通过曲率集区分动态数据

Nadezhda Belova[1], Maxwell Goldberg[1], Facundo M´emoli[1], Sriram Raghunath[1], 和 Andrew Xie[1]

[1]罗格斯大学数学系

2026年3月6日

arXiv:2603.04624v1 [math.AT] 4 Mar 2026

摘要

拓扑数据分析（TDA）技术已被证明能够有效研究动态系统中产生的时变数据，例如动物群聚行为和神经科学中的时空模式。早期算法利用持久图的高效更新来处理动态数据，但难以区分这样一类行为：它们在每个固定时刻都是等距的，却在性质上有所不同。Kim 和 M´emoli 提出了针对动态度量空间的时空持久性框架，从而解决了这一局限，并由此得到多参数持久性模。然而，这些模带来了计算上的挑战。

为解决这一问题，我们借鉴了 G´omez 和 M´emoli 的研究成果。他们指出，对度量空间中由 ($2k+2$) 个点组成的子集（称为主曲率集）计算 Rips 复形的同调，既易于处理，又能提供丰富信息。我们将这一思路推广到动态情境，提出动态曲率集持久同调，并将 Kim 和 M´emoli 的时空框架应用于曲率集。我们证明，由此得到的多参数持久模可分解为区间；事实上，它们还具备一种更强的性质，我们称之为反链可分解。利用这一性质，我们提出了一种新算法，可高效计算任意可反链分解模之间的侵蚀距离 d_E （由 Patel 提出），其中包括但不限于通过我们的方法构造的模。此外，我们的方法对于 Kim 和 M´emoli 提出的时变数据集之间的广义格罗莫夫–豪斯多夫距离具有稳定性。这使我们能够建立一套稳健的计算流程来区分动态数据；在 Boids模型的实验中，我们成功检测到了参数变化。

1 引言

拓扑数据分析已被证明是分析复杂动态系统（尤其是生物学和神经科学领域的系统）的一种相关且有效的工具 [18, 7, 5]。这些方法有一个共同点：使用拓扑不变量（例如条形码）表示系统在各个时间切片上的状态，然后分析不变量的时间序列，以检测行为上的相似性或某些模式/基元的出现。

正如 [11], 所阐明的那样，按时间点计算拓扑不变量会造成信息损失，因为存在这样的动态数据示例：它们在每个固定时间步上都是等距的，但行为在定性上却有所不同。图1展示了一个这样的例子。为了构造信息量更高的不变量， [11] Kim 和 M´emoli 构造了时空持久同调，它生成由时间区间和尺度参数共同索引的多参数模。这提高了区分能力，但代价是产生复杂的多参数模。

为减轻这一多参数构造的计算成本，我们考虑曲率集这一思想，该思想最早由 Gromov 在 [9] 中提出。G´omez 和 M´emoli 在 [10],中运用了这一思想，并证明了对于包含点数足够少的点云，可以快速以闭式形式计算 Rips 过滤的同调，同时仍保留信息量。通过考虑任意规模输入数据中所有“足够小”的子集，可以推导出一种易于处理的不变量。

1.1 贡献

我们提出了一种用于动态数据拓扑分析的构造，利用曲率集的思想，简化了 [11] 中的时空构造的计算。由此得到的不变量稳定且易于计算。



图1：两个表现出截然不同定性行为的动态度量空间。这里，$r > 0$、x_1，x_3，y_1和 y_3 固定在图示坐标值，而 $x_2(t) = r \sin t$ 和 $y_2(t) = r|\sin t|$。对任意固定的 t，所得配置作为度量空间彼此等距（当 x_2 位于 x_1 和 x_3的中点左侧时，将左侧空间关于中点反射即可表明它们等距）。然而，这两种动态行为显然不同。图片由 Kim 和 M´emoli 提供；参见 [11]。

此外，我们证明，这种构造产生的多维持久模是反链可分解的：也就是说，它们是区间可分解的，且区间分解中任意两个不同的直和项在非正式意义上都是“不可比较的”。这些模还具有薄性等其他性质，即逐点维数为 0 或 1，参见定义 2.14。

我们提出了一种新算法，可高效计算任意两个可分解为反链的持久模之间的侵蚀距离 [15] 。随后，我们将此算法推广至计算任意两个具有广义凸性的模之间的侵蚀距离，参见 6.16。值得强调的是，我们提出的侵蚀距离算法并不依赖于我们特定构造的细节。

最后，我们通过分析由 [17] 中介绍的类鸟体模型生成的动态数据，展示了我们构造的可处理性和有效性。

2 预备知识

2.1 动态数据建模

我们沿用 [12] 和 [11]，采用以下动态数据集的数学表示。

定义 2.1（动态度量空间（DMS） [12]）。一个动态度量空间（X，$d_X(\cdot)$）是一个有限集X，并配有一个函数 $d_X\colon \mathbb{R} \times X \times X \to \mathbb{R}_{\geq 0}$，满足

(1) 对每个 $t \in \mathbb{R}$，$(X,\ d_X(t))$ 都是一个伪度量空间；(2) 对每一对 $x,\ x' \in X$，映射 $t \mapsto d(t)(x,\ x')$ 是连续的

定义 2.2（半度量空间）。对于集合 X，函数 $d_X\colon\ X \times X \to \mathbb{R}_{\geq 0}$ 是半度量，如果它满足：

- $d_X(x, x) = 0$ 对所有 $x \in X$
- $d_X(x, x') = d_X(x', x)$ 对所有 $x, x' \in X$

称二元组 (X, d_X) 为半度量空间。

定义 2.3 ([11])。给定任意 DMS $\gamma_X = (X, d_X(\cdot))$ 以及区间 $I \subset \mathbb{R}$，X上的半度量 $d_X(I)(\cdot,\cdot)$ 由 $d_X(I)(x, x') = \inf_{t \in I} d_X(t)(x, x')$ 给出，对任意 $x, x' \in X$。令 $\gamma_X(I) := (X, d_X(I))$ 表示相应的半度量空间。

为了恰当地编码动态数据集的信息，我们的数据表示必须结合不同时间的信息。我们回顾 [11] 中以下一系列定义，以定义 DMS 的时空 Rips 过滤。它是一种不变量，用于记录在时间和尺度变化过程中持续存在的数据特征。

尽管 Rips 复形和过滤通常定义在度量空间上，但也可以在半度量空间上构造。

定义 2.4（Rips 复形）. 设 (X, d_X) 为一个半度量空间。对每个 $\delta \in \mathbb{R}_+$，我们将 X 在尺度 δ 处的*Rips* 复形 定义为 $\mathcal{R}_\delta(d_X)$，即 X 上的一个抽象单纯复形，其中对有限的 $\sigma \subseteq X$，$\sigma \in \mathcal{R}_\delta(d_X)$ 成立当且仅当对所有 $x, y \in \sigma$ 均有 $d_X(x, y) \leq \delta$。

用 **Simp** 表示以单纯映射为态射的抽象单纯复形范畴。另请注意，我们将 $\mathbb{R}$ 视为一个偏序范畴。用 **Vec** 表示向量空间与线性映射的范畴（在固定域 $\mathbb{F}$ 上）。

定义 2.5（Rips 过滤）。设 (X, d_X) 是一个半度量空间。(X, d_X) 的 ***Rips*** 过滤 是函子 $\mathcal{R}_\bullet(d_X) : \mathbb{R}_+ \to \mathbf{Simp}$，定义如下：对每个 $\delta \in \mathbb{R}_+$，赋予 $\mathcal{R}_\delta(d_X)$。对于每个态射$\delta \le \delta' \in \mathbb{R}$，也赋予包含映射 $\mathcal{R}_\delta(d_X) \hookrightarrow \mathcal{R}_{\delta'}(d_X)$。

这里用 **Int** 表示 $\mathbb{R}$ 中由包含关系构成的紧区间偏序集。并且 $\mathbf{Int} \times \mathbb{R}$ 是乘积偏序集，也视为一个偏序范畴。我们将此偏序集记为 **Dyn**。

现在我们已准备好定义时空 Rips 过滤，稍后将用它来构造我们提出的不变量。

定义 2.6（时空 Rips 过滤 [11]）。给定任意 DMS $\gamma_X = (X, d_X(\cdot))$，其时空 ***Rips*** 过滤 是函子 $\mathcal{R}^{\mathrm{lev}}(\gamma_X) : \mathbf{Dyn} \to \mathbf{Simp}$，定义如下：对每个 $(I, \delta) \in \mathbf{Dyn}$，赋予$\mathcal{R}_\delta(\gamma_X(I))$。对于每个态射 $(I, \delta) \le (I', \delta') \in \mathbf{Dyn}$，赋予包含映射 $\mathcal{R}_\delta(\gamma_X(I)) \hookrightarrow \mathcal{R}_{\delta'}(\gamma_X(I'))$。

同样，按照 [11]，我们将DMS（X, γ_X）的第 k个动态Rips持续同调模定义为

$$H_k(\mathcal{R}^{\mathrm{lev}}(X, \gamma_X)) : \mathbf{Dyn} \to \mathbf{Vec}$$

其中 H_k 是第 k个单纯同调函子（在固定域 $\mathbb{F}$上）。这是持久模的一个例子，即从偏序范畴到模范畴的函子。

注记 2.7（**Dyn**-模之间的交错距离）. 所有以**Dyn**为指标的持久模空间由交错距离 $d_I^{\mathbf{Vec}}$赋予度量。关于该交错距离的详细阐述，见 [11, 第4节。1]。

由于t 偏序集**Dyn** 是多维的，标准结果 [3] 表明计算问题 s
关于i 这些持久模之间的交错距离通常难以计算。

2.2 持久模的性质

一种特别重要的持久模是区间可分解持久模。我们在任意偏序集上给出定义：

定义 2.8（序连通性）。设 $(\mathcal{P}, \trianglelefteq)$ 是一个偏序集，且 $S \subset \mathcal{P}$。若存在一个由 S 中元素构成的序列 $p = x_0, \dots, x_n = q$，使得相邻元素均可比较，则称 $p, q \in S$ 在 S 中是序连通的；也就是说，对于所有 $i \in \{0, \dots, n-1\}$，$x_i \ge x_{i+1}$ 或 $x_{i+1} \ge x_i$。

S 中的序连通性是 S 中元素上的一种等价关系。因此，我们可以将 S 划分为等价类，即 $S = \bigsqcup_{\alpha \in A} C_\alpha$。我们称每个等价类 C_α 是 S 的一个序连通分支。

定义 2.9（区间）。对于偏序集 $(\mathcal{P}, \trianglelefteq)$，$\mathcal{P}$ 的一个区间 是满足以下条件的子集 $I \subset \mathcal{P}$：

(i)（连通性）对于任意 $p, q \in I$，元素 p 和 q 在 I中是序连通的。

(ii)（凸性）对于任意 $p, q \in I$ 和 $r \in \mathcal{P}$，若 $p \le r \le q$，则 $r \in I$。

$\mathcal{P}$ 上的所有区间的集合记为**Int**($\mathcal{P}$)。

以下是区间的一种特殊情形，将在后续章节中用到：

定义 2.10（线段）。对于偏序集 $(\mathcal{P}, \trianglelefteq)$以及满足 $p \le q$的某个 $p, q \in \mathcal{P}$，线段 $[p, q]$ 为定义为该集合

$$\{x \in \mathcal{P} : p \le x \le q\}.$$

我们将 $\mathcal{P}$ 上所有线段的集合记为 **Seg**($\mathcal{P}$)。该集合可以看作一个按包含关系排序的偏序范畴，并且同构于偏序集 $\mathcal{P}^{\mathrm{op}} \times \mathcal{P}$。这是因为，当且仅当 $[p_1, q_1] \subset [p_2, q_2]$、$p_1 \ge p_2$ 和 $q_1 \le q_2$。

$\mathcal{P}$ 的某些子集可以用一种自然的方式构造持久模：

定义 2.11（示性模）。设 $\mathcal{P}$ 为一个偏序集，I 为 $\mathcal{P}$ 的一个凸子集（定义 2.9 的性质 **(ii)** 所指的意义下）。示性模 $\mathbb{I}_I : \mathcal{P} \to \mathbf{Vec}$ 是由以下方式定义的函子：

$$\mathbb{I}_I(p) := \begin{cases} \mathbb{F} & p \in I \\ 0 & p \notin I \end{cases} \qquad \mathbb{I}_I(p \le q) := \begin{cases} \mathbf{Id}_{\mathbb{F}} & p, q \in I \\ 0 & \text{otherwise} \end{cases}$$

注意，函子性源于 I 是凸的这一事实。

定义 2.12（区间模）。对于偏序集 $(\mathcal{P}, \le)$，若某个持久模 $\mathbb{V}: \mathcal{P} \to \mathbf{Vec}$ 满足：对于某个 $\mathcal{P}$ 的区间 I，$\mathbb{V} \cong \mathbb{I}_I$，则称其为区间模。

定义 2.13（区间可分解模）。设 $\mathcal{P}$ 为偏序集，$\mathbb{V}: \mathcal{P} \to \mathbf{Vec}$ 为持久模。若存在一组关于 $\mathcal{P}$ 的区间 $\{I_\alpha\}_{\alpha \in A}$，使得

$$\mathbb{V} \cong \bigoplus_{\alpha \in A} \mathbb{I}_{I_\alpha}.$$

另一类重要的模如下所述。

定义 2.14。设 $\mathcal{P}$ 为偏序集，$\mathbb{V}: \mathcal{P} \to \mathbf{Vec}$ 为持久模。若对所有 $p \in \mathcal{P}$，$\dim \mathbb{V}(p) \in \{0, 1\}$，则称 $\mathbb{V}$ 为薄模。

注意，区间模是薄模，但反过来一般不成立。对于薄模，自然可以作如下定义：

定义 2.15（薄模的支撑）。薄模 $\mathbb{V}: \mathcal{P} \to \mathbf{Vec}$ 的支撑记作 $\operatorname{supp}\mathbb{V}$，是子集 $\operatorname{supp}\mathbb{V} := \{p \in \mathcal{P} : \mathbb{V}(p) \cong \mathbb{F}\}$。

现在我们定义本文所分析的主要模类：

定义 2.16 (反链可分解模**)**。设 $\mathcal{P}$ 是一个偏序集，且 $\mathbb{V}: \mathcal{P} \to \mathbf{Vec}$ 是一个持久模。若存在一族区间 $\{S_\alpha\}_{\alpha \in A}$，使得 $\mathbb{V} \cong \bigoplus_{\alpha \in A} \mathbb{I}_{S_\alpha}$ 且满足以下条件，则称 $\mathbb{V}$ 是反链可分解 的：

$$p \in S_\alpha, q \in S_\beta \implies p \not\le q \text{ and } q \not\le p \qquad \forall \alpha \neq \beta.$$

非正式地说，这意味着区间分解中的所有直和项都“不可比较”。

为简洁起见，我们将使用缩写 ACD 来指代反链可分解这一性质 y.

注 2.17 (ACD 模的性质**)**。若 $\mathbb{V}: \mathcal{P} \to \mathbf{Vec}$ 是一个 ACD 模，则以下结论成立 e:

- $\mathbb{V}$ 是一个薄模。反设不然：那么在 $\mathbb{V}$ 的区间分解中必定存在两个直和项 $\mathbb{I}_S$ 和 $\mathbb{I}_T$，满足 $S \cap T \neq \emptyset$。但这样一来，就存在某个 x，使得 $x \in S$ 和 $x \in T$，当然还有 $x \le x$：然而这与 $\mathbb{V}$ 是 ACD 矛盾。因此，$\mathbb{V}$ 必定是薄模。

- $\operatorname{supp}\mathbb{V}$ 是 $\mathcal{P}$ 的一个凸子集。反设不然：那么存在 p，r，$q \in \mathcal{P}$ 满足 $p \le r \le q$，使得 p，$q \in \operatorname{supp}\mathbb{V}$，但 $r \notin \operatorname{supp}\mathbb{V}$。由于区间是凸的，因此在 $\mathbb{V}$ 的区间分解中必定存在两个不同的直和项 $\mathbb{I}_S$，$\mathbb{I}_T$，使得 $p \in S$，$q \in T$。然而，p 和 q 可比较这一事实，与 $\mathbb{V}$ 是一个 ACD 模相矛盾。

请注意，反命题（具有凸支撑的薄模）并不必然意味着 ACD；这类模通常甚至未必是区间可分解的。

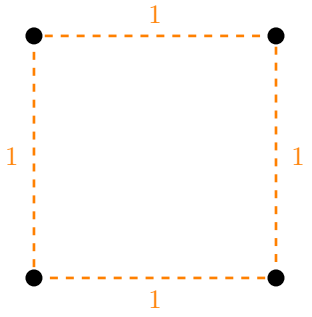
最后，出于计算目的，当指标偏序集是 $\mathbb{R}^d$ 且采用乘积序时，我们引入有界性的概念。

定义 2.18（有界 ACD 模）。对于 ACD 模 $\mathbb{V}: \mathbb{R}^d \to \mathbf{Vec}$，我们称 $\mathbb{V}$ 是有界的 if $\operatorname{supp}(\mathbb{V})$ 是 $\mathbb{R}^d$ 的有界子集。

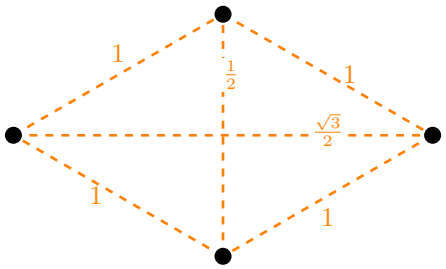
3 持久模的曲率集简化

为了说明我们在一般情形下（针对任意规模的动态数据）的构造，我们首先考虑静态点云的 Rips 持久同调。

设 (X, d_X) 是一个有限度量空间。如果 X 恰好有三个点，即 $|X| = 3$，那么对于任意 $\delta \ge 0$，$H_1(\mathcal{R}_\delta(d_X)) \cong \{0\}$：这是因为 $\mathcal{R}_\delta(d_X)$ 是一个旗复形，因此一旦它包含一个由三条边组成的圈，也就包含一个填充的 2-单形，从而使 $\mathcal{R}_\delta(d_X)$ 可缩。



(a) 一个空间（X，d_X），其中 $H_1(\mathcal{R}_\delta(d_X))$ 对于 $\delta \in [1, \sqrt{2})$。



(b) 一个空间（Y，d_Y），其中 $H_1(\mathcal{R}_\delta(d_Y))$ 对于所有 $\delta \geq 0$。

图2：4点度量空间的示例。

然而，当 $|X| = 4$ 时，1阶同调会出现一些有趣的变化。请看图2所示的点云。

对于图2a，我们可以看到，在 $\delta = 1$ 时，$\mathcal{R}_1(d_X)$中形成了一个圈，它持续存在于所有$\delta \in [1, \sqrt{2})$，直到在 $\delta = \sqrt{2}$时整个正方形被填满，使得 $\mathcal{R}_{\sqrt{2}}(d_X)$可缩。因此，$X$ 的 Rips 过滤在$\delta \in [1, \sqrt{2})$时具有非平凡的 1阶同调。

另一方面，对于图2b，在 Rips 复形 $\mathcal{R}_\delta(d_Y)$中，菱形的短对角线在 $\delta = 1$时被填满，而这正是外边构成的圈形成的尺度。因此，$H_1(\mathcal{R}_\delta(d_Y))$对于所有 $\delta \geq 0$都平凡。

这些空间属于一种普遍模式：对于4点度量空间及其1阶同调，空间的持久图中至多有一个点。若该点存在，其出生时间由四边形所有边长的最大值决定，死亡时间则由短对角线的长度决定。

这一模式可推广到更高维：对于具有（$2k+2$）个点的度量空间的第 k 阶 Rips 持久同调，其持久图中至多有 1 个点；正如 G´omez 和 M´emoli 在 [10] 中推导的那样，可以通过类似的公式以闭式计算。

现在，我们将其推广到动态情境。首先，给出以下定义：

定义 3.1（n-DMS）。对于任意 $n \in \mathbb{N}$，n-DMS 是一个动态度量空间 $\gamma_X = (X, d_X(\cdot))$ with $|X| = n$，即恰好有 n 个点。

我们将研究任意 $k \in \mathbb{Z}_{\geq 0}$ 下，一个 $(2k+2)$-DMS γ_X 的 $H_k(\mathcal{R}^{\text{lev}}(\gamma_X))$。

注 3.2. 若 γ_X 是一个 $(2k+2)$-DMS，则模 $H_k(\mathcal{R}^{\text{lev}}(\gamma_X))$ 必须是薄的。这由 **Dyn** 中每个 $p \in$ 对应的 $\mathcal{R}^{\text{lev}}(\gamma_X)(p)$ 都是恰有 $(2k+2)$ 个点的半度量空间这一事实，立即可由 [10, 定理 4.4(B)] 得出。[10, 定理 4.4(B)] 的简要表述可见 C 节。

特别地，$H_k(\mathcal{R}^{\text{lev}}(\gamma_X))$ 是薄的这一事实意味着其支撑是良定义的。

为了证明定理 3.5——这是关于动态曲率集构造的主要结构结果之一——我们将依赖以下引理：

引理 3.3. 设 $k \in \mathbb{Z}_{\geq 0}$，并令 $\gamma_X = (X, d_X(\cdot))$ 为一个$(2k+2)$-*DMS*。令 $\mathbb{V} = H_k(\mathcal{R}^{\text{lev}}(\gamma_X))$。那么，对任意满足 $p \leq q$ 的 $p, q \in$ 支撑 $\mathbb{V}$，我们有 $\mathcal{R}^{\text{lev}}(\gamma_X)(p) = \mathcal{R}^{\text{lev}}(\gamma_X)(q)$ 作为抽象单纯复形成立。此外，结构映射 $\mathbb{V}(p \leq q)$ 是一个同构。

证明。 根据**Dyn**的定义，存在区间 I_1, I_2 和 $\delta_1, \delta_2 \in \mathbb{R}_+$，使得 $p = (I_1, \delta_1)$ 且$q = (I_2, \delta_2)$。由于 $p, q \in$ supp $\mathbb{V}$,

$$H_k(\mathcal{R}_{\delta_1}(\gamma_X(I_1)) \cong H_k(\mathcal{R}_{\delta_2}(\gamma_X(I_2)) \cong \mathbb{F},$$

这意味着 $\text{dgm}_k^{\text{VR}}(\gamma_X(I_1))$ 和 $\text{dgm}_k^{\text{VR}}(\gamma_X(I_2))$ 均非空。根据 [10, 定理 4.3]（在 C 节中重述），这意味着

$$\mathcal{R}^{\text{lev}}(\gamma_X)(p) = \mathcal{R}_{\delta_1}(\gamma_X(I_1)) \cong \mathfrak{B}_{k+1} \quad \text{and} \quad \mathcal{R}^{\text{lev}}(\gamma_X)(q) = \mathcal{R}_{\delta_2}(\gamma_X(I_2)) \cong \mathfrak{B}_{k+1}$$

作为抽象单纯复形，其中 $\mathfrak{B}_{k+1}$ 是（$k+1$）维交叉多胞体。因此

$$\mathcal{R}^{\text{lev}}(\gamma_X)(p) \cong \mathcal{R}^{\text{lev}}(\gamma_X)(q).$$

但由函子性可知，因为 $p \le q$，我们知道 $\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p) \subseteq \mathcal{R}^{\mathrm{lev}}(\gamma_X)(q)$：结合 $\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p) \cong \mathcal{R}^{\mathrm{lev}}(\gamma_X)(q)$ 这一事实，我们得到 $\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p) = \mathcal{R}^{\mathrm{lev}}(\gamma_X)(q)$。结构映射 $\mathbb{V}(p \le q)$ 由包含映射 ι 诱导：$\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p) \to \mathcal{R}^{\mathrm{lev}}(\gamma_X)(q)$，它是一个单纯复形同构。因此 $\mathbb{V}(p \le q)$ 是一个同构。 □

利用上述描述，我们可以得到以下结论：

注 3.4（支撑集的凸性）。设 γ_X 是一个 $(2k+2)$-DMS。那么，$\mathrm{supp}\ H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$ 是**Dyn**中的一个凸子集。

证明。令 $\mathbb{V} = H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$，并设 $p, q, r \in \mathrm{Dyn}$ 满足 $p \le q \le r$ 且符合 $p, r \in \mathrm{supp}\ \mathbb{V}$。为证明凸性，我们必须证明 $q \in \mathrm{supp}\,\mathbb{V}$。根据前面的假设，结构映射 $\mathbb{V}(p \le r)$ 必须是同构，即秩为 1。但根据函子性，$\mathbb{V}(p \le r) = \mathbb{V}(q \le r) \circ \mathbb{V}(p \le q)$，因此映射 $\mathbb{V}(q \le r) : \mathbb{V}(q) \to \mathbb{V}(r)$ 和 $\mathbb{V}(p \le q) : \mathbb{V}(p) \to \mathbb{V}(q)$ 的秩都必须为 1。因此，$\mathbb{V}(q)$ 不可能是平凡的，从而 $q \in \mathrm{supp}\,\mathbb{V}$。 □

由此得到一个结构定理，描述 $(2k+2)$-DMS γ_X 的 $H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$。特别地，我们证明了区间可分解性，并给出了区间分解的构造性描述。

定理 3.5（区间可分解性）。设 $k \in \mathbb{Z}_{\ge 0}$，且 $\gamma_X = (X, d_X(\cdot))$ 是一个 $(2k+2)$-*DMS*。那么 $H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$ 是区间可分解的。此外，$H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X)) \cong \bigoplus_{\alpha \in A} \mathbb{I}_{S_\alpha}$，其中 $\{S_\alpha\}_{\alpha \in A}$ 是支撑 $H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$ 的序连通分支。

证明。设 $\mathbb{V} = H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$，并设 $\{S_\alpha\}_{\alpha \in A}$ 是 $\mathrm{supp}\ \mathbb{V}$ 的序连通分支分解。我们首先证明 $\mathbb{V} \cong \mathbb{I}_{(\mathrm{supp}\,\mathbb{V})}$。

根据引理 3.3，对任意 $p, q \in \mathrm{supp}\ \mathbb{V}$ 使得 $p \le q$，可得 $\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p) = \mathcal{R}^{\mathrm{lev}}(\gamma_X)(q)$ 作为单纯复形。因此，对任意固定的 $\alpha \in A$，对所有 $p, q \in S_\alpha$，有 $\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p) = \mathcal{R}^{\mathrm{lev}}(\gamma_X)(q)$。由于底层单纯复形相等，我们可以等同它们的链复形。对每个 α，任取一个 $p_\alpha \in S_\alpha$。因为 $\mathbb{V}(p_\alpha) = H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X)(p_\alpha)) \cong \mathbb{F}$，我们可以选取一个生成 $\mathbb{V}(p_\alpha)$ 的代表循环 $[\xi_\alpha]$。从而对所有 $q \in S_\alpha$，$[\xi_\alpha]$ 生成 $\mathbb{V}(q)$。

对任意 $p \in \mathbf{Dyn}$，我们可以如下定义映射 $f_p : \mathbb{V}(p) \to \mathbb{I}_{(\mathrm{supp}\,\mathbb{V})}(p)$ 与 $g_p : \mathbb{I}_{(\mathrm{supp}\,\mathbb{V})}(p) \to \mathbb{V}(p)$

$$f_p(x) = \begin{cases} t & \exists!\, \alpha \in A \text{ s.t. } p \in S_\alpha, x = t[\xi_\alpha] \\ 0 & \text{otherwise} \end{cases} \qquad g_p(t) = \begin{cases} t[\xi_\alpha] & \exists!\, \alpha \in A \text{ s.t. } p \in S_\alpha \\ 0 & \text{otherwise} \end{cases}$$

它们满足自然性，且互为逆映射。因此，$\mathbb{V} \cong \mathbb{I}_{(\mathrm{supp}\,\mathbb{V})}$。由此可得，$\mathbb{V} \cong \mathbb{I}_{(\mathrm{supp}\,\mathbb{V})} \cong \bigoplus_{\alpha \in A} \mathbb{I}_{S_\alpha}$，它是 $\mathbb{V}$ 的区间分解。 □

下　　述结论由定理 3.5 和备注 3.4 可立即得出：

推论 3.6。对于任意 $k \in \mathbb{Z}_{\ge 0}$ 和 γ_X，若其为一个 $(2k+2)$-*DMS*，则模 $H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X))$ 是 *ACD*（定义 *2.16* 意义下）。

证明。已知，对于一族区间 $\{S_\alpha\}_{\alpha \in A}$，$H_k(\mathcal{R}^{\mathrm{lev}}(\gamma_X)) \cong \bigoplus_{\alpha \in A} \mathbb{I}_{S_\alpha}$。剩下只需验证不可比较性。假设这不成立，那么对于某个 $\alpha \ne \beta$，存在 $p \in S_\alpha$，$q \in S_\beta$，使得 $p \le q$。但这是不可能的，因为序连通分支 $\{S_\alpha\}_{\alpha \in A}$ 是极大的，而存在这样的 p，q 意味着 $S_\alpha \cup S_\beta$ 是序连通的，矛盾。 □

一般而言，区间可分解模在多参数持久性模空间中并不典型：特别是，Bauer 和 Scoccola 已在 [1] 中证明，不可分解持久模在多参数持久性模空间中是稠密的。

我们的构造会生成 ACD 模这一事实带来了显著的计算优势，这些优势将在接下来的章节中得到利用。

4 推广到任意大小的数据

为了利用结构定理 Theorem3.5 和推论 3.6 来分析动态数据，我们必须面对这样一个事实：现实世界数据集的基数可以是任意的，对于任何合理的 k 而言，其基数当然都远大于 $(2k+2)$。

为了推广我们的方法，我们将考虑数据集的所有大小为 $(2k+2)$ 的子集。在实践中，我们将从这些子集中抽取样本，并在第 7 节中介绍一种 k 中心方案，用于在尽可能保持代表性的同时，选出数量可控的子集。

定义 4.1（DMS 的曲率集）。设 $\gamma_X = (X,\ d_X)$ 为一个动态度量空间。 n阶曲率集[1] $K_n(\gamma_X)$ 是动态度量空间的集合

$$K_n(\gamma_X) := \{(S, d_X|_S) : S \subseteq X, |S| \leq n\}.$$

随后，我们可以将 Rips时空持久同调应用于曲率集的每个元素。

定义 4.2（DMS 的持久集）。设 $\gamma_X = (X,\ d_X)$ 为一个动态度量空间，且 $k \in \mathbb{Z}_{\geq 0}$。$(n,\ k)$ 阶持久集记为 $D_k^n(\gamma_X)$，其定义为

$$D_k^n(\gamma_X) := \{H_k(\mathcal{R}^{\text{lev}}(\gamma_S)) : \gamma_S \in K_n(\gamma_X)\}.$$

对于动态度量空间 γ_X，$D_k^n(\gamma_X)$ 是我们将要研究的不变量。我们最常考虑 $D_k^{(2k+2)}(\gamma_X)$，因为根据前面的结果，$D_k^{(2k+2)}(\gamma_X)$ 中持久模的结构尤其简单。为减轻记号负担，定义 $D_k(\gamma_X) :\ = D_k^{(2k+2)}(\gamma_X)$。

为了赋予所有持久集构成的空间以度量，我们有如下度量：

定义 4.3（持久集上的豪斯多夫距离）。设 $n,\ k \in \mathbb{Z}_{\geq 0}$，并设 γ_X 和 γ_Y 为两个 DMS。对于 Dyn 索引模之间的任意度量 ρ，我们有由 ρ 诱导的持久集之间的相关豪斯多夫距离：

$$d_H^{\rho}(D_k^n(\gamma_X), D_k^n(\gamma_Y)) = \max\left\{\sup_{\mathbb{V}\in D_k^n(\gamma_X)}\inf_{\mathbb{W}\in D_k^n(\gamma_Y)}\rho(\mathbb{V},\mathbb{W}),\ \sup_{\mathbb{W}\in D_k^n(\gamma_Y)}\inf_{\mathbb{V}\in D_k^n(\gamma(X))}\rho(\mathbb{V},\mathbb{W})\right\}.$$

为简化记号，我们用 $d_H(D_k^n(\gamma_X)\ D_k^n(\gamma_Y))$ 表示 $d_H^{d_I^{\mathbf{Vec}}}(D_k^n(\gamma_X)\ D_k^n(\gamma_Y))$。也就是说，这是由 $d_I^{\mathbf{Vec}}$ 诱导的豪斯多夫距离（注 2.7）。

从给定的动态度量空间构造持久集具有稳定性：

定理 4.4（动态曲率集的稳定性）。设 $\gamma_X = (X, d_X)$ 和 $\gamma_Y = (Y, d_Y)$ 是两个动态度量空间。那么对于任意 $k \in \mathbb{Z}_{\geq 0}$ 和 $n \in \mathbb{N}$,

$$d_H\left(D_k^n(\gamma_X), D_k^n(\gamma_Y)\right) \leq 2 \cdot d_{dyn}(\gamma_X, \gamma_Y).$$

wh其中 d_{dyn} 是从 *[12]* 出发的 $\lambda = 2$ 广义*Gromov–Hausdorff*距离；参见定义 *A*。 *6.*

在下一节中，我们将探讨定义在以 **Dyn** 为指标的模上的另一种度量 d_E，它诱导的豪斯多夫距离 $d_H^{d_E}$ 也具有稳定性；参见推论 5.11。

关于 d_{dyn} 的完整介绍，参见 A节。定理 4.4 的证明见第8节。

5 侵蚀距离

本节将回顾侵蚀距离（最初由 Patel 在 [15] 中提出，后由 Puuska 在 [16]中推广），并说明它与其他度量（例如交错距离 $d_I^{\mathbf{Vec}}$）之间的联系。这将是我们算法工作的主要目标。

为方便起见，在本节中，如果未指定偏序集 $\mathcal{P}$，则应将**Seg** 理解为**Seg**$(\mathbb{R}^d)$，其中 $\mathbb{R}^d$ 被视为带有乘积序的偏序范畴。参见定义 2.10。

5.1 d_E 在 $\mathbb{R}^n$上的侵蚀距离

考虑带有乘积序的 d 维偏序集 $\mathbb{R}^d$。

定义 5.1。线段 $I = [p, q] \in$**Seg** 的 ε-**加厚**（其中 $\varepsilon \in \mathbb{R}$非负）是线段$[p^{-\varepsilon}, q^{+\varepsilon}]$，记作 $I^{+\varepsilon}$。若 $p \in \mathbb{R}^d$，则 $p^{+\varepsilon} = \sum_i (p^{(i)} + \varepsilon)\vec{e}_i$ 且 $p^{-\varepsilon} = \sum_i (p^{(i)} - \varepsilon)\vec{e}_i$。

注意，对于所有 $I \in$**Seg** 都有 $I \leq I^{+\varepsilon}$。现在我们正式定义 $\mathbb{R}^d$ 模的秩不变量 le.

定义 5.2 ([4]**)**。给定一个 $\mathbb{R}^d$ 模 $\mathbb{V} : \mathbb{R}^d \to \mathbf{Vec}$，其秩不变量 是一个函子 $\mathbf{rk}_{\mathbb{V}} : \mathbf{Seg} \to \mathbb{N}^{\text{op}}$，它将 $[p, q] \in \mathbf{Seg}$ 映射到 $\mathbb{V}$ 中从 p 到 q 的结构映射的秩。

[1]该 “曲率集”一词由 Gromov 在 [9]年创造，因为这些子集可用于恢复曲率 re 信息 来自流形。

请注意，我们断言，对于任意 $\mathbb{R}^d$ 模 $\mathbb{V}$ 和 $I \le J \in \mathbf{Seg}$ （尤其是对于 $I \le I^{+\varepsilon}$），都有 $\mathbf{rk}_{\mathbb{V}}(J) \le \mathbf{rk}_{\mathbb{V}}(I) \in \mathbb{N}$。

现在我们将侵蚀距离 [15] 专门化，从而得到 $\mathbb{R}^d$ 模的秩不变量之间的一个度量 e.

定义 5.3。 秩不变量 $\mathbf{rk}_{\mathbb{V}}$ 和 $\mathbf{rk}_{\mathbb{W}}$ 之间的侵蚀距离 d_{E} 定义为

$$d_{\mathrm{E}}(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) := \inf\Big\{\varepsilon > 0 \;\Big|\; \forall I \in \mathbf{Seg}, \mathrm{rk}_{\mathbb{V}}(I^{+\varepsilon}) \le \mathrm{rk}_{\mathbb{W}}(I), \mathrm{rk}_{\mathbb{W}}(I^{+\varepsilon}) \le \mathrm{rk}_{\mathbb{V}}(I)\Big\}.$$

5.2 Dyn 上的侵蚀距离 d_E

现在我们以[11]的形式引入 Dyn 索引模上的侵蚀距离 d_E。这与 $\mathbb{R}^n$类似，因为**Dyn**可以视为 $\mathbb{R}^{\mathrm{op}} \times \mathbb{R} \times \mathbb{R}$中可容许点 的一个子集。

定义 5.4 ([11]**)**。 定义偏序集 $\mathbb{R}^3_\times$ 为带有乘积序的 $\mathbb{R}^{\mathrm{op}} \times \mathbb{R} \times \mathbb{R}$。定义偏序集$\mathbb{R}^6_\times$ 为带有乘积序的 $\mathbb{R} \times \mathbb{R}^{\mathrm{op}} \times \mathbb{R}^{\mathrm{op}} \times \mathbb{R}^{\mathrm{op}} \times \mathbb{R} \times \mathbb{R}$。

直观地说， $\mathbb{R}^6_\times$ 应被视为 $\left(\mathbb{R}^3_\times\right)^{\mathrm{op}} \times \mathbb{R}^3_\times$。

定义 5.5 （可容许点 [11]）**：** 如果点（t_1， t_2， δ） $\in \mathbb{R}^3_\times$ 满足可容许 条件 $t_1 \le t_2$ 且$\delta \ge 0$，则称其为可容许点。我们可以将偏序集 **Dyn** 视为 $\mathbb{R}^3_\times$中由可容许点构成的子偏序集。

如果点 a =（a_1， a_2， a_3， a_4， a_5， a_6）$\in \mathbb{R}^6_\times$ 满足（a_1， a_2， a_3）、（a_4， a_5， a_6）$\in$Dyn（在上述对应关系下），且（a_1， a_2， a_3）$\le$（a_4， a_5， a_6）属于 $\mathbb{R}^3_\times$，则称点 a 为可容许点。注意，如果这样的 a 是可容许点，那么在将（a_1， a_2， a_3）和（a_4， a_5， a_6）分别与 Dyn 中的点对应后，我们可以将 a 视为 Seg(Dyn) 中的线段［（a_1， a_2， a_3）、（b_1， b_2， b_3）］。

如果点 $\mathbf{a} \in \mathbb{R}^6_\times$ 不是可容许的，我们就称**a** 为不可容许。此外，如果不存在可容许的$\mathbf{b} \in \mathbb{R}^6_\times$ 使得$\mathbf{b} < \mathbf{a}$，则称不可容许点$\mathbf{a} \in \mathbb{R}^6_\times$ 为平凡不可容许。

由此，我们现在可以定义 Dyn 上的秩不变量了。

定义 5.6 （Dyn 上的适配秩不变量**Dyn** [11]）。 设 $\mathbb{V}$： **Dyn** $\to$ **Vec** 是一个持久模。Dyn 的适配秩不变量是函子 $\mathbf{rk}_{\mathbb{V}}$： $\mathbb{R}^6_\times \to \mathbb{Z}_+$，定义如下：对于$\mathbf{a}$ =（a_1，.....， a_6），

$$\mathrm{rk}_{\mathbb{V}}(\mathbf{a}) := \begin{cases} \mathrm{rank}(\mathbb{V}(([a_1, a_2], a_3) \le ([a_4, a_5], a_6)) & \mathbf{a} \text{ is admissible} \\ \infty & \mathbf{a} \text{ is trivially non-admissible} \\ 0 & \text{otherwise} \end{cases}$$

$\mathbf{rk}_{\mathbb{V}}$ 是一个函子这一事实可由 [11, 命题 4.3]看出。

为了定义侵蚀距离的概念，我们需要先定义加厚:

定义 5.7。 对于点$\mathbf{p} \in \mathbb{R}^3_\times$ 以及一个 $\varepsilon > 0$，定义平移 $\mathbf{p}^{+\varepsilon}$:= $\mathbf{p} + \varepsilon \cdot$（$-1$，**1**，**1**）。类似地，定义平移$\mathbf{p}^{-\varepsilon}$:= $\mathbf{p} - \varepsilon \cdot$（$-1$，**1**，**1**）。

对于一点$\mathbf{a} \in \mathbb{R}^6_\times$ 和一个 $\varepsilon > 0$，定义平移

$$\mathbf{a}^{+\varepsilon} := \mathbf{a} + \varepsilon \cdot (1, -1, -1, -1, 1, 1)$$

类似地

$$\mathbf{a}^{-\varepsilon} := \mathbf{a} - \varepsilon \cdot (1, -1, -1, -1, 1, 1).$$

请注意，虽然 $+\epsilon$ 在 $\mathbb{R}^6_\times$ 上的记号与 $+\epsilon$在**Seg**上的记号相同，但它们并非同一种运算。在 $\mathbb{R}^6_\times$中指代变量时，我们将使用粗体加以区分。

适配秩不变量的函子性使我们能够在**Dyn**上构造侵蚀距离:

定义 5.8. 对于 $\mathbb{V}, \mathbb{W} : \mathbf{Dyn} \to \mathbf{Vec}$，适配秩不变量 $\mathbf{rk}_{\mathbb{V}}$与 $\mathbf{rk}_{\mathbb{W}}$ 之间的侵蚀距离 d_{E} 定义为

$$d_{\mathrm{E}}(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) := \inf\Big\{\varepsilon > 0 \;\Big|\; \forall I \in \mathbb{R}^6_\times, \mathrm{rk}_{\mathbb{V}}(I^{+\varepsilon}) \le \mathrm{rk}_{\mathbb{W}}(I), \mathrm{rk}_{\mathbb{W}}(I^{+\varepsilon}) \le \mathrm{rk}_{\mathbb{V}}(I)\Big\}.$$

请注意，这与此前在 $\mathbb{R}^d$ 上定义的侵蚀距离类似，只不过为了避开可容许性问题，我们在 $\mathbb{R}^6_\times$ 上而不是在 **Seg(Dyn)** 上进行研究。

5.3 侵蚀距离的应用

注 5.9. 根据定理 8.2], 侵蚀距离是交错距离的下界。更确切地说，对于以 **Dyn** 为指标的模 $\mathbb{V}, \mathbb{W}$ ，有如下结论：

$$d_E(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) \leq d_I^{\mathbf{Vec}}(\mathbb{V}, \mathbb{W}).$$

这使得侵蚀距离成为交错距离的一种计算代理。由于侵蚀距离是 **Dyn** 索引模上的度量，因此我们可以在持久集上定义与之相关的豪斯多夫距离：

定义 5.10. 对于 $n, k \in \mathbb{Z}_+$ 和 γ_X, γ_Y这两个动态度量空间，我们作如下定义

$$d_H^{\mathrm{ero}}(D_k^n(\gamma_X), D_k^n(\gamma_Y)) := d_H^{d_E}(D_k^n(\gamma_X), D_k^n(\gamma_Y)),$$

其中， d_E 是**Dyn** 模的秩不变量之间的侵蚀距离。注意，这是定义 4.3 的一个特例。

由此立即得到如下推论，

推论 5.11. 结合注 *5.9*、定义 *5.10* 和定理 *4.4*可知，对于任意的n、 $k \in \mathbb{Z}_{\geq 0}$和 γ_X、 γ_Y这两个动态度量空间，我们有

$$d_H^{\mathrm{ero}}(D_k^n(\gamma_X), D_k^n(\gamma_Y)) \leq d_H(D_k^n(\gamma_X), D_k^n(\gamma_Y)) \leq 2 \cdot d_{\mathrm{dyn}}(\gamma_X, \gamma_Y).$$

d_E 和 d_E^{ero}的计算将在算法部分和实验部分的后续内容中讨论。我们将讨论计算$\mathbb{R}^d$指标模的秩不变量之间侵蚀距离的情形。

6 反链可分解模的算法结果

本节将介绍一种新算法，用于在离散化后计算[15], 中引入的侵蚀距离的一种变体，该距离适用于任意两个 $\mathbb{R}^d$指标的反链可分解模（见定义 2.16）。具体而言，若离散化定义域为 $[n]^d$，则该算法的运行时间为 $O(n^{d-1}d\log\ n)$。值得强调的是，这一结论适用于任意偏序集指标ACD模，且与（$2k+2$）构造无关。

本节中的所有持久模均假定为反链可分解模（见定义 *2.16*）。特别地，请回顾 *ACD* 模是薄的且具有凸支撑（见备注 *2.17*）。此外，所有 $\mathbb{R}^d$-模和**Dyn**-模均假定在定义 *2.18* 的意义下有界。

注记 6.1. 虽然本节中的结果是针对 ACD 模给出的，但它们也适用于稍广泛的情形。具体而言，由于侵蚀距离是秩不变量上的度量，因此只有结构映射的秩才起作用。因此，第 6.2 节中概述的算法适用于任何具有凸支撑的薄模。第 6.3 节进一步推广了这一结果，放宽了薄性要求。

6.1 d_E的简化

为了理解定理 6.11 中的算法，我们将针对薄模情形，推导一些关于秩不变量的结果。首先，我们重新表述侵蚀距离的定义，以强调 **Seg** 中的每条线段都有一个阈值 ε 值，超过该值后，该线段便满足侵蚀距离定义中的秩不变量不等式。

定义 6.2. 对于 $I \in$**Seg** 以及 ACD $\mathbb{R}^d$ 模 $\mathbb{V}$ 和 $\mathbb{W}$，我们定义

$$\mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = \inf\{\varepsilon > 0 \mid \mathrm{rk}_{\mathbb{V}}(I^{+\varepsilon}) \leq \mathrm{rk}_{\mathbb{W}}(I),\ \mathrm{rk}_{\mathbb{W}}(I^{+\varepsilon}) \leq \mathrm{rk}_{\mathbb{V}}(I)\}.$$

这使得 d_E 的表达式得以简化，即：

注 6.3。我们看到

$$d_E(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) = \sup_{I \in \mathbf{Seg}} \mathcal{E}_{\mathbb{V},\mathbb{W}}(I).$$

定义 6.4. 对于 $I \in$**Seg** 和 ACD $\mathbb{R}^d$ 模 $\mathbb{V}$，我们定义

$$\mathcal{E}_{\mathbb{V}}(I) := \inf\{\varepsilon > 0 \mid \mathrm{rk}_{\mathbb{V}}(I^{+\varepsilon}) = 0\}.$$

现在，我们将分析 $\mathcal{E}_{\mathbb{V},\mathbb{W}}$ 的行为。

引理 6.5. 设 $\mathbb{V}, \mathbb{W}$ 为 *ACD* $\mathbb{R}^d$ 模，且 $I \in$**Seg**。

- 若 $\mathrm{rk}_{\mathbb{V}}(I) = 0$ 且 $\mathrm{rk}_{\mathbb{W}}(I) = 0$，则 $\mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = 0$。
- 若 $\mathbf{rk}_{\mathbb{V}}(I) = 1$ 且 $\mathbf{rk}_{\mathbb{W}}(I) = 0$ 则 $\mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = \mathcal{E}_{\mathbb{V}}(I)$。
- 若 $\mathbf{rk}_{\mathbb{V}}(I) = 0$ 且 $\mathbf{rk}_{\mathbb{W}}(I) = 1$ 则 $\mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = \mathcal{E}_{\mathbb{W}}(I)$。
- 若$\mathrm{rk}_{\mathbb{V}}(I) = 1$ 且$\mathrm{rk}_{\mathbb{W}}(I) = 1$，则$\mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = 0$。

证明。在 $\mathbf{rk}_{\mathbb{V}}(I) = \mathbf{rk}_{\mathbb{W}}(I)$ 的情形下，对于 $\varepsilon = 0$，条件 $\mathbf{rk}_{\mathbb{V}}(I^{+\varepsilon}) \leq \mathbf{rk}_{\mathbb{W}}(I)$，$\mathbf{rk}_{\mathbb{W}}(I^{+\varepsilon}) \leq \mathbf{rk}_{\mathbb{V}}(I)$ 成立。

在 $\mathbf{rk}_{\mathbb{V}}(I) = 1$ 且 $\mathbf{rk}_{\mathbb{W}}(I) = 0$ 的情形下，对于$\varepsilon = 0$，条件 $\mathbf{rk}_{\mathbb{W}}(I^{+\varepsilon}) \leq \mathbf{rk}_{\mathbb{V}}(I)$ 成立。因此，$\mathcal{E}_{\mathbb{V},\mathbb{W}}(I)$ 将取值为满足 $\mathbf{rk}_{\mathbb{V}}(I^{+\varepsilon}) \leq \mathbf{rk}_{\mathbb{W}}(I) = 0$ 或等价地满足 $\mathbf{rk}_{\mathbb{V}}(I^{+\varepsilon}) = 0$ 的最小 ε。这正是 $\mathcal{E}_{\mathbb{V}}(I)$。

最后一种情形可由前一种情形对称地推出。 □

定义 6.6。对于 ACD $\mathbb{R}^d$ 模 $\mathbb{V}$ 和 $\mathbb{W}$，我们定义

$$\mathbf{Seg}_{\mathbb{V},\mathbb{W}} := \{I \in \mathbf{Seg} \mid \mathrm{rk}_{\mathbb{V}}(I) = 1, \mathrm{rk}_{\mathbb{W}}(I) = 0\}.$$

注 6.7。对于 $I \in$**Seg** 和 ACD $\mathbb{R}^d$ 模 $\mathbb{V}$ 和 $\mathbb{W}$，$I \notin \mathbf{Seg}_{\mathbb{V},\mathbb{W}} \cup \mathbf{Seg}_{\mathbb{W},\mathbb{V}}$ 蕴含$\mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = 0$（由引理 6.5），因此根据注 6.3,

$$d_E(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) = \sup_{I \in \mathbf{Seg}} \mathcal{E}_{\mathbb{V},\mathbb{W}}(I) = \sup_{I \in \mathbf{Seg}_{\mathbb{V},\mathbb{W}} \cup \mathbf{Seg}_{\mathbb{W},\mathbb{V}}} \mathcal{E}_{\mathbb{V},\mathbb{W}}(I).$$

因此

$$d_E(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) = \max\left(\sup_{I \in \mathbf{Seg}_{\mathbb{V},\mathbb{W}}} \mathcal{E}_{\mathbb{V},\mathbb{W}}(I), \sup_{I \in \mathbf{Seg}_{\mathbb{W},\mathbb{V}}} \mathcal{E}_{\mathbb{V},\mathbb{W}}(I)\right).$$

再次应用引理 6.5 可知

$$d_E(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) = \max\left(\sup_{I \in \mathbf{Seg}_{\mathbb{V},\mathbb{W}}} \mathcal{E}_{\mathbb{V}}(I), \sup_{I \in \mathbf{Seg}_{\mathbb{W},\mathbb{V}}} \mathcal{E}_{\mathbb{W}}(I)\right).$$

我们现在给出一些额外的改写，最终将使上确界中必须包含的 $\mathbb{R}^d$ 元素子集的维数最小化，从而降低计算复杂度。

定义 6.8。对于 ACD $\mathbb{R}^d$ 模 $\mathbb{V}$、$\mathbb{W}$ 和 p，$q \in \mathbb{R}^d$ 我们定义:

- $\mathbb{R}^d_{\mathbb{V},\mathbb{W}} := \{u \in \mathbb{R}^d \mid \mathbb{V}(u) \cong \mathbb{F}, \mathbb{W}(u) \cong 0\}$,
- $\mathbb{R}^d_{\mathbb{V}} := \{u \in \mathbb{R}^d \mid \mathbb{V}(u) \cong \mathbb{F}\}$,
- $\mathcal{E}_{小\mathbb{V}}(p) :=$ 上标$_{q \geq p \in \mathbb{R}^d} \mathcal{E}_{\mathbb{V}}([p, q])$,
- $\mathcal{E}_{大\mathbb{V}}(q) :=$ 上$_{p \leq q \in \mathbb{R}^d} \mathcal{E}_{\mathbb{V}}([p, q])$。

注 6.9。由凸性可知，对于 $[p,\,q]\in\mathbf{Seg}$ 和 ACD $\mathbb{R}^d$ 模 $\mathbb{V}$ 与 $\mathbb{W}$，$[p,\,q]\in\mathbf{Seg}_{\mathbb{V},\,\mathbb{W}}$当且仅当 $p\in\mathbb{R}^d_{\mathbb{V},\,\mathbb{W}}$ 且 $q\in\mathbb{R}^d_{\mathbb{V}}$，或 $q\in\mathbb{R}^d_{\mathbb{V},\,\mathbb{W}}$ 且 $p\in\mathbb{R}^d_{\mathbb{V}}$。因此

$$\sup_{I\in\mathbf{Seg}_{\mathbb{V},\,\mathbb{W}}}\mathcal{E}_{\mathbb{V}}(I)=\max\left(\sup_{p\in\mathbb{R}^d_{\mathbb{V},\,\mathbb{W}}}\mathcal{E}^{\mathrm{small}}_{\mathbb{V}}(p),\ \sup_{q\in\mathbb{R}^d_{\mathbb{V},\,\mathbb{W}}}\mathcal{E}^{\mathrm{large}}_{\mathbb{V}}(q)\right).$$

应用注 6.7，我们看到，如果我们定义

$$\mathcal{E}^*(\mathbb{V},\,\mathbb{W})=\max\left(\sup_{p\in\mathbb{R}^d_{\mathbb{V},\,\mathbb{W}}}\mathcal{E}^{\mathrm{small}}_{\mathbb{V}}(p),\ \sup_{q\in\mathbb{R}^d_{\mathbb{V},\,\mathbb{W}}}\mathcal{E}^{\mathrm{large}}_{\mathbb{V}}(q)\right)$$

then

$$d_E(\mathrm{rk}_{\mathbb{V}},\,\mathrm{rk}_{\mathbb{W}})=\max\left(\mathcal{E}^*(\mathbb{V},\,\mathbb{W}),\,\mathcal{E}^*(\mathbb{W},\,\mathbb{V})\right).$$

引理 6.10. 对于 $\mathbb{R}^d$ 的任意线段 $[p,q]$，我们有 $\mathcal{E}_{\mathbb{V}}([p,\,p])\geq\mathcal{E}_{\mathbb{V}}([p,\,q])$。类似地，对于这样的 $[p,q]$，我们也有 $\mathcal{E}_{\mathbb{V}}([q,\,q])\geq\mathcal{E}_{\mathbb{V}}([p,\,q])$。

作为直接推论，对于任意 p、$q\in\mathbb{R}^d$，我们有

$$\mathcal{E}^{\mathrm{small}}_{\mathbb{V}}(p)=\mathcal{E}_{\mathbb{V}}([p,p])\qquad\mathcal{E}^{\mathrm{large}}_{\mathbb{V}}(q)=\mathcal{E}_{\mathbb{V}}([q,q]).$$

证明. 我们将证明 $\mathcal{E}_{\mathbb{V}}([p,\,p])\geq\mathcal{E}_{\mathbb{V}}([p,\,q])$，而 $\mathcal{E}_{\mathbb{V}}([q,\,q])\geq\mathcal{E}_{\mathbb{V}}([p,\,q])$ 的证明可类似地进行。

设 $\varepsilon>0$ 为任意的。因为 $p\leq p^{+\varepsilon}\leq q^{+\varepsilon}$，我们有 $[p,p]^{+\varepsilon}\subseteq[p,q]^{+\varepsilon}$ 作为线段。因为 $\mathrm{rk}_{\mathbb{V}}$ 是反序的，这意味着 $\mathrm{rk}_{\mathbb{V}}([p,p]^{+\varepsilon})\geq\mathrm{rk}_{\mathbb{V}}([p,q]^{+\varepsilon})$。因此，如果对任意 ε 有 $\mathcal{E}_{\mathbb{V}}([p,q])\geq\varepsilon$，那么我们也有 $\mathcal{E}_{\mathbb{V}}([p,p])\geq\varepsilon$。对 ε 取下确界得到 $\mathcal{E}_{\mathbb{V}}([p,\,p])\geq\mathcal{E}_{\mathbb{V}}([p,\,q])$。 □

根据本节的结果，我们已成功将 d_E 处理成适合采用算法方法的形式。

6.2 离散化与算法

为了在计算机内存中表示一个反链可分解的 $\mathbb{V}:\ \mathbb{R}^d\to\mathbf{Vec}$ 模的秩信息，我们引入离散化的概念。固定 $\mathbb{R}^d$的一个离散子集：具有乘积序的格$[n]^d$ 即可（或许经过平移或旋转之后）。然后，我们考虑支撑为 $[n]^d\cap\operatorname{supp}\mathbb{V}$的 $[n]^d$索引薄模。这个 $[n]^d$ 模将是我们将要使用的$\mathbb{V}$的离散化版本。

请注意，计算这些离散化模之间的侵蚀距离，总会给出未离散化模之间侵蚀距离的下界：这是因为离散化偏序集中的每条线段同时也是未离散化偏序集中的线段。

为了量化**Dyn**这一近似的精度，我们为离散化过程引入参数。回顾到**Dyn** 具有两个时间维数和一个空间维度，我们将离散样本选取为一个子格，其在空间方向上有 S 个等间距点，在每个时间方向上有 T 个等间距点。因此，该子集的大小为 ST^2。然而，利用凸性，我们可以将其存储在$\mathcal{O}(T^2)$空间中：对于每个 $I\in\mathbf{Seg}(\mathbb{R})$，都存在一个相关联的区间$(s_I,t_I)$（可能为空），使得$(I,t)\in\operatorname{supp}\mathbb{V}$当且仅当 $t\in(s_I,t_I)$。只需存储这些(s_I,t_I)元组即可：也就是说，对每个时间区间，记录它出现在**Dyn**中时的最小与最大空间尺度。

类似地，$\mathbb{R}^d$ 可以由离散子格 $[n]^d$ 近似，而离散化$\mathbb{R}^d$-模的支撑可以用 $\mathcal{O}(n^{d-1})$ 空间存储：对于每个 $x\in[n]^{d-1}$，由凸性可知，只需存储在 x 处属于支撑的第 d 个坐标的最小值和最大值。

定理 6.11. 存在一种 $\mathcal{O}(n^{d-1}d\log\,n)$ 算法，用于计算两个反链可分解 $[n]^d$ 模的秩不变量之间的侵蚀距离。由此可得到一种 $\mathcal{O}(S+T)^2\log(S+T)$算法，用于计算反链可分解 **Dyn** 模之间的侵蚀距离。

我们将在此详细介绍该算法。该算法采用扫描线方法，以下定义将 s
是必 要的。

定义 6.12。对于乘积序下的离散格 $[n]^d$，对角投影是映射 $\pi:\ [n]^d\to\mathbb{Z}^{d-1}$，其定义为

$$(x_1,\ldots,x_{d-1},x_d)\mapsto(x_d-x_1,\ldots,x_d-x_{d-1}).$$

注意，底空间 $\pi([n]^d)$中的每个点都唯一确定 $[n]^d$中的一条直线。对于 $v\in\pi([n]^d)$，我们将 $\pi^{-1}(v)$称为 v的纤维。

至此，我们现在可以构造该算法 m.

算法。函数 ErosionDistance**(**$\mathbb{V}$**,** $\mathbb{W}$**)**$P_{\mathbb{V}} \leftarrow$ **InitProjectedSweepList(**$\mathbb{V}$**)**$P_{\mathbb{W}} \leftarrow$ **InitProjectedSweepList(**$\mathbb{W}$**) ans** $\leftarrow 0$**对于每条直线** L **（在** $[n]^{d-1}$ **中，平行于 (1, . . . , 1)）执行 将集合** $S_{\mathbb{V}}$ **初始化为空。将集合** $S_{\mathbb{W}}$ **初始化为空。对** $v \in L \cap \pi([n]^d)$ **执行** $\triangleright$ **按** $\mathbb{Z}^{d-1}$**上的字典序迭代**$S_{\mathbb{V}}$**.Update(**$P_{\mathbb{V}}$**,** v**)**$S_{\mathbb{W}}$**.Update(**$P_{\mathbb{W}}$**,** v**)**$a \leftarrow$ **ReconstructMin(**$S_{\mathbb{V}}$**, v)**$b \leftarrow$ **ReconstructMax(**$S_{\mathbb{V}}$**, v)**$c \leftarrow$ **ReconstructMin(**$S_{\mathbb{W}}$**, v)**$d \leftarrow$ **ReconstructMax(**$S_{\mathbb{W}}$**, v)**$x \leftarrow$ **max(Erosion1d(**a**,** b**,** c**,** d**),** **Erosion1d(**c**,** d**,** a**,** b**)) ans** $\leftarrow$ **max(ans** x**) 结束循环 结束循环 返回 ans 结束函数**

函数 InitProjectedSweepList**(**$\mathbb{V}$**) 初始化** P**start** P^{end} **为哈希映射，键为** $\pi([n]^d)$**，值为列表，默认均为空。对于** $x \in [n]^{d-1}$ **执行 (**a**,** b**)** $\leftarrow$ **SupportInterval(**$\mathbb{V}$**,** x**) 将** x **追加到** $P^{\text{start}}[\pi((x, a))]$**将** x **追加到** $P^{\text{end}}[\pi((x, b))]$ **结束循环 返回 (**P^{start}**,** P**end) 结束函数**

这依赖于几个辅助函数，详述如下：

- **SupportInterval(**$\mathbb{V}$**,** x**)**。对于一个索引 $x \in [n]^{d-1}$，它表示 $[n]^d$中某个点的前 $(d-1)$ 个坐标，**SupportInterval(**$\mathbb{V}$**,** x**)** 返回一个区间 (a, b)，使得对于所有$t \in (a, b)$，$(x, t) \in \mathbb{R}^d_{\mathbb{V}}$。若对于所有 t，$(v, x) \notin \mathbb{R}^d_{\mathbb{V}}$，则该区间为空。
- 函数 S.**更新(**($P_{\text{开始}}$ $P_{\text{结束}}$)**,** v**)** 会将满足 $x \in P_{\text{开始}}[v]$ 的所有 S 插入其中，并从 S 中删除所有满足 $x \in P_{\text{结束}}[v]$ 的 x。
- 函数 **ReconstructMin(**S, v**)**。对于 S, v，首先在 S 中找到最小的 x（依据 $\mathbb{Z}^{d-1}$ 上的乘积序）。然后找到唯一的 t，使得 (x, t) 位于 v 的纤维上。返回 t。为确保定义良好，必须维持若干不变量，详见分析。
- 函数 **ReconstructMax(**S, v**)** 执行与 **ReconstructMin(**S, v**)** 类似的功能，但寻找的是最大元素而非最小元素。
- 函数 **Erosion1d(**a**,** b**,** c**,** d**)** 返回作为 $\mathbb{R}$ 子集的一维支撑$[a, b] \setminus [c, d]$ 中的最大侵蚀。

分析。此处我们分析第 6.2 节中的算法，并证明其正确性和运行时间保证。

定理 6.13. 对于任意两个反链可分解的 $[n]^d$-模 $\mathbb{V}$、$\mathbb{W}$，第 *6.2* 节中的算法计算 $d_E(\mathbb{V}, \mathbb{W})$。

证明。 注意，对于支撑为凸的薄模（特别是 ACD 模）， $\mathbb{V}$ 和 $\mathbb{W}$ 的支撑沿纤维 v 的切片要么为空，要么是单个区间。根据前面概述的侵蚀

距离的约化，只需证明对于每个纤维 v， $\mathbb{R}^d_{\mathbb{V}}$ 和 $\mathbb{R}^d_{\mathbb{W}}$ 与纤维 v 的交集可计算：在伪代码中，这表示为 a， b， c， d。

假设我们正在考虑最内层循环中的特定纤维 v。对于一点 $x \in [n]^{d-1}$，令 $(a_x, b_x) :=$ **SupportInterval**$(\mathbb{V}, x)$。若纤维 v 与线段$[(x, a_x), (x, b_x)]$相交，我们就称 x 对纤维 v 是活跃的。

我们声称， $S_{\mathbb{V}}$ 恰好是纤维 v 的活跃点集合 $x \in [n]^{d-1}$。对于任意一点 $x \in [n]^{d-1}$，定义 A_x 为所有满足以下条件的 v 的集合： v 与 $[(x, a_x), (x, b_x)]$ 相交（即所有 x 应该活跃的 v 的集合）。注意，由于 π 的线性性，我们有 $A_x = [\pi(x, a_x)\ \pi(x, b_x)]$。因为固定前 $d-1$ 个坐标（即固定 x）时，π 在最后一个坐标上保序，所以这是一个定义良好的线段，其 $\pi((x, a_x)) \le \pi((x, b_x))$。

现在注意，对于任意固定的 x， A_x 是 $[n]^{d-1}$ 中平行于 $(1, \dots, 1)$ 的线段：因此，根据我们的迭代顺序，一旦点 x 变为活跃点，它就会一直保持活跃，直到扫描线迭代 v 等于 $\pi(x, b_x)$。这一点在我们对 $S_{\mathbb{V}}$ 的更新中得到了正确反映。因此，对于任意纤维 v， $S_{\mathbb{V}}$ 都能正确维护所有活跃点构成的集合。

最后，为了证明 **ReconstructMin** 和 **ReconstructMax** 定义良好，我们需要证明 $S_{\mathbb{V}}$ 中的元素始终可比较。假设 $x, y \in S_{\mathbb{V}}$。那么根据构造，必定存在 t_x, t_y，使得 (x, t_x) 和 (y, t_y) 都位于纤维 v 上。于是 $\pi(x, t_x) = \pi(y, t_y) = v$。展开 π，这要求 x 和 y 之差严格为向量 $(1, \dots, 1)$ 的标量倍数。因此，活跃集 $S_{\mathbb{V}}$ 严格构成一个关于乘积序的全序链，所以取最小值和最大值是定义良好的。

因此， $\mathbb{R}^d_{\mathbb{V}}$ 与 $\mathbb{R}^d_{\mathbb{W}}$ 和纤维 v 的交集计算正确，从而证明了该算法的正确性。

□

定理 6.14。 对于任意两个反链可分解的 $[n]^d$-模 $\mathbb{V}$、$\mathbb{W}$，第 *6.2* 节中的算法可在 $O(n^{d-1} d \log n)$ 时间内运行。

证明。 我们将简要分析初始化和扫描线阶段的运行时间。假设 $\mathbb{V}, \mathbb{W}$ 的表示方式使得 **SupportInterval** 是一个 $O(1)$查询。这可以通过一个$(d-1)$维元组数组实现，使得对于每个 $x \in [n]^{d-1}$，元组表示满足$(x, t) \in \mathbb{R}^d_{\mathbb{V}}$的$t$ 范围。

在这些假设下，**InitProjectedSweepList** 是一个 $O(n^{d-1})$计算。生成的 P_{start} P_{end}映射包含扫描线事件列表：所有键总计有 $O(n^{d-1})$个事件。

扫描线迭代处理了总计 $|\pi([n]^d)| = O(n^{d-1})$ 个纤维。对于每个纤维，如果 S用按字典序排列的平衡二叉搜索树表示，则操作 S.**更新** 可在 $O(d \log |S|)$ 时间内完成（注意，由于 $S_{\mathbb{V}}, S_{\mathbb{W}}$ 中的元素在乘积序下总是两两可比较的，因此字典序给出了一个与乘积序“相容”的全序）。由于 $|S_{\mathbb{V}}|$ 和 $|S_{\mathbb{W}}|$ 均为 $O(n)$，该操作的时间复杂度为 $O(d \log n)$。操作 **ReconstructMin**、**ReconstructMax** 和 **Erosion1d** 均为常数时间，因此扫描线迭代的时间复杂度为 $O(n^{d-1} \cdot d \log n)$。

因此，第 6.2 节中的算法运行时间为 $O(n^{d-1} d \log n)$。 □

定理 6.11 的证明。 结合前两个定理即可证明定理 6.11。事实上，注意到 Dyn 索引模的秩不变量以 $\mathbb{R}^6_\times$ 为索引（它与 $\mathbb{R}^6$ 相同，只是排序方式不同），因此，通过适当修改投影函数 π，第 6.2 节中的算法可以应用于**Dyn**索引的 ACD 模。具体而言，以下方法可行：

$$\pi_{\mathbf{Dyn}}(\mathbf{a}) = (\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_1 + \mathbf{a}_3, \mathbf{a}_1 + \mathbf{a}_4, \mathbf{a}_1 - \mathbf{a}_5, \mathbf{a}_1 - \mathbf{a}_6) \qquad \forall \mathbf{a} \in \mathbb{R}^6_\times.$$

注意，由于秩不变量在非平凡不可容许点处被定义为 0，且 $+\infty$ 在平凡可容许点处（见定义 5.6）亦如此，这些点对侵蚀距离定义中的上确界没有贡献。

注记 6.15. 计算秩不变量之间交错距离（等价于侵蚀距离）的算法见 [11, 定理 5.4]。对于 d维偏序集 $[n]^d$ 上的持久模，秩不变量是 $2d$维的，而计算其侵蚀距离的算法 [11, 定理 5.4]达到期望运行时间 $O(n^{2d} \log n)$。此外，由 [13, 第 4] 节，任何计算此类模之间侵蚀距离的基于搜索的算法，其期望代价至少为$O(n^{2d} \log n)$。

将此与定理 6.11 进行比较：对于同一偏序集 $[n]^d$ 上的凸薄持久模，该定理在最坏情况下可在 $O(n^{d-1} d \log n)$ 时间内计算它们的侵蚀距离。

6.3 将 d_E 算法推广到非薄模

在此，我们将第 6.2 节中的算法扩展到更一般的持久模类。这些模不要求是薄的或反链可分解的，但满足凸性的一种推广概念。

定义 6.16（秩极大模）。固定一个偏序集 $\mathcal{P}$。我们称持久模 $\mathbb{V}:\mathcal{P}\to\mathbf{Vec}$是秩极大的，如果对每个 $p,q\in\mathcal{P}$（满足 $p\le q$），下列条件成立：

$$\operatorname{rank}\mathbb{V}(p\le q)=\min\left(\dim\mathbb{V}(p),\dim\mathbb{V}(q)\right).$$

注意，这意味着每个结构映射要么是单射，要么是满射（也可能两者皆是）。

为引出对秩极大模的研究，我们首先回顾如下内容。

定义 6.17（单纯过滤）。偏序集 $\mathcal{P}$ 上的单纯过滤是一个函子 $K\colon\mathcal{P}\to\mathbf{Simp}$，使得对任意 $p,\ q\in\mathcal{P}$（满足 $p\le q$），有 $K(p)\subset K(q)$。

事实上，秩极大模自然地源自固定点集的聚类信息。

命题 6.18. 固定一个偏序集 $\mathcal{P}$。令 $K\colon\ \mathcal{P}\to$**Simp** 为一个不添加点的单纯过滤：也就是说，对于任意 $p\le q$，K 的零骨架都相等，即 $K_0(p)=K_0(q)$。

定义模 $\mathbb{V}\colon\ \mathcal{P}\to\mathbf{Vec}$，其定义为 $\mathbb{V}:=H_0\circ K$，其中 H_0 是零阶同调函子。那么 $\mathbb{V}$是一个秩极大模。

特别地，这表明，不进行曲率集简化的 $k=0$ 时空持久同调 [11] *without* 会产生秩极大模，因此可以受益于本节提出的算法。

定理 6.19. 假设 $\mathbb{V},\mathbb{W}\colon\ [n]^d\to\mathbf{Vec}$ 是秩极大模。令

$$M=\max\Big(\max_{p\in[n]^d}\dim\mathbb{V}(p),\max_{p\in[n]^d}\dim\mathbb{W}(p)\Big).$$

那么 $d_E(\mathrm{rk}_{\mathbb{V}},\mathrm{rk}_{\mathbb{W}})$ 可以在 $O(M\cdot n^{d-1}d\log\ n)$ 时间内计算。

证明。对于 $i\in\{0,\dots,M\}$，定义超水平集 $\mathrm{supp}^{(i)}\,\mathbb{V}$ 和超水平模 $\mathbb{V}^{(i)}$ 如下：

$$\mathrm{supp}^{(i)}\,\mathbb{V}:=\{p\in[n]^d:\dim\mathbb{V}(p)\ge i\}\qquad\mathbb{V}^{(i)}:=\mathbb{I}_{\mathrm{supp}^{(i)}\,\mathbb{V}}$$

对于 $\mathbb{W}$，也类似。注意，由秩极大性可知，$\mathbb{V}^{(i)}$ 和 $\mathbb{W}^{(i)}$ 现在是纤薄凸模。薄性显而易见；为证明凸性，假设 $p,q,r\in$ 属于 $\mathrm{supp}^{(i)}\,\mathbb{V}$，且 $p\le r\le q$。那么映射$\mathbb{V}(p\le q)$经由 $\mathbb{V}(r)$分解，因此由秩极大性可得

$$i\le\min\left(\dim\mathbb{V}(p),\dim\mathbb{V}(q)\right)=\operatorname{rank}\mathbb{V}(p\le q)\le\operatorname{rank}\mathbb{V}(p\le r)\le\dim\mathbb{V}(r),$$

因此 $r\in$ 属于 $\mathrm{supp}^{(i)}\,\mathbb{V}$。

此外，对于任意 $\varepsilon>0$，将秩不变量 $\mathrm{rk}_{\mathbb{V}}^{+\varepsilon}$ 的侵蚀定义为 $\mathrm{rk}_{\mathbb{V}}^{+\varepsilon}([p,q]):=\mathrm{rk}_{\mathbb{V}}([p^{-\varepsilon},q^{+\varepsilon}])$，其中 $[p,q]$ 为任意线段；对于 $\mathrm{rk}_{\mathbb{W}}^{+\varepsilon}$ 也类似。显然，$d_E(\mathrm{rk}_{\mathbb{V}},\mathrm{rk}_{\mathbb{W}})$是所有满足 $\mathrm{rk}_{\mathbb{V}}^{+\varepsilon}\le\mathrm{rk}_{\mathbb{W}}$ 且 $\mathrm{rk}_{\mathbb{W}}^{+\varepsilon}\le\mathrm{rk}_{\mathbb{V}}$ 的 ε 的下确界（我们将任意此类 ε 称为可行侵蚀）。

现在，对于任意线段 $I=[p,q]$，由秩极大性可得：

$$\mathrm{rk}_{\mathbb{V}}(I)=\min\left(\dim\mathbb{V}(p),\dim\mathbb{V}(q)\right)=\sum_{i=1}^{M}\mathrm{rk}_{\mathbb{V}^{(i)}}(I),$$

其中求和是逐点进行的。因此我们有

$$\mathrm{rk}_{\mathbb{V}}=\sum_{i=1}^{M}\mathrm{rk}_{\mathbb{V}^{(i)}}\qquad\mathrm{rk}_{\mathbb{V}}^{+\varepsilon}=\sum_{i=1}^{M}\mathrm{rk}_{\mathbb{V}^{(i)}}^{+\varepsilon}\qquad\mathrm{rk}_{\mathbb{W}}=\sum_{i=1}^{M}\mathrm{rk}_{\mathbb{W}^{(i)}}\qquad\mathrm{rk}_{\mathbb{W}}^{+\varepsilon}=\sum_{i=1}^{M}\mathrm{rk}_{\mathbb{W}^{(i)}}^{+\varepsilon}.$$

结合这些性质可知，如果对所有i，ε 都是配对 $\mathrm{rk}_{\mathbb{V}^{(i)}}$，$\mathrm{rk}_{\mathbb{W}^{(i)}}$ 的可行侵蚀，那么根据上文所示的水平集可加性，它必定也是配对 $\mathrm{rk}_{\mathbb{V}}$，$\mathrm{rk}_{\mathbb{W}}$ 的可行侵蚀。显然，如果 ε 是 $\mathrm{rk}_{\mathbb{V}}$，$\mathrm{rk}_{\mathbb{W}}$ 的可行侵蚀，那么它对每个单独的 $\mathrm{rk}_{\mathbb{V}^{(i)}}$，$\mathrm{rk}_{\mathbb{W}^{(i)}}$ 也必定是可行的。

因此，我们已经证明

$$d_E(\mathrm{rk}_{\mathbb{V}}, \mathrm{rk}_{\mathbb{W}}) = \max_{i\in\{1,\dots,M\}} \beta_i(\mathbb{V},\mathbb{W}) \qquad \beta_i(\mathbb{V},\mathbb{W}) := d_E(\mathrm{rk}_{\mathbb{V}^{(i)}}, \mathrm{rk}_{\mathbb{W}^{(i)}}).$$

每个 $\beta_i(\mathbb{V},\mathbb{W})$均可由第 6.2 节中的算法在 $O(n^{d-1}d\log\ n)$时间内计算：特别地，请注意我们无需在内存中实际构造 $\mathbb{V}^{(i)}$ 和 $\mathbb{W}^{(i)}$，因为我们只需要 $\mathbb{V},\mathbb{W}$ 的一种表示，使得对所有 i, v，$\mathbf{SupportInterval}(\mathbb{V}^{(i)}, v)$都是 $O(1)$操作。我们需要进行 M次此类计算，因此这一最终复杂度为 $O(M \cdot n^{d-1}d\log\ n)$，即为所需结果。 □

7 实验结果

为验证本方法的可实现性与区分能力，我们分析由Boids模型 [17]生成的动态度量空间。支持本实验的代码仓库位于 [2], ，其中既包括距离计算，也包括类鸟体的固定种子伪随机生成。

该pa 所采用的Boids模型的rticular变体模拟在fla上运动的有限点集 t 2-环面 $\mathbb{T}^2$，以下参数控制每个点的行为：

- 凝聚：类鸟体被邻近类鸟体吸引的倾向。
- 分离：类鸟体避开距离过近的邻居的倾向。
- 对齐：邻近类鸟体匹配速度向量的倾向。

通过调整这些参数，产生了 5 种不同的行为，并针对每种行为选取了 10 个随机生成的群体。每个群体运行 $t = 600$ 个时间步，从而总共得到 50 个动态度量空间。

随后，我们按照定义 4.2 计算了每个群体的持久集不变量。对于每个群体，共选取了 $15,000$ 个 $2k+2$大小的子集，并计算了它们的时空持久同调模：由此得到该 DMS 的近似持久集。我们使用计算所得模上的 d_E 度量运行近似k中心算法（Gonzalez算法），选取了 2,000 个“代表性”样本。

为了给两个群体之间的距离 d_{dyn} 设定下界，我们针对每个 k，改为计算与这两个群体相关联的持久集之间的度量 d_H^{ero}（定义 5.10）。定义 d_H^{ero}中涉及的豪斯多夫距离，需要计算模之间所有成对的侵蚀距离，但通过并行计算可以高效完成。

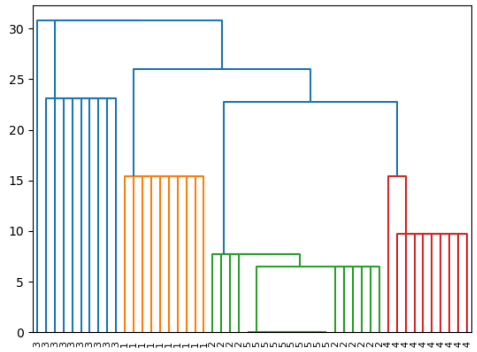
这一过 程针对 $k = 0$和 $k = 1$ 度同调进行，最终得到的距离矩 ix 群体之间 的距离是 $k = 0$ 和 $k = 1$ 距离矩阵逐元素取最大值的结果。

使用单链接层次聚类（图3a）和经典多维缩放（图3b）对该距离矩阵进行了可视化。结果表明，各种行为之间分离效果良好（图3）。

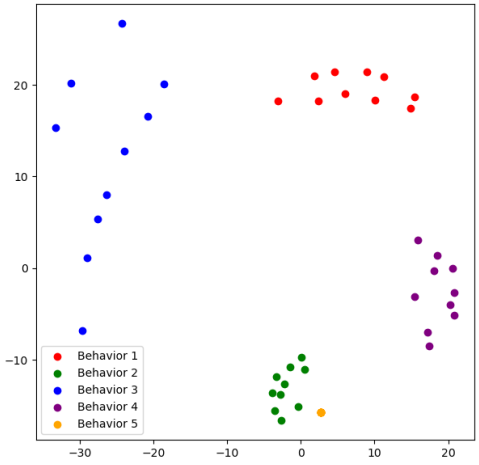
此外，为评估聚类质量，我们计算了1近邻分类器的平均误分类率。对于每种行为，我们随机选择一个群集作为“代表”，并根据由这些代表生成的1近邻分类器对其余所有群集进行分类。该过程重复 10^5 次，并对每次试验的误分类率取平均值，从而得到最终的1近邻误分类率。作为基线，我们将结果与PHoDMSs代码库（[8]）的结果进行比较，该代码库实现了 [11]中描述的算法。相应的1-NN准确率如下：

- $\mathrm{NN}_1(d_E) = 98.57\%$
- $\mathrm{NN}_1(\mathrm{PHoDMSs}) = 72.23\%$

这些实验在一台配备 32 GB 内存的 Intel(R) Xeon(R) Gold 6338 CPU 机器上运行。实验的 d_E 阶段耗时 63 分钟墙钟时间，PHoDMSs 基线阶段耗时 31 小时。



(a) 由 d_E 生成的群体间距离矩阵的单链接树状图。请注意，对于行为 5，任意一对点之间的距离均为 0。



(b) 由 d_E 诱导的群体间距离矩阵的经典多维尺度分析图。请注意，行为 5 的所有 10 个点重合。

图3：由 d_E 诱导的距离生成的群体间距离矩阵的可视化 e

8 稳定性不等式的证明

本节利用结果 [11, 定理 4.1]，确立我们构造的稳定性。对于动态度量空间 $\gamma_X = (X,\ d_X)$ 和子集 $X' \subseteq X$，令 $\gamma_{X'}$ 表示子空间 $\gamma_{X'} = (X',\ d_X|_{X'})$。

引理 8.1。设 $\gamma_X = (X, d_X)$ 和 $\gamma_Y = (Y, d_Y)$ 是两个动态度量空间，并令 $\delta > 0$ 为任意对象。那么对于所有 $X' \subseteq X$，存在一个 $Y' \subseteq Y$，使得 $|Y'| \leq |X'|$，且 $d_{dyn}(\gamma_{X'}, \gamma_{Y'}) \leq d_{dyn}(\gamma_X, \gamma_Y) + \delta$

关于该引理的证明，请参见附录 B。现在我们可以证明：

定理 4.4（动态曲率集的稳定性）。设 $\gamma_X = (X, d_X)$ 和 $\gamma_Y = (Y, d_Y)$ 是两个动态度量空间。那么，对于任意 $k \in \mathbb{Z}_{\geq 0}$ 和 $n \in \mathbb{N}$,

$$d_H\left(D_k^n(\gamma_X), D_k^n(\gamma_Y)\right) \leq 2 \cdot d_{dyn}(\gamma_X, \gamma_Y).$$

其中，d_{dyn} 是从 *[12]*出发的 $\lambda = 2$ 广义*Gromov–Hausdorff*距离；参见定义*A.6*。

证明。为方便记号，定义 $P = D_k^n(\gamma_X)$ 和 $Q = D_k^n(\gamma_Y)$。

任取 $\delta > 0$，并且 $\mathbb{V} \in P$。根据定义 4.2，存在相应的 $S \subseteq X$ 和$\gamma_S \in K_n(\gamma_X)$，使得 $\mathbb{V} = H_k(\mathcal{R}^{\text{lev}}(\gamma_S))$。根据引理 8.1，我们可以选取一个 $T \subseteq Y$和一个 $\gamma_T \in K_n(\gamma_Y)$，使得

$$d_{\text{dyn}}(\gamma_S, \gamma_T) \leq d_{\text{dyn}}(\gamma_X, \gamma_Y) + \delta.$$

定义 $\mathbb{W} := H_k(\mathcal{R}^{\text{lev}}(\gamma_T))$，由定义 4.2 可得 $\mathbb{W} \in Q$。由 [11, 定理 4.1], 我们有不等式 $d_I^{\mathbf{Vec}}(\mathbb{V}, \mathbb{W}) \leq 2 \cdot d_{\text{dyn}}(\gamma_S, \gamma_T)$，因此对任意 $\mathbb{V} \in P$我们找到了相应的$\mathbb{W} \in Q$ 使得

$$d_I^{\mathbf{Vec}}(\mathbb{V}, \mathbb{W}) \leq 2 \cdot d_{\text{dyn}}(\gamma_S, \gamma_T) \leq 2 \cdot d_{\text{dyn}}(\gamma_X, \gamma_Y) + 2\delta.$$

由于 $\mathbb{V}$ 与 $\delta > 0$ 是任意的，由此可得

$$\max_{\mathbb{V} \in P} \min_{\mathbb{W} \in Q} d_I^{\mathbf{Vec}}(\mathbb{V}, \mathbb{W}) \leq 2 d_{\text{dyn}}(\gamma_X, \gamma_Y).$$

通过交换 γ_X 与 γ_Y，我们得到对称的陈述

$$\max_{\mathbb{V} \in Q} \min_{\mathbb{W} \in P} d_I^{\mathbf{Vec}}(\mathbb{V}, \mathbb{W}) \leq 2 d_{\text{dyn}}(\gamma_X, \gamma_Y),$$

这证明了 $d_H(D_k^n(\gamma_X)\ D_k^n(\gamma_Y)) \leq 2 \cdot d_{\text{dyn}}(\gamma_X,\ \gamma_Y)$。 □

9 结论

我们提出了一种用于分析动态数据的构造方法，通过处理原始数据中的许多足够小的子集，应对多维持久性带来的计算挑战。由此得到的区间可分解模可通过高效算法计算度量，例如侵蚀距离；该算法可应用于任意凸薄模，不依赖于我们的构造方法。

这些结果还有若干改进方向。尤其是，关于我们构造的模，一个有趣的问题是能否对这类模给出一种“综合性”定义。我们已经得到了模 $\mathbb{V}$ 是由我们的构造产生的必要条件：它必须可分解为反链。但这并非充分条件：在动态度量空间 (X, d_X) 中，任意两点 $x, y \in X$ 之间的距离函数必须随时间连续，这对这类模的结构提出了额外要求。此外，Rips 过滤的组合性质也会施加进一步约束。在什么条件下，一个由$\mathbf{Dyn}$索引的持久模可以通过对一个 $(2k+2)$-DMS 取 Rips时空持久同调来实现?

此外，这类模在实际处理方面还有提升空间。目前，所有模的支撑都以可能十分稠密的矩阵形式存储。有没有办法高效地对支撑进行“草图化”，即让支撑稀疏化，或用更简单的基本对象来近似它？这将显著提高存储和计算效率。此外，目前计算持久集之间的豪斯多夫距离时，需要计算所有模之间的成对侵蚀距离。更复杂的数据结构或许能大幅减少计算次数，从而提升实际运行性能。

最后，自然会想到，我们的框架能否用于检测和刻画动态度量空间中的基元，即较大 DMS 的许多子 DMS 所呈现的重复模式或行为。具体而言，从基元的视角解读（$2k+2$）持久集，以及高效存储和聚类此类持久集元素，都是未来的研究课题。这与图和网络中的模体发现类似，该概念由 [14]提出。

参考文献

[1] Ulrich Bauer 和 Luis Scoccola。多参数持久模一般不可分解。国际数学研究通讯，2025(5):rnaf034，2025。doi:10.1093/imrn/ rnaf034.[2] Nadezhda Belova、Maxwell Goldberg 和 Andrew Xie。动态曲率集的实现。https://github.com/ndrewxie/dyncurv，2024.[3] H˚avard Bakke Bjerkevik、Magnus Bakke Botnan 和 Michael Kerber。计算交错距离是 NP难。计算数学基础，20，2019。doi:10.1007/ s10208-019-09419-x.[4] Gunnar Carlsson 和 Afra Zomorodian。多维持久性的理论。《离散与计算几何》，2009(5)，2009。doi:10.1007/s00454-009-9176-0.[5] M. J. Catanzaro、S. Rizzo、J. Kopchick、A. Chowdury、D. R. Rosenberg、P. Bubenik 和 V. A. Diwadkar。拓扑数据分析捕捉了个体参与者中由任务驱动的功能磁共振成像特征：一种基于持久性的分类流程。神经信息学，22(1):45–62，2024。doi:10.1007/ s12021-023-09645-3.[6] Samir Chowdhury 和 Facundo M´emoli。网络之间的距离与同构：网络不变量的稳定性与收敛性。《应用与计算拓扑学杂志》，7(2):243– 361，2023.[7] Maria-Veronica Ciocanel、Riley Juenemann、Adriana T. Dawes 和 Scott A. McKinley。运用拓扑数据分析方法揭示丝状网络中环结构出现的时机。数学生物学通报，83(3):21，2021年1月。doi:10.1007/s11538-020-00847-3.[8] Nate Clause 和 Woojin Kim。PHoDMSs：用于动态数据的时空持久同调。https://github.com/ndag/PHoDMSs，2024。

[9]米哈伊尔·格罗莫夫。黎曼与非黎曼空间的度量结构。Birkh¨auser Boston，1999。[10] Mario G´omez 和 Facundo M´emoli。持久图上的曲率集。《离散与计算几何》，72:91–180，2024。doi:10.1007/s00454-024-00634-0。[11] Woojin Kim 和 Facundo M´emoli。动态度量空间的时空持久同调。《离散与计算几何》，66:831–875，2021。doi:10.1007/s00454-019-00168-w。[12] Woojin Kim 和 Facundo M´emoli。通过莫比乌斯反演从动态数据中提取持久簇。arXiv 预印本 arXiv:1712.04064，2022。arXiv:1712.04064。[13] William J. Knight。具有可变探测成本的有序数组搜索。《*SIAM*计算杂志》，17(6):1203–1214，1988。doi:10.1137/0217076。[14] R. Milo、S. Shen-Orr、S. Itzkovitz、N. Kashtan、D. Chklovskii 和 U. Alon。网络模体：复杂网络的简单构件。《科学》，298(5594):824–827，2002。doi:10.1126/science. 298.5594.824。[15] Amit Patel。广义持久图。《应用与计算拓扑学杂志》，1，2018。doi:10.1007/s41468-018-0012-6。[16] Ville P uuska。广义持久模的侵蚀距离。arXiv 预印本 arXiv:1710.01577，2017。arXiv:1710.01577。[17] Craig W. Reynolds。群体、兽群与鱼群：一种分布式行为模型。载于第*14*届计算机图形学与交互技术年度会议论文集，SIGGRAPH '87，第25–34页。计算机协会，1987。doi:10.1145/37401.37406。[18] Tananun Songdechakraiwut 和 Moo K. Chung。面向功能性脑信号的动态拓扑数据分析。载于*2020*年*IEEE*第*17*届生物医学成像国际研讨会专题讨论会（*ISBI*专题讨论会），第1–4页，2020。doi:10.1109/ISBIWorkshops50223.2020.9153431。

A λ-松弛交错距离

我们提出 λ-松弛交错距离，它是格罗莫夫–豪斯多夫距离向动态情境的推广，由 [12]引入。它们是动态度量空间（模强同构）空间上的一族度量，由数 $\lambda \in [0, \infty$参数化）。

定义 A.1（三脚架 [12]）。对于两个集合 X 与 Y，三脚架是一个集合 Z 以及两个满射$\varphi_X\colon\ Z \to X$ 与 $\varphi_Y\colon\ Z \to Y$。我们将三脚架 R 记为 $X \xleftarrow{\varphi_X} Z \xrightarrow{\varphi_Y} Y$

定义 A.2（通过三脚架比较函数 [12]）。设 X 与 Y 为任意两个集合，并带有映射$d_1\colon\ X \times X \to \mathbb{R}_+$ 与 $d_2\colon\ Y \times Y \to \mathbb{R}_+$。给定三脚架 $R\colon\ X \xleftarrow{\varphi_X} Z \xrightarrow{\varphi_Y} Y$，我们称 $d_1 \leq_R d_2$ 当且仅当

$$\varphi_X^* d_1(z, z') \leq \varphi_Y^* d_2(z, z') \qquad \forall z, z' \in Z$$

。这里，$\varphi_X^* d_1(z,\ z')$ 表示 d_1 的拉回，即 $\varphi_X^* d_1(z,\ z') = d_1(\varphi_X(z)\ \varphi_X(z'))$

对于任意 $t \in \mathbb{R}$、$\varepsilon > 0$，我们令 $[t]^{+\varepsilon} := [t - \varepsilon, t + \varepsilon]$。此外，我们将定义

定义 A.3 ([12])。设 X 是一个集合，且 $d_X\colon\ X \times X \to \mathbb{R}_+$。那么对于任意 $t \in \mathbb{R}$ 和 $\varepsilon > 0$，定义 $\bigvee_{[t]^{+\varepsilon}} d_X$ 为

$$\left(\bigvee_{[t]^{+\varepsilon}} d_X\right)(x, x') := \min_{t' \in [t]^{+\varepsilon}} d_X(t')(x, x') \qquad \forall x, x' \in X.$$

借助这个定义，我们现在可以引入 λ-松弛畸变，它用于构造 λ-松弛交错距离。

定义 A.4（三脚架 [12] 的 λ-畸变）。设 $\gamma_X = (X, d_X)$ 和 $\gamma_Y = (Y, d_Y)$ 是两个动态度量空间。设 $R\colon\ X \xleftarrow{\varphi_X} Z \xrightarrow{\varphi_Y} Y$ 是一个三脚架。我们称 R 为一个（λ, ε）-三脚架，当且仅当

$$\text{For all } t\in\mathbb{R},\ \bigvee_{[t]^{+\varepsilon}} d_X \le_R d_Y(t)+\lambda\cdot\varepsilon \text{ and } \bigvee_{[t]^{+\varepsilon}} d_Y \le_R d_X(t)+\lambda\cdot\varepsilon$$

λ-畸变 $\mathrm{dis}_\lambda^{\mathrm{dyn}}(R)$ 是所有满足 R 是一个（λ,ε）-三脚架的 $\varepsilon>0$ 的下确界。

定义A.5（动态度量空间 [12]之间的 λ-松弛交错距离）。对于任意 $\lambda\ge 0$ 和动态度量空间 γ_X，γ_Y，定义 λ-松弛交错距离如下：

$$d_{\mathrm{dyn},\lambda}(\gamma_X,\gamma_Y)=\min_R \mathrm{dis}_\lambda^{\mathrm{dyn}}(R)$$

其中，我们对所有三脚架 R： $X \xleftarrow{\varphi_X} Z \xrightarrow{\varphi_Y} Y$ 取最小值。

为简洁起见，我们作如下定义：

定义A.6（动态度量空间 [12]之间的 d_{dyn}）。我们将 $\lambda=2$ 松弛交错距离 $d_{\mathrm{dyn},\,2}$称为 d_{dyn}。

B 引理 8.1 的证明

我们将证明：

引理 8.1. 设 $\gamma_X=(X,d_X)$ 和 $\gamma_Y=(Y,d_Y)$ 是两个动态度量空间，并设 $\delta>0$为任意元素。那么对所有 $X'\subseteq X$，存在一个 $Y'\subseteq Y$使得 $|Y'|\le|X'|$，且 $d_{dyn}(\gamma_{X'},\gamma_{Y'})\le d_{dyn}(\gamma_X,\gamma_Y)+\delta$

证明。为方便记号，令 $\Delta=d_{\mathrm{dyn}}(\gamma_X,\gamma_Y)$。

根据定义 A.5，存在一个动态度量空间 (Z,d_Z) 和一个满射三脚架 R，由 $X\xleftarrow{\varphi_X}Z\xrightarrow{\varphi_Y}Y$ 给出，使得 $\mathrm{dis}^{\mathrm{dyn}}(R)<\Delta+\delta$。

由 R 的满射性，对每个 $x\in X'$，存在相应的 z_x 使得 $\varphi_X(z_x)=x$。

令 $Z'=\{z_x:\ x\in X'\}$，并定义 $Y'=\varphi_Y(Z')$。

由此得到一个满射三脚架 R'，记作 $X'\xleftarrow{\varphi_X|_{Z'}}Z'\xrightarrow{\varphi_Y|_{Z'}}Y'$。考虑动态度量空间 $\gamma_{X'},\gamma_{Y'},\gamma_{Z'}$。我们声称 $\mathrm{dis}^{\mathrm{dyn}}(R)<\Delta+\delta$。反设不然：那么根据定义 A.5，$R'$ 对于任意 $\varepsilon<\Delta+\delta$ 都不是一个 ε-三脚架。因此，必定存在某个 $t\in\mathbb{R}$ 和一对 $z,z'\in Z'$，使得以下情况之一成立：

$$\begin{aligned}\left(\varphi_{X'}^*\,d_{X'}([t]^{+\varepsilon})\right)(z,z')&>\left(\varphi_{Y'}^*\,d_{Y'}([t]^{+\varepsilon})\right)(z,z')+2\varepsilon,\ \text{or}\\ \left(\varphi_{Y'}^*\,d_{Y'}([t]^{+\varepsilon})\right)(z,z')&>\left(\varphi_{X'}^*\,d_{X'}([t]^{+\varepsilon})\right)(z,z')+2\varepsilon.\end{aligned}$$

但由于 $Z'\subseteq Z$，相同的 t 和 $z,z'\in Z$ 表明 R 不是一个 ε-三脚架。但这与 $\mathrm{dis}^{\mathrm{dyn}}(R)<\Delta+\delta$ 的事实矛盾。

因此，我们有一个三脚架 R'，使得 $\mathrm{dis}^{\mathrm{dyn}}(R')<\Delta+\delta$，从而我们选取的 Y' 满足$d_{\mathrm{dyn}}(\gamma_{X'},\gamma_{Y'})\le d_{\mathrm{dyn}}(\gamma_X,\gamma_Y)+\delta$。 □

C 来自 [10]的必要结果陈述

尽管这些结果最初是针对度量空间陈述的，但在半度量情形下同样成立。事实上，曲率集甚至可以在更一般的网络情形下定义，见 [6]。

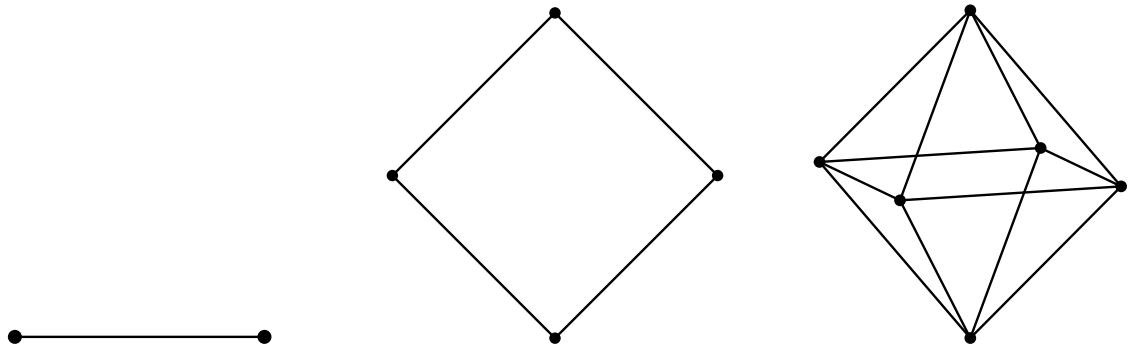
定理 C.1 ([10, 定理 4.4(B)]). 设 (X,d_X) 是一个含有 n 个点的度量空间。假设 n 为偶数，且 $k\in\mathbb{Z}_+$ 满足 $n=2k+2$。令 $\mathrm{dgm}_k^{\mathrm{VR}}(X)$ 表示 X 的 *Rips* 过滤的 k 次持久图。

那么，$\mathrm{dgm}_k^{\mathrm{VR}}(X)$ 当且仅当 $t_b(X)<t_d(X)$ 时由单个点 $(t_b(X)\ t_d(X))$ 构成，否则为空；其中 $t_b(X)\ t_d(X)$ 是定义于定义 *4.1]* 中的函数。

事实上，[10]给出了更强的结论：

命题 C.2 ([10, 命题 4.3]). 设 (X,d_X) 是一个含有 n 个点的度量空间。假设 n 为偶数，且 $k\in\mathbb{Z}_+$ 满足 $n=2k+2$。

如果存在一个 $r\in\mathbb{R}_+$，使得 维数 $H_k(\mathcal{R}_r(d_X))=1$，那么作为单纯复形，$\mathcal{R}_r(d_X)$ 必定与交叉多胞体 $\mathfrak{B}_{k+1}$ 同构（见图 *4*）。



(a) 交叉多胞体 $\mathfrak{B}_1$　(b) 交叉多胞体 $\mathfrak{B}_2$　(c) 交叉多胞体 $\mathfrak{B}_3$

图 4：一些交叉多胞体