arXiv:2602.03163v1 [math.AT] 3 Feb 2026

持久相对同调的 U-MATCH 算法

Christian Lentz塔夫茨大学，马萨诸塞州梅德福
christian.lentz@tufts.edu

Gregory Henselman-Petrusek太平洋西北国家实验室，美国
gregory.roek@pnnl.gov

Lori Ziegelmeier麦卡莱斯特学院，明尼苏达州圣保罗
lziegel1@macalester.edu

摘要

数据驱动的科学探究中的一个核心问题，是如何解读含噪高维数据中的结构。拓扑数据分析（TDA）通过持久同调提供了一种解决方案，它将感兴趣的特征编码为数据滤过中的拓扑孔洞。本研究将这一框架扩展到一种相关不变量，该不变量揭示空间相对于子空间的拓扑结构：持久相对同调（PRH）。我们展示了如何利用两步矩阵约简技术，以简单、透明且通用的方式计算这一不变量，其最坏情况下的时间复杂度与普通持久同调相当。除提供证明以说明该方法的正确性和计算复杂度外，我们还针对一种特殊情形实现了性能优化。

关**键词** 持久同调、相对同调、拓扑数据分析

1 引言

持久同调刻画了数据的拓扑结构，以及该结构随着一个或多个参数值 [6, 17, 15, 5, 14] 的选择而发生的变化。持久同调理论与计算方面的诸多突破，为拓扑数据分析 （TDA）这一领域铺平了道路。目前，拓扑数据分析是一个不断发展的领域，在机器学习 [23], 、网络科学 [2], 和机器人学 [3] 等领域的应用日益广泛，也展现出巨大的跨学科潜力。对数据进行拓扑洞察的需求日益增长，加之持久同调范式的成功，促使人们进一步探索拓扑不变量，将其作为从科学数据中提取洞见的工具。

然而，拓扑不变量的计算是一项重大瓶颈。事实上，拓扑数据分析领域面临的一项重大挑战，是确定在科学应用场景中哪些拓扑描述符能够被高效计算，以及如何在代码中可靠地实现这些计算。本研究旨在针对一种可用于科学数据的最基本拓扑描述符之一——持久相对同调——来应对这一挑战。具体而言，我们希望更深入地了解通用算法的能力与局限；这类算法几乎可应用于任何数据集，对数据格式和结构的要求极低。

问题陈述。 考虑一对滤过拓扑空间 $F_0X \subseteq \cdots \subseteq F_NX$和 $G_0Y \subseteq \cdots \subseteq G_NY$，满足对于任意 t, $G_tY \subseteq F_tX$。固定系数域 K，我们的目标是计算一个持久模的条形码

$$H_*(F_0X/G_0Y) \to \cdots \to H_*(F_NX/G_NY)$$

并为条形码中的每个条计算一个（相对）循环代表元。

贡献。 考虑计算绝对持久同调模$H_*(F_0K) \to \cdots \to H_*(F_NK)$条形码的经典方法。我们构造微分矩阵或边界矩阵 D，并计算所谓的 $R = DV$分解 [9]。随后可以从 R的稀疏模式中读出绝对条形码，并可由 R 和 V 的列得到相应的循环代表元。

我们提供了一个理论框架，说明类似的方法可用于计算 PRH 模块的条形码。事实上，我们只需进行一次基变换。我们首先计算 U-match 分解$\mathcal{T}M = \mathcal{D}\mathcal{S}$，其中 $\mathcal{D}$ 是边界矩阵 D [18]的一个排列。随后，我们对 $\mathcal{T}$和 $\mathcal{S}$ 的列进行排列，分别得到矩阵 $\mathcal{A}$ 和 $\mathcal{B}$，并得到第二个 U-match 分解 $\tilde{\mathcal{T}}\tilde{M} = (\mathcal{A}^{-1}\mathcal{B})\tilde{\mathcal{S}}$。所需的条形码

˜可由 M的稀疏模式得到，相应的相对循环代表元可由矩阵 $\mathcal{AT}$的列˜得到。该方法适用于任意系数域上的任意过滤胞腔复形，并允许任意滤过 $F_\bullet$ 和 $G_\bullet$。

U-match 丰富的结构使我们的工作直截了当、易于理解，并确保正确性证明只需用到线性代数和代数拓扑的基本定义。事实上，U-match 的性质还将为这种方法带来两项关键优势：应用上更加灵活，以及能够直观地解释相对同调类的生成元和边界链。我们在第7节中介绍一个经过性能优化且易于解释的实现案例，以作说明。

据我们所知，本文提出的算法独具特色，因为它结合了两项特性：(A)$F_\bullet$ 和 $G_\bullet$ 可以是有限胞腔复形的任意滤过；(B)该过程可通过两次矩阵分解完成，不仅揭示条形码，还能给出持久相对循环代表元。先前的工作已经提供了适用于任意滤过的算法，但需要进行五次矩阵分解；详情请参阅先前及相关工作。

先前及相关工作。 在 [22], 中，Morozov 提出了一种算法，用于构造与像、核和余核的嵌套序列相对应的持久图。该方法的计算复杂度关于单形数量为三次，基于 [9], 的标准矩阵约化，并要求：(1) 单纯复形 $K_0 \subseteq K$；(2) 一个单射$f\colon K \to \mathbb{R}$，其次水平集是 K的子复形；以及(3) 一个单射 g，它是 f 到K_0的限制。该算法通过一系列五次矩阵分解来执行，并可借助 [22]的映射柱引理 扩展到过滤空间对的核、像和余核的嵌套序列。 [8]中可以找到对这项工作的精简介绍。

受计算本质同调的启发， [7]中的作者引入了扩展持久性，利用*Poincar´e*和*Lefschetz*对偶性，设计了一种矩阵约简算法，通过构造所谓的锥复形来计算相对同调。在 [12], 中，作者引入了逐点和全局对偶性。他们证明，在计算持久相对同调（PRH）时， [7] 的锥复形并非必要；并且四种标准的持久（上）同调模中的任意一种都可由其余各项推得，因此至多通过两步矩阵约简即可完成计算。最近， [4] 中提出了相对 Delaunay-ˆCech 复形 $Del\,\check{C}(X, A)$，以便在低维欧几里得空间中高效计算相对 ˆCech 持久同调。具体而言，给定 $A \subseteq X \subseteq \mathbb{R}^d$，作者定义了$Del\,\check{C}(X, A)$，并证明其底层过滤单纯复形的持久同调与过滤 ˆCech 复形ˆ$C_\bullet(X)/\check{C}_\bullet(A)$的同调同构。

局部同调 研究拓扑空间 X 中的单个点 $x \in X$ 这一特殊情形 $H_p(X,\ X \setminus x)$。Morozov 在 [22] 的工作也描述了这个问题。最近，Kerber 和 S¨ols 提出了局部双过滤 [21] ，其中数据按尺度进行过滤，第二个过滤参数则描述尺度过滤相对于单个点的局部性。在 [24], 中，作者描述了抽象单纯复形的局部同调，并将其应用于图和超图的分析。在 [25], 中，作者提出了一个理论框架，用于识别网络中的关键节点或瓶颈，以局部同调方法取代计算开销高昂的图算法。

总体而言，持久相对同调算法可分为三类：(A) 如 [8], 中所示的像与核持久性；这类方法虽不能与相对同调完全对应，但包含相似的概念；(B) 形式为 $H_*(X,\ F_1X) \to \cdots \to H_*(X,\ F_NX)$ 的模，见于 [7, 12], ；以及 (C) 形式为 $H_*(F_0X,\ G_0Y) \to$ 的模

$\cdots \to H_*(F_NX,\ G_NY)$，其中 $G_tY = (F_tX) \cap Y$。 [8] 的作者也考虑了情形 (C)，据我们所知，这是唯一一项还考虑了任意形式为 $H_*(F_0X,\ G_0Y) \to \cdots \to H_*(F_NX,\ G_NY)$，且不施加 (B) 或 (C) 类限制的模的工作。这是借助映射柱引理实现的，该方法通过依次应用至多五次矩阵分解来完成。本工作的主要贡献是一种同样具有一般性的算法，至多使用两次矩阵分解。尽管 [12] 也实现了至多两步的矩阵分解，但这属于情形 (B)。此外，计算持久相对循环代表元也是本工作的重点，而现有方法和文献中并未涉及这一点。

2 预备知识

如需了解更多背景知识，相关的代数拓扑定义请参阅 [14, 19]，单纯同调入门请参阅 [10] ，持久同调综述请参阅 [5, 6]，计算技术入门请参阅 [13] 。

链复形与同调。 持久同调通过对原始数据进行一系列细化来计算。首先，将数据编码为如下形式的链复形

$$\mathcal{C}_\bullet := \ldots\ C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \ \ldots\ \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0,$$

或者满足 $\partial_p\partial_{p+1}=0$ 的一系列线性映射和向量空间。我们将每个 C_p 称为链向量空间，并将$C=\bigoplus_p C_p$ 记作所有链构成的空间。我们将这些同态称为边界算子。[1] 链复形可以用一个线性变换简洁地描述，其矩阵为

$$D=\begin{pmatrix} 0 & \partial_1 & & & \\ & 0 & \partial_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & \partial_n \\ & & & & 0 \end{pmatrix},$$

或分块边界矩阵。很容易证明，命题 $\partial_p\partial_{p+1}=0$ 等价于命题 $\mathrm{im}(\partial_{p+1})\subseteq\ker(\partial_p)$，其中，我们将向量空间 $\mathrm{im}(\partial_{p+1})$ 的任意元素称为 p-边界，将向量空间 $\ker(\partial_p)$ 的任意元素称为 p-圈。由此，我们将链复形 $\mathcal{C}_\bullet$ 的 p^{th} 同调向量空间定义为商空间 $H_p(C_\bullet)=\ker(\partial_p)/\mathrm{im}(\partial_{p+1})$，其中，非平凡等价类表示一组不是 p-边界的 p-圈。等价类 $[\xi]=\xi+\mathrm{im}(\partial_{p+1})$ 的元素 ξ 称为该类的循环代表元。对于聚合了所有维度的全部圈和边界空间，我们分别记作 $Z:=\ker(D)\cong\oplus_p\ker(\partial_p)$ 和 $B:=\mathrm{im}(D)\cong\oplus_p\mathrm{im}(\partial_p)$。同样地，对于所有同调群组成的总体空间，我们记作 $H_*(C_\bullet):=Z/B\cong\bigoplus_p H_p(C_\bullet)$。

过滤空间。我们的讨论将聚焦于由滤过产生的链复形，即拓扑胞腔复形的嵌套序列 $F_\bullet X:=F_1X\subseteq F_2X\subseteq\cdots\subseteq F_NX$ 其中 $F_NX=X$。在实践中，X 几乎总是单纯复形或立方复形。对于每个子复形 F_tX，我们关联一个链复形 $\mathcal{C}_\bullet(F_tX)$。为方便起见，并且在上下文不会引起混淆的情况下，我们将 F_tX 与 $C(F_tX)$交替使用，后者表示 $C_\bullet(F_tX)$中所有链构成的空间。胞腔 $\sigma\in X$ 的出生过滤值 定义为使得 $\sigma\in F_tX$的最小 t；我们将该值记为 $b_F(\sigma)=\min\{t\colon\sigma\in F_tX\}$.[2] 给定一个域 K 和一个 p-链 $\alpha=\sum a_i\sigma_k$，满足 $a_i\in\mathrm{K}\setminus\{0\}$ 和$\sigma_k\in C_p$，我们将 α 的出生过滤值 定义为 $b_F(\alpha):=\max\{b_F(\sigma_k)\}$。

持久性。设 $\mathcal{C}_\bullet(X)$ 是滤过拓扑空间 $F_\bullet X$ 上的链复形。每个 $C_p(X)$ 都由胞腔复形中 p 维单元的标准基生成，并且每个线性变换 ∂_{p+1} 将一个（$p+1$）维单元映射为若干个 p 维单元的线性组合，或者映射为一个 p-链，即其边界。[3] 图 1 给出了一个由一系列链复形及其包含映射构成的示例，这些包含映射称为链映射，并作用于相应复形的链向量空间。

$$\begin{array}{ccccccccc} 0 & \xrightarrow{\partial_{n+1}} & C_n(F_aX) & \xrightarrow{\partial_n} & \cdots & \xrightarrow{\partial_1} & C_0(F_aX) & \xrightarrow{\partial_0} & 0 \\ & & \downarrow{\scriptstyle f_n} & & & & \downarrow{\scriptstyle f_0} & & \\ 0 & \xrightarrow[\partial_{n+1}]{} & C_n(F_bX) & \xrightarrow[\partial_n]{} & \cdots & \xrightarrow[\partial_1]{} & C_0(F_bX) & \xrightarrow[\partial_0]{} & 0 \end{array}$$

图 1：滤过拓扑空间 $F_\bullet X$ 上滤过链复形的两个层级，其链映射的形式为 $f_p\colon C_p(F_aX)\hookrightarrow C_p(F_bX)$，其中 $a\le b$。

这些链映射在同调上诱导出线性变换 $f_p^*\colon H_p(F_aX)\to H_p(F_bX)$，其中$a\le b$，并可将其排列成序列 $H_*(F_\bullet X):=H_*(F_1X)\to\cdots\to H_*(F_NX)$，称为（分次）持久同调模。Gabriel 的一个结果 [16] 表明，每个有限胞腔复形上的持久同调模都可作如下直和分解：$H_*(F_\bullet X)\cong\oplus_\xi\mathrm{I}\langle b_\xi,\ d_\xi\rangle$，其中每个 $\langle b_\xi,\ d_\xi\rangle$ 都是所谓的区间模，其形式为 $0\to\cdots\to0\to\mathrm{K}\xrightarrow{id}\cdots\xrightarrow{id}\mathrm{K}\to0\to\cdots\to0$，描述某个循环代表元 $\xi\in H_*(F_\bullet X)$ 的诞生时间 b_ξ 和消亡时间 d_ξ。区间模的多重集给出了持久模$H_*(F_\bullet X)$的条形码 。

3 持久相对同调

现在，我们将上一节的定义和记号扩展到持久相对同调（PRH）。关于商拓扑的更多内容，请参阅 [11]；关于相对同调的更多内容，请参阅 [19] 。我们考虑一对拓扑空间的滤过对 (X,Y)

[1]边界算子一词常与微分算子互换使用。[2]在实践中，序列 $F_1X\subseteq F_2X\subseteq\cdots\subseteq F_NX$通常对应于实数序列 $t_1\le\cdots\le t_N$，这些实数通常称为过滤值。[3]链 $\alpha\in C_p$ 的边界是 $\partial_p(\alpha)$，其中微分映射 ∂_p 针对 C_p 中的胞腔定义，并通过线性扩展作用于 C_p中的所有链。

它们分别配备了滤过 $F_\bullet$ 和 $G_\bullet$。我们将这对空间及其滤过记作 $(F_\bullet X, G_\bullet Y)$，并要求它满足：(1) $F_N X = X$，(2) $G_N Y = X$，以及对于每个 $t \in \{1, \ldots, N\}$ 和 $N \in \mathbb{N}$，(3) $G_t Y \subseteq F_t X$。注意，条件 (2) 是为了确保滤过终止。

商链复形。持久相对同调可以视为一种特定类型滤过链复形的同调，该滤过链复形源自一系列相对链向量空间，描述了商空间中所有相对链的张成空间 X/Y。在维数 p 中，该链向量空间是商 $C_p(X, Y) = C_p(X)/C_p(Y)$。由于 $\partial_p(C_p(Y)) \subseteq C_{p-1}(Y)$，因此无需定义所谓的相对边界算子。因此，我们使用微分 ∂_p 以及形如 $C_p(X, Y)$ 的一系列链向量空间，在 $(F_\bullet X, G_\bullet Y)$ 上定义一个商链复形 $\mathcal{C}_\bullet(X/Y)$。

相对同调。在绝对同调中，我们将任意 $\alpha \in C_p(X)$ 解释为 p-循环，当且仅当 $\partial_p(\alpha)$ 是平凡的。相对 p-循环的定义是将 $C_{p-1}(Y)$ 的任意链视为平凡。同样，在相对同调的语境中，边界的概念也有所修改。

定义 *1.* 相对 p-循环是 $\partial_p^{-1}(C_{p-1}(Y)) = \{\alpha \in C_p(X)$ 中的一个元素：$\partial_p(\alpha) \in C_{p-1}(Y)\}$。对于所有$p$，我们将相对循环的向量空间记为 $Z(X,Y) = C(X) \cap \partial^{-1}(Y)$。类似地，相对 p-边界是任意链 $\alpha = \partial_{p+1}(\beta) + \gamma$，其中某个 $\beta \in C_{p+1}(X)$ 和 $\gamma \in C_p(Y)$，而相对边界的向量空间为$B(X,Y) = B(X) + C(Y)$。

由于 $\vec{0} \in C_{p-1}(Y)$显然成立，因此任何绝对 p-循环必定是相对 p-循环。另一方面，如果相对 p-循环 α 与某个绝对 p-边界 $\partial_{p+1}(\beta)$ 相差一个（可能为零的） p-链 $\gamma \in C_p(K_0)$，那么 α 也必定是相对 p-边界。由于任何绝对 p-边界都可以写作 $\alpha = \partial_{p+1}(\beta) + \vec{0}$，因此任何绝对 p-边界也必定是相对 p-边界。(X, Y) 的相对同调是商向量空间 $H_*(X, Y) = Z(X, Y)/B(X, Y)$。

定义 2。链 $\xi \in C_p$（X）的一个相对边界链 是指满足 $\xi - \partial_{p+1}$（β） $\in C_p$（Y）的链 $\beta \in C_{p+1}$（X）。换言之，我们称 ξ 是相对于 Y的 β 的边界。

下面的引理 [19] 概述了相对同调的结构。

引理 3。如果空间对 （X、Y）和（K、L）满足 $X \subseteq K$和 $Y \subseteq L$，则链映射 g_p： C_p（X、Y） $\to C_p$（K、L）在相对同调上诱导出同态 g_p^*： H_p（X、Y） $\to H_p$（K、L）。

持久性。引理 3 表明，包含映射 $G_t Y \subseteq G_{t+1} Y$、$F_t X \subseteq F_{t+1} X$ 和 $G_t Y \subseteq F_t X$ 在相对同调上产生了形式为 H_p（F_1， G_1） $\to H_p$（F_2， G_2） $\to \cdots \to H_p$（F_N， G_N）的诱导映射。这就是（维数 p）持久相对同调模。对所有 p而言，这产生了一个分次持久相对同调模 H_*（$F_\bullet X$， $G_\bullet Y$），根据 [16]，它具有唯一的条形码分解。

4 U-match 分解

U-match 分解是矩阵分解的一种形式，最初用于从结构和算法两个层面研究持久同调及其变体（持久上同调、扩展持久性等） [18]。每个矩阵都存在 U-match 分解，并且可以在立方时间内计算得到 [18]。

定义 4。对于一个域 $\mathbb{K}$，令 $\mathbb{K}^{n \times m}$ 表示所有系数来自$\mathbb{K}$的 $n \times m$ 矩阵的集合。矩阵 $D \in \mathbb{K}^{n \times m}$ 的一个 ***U-match*** 分解 由矩阵元组$(\mathcal{T}, M, D, \mathcal{S})$ 给出，并满足以下三个条件：

- $\mathcal{T} M = D \mathcal{S}$。
- $\mathcal{T} \in \mathbb{K}^{n \times n}$ 和 $\mathcal{S} \in \mathbb{K}^{m \times m}$ 都是上三角矩阵，且对角线元素均为 *1*。
- $M \in \mathbb{K}^{n \times m}$ 是一个广义匹配矩阵。具体而言，这意味着 M 每行至多有一个非零元素，每列至多有一个非零元素。

这种分解在多种情境中都很有用，因为矩阵 $\mathcal{S}$ 和 $\mathcal{T}$包含了有关 $\mathcal{D}$ 的若干行子矩阵和列子矩阵的核与像的丰富信息，而这些信息与持久性计算相关。U-match 与众所周知的 $R = DV$ 分解密切相关，所谓的适当的 U-match 分解与 适当的 $R = DV$分解之间存在一一对应关系；详情请见 [18] 。

5 通过 U-match 计算持久相对同调

这里介绍主要结果：一种通过矩阵分解计算持久相对同调条形码分解的过程。我们还证明了其正确性，并给出了立方复杂度界。

5.1 主算法

算法 1 给出了主要算法的伪代码。为妥善处理某些边界情况，该过程不失一般性地假设子空间 Y 等于 X；也就是说，$G_N = X = F_N$。如果不满足这一条件，则始终可以通过 $G_{N+1}X = X = F_{N+1}X$ 扩展 $F_\bullet$ 和 $G_\bullet$。

从整体上看，该算法包含几个关键步骤。首先，通过以与滤过 $G_\bullet$ 相容的方式（从上到下取值递增）对边界矩阵 D 的行进行置换，并以与滤过 $F_\bullet$ 相容的方式（从左到右取值递增）对列进行置换，构造矩阵 $\mathcal{D}$。然后，我们计算一个 U-match 分解 $\mathcal{T}M = \mathcal{D}\mathcal{S}$。正如稍后将看到的，$\mathcal{T}$ 的列和 $\mathcal{S}$ 的列分别包含相对边界的一组基和相对循环的一组基。接着，我们计算第二个 U-match 分解，得到一个矩阵，其列包含相对循环空间和相对边界空间的公共基。条形码和相对循环代表元都可以从这个第二个 U-match 中的矩阵确定。

伪代码中使用了两个特殊函数。相对循环次水平集函数是在 $C(X)$ 上定义的唯一整数值函数 f_{ker}，满足对于所有 t，$Z(F_tX, G_tX) = \{c \in C(X) : f_{ker}(c) \leq t\}$。也就是说，$c \in Z(F_tX, G_tX) \iff f_{ker}(c) \leq t$。直观地说，我们将 $f_{ker}(c)$ 视为 c 成为相对循环的过滤值，即 $[c]$ 作为相对同调类诞生时的过滤值。类似地，相对边界次水平集函数是在 $C(X)$ 上定义的唯一整数值函数 f_{im}，满足对于所有 t，$B(F_tX, G_tX) = \{c \in C(X) : f_{im}(c) \leq t\}$。也就是说，$c \in B(F_tX, G_tX) \iff f_{im}(c) \leq t$。直观地说，我们将 $f_{im}(c)$ 视为 c 成为相对边界的过滤值，即 $[c]$ 作为相对同调类消亡时的过滤值。

算法 1 持久相对同调

要求： 一个有限胞腔复形 X，配备滤过 $F_\bullet$：$F_1X \subseteq \cdots \subseteq F_NX$ 和 $G_\bullet$：$G_1X \subseteq \cdots G_NX$。要求对所有 k，均有 $G_NX = X = F_NX$ 和 $G_kX \subseteq F_kX$。**确保：** 持久相对同调模 $H_*(F_1X, G_1X) \to$ 的区间分解 $\cdots \to H_*(F_NX, G_NX)$。该分解编码为集合 I，其中每个条形码中的半开区间 $[birth, death)$ 对应一个三元组 $(birth, death, representative)$，其中 $representative$ 是相应的相对循环代表元。1: $\sigma_1, \ldots,$ $\sigma_m \leftarrow$ 是 X 中胞腔的一个排列，使得 $b_F(\sigma_1) \leq \cdots \leq b_F(\sigma_m)$ 2: $\tau_1, \ldots, \tau_m \leftarrow$ 是 X 中胞腔的一个排列，使得 $b_G(\tau_1) \leq \cdots \leq b_G(\tau_k)$ 3: $\mathcal{D} \leftarrow$ 是 $\mathcal{C}_\bullet(X)$ 的微分矩阵，行和列按顺序排列，使得 $\partial(\sigma_j) = \sum_i \mathcal{D}_{ij}\tau_i$。4: 定义 $\mathcal{D}_{\tau_i,\sigma_j} = \mathcal{D}_{ij}$（对所有 i, j；仅用于下游分析）。5: $(\mathcal{T}, M, \mathcal{D}, \mathcal{S}) \leftarrow$ 是 $\mathcal{D}$ 的 U-match 分解。6: 定义 $\mathcal{T}_{\tau_i\tau_j} = \mathcal{T}_{ij}$，以及 $\mathcal{S}_{\sigma_i\sigma_j} = \mathcal{S}_{ij}$（对所有 i, j；仅用于下游分析）。7: $\beta_1, \ldots, \beta_m \leftarrow$ 是 X 中胞腔的排列，使得 $f_{im}(COL_{\beta_1}(\mathcal{T})) \leq \cdots \leq f_{im}(COL_{\beta_m}(\mathcal{T}))$。8: $\gamma_1, \ldots, \gamma_m \leftarrow$ 是 X 中胞腔的排列，使得 $f_{ker}(COL_{\gamma_1}(\mathcal{S})) \leq \cdots \leq f_{ker}(COL_{\gamma_m}(\mathcal{S}))$。9: $\mathcal{A} \leftarrow$ 是 $\mathcal{T}$ 的排列，使得 $\mathcal{A}_{ij} = \mathcal{T}_{\tau_i,\beta_j}$。10: $\mathcal{B} \leftarrow$ 是 $\mathcal{S}$ 的排列，使得 $\mathcal{B}_{ij} = \mathcal{S}_{\tau_i,\gamma_j}$。11: 定义 $\mathcal{A}_{\tau_i,\beta_j} = \mathcal{T}_{\tau_i,\beta_j}$ 以及 $\mathcal{B}_{\tau_i,\gamma_j} = \mathcal{S}_{\tau_i,\gamma_j}$（对所有 i, j；仅用于下游分析）。12: $(\tilde{\mathcal{T}}, \tilde{M}, \mathcal{A}^{-1}\mathcal{B}, \tilde{\mathcal{S}}) \leftarrow$ 是 $\mathcal{A}^{-1}\mathcal{B}$ 的 U-match 分解。13: 定义 $\tilde{\mathcal{T}}_{\beta_i\beta_j} = \mathcal{T}_{ij}$，以及 $\tilde{\mathcal{S}}_{\gamma_i\gamma_j} = \mathcal{S}_{ij}$（对所有 i, j；仅用于下游分析）。14: $I \leftarrow \emptyset$ 15: 对矩阵 $\tilde{M}$ 的每个非零元素 M_{ij} 执行 16: $representative \leftarrow \sum_k (\mathcal{A}\mathcal{T})_{ki}\gamma_k$ 17: $birth \leftarrow f_{ker}(representative)$ 18: $death \leftarrow f_{im}(representative)$ 19: 如果 $birth \neq death$，则 20: 将 $(birth, death, representative)$ 插入 I 21: 结束如果 22: 结束循环 23: 返回 I

该算法会生成若干不同的对象，每个对象都可以用整数或胞腔进行索引。（详见算法 1）：

$$\mathcal{A}_{\tau_i\beta_j} = \mathcal{A}_{ij} \quad \mathcal{B}_{\tau_i\gamma_j} = \mathcal{B}_{ij} \quad M_{\tau_i\sigma_j} = M_{ij} \quad \mathcal{S}_{\sigma_i\sigma_j} = \mathcal{S}_{ij} \quad \mathcal{T}_{\tau_i\tau_j} = \mathcal{T}_{ij} \quad (\mathcal{A}\tilde{\mathcal{T}})_{\tau_i\gamma_j} = (\mathcal{A}\tilde{\mathcal{T}})_{ij}$$

我们将默认把这些矩阵的列与相应胞腔的线性组合对应起来。例如，

$$COL_\eta(\mathcal{A}) \equiv \sum \mathcal{A}_{\sigma\eta} \cdot \sigma$$

>其中 σ 遍历 X中的所有胞腔。注意，我们不必在变量$^\sigma\sigma$ 中添加下标 i，因为胞腔本身就是矩阵 $\mathcal{A}$的有效指标。

还要注意，$\mathcal{A}$ 和 $\mathcal{T}$是“等价置换”，也就是说，对于所有胞腔 η，ρ，$\mathcal{A}_{\eta\rho} = \mathcal{T}_{\eta\rho}$，尽管一般而言$\mathcal{A}_{ij} \neq \mathcal{T}_{ij}$。同样，$\mathcal{B}$ 和 $\mathcal{S}$ 也是等价置换，也就是说，对于所有胞腔 η，ρ，$\mathcal{B}_{\eta\rho} = \mathcal{S}_{\eta\rho}$，尽管一般而言 $\mathcal{B}_{ij} \neq \mathcal{S}_{ij}$。

5.2 相容的有序基

向量空间 V的一个有序基 是一个有序（不重复）向量序列 $b = (b_1, \ldots, b_k)$，其元素构成 V的一个基。我们称 b与子空间滤过 FV相容：$F_1V \subseteq F_2V \subseteq \cdots \subseteq F_NV$，如果对每个 k，都存在一个 ℓ，使得 $(b_1, \ldots, b_\ell)$构成子空间 F_kV的一个基。在本小节中，我们将证明，算法 1 为持久相对同调中的若干基本滤过生成相容的有序基。稍后的章节中，我们将利用这些基证明算法的正确性。

定理 5. 矩阵 $\mathcal{S}$ （分别地，矩阵 $\mathcal{B}$）的列包含每个相对循环空间$Z(F_tX, G_tX)$的一组基。事实上，序列 $COL_1(\mathcal{B}) \ldots, COL_m(\mathcal{B})$ 是滤过$Z(F_1X, G_1X) \subseteq \cdots \subseteq Z(F_NX, G_NX)$的相容有序基。

证明。 空间 $Z(F_tX, G_tX)$ 是所有满足 $\partial c \in G_tX$ 的链 $c \in C(F_tX)$ 的集合。该空间可以写成两个子空间的交集：(A) 链 $C(F_tX)$ 所构成的空间，它对应于某个 ℓ 下 $\mathcal{D}$ 的前 ℓ 列；(B) 原像 $\partial^{-1}(G_tX)$ 所构成的空间，其中 G_tX 对应于某个 ℓ 下 $\mathcal{D}$ 的前 ℓ 行。因此，[18, 定理 10] 表明，$\mathcal{S}$ 的列包含每个子空间 $Z(F_tX, G_tX)$ 的一组基。对于 $\mathcal{B}$ 的列，同样的结论也成立，因为 $\mathcal{B}$ 是 $\mathcal{S}$ 的一个等价排列（参见第 5.1 节）；特别地，由 $\mathcal{B}$ 的列所表示的链的无序集合，等于由 $\mathcal{S}$ 的列所表示的链的无序集合。由于 $\mathcal{B}$ 的列按照函数 f_{ker} 的值递增排列，因此，对于某个 ℓ，每个 $Z(F_tX, G_tX)$ 都是 $\mathcal{B}$ 的前 ℓ 列的张成空间。 □

定理 6. 矩阵 $\mathcal{T}$ （分别地，矩阵 $\mathcal{A}$）的列包含每个相对边界空间$B(F_tX, G_tX)$的一组基。事实上，序列 $COL_1(\mathcal{A}) \ldots, COL_m(\mathcal{A})$ 是滤过$B(F_1X, G_1X) \subseteq \cdots \subseteq B(F_NX, G_NX)$的一个相容有序基。

证明. 空间 $B(F_tX, G_tX)$ 是两个子空间之和：(A) 边界空间 $\partial(F_tX)$，它对应于 $\mathcal{D}$ 的前 ℓ 列，其中 ℓ取某个值；以及 (B) 链空间 $C(G_tX)$，其中 G_tX 对应于 $\mathcal{D}$的前 ℓ行，其中 ℓ取某个值。因此，定理 10] 说明，$\mathcal{T}$ 的列包含子空间$B(F_tX, G_tX)$的一组基。对于 $\mathcal{A}$的列，同样的结论也成立，因为 $\mathcal{A}$ 是 $\mathcal{T}$的一个等价排列（参见第 5.1 节）；特别地，由 $\mathcal{A}$ 的列表示的链的无序集合，等于由 $\mathcal{T}$的列表示的链的无序集合。由于 $\mathcal{A}$ 的列按照函数 f_{im}的值递增排列，因此，对于某个 ℓ，每个 $B(F_tX, G_tX)$ 都是 $\mathcal{A}$的前 ℓ 列所张成的空间。 □

定理 7。 序列 $COL_1(\mathcal{AT}) \ldots, COL_m(\mathcal{AT})$ 是滤过$B(F_1X, G_1X) \subseteq \cdots \subseteq B(F_NX, G_NX)$的相容有序基。

证明。 由定理 6 可知这一点，因为 $\mathcal{T}$ 可逆且为上三角矩阵。特别地，每一列$COL_k(\mathcal{AT})$都是 $\mathcal{A}$的前 k 列的非平凡线性组合，因此属于与$COL_k(\mathcal{A})$相同的滤过层级。 □

由于 $\tilde{M} = \tilde{\mathcal{T}}^{-1}(\mathcal{A}^{-1}\mathcal{B})\tilde{\mathcal{S}}$ 是可逆矩阵的乘积，$\tilde{M}$ 也可逆。由于 $\tilde{M}$ 是广义匹配矩阵，因此 M 的每一行（分别地，每一列）都恰好包含一个非零元素。设 $\pi =$ $(\pi_1, \pi_2, \ldots)$ 是唯一的排列，使得 $M_{\pi_1,1}, M_{\pi_2,2}, \ldots$ 是 M的非零元素。

定理 8。 序列 $COL_{\pi_1}(\mathcal{AT}) \ldots COL_{\pi_m}(\mathcal{AT})$ 是滤过$Z(F_1X, G_1X) \subseteq \cdots \subseteq Z(F_NX, G_NX)$的相容有序基。

证明。 设给定 k 。因为 M 是广义匹配矩阵，对每个 i 存在非零标量 α，使得 $COL_{\pi_i}(\mathcal{AT}) = \alpha \cdot COL_i(\mathcal{AT}M)$。因此，序列$(COL_{\pi_1}(\mathcal{A}\tilde{\mathcal{T}})\ COL_{\pi_2}(\mathcal{A}\tilde{\mathcal{T}}) \ldots)$中前 k 个向量的张成空间等于 $\mathcal{A}\tilde{\mathcal{T}}\tilde{M}$的前 k列的张成空间。因为 $\tilde{\mathcal{T}}\tilde{M} = (\mathcal{A}^{-1}\mathcal{B})\,\tilde{\mathcal{S}}$，$\mathcal{A}\tilde{\mathcal{T}}\tilde{M}$ 的前 k 列的张成空间等于 $\mathcal{A}(\mathcal{A}^{-1}\mathcal{B})\,\tilde{\mathcal{S}} = \mathcal{B}\tilde{\mathcal{S}}$的前 k 列的张成空间。由于 $\tilde{\mathcal{S}}$ 是可逆且上三角的，$\mathcal{BS}$ 的前 k 列的张成空间等于 $\mathcal{B}$的前 k 列的张成空间。因此

~ ~$(COL_{\pi_1}(\mathcal{AT})\ COL_{\pi_2}(\mathcal{AT})\ldots)$中前 k 个向量的张成空间等于 B的前 k 列的张成空间。因此，所需结论由定理 5 得出，该定理指出 $\mathcal{B}$ 的列构成滤过 $Z(F_1X, G_1X) \subseteq \cdots \subseteq Z(F_NX, G_NX)$的相容有序基。

□

5.3 出生值和死亡值

~这里，我们通过说明如何计算 $f_{im}(COL_i(\mathcal{AT}))$~和 $f_{ker}(COL_i(\mathcal{AT}))$（对于任意 i），给出实现算法 1 所需的最后一个细节。事实上，这一计算只需要少量的查找操作。我们首先将问题归结为关于 $\mathcal{S}$ 和 $\mathcal{T}$的问题。

~**定理 9**。 **</style}我们有 (**$COL_j(\mathcal{AT})) = f_{im}(COL_j(\mathcal{A})) = f_{im}(COL_{\beta_j}(\mathcal{T}))$ **</style}**对所有 j均成立。

~证明。 **</style}**方程 **(()) (())** 由定理 *6* 和定理 *7* 得出，它们指出，序列 **(()() . . .)** 和 $(COL_1(\mathcal{AT})\ COL_2(\mathcal{AT})\ldots)$ 都是滤过 $B(F_1X, G_1X) \subseteq \cdots \subseteq B(F_NX, G_NX)$ 的相容有序基。方程 $f_{im}(COL_j(\mathcal{A})) = f_{im}(COL_{\beta_j}(\mathcal{T}))$ 源于以下事实：根据 $\mathcal{A}$ 的定义， $\mathcal{A}$ 的第 j 列表示的链与 $\mathcal{T}$ 的第 β_j 列表示的是同一条链。 □

~**定理 10.** 对所有 i，我们有 $f_{ker}(COL_{\pi_i}(\mathcal{AT})) = f_{ker}(COL_{\gamma_i}(\mathcal{S}))$。

~ ~证明. 定理 5 和定理 8 表明，序列 $(COL_{\pi_1}(\mathcal{AT})\ COL_{\pi_2}(\mathcal{AT})\ldots)$ 和序列 $(COL_1(\mathcal{B})\ COL_2(\mathcal{B})\ldots)$ 都是圈滤过 $Z(F_1X, G_1X) \subseteq \cdots \subseteq Z(F_NX, G_NX)$ 的相容有序基。因此，对所有 i， $f_{ker}(COL_{\pi_i}(\mathcal{AT})) = f_{ker}(COL_i(\mathcal{B}))$。所需结论由此得出，因为根据 $\mathcal{B}$ 的定义， $COL_i(\mathcal{B})$ 和 $COL_{\gamma_i}(\mathcal{S})$ 在 $C(X)$ 中表示同一条链。

□

~现在，我们已将计算 $\mathcal{AT}$ 的列的出生值和死亡值，简化为计算 $\mathcal{S}$ 和 $\mathcal{T}$ 的列。这些值又可以通过简单的查找公式计算得出:

定理 11。 对于任意胞腔 α，有 $f_{ker}(COL_\alpha(\mathcal{S})) =$最大 $\{b_G(COL_\alpha(M))\ b_F(\alpha)\}$。

证明。 令 $\eta = COL_\alpha(\mathcal{S})$。回想一下，我们将此向量与链 $\sum_\rho \mathcal{S}_{\rho\alpha} \cdot \rho$ 隐式地视为相同，其中 ρ 遍历 X 中的所有胞腔。当且仅当 $\eta \in F_tX$ 且 $\partial(\eta) \in G_tX$ 时，链 η 是 F_tX/G_tX 中的相对循环。因此$f_{ker}(\eta) =$最大$\{b_G(\partial\eta)\ b_F(\eta)\}$。

由于 $\mathcal{S}$ 是对角线上为 1 的上三角矩阵，列 $COL_\alpha(\mathcal{S})$ 是形如 $(\mathcal{S}_{\sigma_1\alpha}, \ldots\ldots, \mathcal{S}_{\sigma_k\alpha}, 0, \ldots\ldots, 0)$ 的标量序列，其中 $\sigma_k = \alpha$ 且 $\mathcal{S}_{\sigma_k\alpha} = 1$。因此，事实上，$\eta \equiv \sum_{i\le k} \mathcal{S}_{\sigma_i\alpha}\sigma_i$。由于我们在构造序列 $\sigma_1, \ldots\ldots, \sigma_m$ 时要求 $b_F(\sigma_1) \le \cdots \le b_F(\sigma_m)$，所以 $b_F(\eta) = b_F(\alpha)$。

现在考虑 $\partial\eta = \mathcal{D}COL_\alpha(\mathcal{S}) = COL_\alpha(\mathcal{DS}) = COL_\alpha(\mathcal{T}M) = \mathcal{T}COL_\alpha(M)$。若 $COL_\alpha(M) = 0$ 则$\partial\eta = 0$，那么 $b_G(\partial\eta) = b_G(0) = b_G(COL_\alpha(M))$。另一方面，若 $COL_\alpha(M) \neq 0$，则存在唯一的胞腔 ρ使得 $M_{\rho\alpha} \neq 0$。在这种情况下， $\mathcal{T}COL_\alpha(M)$ 是 $COL_\rho(\mathcal{T}) = (\mathcal{T}_{\tau_1,\rho}, \ldots, \mathcal{T}_{\tau_k,\rho}, 0, \ldots, 0)$ 的非零标量倍，其中 $\tau_k = \rho$ 且 $\mathcal{T}_{\tau_k,\rho} = 1$。由于定义排列 $\tau_1, \ldots, \tau_m$ 时，我们要求 $b_G(\tau_1) \le \cdots \le b_G(\tau_m)$，因此可得 $b_G(\partial\eta) = b_G(\rho) = b_G(COL_\alpha(M))$。

将计算得到的 $b_G(\partial\eta)$ 和 $b_F(\eta)$ 的值代入表达式 $f_{ker}(\eta) = \max\{b_G(\partial\eta)\ b_F(\eta)\}$，即可得到所需结果。

□

定理 12. 我们有

$$f_{im}(COL_\alpha(\mathcal{T})) = \begin{cases} \min\{b_G(\alpha), b_F(\eta)\} & \textit{if } \exists \textit{ a (unique) cell } \eta \textit{ such that } M_{\alpha,\eta} \neq 0 \\ b_G(\alpha) & \textit{otherwise} \end{cases} \qquad \textit{Span}$$

证明。 根据 $\tau_1, \ldots, \tau_m$ 的构造， $b_G(\tau_1) \le \cdots \le b_G(\tau_m)$，因此每个子空间 G_tX 都可以表示为某个形如 $\{\tau_1, \ldots, \tau_\ell\}$ 的集合的张成空间，其中 ℓ 满足某条件。由于 $\mathcal{T}$ 是上三角的， $\mathcal{T}$ 的第 ℓ 列表示胞腔 $\tau_1, \ldots, \tau_j$ 的线性组合；因此 $\{COL_1(\mathcal{T}) \ldots, COL_\ell(\mathcal{T})\}$ 与 $\{\tau_1, \ldots, \tau_\ell\}$ 的张成空间相同。由此可得：(A) $\mathcal{T}$ 的列包含每个子空间 G_tX 的一组基；(B) 对所有 ℓ，过滤值 $b_G(COL_\ell(\mathcal{T}))$ 等于过滤值 $b_G(\tau_\ell)$。

由于根据构造， $b_F(\sigma_1) \le \cdots \le b_F(\sigma_m)$，对每个 t，都存在一个 ℓ，使得 $\sigma_1, \ldots, \sigma_\ell$ 张成 F_tX。由于 $\mathcal{D}$的列由 $\sigma_1, \ldots, \sigma_m$索引，因此 $\mathcal{D}$的前 ℓ 列张成 $\partial(F_tX)$。由于 $\mathcal{S}$是上三角的， D 的前 ℓ 列的张成空间等于 $\mathcal{DS} = \mathcal{T}M$的前 ℓ 列的张成空间。由于 M是广义匹配矩阵， $\mathcal{T}M$ 的这前 ℓ 列中的每一列要么为零，要么是 $\mathcal{T}$ 中某一（唯一）列的标量倍。因此， $\mathcal{T}$的（线性无关）列包含 $\partial(F_tX)$ 的一组基。此外，列 $COL_i(\mathcal{T})$ 是 $\partial(F_tX)$ 中的边界，当且仅当 $\mathcal{T}M$ 的前 ℓ 列包含$COL_i(\mathcal{T})$ 的一个非零标量倍；这等价于存在某个 $j \le \ell$，使得 $M_{ij} \neq 0$。回想一下，根据定义，$M_{ij} = M_{\tau_i,\sigma_j}$，并注意到 $\{\sigma_1, \ldots, \sigma_\ell\} = \{\sigma_p\colon\ b_F(\sigma_p) \le t\}$，可知该条件等价于存在一个胞腔 σ_j，使得 $b_F(\sigma_j) \le t$ 且 $M_{i,j} \neq 0$。因此，对于每一列 $COL_i(\mathcal{T})$，以下两个条件之一

必须成立：(A) M 的第 i 行为零，此时对于任意 t，$COL_i(T) \notin \partial(F_tX)$；或者 (B) 存在一个 j，使得$M_{ij} \neq 0$，此时 $COL_i(T) \in \partial(F_t) \iff b_F(\sigma_j) \leq t$。

我们现在已经证明，$\mathcal{T}$的列同时包含 (A) 对所有 t，$\partial(F_tX)$ 的一组基；以及 (B) 对所有 t，G_tX 的一组基。因此，对所有 t，$\mathcal{T}$的列包含 $B(F_tX, G_tX) = \partial(F_tX) + G_tX$ 的一组基。由此可知，经过初等推导，$COL_i(\mathcal{T}) \in B(F_tX, G_tX)$ 当且仅当 (I) $COL_i(\mathcal{T}) \in \partial(F_tX)$，或 (II) $COL_i(\mathcal{T}) \in G_tX$。过滤值 $f_{im}(COL_i\mathcal{T})$ 是使这两个条件之一成立的最小 t。对所有 $t \geq b_G(COL_i(\mathcal{T}))$，条件 (II) 均成立；并且我们已经看到，$b_G(COL_i(\mathcal{T})) = b_G(\tau_i)$。对所有 $t \geq t_0$，条件 (I) 均成立，其中 t_0 要么是 $b_F(\sigma_j)$（若存在 σ_j 使得 $M_{\tau_i,\sigma_j} \neq 0$），要么是 ∞（若 M 的第 j 行为零）。令 $\alpha = \tau_i$ 和 $\eta = \sigma_j$，即可得到所需结论。 □

5.4 正确性证明

~设 U 表示 $\mathcal{AT}$的列所构成的无序集合，这些列被视为链。定理 7 和定理 8 指出，U 包含相对循环的空间 $Z(F_tX, G_tX)$以及相对边界的空间 $B(F_tX, G_tX)$的一个基，对于每个 t。因此，线性代数中的一个初等练习表明，集合差 $\Delta_t := \Big(U \cap Z(F_tX, G_tX)\Big) \setminus \Big(U \cap B(F_tX, G_tX)\Big)$ 构成商空间 $H_*(F_tX, G_tX) = {}^{Z}(F_tX, G_tX)\, B(F_tX, G_tX$ 的一个基。根据 f_{ker}的定义，元素 $u \in U$属于 $Z(F_tX, G_tX)$当且仅当 $f_{ker}(u) \leq t$。同样，u属于 $B(F_tX, G_tX)$当且仅当 $f_{im}(u) \leq t$。因此u 属于基 Δ_t 当且仅当 $f_{ker}(u) \leq t < f_{im}(u)$。

这些观察综合表明，对于 $(F_\bullet, G_\bullet)$，每个 $u \in U$ 都对应一个形如 $[f_{ker}(u)\ f_{im}(u))$ 的条形码条，而与该条对应的同调类具有相对循环代表元 u。因此，集合$I = \{(f_{ker}(u)\ f_{im}(u)\ u) : u \in U\}$ 是条形码的一种有效表示，并附有代表元。这与算法 1 返回的集合 I 相同。因此，该算法是正确的。

5.5 复杂度

该算法包含若干查找操作（线性时间）、排序操作（$O(m\log\ m)$ 时间）和 U-match 分解（$O(m^3)$ 时间，参见 [18]），其中 m 是胞腔复形中的胞腔数。因此，算法 1 的复杂度为 $O(m^3)$。

6 滞后滤过的性能优化

大多数持久同调计算面临的一个根本挑战，是对矩阵进行分解的问题：矩阵规模会随输入增大而急剧膨胀，通常有数千亿行和列。在计算拓扑领域，人们已在矩阵为 $\mathcal{D}$（即立方复形或单纯复形的微分）的特殊情形下，取得了应对这一挑战的重大进展。然而，尽管许多进展可加以改造，用于支持算法 1 中第一个 U-match 分解（$\mathcal{T}, M, \mathcal{D}, \mathcal{S}$）的计算，但它们无法帮助进行第二个分解（$\tilde{\mathcal{T}}, \tilde{M}, \mathcal{A}^{-1}\mathcal{B}, \tilde{\mathcal{S}}$）。

这一挑战自然促使人们寻找能够加速持久相对同调计算的特殊情形。事实上，计算相对同调领域的大量现有文献，都可以视为试图找出可应用有效性能优化的特定数据类型。

在此，我们提出一类新的例子及其对应的 U-match 优化：滞后过滤。令 X 为配备过滤 $F_\bullet$：$F_1X \subseteq \cdots \subseteq F_NX = X$和 $G_\bullet$：$G_1X \subseteq \cdots \subseteq G_NX = X$的有限胞腔复形。我们称 $G_\bullet$ 为 $F_\bullet$ 的滞后过滤，若存在常数 l，使得对所有 $1 \leq t \leq N$都有 $G_tX = F_{t-l}X$；按照约定，当 $t \leq 0$时，$F_tX := 0$。在这种情况下，存在一个形式为 $(J,\ M,\ \mathcal{D},\ J)$ 的 U-match 分解，其中 $\mathcal{D}$是微分矩阵的一个副本，其行和列按照出生函数 b_F [18, 对应的过滤顺序排列（该顺序可能不唯一）。备注 8]。事实上，可以使用任何能产生第 1.3 节和引理 16]中有效 $R = DV$ 分解的算法来计算此分解 $\mathcal{D}$ [18,。

可逆上三角矩阵 J 的列包含形式为(A) F_tX、(B) $\partial(F_tX)$或(C) $\partial^{-1}(F_tX) = \{c \in C(X) : \partial c \in F_tX\}$的每个子空间的一个基；此外，$J$ 的列包含任何可由形式(A)、(B)或(C)的子空间通过求和或交得到的子空间的一个基 [18, 定理 25]。特别地，J的列包含每个相对循环空间 $Z(F_tX, G_tX) = F_tX \cap \partial^{-1}(G_tX)\, F_tX \cap \partial^{-1}\ F_{t-l}X$ $G_tX + \partial\ F_tX = F_{t-l}X + \partial\ F_tX$ ( )以及每个相对边界空间( ) ( )的一个基。

这一最后的观察表明，我们可以遵循第5.4节中所述的精确步骤，计算$(F_\bullet, G_\bullet)$的~持久相对同调条形码；只需用矩阵 J 替换矩阵 $\mathcal{AT}$。相对边界次水平集函数 f_{im} 与相对循环次水平集函数 f_{ker} 对每一列 $COL_j(J)$ 的值，可以按照与第5.3节所述类似的方式计算。

综合来看，这些观察表明，在滞后过滤这一特殊情形下，计算持久相对同调条形码的成本可与 $R = DV$ 分解相当。尽管 $R = DV$ 的渐近复杂度与算法 1 相同，但在实践中，由于采用了高度优化的 $R = DV$ 求解器，性能得到了大幅提升。我们认为，这可能是该问题领域中众多潜在性能优化之一。

7 实现

我们提供了一个针对滞后过滤（参见第 6 节）优化的持久相对同调条形码计算流程实现，可从 [1] 获取。该实现扩展自 Open Applied Topology [20]，这是一个开源计算拓扑学软件包，采用 Rust 编写底层后端，并以 Python 构建高层前端。

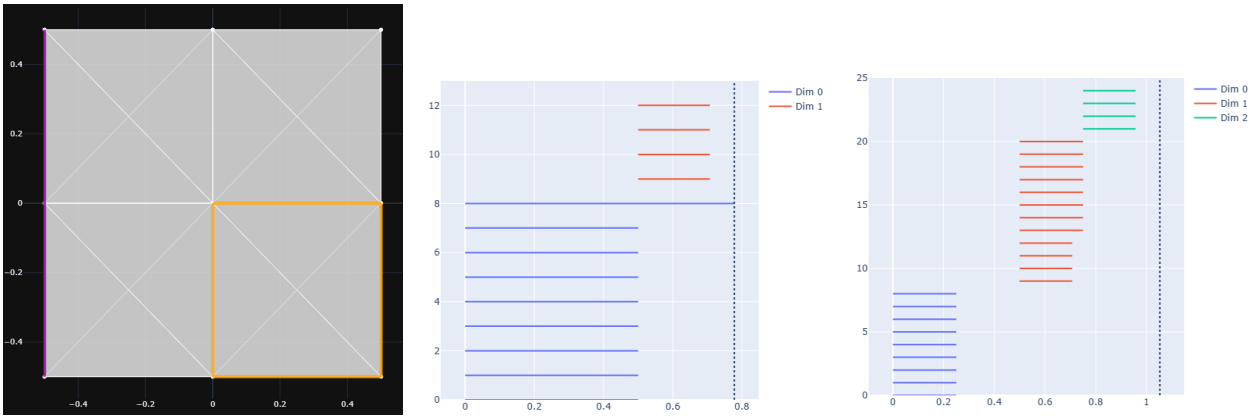

图 2：由位于 3×3 网格上的节点组成的点云，其中网格上相邻节点满足 a，b，而对角相邻节点满足 $\|a - b\|_2 = 1/2$。图中展示了相对（紫色）和绝对（橙色）1-循环的示例（**左**），以及对应于滞后为 $l = 1/4$ 的滤过的标准持久同调条形码（**中**）和持久相对同调条形码（**右**）。注意右图中，没有任何同调类的寿命能超过 l，包括本质0-同调。

图 2 对基于滞后为 $l = 1/4$ 的网格节点点云构建 Vietoris-Rips复形作了概念性说明。注意，有四个绝对 H_1 类，其区间模的形式为 $\mathbb{I}\langle 1/2,\ \sqrt{2}/2\rangle$（中）。有八个相对1-同调类，其区间模的形式为 $\mathbb{I}\langle 1/2,\ \overline{3/4}\rangle$（右）。其中四个在过滤值 $\sqrt{2}/2$（对角边的长度）处消亡，其余则持续到所有外部边进入子复形（在 0.75 处）。此外还有四个相对2-同调类。在过滤值$\sqrt{2}/2$ 处出现的2-单形构成这些球面的表面，而在 $3/4$ 处进入子复形的1-单形则是它们的边界。

图 3 中的示例针对欧氏平面中的随机点云，计算了若干不同滞后参数 l 对应的条形码。随着 l 增大，条形码的初始部分收敛于数据的绝对持久同调条形码。当 l 较小时，会出现各种中间相对特征。我们认为，这些特征可能对关注数据局部结构的科学应用具有价值，例如材料科学。

8 结论

我们提出了两种通过U-match 分解计算持久相对同调的新方法，着眼于拓扑数据分析中的应用。这些方法可借助初等线性代数以高度透明的方式加以分析。本文的主要贡献是一种至多使用两次矩阵分解即可运行的算法。我们不仅计算持久相对同调的条形码，还计算持久相对循环代表元。我们还针对滞后过滤这一特殊情形给出了一种性能优化方法，并相信未来仍有许多提升性能的机会。

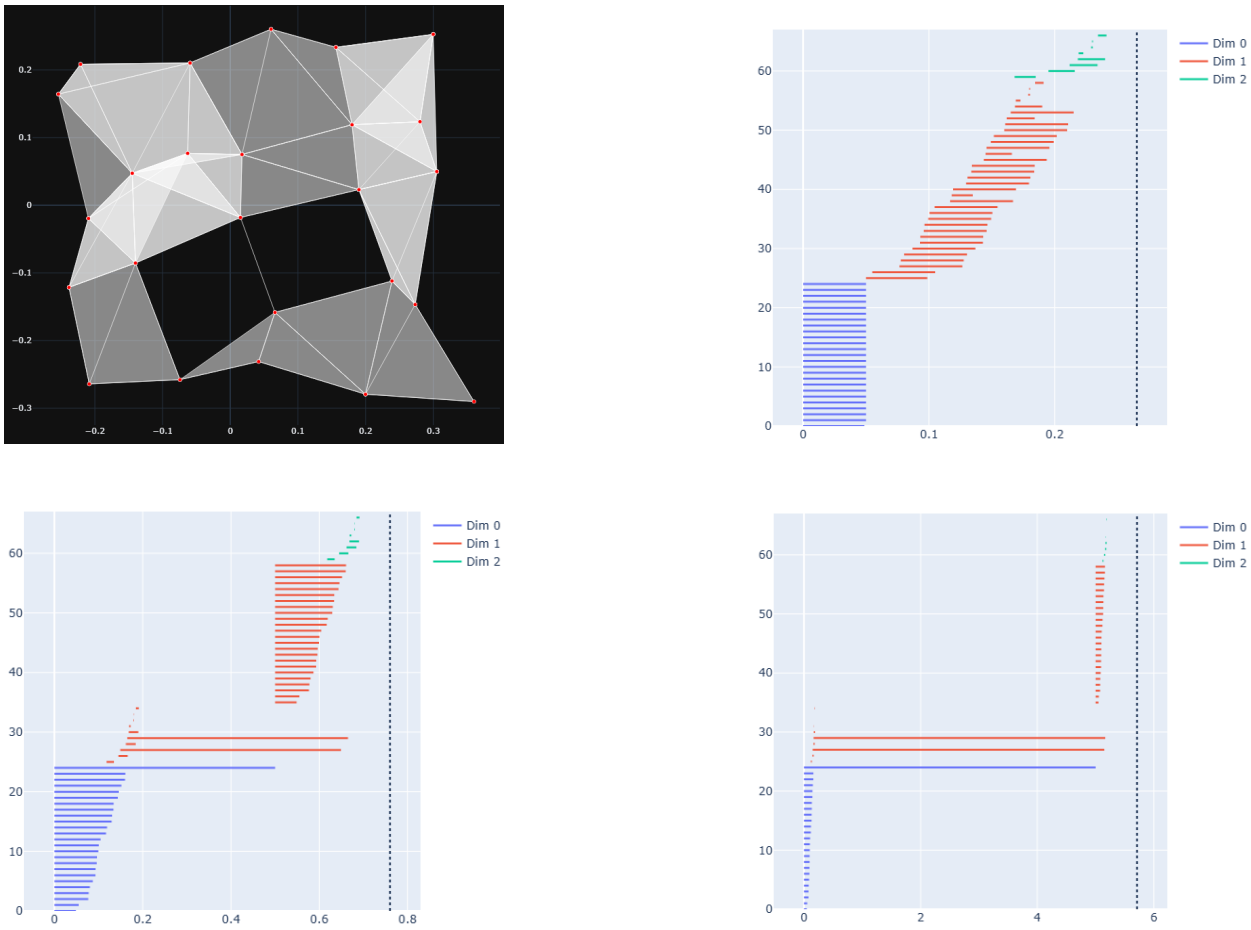

图3：固定过滤值 ε （**上行，左**）下的点云及（叠加的）Vietoris-Rips复形，以及使用滞后参数 $\delta_1 = 0.05$ （**上行，右**）、$\delta_2 = 0.5$ （**下行，左**）和 $\delta_3 = 5$（**下行，右**）计算得到的三个持久相对同调条形码序列。

参考文献

[1] U-match持久相对同调，2025年11月。URL：https://anonymous.4open.science/r/ Umatch_PersistentRelativeHomology/README.md.[2] Mehmet E. Aktas、Ahmed El Fatmaoui与Esra Akbas。网络的持久同调：方法与应用。*Applied Network Science*，2019年。doi:10.1007/s41109-019-0179-3.[3] Subhrajit Bhattacharya、Robert Ghrist与Vijay Kumar。不确定环境中路径规划的持久同调。*IEEE*机器人学汇刊，31(3):578–590，2015年。doi:10.1109/TRO.2015.2412051.[4] Nello Blaser与Morten Brun。相对持久同调。离散与计算几何，68(4):949– 963，2022年。[5] G. Carlsson。拓扑学与数据。美国数学会通报，46:255–308，2009年。[6] Gunnar Carlsson与Afra Zomorodian。计算持久同调。载于第二十届计算几何年度研讨会论文集，第347–356页，2004年。[7] David Cohen-Steiner、Herbert Edelsbrunner与John Harer。利用poincar´e与Lefschetz对偶扩展持久性。计算数学基础，9:79–103，2009年。doi:10.1007/s10208-008-9027-z.[8] David Cohen-Steiner、Herbert Edelsbrunner、John Harer与Dmitriy Morozov。核、像与余核的持久同调。见第二十届*ACM-SIAM*离散算法年度研讨会论文集，第1011–1020页。SIAM，2009年。

[9] David Cohen-Steiner、Herbert Edelsbrunner 与 Dmitriy Morozov。通过线性时间更新持久性得到的藤蔓与葡萄园。载于第二十二届计算几何年度研讨会论文集，第119–126页。计算机协会，2006年。[10] Martin D Crossley。基本拓扑学，第9章。施普林格科学与商业媒体，2006年。[11] Martin D Crossley。基本拓扑学，第5章。施普林格科学与商业媒体，2006年。[12] Vin De Silva、Dmitriy Morozov 与 Mikael Vejdemo-Johansson。持久（上）同调中的对偶性。《逆问题》，27(12):124003，2011年。[13] Tamal Krishna Dey 与 Yusu Wang。面向数据分析的计算拓扑学，第3-4章。剑桥大学出版社，2022年。[14] Herbet Edelsbrunner 与 John L. Harer。《计算拓扑学：导论》。美国数学学会，2022年。[15] Herbet Edelsbrunner、David Letscher 与 Afra Zomorodian。拓扑持久性与简化。《离散与计算几何》，第511–533页，2002年。[16] Peter Gabriel。不可分解表示 i。《数学手稿》，6:71–103，1972年。[17] Robert Ghrist。条形码：数据的持久拓扑。《美国数学学会通报》（新系列），45，2008年2月。doi:10.1090/S0273-0979-07-01191-3。[18] Haibin Hang、Chad Giusti、Lori Ziegelmeier 与 Gregory Henselman-Petrusek。U-match 分解：稀疏同调代数、惰性循环代表元，以及持久（上）同调中的对偶性。2021年。arXiv:2108.08831。[19] Allen Hatcher。《代数拓扑》。剑桥大学出版社，2002年。[20] Gregory Henselman-Petrusek、Ziegelmeier Lori、Chad Giusti 与 Haibin Hang。Open Applied Topology，2023年10月。URL: https://openappliedtopology.github.io。[21] Michael Kerber 与 Matthias S¨ols。局部化球并双过滤。载于第*39*届国际计算几何研讨会（*SoCG 2023*），第45–1页。Schloss Dagstuhl–Leibniz-Zentrum f¨ur Informatik，2023年。[22] Dmitriy Morozov。持久性与稳定性的同调幻象。博士论文，杜克大学，2008年。[23] Chi Seng Pun、Kelin Xia 与 Si Xian Lee。基于持久同调的机器学习及其应用——综述。可于 *SSRN 3275996* 获取，2018年。[24] Michael Robinson、Chris Capraro、Cliff Joslyn、Emilie Purvine、Brenda Praggastis、Stephen Ranshous 与 Arun Sathanur。抽象单纯复形的局部同调，2018年。arXiv:1805.11547。[25]Michael Robinson 与 Jimmy Palladino。协议无关的关键节点检测。*arXiv* 预印本 *arXiv:1607.06022*，2016年。