An Introduction to Multiparameter Persistence

Magnus Bakke Botnan and Michael Lesnick

Abstract

In topological data analysis (TDA), one often studies the shape of data by constructing a filtered topological space, whose structure is then examined using persistent homology. However, a single filtered space often does not adequately capture the structure of interest in the data, and one is led to consider multiparameter persistence, which associates to the data a space equipped with a multiparameter filtration. Multiparameter persistence has become one of the most active areas of research within TDA, with exciting progress on several fronts. In this article, we introduce multiparameter persistence and survey some of this recent progress, with a focus on ideas likely to lead to practical applications in the near future.

Contents

1 Introduction

1.1 Overview

Topological data analysis (TDA) is a branch of data science which applies topology to study the *shape* of data, i.e., the coarse-scale, global, non-linear geometric features of data. Examples of such features include clusters, loops, and tendrils in point cloud data, as well as modes and ridges in functional data. While the history of TDA dates back to the 1990's, in recent years the field has advanced rapidly, leading to a rich theoretical foundation, highly efficient algorithms and software, and many applications [152, 45, 158, 106].

On a high level, the TDA pipeline generally involves two steps: Given a data set X (e.g., a finite set of points in $\mathbb{R}^n$), we

1. construct a commutative diagram of topological spaces $F(X)$ whose topological structure encodes information about the shape of X, and then

2. analyze the structure of $F(X)$ using tools from topology and abstract algebra.

The most common instance of this pipeline is *(1-parameter) persistent homology.* Here the diagram $F(X)$ is assumed to be indexed by a totally ordered set A (e.g., $A = \mathbb{R}$, or $A = \mathbb{N}$), and one applies homology with field coefficients, along with an algebraic structure theorem, to obtain signatures of data called *barcodes.* We give a detailed overview in Section 3.

However, there are many settings in TDA where one wishes to construct a diagram of spaces $F(X)$ indexed by a poset that is not totally ordered. In particular, it is often natural to consider a diagram indexed by a product A of n totally ordered sets, e.g. $A = \mathbb{R}^n$ or $A = \mathbb{N}^{\mathrm{op}} \times \mathbb{R}$; we call these *(n-parameter) filtrations*, or *multifiltrations.* The branch of TDA which studies data via multifiltrations is called *multiparameter persistence* or (when homology is used to study these diagrams) *multiparameter persistent homology* (MPH). The cases $n = 2, 3$ are of primary interest and the $n = 2$ case has received the most attention thus far. As we will see in Sections 4 and 5, the use of 2- and 3-parameter persistent homology is natural when working with

- data with outliers or variations in density (indeed, standard constructions of 1-parameter persistent homology are notoriously unstable to outliers, which motivates a 2-parameter approach; see Section 3.4),

- time-varying data, or more generally, data equipped with a real-valued function (e.g., the partial charge function on the atom centers of a protein),

- data with *tendrils* emanating from a central core,

- functional data (e.g. image data) with noise that is large in magnitude but localized ("spike noise").

Multiparameter persistence was introduced by Frosini and Mullazani, who considered multiparameter persistent homotopy groups [97]. Multiparameter persistent homology was first considered by Carlsson and Zomorodian [47]; several of the key ideas about MPH we will encounter in this introduction first appeared in [47].

MPH provides algebraic invariants of data called *(n-parameter) persistence modules*, simply by applying homology with field coefficients to a multifiltration; in the case that $A = \mathbb{Z}^n$, these are $\mathbb{Z}^n$-graded modules over a polynomial ring in n variables, objects classically studied in commutative algebra [144]. Persistence modules can also be viewed as commutative diagrams of vector spaces, which have been studied extensively in quiver representation theory [78]. Thus, classical work in algebra offers a highly developed set of mathematical tools for working with MPH. However, the sorts of mathematical questions about these objects that arise in TDA are often very different than the ones encountered in classical work in algebra; as we will see, these questions have stimulated the development of a rich theory for MPH, distinct from classical work.

When $n \geq 2$, there are difficulties with defining the barcode of a n-parameter persistence module; see Section 4.1 for a detailed discussion. Quiver representation theory provides an illuminating perspective on this: multiparameter persistence modules have *wild representation type.* In brief, this means that in

contrast to the 1-parameter case, the natural generalization of a barcode in the n-parameter setting is enormously complicated, and in general far too complex to work with in practice. We give an overview of this quiver representation theory and its connections to MPH in Section 8.

Because of these difficulties, the standard methodology for data analysis using 1-parameter persistence tends not to extend naively to the multiparameter setting, and new ideas are needed to make practical use of multiparameter persistence. Recently, there has been substantial progress on the development of such ideas. These advances suggest that MPH is an especially natural approach to the study of the shape of data, and one with potential to have substantial impact on data science. Yet, for this potential to be reached, much critical work remains.

In this article, we introduce multiparameter persistent homology, focusing on some of the topics that we feel are most fundamental and most likely to lead to practical applications in the near future. We hope that this article will serve as an invitation for others–mathematicians, computer scientists, statisticians, and application specialists—to join the effort to realize the potential of this approach to data analysis. While we have tried to make this article as accessible as possible, we do assume some familiarity with a few elementary mathematical topics that arise in the study of 1-parameter persistence, such as basic category theory, homology, and abstract algebra.

We mention three other introductory resources on multiparameter persistence: First, Karagüler's recent M.S. thesis [119] provides an introduction to the theoretical aspects of multiparameter persistence, covering some of the same territory as our article, as well as some topics not discussed here. Second, two chapters in Dey and Wang's textbook [82] focus on multiparameter persistence, particularly its computational aspects. Third, Hal Schenck's recent book [163] has some material on the algebraic aspects of multiparameter persistence.

1.2 Main Themes

This article explores six main themes which underly much of the recent progress in MPH:

Invariants Though defining multiparameter barcodes is problematic, many simple invariants of persistence modules are available to us, which can serve as a surrogate for the barcode in applications. Three simple invariants which play a central role in the MPH literature are the *Hilbert function*, the *rank invariant*, and the *multigraded Betti numbers*. As we will explain, the rank invariant is equivalent to each of two different barcode-like invariants, the *fibered barcode* [134] and the *signed barcode* [36].

This signed barcode in fact belongs to a family of closely related generalized barcode constructions, which can be defined either in terms of Möbius inversion or relative homological algebra [18, 139, 155, 37]. In particular, the rank invariant admits a natural extension, the *generalized rank invariant* [127], which has an analogous signed barcode [2, 127, 36].

We introduce invariants of multiparameter persistence modules in Section 4.2, and consider signed barcodes in detail in Section 9.

4

Visualization Visualization of barcodes has been critical to the practical success of 1-parameter persistence. The problem of visualizing (invariants of) MPH in a practical, computationally efficient way seems to be of similar importance to the success of MPH. RIVET [134, 79], a practical software for working with 2-parameter persistent homology, introduces a novel interactive visualization of the Hilbert function, the bigraded Betti numbers, and the fibered barcode. We illustrate RIVET's visualization on several examples in Section 4.3. Other ideas for visualizing MPH include the signed barcode [36], the *multiparameter persistent landscape* [178] and (in an important special case) the *staircase code* of [44].

Metrics and Stability Metrics on barcodes play an especially important role in both theory and applications of 1-parameter persistence. Analogously, to develop the theory and applications of MPH, we need good metrics in the multiparameter setting. Perhaps surprisingly, though defining barcodes for MPH is problematic, we have well-behaved extensions of the standard metrics on barcodes to the multiparameter persistence modules. Of these, the best known is the *interleaving distance*, which extends the *bottleneck distance* on persistence barcodes. The interleaving distance is useful in the theory of MPH, e.g., in establishing that several MPH constructions are robust to outliers [31]. It has been shown that computing the interleaving distance on n-parameter persistence module is NP-hard for $n \geq 2$ [25], which motivates the search for a more computable surrogate. The *matching distance* [54] has emerged as one attractive choice.

The interleaving distance is an ℓ^∞-type distance, i.e., it can be defined in terms of the ℓ^∞ distance on Euclidean spaces [27]. As such, it is insensitive to small-scale algebraic structure that can be important in applications. Several recent works explore the problem of defining ℓ^p-type distances on multiparameter persistence modules.

We discuss metrics on persistence modules and the stability theory for MPH in Section 6.

Computation and Software Efficient algorithms are a critical prerequisite to practical applications of MPH, as is user-friendly software implementing such algorithms. Activity in the computational aspects of MPH has accelerated in recent years, especially in the 2-parameter setting, leading to several important advances which have lowered the barrier to applications. Arguably, the most fundamental computational problem in MPH is *minimal presentation computation*. Recent work shows that the standard algorithm for computing barcodes in the 1-parameter setting extends to a similarly efficient algorithm for computing minimal presentations of 2-parameter persistence modules, which has been implemented in [135] and (in a substantially improved form) in [126]. Other notable progress on the computational aspects of MPH concerns the computation of certain bifiltrations [71, 136, 87, 88], the computation of distances between multiparameter persistence modules [125, 25, 123, 80], and decomposition of MPH modules into indecomposables [81]. We discuss computation in several places where related concepts arise, specifically in Sections 5.4, 6.1, 6.2, 7.3, 8.5 and 9.1.

Barcodes in Special Cases In some special cases of interest, we do have well-defined barcodes that are simple enough to work with. In particular, a fundamental 2-parameter persistence invariant of $\mathbb{R}$-valued functions called *inter-level set persistence* has simple barcodes analogous to those in the 1-parameter setting [66, 14, 32]. A part of the MPH literature focuses on developing a theoretical and computational understanding of such special cases, and some interesting discoveries have been made. This will be discussed in Section 10.

Applications There have been several efforts to develop practical applications of MPH, e.g., to image analysis, computational chemistry, and shape analysis, which hint at the promise of multiparameter persistence as a practical approach to data analysis [122, 1, 177, 17, 21, 164, 51, 65, 13, 178]. Nevertheless, in spite of the widespread interest in MPH among the TDA community and very encouraging recent progress in the field, applications of MPH are still in their infancy. Arguably, this is in large part because the algorithms and software tools needed to study data at scale using MPH have been introduced only very recently, and are still very much under development. We discuss two recent applications of MPH to cancer imaging in Section 4.4.

1.2.1 Other Key Themes

The literature of multiparameter persistence is growing rapidly; we have made no attempt to discuss all of the interesting and promising recent research, and the choice of topics here is biased towards our own research interests. We draw the reader's attention to two interesting and important themes which are not covered in this article:

Sheaf-theoretic Viewpoints on Multiparameter Persistence There is a large literature on connections between sheaves and generalized persistence, which is very closely related to MPH, e.g., [16, 120, 77, 75]. The sheaf-theoretic framework naturally extends the multiparameter persistence framework, and offers powerful tools (e.g., sheafification) that promise to be useful in some of the settings considered in our paper.

Infinitely Presented but Tame Modules The diagrams of spaces one deals with in applications of multiparameter persistence are often indexed by $\mathbb{R}^n$. In many cases, the homology modules of these diagrams are finitely presented, in which case their study essentially reduces to the study of persistence modules indexed by $\mathbb{N}^n$. But it is also common to encounter settings where the persistence modules are not finitely presented, yet satisfy a weaker finiteness property called *tameness*, where generators and relations appear not only at single points in $\mathbb{R}^n$ but along continuous curves. The algebraic properties of tame modules have been explored in depth in recent work of Miller [142], whose work was motivated in part by the study of fly wing morphology [140]; see also [141, 143]. Two recent (independent) works [6, 41] draw on ideas from singularity theory [180] to show that for M a compact manifold, the 2-parameter sublevel persistent homology (see Section 3.1) of a generic smooth function $f : M \to \mathbb{R}^2$ is tame; in fact, a precise description of the topology of the sublevel filtration can be given in terms of the singularities of f. We believe that these ideas are

6

likely to be useful in both theory and applications of multiparameter persistent homology. Closely related ideas also appear in [56, Section 3].

1.3 Outline

Section 2 introduces multiparameter filtrations and persistence modules, the main objects of study in multiparameter persistence. Section 3 reviews 1-parameter persistent homology and its stability properties. Section 4 introduces multiparameter persistent homology: Section 4.1 discusses the difficulty of defining barcodes of multiparameter persistence modules, Section 4.2 discusses invariants of such modules, Section 4.3 gives an example of RIVET's visualization, and Section 4.4 discusses two applications of MPH to cancer imaging. Section 5 introduces several constructions of multifiltrations from data, and discusses the computation of some of these. Section 6 considers metrics on multiparameter persistence modules and stability results for MPH. Section 7 discusses the problem of computing minimal presentations and resolutions of persistence modules. Section 8 provides a brief introduction to quiver representation theory and its connections to MPH. Section 9 introduces signed barcodes of multiparameter persistence modules. Section 10 considers special settings where bipersistence modules decompose into *interval modules*, and thus have nice barcodes, including the important case of *interlevel persistent homology*.

Acknowledgements

2 Filtrations and Persistence Modules

2.1 Posets

Given posets P and Q, we let $P \times Q$ denote the *product poset*, i.e. the partial order on $P \times Q$ is given by $(x_1, y_1) \leq (x_2, y_2)$ if and only if $x_1 \leq x_2$ and $y_1 \leq y_2$. More generally, the product of n posets

$$P_1 \times P_2 \times \cdots \times P_n$$

is defined analogously.

The *opposite poset* of P, denoted P^{op}, is the poset with the same underlying set, with $x \leq y$ in P^{op} if and only if $y \leq x$ in P.

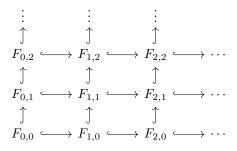**Definition 2.1.** An *interval* in a poset P is a non-empty subset I of P satisfying the two following conditions:

1. If $s, t \in I$ and $s \leq u \leq t$, then $u \in I$,

2. If $s, t \in I$, then there are $s = u_0, \ldots, u_m = t \in I$ such that u_i and u_{i+1} are comparable for all $0 \leq i < m$.

A poset P can be considered as a category with objects the elements of P and morphisms:

$$\mathrm{Hom}(x, y) = \begin{cases} \{*\} & (x \leq y), \\ \varnothing & (x \not\leq y). \end{cases}$$

2.2 Filtrations

Let **Top** denote the category of topological spaces and continuous maps. A *(P-indexed) filtration* is functor $F : P \to \mathbf{Top}$ such that if $x \leq y$, then $F_x \subseteq F_y$ and the map $F_{x,y} : F_x \to F_y$ is the inclusion. In plainer language, F is a choice of a topological space F_x for each $x \in P$ such that $F_x \subseteq F_y$ whenever $x \leq y$. If $P = T_1 \times \cdots \times T_n$ where each T_i is a totally ordered set, then we call F a *multiparameter or n-parameter filtration*. When $n = 2$, we often call F a *bifiltration*. For example, a bifiltration $F : \mathbb{N} \times \mathbb{N} \to \mathbf{Top}$ is a diagram of topological spaces of the form

$$
\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\Big\updownarrow & & \Big\updownarrow & & \Big\updownarrow & & \\
F_{0,2} & \hookrightarrow & F_{1,2} & \hookrightarrow & F_{2,2} & \hookrightarrow & \cdots \\
\Big\updownarrow & & \Big\updownarrow & & \Big\updownarrow & & \\
F_{0,1} & \hookrightarrow & F_{1,1} & \hookrightarrow & F_{2,1} & \hookrightarrow & \cdots \\
\Big\updownarrow & & \Big\updownarrow & & \Big\updownarrow & & \\
F_{0,0} & \hookrightarrow & F_{1,0} & \hookrightarrow & F_{2,0} & \hookrightarrow & \cdots
\end{array}
$$

We will often consider filtrations valued in the category of simplicial complexes **Simp**, which we regard as a subcategory of **Top** via geometric realization.

For fixed P, the P-indexed filtrations form a category whose morphisms are natural transformations. That is, a morphism $\eta : F \to G$ in this category is a choice of a continuous map $\eta_x : F_x \to G_x$ for each $x \in P$, such that for all $x \leq y \in P$, the following diagram commutes:

$$
\begin{array}{ccc}
F_x & \hookrightarrow & F_y \\
\eta_x \downarrow & & \downarrow \eta_y \\
G_x & \hookrightarrow & G_y
\end{array}
$$

For example, given filtrations $F, G : \mathbb{N} \to \mathbf{Top}$, a morphism is an extension of F and G to a commutative diagram of the following shape:

$$
\begin{array}{ccccccc}
F_0 & \hookrightarrow & F_1 & \hookrightarrow & F_2 & \hookrightarrow & \cdots \\
\downarrow & & \downarrow & & \downarrow & & \\
G_0 & \hookrightarrow & G_1 & \hookrightarrow & G_2 & \hookrightarrow & \cdots
\end{array}
$$

Sublevel Filtrations

Definition 2.2. Given a poset P, a topological space W, and a (not necessarily continuous) function $\gamma\colon W \to P$, we define the *sublevel filtration of γ* as
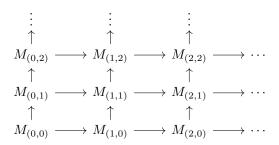
$$\mathcal{S}^{\uparrow}(\gamma)\colon P \to \textbf{Top},$$
$$\mathcal{S}^{\uparrow}(\gamma)_p = \{w \in W \mid \gamma(w) \leq p\}.$$

In the case $P = \mathbb{R}^n$, this construction was introduced in [97]. A filtration isomorphic to a sublevel filtration is said to be *1-critical* [50, 30], whereas a filtration that is not 1-critical is said to be *multicritical*. Note that if $F\colon P \to \textbf{Top}$ is 1-critical, then each element of $\bigcup_{p \in P} F_p$ first appears in the filtration at some unique minimal index in P, whereas this need not be the case for multicritical filtrations. We will see later that 1-critical and multicritical bifiltrations both arise naturally from data.

2.3 Persistence Modules

Fix a field k, for example, $k = \mathbb{Z}/2\mathbb{Z}$, and let **Vect** denote the category of k-vector spaces and linear maps. For P a poset, a *(P-indexed) persistence module*, or simply a *P-module*, is a functor $M\colon P \to \textbf{Vect}$. For $x \leq y$ we let $M_{x,y}$ denote the morphism $M_x \to M_y$. If $P = T_1 \times \cdots \times T_n$ where each T_i is a totally ordered set, then M is called a *multiparameter or n-parameter persistence module*. When $n = 2$, we also call M a *bipersistence module*.

For example, a bipersistence module $M\colon \mathbb{N} \times \mathbb{N} \to \textbf{Vect}$ is a commutative diagram of k-vector spaces of the form:

$$
\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \\
M_{(0,2)} & \longrightarrow & M_{(1,2)} & \longrightarrow & M_{(2,2)} & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \\
M_{(0,1)} & \longrightarrow & M_{(1,1)} & \longrightarrow & M_{(2,1)} & \longrightarrow & \cdots \\
\uparrow & & \uparrow & & \uparrow & & \\
M_{(0,0)} & \longrightarrow & M_{(1,0)} & \longrightarrow & M_{(2,0)} & \longrightarrow & \cdots
\end{array}
$$

A persistence module is said to be *pointwise finite-dimensional*, or *p.f.d.* if $\textbf{dim}(M_x) < \infty$ for all $x \in P$.

Applying i^{th} homology with coefficients in k to each space and each inclusion map in a filtration $F\colon P \to \textbf{Top}$ yields a persistence module $H_i(F)\colon P \to \textbf{Vect}$.

For fixed P, the P-modules form a category $\textbf{Vect}^P$ whose morphisms are the natural transformations, just as for filtrations. By working pointwise, the category $\textbf{Vect}^P$ inherits direct sums (i.e., coproducts), images, kernels, cokernels from **Vect**. In fact, $\textbf{Vect}^P$ is an *abelian category*. For example, given a morphism $f\colon M \to N$ of P-modules, $\ker f$ is the P-module defined by $(\ker f)_x :=$ $\ker(f_x)$ for all $x \in P$, with the internal maps of $\ker f$ given by restricting $M_{x,y}$ to $(\ker f)_x$. We say that $M\colon P \to \textbf{Vect}$ is *indecomposable* if $M \cong M' \oplus M''$ implies that $M' = 0$ or $M'' = 0$.

Definition 2.3. For an interval I in a poset P, the *interval module* k_I is defined by

$$(k_I)_x = \begin{cases} k & \text{if } x \in I, \\ 0 & \text{otherwise,} \end{cases} \qquad (k_I)_{x,y} = \begin{cases} \text{Id}_k & \text{if } x \leq y \in I, \\ 0 & \text{otherwise.} \end{cases}$$

It is not hard to verify that k_I is indecomposable; see e.g. [33, Proposition 2.2].

2.4 Persistence Modules as Multigraded Modules

As the name suggests, persistence modules can be seen as modules over a ring, and sometimes this perspective is very useful. We explain this just in the case where $P = \mathbb{Z}^n$, though the story is much the same when $P = \mathbb{R}^n$ [133].

Definition 2.4. Let $\mathbf{e}_i$ denote the ith standard basis vector in $\mathbb{Z}^n$. An *n-grading* on a $k[x_1, \ldots, x_n]$-module M is a vector space decomposition

$$M = \oplus_{z \in \mathbb{Z}^n} M_z$$

such that $x_i M_z \subset M_{z+\mathbf{e}_i}$ for all $z \in \mathbb{Z}^n$ and $i \in \{1, \ldots, n\}$. A $k[x_1, \ldots, x_n]$-module M is said to be *n-graded* if it comes equipped with an n-grading.

A *morphism* $f : M \to N$ of n-graded modules is a module homomorphism (in the usual sense) such that $f(M_z) \subset N_z$ for all $z \in \mathbb{N}^n$. With these morphisms, the n-graded modules form a category Mod.

Proposition 2.5 (Carlsson, Zomorodian [47, 70]). *The categories* $\mathbf{Vect}^{\mathbb{Z}^n}$ *and Mod are equivalent.*

$k[x_1, \ldots, x_n]$-modules are the basic objects of study in commutative algebra, and thus have a rich and highly developed theory [90]. Proposition 2.5 allows us to adapt standard language and constructions for $k[x_1, \ldots, x_n]$-modules to the study of persistence modules, provided those constructions make sense in the n-graded setting. Fortunately, as a rule of thumb, definitions and results about $k[x_1, \ldots, x_n]$-modules do tend to carry over to the n-graded setting, and often become simpler there.

3 1-Parameter Persistent Homology

In this section, we briefly review 1-parameter persistent homology. For a more thorough introduction, see, e.g., [154, 86].

As mentioned the introduction, 1-parameter persistent homology provides invariants of data called *barcodes*. The barcode is a collection of intervals lying in a fixed totally ordered set, e.g., $\mathbb{R}$ or $\mathbb{N}$. Intuitively, intervals in the barcode represent geometric features of the data, and the length of the interval is a measure of size of the corresponding feature. For instance, the circular structure in Fig. 3.2a is represented by the long interval in Fig. 3.2c.

3.1 The Persistent Homology Pipeline

Let T be a totally ordered set. The standard pipeline for constructing barcodes proceeds in three steps, as illustrated by the following diagram:

$$\boxed{\text{Data}} \Rightarrow \boxed{T\text{-Filtration}} \xRightarrow{\text{Homology}} \boxed{T\text{-Module}} \xRightarrow{\text{Decomposition}} \boxed{\text{Barcode}}$$

Depending on the type of data and the information one wants to capture, the construction of a filtration from data can take many forms. The following examples are standard:

Example 3.1. For $X \subset \mathbb{R}^n$, define the *offset filtration* of X to be $\mathcal{O}(X) := \mathcal{S}^{\uparrow}(d_X)$, where $d_X \colon \mathbb{R}^n \to [0, \infty)$ is the distance function to X, i.e.,

$$d_X(y) = \inf_{x \in X} \|y - x\|.$$

Note that $\mathcal{O}(X)_r$ is simply the union of the closed balls of radius r centered at the points of X.

Remark 3.2. The offset filtration is topologically equivalent to two well-known simplicial filtrations, the *Čech filtration*, and a sub-filtration thereof called the *Delaunay filtration* (also known as the *α-filtration*) [84]. The Delaunay filtration is used frequently in persistent homology computations involving points in low-dimensional Euclidian space (particularly in $\mathbb{R}^3$).

Definition 3.3. For a finite metric space X and $r \in [0, \infty)$, let $N(X)_r$ denote the neighborhood graph of X, i.e., the graph with vertex set X containing the edge $[x, y]$ if and only if $d(x, y) \leq r$. Let $\mathrm{Rips}(X)_r$ denote the *clique complex* on $N(X)_r$, i.e., the largest simplicial complex with 1-skeleton $N(X)_r$. These simplicial complexes assemble into a filtration $\mathrm{Rips}(X) \colon [0, \infty) \to \mathbf{Simp}$, called the *(Vietoris-)Rips filtration*.

In view of computational considerations, Rips filtrations are a natural construction when working with high-dimensional or non-Euclidean data. Fig. 3.1 illustrates the Rips filtration of a set of points in the plane.

Figure 3.1: The Rips filtration (Definition 3.3).

By applying the i-th homology functor with coefficients in the field k to either an offset or Rips filtration, one obtains a $[0, \infty)$-persistence module. The following theorem tells us that this module has a well-defined barcode.

Theorem 3.4 (Structure of Persistence Modules [32, 73])**.** *If M is p.f.d. persistence module indexed by a totally ordered set T, then there exists a unique multiset $\mathcal{B}(M)$ of intervals in T, such that*

$$M \cong \bigoplus_{I \in \mathcal{B}(M)} k_I.$$

We call $\mathcal{B}(M)$ the *barcode of M*. For F a filtration such that $H_i(F)$ is p.f.d., we write $\mathcal{B}(H_i(F))$ simply as $\mathcal{B}_i(F)$.

Remark 3.5 (History)**.** Theorem 3.4 was proven in the case $T = \mathbb{Z}$ by Webb in 1985 [179], for all T with a countable dense subset by Crawley-Boevey in 2012 [73], and for all totally ordered sets T by Botnan and Crawley-Boevey in 2018 [32]. The important special case of finitely presented modules [181] is a slight variant of the standard structure theorem for finitely generated modules over a PID, which can be found in many undergraduate abstract algebra textbooks.

Computation   The barcode of a filtered simplicial complex containing n simplices can be computed in $O(n^3)$ using a variant of Gaussian elimination [85]. There has been a great deal of work in the past two decades on optimizing this basic approach, leading to dramatic performance gains for important classes of filtrations [7, 152].

Applications   Persistent homology has been applied in many areas, including computational chemistry [149], materials science [132, 114], neuroscience [170, 108, 162, 105], and bioinformatics [158]. It is useful for exploratory data analysis [114, 59, 46], and is commonly integrated into pipelines for supervised learning [113].

Vectorization of Barcodes   In applications to machine learning or statistics, one often wants to work with features of data taking values in a vector space, since many standard machine learning methods and statistical tests require their input to be of this form. Yet the space of barcodes does not naturally have the structure of a vector space, and indeed its geometry is rather complicated [175, Theorem 2.5]. A simple and practical solution is to fix a map from barcode space to a vector space, and work with the images of barcodes under this map. Many such maps have been proposed and applied to problems in machine learning. See [113] for an overview.

3.2   The Bottleneck Distance

The *bottleneck distance d_B* is the standard metric on barcodes in the persistence theory. Intuitively, for barcodes $\mathcal{C}$ and $\mathcal{D}$, $d_B(\mathcal{C}, \mathcal{D})$ is the maximum distance we must perturb an endpoint on of an interval in $\mathcal{C}$ to transform $\mathcal{C}$ into $\mathcal{D}$. While d_B can be defined between any two barcodes whose intervals lie in $\mathbb{R}$ [11], for simplicity's sake we give the definition only for barcodes with finitely many intervals, each of the form $[a, b)$, where $a < b \in \mathbb{R}$.

Let $\mathcal{C}$ and $\mathcal{D}$ be two such barcodes. We define a *matching between $\mathcal{C}$ and $\mathcal{D}$* to be a collection of pairs $\chi = \{(I, J) \in \mathcal{C} \times \mathcal{D}\}$ where each I and each J appears in at most one pair. If $I \in C \cup \mathcal{D}$ does not appear in a pair, then we say that I is *unmatched*.

Given intervals $I = [a, b)$ and $J = [c, d)$, let

$$c(I, J) = \max(|c - a|, |d - b|), \qquad c(I) = (b - a)/2.$$

For a matching χ between $\mathcal{C}$ and $\mathcal{D}$, we define

$$\text{cost}(\chi) := \max \left(\max_{(I,J) \in \chi} c(I, J), \max_{\substack{I \in \mathcal{C} \cup \mathcal{D} \text{ unmatched}}} c(I) \right). \qquad (3.1)$$

Definition 3.6. The *bottleneck distance* between $\mathcal{C}$ and $\mathcal{D}$ is

$$d_B(\mathcal{C}, \mathcal{D}) = \inf \{\text{cost}(\chi) : \; \chi \text{ is a matching between } \mathcal{C} \text{ and } \mathcal{D}\}.$$

Remark 3.7. The bottleneck distance $d_B(\mathcal{C}, \mathcal{D})$ can be computed in time $O(n^{1.5} \log n)$ where $n = |\mathcal{C}| + |\mathcal{D}|$ [124, 89]. An efficient implementation is available in the Hera software [124].

3.3 Stability of Persistent Homology

The *stability theorem* for persistent homology asserts that the map from data to barcodes is 1-Lipschitz continuous with respect to suitable choices of metrics. This plays an important role in both the statistical foundations of PH [94] and the computational theory [168]. And as we will see, it is also the starting point for some of the central ideas in multiparameter persistence. In fact, there are many variants of the stability theorem in the literature [67, 63, 10, 33, 27]. Here we will state a version of stability for sublevel filtrations, as well as corollaries of this concerning the offset and Rips filtrations. In Section 6.1, we will also discuss a more general algebraic formulation of stability called the *isometry theorem*.

Definition 3.8.

(i) Given $X, Y \subset \mathbb{R}^n$, the *Hausdorff distance* between X and Y is given by

$$d_H(X, Y) = \inf \{r \geq 0 \mid X \subset \cup_{p \in Y} B(p, r), \; Y \subset \cup_{p \in X} B(p, r)\},$$

where $B(p, r)$ denotes the closed ball of radius r centered at p.

(ii) Given metric spaces X, Y, the *Gromov-Hausdorff* distance between X and Y is given by

$$d_{GH}(X, Y) = \inf_{\gamma, \kappa} d_H(\gamma(X), \kappa(Y)),$$

where $\gamma : X \to Z$ and $\kappa : Y \to Z$ range over all isometric embeddings of X and Y into a common metric space Z.

Theorem 3.9 (Stability of Persistent Homology [67, 61, 62, 64])**.**

(i) *For any topological space W, functions $\gamma, \kappa : W \to \mathbb{R}$, and $i \in \mathbb{N}$ such that $H_i(S^{\uparrow}(\gamma))$ and $H_i(S^{\uparrow}(\kappa))$ are p.f.d., we have*

$$d_B(\mathcal{B}_i(S^{\uparrow}(\gamma)), \mathcal{B}_i(S^{\uparrow}(\kappa))) \leq \sup_{w \in W} |\gamma(w) - \kappa(w)|.$$

(ii) *For finite $X, Y \subset \mathbb{R}^n$ and $i \in \mathbb{N}$,*

$$d_B(\mathcal{B}_i(\mathcal{O}(X)), \mathcal{B}_i(\mathcal{O}(Y))) \leq d_H(X, Y).$$

(iii) *For finite metric spaces X, Y and $i \in \mathbb{N}$,*

$$d_B(\mathcal{B}_i(\text{Rips}(X)), \mathcal{B}_i(\text{Rips}(Y))) \leq d_{GH}(X, Y).$$

We will see multiparameter analogues of each of these three stability statements in Section 6.

3.4   Ordinary Persistent Homology is Not Robust

Theorem 3.9 (ii) and (iii) notwithstanding, both the offset and Rips construc-
tions of persistent homology are highly unstable to outliers [29, Section 4], and
both constructions can be insensitive to geometric structure in high-density re-
gions of the data. For example, consider the two data sets shown in Fig. 3.2, and
note that the second differs from the first only by the addition of a few points.
While the H_1 barcode of the first data set has a long interval, representing a
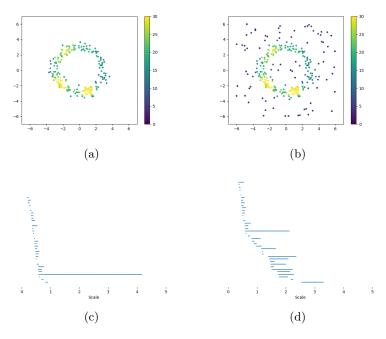loop in the data, this signal is lost in the barcode of the second data set.

Figure 3.2: **(a)** A data set with a circular shape colorized by a local density
estimate. **(b)** The data from (a) with added noise and colorized by a local
density estimate. **(c)** The barcode of the data in (a). **(d)** The barcode of the
data in (b). The barcodes were computed using Ripser [7].

There have been many proposals for dealing with such issues within the
framework of 1-parameter persistence; see [31] for a recent overview. However,
these approaches share important disadvantages, which we now explain, closely
following [31]. First, they require the user to fix one or more parameters, typi-
cally either a spatial scale or a density threshold. The suitability of a particular
parameter choice depends on the data, and it can be unclear how to best choose
such a parameter. Moreover, for data with interesting features across a range
of scales and densities, there may be no single parameter choice that fully cap-
tures the structure of interest in the data. In addition, strategies that fix a scale
parameter do not distinguish small features in the data from large features, and
strategies that fix a density threshold do not distinguish features appearing at
high densities from those appearing at low densities.

A natural solution to these problems is to consider 2-parameter persistent
homology, where one of the parameters is a scale parameter, as in the Rips

filtration, and the other parameter is a density threshold. In fact, as we will see in Section 5.4, there are several natural constructions in this spirit with different advantages and disadvantages.

4 Multiparameter Persistent Homology

To develop multiparameter persistent homology, one would like to generalize the persistent homology pipeline of Section 3.1. The first two arrows of the pipeline generalize without difficulty, as follows:

$$\boxed{\text{Data}} \Rightarrow \boxed{\text{Multifiltration}} \xRightarrow{\text{Homology}} \boxed{\text{Multiparameter Persistence Module}}$$

Indeed, applying homology to each space and each map in a multifiltration yields a multiparameter persistence module, exactly as in the 1-parameter case. Furthermore, there are many natural ways of constructing multifiltrations from data; several of these will be discussed in Section 5. As one example, we have the following density-sensitive extension of the Rips filtration:

Definition 4.1. For X a metric space, $r \geq 0$, and $d > 0$, let $\text{DRips}(X)_{d,r}$ be the maximal subcomplex of $\text{Rips}(X)_r$ whose vertices have degree at least $d - 1$ in the 1-skeleton of $\text{Rips}(X)_r$. Varying r and d, we obtain a bifiltration $\text{DRips}(X)$, the *degree-Rips bifiltration* [134].

Fig. 4.1a illustrates part of the degree-Rips bifiltration of the example from Fig. 3.1.

4.1 The Difficulty of Defining Barcodes of Multiparameter Persistence Modules

We now turn to the salient question: Is there any good way to define the barcode of a multiparameter persistence module? As we will now see, there are several illuminating ways one can approach this question.

We begin with some good news: As in the 1-parameter case, p.f.d. multiparameter persistence modules decompose into indecomposable summands in an essentially unique way. In fact, this is true for persistence modules indexed by any poset:

Theorem 4.2. *For a poset P and persistence module $M \colon P \to \mathbf{Vect}$,*

(i) *If M is pointwise finite-dimensional, then there exists a collection of indecomposables $\{M^\lambda\}_{\lambda \in \Lambda}$ such that*

$$M \cong \bigoplus_{\lambda \in \Lambda} M^\lambda.$$

(ii) *If*

$$M \cong \bigoplus_{\lambda \in \Lambda} M^\lambda \cong \bigoplus_{\gamma \in \Gamma} M^\gamma$$

with each M^λ and M^γ indecomposable, then there exists a bijection $\sigma \colon \Lambda \to \Gamma$ such that $M^\lambda \cong M^{\sigma(\lambda)}$ for all $\lambda \in \Lambda$.

In view of this result, understanding the algebraic structure of p.f.d. persistence modules boils down to understanding the structure of the indecomposable modules.

Remark 4.3. For finite posets, the proof of Theorem 4.2 (i) is elementary and can be found in many introductory texts on quiver representations; see e.g. [5, Chapter 4.1]. In the generality given here, Theorem 4.2 (i) was sketched in work of Gabriel and Roiter in 1992 [99], follows from work of Crawley-Boevey from 1994 [72], and was given a short direct proof in recent work of Crawley-Boevey and Botnan from 2018 [32]. Theorem 4.2 (ii) is the Azumaya–Krull–Remak–Schmidt theorem [4].

Definition 4.4. A P-module M is *interval-decomposable* if there exists a multiset $\mathcal{B}(M)$ of intervals in P such that

$$M \cong \bigoplus_{I \in \mathcal{B}(M)} k_I.$$

We call $\mathcal{B}(M)$ the *barcode* of M.

By virtue of Theorem 4.2 (ii), the barcode of an interval-decomposable is well-defined whenever it exists. However, it turns out that for $n \geq 2$, not all n-parameter persistence modules are interval-decomposable; drawing on classical ideas from quiver representation theory, we shall see in Section 8 that the space of indecomposables that can arise in the multiparameter setting is enormously complex. In particular, there is no way of parameterizing this space by collections of nice regions in the parameter space, as in the 1-parameter setting. Thus, while one could define the barcode of a multiparameter parameter persistence module to be the multiset of isomorphism types of its incomposable summands, this object is generally too complex to work with in practice.

One might nevertheless hope that there is a good way to define the barcode of a multiparameter persistence module $M : P \to \mathbf{Vect}$ as a collection of regions in P. But it turns out not to be possible, in the following sense.

Definition 4.5. A multiset B of subsets of P is a *good barcode of M* if for all $x \leq y \in P$ we have

$$\mathrm{Rk}(M_x \to M_y) = |\{S \in B : x, y \in S\}|,$$

i.e., the rank of the map $M_x \to M_y$ is the number of elements of B containing both x and y.

Given how barcodes of 1-parameter persistence modules are usually interpreted and used in TDA, the goodness condition of Definition 4.5 is quite natural. However, the next example shows that a good barcode of M need not exist.

Example 4.6. Let $P = \{0, 1, 2\} \times \{0, 1, 2\}$ and let M be the following P-module, where $=$ denotes identity maps:

$$
\begin{array}{ccccc}
k & \xrightarrow{\;=\;} & k & \longrightarrow & 0 \\
{\scriptstyle =}\big\uparrow & & {\scriptstyle [1,0]}\big\uparrow & & \big\uparrow \\
k & \xrightarrow{[1,0]^T} & k^2 & \xrightarrow{[1,1]} & k \\
\big\uparrow & & {\scriptstyle [0,1]^T}\big\uparrow & & {\scriptstyle =}\big\uparrow \\
0 & \longrightarrow & k & \xrightarrow{\;=\;} & k
\end{array}
\tag{4.1}
$$

16

A simple argument by contradiction shows that M cannot have a good barcode: If B is a good barcode of M, then since

$$\mathrm{Rk}(M_{(0,1)} \to M_{(2,1)}) = \mathrm{Rk}(M_{(0,1)} \to M_{(1,2)}) = \mathrm{Rk}(M_{(1,0)} \to M_{(2,1)}) = 1,$$

B must contain intervals I, J, K (not necessarily distinct) with

$$(0,1),(2,1) \in I, \qquad (0,1),(1,2) \in J, \qquad (1,0),(2,1) \in K.$$

But since $\mathbf{dim}\, M_{0,1} = \mathbf{dim}\, M_{2,1} = 1$, we then have $I = J = K$, implying that $(1,0),(1,2) \in I$. This contradicts the fact that $\mathrm{Rk}(M_{(1,0)} \to M_{(1,2)}) = 0$.

Perhaps surprisingly, recent work [36] has shown that if we allow the elements of the barcode to be *signed* subsets of P (i.e., to be labeled positive or negative) then it is possible to give a well-behaved definition of the barcode, which encodes ranks in a way analogous to Definition 4.5. We shall discuss this in detail in Section 9.

4.2 Invariants

Invariants of multiparameter persistence modules play a role analogous to barcodes in 1-parameter persistence: They can be used to visualize persistence modules and can be fed as input to machine learning methods and statistical tests. Many invariants of persistence modules have been proposed in the TDA literature, and one can find yet more in the classical literature on commutative algebra and representation theory. The main question for TDA is which such invariants can be useful in the development of data analysis methodology. We are still in the early stages of understanding this.

In this section, we will introduce several simple and well-known invariants of multiparameter persistence modules. Signed barcodes, a class of potentially useful invariants which are equivalent to some of the invariants discussed in this section, will be introduced later, in Section 9.

Three Simple Invariants We begin by introducing three simple invariants which arise frequently in the MPH literature. These are a good starting point for thinking about how to do practical work with MPH. Let P be a poset.

1. The *Hilbert function* of a p.f.d. persistence module $M\colon P \to \mathbf{Vect}$ is the function $P \to \mathbb{N}$ sending z to $\mathbf{dim}\, M_z$.

2. The *rank invariant* of M [47] records the rank of the linear map $M_a \to M_b$ for every $a \le b \in P$. Clearly, this is a refinement of the Hilbert function. When $P = \mathbb{N}$ or $\mathbb{R}$, the rank invariant is *complete*, i.e., it determines the isomorphism type of M; see [47, Theorem 12] and [74]. However, for $P = \mathbb{N}^2$, the rank invariant is incomplete.

3. For $M\colon \mathbb{R}^n \to \mathbf{Vect}$ finitely presented (see Section 7.2) and $z \in \mathbb{R}^n$, the *(multi-graded) Betti numbers of M* at z are natural numbers

$$\beta_0^M(z), \beta_1^M(z), \ldots, \beta_n^M(z).$$

Informally, $\beta_0^M(z)$ and $\beta_1^M(z)$ count the number of generators and relations in M at z, respectively, while for $i > 1$, $\beta_i^M(z)$ counts higher-order relations. More formally, given a minimal free resolution of M

$$0 \to F_n \to \cdots \to F_1 \to F_0 \to M \to 0$$

(see Definition 7.5), $\beta_i^M(z)$ is the number of elements at bigrade r in a basis for F_i [91, 134]. These are standard invariants in commutative algebra [144].

The Fibered Barcode We say an affine line $L \subset \mathbb{R}^n$ is *admissible* if the product partial order on $\mathbb{R}^n$ restricts to a total order on L. Note that the admissible lines in $\mathbb{R}^2$ are those with non-negative slope. The *fibered barcode* of an $\mathbb{R}^n$-indexed module M [54] is the function which maps each admissible line L to the barcode $\mathcal{B}(M \circ L)$. It is equivalent to the rank invariant of M, but more convenient for visualization and stability analysis; see Sections 4.3 and 6.2.

The Generalized Rank Invariant The following extension of the rank invariant was proposed in [127].

Definition 4.7. Let $M \colon P \to \mathbf{Vect}$ be p.f.d. Given an interval $I \subseteq P$, the *generalized rank* of M over I, denoted by $\mathrm{Rk}_I M$, is defined by:

$$\mathrm{Rk}_I M = \mathrm{rank}\left[\varprojlim M|_I \to \varinjlim M|_I \right].$$

Given a collection $\mathcal{I}$ of intervals, the *generalized rank invariant* of M over $\mathcal{I}$ is the function $\mathrm{rank}_{\mathcal{I}} M \colon \mathcal{I} \to \mathbb{N} \cup \{\infty\}$ defined by $\mathrm{rank}_{\mathcal{I}} M(I) = \mathrm{Rk}_I M$.

Remark 4.8. Kim and Memoli [127] proved that for interval-decomposable modules indexed by *essentially finite* posets, the generalized rank invariant is complete. This result was later extended slightly in [36]. Moreover, Kim and Moore [129] showed that the generalized rank invariant determines the Betti numbers of M.

Vectorizations of Persistence Modules As noted in Section 3.1, many applications of 1-parameter persistent homology involve mapping barcodes into linear spaces and then applying standard statistics and machine learning methods. It is natural to pursue the same idea in the multiparameter setting. To this end, several novel maps from the space of persistence modules into linear spaces have been proposed and applied to data [177, 51, 69, 178, 13]. Most such maps proposed so far depend only on the rank invariant, and several depend only on a part of the rank invariant.

Invariants from Metrics A choice of metric d on persistence modules can be used to define invariants. For example, fixing a persistence module N, we may consider the map $M \mapsto d(M, N)$.

For a more interesting example, suppose we are given an $\mathbb{R}$-valued invariant f of multiparameter persistence modules; for example, we may take $f(M)$ to be the sum of all 0th Betti numbers of M. Scolamiero et. al [166] define an

invariant $\hat{f}$, called the *hierarchical stabilization* of f, taking values in the set of decreasing functions from $[0, \infty)$ to $\mathbb{R}$. Specifically, they define

$$\hat{f}(M)(r) := \inf \{f(N) \colon d(N, M) \leq r\}.$$

It is shown that the map $M \mapsto \hat{f}(M)$ is 1-Lipchitz continuous with respect to d and an interleaving metric on the functions. This construction in fact extends immediately from persistence modules to any metric space; a variant was previously considered in the context of mode detection for $\mathbb{R}$-valued functions [12]. While mathematically attractive, computation of hierarchical stabilization in the multiparameter setting can be hard [102].

Other Work on Invariants of MPH We briefly mention a few other works which consider invariants of multiparameter persistence modules in the context of TDA; we make no attempt to be exhaustive. Harrington et al. [111] study several classical invariants from commutative algebra in the multigraded setting, namely, Hilbert series, associated primes, and 0th local cohomology. Extending an idea of Patel in the 1-parameter setting [155], several papers [139, 2, 127] have explored the idea of defining signed barcodes in the multiparameter setting via *Möbius inversion* [161]. This work sets the stage for the results of [36] discussed above on signed decompositions of the rank invariant; see Remark 9.3. Cai et al. [44] introduce a well-behaved and computable barcode, called the *staircase code*, for the 0th homology modules of superlevel-Rips bifiltrations (Definition 5.1).

4.3   Visualization

We visualize the Hilbert function, fibered barcode, and bigraded Betti numbers in two examples using RIVET [134, 79]. Figures 4.1b and 4.2 show (part of) RIVET's visualization of degree-Rips bifiltrations: As a simple example, Fig. 4.1b visualizes the 1st PH (i.e., loop structure) of the degree-Rips bifiltration of the data of Fig. 3.1. Fig. 4.2 visualizes the 0th PH (i.e., cluster structure) of the degree-Rips bifiltration of 1088 pre-aligned HIV-1 genomes from [59], metrized using the Hamming distance.[1]

To explain the figures, first note that the x-axis is mirrored in each figure so that values decrease from left to right. The Hilbert function is represented by greyscale shading: In each figure, the darkness of shading is proportional to the vector space dimension, and the lightest non-white shade represents a value of 1. The bigraded Betti numbers are represented by translucent colored dots whose area is proportional to the value; the 0th, 1st, and 2nd Betti numbers are shown in green, red, and yellow. For the fibered barcode visualization, the query line L is shown in blue, and the barcode is shown in purple, with each interval offset perpendicularly from L.

Like other viruses, HIV has a rich and epidemiologically important subtype structure [112]; Fig. 4.2 indicates that the degree-Rips PH is able to see key aspects of this structure, without any data preprocessing or parameter choices that may bias the results. In contrast, the 1-parameter Rips PH of this data

---

[1]To control the size of this bifiltration, RIVET coarsens it slightly so that all simplices are born on a 250×250 grid. This coarsening is *stable*, i.e., it changes the modules only slightly in the interleaving distance.
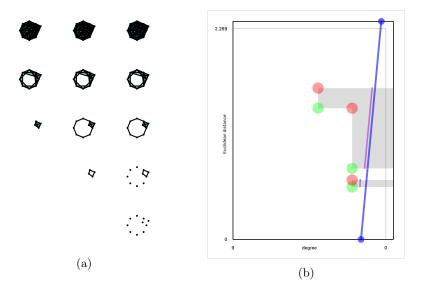
Figure 4.1: (a) The Degree-Rips Bifiltration of the same data set considered in Fig. 3.1, for a few choices of the degree and scale parameter: The left, middle, and right columns correspond to the degree parameters 4, 2, and 0, respectively. (b) The corresponding RIVET visualization in homology dimension 1.
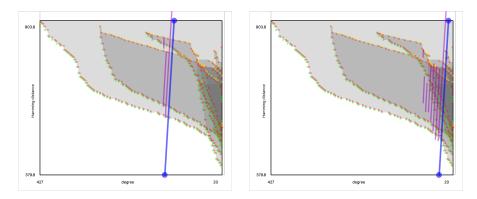
Figure 4.2: RIVET's visualization of the 0th degree-Rips persistent homology of a data set of 1088 HIV-1 genomes, for two different choices of the line L (shown in blue). The visualization indicates the presence of two major clusters in the data (large grey regions to the left), each with several hundred points, as well as 5 smaller clusters of less than 40 points (darker grey regions to the right). Beyond this, the plots of the Hilbert functions and bigraded Betti numbers exhibit interesting geometry which encodes subtle information about the size and shape of the clusters.

(not shown) sees no cluster structure, because of the presence of low density outliers between the clusters. [91, 134].

Remark 4.9. One of the main features of RIVET's visualization is an interactive scheme for visualizing the fibered barcode: The user selects the line L by clicking and dragging the mouse, and the display of the barcode $\mathcal{B}(M \circ L)$ up-

20

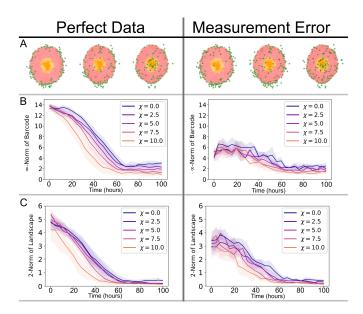Figure 4.3: (A) Cell distributions at three time points from a simulation of immune cells infiltrating a tumor. Green, red, and orange represent macrophages, viable tumor cells, and necrotic tumor cells, respectively. There are ∼100 macrophages and ∼20 tumor cells falsely registered as macrophages. (B) Decay curves tracking the length of the longest bar in the H_1 barcode against time for simulations with different chemotaxis parameter χ. Average curves with standard deviation bands are shown, computed with five simulations for each value of χ. (C) MPH decay curves tracking the 2-norm of the MPH landscape against time. An extended version of this figure can be found in [178, Fig. 1].

dates in realtime. To support this real-time interactivity, RIVET precomputes a data structure called the *augmented arrangement,* which can be queried for the barcode $\mathcal{B}(M \circ L)$ along a generic line L in time $O(|\mathcal{B}(M \circ L)| + \log n)$, where n is the size of a grid containing the supports M [134].

Remark 4.10. A different approach to visualizing the rank invariant arises from the idea of signed barcodes, discussed in Section 9.

4.4 Applications

We briefly discuss two recent applications of MPH to cancer imaging.

Spatial Patterns of Immune Cells in Tumors Vipond et al. [178] use a vectorization of MPH called the *multiparameter persistence landscape* [177] to analyze the spatial patterns of immune cells in cancerous tumors. This work uses RIVET's computational backend, as well as Vipond's code for multiparameter persistence landscapes [176].

Information about the spatial distribution of immune cells in and around a tumor is useful both for prognosis and for guiding treatment [96, 103, 121, 147, 43]. The extent to which immune cells infiltrate the tumor is of particular importance. In [178], the positions of the immune cells are determined from

21

chromogenic immunohistochemistry images of the cancer tissue via a semiauto-mated method.

To quantify the spatial patterns of immune cells, prior approaches have compared the densities of immune cells at the boundary and inner core of the tumor, or used established spatial statistics. However, these approaches capture limited information about the spatial patterns of the cells; one hopes that the geometric information provided by TDA would be of clinical use. A key challenge is that the analysis of spatial distributions of immune cells is complicated by noise in the data: the classification of cells in a tumor image according to cell type is prone to errors, which is expected to cause problems for the standard 1-parameter constructions of persistent homology. It is thus natural to consider MPH in this setting.

In [178], Vipond et al. analyze both synthetic data and real histological data from head and neck tumors. The synthetic data is generated using a 2-dimensional agent-based simulation of macrophages infiltrating a disc-shaped tumor. Initially, the macrophages are distributed along the boundary of the tumor; see Fig. 4.3 (A). As time evolves, the macrophages move inward towards the center of the tumor, driven by a gradient in the concentration of a chemical attractant. The speed of movement is determined by a *chemotaxis* parameter χ. The authors track the dynamics topologically, using Vietoris–Rips persistent homology on the point cloud determined by the locations of the macrophages: the long bar present in the H_1 barcode at time 0 gradually shrinks as time evolves. In order to quantify this, the authors compute a persistent homology *decay curve*, which tracks the length of the longest bar in the barcode as a function of time. It is observed that different choices of χ give rise to distinct decay curves, with larger values of χ yielding steeper curves. However, when a biologically realistic amount of noise is added to the simulation by misclassifying some tumor cells as macrophages, the decay curves change qualitatively, muddying the relationship between χ and the shape of the decay curve; see Fig. 4.3 (B). Such noise causes similar problems for an analogue of the decay curve constructed using an established spatial statistic called the *pair correlation function*. To rectify this, the authors use a *codensity-Rips bifiltration*, a variant of the density-Rips bifiltration introduced in Example 5.2 below, where one filters by both radius and a nearest-neighbors codensity function. The resulting *multiparameter decay curves*, constructed from the 2-norm of the H_1 multiparameter persistence landscape [177], are observed to be more robust to noise than their 1-parameter counterparts; see Fig. 4.3 (C). This finding serves as a proof of concept for the application of multiparameter persistence landscapes to immunohistological data.

The authors next study the spatial patterns of three types of immune cells (namely, the T-cells CD8$^+$ and FoxP3$^+$, and the macrophage CD68$^+$) in images of head and neck tumors. 16 tumors are considered in total, though the key analyses and findings concern smaller subsets of these tumors. For each image and cell type, the cell locations determine a point cloud. The image is decomposed into regions of a suitable fixed size and a MPH landscape is computed for the restriction of the point cloud to each region. This yields an ensemble of landscapes for each tumor and cell type. Each landscape can be loosely interpreted as a measure of the extent to which the immune cells infiltrate the corresponding region; intuitively, regions which have few immune cells lead to features of large persistence.

An analysis of the landscapes of five of the largest tumors provides insight into the patterns of tumor infiltration for each cell type, and into the relationship between these patterns. For instance, the authors observe via PCA and linear discriminant analysis that the landscape ensembles of different cell types have qualitatively distinct distributions, reflecting differences in the extent and pattern of immune cell infiltration [178, Figures 3, S10]. In addition, statistics constructed from the landscapes indicate that FoxP3$^+$ cells usually infiltrate the tumors less than either CD8$^+$ or CD68$^+$ cells [178, Tables S4, S7]. Moreover, it is observed that a landscape-based measure of how FoxP3$^+$ cells infiltrate the tumor is correlated with oxygen levels of the tumor [178, Supporting information, lines 75-85 and Tables S4, S7]. The sample sizes considered are too small to establish the statistical significance or biological reproducibility of these findings, but the results suggest that it may be fruitful to conduct similar analyses with a larger cohort.

While [178] does not make a direct connection between topological invariants and clinically important variables such as disease prognosis, it demonstrates that MPH is a viable tool in the study of this type of data, capable of providing biological insights.

Classification of Breast Cancer Tissue As part of a larger investigation of feature maps for MPH in machine learning, Carrière and Blumberg [51] use MPH in the study of *quantitative immunofluorescence images* of breast cancer tissue. For each tissue sample, one has two images with the same domain; in one, pixel intensity corresponds to the number of immune cells at that pixel and in the other, pixel intensity corresponds to the number of cancer cells; see Fig. 4.4. The objective is to train a classifier that predicts patient survival, using only these images.
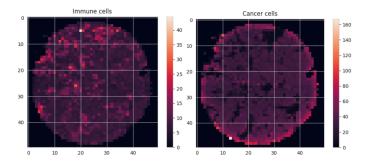
Figure 4.4: Quantiative imunofluoresence images of a breast cancer tissue sample. Left: pixel intensity corresponds to number of immune cells. Right: pixel intensity corresponds to number of cancer cells. The figure is taken from [51].

As explained in [51], the usual approach to this problem begins by thresholding the intensities, in order to identify one subset of the pixels as "immune pixels," and another subset as "cancer pixels." For each cancer pixel the distance to its nearest immune pixel is computed, yielding a multiset of distance values. Symmetrically, for each immune pixel the distance to its nearest cancer pixel is computed, yielding a second multiset of distance values. The means and variances of these two multisets serve as features for classification. How-

ever, this featurization method is unstable, requires manual thresholding, and is insensitive to geometric information at larger spatial scales.

To address these shortcomings, the authors construct a 2-parameter sublevel filtration from the two images associated to each tissue sample. They then compute the H_1 persistent homology module of this filtration, and vectorize this using each of three different MPH feature maps: the *multiparameter persistence image* introduced in the same paper; the multiparameter persistence kernel of [69] (using sliced Wasserstein kernels [52]); and the multiparameter persistence landscape [177]. Several 1-parameter persistence feature maps are also considered; these are obtained by restricting the bipersistence module to a single diagonal line. Comparing the performance of these different feature maps in the classification task, the authors find that the three multiparameter approaches perform similarly to each other, and far better than either the 1-parameter persistence approaches or the usual approach based on nearest neighbor distributions.

5   Multifiltrations

The *degree-Rips* bifiltration of Definition 4.1 is just one of many natural constructions of bifiltrations from data. Here, we introduce several interesting constructions of bifiltrations and trifiltrations that have either already played a significant role in MPH or that we think might in the future.

5.1   Superlevel-Rips Bifiltrations

Definition 5.1 ([47])**.** Consider a finite metric space X and function $\gamma\colon X \to \mathbb{R}$. The *superlevel-Rips bifiltration* $\mathrm{Rips}^{\downarrow}(\gamma)\colon \mathbb{R}^{\mathrm{op}} \times [0, \infty) \to \mathbf{Simp}$ is given by

$$\mathrm{Rips}^{\downarrow}(\gamma)_{a,r} = \mathrm{Rips}(\gamma^{-1}[a, \infty))_r.$$

The sublevel-Rips bifiltration can be defined analogously, and Čech complexes can also be used in place of Rips complexes. In [47], several choices of the function γ are proposed; we mention two of these.

Example 5.2 (Density-Rips bifiltrations)**.** Informally, a *density function* on a metric space X is a function $\gamma\colon X \to [0, \infty)$ whose value is high in dense regions of the data and low near sparse regions of the data. Standard examples include kernel density functions, e.g., with a Gaussian or disk kernel, and nearest neighbor density functions.

Given a density function $\gamma\colon X \to [0, \infty)$, we call the superlevel-Rips bifiltration $\mathrm{Rips}^{\downarrow}(\gamma)$ a *density-Rips* bifiltration. This is a practical, computationally tractable choice of density-sensitive bifiltration on a finite metric space. We have already discussed an application to cancer imaging in Section 4.4.

Example 5.3 (Eccentricity-Rips bifiltrations)**.** Define $\gamma\colon X \to [0, \infty)$ by

$$\gamma(x) = \frac{1}{|X|} \sum_{y \in X} d(x, y),$$

i.e., $\gamma(x)$ is the average distance of x to points in X. We call γ an *eccentricity* function, and $\mathrm{Rips}^{\downarrow}(\gamma)$ the *eccentricity-Rips* bifiltration of X. When X has

the structure of multiple tendrils emanating radially from a central core, the superlevel sets of γ decompose into clusters in corresponding to the tendrils of X, which can be detected as persistent features in $H_0(\mathrm{Rips}^{\downarrow}(\gamma))$.

Example 5.4. In some applications, the data set X comes equipped with a function $\gamma\colon X \to \mathbb{R}$, and we can consider the superlevel-Rips bifiltration of this. For one example, in the study of time-varying data, $\gamma(x)$ may be the time of appearance of a data point. As a second example, in applications to TDA computational chemistry, X is taken to be the atom centers of a ligand (drug candidate), with the Euclidean distance and γ is taken to be the partial charge function [122].

5.2   Interlevel-Rips Trifiltrations

In defining the superlevel-Rips bifiltration, we have made a choice of filtration direction: we have chosen to filter X by the superlevel sets of γ rather than by sublevel sets. This is natural for some examples of γ, e.g., those of Example 5.2 and Example 5.3. But for other examples of γ, such as those of Example 5.4, imposing such a choice of direction can be unnatural. For such examples, it is arguably more natural to consider a direction-agnostic variant of the construction, which we will call the *interlevel-Rips trifiltration* [101]. The idea is to filter X not by superlevel sets, but by *interlevel-sets*, i.e., the inverse images of intervals. In the following $\overline{U}$ denotes the subposet of $\mathbb{R}^{\mathrm{op}} \times \mathbb{R}$ given by $\overline{U} = \{(a,b) : a \leq b\}$.

Definition 5.5. Consider a finite metric space X and function $\gamma\colon X \to \mathbb{R}$. The *interlevel-Rips trifiltration* $\overline{U} \times [0,\infty) \to \mathbf{Simp}$ is defined by

$$\mathrm{Rips}(\gamma)_{(a,b),r} = \mathrm{Rips}(\gamma^{-1}[a,b])_r.$$

Observe that for γ bounded, $\mathrm{Rips}(\gamma)$ is a refinement of $\mathrm{Rips}^{\downarrow}(\gamma)$.

In Section 10.2, we will consider a similar construction, the *interlevel bifiltration*, which is defined for any real-valued function on a topological space. In the case where one has a 1-parameter family of metrics $d = \{d_t\}_{t \in [0,1]}$ on a fixed finite set S, Mémoli and Kim give a different construction of an interlevel-Rips trifiltration [128], which is perhaps more natural in that setting. Such 1-parameter families arise in the study of collective motion of animals, e.g., flocking and swarming [19, 150, 20, 174], and also in the topological study of time series via sliding window embeddings [157].

5.3   Multifiltrations from Images

Many applications of 1-parameter persistent homology concern image analysis, where sublevel filtrations are often used. There is not yet a consensus on what the most natural or useful multifiltrations are for image analysis, but one promising idea is that a second persistence parameter can be used to thicken sublevel or superlevel sets, thereby introducing some sensitivity to the *width* of features that the ordinary sublevel and superlevel filtrations lack.

One construction along these lines is due to Chung, Day, and Hu [65]: For (X,d) a metric space and a function $f\colon X \to \mathbb{R}$, define a bifiltration $F\colon \mathbb{R}^2 \to$

Top by

$$F_{a,t} = \begin{cases} \{y \in X : d\left(y, f^{-1}[a, \infty)\right) \le t\} & t \ge 0, \\ \{y \in f^{-1}[a, \infty) : d\left(y, X \setminus f^{-1}[a, \infty)\right) \ge -t\} & t < 0. \end{cases}$$

By considering interlevel sets rather than superlevel sets, one can also define a more symmetric 3-parameter analogue of this filtration.

Intuitively, these multifiltrations should, in some sense, exhibit better robustness to "spike noise" (i.e., noise of large magnitude but small width) than the ordinary sublevel or interlevel filtrations, but to the best of our knowledge there is no theoretical result along these lines.

5.4 Parameter-Free, Density-Sensitive Multifiltrations

One important drawback of the density-Rips bifiltration introduced above is that, as a rule, the density function γ depends on choice of *bandwidth parameter*, e.g., the variance of the Gaussian kernel, or the number of nearest neighbors. As suggested in Section 3.4, it is generally preferable to work with constructions which do not require us to choose parameters.

The degree-Rips bifiltration of Definition 4.1 is one natural choice of density-sensitive bifiltration whose construction requires no parameters. Its simplicity is appealing, and as we discuss below, it has the advantage of being computable. However, there are other natural alternatives, some of which have superior robustness properties (Theorem 6.16). We now discuss several of these.

The Multicover Bifiltration   Given $X \subset \mathbb{R}^n$, let μ_X denote the counting measure of X.

Definition 5.6. Given any metric space X and measure μ on X, the *measure bifiltration* of μ is the $(0, \infty)^{\mathrm{op}} \times [0, \infty)$-indexed bifiltration $\mathcal{M}(\mu)$ defined as

$$\mathcal{M}(\mu)_{m,r} = \{x \in X : \mu(B(x, r)) \ge m\},$$

where as in Definition 3.8, $B(y, r)$ denotes the closed ball in X of radius r centered at x. For $X \subset \mathbb{R}^n$, the bifiltration $\mathcal{M}(\mu_X)$ is called the *multicover bifiltration* [60, 167].

Fixing $m = 1$, we recover the offset filtration of Example 3.1; the multicover bifiltration can thus be seen as a density-sensitive refinement of this.

Subdivision Bifiltrations   Let bary(T) denote the barycentric subdivision of a simplicial complex T, i.e., the abstract simplicial complex whose simplices are sets $\{\sigma_1, \ldots, \sigma_l\}$ of simplices in T such that $\sigma_i \subset \sigma_{i+1}$ for $\in 1, \ldots, l-1$. Note that vertices of bary(T) correspond to simplices of T. For $m \in (0, \infty)$, let $\mathcal{S}(T)_m$ denote the maximal subcomplex of bary(T) whose vertices correspond to simplices in T of dimension at least $m-1$. These subcomplexes assemble into a $(0, \infty)^{\mathrm{op}}$-filtration $\mathcal{S}(T)$ of bary(T). Moreover, for any simplicial filtration F, applying this construction at each index yields a bifiltration.

Definition 5.7 ([167])**.** When the above construction is applied to the filtration $F = \mathrm{Rips}(X)$, we call the resulting $(0, \infty)^{\mathrm{op}} \times [0, \infty)$-indexed bifiltration the *subdivision-Rips bifiltration* of X, and denote it as $\mathrm{SRips}(X)$. A subdivision-Čech bifiltration can be defined analogously.

Vertices of $\mathrm{SRips}(X)_{m,r}$ correspond to cliques in $\mathrm{Rips}(X)_r$ of size at least m, which for large m represent dense regions in X at scale r. In this sense, this bifiltration is indeed density sensitive.

It follows from work of Sheehy and Cavanna [53, 167] that slight variants of the subdivision-Čech and the multicover bifiltrations are topologically equivalent (more precisely, weakly equivalent [30]) and therefore have isomorphic persistent homology modules; see also [31, Section 4]. However, both the Rips and Čech subdivision filtrations have exponentially many vertices with respect to $|X|$, and are therefore too large to handle directly in practical computations.

The Rhomboid Bifiltration For $X \subset \mathbb{R}^n$ finite, Corbet et al. [71] showed that a polyhedral bifiltration in $\mathbb{R}^{n+1}$ called the *rhomboid bifiltration*, introduced by Edelsbrunner and Osang [88], is topologically equivalent to the multicover bifiltration of X. The rhomboid bifiltration is a density-sensitive extension of the Delaunay filtration. We refer the reader to [71, 88] for the definition. The rhomboid bifiltration has size $O(|X|^{n+1})$.

The Kernel Bifiltration Given a *kernel function* $K \colon [0, \infty) \to [0, \infty)$ satisifying mild conditions, Rolle and Scoccola extend the definition of the associated density-Rips bifiltration to obtain a 3-parameter filtration, essentially by considering all possible values of the kernel bandwidth [160]. The construction thus no longer depends on a persistence parameter. For a suitable choice of K, this trifiltration also extends the degree-Rips bifiltration.

5.5   Computation of Density-Sensitive Bifiltrations

For a simplicial multifiltration F indexed by P and $a \in P$, let $\mathrm{births}(a)$ denote the number of simplices in F_a not present in F_b for any $b < a$. Define the *size* of F to be $\sum_{a \in P} \mathrm{births}(a)$.

For fixed K, the K-dimensional skeleton of the degree-Rips bifiltration of a metric space X has size $O(|X|^{K+2})$. However, if the bifiltration is coarsened to lie on a grid of constant size, then the K-skeleton has size $O(|X|^{K+1})$, which agrees (asymptotically) with that of the ordinary Rips filtration. Using a line sweep algorithm designed by Roy Zhao and implemented in RIVET, these low-dimensional skeleta are readily computed in practice [136].

Edelsbrunner and Osang [87] recently introduced a clever and relatively simple algorithm for computing the rhomboid bifiltration. The algorithm and its complexity analysis depend on a choice of algorithm for computing weighted Delaunay bifiltrations. For a suitable such choice, the algorithm computes the rhomboid bifiltration of $X \subset \mathbb{R}^3$ in time $O(|X|^5)$. An implementation of the algorithm is available [151], and can compute the full rhomboid bifiltration of at least 200 points in $\mathbb{R}^3$; examples and timing results appeared in [71] and [87]. This brings us close to practical computation of multicover PH for low-dimensional data sets, though it seems that further algorithmic work is needed to fully realize the practical potential of these ideas.

The problem of computing the homology of the subdivision-Rips bifiltration is not yet understood. As indicated above, the subdivision-Rips bifiltration is too large to be directly computed, but it may be that a smaller bifiltration has (exactly or approximately) the same homotopy type.

6 Metrics and Stability

There have been many proposals for metrics on multiparameter persistence modules. In this section we discuss a few of these and use them to formulate stability theorems. We focus primarily on the interleaving and matching distances, but also briefly discuss several other distances.

6.1 The Interleaving Distance

The interleaving distance [61, 133, 63], an algebraic generalization of the bottleneck distance, is the best known and most thoroughly studied distance on multiparameter persistence modules. As we will see, it is a natural choice of distance for formulating stability and inference results for multiparameter persistent homology.

For intuition, let us first consider a simple case: A *1-interleaving* between $\mathbb{Z}$-persistence modules M, N is a commutative diagram of vector spaces of the following form, extending M and N:

$$\cdots \longrightarrow M_{-2} \longrightarrow M_{-1} \longrightarrow M_0 \longrightarrow M_1 \longrightarrow M_2 \longrightarrow \cdots$$

$$\cdots \longrightarrow N_{-2} \longrightarrow N_{-1} \longrightarrow N_0 \longrightarrow N_1 \longrightarrow N_2 \longrightarrow \cdots$$

Similarly, a *2-interleaving* between M and N is a commutative diagram

$$\cdots \longrightarrow M_{-2} \longrightarrow M_{-1} \longrightarrow M_0 \longrightarrow M_1 \longrightarrow M_2 \longrightarrow \cdots$$

$$\cdots \longrightarrow N_{-2} \longrightarrow N_{-1} \longrightarrow N_0 \longrightarrow N_1 \longrightarrow N_2 \longrightarrow \cdots$$

For any integer $p \geq 0$, we can define a p-interleaving between M and N analogously, taking the arrows between M and N to increase indices by p. Moreover, the definition extends to $\mathbb{Z}^n$, by taking the arrows to increase indices by the vector $\vec{p} := (p, \ldots, p)$.

We can also define interleavings between $\mathbb{R}^n$-indexed modules in essentially the same way. We now give the formal definition. For $u \in [0, \infty)^n$, define the *u-shift* $(-)(u)$ to be the endofunctor on $\mathbb{R}^n$-modules given on objects by $M(u)_a = M_{u+a}$ and $M(u)_{a,b} = M_{a+u,b+u}$, and on morphisms by $f(u)_a = f_{u+a}$. Furthermore, let $\phi^u_M \colon M \to M(u)$ be the morphism whose restriction to each M_a is the linear map $M_{a,a+u} \colon M_a \to M_{a+u}$.

Definition 6.1. Given $\epsilon \in [0, \infty)$, we say that persistence modules $M, N \colon \mathbb{R}^n \to$ **Vect** are *ϵ-interleaved* if there exist morphisms

$$f \colon M \to N(\vec{\epsilon}) \qquad g \colon N \to M(\vec{\epsilon})$$

such that

$$g(\vec{\epsilon}) \circ f = \phi^{2\vec{\epsilon}}_M \qquad f(\vec{\epsilon}) \circ g = \phi^{2\vec{\epsilon}}_N.$$

The *interleaving distance* is defined by

$$d_I(M, N) = \inf \left\{ \epsilon \geq 0 \colon M \text{ and } N \text{ are } \epsilon\text{-interleaved.} \right\}.$$

Remark 6.2. Clearly, Definition 6.1 generalizes to functors $\mathbb{R}^n \to \mathbf{C}$, for any category $\mathbf{C}$. The cases $\mathbf{C} = \mathbf{Top}$ and $\mathbf{C} = \mathbf{Set}$ play important roles in TDA.

The following theorem tells us that the interleaving distance is an extension of the bottleneck distance to multiparameter persistence modules.

Theorem 6.3 (The isometry theorem). *For any p.f.d. persistence modules* $M, N \colon \mathbb{R} \to \mathbf{Vect}$,
$$d_I(M, N) = d_B(\mathcal{B}(M), \mathcal{B}(N)).$$

The inequality $d_I(M, N) \geq d_B(\mathcal{B}(M), \mathcal{B}(N))$, known as the *algebraic stability theorem*, was first proven by Chazal et al. [61, 63]. Other proofs have appeared in [10, 11, 23]. The converse inequality first appeared in [133].

Multiparameter Algebraic Stability The bottleneck distance d_B on intervals in $\mathbb{R}$ admits a natural extension to collections of intervals in $\mathbb{R}^n$. It is convenient to define this via the interleaving distance:

Definition 6.4 (Bottleneck Distance on Barcodes in $\mathbb{R}^n$ [33]). For a matching χ between barcodes $\mathcal{C}$ and $\mathcal{D}$ in $\mathbb{R}^n$, let

$$\mathrm{cost}(\chi) := \max \left(\max_{(X,Y) \in \chi} d_I(k_X, k_Y), \ \max_{X \in \mathcal{C} \cup \mathcal{D} \ \mathrm{unmatched}} d_I(k_X, 0) \right). \quad (6.1)$$

We then define the $d_B(\mathcal{C}, \mathcal{D})$ to be the infimal cost of a matching between $\mathcal{C}$ and $\mathcal{D}$, exactly as in Definition 3.6.

For certain types of intervals arising naturally in TDA, e.g., those considered in Theorem 6.5 below, one has simple, explicit formulae for the interleaving distance between interval modules. On barcodes consisting of such intervals, these formulae yield simpler formulations of Definition 6.4 [33, 23, 37].

Given that d_I and d_B are both defined the n-parameter setting, one might hope that the isometry theorem extends to interval-decomposable $\mathbb{R}^n$-persistence modules. The following summarizes what is known about this:

Theorem 6.5 ([33, 23, 37]). *Fix* $n \geq 2$, *and assume that* M *and* N *are finitely presented, interval-decomposable* $\mathbb{R}^n$*-modules.*

(i) *There is no constant* c *such that*
$$d_B(\mathcal{B}(M), \mathcal{B}(N)) \leq c \, d_I(M, N)$$
for all such M *and* N.

(ii) *If all intervals in* $\mathcal{B}(M)$ *and* $\mathcal{B}(N)$ *are* blocks *(Definition 10.3), then*
$$d_I(M, N) = d_B(\mathcal{B}(M), \mathcal{B}(N)).$$

(iii) *If* M *and* N *are free (Definition 7.1), then*
$$d_I(M, N) \leq d_B(\mathcal{B}(M), \mathcal{B}(N)) \leq (n-1)d_I(M, N).$$

These bounds are tight for $n = 2$ *and* $n = 4$.

(iv) If all intervals in $\mathcal{B}(M)$ and $\mathcal{B}(N)$ are rectangles (Definition 9.1) or all are hooks (Definition 9.9), then

$$d_I(M, N) \leq d_B(\mathcal{B}(M), \mathcal{B}(N)) \leq (2n - 1)d_I(M, N).$$

These bounds are tight for $n = 2$.

Universality

Definition 6.6. We will say that a distance d (i.e., extended pseudometric) on multiparameter persistence modules is *stable* if for all topological spaces W, functions $\gamma, \kappa : W \to \mathbb{R}^n$, and $i \geq 0$, we have

$$d(H_i(\mathcal{S}^\uparrow(\gamma)), H_i(\mathcal{S}^\uparrow(\kappa))) \leq \sup_{w \in W} \|\gamma(w) - \kappa(w)\|_\infty.$$

Theorem 6.7 (Universality of the Interleaving Distance [133])**.**

(i) *The interleaving distance d_I is stable.*

(ii) *Assume that the field k is prime, i.e., $k = \mathbb{Q}$ or $k = \mathbb{Z}/p\mathbb{Z}$ for some prime p. Then for any other stable distance d on multiparameter persistence modules, we have $d \leq d_I$.*

Remarks 6.8.

1. A version of Theorem 6.7 for the special case of 1-parameter persistence and 0th homology first appeared in work by d'Amico et al. [76].

2. The proof of Theorem 6.7 (i) turns out to be trivial, but the proof of Theorem 6.7 (ii) is not.

3. In view of Theorem 6.3, Theorem 6.7 (i) extends Theorem 3.9 (i).

4. The generalization of Theorem 6.7 to arbitrary fields is an open question.

5. A similar universality result is given for filtrations in [30], using the *homotopy interleaving distance*.

6.1.1   Computation

While the interleaving distance has good theoretical properties, it turns out to be hard to compute, even for very simple bipersistence modules.

Theorem 6.9 ([25])**.** *For k a finite field, it is NP-hard to approximate the interleaving distance on 2-parameter persistence modules within a factor of 3.*

Theorem 6.9 is proved by considering interleavings of *staircase*-decomposable modules, where a staircase is a particularly simple type of interval. That said, the hardness result still holds if all modules are assumed to be indecomposable.

The proof of Theorem 6.9 is a reduction from the following problem [24], which is shown to be NP-complete in [25]: let A and B be $n \times n$ matrices of distinct variables and suppose that we assign the value 0 to each variable in a given subset of the variables. Is it possible to assign a value in k to each of remaining variables such that the resulting product AB is the identity matrix I_n?

Example 6.10. Let A and B be 3×3 matrices with entries set to 0 as follows

$$A = \begin{bmatrix} * & * & * \\ * & 0 & * \\ * & * & 0 \end{bmatrix} \qquad B = \begin{bmatrix} * & * & * \\ * & * & 0 \\ * & 0 & * \end{bmatrix}$$

In this case the decision problem has a positive answer:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = I_3$$

6.2   The Matching Distance

Theorem 6.9 motivates the search for a more computable surrogate for the interleaving distance. The matching distance [54] has emerged as a popular choice. It is defined by restricting the modules to (suitably parameterized) affine lines with positive slope, and taking bottleneck distances:

Definition 6.11. The **matching distance** between p.f.d. $\mathbb{R}^n$-indexed modules M and N is given by:

$$d_{\mathrm{match}}(M, N) = \sup_L \ d_B(\mathcal{B}(M \circ L), \mathcal{B}(N \circ L)),$$

where $L : \mathbb{R} \to \mathbb{R}^n$ ranges over parameterized lines of the form $L(t) = vt + b$ where $v \in [1, \infty)^n$ and $b \in \mathbb{R}^n$.

Note that $d_M(M, N)$ depends only on the fibered barcodes of M and N. The following is an easy consequence of the algebraic stability theorem:

Proposition 6.12 ([131, 54]). *For all p.f.d. $\mathbb{R}^n$-persistence modules M and N,*

$$d_{\mathrm{match}}(M, N) \leq d_I(M, N).$$

6.2.1   Computation

In 2011, Biasotti et al. [22] gave a quad tree-based algorithm to approximate the matching distance between bipersistence modules. More recently, Kerber and Nigmetov [125] revisited their approach, arriving at a more efficient version, and released an implementation as part of the Hera software package.

The matching distance on bipersistence modules can in fact be exactly computed in polynomial time [123, 26]; the following describes the state of the art.

Theorem 6.13 ([26]). *Given minimal free presentations of $\mathbb{R}^2$-indexed persistence modules M and N (see Section 7.2), $d_{\mathrm{match}}(M, N)$ can be computed deterministically in time $O(n^6 \log n)$ and via randomization in time $O(n^5 \log^3 n)$, where n denotes the total number of generators and relations in the two presentations.*

There is no literature on computing the matching distance between higher-parameter persistence modules, but it seems likely that the known approaches in the two-parameter case can be extended.

6.3 Stability of Density Sensitive Bifiltrations

The stability and robustness of the multicover, subdivision, and degree-Rips bifiltrations were recently studied in [31], using the *Prohorov* and *Gromov-Prohorov* distances on measures and interleavings on persistence modules. The Prohorov distance d_{Pr} is a classical distance between measures on the same metric space, which can be thought of as a measure-theoretic analogue of the Hausdorff distance; the *Gromov-Prohorov distance* d_{GPr} is an extension of d_{Pr} to measures on different metric spaces [110, 117]:

Definition 6.14. The *Prohorov distance* between measures μ and η on a metric space (Z, d_Z) is given by

$$d_{Pr}(\mu, \eta) = \sup_A \inf\{\delta \geq 0 : \mu(A) \leq \eta(A^\delta) + \delta \text{ and } \eta(A) \leq \mu(A^\delta) + \delta\},$$

where $A \subset Z$ ranges over all closed sets and

$$A^\delta = \{y \in Z : d_Z(y, a) < \delta \text{ for some } a \in A.\}$$

Definition 6.15 ([110])**.** The *Gromov-Prohorov* distance between measures μ_X and μ_Y on metric spaces X and Y is given by

$$d_{GPr}(\mu_X, \mu_Y) = \inf_{\varphi, \psi} d_{Pr}(\varphi_*(\mu_X), \psi_*(\mu_Y)),$$

where $\varphi \colon X \to Z$ and $\psi \colon Y \to Z$ range over all isometric embeddings into a common metric space Z.

A important property of these distances is that they are stable to the addition of outliers, in a strong sense [31, Remark 2.17].

Recall that for $X \subset \mathbb{R}^n$, we let μ_X denote the counting measure of X, seen as a measure on $\mathbb{R}^n$. Similarly, for X a finite metric space, we let η_X denote the counting measure of X, seen as a measure on X.

Theorem 6.16 (Stability of the Multicover and Subdivision Bifiltrations [31])**.**

(i) For all $X, Y \subset \mathbb{R}^n$ finite and $i \in \mathbb{N}$,

$$d_I(H_i(\mathcal{M}(\mu_X)), H_i(\mathcal{M}(\mu_Y))) \leq d_{Pr}(\mu_X, \mu_Y).$$

(ii) For all finite metric spaces X, Y and $i \in \mathbb{N}$,

$$d_I(H_i(\mathrm{SRips}(X)), H_i(\mathrm{SRips}(Y))) \leq d_{GPr}(\eta_X, \eta_Y).$$

Theorem 6.16 is a measure-theoretic analogue of items (ii) and (iii) of Theorem 3.9, and is proven using a closely analogous argument. A variant of this result using normalized counting measures also holds, and implies a version which uses the Wasserstein distance on the measures in place of the Prohorov distance.

In the same paper, an analogue of these results for degree-Rips bifiltrations is also given [31, Theorem 1.6], using the Gromov-Prohorov distance and a generalized definition of interleavings. This turns out to be a qualitatively weaker

kind of stability than the 1-Lipschitz stability of Theorem 6.16, but is tight (in a reasonable sense). The upshot is that, as measured by interleavings, the degree-Rips bifiltration is qualitatively less robust to outliers than the subdivision-Rips bifiltration, but qualitatively more robust to outliers than the ordinary Rips filtration.

In a similar vein, Rolle and Scocolla have given a Lipschitz stability result for the kernel trifiltration (see Section 5.4) using the *Gromov-Hausdorff-Prohorov distance,* an upper bound for the Gromov-Prohorov which is density-sensitive but unstable to the addition of outliers [160]. This result implies a corresponding Lipschitz stability result for the degree-Rips bifiltration. These ideas are applied in [160] to obtain a stable and consistent clustering scheme based on 3-parameter persistence.

Remark 6.17 (The Tension Between Robustness and Computability)**.** Together, the stability results discussed above and the discussion of computation in Section 5.4 reveal a tension between robustness and computability for the 2-parameter PH of metric data: The multicover and subdivision-Rips bifiltrations have excellent robustness properties, but we do not yet have a practical computational framework for handling either of them. (However, for the multicover bifiltration of low dimensional data, we are already close.) The degree-Rips bifiltrations, on the other hand, are computable, but the present theory guarantees robustness only in a far weaker sense.

This raises the following critical question: Can we develop a framework for MPH of metric data that is both computationally efficient and provably robust?

6.4 ℓ^p-Metrics on Multiparameter Persistence Modules

The interleaving and matching distances are both ℓ^∞-distances, i.e., they can be defined in terms of ℓ^∞-metrics on Euclidean spaces [27]. As such, they are insensitive to certain small scale differences between modules, which can be undesirable in both theory and applications. To address this, it is natural to consider ℓ^p-metrics on multiparameter persistence modules. Several recent works have explored the question of defining such metrics [171, 27, 39, 107, 173].

In the 1-parameter case, there is a standard ℓ^p-generalization of the bottleneck distance, called the *p-Wasserstein distance* and denoted W_p^q; here $q \in [0, 1]$ is a parameter specifying an ℓ^q-metric on $\mathbb{R}^2$ used in the definition of the distance. W_∞^∞ is equal to the bottleneck distance, and varying q changes the distance W_p^q by at most a factor of 2. The distances W_p^q are used in many applications and in some TDA theory; see [27, Section 1].

The problem of extending the Wasserstein distance to multiparameter persistence modules was first considered by Bubenik, Scott, and Stanley, who generalized W_p^1 to persistence modules indexed by an arbitrary abelian category [39]. However, in the case $p = \infty$, their distance is equal to neither the interleaving distance nor the matching distance, and has different qualitative properties.

More recently, Bjerkevik and Lesnick [27] introduced ℓ^p-extensions of the interleaving and matching distance to multiparameter persistence modules, called the *p-presentation distance* d_I^p and the *p-matching distance* d_{match}^p, respectively. In the 1-parameter case, these distances are equal to W_p^p, and they share several of the desirable properties of the interleaving and matching distances discussed earlier in this section. Most notably, on n-parameter persistence modules with

$n \in \{1, 2\}$, d_I^p satisfies a universal property closely analogous to Theorem 6.7. This result extends a fundamental ℓ^p-stability result for the 1-parameter persistent homology of cellular filtrations, due to Skraba and Turner [171, Theorem 4.7].

6.5 Other Metrics

Besides the works referenced above, several other papers have introduced metrics for multiparameter persistence and developed theory for these; we briefly mention three: Scolamiero et. al [166] introduce the formalism of *noise systems* and use this to define a family of metrics on multiparameter persistence modules. Cerri et al. [56, 55] study a variant of the matching distance requiring the matchings along different affine lines to be chosen coherently, and they prove a stability result for this. McCleary and Patel [139] use Möbius inversion to define a distance on the homology modules of certain lattice-indexed filtrations, and they establish functoriality and stability results for their approach.

7 Minimal Presentations and Resolutions

In this section, we introduce minimal presentations and resolutions of persistence modules. We then discuss the problem of computing minimal presentations, focusing on recent progress in the 2-parameter case.

Minimal presentations are particularly important in computation because they provide a memory-efficient representation of a persistence module, which then can be used as input to algorithms for computing invariants or distances. Minimal presentations of persistent homology modules are often quite small in practice, in comparison to the filtrations from which they arise [126], which makes them convenient for computation. However, minimal presentations are usually not unique, so they cannot be directly used in TDA the ways barcodes are used, e.g., as input to machine learning algorithms or statistical tests.

7.1 Free Modules

Let $P = T_1 \times \cdots \times T_n$, where each T_i is a totally ordered set, and for $z \in P$, let $\langle z \rangle$ denote the interval $\{p \in P \colon p \geq z\}$.

Definition 7.1. A P-persistence module F is *free* if it is interval-decomposable and all intervals in the barcode $\mathcal{B}(F)$ are of the form $\langle z \rangle$.

For F free, we let $|F| = |\mathcal{B}(F)|$, where $|\cdot|$ denotes the cardinality of a multiset.

Given a P-persistence module M and $v \in M_z$, we write $\mathrm{gr}(v) = z$. We say that $S \subset \bigcup_{z \in P} M_z$ is *a set of generators* for M if for any $v \in \bigcup_{z \in P} M_z$,

$$v = \sum_{i=1}^{m} c_i M_{\mathrm{gr}(v_i), \mathrm{gr}(v)}(v_i)$$

for some $v_1, v_2, \ldots, v_m \in S$ and scalars $c_1, \ldots, c_m \in k$. We say M is *finitely generated* if there exists a finite set of generators for M.

A *basis* of a free module F is a minimal generating set. It can be shown via elementary linear algebra that if B is any basis for a free module F, then for all $z \in P$, the number of elements in B of grade z is equal to the number of copies of $\langle z \rangle$ in $\mathcal{B}(F)$.

Matrix Representation of Morphisms of Free Modules Let $\gamma \colon F' \to F$ be a morphism of finitely generated free modules, and let $B' = \{B_i'\}_{i=1}^{d'}$ and $B = \{B_i\}_{i=1}^{d}$ be ordered bases of F' and F, respectively. In analogy with ordinary linear algebra, we can represent γ with respect to these bases via a matrix $[\gamma]$ with coefficients in the field k, together with a P-valued label for each row and each column of the matrix. To explain the details, for $z \in P$, we represent $v \in F_z$ with respect to B as a vector $[v]^B \in k^{|B|}$, by taking $[v]^B$ to be the unique vector such that $[v]_i^B = 0$ if $\mathrm{gr}(B_i) \not\leq z$ and

$$v = \sum_{i \colon \ \mathrm{gr}(B_i) \leq z} [v]_i^B F_{\mathrm{gr}(B_i), z}(B_i).$$

Thus, $[v]^B$ records the field coefficients in the linear combination of B giving v. We now define $[\gamma]$ as follows:

- the jth column is $[\gamma(B_j')]^B$,

- the label of the jth column is $\mathrm{gr}(B_j')$,

- the label of the ith row is $\mathrm{gr}(B_i)$.

In the literature on multi-graded commutative algebra, the matrix $[\gamma]$ is called a *monomial matrix* [144].

7.2   Free Presentations and Resolutions

Definition 7.2.

(i) A *(free) presentation* $\mathcal{F}$ of a persistence module M is a morphism of free modules

$$F_1 \xrightarrow{\phi_1} F_0 \tag{7.1}$$

with $\operatorname{coker} \phi_1 \cong M$.

(ii) A *(free) resolution* $\mathcal{F}$ of M is an exact sequence of free modules

$$\cdots \xrightarrow{\phi_3} F_2 \xrightarrow{\phi_2} F_1 \xrightarrow{\phi_1} F_0$$

with $\operatorname{coker} \phi_1 \cong M$.

M is said to be *finitely presented* if there exists a presentation $\mathcal{F}$ of M with F_0 and F_1 finitely generated.

Remark 7.3. In view of the discussion Section 7.1, a free presentation $\gamma \colon F_0 \to F_1$ with each F_i finitely generated can be represented with respect to a choice of bases for F_0 and F_1 as a labeled matrix; we call this a *presentation matrix*, or sometimes (by a slight abuse of terminology) simply a *presentation*. Similarly, a resolution can be represented as a sequence of matrices.

Of particular interest are minimal presentations and resolutions. Several equivalent definitions can be given; the following approach is perhaps the most transparent.

Definition 7.4.

(i) A *trivial resolution* is a direct sum of free resolutions the form

$$\cdots 0 \to 0 \to F \xrightarrow{\mathrm{Id}_F} F \to 0 \to 0 \to \cdots \to 0.$$

where the two copies of F may appear at any two consecutive indices.

(ii) A *trivial presentation* is a free presentation of the form

$$F \oplus F' \xrightarrow{\mathrm{Id}_F \oplus 0} F.$$

Definition 7.5.

(i) A resolution $\mathcal{F}$ of a persistence module M is *minimal* if any resolution $\mathcal{F}'$ of M is isomorphic to the direct sum of $\mathcal{F}$ with a trivial resolution.

(ii) A presentation $\mathcal{F}$ of M is *minimal* if any presentation $\mathcal{F}'$ of M is isomorphic to the direct sum of $\mathcal{F}$ with a trivial presentation.

Theorem 7.6 (Structure of Minimal Resolutions). *Let M be a finitely presented n-parameter persistence module.*

(i) *A minimal free resolution $\mathcal{F}$ of M exists and is unique up to natural isomorphism.*

(ii) *Each module F_i in $\mathcal{F}$ is finitely generated and $F_i = 0$ for $i \geq n + 1$.*

Remarks 7.7.

(i) Theorem 7.6 (i) is variant of a standard result in commutative algebra. Using Azumaya's theorem (Theorem 4.2 (ii)), the proof given in [156] for the case of finitely generated $\mathbb{Z}$-graded $\mathbb{Z}^n$-modules adapts readily to a proof of our version.

(ii) Theorem 7.6 (ii) is a variant of the well-known *Hilbert's basis theorem*; the standard proofs adapt to our multigraded setting.

(iii) Theorem 7.6 implies that a minimal presentation of M also exists and is unique up to isomorphism.

The following easy corollary of Theorem 7.6, due to Scolamiero et al. [166], turns out to be very useful in the study of 2-parameter persistent homology.

Corollary 7.8. *If $f : F_1 \to F_0$ is a morphism of free, finitely presented bipersistence modules, then $\ker f$ is free.*

Example 7.9. Consider the following indecomposable $\mathbb{N}^2$-module M

$$
\begin{array}{ccccccc}
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
k & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 \\
{\scriptstyle [1,0]}\uparrow & & \uparrow & & \uparrow & & \uparrow \\
k^2 & \xrightarrow{[1,1]} & k & \longrightarrow & 0 & \longrightarrow & 0 \\
{\scriptstyle =}\uparrow & {\scriptstyle [1,1]}\uparrow & & \uparrow & & \uparrow \\
k^2 & \xrightarrow{=} & k^2 & \xrightarrow{[0,1]} & k & \longrightarrow & 0
\end{array}
$$

A minimal free resolution of M is given is given in matrix form by

$$
A_1 = \begin{pmatrix} \overset{(0,2)}{1} & \overset{(0,3)}{0} & \overset{(1,1)}{1} & \overset{(2,0)}{0} & \overset{(3,0)}{1} \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix} \begin{matrix} (0,0) \\ (0,0) \end{matrix}
$$

$$
A_2 = \begin{pmatrix} \overset{(1,3)}{1} & \overset{(2,2)}{1} & \overset{(3,1)}{0} \\ 1 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} (0,2) \\ (0,3) \\ (1,1) \\ (2,0) \\ (3,0) \end{matrix}
$$

7.3 Computing Minimal Presentations and Resolutions

Considerations of efficiency aside, minimal presentations and resolutions can be computed using Gröbner basis techniques such as *Schreyer's algorithm* and its variants [130, 165, 92], which in fact work in much greater generality. The application of Schreyer's algorithm to MPH was first explored in [50]. However, the problem instances one considers in TDA have very special structure: the modules are multi-graded (see Section 2.4), the number of persistence parameters is very small (usually at most three), and the problems are very large but very sparse. Thus one would expect that on TDA problems, specialized algorithms and implementations would far outperform more classical approaches designed and optimized for other purposes.

Short Chain Complexes as Input to the Computations Let P be a product of finite totally ordered sets with maximum element $P_{\max}$, and let $F\colon P \to \mathbf{Simp}$ be a simplicial filtration such that $F_{P_{\max}}$ is a finite simplicial complex. We will consider the problem of computing a presentation of $H_i(F)$ for fixed i.

In analogy with ordinary simplicial homology, F has an associated chain complex of free P-modules

$$
\cdots \xrightarrow{\partial_{i+1}} C_i(F) \xrightarrow{\partial_i} C_{i-1}(F) \xrightarrow{\partial_{i-1}} \cdots \xrightarrow{\partial_1} C_0(F) \to 0,
$$

where for $z \in P$, $C_i(F)_z := C_i(F_z; k)$ is the usual simplicial chain vector space with coefficients in k, and the internal maps in $C_i(F)$ are inclusions. The ith

homology module of this chain complex is exactly $H_i(F)$. Given the filtration F, constructing this chain complex (or portions thereof) is generally straight-forward.

It is not hard to verify that $C_i(F)$ is free if and only if F is 1-critical (i.e., if each simplex has a unique birth index; see Section 2.2). If F is not 1-critical, then we may use a simple construction due to Scolamiero et al. [58] to convert the short chain complex

$$C_{i+1}(F) \xrightarrow{\partial_{i+1}} C_i(F) \xrightarrow{\partial_i} C_{i-1}(F)$$

into a short chain complex of free P-modules

$$X \xrightarrow{f} Y \xrightarrow{g} Z \tag{7.2}$$

such that $H_i(F) \cong \ker g / \operatorname{im} f$. This construction is implemented in RIVET in the 2-parameter case, where it is used to compute minimal presentations of degree-Rips bifiltrations.

Thus, whether or not F is 1-critical, we may assume the input to our presentation computation is a short chain complex of free P-modules as in Eq. (7.2) whose homology module is isomorphic to $H_i(F)$.

Algorithms for Bipersistence Modules An efficient algorithm for computing a minimal presentation of a bipersistence module from a short chain complex was introduced in [135] and subsequently improved in [126]. Extensive computational experiments described in [126] demonstrate that this approach scales well with the size of the filtrations, and is efficient enough for practical use in TDA. In fact, this algorithm underlies the MPH computations of [178] which we discussed Section 4.4.

The core computational engine behind this approach is a simple matrix reduction algorithm called the *bigraded reduction*. It is very similar to the standard matrix reduction algorithm for computing persistent homology [181], but instead of reducing an entire matrix in one pass, it proceeds by reducing submatrices of increasing size. Slight variants of the bigraded reduction solve three fundamental algebra problems involving free bipersistence modules [135]: One variant computes a basis for the kernel of a map $\gamma \colon F \to G$ of free modules; a second variant minimizes a set of generators of a persistence module; and a third computes a minimal Gröbner basis for a submodule of a free module. As explained in [135], the first two variants are asymptotically more memory efficient than their more general classical counterparts, which first compute a Gröbner basis of $\operatorname{im} \gamma$.

The first two variants are used in [135, 126] to compute presentations. Specifically, given a short chain complex $X \xrightarrow{f} Y \xrightarrow{g} Z$ of free modules, and letting $M = \ker g / \operatorname{im} f$, one computes a presentation matrix Q for M in three simple steps:

1. compute a minimal set of generators $S = \{s_1, \ldots, s_m\}$ for $\operatorname{im} f$,

2. compute an ordered basis B for $\ker g$, and

3. using ordinary linear algebra, express each $s_i \in S$ as a linear combination of elements of B.

One then takes Q to be the $|B| \times |S|$ matrix with $Q(*, i) = [s_i]^B$. The row and column labels of Q are taken to be the grades of the elements of B and S, respectively.

This presentation will usually not be minimal, but it is *semi-minimal*, in the sense that there is no presentation matrix with the same number of rows and fewer columns. A standard procedure from commutative algebra for minimizing resolutions (essentially, a variant of Gaussian elimination) adapts to minimize such a presentation [109, pages 127 and 166]. An efficient variant of this minimization algorithm in the setting of sparse matrices has been introduced in [98]. This algorithm in fact works for resolutions of multiparameter persistence modules with arbitrary numbers of parameters and, moreover, can be used more generally to minimize a chain complex of free modules.

Letting $l = |X| + |Y| + |Z|$, the above algorithm for computing a minimal presentation runs in time $\Theta(l^3)$ and memory $\Theta(l^2)$; perhaps surprisingly, this matches the asymptotic runtime and memory cost of the ordinary persistence algorithm in the 1-parameter setting [181].

The approach adapts readily to compute a minimal resolution of a bipersistence module. In view of Theorem 7.6 (ii), one needs only to perform a single additional kernel computation.

Further Directions The surprisingly good performance of these minimal presentation computations is cause for optimism about the prospects for applications of MPH. Yet the speed and scalability of 2-parameter persistence computations is still well behind that of 1-parameter computations. Recent approaches to 1-parameter computations use several clever optimizations that together make a huge difference in speed and scalability, as e.g., in [7]. The incorporation of such optimizations into the multiparameter setting is an important direction for future research.

The approach to minimal presentation computation described above makes essential use of special structure in the 2-parameter setting. Gäfvert's recent Ph.D. thesis [101] provides a preliminary account of work with Bender and Lesnick on extending the approach of [135, 126] to multiple parameters, drawing on Faugére's well-known F5 algorithm for computing Gröbner bases [95] and on ideas from the Ph.D. thesis of Skryzalin [172].

8 The Representation Theory of Multiparameter Persistence

In this section, we use quiver representation theory to precisely describe the algebraic complexity of bipersistence modules indexed by the grid $[m] \times [n]$, for all $m, n \geq 1$. We begin in Section 8.1 with a discussion of the relevant representation theory. Then in Section 8.2, we explain how this theory applies to the case of $[m] \times [n]$-modules. Sections 8.3 and 8.4 consider the homology modules of filtrations. Section 8.5 briefly discusses the computation of direct sum decompositions.

8.1 Posets of Finite, Tame, and Wild Type

Drozd's theorem, a foundational result from quiver representation theory, provides a classification of finite posets into into three types (*finite*, *tame*, and *wild*), according to the complexity of their spaces of persistence modules. In what follows, we provide a brief introduction to these ideas. We refer the reader to [93] for an introduction to quiver representation theory catered to a TDA audience, or to [5, 169] for a thorough treatment.

We say that a finite poset P is of *finite representation type* if there exists a finite number of indecomposable P-modules up to isomorphism.

Example 8.1. A poset of the form $\bullet \leftrightarrow \bullet \leftrightarrow \cdots \leftrightarrow \bullet$, where each $\leftrightarrow$ denotes a single arrow pointing either to the left or to the right, is of finite representation type; the indecomposables are precisely the interval modules. This a special case of *Gabriel's theorem* [100]; see [159] for a direct and elementary proof.

Example 8.2. The poset $P = [2] \times [2]$ is also of finite type, and the indecomposables are precisely the interval modules.

Example 8.3. The poset $P = [3] \times [2]$ is of finite type, but the indecomposables do not correspond to intervals. For example, the following P-module is indecomposable:
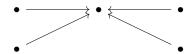
$$
\begin{array}{ccccc}
k & \xrightarrow{[0,1]^T} & k^2 & \xrightarrow{[1,1]} & k \\
\uparrow & & {\scriptstyle [1,0]^T}\uparrow & & {\scriptstyle =}\uparrow \\
0 & \longrightarrow & k & \xrightarrow{=} & k
\end{array}
$$

See [93] for a discussion on how such indecomposables can be used to infer common topological structure in pairs of data sets, e.g. the atomic arrangements of silica glasses.

The following gives an example of a poset P that is *not* of finite type.

Example 8.4. Let P be the 4-*star* poset:

Let n be a positive integer, $\lambda \in k$, and $J(\lambda)$ the $n \times n$ Jordan block over k:

$$
J(\lambda) = \begin{bmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
0 & \lambda & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \lambda
\end{bmatrix}
$$

Define M^λ to be the P-module

where the block matrices are $A = \begin{pmatrix} I_n \\ \mathbf{0} \end{pmatrix}$, $B = \begin{pmatrix} \mathbf{0} \\ I_n \end{pmatrix}$, $C = \begin{pmatrix} I_n \\ I_n \end{pmatrix}$ and $D = \begin{pmatrix} I_n \\ J(\lambda) \end{pmatrix}$, and I_n is the $n \times n$ identity matrix. One can verify that M^λ is indecomposable and that $M^\lambda \not\cong M^{\lambda'}$ for $\lambda \neq \lambda'$. In fact, for k is algebraically closed and n fixed, all but a finite number of indecomposables appear in one of finitely many such 1-parameter families; see [148] for a complete classification of the indecomposables. Furthermore, if we work over F_q, the finite field with q elements, then the number of isomorphism classes of indecomposables is a polynomial in q of degree n [118]; see Remark 8.6 below.

The poset in the previous example is of *tame type*. Roughly speaking, this means that once the vector spaces dimensions have been fixed, all but finitely many of the indecomposables appear in one of finitely many "1-parameter families" of indecomposables. We refer to [5, Section 9.5] for a precise definition.

We next observe that a small modification of the previous example can lead to a dramatic increase in complexity.

Example 8.5. Let $\lambda, \gamma \in k$, and consider the following two-parameter family of persistence modules indexed by the 5-*star poset*:

$$
M^{\lambda,\gamma} = \quad
\begin{array}{ccccc}
 & & k^n \xrightarrow{\;\;A\;\;} k^n \oplus k^n \xleftarrow{\;\;C\;\;} k^n & & \\
 & \nearrow^{B} & \uparrow^{E} & \nwarrow^{D} & \\
k^n & & & & k^n \\
 & & k^n & &
\end{array}
\tag{8.1}
$$

where A, B, C and D are given in Example 8.4 and $E = \begin{pmatrix} I_n \\ J(\gamma) \end{pmatrix}$. Then, one can check that $M^{\lambda,\gamma}$ is indecomposable and that $M^{\lambda,\gamma} \not\cong M^{\lambda',\gamma'}$ for $(\lambda,\gamma) \neq (\lambda',\gamma')$.

In contrast with Example 8.4, however, the collection of such two-parameter families does not come close to describing all isomorphism classes of indecomposables. For instance, we can explicitly describe an $(n-1)$-parameter family of non-isomorphic indecomposables as follows: Fix $\lambda = 0$ and replace $J(\gamma)$ in the matrix E by the $n \times n$ matrix $K(\vec{\alpha})$ whose entries on the super-diagonal are given by a vector $\vec{\alpha} \in k^{n-1}$, with all other entries of $K(\vec{\alpha})$ set to 0. It can be shown that different choices of $\vec{\alpha}$ yield non-isomorphic indecomposables.

To give a further sense of the increase in complexity when passing from the 4-star poset to the 5-star poset, let $I(n,q)$ denote the number of isomorphism classes of indecomposables over F_q for which the dimensions of the vector spaces are as in Eq. (8.1). Then $I(n,q)$ is a polynomial in q of degree $1 + n^2$ [118].

Remark 8.6. More generally, Kac [118] proved that if $\bar{d}$ is the *dimension vector* of an indecomposable quiver representation over F_q, then the total number of indecomposables with dimension vector $\bar{d}$ is a polynomial $f \in \mathbb{Q}[q]$ with

$$
\text{degree}(f) = \begin{cases} 1 - q_Q(\bar{d}) & \text{if} \quad q_Q(\bar{d}) < 0, \\ \gcd(\bar{d}) & \text{if} \quad q_Q(\bar{d}) = 0, \\ 0 & \text{if} \quad q_Q(\bar{d}) > 0, \end{cases}
$$

where q_Q is the *Tits form*. A simple computation of the Tits form recovers the degrees of the polynomials mentioned in Examples 8.4 and 8.5. Computing the lower-order coefficients of the polynomial is difficult, in general; explicit examples for small quivers can be found in [116].

The 5-star poset is an example of a poset of *wild type*. There are several equivalent definitions of wild type [169, Chapter XIX, Theorem 1.11]. We give three, treating the first two informally:

Definition 8.7. A finite poset P is of *wild type* if P satisfies any of the following (equivalent) conditions:

(i) There exists a 2-parameter family of indecomposable P-modules,

(ii) There exists an n-parameter family of indecomposable P-modules for all $n \in \mathbb{N}$,

(iii) For every finite poset Q, there exists an exact functor $F \colon \mathbf{vect}^Q \to \mathbf{vect}^P$ such that

    1. if M is indecomposable, then so is $F(M)$,

    2. $F(M) \cong F(N)$ if and only if $M \cong N$,

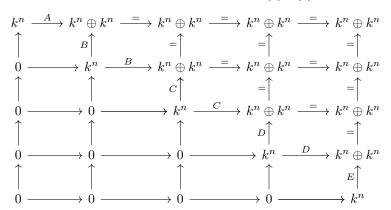   3. $F(cf + dg) = cF(f) + dF(g)$ for all $f, g \colon M \to N$ and $c, d \in k$.

Thus, if P is wild, a full classification of the indecomposable P-modules would yield a classification of the indecomposable Q-modules for any finite poset Q. Obtaining such a classification is thought to be a hopelessly difficult task.

The following "trichotomy" result is remarkable and surprising.

Theorem 8.8 (Drozd [83])**.** *Over an algebraically closed field k, any finite poset is of finite, tame, or wild representation type.*

8.2 The Representation Theory of $[m] \times [n]$

It turns out that the representation type of a commutative grid is wild even for small grid sizes. To see this, observe that any module over the 5-star poset embeds as multiparameter persistence module over $[5] \times [5]$:

$$
\begin{array}{ccccccccc}
k^n & \xrightarrow{A} & k^n \oplus k^n & \xrightarrow{=} & k^n \oplus k^n & \xrightarrow{=} & k^n \oplus k^n & \xrightarrow{=} & k^n \oplus k^n \\
\uparrow & & {\scriptstyle B}\uparrow & & {\scriptstyle =}\uparrow & & {\scriptstyle =}\uparrow & & {\scriptstyle =}\uparrow \\
0 & \longrightarrow & k^n & \xrightarrow{B} & k^n \oplus k^n & \xrightarrow{=} & k^n \oplus k^n & \xrightarrow{=} & k^n \oplus k^n \\
\uparrow & & \uparrow & & {\scriptstyle C}\uparrow & & {\scriptstyle =}\uparrow & & {\scriptstyle =}\uparrow \\
0 & \longrightarrow & 0 & \longrightarrow & k^n & \xrightarrow{C} & k^n \oplus k^n & \xrightarrow{=} & k^n \oplus k^n \\
\uparrow & & \uparrow & & \uparrow & & {\scriptstyle D}\uparrow & & {\scriptstyle =}\uparrow \\
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & k^n & \xrightarrow{D} & k^n \oplus k^n \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & {\scriptstyle E}\uparrow \\
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & k^n
\end{array}
$$

In particular, $[m] \times [n]$ must be of wild type for $m, n \geq 5$. The following theorem gives a complete description of the representation type of finite grids of varying sizes.
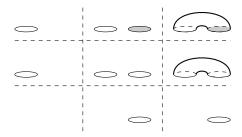
Figure 8.1: Applying $H_1(-;k)$ to the above bifiltration yields the indecomposable bipersistence module in Eq. (8.2).

Theorem 8.9 ([138, Theorem 5], [137, Theorem 2.5]). *The poset $[m] \times [n]$ is of finite type if*

- $m = 1$ *or* $n = 1$,

- $(m, n) \in \{(2, 2), (2, 3), (2, 4), (3, 2), (4, 2)\}$.

It is of tame type if

- $(m, n) \in \{(2, 5), (3, 3), (5, 2)\}$.

It is of wild type in all other cases.

8.3 Homology in Dimension ≥ 1

The previous discussion shows that the representation theory becomes very complicated even for small grids. Moreover, Carlsson and Zomorodian have observed that, when working over a prime field k, one can realize any persistence module over a finite grid by applying $H_1(-;k)$ to a cellular or simplicial sublevel filtration. As an example, Fig. 8.1 illustrates how one can realize the following indecomposable persistence module, previously considered in Example 4.6:

$$
\begin{array}{ccccc}
k & \xrightarrow{\ =\ } & k & \longrightarrow & 0 \\
{\scriptstyle =}\big\uparrow & & {\scriptstyle [1,0]}\big\uparrow & & \big\uparrow \\
k & \xrightarrow{[1,0]^T} & k^2 & \xrightarrow{[1,1]} & k \\
\big\uparrow & & {\scriptstyle [0,1]^T}\big\uparrow & & {\scriptstyle =}\big\uparrow \\
0 & \longrightarrow & k & \xrightarrow{\ =\ } & k
\end{array}
\tag{8.2}
$$

It is shown in [40] that infinite families of indecomposables on a finite grid can arise as the homology modules of function-Rips bifiltrations.

8.4 Homology in Dimension 0

Several bifiltrations that appear naturally in multiparameter persistence have the property that their 0-th homology modules are surjective in at least one parameter direction. As an example, for $\gamma \colon P \to \mathbb{R}$ as in Definition 5.1,

$$
H_0(\mathrm{Rips}^{\downarrow}(\gamma)_{(a,r),(a,s)}) \colon H_0(\mathrm{Rips}^{\downarrow}(\gamma)_{a,r}) \to H_0(\mathrm{Rips}^{\downarrow}(\gamma)_{a,s})
$$

is surjective for all $r \leq s$.

Remark 8.10. H_0 factors as $H_0(-; k) \cong \text{Free} \circ \pi_0$, where π_0 denotes path-components, and Free: $\mathbf{Set} \to \mathbf{Vect}_k$ is the free functor sending a set X to the free k-vector space with basis X. The fact that $H_0(\text{Rips}^{\downarrow}(\gamma))_{(a,r),(a,s)}$ is an epimorphism for $r \leq s$ is a direct consequence of this factorization. See [58, 38] for a further discussion of persistence modules factoring through the category of sets.

A natural question is to what extent such additional structure reduces the algebraic complexity. For an instance of a result in this direction, one can show that that if M is an $[m] \times [n]$-module with surjections in one direction and injections in the other, then M is interval-decomposable; see Theorem 10.4. On the other hand, the two-parameter families of persistence modules from Example 8.5 all embed as $[5] \times [5]$-modules with all morphisms surjective. This suggests that the reduction in complexity can only be rather limited. The following theorem confirms that.

Theorem 8.11 ([8])**.** *The category of $[m] \times [n]$-modules whose horizontal maps are surjective is equivalent to the category of $[m] \times [n-1]$-modules, up to a finite, explicit list of indecomposables.*

An analogue of Theorem 8.11 for the case where all morphisms are surjective is also given in [8]. Furthermore, Theorem 8.11 and Theorem 8.9 together immediately imply the following:

Corollary 8.12. *For $[m] \times [n]$-modules whose horizontal maps are surjective, the cases of finite type are:*

- *$m = 1$ or $n \leq 2$ (all modules are interval-decomposable)*

- *$(m, n) \in \{(2, 3), (2, 4), (2, 5), (3, 3), (4, 3)\}$,*

and the tame cases are: $(m, n) \in \{(2, 6), (3, 4), (5, 3)\}$.

8.5 Computing Decompositions

While explicitly classifying the indecomposable multiparameter persistence modules is hopeless, one can still compute the indecomposable summands of any given persistence module. General purpose polynomial-time algorithms for decomposing modules over finite-dimensional algebras exist in the literature and implementations are readily available; see e.g,. the MeatAxe algorithm [115], which computes a decomposition of a P-module in time $\tilde{O}\left(\left(\sum_{p \in P} \dim M_p\right)^6\right)$, and is implemented in GAP [104]. However, such algorithms are impractical in the context of TDA.

Dey and Xin [81] have recently provided a decomposition algorithm tailored to multiparameter persistence. Their algorithm takes as input a presentation matrix of a d-parameter persistence module M. Under the assumption that no two rows or columns have the same label, the algorithm outputs a decomposition of M into indecomposables. When this assumption does not hold, the algorithm still provides a decomposition of M into summands, but the summands may not be indecomposabale. For an input presentation matrix with a total of N rows

and columns, the algorithm runs in time $O(N^{(d-1)(2\omega+1)})$, where $\omega < 2.373$ is the exponent for matrix multiplication. While this work represents significant progress on the efficient decomposition of MPH modules, new ideas are needed for how to use such decompositions in the context of data analysis.

9   Signed Barcodes

Recall that the rank invariant $\text{Rk}\,M$ of a persistence module M is the function $(s,t) \mapsto \text{Rk}(M_s \to M_t)$, where $s \le t$. As noted in Section 4.2, the rank invariant is a complete invariant of a 1-parameter persistence module. Thus, the barcode and rank invariant of a 1-parameter persistence module determine each other. Though the rank invariant is incomplete outside of the 1-parameter setting, one may hope that the correspondence between barcodes and rank invariants can be extended to the multiparameter setting. As we will explain, this is indeed possible if one considers *signed barcodes*.

In this section we will consider three signed barcode constructions appearing in the recent TDA literature, which we will call the *rectangle*, *interval*, and *hook* barcodes. The rectangle barcode [36] and hook barcodes [36, 37] are both equivalent to the rank invariant; the former can be constructed via Möbius inversion, while the latter is constructed via relative homological algebra. The hook barcode has the advantage of being stable in a stronger sense than the rectangle barcode. The interval barcode (also known as the *generalized persistence diagram*) [36, 127, 2, 127] is an analogue of the rectangle barcode which is equivalent to the generalized rank invariant (Definition 4.7).

9.1   Rectangle Barcodes via Decomposition of the Rank Invariant

In Example 4.6, we saw a module M whose rank invariant is not equal to the rank invariant of any interval-decomposable module. On the other hand, the rank invariant of M can be expressed as the *difference* between the rank invariants of two interval-decomposable modules:

$$\tag{9.1}$$

Such a decomposition need not be unique:

$$
\mathrm{Rk}\left(\begin{array}{c}\mathbf{k}\xrightarrow{\;\mathrm{id}\;}\mathbf{k}\longrightarrow 0\\ {\scriptstyle\mathrm{id}}\Big\downarrow\quad\Big\downarrow{\scriptstyle[1\,0]}\quad\Big\downarrow\\ \mathbf{k}\xrightarrow{\;[{1\atop 0}]\;}\mathbf{k}^2\xrightarrow{[1\,1]}\mathbf{k}\\ \Big\downarrow\quad\Big\downarrow{\scriptstyle[{0\atop 1}]}\quad\Big\downarrow{\scriptstyle\mathrm{id}}\\ 0\longrightarrow\mathbf{k}\xrightarrow{\;\mathrm{id}\;}\mathbf{k}\end{array}\right)=\mathrm{Rk}\left(\;\blacksquare\;\oplus\;\blacksquare\;\oplus\;\blacksquare\;\oplus\;\blacksquare\;\right)
$$

$$
-\;\mathrm{Rk}\left(\;\blacksquare\;\oplus\;\blacksquare\;\right)
$$

(9.2)

It turns out, however, that the decomposition is unique if we restrict to rectangles.

Definition 9.1. A *rectangle* in a product of totally ordered sets $T_1 \times \cdots \times T_n$ is a poset of the form $I_1 \times \cdots \times I_n$, where each I_j is an interval in T_j.

Theorem 9.2 ([36]). *Let $P = T_1 \times \cdots \times T_n$, where each $T_i \subseteq \mathbb{R}$, and let M be a finitely presented P-module. There exists a* minimal *pair $(\mathcal{R}, \mathcal{S})$ of finite multisets of rectangles in P satisfying*

$$
\mathrm{Rk}\,M = \mathrm{Rk}\left(\bigoplus_{I \in \mathcal{R}} k_I\right) - \mathrm{Rk}\left(\bigoplus_{J \in \mathcal{S}} k_J\right); \tag{9.3}
$$

here minimal means that if $(\mathcal{R}', \mathcal{S}')$ is any other such pair satisfying Eq. (9.3), then $\mathcal{R} \subseteq \mathcal{R}'$ and $\mathcal{S} \subseteq \mathcal{S}'$.

We shall refer to the pair $(\mathcal{R}, \mathcal{S})$ as the *rectangle barcode* of $\mathrm{Rk}\,M$. (It is called a *minimal rank decomposition* in [36].) Note that if $d = 1$, then by Theorem 3.4 and the uniqueness of the signed barcode, we must have that $\mathcal{S} = \emptyset$.

The rectangle barcode can be constructed by taking a Möbius inversion of the restriction of the rank invariant to a finite grid [36].
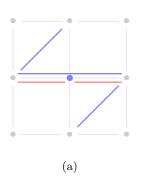
Remark 9.3. The idea of using Möbius inversion to associate a signed barcode to the rank invariant was introduced by Patel [155], who worked in the setting of 1-parameter persistence modules taking values in certain symmetric monoidal categories. Subsequently, Möbius inversions have been applied to persistence modules in several settings [139, 2, 127, 18]. See [36, Section 1] for a detailed discussion of the connection between the various approaches.

The 1-parameter slices of rectangle barcodes are stable, in the sense of the following theorem:

Theorem 9.4 ([36]). *Given signed barcodes $(\mathcal{R}, \mathcal{S})$ and $(\mathcal{R}', \mathcal{S}')$ of finitely presented $\mathbb{R}^d$-persistence modules M and M', we have:*

$$
d_{\mathrm{match}}\left(\bigoplus_{I \in \mathcal{R}} k_I \oplus \bigoplus_{J' \in \mathcal{S}'} k_{J'},\; \bigoplus_{I' \in \mathcal{R}'} k_{I'} \oplus \bigoplus_{J \in \mathcal{S}} k_J\right) \leq d_{\mathrm{match}}(M, M').
$$

One may ask whether Theorem 9.4 still holds if one replaces d_{match} with the generalized bottleneck distance d_B (Definition 6.4). The answer is no, even up to a constant factor [37, Proposition 3.1]. In contrast, the hook barcode, which we will introduce in Section 9.3, is Lipschitz stable with respect to d_B.
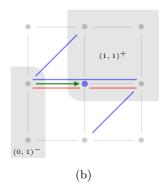
46

(a) (b)

Figure 9.1: (a) Visualization of the signed barcode of Eq. (9.2). Rectangles with positive and negative sign are represented using blue and red bars, respectively. (b) The rank of the map $M_{(0,1)} \to M_{(1,1)}$ (represented by the green arrow) is 1; This is the number of blue bars intersecting both gray regions minus the number of red bars intersecting both regions.

Visualization   Visualizing the rectangle barcode of a bipersistence module by directly plotting the rectangles can be messy for examples of realistic size, since many rectangles may overlap. To obtain a cleaner visualization, [36] proposes to instead represent each rectangle via a line segment ("bar") connecting its infimum to its supremum, as illustrated in Fig. 9.1 (a). Examples involving larger persistence modules can be found in [36].

One can read the rank invariant off of such a visualization: $\mathrm{Rk}(M_s \to M_t)$ is the number of bars in $\mathcal{R}$ intersecting both the sets $s^- = \{u \in P \colon u \le s\}$ and $t^+ = \{u \in P \colon u \ge t\}$, minus the number of bars in $\mathcal{S}$ intersecting both sets; see Fig. 9.1 (b).

Computation   The rectangle barcode of a p.f.d. persistence module M indexed over a finite grid $\mathcal{G} = \prod_{i=1}^d [1, n_i]$ can be computed by means of a simple inclusion-exclusion formula. Let $\alpha_{[s,t]}$ denote the multiplicity of the rectangle

$$[s, t] = \{u \in \mathbb{R}^d \mid s \le u \le t\},$$

in the rectangle barcode $(\mathcal{R}, \mathcal{S})$. That is, if $\alpha_{[s,t]} > 0$ then $[s, t]$ appears with multiplicity $\alpha_{[s,t]}$ in $\mathcal{R}$, and if $\alpha_{[s,t]} < 0$ then $[s, t]$ appears with multiplicity $-\alpha_{[s,t]}$ in $\mathcal{S}$. Then we have:

$$\alpha_{[s,t]} = \sum_{\substack{s' \le s \\ \|s'-s\|_\infty \le 1}} \sum_{\substack{t' \ge t \\ \|t'-t\|_\infty \le 1}} (-1)^{\|s'-s\|_1 + \|t'-t\|_1} \ \mathrm{Rk}\, M(s', t'). \qquad (9.4)$$

Remark 9.5. The case $d = 1$ gives the well-known inclusion-exclusion formula relating the persistence diagram of a one-parameter persistence module to its rank invariant [67]. The case $d = 2$ specializes to the inclusion-exclusion formula for computing the multiplicities of summands in rectangle-decomposable 2-parameter persistence modules [35].

A simple inspection of the formula in Eq. (9.4) reveals that computation is bounded in time $O\left(2^{2d} \prod_{i=1}^d n_i^2\right)$, assuming constant-time access to the ranks

Rk $M(s', t')$ and constant-time arithmetic operations. When the module M comes from a simplicial filtration over the grid $\mathcal{G}$ with $n = \max_i n_i$ simplices in total, the rank invariant itself can be pre-computed and stored, e.g., by naively computing the ranks Rk $M(s, t)$ for each pair $s \leq t \in \mathcal{G}$ independently, which takes $O(n^{2d+\omega})$ time in total, where $\omega < 2.373$ is the exponent for matrix multiplication [145]. Adding in the computation time for the signed barcode yields a bound in $O(n^{2d+\omega} + (2n)^{2d})$. In the special case where $d = 2$, assuming the filtration is 1-critical (i.e. each simplex has a unique minimal index of appearance in the filtration), there is an $O(n^4)$-time algorithm to compute the rank invariant [146, 35], and computing its rectangle barcode also takes $O(n^4)$ time.

9.2 Interval Barcodes via Decomposition of the Generalized Rank Invariant

While a smaller signed barcode using intervals of other shapes than rectangles is sometimes possible, it comes at the expense of uniqueness, as illustrated by Eq. (9.1) and Eq. (9.2). It is however possible to retain uniqueness if one instead decomposes the generalized rank invariant (Definition 4.7).

Let P be an arbitrary poset. We say a collection $\mathcal{I}$ of intervals in P is *locally finite* if for any $I, K \in \mathcal{I}$ the set $\{J \in \mathcal{I} : I \subseteq J \subseteq K\}$ is finite.

Theorem 9.6 ([36])**.** *Let $\mathcal{I}$ be a locally finite collection of intervals in P, and let M be a p.f.d. P-module. Then there exists a pair $(\mathcal{R}, \mathcal{S})$, where $\mathcal{R}$ and $\mathcal{S}$ are finite multi-sets of intervals in $\mathcal{I}$, such that*

$$\mathrm{Rk}_{\mathcal{I}} M = \mathrm{Rk}_{\mathcal{I}} \left(\bigoplus_{I \in \mathcal{R}} k_I \right) - \mathrm{Rk}_{\mathcal{I}} \left(\bigoplus_{J \in \mathcal{S}} k_J \right), \tag{9.5}$$

and if $(\mathcal{R}', \mathcal{S}')$ is any other such pair satisfying Eq. (9.5), then $\mathcal{R} \subseteq \mathcal{R}'$ and $\mathcal{S} \subseteq \mathcal{S}'$.

Remark 9.7. When $\mathcal{I}$ consists of all intervals in an essentially finite poset P, the *generalized persistence diagram* of [127], constructed via Möbius inversion of the generalized rank invariant, is essentially the same as the decomposition $(\mathcal{R}, \mathcal{S})$ of Theorem 9.6.

9.3 Hook Barcodes via Relative Homological Algebra

Several recent works use *relative homological algebra* to study generalized persistence. In brief, the idea is to identify an *exact structure* for a given invariant, i.e., a distinguished class of exact sequences which respects the structure of the invariant. There is a notion of *relative minimal projective resolution* associated to any exact structure, which is used to define a signed barcode. This approach to generalized persistence was initiated for the rank invariant in [36], and in parallel for other invariants in [28]; see also [153, 3, 57, 37] for other recent work, and [42] for an excellent introduction to exact categories.

In the special case of the rank invariant, this approach yields the *hook barcode*, a signed barcode which is bottleneck-stable with respect to the interleaving distance [37]. We now explain this case in more detail.

Definition 9.8.

(i) A short exact sequence

$$0 \to M \to M' \to M'' \to 0$$

of $\mathbb{R}^n$-persistence modules is *rank exact* if $\operatorname{Rk} M' = \operatorname{Rk} M + \operatorname{Rk} M''$,

(ii) A long exact sequence

$$\cdots \xrightarrow{f_{i+1}} X_i \xrightarrow{f_i} X_{i-1} \xrightarrow{f_{i-1}} \cdots$$

is *rank exact* if $\operatorname{im} f_i \to X_{i-1} \to \operatorname{im} f_{i-1}$ is rank exact for all i,

(iii) An $\mathbb{R}^n$-persistence module M is *rank projective* if $\operatorname{Hom}(M, -)$ takes each rank exact sequence to an exact sequence of vector spaces.

Definition 9.9. A *hook* is an interval in $\mathbb{R}^n$ of the form $\{p \in \mathbb{R}^n : r \le p \not\ge s\}$ for $r < s \in \mathbb{R}^n \cup \{\infty\}$.

One can show that if $M \colon \mathbb{R}^n \to \mathbf{Vect}$ is finitely presented and rank projective, then M is *hook-decomposable*, i.e., M is interval-decomposable and each interval in $\mathcal{B}(M)$ is a hook.

Theorem 9.10 ([36, 37]). *A finitely presented $\mathbb{R}^n$-persistence module M admits a minimal rank projective resolution*

$$0 \to X_k \to X_{k-1} \to \cdots \to X_1 \to X_0 \to M,$$

with each X_i finitely presented (and hook-decomposable) and $k \le 2n - 2$.

It follows from the theory of exact categories that a minimal rank projective resolution is unique up to isomorphism. Given a finitely presented $\mathbb{R}^n$-persistence module M, let $\beta_i(M) = \mathcal{B}(X_i)$, where X_i is as in Theorem 9.10.

Definition 9.11. We define the *hook barcode* of M to be the pair

$$(\beta_{2\mathbb{N}}(M), \beta_{2\mathbb{N}+1}(M)) := \left(\bigcup_{i \text{ even}} \beta_i(M), \bigcup_{i \text{ odd}} \beta_i(M) \right).$$
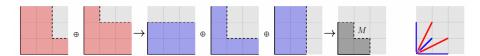
Remarks 9.12.

(i) By virtue of the exact structure, one has that

$$\operatorname{Rk} M = \sum_{i=0}^{k} (-1)^i \operatorname{Rk} X_i = \sum_{i \text{ even}} \operatorname{Rk} X_i - \sum_{i \text{ odd}} \operatorname{Rk} X_i.$$

This implies that we can interpret the hook barcode of M as a decomposition of $\operatorname{Rk} M$, as for the rectangle barcode.

(ii) In contrast to the rectangle barcode, the two components of the hook barcode may have intervals in common. Removing common hooks from the decomposition in (i) yields a minimal decomposition of the rank function in terms of hook modules. However, the main stability property of the hook barcode (Theorem 9.13 below) does not hold for this minimal decomposition [37, Remark 3.2].

Like the rectangle barcode, the hook barcode can be visualized by a collection of line segments, as the following rank projective resolution illustrates.

The hook barcode is stable in the following sense:

Theorem 9.13 ([37]). *If M and N are finitely presented $\mathbb{R}^n$-persistence modules, then*

$$d_B\left(\beta_{2\mathbb{N}}(M) \cup \beta_{2\mathbb{N}+1}(N), \beta_{2\mathbb{N}}(N) \cup \beta_{2\mathbb{N}+1}(M)\right) \leq (2n-1)^2 d_I(M, N).$$

The proof is an application of the algebraic stability of hook-decomposable modules (Theorem 6.5 (iv)).

Remark 9.14. Computation of minimal resolutions and *relative Betti numbers* in exact categories is an area of active research. Recently, Chachólski et al. [57] provided a framework for computing relative Betti numbers of exact structures on upper semi-lattices.

10 Local Conditions for Interval Decomposability

In this section, we introduce local conditions on a bipersistence module M which force M to be interval-decomposable, with each interval in $\mathcal{B}(M)$ of a particularly simple form. We will see that such conditions arise naturally in the study of the *interlevel persistence* of $\mathbb{R}$-valued functions.

10.1 Decomposition into Blocks and Rectangles

Let S and T be totally ordered sets.

Definition 10.1. Given $P \subseteq S \times T$ and a persistence module $M \colon P \to \mathbf{Vect}$,

(i) M is *middle-exact* [32, 66, 49, 48] if

$$M_a \xrightarrow{M_{a,b} \oplus M_{a,c}} M_b \oplus M_c \xrightarrow{(M_{b,d} - M_{c,d})} M_d \qquad (10.1)$$

is exact (i.e., the image of the first map equals the kernel of the second map) for all $a, b, c, d \in P$ of the form

$$a = (x, y), \quad b = (x, y'), \quad c = (x', y), \quad d = (x', y').$$

(ii) M is *weakly-exact* [34, 35] if for all a, b, c, d as above,

$$\mathrm{Im}(M_a \to M_d) = \mathrm{Im}(M_b \to M_d) \cap \mathrm{Im}(M_c \to M_d),$$
$$\mathrm{Ker}(M_a \to M_d) = \mathrm{Ker}(M_a \to M_b) + \ker(M_a \to M_c).$$

One can check that if M is middle-exact then it is weakly-exact.

Recall from Example 8.2 that a persistence module over a commutative square is interval-decomposable. A simple case-by-case inspection of the intervals gives the following characterization of middle-exact and weakly-exact persistence modules.

Proposition 10.2.

(i) *M is middle-exact if and only if its restriction to any square*

$$
\begin{array}{ccc}
M_c & \longrightarrow & M_d \\
\uparrow & & \uparrow \\
M_a & \longrightarrow & M_b
\end{array}
$$

decomposes into interval modules supported on the set

$$B := \{\{a\}, \{d\}, \{a, b\}, \{a, c\}, \{b, d\}, \{c, d\}, \{a, b, c, d\}\}.$$

(i) *M is weakly-exact if and only if its restriction to any square as above decomposes into interval modules supported on the set*

$$B \cup \{\{b\}, \{c\}\}.$$

Definition 10.3. An interval $I \subseteq S \times T$ is called a *block* if I can be written in one of the following ways:

1. $I = J_S \times J_T$ for downsets J_S and J_T,

2. $I = J_S \times J_T$ for upsets J_S and J_T,

3. $I = J_S \times T$ for an interval J_S,

4. $I = S \times J_T$ for an interval J_T.

More generally, recall from Definition 9.1 that a rectangle in $S \times T$ is an interval of the form $I \times J$, where I and J are intervals in S and T. We say that M is *block-decomposable* (respectively, *rectangle-decomposable*) if M is interval-decomposable and each interval of $\mathcal{B}(M)$ is a block (respectively, a rectangle).

The following theorems tell us that a module is block-decomposable or rectangle-decomposable if and only if its restriction to each square is.

Theorem 10.4 ([32]). *A p.f.d. $S \times T$-module is middle-exact if and only if it is block-decomposable.*

Given a totally ordered set T, a subset $T' \subset T$ is said to be *coinitial* if for all $t \in T$ there exists $t' \in T'$ with $t' \leq t$. For example, $\mathbb{Z}$ is a coinitial subset of $\mathbb{R}$.

Theorem 10.5 ([34]). *Suppose that S and T are totally ordered sets such that every interval in S or T admits a countable coinitial subset. Then a p.f.d. $S \times T$-module is weakly-exact if and only if it is rectangle-decomposable.*

Remark 10.6. In the case that S and T satisfy the conditions of Theorem 10.5, Theorem 10.4 follows easily from Theorem 10.5.

Remark 10.7. Proofs of Theorem 10.4 for the special cases $S \times T$ finite and $S \times T = \mathbb{R}^2$ first appeared in [14] and [66], respectively; see also [49, 48], which had a strong influence on these results. For the case $S \times T$ finite, an elementary proof of Theorem 10.5 appeared in [35].

As in Section 5.2, let

$$\overline{U} := \{(a,b) \in \mathbb{R}^{\mathrm{op}} \times \mathbb{R} \colon a \leq b\},$$

and let

$$U := \{(a,b) \in \mathbb{R}^{\mathrm{op}} \times \mathbb{R} \colon a < b\}.$$

As a (non-trivial) application of Theorem 10.4, [32] establishes the following result.

Theorem 10.8 ([32])**.** *Let P be U or $\overline{U}$. Any p.f.d. middle-exact P-module M is interval-decomposable, and each interval of $\mathcal{B}(M)$ is of the form $P \cap I$, where I is a block in $\mathbb{R}^{\mathrm{op}} \times \mathbb{R}$.*

We shall also refer to the intervals $P \cap I$ arising in Theorem 10.8 as *blocks (in P)*.

10.2   Interlevel Persistent Homology

One can use Theorem 10.8 to define *interlevel barcodes*, fundamental invariants of $\mathbb{R}$-valued functions which refine the standard sublevel barcodes. Moreover, these invariants admit extensions given using relative homology, and the invariants in different homological degrees assemble nicely into a single object, the *Mayer-Vietoris strip*. We will now explain this.

To formulate the definitions, one has to choose between using open intervals or closed intervals in $\mathbb{R}$. For consistency with our treatment of sublevel filtrations (Definition 2.2), we will use closed intervals here, though using open intervals is arguably more convenient. We briefly discuss the approach via open intervals in Remark 10.15 below.

Given a continuous function $\gamma \colon W \to \mathbb{R}$, the *interlevel bifiltration of γ* is the functor

$$\mathcal{S}(\gamma) \colon \overline{U} \to \mathbf{Top}, \quad \mathcal{S}(\gamma)(s,t) = \{x \in X : s \leq \gamma(x) \leq t\}.$$

It follows from the Mayer–Vietoris sequence that for all $i \geq 0$, the restriction of $H_i \mathcal{S}(\gamma)$ to U is middle-exact. Thus, if $H_i \mathcal{S}(\gamma)$ is p.f.d., then Theorem 10.8 provides a barcode of the restriction of $H_i \mathcal{S}(\gamma)$ to U.

Furthermore, assuming that γ is a Morse function or, more generally, *of Morse type* [49, Section 2], $H_i \mathcal{S}(\gamma)$ is itself middle-exact, and Theorem 10.8 then provides a barcode $\mathcal{B}(H_i \mathcal{S}(\gamma))$ of $H_i \mathcal{S}(\gamma)$ without having to restrict to U.

Taking the intersection of each interval of $\mathcal{B}(H_i \mathcal{S}(\gamma))$ with $\partial U = \{(x,x) \mid x \in \mathbb{R}\}$ and projecting onto the first coordinate, we recover the ith *levelset barcode* of γ, as introduced in [49]; we denote this as $\mathcal{L}_i(\gamma)$. The intervals in this barcode encode how the homology of the fibers $\gamma^{-1}(r)$ changes as r varies.

Example 10.9. Fig. 10.1 illustrates a function $\gamma \colon W \to \mathbb{R}$ and the barcodes $\mathcal{B}(H_i \mathcal{S}(\gamma))$. By considering how the intervals in these barcodes intersect ∂U, we see that

$$\mathcal{L}_0(\gamma) = \{[a_1, a_6], [a_2, a_3), (a_4, a_5)\}$$
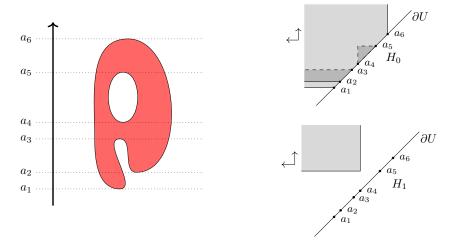
and $\mathcal{L}_i(\gamma) = \emptyset$ for $i > 0$.

Figure 10.1: The interlevel homology of the function $\gamma\colon W \to \mathbb{R}$ given by $\gamma(x, y) = y$, where $W \subset \mathbb{R}^2$ is as shown in red.
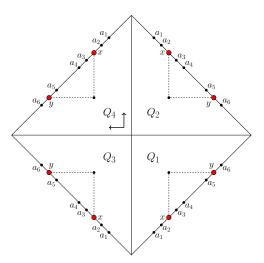

Figure 10.2: Visualization of the poset Q indexing relative interlevel homology.

Relative Interlevel Homology   For each $i \geq 0$, we can extend the above construction to a relative version as follows: consider the set

$$Q = Q_1 \sqcup Q_2 \sqcup Q_3 \sqcup Q_4$$

of pairs of subsets of $\mathbb{R}$, where

$$Q_1 = \{([x,y], \emptyset) \colon x \leq y\} \qquad Q_2 = \{([x,\infty), [y,\infty)) \colon x \leq y\}),$$
$$Q_3 = \{((-\infty, y], (-\infty, x]) \colon x \leq y\}, \quad Q_4 = \{(\mathbb{R}, (-\infty, x] \cup [y, \infty)) \colon x \leq y\}.$$

We regard Q as poset by taking $(A, B) \leq (A', B')$ if and only if $A \subseteq A'$ and $B \subseteq B'$. Let $2\mathbb{R}$ denote the totally ordered set consisting of two disjoint copies of $\mathbb{R}$, with all elements in the first copy less than all elements in the second copy. We have a natural identification of Q with an interval in $(2\mathbb{R}^{\mathrm{op}}) \times 2\mathbb{R}$.

Given a continuous function $\gamma \colon W \to \mathbb{R}$, we define a functor $\mathcal{S}^\bullet(\gamma)$ from Q to the category of pairs of topological spaces by

$$\mathcal{S}(\gamma)^\bullet(A, B) = (\gamma^{-1}(A), \gamma^{-1}(B)).$$

If γ is of Morse type, then $H_i \mathcal{S}^\bullet(\gamma)$ is interval-decomposable, where each interval is the intersection of a rectangle in $2\mathbb{R}^{\mathrm{op}} \times 2\mathbb{R}$ with Q [14, 35]. It is helpful to visualize Q schematically as the decomposition of a tilted square into four triangles, as shown in Fig. 10.2; we can make formal sense of this schematic either by reparameterizing Q as in [9], or by assuming that γ is bounded above and below, in which case we may replace Q with a subposet where x and y are correspondingly bounded, as in [49].

Example 10.10. Fig. 10.3 shows the decompositions of the modules $H_0 \mathcal{S}^\bullet(\gamma)$ and $H_1 \mathcal{S}^\bullet(\gamma)$ for γ the function of Theorem 10.9 and Fig. 10.2. Observe that the off-diagonal block in $H_1 \mathcal{S}(\gamma)$ now extends all the way to the top of Q. Moreover, if we reflect the block around the anti-diagonal passing through the center of Q, then we see that the boundary of the block aligns perfectly with the boundary of the triangle from $H_0 \mathcal{S}(\gamma)$. Geometrically, this makes sense, as the circle in $\gamma^{-1}[a_4, a_5]$ gives rise to two connected components in all pre-images $\gamma^{-1}[x, y]$ where $[x, y] \subseteq (a_4, a_5)$.

The Mayer–Vietoris Strip   In the previous example, there is a natural matching between $\mathcal{B}(H_0 \mathcal{S}^\bullet(\gamma))$ and $\mathcal{B}(H_1 \mathcal{S}^\bullet(\gamma))$. It turns out that one always has such a matching. To explain, using excision and the boundary map in the relative Mayer–Vietoris sequence, one obtains maps connecting the diagrams $H_i \mathcal{S}^\bullet(\gamma)$ and $H_{i+1} \mathcal{S}^\bullet(\gamma)$ for all i. Using these maps and reflecting every other diagram $H_i \mathcal{S}^\bullet$ as shown in Fig. 10.3, the diagrams $\{H_i \mathcal{S}^\bullet(\gamma)\}_{i \in \mathbb{N}}$ assemble into one big persistence module called the *Mayer–Vietoris strip* of γ. Let Q' denote the indexing poset of this persistence module. Like Q, the poset Q' is naturally identified with an interval in a product of two totally ordered sets, so rectangles in Q' are well defined.

It turns out that the Mayer–Vietoris strip is middle-exact. Using this, one can prove the following structure theorem.

Theorem 10.11 ([9, 14, 34]). *If γ is of Morse type, then the Mayer–Vietoris strip of γ is interval-decomposable, where each interval is a maximal rectangle in Q'.*

Example 10.12. The Mayer–Vietoris strip of the function in Fig. 10.2 (a) is shown in Fig. 10.3.

Remark 10.13 (Computation). The barcode of the Mayer–Vietoris strip of γ is fully determined by the levelset barcodes of γ in all homological degrees, and is also fully determined by the *extended persistence* barcodes [68] of γ in all homological degrees; this is implicit in [49]. It follows that in practice, one can obtain the barcode of the Mayer–Vietoris strip by either a (1-parameter) extended persistence computation or a zigzag persistence computation. Conversely, the barcode of the Mayer-Vietoris strip determines both the levelset and extended persistence barcodes.

Remark 10.14 (Stability). The bottleneck distance (Definitions 3.6 and 6.4) generalizes straightforwardly to barcodes consisting of rectangles in Q', and it can be shown that the barcode of the Mayer–Vietoris strip is stable with respect to this and the L^∞-distance on functions [49, 9]. This is a direct consequence of the stability of extended persistent homology [68].

Remark 10.15 (Interlevel Persistence with Open Intervals). We now briefly discuss the corresponding interlevel persistence theory using open intervals. Given a continuous function, $\gamma : W \to \mathbb{R}$, we can define

$$\mathcal{S}(\gamma)^o \colon U \to \mathrm{Top}, \qquad \mathcal{S}(\gamma)^o(s,t) = \{x \in X \mid s < \gamma(x) < t\}.$$

By the Mayer–Vietoris sequence, $\mathcal{S}(\gamma)^o$ is middle-exact. Thus by Theorem 10.8, if $H_i \mathcal{S}(\gamma)^o$ is p.f.d., then it is interval-decomposable. Analogously, we can define a variant of the Mayer–Vietoris strip using open intervals, and we have a structure theorem [34, 9] which says that if this persistence module is p.f.d., then it is interval-decomposable where the intervals are maximal rectangles. As in the case of closed intervals, this decomposition is stable.

Remark 10.16 (Related Results). See Bauer et al. [9] for a careful construction of the Mayer–Vietoris strip for an arbitrary cohomology theory, and a corresponding structure theorem in the p.f.d. setting. A structure theorem for middle-exact p.f.d. modules over the Mayer–Vietoris strip not necessarily coming from functions can be found in [34]. A related construction and structure theorem in the context of derived sheaf theory can be found in [15].

References

[1] Aaron Adcock, Daniel Rubin, and Gunnar Carlsson. Classification of hepatic lesions using the matching metric. *Computer vision and image understanding*, 121:36–42, 2014.

[2] Hideto Asashiba, Emerson G Escolar, Ken Nakashima, and Michio Yoshiwaki. On approximation of 2 d persistence modules by interval-decomposables. *arXiv preprint arXiv:1911.01637*, 2019.

[3] Hideto Asashiba, Emerson G Escolar, Ken Nakashima, and Michio Yoshiwaki. Approximation by interval-decomposables and interval resolutions of persistence modules. *arXiv preprint arXiv:2207.03663*, 2022.
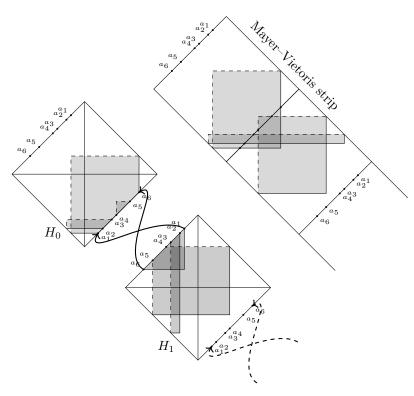
Figure 10.3: Relative interlevel set modules $H_i \mathcal{S}^\bullet(\gamma)$ and the corresponding Mayer–Vietoris strip of the function γ from Example 10.9; see Examples 10.10 and 10.12 and the surrounding discussion.

[4] Gorô Azumaya. Corrections and supplementaries to my paper concerning Krull–Remak–Schmidt's theorem. *Nagoya Mathematical Journal*, 1:117–124, 1950.

[5] Michael Barot. *Introduction to the representation theory of algebras*. Springer, 2015.

[6] Yuliy Baryshnikov et al. Biparametric persistence for smooth filtrations. *arXiv preprint arXiv:2110.09602*, 2021.

[7] Ulrich Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, pages 1–33, 2021.

[8] Ulrich Bauer, Magnus B Botnan, Steffen Oppermann, and Johan Steen. Cotorsion torsion triples and the representation theory of filtered hierarchical clustering. *Advances in Mathematics*, 369:107171, 2020.

[9] Ulrich Bauer, Magnus Bakke Botnan, and Benedikt Fluhr. Structure and interleavings of relative interlevel set cohomology. *arXiv preprint arXiv:2108.09298*, 2021.

[10] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *Journal of Computational Geometry*, 6(2):162–191, 2015.

[11] Ulrich Bauer and Michael Lesnick. Persistence diagrams as diagrams: A categorification of the stability theorem. In *Topological Data Analysis*, pages 67–96. Springer, 2020.

[12] Ulrich Bauer, Axel Munk, Hannes Sieling, and Max Wardetzky. Persistence barcodes versus kolmogorov signatures: detecting modes of one-dimensional signals. *Foundations of Computational Mathematics*, 17(1):1–33, 2017.

[13] Gabriele Beltramo, Rayna Andreeva, Ylenia Giarratano, Miguel O Bernabeu, Rik Sarkar, and Primoz Skraba. Euler characteristic surfaces. *arXiv preprint arXiv:2102.08260*, 2021.

[14] Paul Bendich, Herbert Edelsbrunner, Dmitriy Morozov, and Amit Patel. Homology and robustness of level and interlevel sets. *Homology, Homotopy and Applications*, 15(1):51–72, 2013.

[15] Nicolas Berkouk, Grégory Ginot, and Steve Oudot. Level-sets persistence and sheaf theory. *arXiv preprint arXiv:1907.09759*, 2019.

[16] Nicolas Berkouk and François Petit. Ephemeral persistence modules and distance comparison. *Algebraic & Geometric Topology*, 21(1):247–277, 2021.

[17] Catalina Betancourt, Mathieu Chalifour, Rachel Neville, Matthew Pietrosanu, Mimi Tsuruga, Isabel Darcy, and Giseon Heo. Pseudo-multidimensional persistence and its applications. In *Research in Computational Topology*, pages 179–202. Springer, 2018.

[18] Leo Betthauser, Peter Bubenik, and Parker B Edwards. Graded persistence diagrams and persistence landscapes. *Discrete & Computational Geometry*, 67(1):203–230, 2022.

[19] Dhananjay Bhaskar, Angelika Manhart, Jesse Milzman, John T. Nardini, Kathleen M. Storey, Chad M. Topaz, and Lori Ziegelmeier. Analyzing collective motion with machine learning and topology. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(12):123125, 2019.

[20] Dhananjay Bhaskar, Angelika Manhart, Jesse Milzman, John T Nardini, Kathleen M Storey, Chad M Topaz, and Lori Ziegelmeier. Analyzing collective motion with machine learning and topology. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(12):123125, 2019.

[21] S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, and C. Landi. Multidimensional size functions for shape comparison. *Journal of Mathematical Imaging and Vision*, 32(2):161–179, 2008.

[22] Silvia Biasotti, Andrea Cerri, Patrizio Frosini, and Daniela Giorgi. A new algorithm for computing the 2-dimensional matching distance between size functions. *Pattern Recognition Letters*, 32(14):1735–1746, 2011.

[23] Håvard Bakke Bjerkevik. On the stability of interval decomposable persistence modules. *Discrete & Computational Geometry*, pages 1–30, 2021.

[24] Håvard Bakke Bjerkevik and Magnus Bakke Botnan. Computational complexity of the interleaving distance. *arXiv preprint arXiv:1712.04281*, 2017.

[25] Håvard Bakke Bjerkevik, Magnus Bakke Botnan, and Michael Kerber. Computing the interleaving distance is NP-hard. *Foundations of Computational Mathematics*, pages 1–35, 2019.

[26] Håvard Bakke Bjerkevik and Michael Kerber. Asymptotic improvements on the exact matching distance for 2-parameter persistence. *arXiv preprint arXiv:2111.10303*, 2021.

[27] Håvard Bakke Bjerkevik and Michael Lesnick. ℓ^p-distances on multiparameter persistence modules. *arXiv preprint arXiv:2106.13589*, 2021.

[28] Benjamin Blanchette, Thomas Brüstle, and Eric J. Hanson. Homological approximations in persistence theory. *Canadian Journal of Mathematics*, page 1–38, 2022.

[29] Andrew J Blumberg, Itamar Gal, Michael A Mandell, and Matthew Pancia. Robust statistics, hypothesis testing, and confidence intervals for persistent homology on metric measure spaces. *Foundations of Computational Mathematics*, 14(4):745–789, 2014.

[30] Andrew J Blumberg and Michael Lesnick. Universality of the homotopy interleaving distance. *arXiv preprint arXiv:1705.01690*, 2017.

[31] Andrew J Blumberg and Michael Lesnick. Stability of 2-parameter persistent homology. *arXiv preprint arXiv:2010.09628*, 2020.

[32] Magnus Botnan and William Crawley-Boevey. Decomposition of persistence modules. *Proceedings of the American Mathematical Society*, 148(11):4581–4596, 2020.

[33] Magnus Botnan and Michael Lesnick. Algebraic stability of zigzag persistence modules. *Algebraic & geometric topology*, 18(6):3133–3204, 2018.

[34] Magnus B Botnan, Vadim Lebovici, and Steve Oudot. Local characterizations for decomposability of 2-parameter persistence modules. *Algebras and Representation Theory*, pages 1–44, 2023.

[35] Magnus Bakke Botnan, Vadim Lebovici, and Steve Oudot. On Rectangle-Decomposable 2-Parameter Persistence Modules. In Sergio Cabello and Danny Z. Chen, editors, *36th International Symposium on Computational Geometry (SoCG 2020)*, volume 164 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 22:1–22:16, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[36] Magnus Bakke Botnan, Steffen Oppermann, and Steve Oudot. Signed barcodes for multi-parameter persistence via rank decompositions and rank-exact resolutions. *arXiv preprint arXiv:2107.06800*, 2021.

[37] Magnus Bakke Botnan, Steffen Oppermann, Steve Oudot, and Luis Scoccola. On the bottleneck stability of rank decompositions of multi-parameter persistence modules. *arXiv preprint arXiv:2208.00300*, 2022.

[38] Jacek Brodzki, Matthew Burfitt, and Mariam Pirashvili. On the complexity of zero-dimensional multiparameter persistence. *arXiv preprint arXiv:2008.11532*, 2020.

[39] Peter Bubenik, Jonathan Scott, and Donald Stanley. An algebraic wasserstein distance for generalized persistence modules. *arXiv preprint arXiv:1809.09654*, 2018.

[40] Mickaël Buchet and Emerson G Escolar. Realizations of indecomposable persistence modules of arbitrarily large dimension. *arXiv preprint arXiv:1803.05722*, 2018.

[41] Ryan Budney and Tomasz Kaczynski. Bi-filtrations and persistence paths for 2-morse functions. *arXiv preprint arXiv:2110.08227*, 2021.

[42] Theo Bühler. Exact categories. *Expositiones Mathematicae*, 28(1):1–69, 2010.

[43] Joshua A Bull, Philip S Macklin, Tom Quaiser, Franziska Braun, Sarah L Waters, Chris W Pugh, and Helen M Byrne. Combining multiple spatial statistics enhances the description of immune cell localisation within tumours. *Scientific reports*, 10(1):1–12, 2020.

[44] Chen Cai, Woojin Kim, Facundo Mémoli, and Yusu Wang. Elder-rule-staircodes for augmented metric spaces. *arXiv preprint arXiv:2003.04523*, 2020.

[45] G. Carlsson. Topology and data. *American Mathematical Society*, 46(2):255–308, 2009.

[46] G. Carlsson, T. Ishkhanov, V. De Silva, and A. Zomorodian. On the local behavior of spaces of natural images. *International Journal of Computer Vision*, 76(1):1–12, 2008.

[47] G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. *Discrete and Computational Geometry*, 42(1):71–93, 2009.

[48] Gunnar Carlsson and Vin De Silva. Zigzag persistence. *Foundations of computational mathematics*, 10(4):367–405, 2010.

[49] Gunnar Carlsson, Vin de Silva, and Dmitriy Morozov. Zigzag persistent homology and real-valued functions. In *Proceedings of the 25th Annual Symposium on Computational Geometry*, SCG '09, pages 247–256, New York, NY, USA, 2009. ACM.

[50] Gunnar Carlsson, Gurjeet Singh, and Afra J Zomorodian. Computing multidimensional persistence. *Journal of Computational Geometry*, 1(1):72–100, 2010.

[51] Mathieu Carrière and Andrew Blumberg. Multiparameter persistence images for topological machine learning. In *Conference on Neural Information Processing Systems*, 2020.

[52] Mathieu Carriere, Marco Cuturi, and Steve Oudot. Sliced wasserstein kernel for persistence diagrams. In *International conference on machine learning*, pages 664–673. PMLR, 2017.

[53] Nicholas J Cavanna, Kirk P Gardner, and Donald R Sheehy. When and why the topological coverage criterion works. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2679–2690. SIAM, 2017.

[54] A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, and C. Landi. Betti numbers in multidimensional persistent homology are stable functions. *Mathematical Methods in the Applied Sciences*, 36(12):1543–1557, 2013.

[55] Andrea Cerri, Marc Ethier, and Patrizio Frosini. The coherent matching distance in 2d persistent homology. In *International Workshop on Computational Topology in Image Context*, pages 216–227. Springer, 2016.

[56] Andrea Cerri, Marc Ethier, and Patrizio Frosini. On the geometrical properties of the coherent matching distance in 2d persistent homology. *Journal of Applied and Computational Topology*, 3(4):381–422, 2019.

[57] Wojciech Chacholski, Andrea Guidolin, Isaac Ren, Martina Scolamiero, and Francesca Tombari. Effective computation of relative homological invariants for functors over posets. *arXiv preprint arXiv:2209.05923*, 2022.

[58] Wojciech Chachólski, Martina Scolamiero, and Francesco Vaccarino. Combinatorial presentation of multidimensional persistent homology. *Journal of Pure and Applied Algebra*, 221(5):1055–1075, 2017.

[59] Joseph Minhow Chan, Gunnar Carlsson, and Raul Rabadan. Topology of viral evolution. *Proceedings of the National Academy of Sciences*, 110(46):18566–18571, 2013.

[60] F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Foundations of Computational Mathematics*, pages 1–19, 2011.

[61] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J. Guibas, and Steve Y. Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the 25th Annual Symposium on Computational Geometry*, SCG '09, pages 237–246, New York, NY, USA, 2009. ACM.

[62] Frédéric Chazal, David Cohen-Steiner, Leonidas J. Guibas, Facundo Mémoli, and Steve Y. Oudot. Gromov–Hausdorff stable signatures for shapes using persistence. In *Proceedings of the Symposium on Geometry Processing*, SGP '09, pages 1393–1403, Aire-la-Ville, Switzerland, Switzerland, 2009. Eurographics Association.

[63] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The Structure and Stability of Persistence Modules*. Springer International Publishing, 2016.

[64] Frédéric Chazal, Vin De Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173(1):193–214, 2014.

[65] Yu-Min Chung, Sarah Day, and Chuan-Shen Hu. A multi-parameter persistence framework for mathematical morphology. *arXiv preprint arXiv:2103.13013*, 2021.

[66] Jérémy Cochoy and Steve Oudot. Decomposition of exact pfd persistence bimodules. *Discrete & Computational Geometry*, 63(2):255–293, 2020.

[67] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007.

[68] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Extending persistence using poincaré and lefschetz duality. *Foundations of Computational Mathematics*, 9(1):79–103, 2009.

[69] René Corbet, Ulderico Fugacci, Michael Kerber, Claudia Landi, and Bei Wang. A kernel for multi-parameter persistent homology. *Computers & Graphics: X*, 2:100005, 2019.

[70] René Corbet and Michael Kerber. The representation theorem of persistence revisited and generalized. *Journal of Applied and Computational Topology*, 2(1):1–31, 2018.

[71] René Corbet, Michael Kerber, Michael Lesnick, and Georg Osang. Computing the multicover bifiltration. *To appear in the 37th International Symposium on Computational Geometry (SoCG 2021). arXiv preprint arXiv:2103.07823*, 2021.

[72] William Crawley-Boevey. Locally finitely presented additive categories. *Communications in Algebra*, 22(5):1641–1674, 1994.

[73] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications*, 14(05):1550066, 2015.

[74] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and Its Applications*, 14(05):1550066, 2015.

[75] Justin Curry. *Sheaves, Cosheaves and Applications*. PhD thesis, University of Pennsylvania, December 2014.

[76] M. d'Amico, P. Frosini, and C. Landi. Natural pseudo-distance and optimal matching between reduced size functions. *Acta applicandae mathematicae*, 109(2):527–554, 2010.

[77] Vin De Silva, Elizabeth Munch, and Amit Patel. Categorified reeb graphs. *Discrete & Computational Geometry*, 55(4):854–906, 2016.

[78] Harm Derksen and Jerzy Weyman. Quiver representations. *Notices of the AMS*, 52(2):200–206, 2005.

[79] The RIVET Developers. Rivet, version 1.1. https://github.com/rivetTDA/, 2020.

[80] Tamal K Dey and Cheng Xin. Computing bottleneck distance for multiparameter interval decomposable persistence modules. *arXiv preprint arXiv:1803.02869*, 2018.

[81] Tamal K. Dey and Cheng Xin. Generalized persistence algorithm for decomposing multiparameter persistence modules. *Journal of Applied and Computational Topology*, 2022.

[82] T.K. Dey and Y. Wang. *Computational Topology for Data Analysis*. Cambridge University Press, 2022.

[83] Ju A Drozd. Tame and wild matrix problems. In *Representation theory II*, pages 242–258. Springer, 1980.

[84] H. Edelsbrunner and J. Harer. *Computational topology: an introduction*. American Mathematical Society, 2010.

[85] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, 2002.

[86] Herbert Edelsbrunner and Dmitriy Morozov. Persistent homology. In *Handbook of Discrete and Computational Geometry*, pages 637–661. Chapman and Hall/CRC, 2017.

[87] Herbert Edelsbrunner and Georg Osang. A simple algorithm for higher-order delaunay mosaics and alpha shapes. *arXiv preprint arXiv:2011.03617*, 2020.

[88] Herbert Edelsbrunner and Georg Osang. The multi-cover persistence of euclidean balls. *Discrete & Computational Geometry*, 65(4):1296–1313, 2021.

[89] Alon Efrat, Alon Itai, and Matthew J Katz. Geometry helps in bottleneck matching and related problems. *Algorithmica*, 31(1):1–28, 2001.

[90] D. Eisenbud. *Commutative algebra with a view toward algebraic geometry*. Springer, 1995.

[91] David Eisenbud. *The geometry of syzygies: a second course in algebraic geometry and commutative algebra*, volume 229. Springer Science & Business Media, 2005.

[92] Burçin Eröcal, Oleksandr Motsak, Frank-Olaf Schreyer, and Andreas Steenpaß. Refined algorithms to compute syzygies. *Journal of Symbolic Computation*, 74:308–327, 2016.

[93] Emerson G Escolar and Yasuaki Hiraoka. Persistence modules on commutative ladders of finite type. *Discrete & Computational Geometry*, 55(1):100–157, 2016.

[94] Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan, Aarti Singh, et al. Confidence sets for persistence diagrams. *The Annals of Statistics*, 42(6):2301–2339, 2014.

[95] Jean Charles FaugÃšre. A New Efficient Algorithm for Computing Gröbner Bases Without Reduction to Zero (F5). In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, ISSAC '02, pages 75–83, New York, NY, USA, 2002. ACM.

[96] Wolf H Fridman, Laurence Zitvogel, Catherine Sautès-Fridman, and Guido Kroemer. The immune contexture in cancer prognosis and treatment. *Nature reviews Clinical oncology*, 14(12):717–734, 2017.

[97] P. Frosini and M. Mulazzani. Size homotopy groups for computation of natural size distances. *Bulletin of the Belgian Mathematical Society Simon Stevin*, 6(3):455–464, 1999.

[98] Ulderico Fugacci and Michael Kerber. Chunk reduction for multiparameter persistent homology. In *35th International Symposium on Computational Geometry (SoCG 2019)*, volume 129, pages 37:1–37:14, 2019.

[99] P Gabriel and AV Roiter. Representations of finite-dimensional algebras, with a chapter by b. keller, encyclopedia math. sci. 73, algebra viii, 1–177, 1992.

[100] Peter Gabriel. Unzerlegbare darstellungen i. *Manuscripta mathematica*, 6(1):71–103, 1972.

[101] Oliver Gäfvert. *Topological and geometrical methods in data analysis*. PhD thesis, KTH Royal Institute of Technology, 2021.

[102] Oliver Gäfvert and Wojciech Chachólski. Stable invariants for multiparameter persistence. *arXiv preprint arXiv:1703.03632*, 2017.

[103] Jérôme Galon, Anne Costes, Fatima Sanchez-Cabo, Amos Kirilovsky, Bernhard Mlecnik, Christine Lagorce-Pagès, Marie Tosolini, Matthieu Camus, Anne Berger, Philippe Wind, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*, 313(5795):1960–1964, 2006.

[104] The GAP Group. *GAP – Groups, Algorithms, and Programming, Version 4.11.1*, 2021.

[105] Richard J Gardner, Erik Hermansen, Marius Pachitariu, Yoram Burak, Nils A Baas, Benjamin A Dunn, May-Britt Moser, and Edvard I Moser. Toroidal topology of population activity in grid cells. *Nature*, pages 1–6, 2022.

[106] Barbara Giunti and Jānis Lazovskis. TDA-Applications (an online database of papers on applications of TDA outside of math). https://www.zotero.org/groups/2425412/tda-applications, 2021.

[107] Barbara Giunti, John S Nolan, Nina Otter, and Lukas Waas. Amplitudes in abelian categories. *arXiv preprint arXiv:2107.09036*, 2021.

[108] Chad Giusti, Robert Ghrist, and Danielle S Bassett. Two's company, three (or more) is a simplex. *Journal of computational neuroscience*, 41(1):1–14, 2016.

[109] Gert-Martin Greuel and Gerhard Pfister. *A Singular introduction to commutative algebra*. Springer Science & Business Media, 2012.

[110] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces (λ-coalescent measure trees). *Probability Theory and Related Fields*, 145(1-2):285–322, 2009.

[111] Heather A Harrington, Nina Otter, Hal Schenck, and Ulrike Tillmann. Stratifying multiparameter persistent homology. *SIAM Journal on Applied Algebra and Geometry*, 3(3):439–471, 2019.

[112] Joris Hemelaar, Ramyiadarsini Elangovan, Jason Yun, Leslie Dickson-Tetteh, Isabella Fleminger, Shona Kirtley, Brian Williams, Eleanor Gouws-Williams, Peter D Ghys, Abimiku Alash'le G, et al. Global and regional molecular epidemiology of hiv-1, 1990–2015: a systematic review, global survey, and trend analysis. *The Lancet infectious diseases*, 19(2):143–155, 2019.

[113] Felix Hensel, Michael Moor, and Bastian Rieck. A survey of topological machine learning methods. *Frontiers in Artificial Intelligence*, 4:52, 2021.

[114] Yasuaki Hiraoka, Takenobu Nakamura, Akihiko Hirata, Emerson G Escolar, Kaname Matsue, and Yasumasa Nishiura. Hierarchical structures of amorphous solids characterized by persistent homology. *Proceedings of the National Academy of Sciences*, 113(26):7035–7040, 2016.

[115] Derek F Holt. The meataxe as a tool in computational group theory. *London Mathematical Society Lecture Note Series*, pages 74–81, 1998.

[116] Jiuzhao Hua. Counting representations of quivers over finite fields. *Journal of Algebra*, 226(2):1011–1033, 2000.

[117] Svante Janson. On the Gromov-Prohorov distance. *arXiv preprint arXiv:2005.13505*, 2020.

[118] Victor G Kac. Infinite root systems, representations of graphs and invariant theory, ii. *Journal of algebra*, 78(1):141–162, 1982.

[119] Dilan Karagüler. A survey on multidimensional persistence theory. Master's thesis, Middle East Technical University, 2021.

[120] Masaki Kashiwara and Pierre Schapira. Persistent homology and microlocal sheaf theory. *Journal of Applied and Computational Topology*, 2(1):83–113, 2018.

[121] Jakob Nikolas Kather, Meggy Suarez-Carmona, Pornpimol Charoentong, Cleo-Aron Weis, Daniela Hirsch, Peter Bankhead, Marcel Horning, Dyke Ferber, Ivan Kel, Esther Herpel, et al. Topography of cancer-associated immune cells in human solid tumors. *Elife*, 7:e36967, 2018.

[122] Bryn Keller, Michael Lesnick, and Theodore L Willke. Persistent homology for virtual screening. *ChemRxiv*, 2018.

[123] Michael Kerber, Michael Lesnick, and Steve Oudot. Exact computation of the matching distance on 2-parameter persistence modules. In *35th International Symposium on Computational Geometry (SoCG 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[124] Michael Kerber, Dmitriy Morozov, and Arnur Nigmetov. Geometry helps to compare persistence diagrams. *Journal of Experimental Algorithmics (JEA)*, 22:1–4, 2017.

[125] Michael Kerber and Arnur Nigmetov. Efficient approximation of the matching distance for 2-parameter persistence. In *36th International Symposium on Computational Geometry (SoCG 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

[126] Michael Kerber and Alexander Rolle. Fast minimal presentations of bigraded persistence modules. In *2021 Proceedings of the Workshop on Algorithm Engineering and Experiments (ALENEX)*, pages 207–220. SIAM, 2021.

[127] Woojin Kim and Facundo Memoli. Generalized persistence diagrams for persistence modules over posets. *arXiv preprint arXiv:1810.11517*, 2018.

[128] Woojin Kim and Facundo Mémoli. Spatiotemporal persistent homology for dynamic metric spaces. *Discrete & Computational Geometry*, 66(3):831–875, 2021.

[129] Woojin Kim and Samantha Moore. Bigraded betti numbers and generalized persistence diagrams. *arXiv preprint arXiv:2111.02551*, 2021.

[130] Roberto La Scala and Michael Stillman. Strategies for computing minimal free resolutions. *Journal of Symbolic Computation*, 26(4):409–431, 1998.

[131] C. Landi. The rank invariant stability via interleavings. In *Research in Computational Topology*, pages 1–10. Springer International Publishing, 2018.

[132] Yongjin Lee, Senja D Barthel, Paweł Dłotko, S Mohamad Moosavi, Kathryn Hess, and Berend Smit. Quantifying similarity of pore-geometry in nanoporous materials. *Nature communications*, 8(1):1–8, 2017.

[133] Michael Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics*, 15(3):613–650, 2015.

[134] Michael Lesnick and Matthew Wright. Interactive visualization of 2-d persistence modules. *arXiv preprint arXiv:1512.00180*, 2015.

[135] Michael Lesnick and Matthew Wright. Computing minimal presentations and bigraded betti numbers of 2-parameter persistent homology. *arXiv preprint arXiv:1902.05708*, 2019.

[136] Michael Lesnick and Roy Zhao. Computing the degree-Rips bifiltration. In preparation.

[137] Zbigniew Leszczyński. On the representation type of tensor product algebras. *Fund. Math.*, 144(2):143–161, 1994.

[138] Zbigniew Leszczyński and Andrzej Skowroński. Tame triangular matrix algebras. *Colloq. Math.*, 86(2):259–303, 2000.

[139] Alexander McCleary and Amit Patel. Edit distance and persistence diagrams over lattices. *arXiv preprint arXiv:2010.07337*, 2020.

[140] Ezra Miller. Fruit flies and moduli: interactions between biology and mathematics. *Notices of the AMS*, 62(10):1178–1184, 2015.

[141] Ezra Miller. Essential graded algebra over polynomial rings with real exponents. *arXiv preprint arXiv:2008.03819*, 2020.

[142] Ezra Miller. Homological algebra of modules over posets. *arXiv preprint arXiv:2008.00063*, 2020.

[143] Ezra Miller. Primary decomposition over partially ordered groups. *arXiv preprint arXiv:2008.00093*, 2020.

[144] Ezra Miller and Bernd Sturmfels. *Combinatorial commutative algebra*, volume 227. Springer Science & Business Media, 2004.

[145] N. Milosavljevic, D. Morozov, and P. Skraba. Zigzag Persistent Homology in Matrix Multiplication Time. *Proceedings of the Annual Symposium on Computational Geometry*, 2011.

[146] Dmitriy Morozov. *Homological illusions of persistence and stability*. Duke University, 2008.

[147] Sidra Nawaz, Andreas Heindl, Konrad Koelble, and Yinyin Yuan. Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. *Modern Pathology*, 28(6):766–777, 2015.

[148] Liudmila A Nazarova. Representations of quivers of infinite type. *Mathematics of the USSR-Izvestiya*, 7(4):749, 1973.

[149] Duc Duy Nguyen, Zixuan Cang, Kedi Wu, Menglun Wang, Yin Cao, and Guo-Wei Wei. Mathematical deep learning for pose and binding affinity prediction and ranking in d3r grand challenges. *Journal of computer-aided molecular design*, 33(1):71–82, 2019.

[150] Reza Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on automatic control*, 51(3):401–420, 2006.

[151] Georg Osang. Rhomboid tiling and order-k delaunay mosaics. https://github.com/geoo89/rhomboidtiling, 2020.

[152] Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6:1–38, 2017.

[153] Steve Oudot and Luis Scoccola. On the stability of multigraded betti numbers and hilbert functions. *arXiv preprint arXiv:2112.11901*, 2021.

[154] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, 2015.

[155] Amit Patel. Generalized persistence diagrams. *Journal of Applied and Computational Topology*, 1(3):397–419, 2018.

[156] Irena Peeva. *Graded syzygies*. Springer-Verlag, 2011.

[157] Jose A Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. *Foundations of Computational Mathematics*, 15(3):799–838, 2015.

[158] Raúl Rabadán and Andrew J Blumberg. *Topological data analysis for genomics and evolution: topology in biology*. Cambridge University Press, 2019.

[159] Claus Michael Ringel. Representation theory of dynkin quivers. three contributions. *Frontiers of Mathematics in China*, 11(4):765–814, 2016.

[160] Alexander Rolle and Luis Scoccola. Stable and consistent density-based clustering. *arXiv preprint arXiv:2005.09048*, 2020.

[161] Gian-Carlo Rota. On the foundations of combinatorial theory i. theory of möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2(4):340–368, 1964.

[162] Erik Rybakken, Nils Baas, and Benjamin Dunn. Decoding of neural data using cohomological feature extraction. *Neural computation*, 31(1):68–93, 2019.

[163] Hal Schenck. Mathematical foundations for topological data analysis. 2022.

[164] Yair Schiff, Vijil Chenthamarakshan, Karthikeyan Natesan Ramamurthy, and Payel Das. Characterizing the latent space of molecular deep generative models with persistent homology metrics, 2021.

[165] Frank-Olaf Schreyer. *Die berechnung von syzygien mit dem verallgemeinerten weierstraßschen divisionssatz und eine anwendung auf analytische cohen-macaulay stellenalgebren minimaler multiplizität.* PhD thesis, 1980.

[166] Martina Scolamiero, Wojciech Chachólski, Anders Lundman, Ryan Ramanujam, and Sebastian Öberg. Multidimensional persistence and noise. *Foundations of Computational Mathematics*, 17(6):1367–1406, 2017.

[167] Donald R Sheehy. A multicover nerve for geometric inference. In *CCCG*, pages 309–314, 2012.

[168] Donald R Sheehy. Linear-size approximations to the Vietoris–Rips filtration. *Discrete & Computational Geometry*, 49(4):778–796, 2013.

[169] Daniel Simson and Andrzej Skowronski. *Elements of the Representation Theory of Associative Algebras: Volume 3, Representation-infinite Tilted Algebras*, volume 72. Cambridge University Press, 2007.

[170] Ann E Sizemore, Chad Giusti, Ari Kahn, Jean M Vettel, Richard F Betzel, and Danielle S Bassett. Cliques and cavities in the human connectome. *Journal of computational neuroscience*, 44(1):115–145, 2018.

[171] Primoz Skraba and Katharine Turner. Wasserstein stability for persistence diagrams. *arXiv preprint arXiv:2006.16824*, 2020.

[172] Jacek Skryzalin. Numeric invariants from multidimensional persistence. *Ph.D. Dissertation, Stanford University*, 2016.

[173] Ashleigh Linnea Thomas. *Invariants and Metrics for Multiparameter Persistent Homology.* PhD thesis, Duke University, 2019.

[174] John Toner and Yuhai Tu. Flocks, herds, and schools: A quantitative theory of flocking. *Physical review E*, 58(4):4828, 1998.

[175] Katharine Turner, Yuriy Mileyko, Sayan Mukherjee, and John Harer. Fréchet means for distributions of persistence diagrams. *Discrete & Computational Geometry*, 52(1):44–70, 2014.

[176] Oliver Vipond. Multiparameter persistence landscapes. https://github.com/OliverVipond/Multiparameter_Persistence_Landscapes.

[177] Oliver Vipond. Multiparameter persistence landscapes. *Journal of Machine Learning Research*, 21(61):1–38, 2020.

[178] Oliver Vipond, Joshua A Bull, Philip S Macklin, Ulrike Tillmann, Christopher W Pugh, Helen M Byrne, and Heather A Harrington. Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors. *Proceedings of the National Academy of Sciences*, 118(41), 2021.

[179] Cary Webb. Decomposition of graded modules. *Proceedings of the American Mathematical Society*, 94(4):565–571, 1985.

[180] Hassler Whitney. On singularities of mappings of euclidean spaces. i. mappings of the plane into the plane. *Annals of Mathematics*, pages 374–410, 1955.

[181] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete and Computational Geometry*, 33(2):249–274, 2005.