学无法吸收反对称的规范污染。在某些情形下，对齐问题可以得到求解，我们在一个带有锚点节点的可处理特例中证明了这一点（第7.5节）。当这一点成为可能时，下游的 UDE 流程便在实践上可行且具有吸引力（正如数值结果所示）。

然而，离散的有限样本情形正是理论触及自身极限之处。有限样本谱偏置与噪声会与对齐相互作用；表达性动力学族可能拟合伪迹；而和乐会使全局规范一致性在回路上变得不可能，即便局部对齐是完美的。弥合这一间隙是一个开放问题。尤其值得刻画的是，结构约束对齐在何种情况下是稳定的，又在何种情况下从根本上是不稳定的。

若干途径看来有望缩小构造差距：既可在可能时绕过规范，也可在无法避免时量化几何效应。

1. 尽可能在 P空间中工作。 P动力学观点（第5.2节）原则上绕过规范。出自Athreya 等人 [2025] 的欧氏镜像框架表明， P轨迹的规范不变标量摘要（网络状态之间的谱范数普罗克鲁斯特斯距离）可以被一致估计，并且对变点检测具有实际用途。然而，动力学学习需要满矩阵 P，而不仅仅是标量时间间距离。目前，我们缺乏从伯努利样本中恢复 P并反演李雅普诺夫结构的估计理论。 $1/(\lambda_\iota + \lambda_\gamma)$ 放大表明，极小极大速率将强烈依赖于谱隙；明确这一点将阐明何时 P层级的动力学推断在实践中可行，以及镜像的估计保证能否扩展到此处所需的更丰富的微分结构。

2. 量化和乐，而不仅仅是其存在性。除一般非平凡性之外，人们还希望得到依赖于轨迹的“和乐预算”：在一个环上预期会出现多大的规范漂移，它如何作为曲率以及接近秩亏缺程度的函数？

3. 克拉美-罗下界的可达性。对于多项式动力学（平凡和乐、显式的特征值常微分方程），通过基于似然的估计器在 P-轨迹上达到相匹配的上界，似乎是可行的。对于拉普拉斯动力学，和乐表明，要达到克拉美-罗下界，估计器可能需要显式考虑规范输运。

4. 稀疏性改变一切。在稀疏区域 $P = \rho_n XX^\top$ 且满足 $\rho_n \to 0$ 时，单射性会减弱，曲率和 ASE误差都会恶化。这些因素对动力学恢复的综合影响仍鲜有研究。

我们的结果在应用方面具有直接意义。从网络数据中学习可解释的动力学，有望应用于神经科学、生态学和社会系统。阐明框架内 可能实现哪些动力学，有助于指导建模决策。与任何模型一样，随机点积图假设在结构上都很强，因此应将学习到的动力学视为假设，并结合领域知识和模型设定错误进行压力测试。

致谢

我们衷心感谢我们的学生 Connor Stirling James Smith。他在学位论文研究中开创了这项工作，并在数学设定尚未建立之前，便不得不应对非规范等变动力学带来的挑战。

代码可用性

用于复现结果和数据的 Julia 代码可在以下网址获取：https://github.com/gvdr/RDPG_位于_微分_几何。

参考文献

P-A Absil、Robert Mahony 和 Rodolphe Sepulchre。《矩阵流形上的优化算法》。普林斯顿大学出版社，2008 年。数字对象唯一标识符：10.1515/9781400830244。Joshua Agterberg、Qiaosheng Yan 和 Keith Levin。广义随机点积图中潜在位置估计的极小极大速率。*arXiv* 预印本 *arXiv:2307.01942*，2023年。Jesús Arroyo、Avanti Athreya、Joshua Cape、Guodong Chen、Carey E Priebe 和 Joshua T Vogelstein。具有公共不变子空间的多个异质网络的推断。《机器学习研究杂志》，22 (142):1–49，2021年。Avanti Athreya、Carey E Priebe、Minh Tang、Vince Lyzinski、David J Marchette 和 Daniel L Sussman。随机点积图缩放特征向量的极限定理。*Sankhya A*，78(1):1–18，2016年。Avanti Athreya、Donniell E Fishkind、Minh Tang、Carey E Priebe、Youngser Park、Joshua T Vogelstein、Keith Levin、Vince Lyzinski、Yichen Qin 和 Daniel L Sussman。随机点积图的统计推断：综述。《机器学习研究杂志》，18(1):8393–8484，2018年。Avanti Athreya、Zachary Lubberts、Youngser Park 和 Carey E. Priebe。欧几里得镜像与网络时间序列中的动力学。《美国统计学会杂志》，120(550):1025–1036，2025年。数字对象唯一标识符：10.1080/01621459. 2024.2392912。Joshua Cape、Minh Tang 和 Carey E Priebe。二到无穷范数与奇异子空间几何及其在高维统计中的应用。《统计年刊》，47(5):2405–2439，2019。Miles Cranmer。SymbolicRegression.jl：在 Julia 中进行分布式高性能符号回归。https://astroautomata.com/SymbolicRegression.jl/dev/，2023年。Giulio V Dalla Riva 和 Daniel B Stouffer。利用功能性状探索食物网骨干的演化特征。*Oikos*，125(4):446–456，2016年。数字对象唯一标识符：10.1111/oik.02305。Mark I Freidlin 和 Alexander D Wentzell。动力系统的随机扰动。《数学科学基础丛书》，260，1998。Ian Gallagher、Andrew Jones 和 Patrick Rubin-Delanchy。具有稳定性保证的动态网络谱嵌入。载于《神经信息处理系统进展》，第34卷，第10158–10170页，2021年。Markus Heinonen、Cagatay Yildiz、Henrik Mannerström、Jukka Intosalmi 和 Harri Lähdesmäki。利用高斯过程学习未知常微分方程模型。国际机器学习大会，第1959–1968页，2018年。Michel Journée、Francis Bach、P-A Absil 和 Rodolphe Sepulchre。半正定矩阵锥上的低秩优化。《*SIAM*优化杂志》，20(5):2327–2351，2010年。数字对象唯一标识符：10.1137/080731359。Shoshichi Kobayashi 和 Katsumi Nomizu。《微分几何基础》，第1卷。威利出版社，1963年。

Keith Levin、Avanti Athreya、Minh Tang、Vince Lyzinski、Youngser Park 和 Carey E Priebe。多个随机图的 Omnibus 嵌入的中心极限定理及其对多尺度网络推断的启示。*arXiv* 预印本 *arXiv:1705.09355*，2017年。Joshua Daniel Loyal。具有缩放正态混合分布的广义随机点积图的贝叶斯推断。*arXiv* 预印本 *arXiv:2509.19748*，2025年。Estelle Massart 和 P-A Absil。固定秩半正定矩阵流形上具有简单测地线的商几何。*SIAM* 矩阵分析与应用杂志，41(1):171–198，2020 年。数字对象唯一标识符：10.1137/ 18M1231389。Estelle Massart、Julien M Hendrickx 和 P-A Absil。赋予 Bures–Wasserstein 度量的固定秩半正定矩阵流形的曲率。载于信息几何科学（*GSI 2019*），计算机科学讲义第11712卷，第739–748页。Springer，2019。Barrett O'Neill。浸没的基本方程。密歇根数学杂志，13(4):459–469，1966。数字对象唯一标识符：10.1307/mmj/1028999604。Timothée Poisot、Daniel B Stouffer 和 Dominique Gravel。超越物种：生态相互作用网络为何随空间和时间而变化。*Oikos*，124(3):243–251，2015。数字对象唯一标识符：10.1111/oik.01719。Christopher Rackauckas、Yingbo Ma、Julius Martensen、Collin Warner、Kirill Zubov、Rohit Supekar、Dominic Skinner、Ali Ramadhan 和 Alan Edelman。用于科学机器学习的通用微分方程。*arXiv* 预印本 *arXiv:2001.04385*，2020 年。Patrick Rubin-Delanchy、Joshua Cape、Minh Tang 和 Carey E Priebe。谱嵌入的统计学解释：广义随机点积图。英国皇家统计学会杂志：*B*辑，84(4): 1446–1473，2022。Tanya Strydom、Salomé Bouskila、Francis Banville、Ceres Barros、Dominique Caron、Maxwell J Farrell、Marie-Josée Fortin、Victoria Hemming、Benjamin Mercier、Laura J Pollock、Rogini Runghen、Giulio V Dalla Riva 和 Timothée Poisot。通过系统发育转移低秩网络表示来重建食物网。生态学与进化方法，13(12):2838–2849，2022。数字对象唯一标识符：10.1111/2041-210X.13835。Minh Tang、Joshua Cape 和 Carey E Priebe。随机块模型图的特征值。*arXiv* 预印本 *arXiv:1803.11551*，2018。Minh Tang、Joshua Cape 和 Carey E Priebe。随机块模型的渐近有效估计量：朴素最大似然估计、秩约束最大似然估计和谱估计量。伯努利，28(2):1049–1073，2022。Fangzheng Xie 和 Yanxun Xu。随机点积图的最优贝叶斯估计。*Biometrika*，107(4): 875–889，2020。Fangzheng Xie 和 Yanxun Xu。通过一步法有效估计随机点积图。《美国统计学会杂志》，118(541):651–664，2023年。Stephen J Young 和 Edward R Scheinerman。社会网络的随机点积图模型。载于网络图算法与模型：第五届国际研讨会 *WAW 2007*，第138–149页。Springer，2007。

延后给出的证明

命题5（测度为零）的证明。假设 $\mathcal{F}$ 是一个由 p 个参数构成的有限维族（例如，具有 $p = K + 1$ 个系数的多项式动力学），并将讨论限制在这样一个参数域上：解在 $[0, T]$ 上存在，且随参数和初始条件光滑变化（在 X 中满足局部耐利普希茨条件且参数具有光滑性时，通常可保证这一点；参见 Lee 的《光滑流形导论》第2版，2013，教材级参考文献）。对于每个 $(f, X_0) \in \mathcal{F} \times \mathbb{R}^{n\times d}$，皮卡–林德勒夫定理保证存在唯一解 $X(\cdot; f, X_0) \in C^k([0, T], \mathbb{R}^{n\times d})$（同一参考文献）。解映射 $\Phi : (f, X_0) \mapsto X(\cdot; f, X_0)$ 是光滑的，因此 $\mathcal{M}_{\mathcal{F}} = \mathrm{im}(\Phi)$ 是从一个 m 维流形（带有 $m = p + nd$）到路径空间 $\mathcal{X} = C^k([0, T], \mathbb{R}^{n\times d})$ 的光滑映射的像。

选择 $m + 1$ 连续线性泛函 $\ell_1, \ldots, \ell_{m+1}$ 作用于 $\mathcal{X}$（例如，在不同时刻对特定坐标求值），并定义 $L = (\ell_1, \ldots, \ell_{m+1}) : \mathcal{X} \to \mathbb{R}^{m+1}$。选择这些泛函，使前推 $L_*\mu$ 非退化（等价地，具有满秩协方差），因而相对于 $\mathbb{R}^{m+1}$ 上的勒贝格测度绝对连续。于是，$L \circ \Phi : \mathbb{R}^m \to \mathbb{R}^{m+1}$ 是有限维空间之间的光滑映射，且 $m < m+1$。Sard 定理的一个标准推论（等价地，面积公式）是，这类映射的像在 $\mathbb{R}^{m+1}$ 中的勒贝格测度为零（可参阅教科书级资料：Milnor，《从微分观点看拓扑》，1997 年版，Sard 定理一节）。因此，$L(\mathcal{M}_{\mathcal{F}}) \subseteq \mathrm{im}(L \circ \Phi)$ 的勒贝格测度也为零，故 $\mu(\mathcal{M}_{\mathcal{F}}) \le (L_*\mu)(L(\mathcal{M}_{\mathcal{F}})) = 0$。 □

推论的证明 3（线性动力学：显式费舍尔信息）。 对于线性动力学 $\dot{X} = \alpha_0 X$，解为 $X(t) = e^{\alpha_0 t}X(0)$，从而得到 $P(t) = e^{2\alpha_0 t}P(0)$。灵敏度为 $\partial P_{ij}(t)/\partial\alpha_0 = 2te^{2\alpha_0 t}P_{ij}(0) = 2tP_{ij}(t)$。由于只有一个标量参数，费舍尔信息为：

$$\mathcal{I}(\alpha_0) = \sum_{t=0}^{T}\sum_{i<j}\left(\frac{\partial P_{ij}(t)}{\partial\alpha_0}\right)^2 \cdot \frac{1}{P_{ij}(t)(1 - P_{ij}(t))} = 4\sum_{t=0}^{T} t^2 \sum_{i<j}\frac{P_{ij}(t)}{1 - P_{ij}(t)} \tag{79}$$

关于缩放性论断：当 $P_{ij}(t) \in [\epsilon, 1 - \epsilon]$ 对某个 $\epsilon > 0$ 成立时，内层求和 $\sum_{i<j} P_{ij}/(1 - P_{ij})$ 为 $\Theta(n^2)$（共有 $\binom{n}{2}$ 项，每一项都有界且远离零与无穷）。外层求和 $\sum_{t=0}^{T} t^2 = T(T + 1)(2T + 1)/6 = \Theta(T^3)$。因此 $\mathcal{I}(\alpha_0) = \Theta(T^3 \cdot n^2)$，从而给出 Cramér–Rao 下界 $\mathrm{Var}(\hat{\alpha}_0) \ge \mathcal{I}^{-1} = \Omega(1/(T^3n^2))$。 □

B 到纤维的垂直投影

我们推导将切向量 $Z \in T_X\mathbb{R}^{n\times d}_*$ 投影到垂直子空间 $\mathcal{V}_X = \{X\Omega : \Omega \in \mathfrak{so}(d)\}$ 的公式。

垂直分量 $Z^{\mathcal{V}} = X\Omega^*$ 由正交条件 $Z - X\Omega^* \in \mathcal{H}_X$ 定义（其余部分为水平分量）。根据命题 2，向量 W 在 X 处为水平向量，当且仅当 $X^\top W$ 是对称的。将此应用于 $W = Z - X\Omega^*$：

$$X^\top(Z - X\Omega^*) = X^\top Z - G\Omega^* \quad \text{must be symmetric} \tag{80}$$

其中 $G = X^\top X$ 是正定的。将 $X^\top Z$ 分解为对称部分和反对称部分，$X^\top Z = \mathrm{sym}(X^\top Z) + \mathrm{skew}(X^\top Z)$，条件变为：$\mathrm{skew}(G\Omega^*) = \mathrm{skew}(X^\top Z)$。

由于 Ω^* 是反对称的，且 G 是对称的，因此 $G\Omega^*$ 的反对称部分为：

$$\mathrm{skew}(G\Omega^*) = \frac{1}{2}(G\Omega^* - (G\Omega^*)^\top) = \frac{1}{2}(G\Omega^* + \Omega^* G) \tag{81}$$

其中最后一步用到了 $\Omega^{*\top} = -\Omega^*$ 和 $G^\top = G$。令两者相等：

$$G\Omega^* + \Omega^* G = 2\,\mathrm{skew}(X^\top Z) \tag{82}$$

这是 Ω^* 中的李雅普诺夫方程。由于 G 是正定的，所有的和 $\lambda_\iota + \lambda_\gamma > 0$，因此该方程有唯一解。在 $G = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ 的特征基中，解的各元素为：

$$\Omega^*_{\iota\gamma} = \frac{2\,\mathrm{skew}(X^\top Z)_{\iota\gamma}}{\lambda_\iota + \lambda_\gamma} \tag{83}$$

对于命题 9中的应用，$Z = SX$ 满足 S 反对称。于是 $X^\top Z = X^\top SX$，它已经是反对称的（因为 $(X^\top SX)^\top = X^\top S^\top X = -X^\top SX$），因此 $\mathrm{skew}(X^\top SX) = X^\top SX$，李雅普诺夫方程变为 $G\Omega^* + \Omega^* G = 2X^\top SX$。

垂直括号分量的 C 范数

我们推导出 $\|[\bar{\xi}_1, \bar{\xi}_2]^{\mathcal{V}}\|^2$的显式公式，该公式在命题 9中给出。

垂直分量为 $X\Omega^*$，其中 $\Omega^*_{\iota\gamma} = 2(X^\top SX)_{\iota\gamma}/(\lambda_\iota + \lambda_\gamma)$。其中 $S = -[M_1, M_2]$（第 B 节）。其在环境欧几里得度量下的平方范数为:

$$\|X\Omega^*\|^2 = \mathrm{tr}((X\Omega^*)^\top X\Omega^*) = \mathrm{tr}(\Omega^{*\top} G\Omega^*) \tag{84}$$

在 $G = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ 的特征基下计算:

$$\mathrm{tr}(\Omega^{*\top} G\Omega^*) = \sum_{\iota,\gamma} \lambda_\iota (\Omega^*_{\iota\gamma})^2 \tag{85}$$

（利用当 $B = \mathrm{diag}(b_1, \ldots, b_d)$ 时 $\mathrm{tr}(A^\top BA) = \sum_a b_a \sum_b A^2_{ab}$；由于 $\Omega^*_{\iota\iota} = 0$，对角项消失）。将有序对 (ι, γ) 和 (γ, ι) 配对，其中 $\iota < \gamma$，并利用反对称性（$\Omega^*_{\gamma\iota}$）$^2 = (\Omega^*_{\iota\gamma})^2$:

$$= \sum_{\iota<\gamma} (\lambda_\iota + \lambda_\gamma)(\Omega^*_{\iota\gamma})^2 \tag{86}$$

代入 $\Omega^*_{\iota\gamma} = 2(X^\top SX)_{\iota\gamma}/(\lambda_\iota + \lambda_\gamma)$:

$$= \sum_{\iota<\gamma} (\lambda_\iota + \lambda_\gamma) \cdot \frac{4(X^\top SX)^2_{\iota\gamma}}{(\lambda_\iota + \lambda_\gamma)^2} = 4 \sum_{\iota<\gamma} \frac{(X^\top SX)^2_{\iota\gamma}}{\lambda_\iota + \lambda_\gamma} \tag{87}$$

剩下的工作是将 $(X^\top SX)_{\iota\gamma}$表示为 $(U^\top [M_1, M_2]U)_{\iota\gamma}$的形式。将 $X = U\Lambda^{1/2}$写成典范规范（其中 U的正交归一列张成 $\mathrm{col}(X)$，且 $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$）：$X^\top SX = \Lambda^{1/2} U^\top SU \Lambda^{1/2}$，因此 $(X^\top SX)_{\iota\gamma} = \sqrt{\lambda_\iota}(U^\top SU)_{\iota\gamma}\sqrt{\lambda_\gamma}$。由于 $S = -[M_1, M_2]$: $(X^\top SX)^2_{\iota\gamma} = \lambda_\iota \lambda_\gamma (U^\top [M_1, M_2]U)^2_{\iota\gamma}$。代入后:

$$\|X\Omega^*\|^2 = 4 \sum_{\iota<\gamma} \frac{\lambda_\iota \lambda_\gamma}{\lambda_\iota + \lambda_\gamma} [(U^\top [M_1, M_2]U)_{\iota\gamma}]^2 \tag{88}$$

这就是命题9中给出的公式。 $1/(\lambda_\iota + \lambda_\gamma)$ 这种加权方式表明，当一对特征值都较小时，曲率会被放大，这与纤维丛的联络系数结构相符。

D 有向图扩展

本附录旨在概述该框架如何适用于有向图，具体细节从略。

对于有向图，每个节点都有源位置 g_i 和目标位置 r_i。概率矩阵为 $P = GR^\top$（不对称）。

规范群变为（G、R）$\sim$（GQ、RQ），对于 $Q \in O(d)$，不可见动力学为 $G = GA$、$R = RA$，其中 $A \in \mathfrak{so}(d)$。

规范一致架构推广为:

$$\dot{G} = N_G(P)G, \quad \dot{R} = N_R(P)R \tag{89}$$

对（N_G、N_R）施加适当的对称性约束，以确保 $P = GR^\top$ 上的诱导动力学定义良好。纤维丛框架仍然适用：全空间是成对的（G、R）$\in \mathbb{R}^{n\times d}_* \times \mathbb{R}^{n\times d}_*$，结构群为 $O(d)$ 并以对角方式作用；联络、曲率和整体单值理论同样适用，并具有相同的李雅普诺夫型结构，不过 P的不对称性会给谱分解引入额外的复杂性。