OPTIMIZING FUNCTIONS BASED ON SIGNED BARCODES

Differentiability of functions based on signed barcodes arising from multiparameter persistence modules

A thesis submitted to attain the degree of

MASTER OF SCIENCES

(MSc ETH Zürich Mathematics)

presented by

Siddharth Setlur

supervised by

Dr. Sara Kališnik Hintz (ETH Zürich)

Dr. Steve Oudot (Inria Saclay & École Polytechnique)

October 4, 2023

ETH zürich

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

OPTIMIZING FUNCTIONS BASED ON SIGNED BARCODES. DIFFERENTIABILITY OF FUNCTIONS BASED ON SIGNED BARCODES ARISING FROM MULTIPARAMETER PERSISTENCE MODULES.

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
SETLUR	SIDDHARTH

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Zürich, 25/08/2023	Siddh

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Abstract

Single parameter persistent homology is a tool that captures the underlying topological features of a data set by analyzing how its topology varies along a single parameter filtration function. Leveraging the full power of persistence involves incorporating it into loss functions in machine learning algorithms. This has proven wildly successful in a wide range of domains and has motivated the study of multiparameter persistence, a generalization that allows for tracking the evolution of topology over multiple filtration functions rather than a single one. Although multiparameter persistence captures more information than single parameter persistence, the representation theory of multiparameter persistence modules is wild and so in contrast to the single parameter setting there is no complete invariant. However, due to the potential of multiparameter persistence, a large body of work has been devoted to studying incomplete invariants. One promising invariant is the recently introduced signed barcode, which is obtained via a signed decomposition of the rank invariant of a persistence module. The main contribution of this thesis is a novel result that proves that stochastic gradient descent converges almost surely on definable, locally Lipschitz loss functions on signed barcodes. This enables signed barcodes to be incorporated into machine learning algorithms, just like their single parameter counterpart. In order to make the thesis as self-contained as possible, we provide a brief review of single parameter persistence as well as a slightly altered proof that stochastic gradient descent converges on loss functions based on standard persistence barcodes. We also provide an overview of the theory of rank decompositions and signed barcodes.

Acknowledgements

I am fortunate to have been advised by Dr. Sara Kališnik Hintz and Dr. Steve Oudot. It was Sara, who introduced me to the fascinating world of topological data analysis and supervised my first foray into mathematical writing in the form of my BSc thesis. Her mentorship over the past two years has been invaluable. I was then able to continue working in TDA as a research intern under Steve's supervision. I learned a tremendous amount from Steve during my semester in Paris. It was my first experience working on original research and Steve's guidance on the multiple occasions when I was stuck on a problem was extremely helpful.

Finally, I would like to thank all of my friends and of course, my parents who have supported me throughout.

Contents

Important Symbols

$H_n(X)$	n-dimensional homology of a topological space X over a field k
$\mathbf{P}$	Poset category
$\mathbf{Int}(\mathbf{P})$	Collection of *all* intervals of $\mathbf{P}$
$\mathcal{I}$	Collection of intervals
K	Simplicial complex
σ	Simplex in a simplicial complex
$VR(X, r)$	Vietoris Rips (VR) complex of a point cloud X at radious $r \in \mathbb{R}_{\geq 0}$
Δ_n	Standard n simplex
Vec	Category of vector spaces
vec	Category of finite dimensional vector spaces
$\mathcal{M}$	Persistence module
k_I	Interval module supported on the interval I
$\phi : K \to \mathbb{R}$	1-critical filtration of a simplicial complex K
$\mathbb{R}^K$	Free vector space whose basis is generated by simplices of K
Pers	Persistence map that assigns to a point cloud its persistence diagram
$\Phi : A \to \mathbb{R}^K$	Parameterized family of filtrations
Filt_K	Subspace of vectors in $\mathbb{R}^K$ that define a filtration of K
filt_K	Subspace of vectors in $\mathbb{N}^K$ that define a filtration of K
$\mathrm{Rk}_{\mathcal{I}}\mathcal{M}$	Rank invariant of a persistence module $\mathcal{M}$ over a collection of intervals $\mathcal{I}$
$\langle s, t \rangle$	Rectangle or segment spanned by $s \leq t$
$(\mathcal{R}, \mathcal{S})$	Rank decomposition
$\mathrm{Hil}(\mathcal{M})$	Hilbert function of a persistence module $\mathcal{M}$

Chapter 1

Introduction

Over the past two decades, topological data analysis (TDA) has emerged as a powerful new tool that leverages tools from algebraic topology to study the underlying "shape of data".

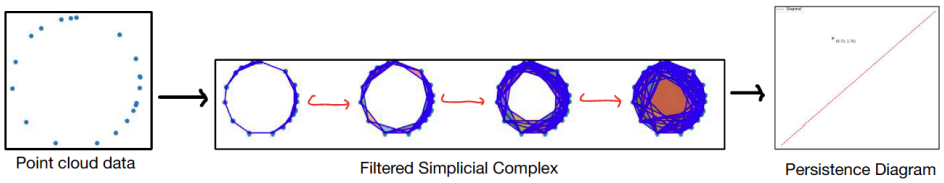


Figure 1.1: The single parameter persistence pipeline.

Classical or single parameter persistent homology was developed in the early 2000's by Carlsson, Edelsbrunner, and Harer as an extension simplicial homology, which is a robust invariant that captures geometric features of simplicial complexes, to discrete data sets that one encounters in practice. It accomplishes this by constructing a filtered simplicial complex and tracks the evolution of simplical homology across the filtration. This information is encoded in a *persistence module*, which can then be characterized up to isomorphism by a complete invariant known as a *persistence barcode or diagram*. The first successful applications of TDA involved using the persistence diagram to infer geometric characteristics of the underlying data. We point the reader to [CVJ21] for a comprehensive survey of single parameter persistence and the first significant applications. For sake of completeness, we provide a brief overview of single parameter persistence in Chapter 2.

Although the pipeline depicted in Figure 1.1 proved effective in multiple domains, it became apparent that incorporating persistence diagrams into machine learning algorithms would unlock a wide range of new applications. Multiple works such as [GHO16] and [PSO18a] began defining functions on the space of persistence diagrams and sought to optimize them, i.e. find minima or maxima. These first papers were tailored to their respective applications, namely physics and shape matching. While the language and hypotheses used in each setting was different, the key conclusion that both papers arrived at is that the pipeline depicted in Figure 3.2 is differentiable. [Car+21] formalized this notion and used classical results from o-minimal theory to prove that the map associating a point cloud to its persistence diagram is differentiable almost everywhere. Moreover, they proved that the map fulfilled the conditions set out in [Dav+20] and were thus able to guarantee the convergence of stochastic gradient descent on loss functions defined on persistence diagrams. Since the space of persistence diagrams comes equipped with a metric,

most applications use the distance to a fixed persistence diagram as a loss function. Intuitively, one can design a loss function that takes a point cloud as input and moves its points to form a desired shape by encouraging the formation of cycles. [San+23] and [DCL22] are two examples of successful real world applications of this framework. In Chapter 3, we give a slightly altered version of the proof of the main result of [Car+21], namely that stochastic gradient descent converges on loss functions defined on persistence diagrams.

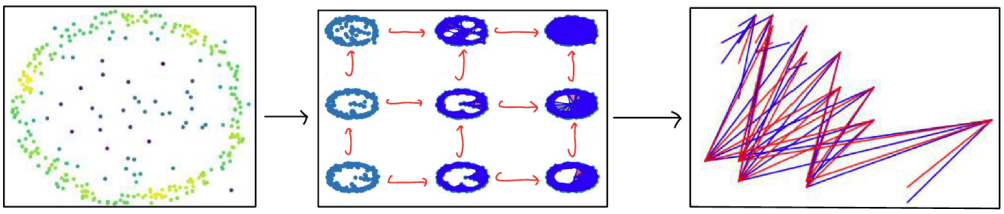


Figure 1.2: The multiparameter persistence pipeline.

Single parameter persistence is now reasonably well-understood and software packages such as [EB] and [Bau21] among others have made applying it relatively easy. Consequently, a significant portion of current research in persistence revolves around *multiparameter persistence*. Often, data comes with multiple functions of interest that one could use to construct a filtered simplicial complex. However, single parameter persistence forces the user to pick one of these functions and ignore the others. In order to address this, Carlsson and Zomorodian introduced multiparameter persistence in [CZ09]. Unfortunately, it turns out that when simplical homology is applied to a multifiltered simplicial complex, such as the one depicted in the second step of Figure 1.2, the resultant persistence modules cannot be classified up to isomorphism like the single parameter persistence modules. In other words, there is no complete invariant analogous to the persistence diagram in Figure 1.1. One of the central problems is to define incomplete invariants in the multiparameter setting that share some of the same desirable properties. One such invariant is the *signed barcode* which was introduced by Botnan, Oppermann, and Oudot in [BOO22]. Intuitively, the persistence barcode in the single parameter case is equivalent to expressing the rank invariant of a persistence module as a sum of rank invariants of indicator modules supported on intervals. [BOO22] generalizes this to the multiparameter setting by allowing negative signs in the sum, i.e. the signed barcode is equivalent to expressing the rank invariant of a multiparameter persistence module as a $\mathbb{Z}$-linear combination of indicator modules supported on intervals. This results in the multiparameter persistence pipeline depicted in Figure 1.2. In Chapter 4, we follow [BOO22] and review the theory of rank decompositions and signed barcodes.

Just as with the single parameter persistence pipeline, we would like to incorporate loss functions based on signed barcodes into machine learning algorithms. This is the content of Chapter 5, which contains two novel mathematical results. We begin by proving that a coarser invariant, namely the Betti signed barcode, of multiparameter persistence modules is differentiable almost everywhere. This result is useful in its own right, since computing signed barcodes is, at least at the time of writing, rather slow. Computing the Betti signed barcode on the other hand is much faster, since it is derived from decompositions of the Hilbert function of a persistence module and is a strictly coarser invariant. A similar approach can then be used to prove that the map depicted in Figure 1.2 is differentiable almost everywhere. Due to the results in [Dav+20], we can then conclude that stochastic gradient descent converges on appropriate loss functions defined on signed barcodes.

Chapter 2

A Brief Review of Single Parameter Persistence

In its most common form, persistent homology aims to use simplicial homology to study the underlying geometric features of point cloud data. It does so by producing a persistence diagram (or equivalently a barcode) for a given point cloud $X \subset \mathbb{R}^d$. Before delving into the full generality, we introduce this pipeline using an example. For a more comprehensive introduction to persistent homology, we refer the reader to [Car09], [EM13] or [Oud15].

The prototypical example used to demonstrate the persistent homology pipeline is the dataset $X \subset \mathbb{R}^2$ depicted in Figure 2.1. Visually, we see immediately that X appears to be sampled from a circle. Algebraic topology and in particular simplicial homology [Hat02] is the perfect tool to formalize this intuition. It is easy to compute, especially over $\mathbb{Z}_2$, and has many desirable properties such as functoriality and homotopy invariance.

Unfortunately, simply interpreting X as a set of 0-simplices and applying simplicial homology delivers the following underwhelming result

$$H_n(X) = \begin{cases} \mathbb{Z}_2^{\#\text{ of points}} & n = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Ideally, we would like to see a single connected component being captured by 0-dimensional

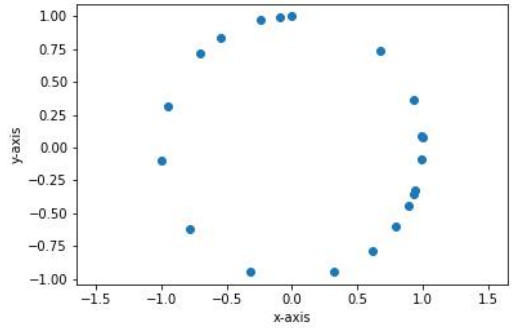


Figure 2.1: Data set X consisting of 20 points sampled from the unit circle.

homology and non-trivial 1-dimensional homology to reflect the circular geometry of X.

2.1 Filtrations and posets

Instead of naively applying simplicial homology to a point cloud X, we first construct a *filtered simplicial complex* using the points of X. The idea here is to track the evolution of topological features as we change a certain parameter. We can then discern important topological features of the data from those arising from noise by looking at which features *persist* through the filtration. In the 1-parameter case, one works with filtrations indexed over the posets $\mathbb{R}, \mathbb{Z}$ or $\mathbb{N}$. However, seeing as the main objective of this thesis is to work with multiparameter filtrations, we introduce the notion of filtrations indexed over arbitrary posets. Recall, first the definition of a poset category.

Definition 2.1 (Poset category). *A poset is a set $\boldsymbol{P}$ endowed with a partial order relation $\leq$ which satisfies the following axioms:*

1. *Reflexivity: $a \leq a \; \forall a \in \boldsymbol{P}$;*
2. *Antisymmetry: $a \leq b$ and $b \leq a \implies a = b$;*
3. *Transitivity: $a \leq b$ and $b \leq c \implies a \leq c$.*

If $a \leq b$ or $b \leq a$, we say that a and b are comparable.

A poset $\boldsymbol{P}$ can be interpreted as a category, whose objects are simply the elements of $\boldsymbol{P}$ with a unique morphism between $a, b \in \boldsymbol{P}$ if and only if $a \leq b$. We will denote this category by $\boldsymbol{P}$.

Definition 2.2 (Filtered simplical complex). *Let K be a finite simplicial complex and $\boldsymbol{P}$ be a poset. A $\boldsymbol{P}$-indexed* filtration *of K is an increasing sequence $(K_r)_{r \in \boldsymbol{P}}$ of subcomplexes of K with respect to inclusion, i.e. $K_r \subset K_s$ for all $r \leq s$ and $K = \bigcup_{r \in \mathbb{P}} K_r$. A filtration is 1-*critical *if for each $\sigma \in K$, there is a unique minimal $\phi_\sigma \in \mathbb{P}$ such that $\sigma \in K_r$ for all $r \geq \phi_\sigma$. We refer to ϕ_σ as the* birth index *of σ.*

Combining Definition 2.1 and Definition 2.2 leads to the following lemma.

Lemma 2.1 (Filtrations are functors). *A $\boldsymbol{P}$-indexed filtration of a simplicial complex K is a functor*

$$\begin{aligned} \boldsymbol{P} &\to \boldsymbol{Simp} \\ r &\mapsto K_r \\ r \leq s &\mapsto K_r \hookrightarrow K_s, \end{aligned}$$

such that K_r is a subcomplex of K for all r and $K = \cup_r K_r$.

There are multiple ways to construct a filtered simplicial complex from point cloud data, but the most common one is the *Vietoris Rips (VR) complex.*

Definition 2.3 (Vietoris Rips (VR) complex). *Let X be a finite subset of some metric space and $r \in \mathbb{R}_{\geq 0}$. Then, the* VR complex *$VR(X, r)$ of X at radius r is a simplicial complex defined as follows:*

1. *add one vertex for each point of X;*
2. *the n-simplex spanned by the vertices $x_0, \ldots, x_n$ is in $VR(X, r)$ if and only if $d(x_i, x_j) < r$ for all $0 \leq i, j \leq r$.*

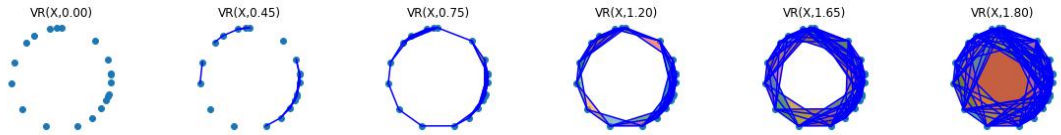


Figure 2.2: The 0-, 1-, and 2-simplicies of the VR complex of a noisy circle.

Clearly, $VR(X,r) \subset VR(X,s)$ *for all* $r \leq s$.

Intuitively, one begins with the points as vertices. An edge connecting two points is then added to $VR(X,r)$ if and only if the two points are within a distance r, the triangle spanning points x, y, and z is included if and only if all the points lie within r of each other, and so on.

Lemma 2.2 (VR complexes form filtrations). *Let* X *be a finite metric space composed of* n*-elements. Then sending* $r \to VR(X,r)$ *for all* $r \geq 0$ *and* $r \to \emptyset$ *for all* $r < 0$ *defines an* $\mathbb{R}$*-indexed filtration of* Δ_{n-1}*, which is the simplicial complex consisting of an* $n-1$*-simplex and all of its faces.*

Proof. The fact that the VR-complexes form a filtration is clear from the definition. If X is finite, then for $R = \max\{d(x,y)|x,y \in X\}$ we have $VR(X,R) = VR(X,r)$ for all $r \geq R$ since every pair of points is within R of each other and so no new simplices can be added. In particular, the VR-complexes define a finite filtration. Given that every pair of points is within R of each other, $VR(X,R)$ must be the simplicial complex containing all possible faces of a simplex with n vertices Δ_{n-1}. ■

The Vietoris-Rips complex of the dataset X from Figure 2.1 is illustrated in Figure 2.2. Note that since we are solely interested in 1-dimensional persistent homology, only the 1- and 2-simplices are displayed. The central topological feature of the data - the central circle - *persists* through a significant portion of the filtration.

2.2 Persistence Modules

To track the evolution of topological features across a filtration, we apply the n-dimensional homology functor (over an arbitrary but fixed field k) to each subcomplex in the filtration. In the 1-parameter case, where $\mathbf{P} = \mathbb{N}$, we obtain a sequence of vector spaces

$$H_n(K_0) \to H_n(K_1) \to \cdots \to H_n(K),$$

where the maps are inclusion induced. This is an example of a persistence module.

Definition 2.4 (Persistence modules). *A* persistence module *is a functor* $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{Vec}$ *from an arbitrary poset to the category of vector spaces. We will restrict ourselves to functors to the category* ***vec*** *of* finite dimensional *vector spaces, which are known as* pointwise finite dimensional *persistence modules.*

Remark 2.1 (Persistence modules as compositions). As hinted at above, persistence modules are often the composition of two functors. The first functor takes us from the poset to the category of topological spaces, or as is often the case in practice, to the subcategory of simplicial complexes. We then apply the homology functor over a fixed field to end up in the category of vector spaces.

Remark 2.2 (Persistence modules are a functor category). Let $\mathbf{P}$ be a poset. Then the category of persistence modules from $\mathbf{P}$ is a *functor category*. As alluded to by the name, the objects of this category are functors, in this case from $\mathbf{P}$ to **vec**. The morphisms are natural transformations. Moreover, since we land in the category of vector spaces, we can talk about direct sums of persistence modules.

For details regarding functor categories and natural transformations, we point the reader to [ML78]. Instead, we present a relevant example.

Example 1 (Morphisms of 1-parameter persistence modules). Let $\mathcal{M}, \mathcal{N} : \mathbb{R} \to \mathbf{vec}$ be persistence modules. Then, a natural transformation $\eta : \mathcal{M} \to \mathcal{N}$ is a sequence of maps $\eta_r : \mathcal{M}(r) \to \mathcal{N}(r)$ for all $r \in \mathbb{R}$ such that the following diagram commutes

$$\begin{array}{ccccccccc}
\ldots & \longrightarrow & \mathcal{M}(r) & \longrightarrow & \mathcal{M}(s) & \longrightarrow & \mathcal{M}(t) & \longrightarrow & \ldots \\
 & & \downarrow{\scriptstyle \eta_r} & & \downarrow{\scriptstyle \eta_s} & & \downarrow{\scriptstyle \eta_t} & & \\
\ldots & \longrightarrow & \mathcal{N}(r) & \longrightarrow & \mathcal{N}(s) & \longrightarrow & \mathcal{N}(t) & \longrightarrow & \ldots
\end{array}$$

for all $r \leq s \leq t$. If each η_r is an isomorphism, then we say $\mathcal{M} \cong \mathcal{N}$.

Definition 2.5 (Indecomposable persistence modules). *We say that a persistence module $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ is* indecomposable *if $\mathcal{M} \cong \mathcal{M}' \oplus \mathcal{M}''$ implies $\mathcal{M}' = 0$ or $\mathcal{M}'' = 0$.*

Applying 1-dimensional homology over $\mathbb{Z}_2$ to the filtration in Figure 2.2, we obtain

$$H_1(K_r) = \begin{cases} \mathbb{Z}_2 & 0.85 < r < 1.79 \\ 0 & \text{otherwise} \end{cases}.$$

This is a special kind of persistence module known as an *interval module*, since it is an indicator module supported on an interval. While the notion of an interval is clear in this setting, we take this opportunity to define intervals for arbitrary posets. This will be important when we discuss multiparameter persistence.

Definition 2.6 (Connected posets, locally finite posets, and intervals). *A poset $\boldsymbol{P}$ is* connected *if for every $p, q \in \boldsymbol{P}$, there exists a sequence*

$$p = p_0, p_1, \ldots, p_n = q$$

such that p_i and p_{i+1} are comparable *for all i.*

A poset $\boldsymbol{P}$ is locally finite *if for all $p \leq q \in P$, the set $[p, q] = \{r \in P \mid p \leq r \leq q\}$ is finite.*

A subset $I \subset \boldsymbol{P}$ is convex*, if $r, t \in I$ and $s \in \boldsymbol{P}$ with $r \leq s \leq t$, then $s \in I$.*

Let $\boldsymbol{P}$ be a poset, then a subset $I \subset \boldsymbol{P}$ is an interval*, if I is connected and convex. We denote by $\boldsymbol{Int}(\boldsymbol{P})$, the collection of all intervals of $\boldsymbol{P}$.*

For further details and examples of the objects in Definition 2.6, see section 2.3 of [Set].

Definition 2.7 (Interval modules). *Let $\boldsymbol{P}$ be a poset and $I \subset \boldsymbol{P}$ be an interval. Then the associated interval module is the functor*

$$k_I : \boldsymbol{P} \to \boldsymbol{vec}$$

$$k_I(p) = \begin{cases} k & p \in I \\ 0 & \textit{otherwise.} \end{cases}$$

with identity morphisms between two copies of k and zero morphisms to, from or between 0-spaces.

The identity morphisms between copies of k at every element of a convex and connected set make k_I indecomposable. Due to the structure theorem for 1-parameter persistence modules, interval modules are an important class of indecomposables.

Theorem 2.1 (Structure theorem for single parameter persistence modules). *Let $\mathcal{M} : \mathbb{R} \to \textbf{\textit{vec}}$ be a persistence module. Then, there exists a collection of intervals $\mathcal{I}$ in $\mathbb{R}$ such that*

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

This collection is unique and is called the barcode *of $\mathcal{M}$.*

In other words, Theorem 2.1 states that barcodes are a complete invariant of persistence modules. There are different (but equivalent) ways to prove Theorem 2.1 owing to the different ways one can view persistence modules. While we defined persistence modules as functors from posets, they were first viewed as finitely generated graded modules (see [ZC05]). In this case, the corresponding structure theorem from classical commutative algebra guaranteed to decomposition. Alternatively one can view persistence modules as *quiver representations* and under this framework, single parameter persistence modules are representations of A_n quivers (see [Oud15]). A classical theorem from representation theory - Gabriel's theorem - then delivers a similar decomposition. Unfortunately, Gabriel's theorem also tells us that there is no hope for a decomposition into interval modules for other indexing posets such as $\mathbb{R}^2$ or even $\mathbb{Z}^2$. For a detailed exposition on the different perspectives and the proofs of the decomposition under each framework see [Set].

2.3 Persistence Barcodes and Diagrams

The barcode provides two equivalent visual summaries that are used in practice to interpret the topological features of the data. The barcode is obtained by simply plotting the intervals obtained from the decomposition. In this case, each bar corresponds to the lifespan of a cycle and hence longer bars represent features that characterize the geometry of the data while shorter bars may be attributed to noise.

Alternatively, given an interval $I = (a, b)$ in the barcode, one can plot the point $(a, b) \in \mathbb{R}^2$. Doing this for each interval in the barcode results in the *persistence diagram*. Note that all points in the persistence diagram must lie above the diagonal, since a feature must have a nonnegative lifespan. Points far above the diagonal represent persistent features, while those lying near the diagonal are short-lived features attributable to noise. The 1-dimensional persistence barcode and diagram of the noisy circle is depicted in Figure 2.3 and as desired the circular nature of the dataset has been isolated. We call the map that associates to a filtration of a finite simplicial complex its persistence diagram or barcode, the *persistence map* and denote it by Pers.

2.4 Metrics on the space of persistence diagrams

The space of persistence diagrams (or equivalently barcodes) can be equipped with a metric.

Definition 2.8 (Wasserstein and bottleneck distances). *Let D, D' be persistence diagrams interpreted as finite multisets of $\mathbb{R}^2 \cup \infty$. A partial matching between D and D' is a subset $\chi \subset D \times D'$ such that*

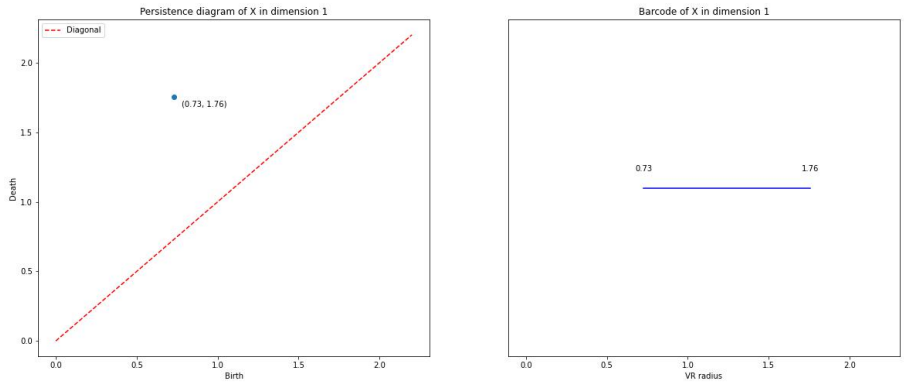


Figure 2.3: The 1-dimensional persistence diagram and barcode of the noisy circle from Figure 2.1 (computed using Gudhi[EB]).

1. *each point in D is matched with at most one point in D', i.e. for each $a \in D$ there is at most one $(a, b) \in \chi$.*
2. *each point in D' is matched with at most one point in D, i.e. for each $b \in D'$ there is at most one $(a, b) \in \chi$.*

The cost of a matching χ is

$$c(\chi) = \Big(\sum_{(a,b)\in\chi} \|a-b\|_\infty^p + \sum_{a\in D\cup D' \text{ unmatched}} \|\frac{a_x - a_y}{2}\|_\infty^p \Big)^{1/p}$$

where $p > 0$ and $\|\|_\infty$ denotes the l^∞-distance. The degree-p Wasserstein *distance is then*

$$W_p(D, D') = \inf_\chi (c(\chi))$$

The bottleneck distance *between is obtained as the limit* $\lim_{p\to\infty} W_p(D, D')$ *and is denoted by* d_B.

The stability theorem for persistent homology states that the persistence map is 1-Lipschitz with respect to the bottleneck distance. Observe first that a 1-critical filtration of a simplicial complex K is simply a function

$$\begin{aligned} \phi : K &\to \mathbb{R} \\ \sigma &\mapsto \phi_\sigma \end{aligned}$$

where ϕ_σ denotes the birth index of σ. This allows us to interpret a filtration of a simplicial complex K as an element of the free vector space $\mathbb{R}^K$. Since the basis of $\mathbb{R}^K$ is composed of the simplices of K, a filtration is a vector in $\mathbb{R}^K$. This interpretation will play an important role in Definition 3.7.

Theorem 2.2 (Stability theorem). *Let K be a simplicial complex and $\phi, \omega : K \to \mathbb{R}$ be filtrations. Then,*

$$d_B(Pers(\phi), Pers(\omega)) \leq \|\phi - \omega\|_\infty = \sup_{\sigma\in K} |\phi_\sigma - \omega_\sigma|$$

Remark 2.3. Theorem 2.2 holds for an arbitrary topological space X and functions $f, g : X \to \mathbb{R}$. In this case one constructs filtrations using sublevel sets of f and g, i.e. $X_r = \{x \in X | f(x) < r\}$ and similarly using g. Computing the homology over the filtration in a given dimension then yields a persistence module and consequently both f and g induce persistence diagrams for which Theorem 2.2 holds.

Chapter 3

Optimizing Functions Based on Standard Persistence Diagrams

One of the earliest and most intuitive motivating works that sought to optimize loss functions defined on persistence diagrams was [GHO16]. They wished to study an important question in physics, namely the geometric origins of rigidity in solids in an amorphous state. As pointed out in [Wya05], statistical physics can provide a description of different states of matter such as liquid or solid states. Crucially however, amorphous solids exhibit a number of anomalous characterstics and are poorly understood by physics. Given that amorphous solids are ubiqutious (glass is a prime example of an amorphous solid), a better understanding of the geometric properties of these solids is extremely important. In [Hir+16], applied the classical persistent homology pipeline to this question by interpreting atomic configurations as point clouds. They were able to use persistence diagrams to discern molecules in an amorphous state from those in crystalline and liquid states. In particular, they noted that the persistence diagrams (see Figure 3.1) of atomic configurations of atoms in an amorphous state exhibited characteristic curves while those from liquid and crystalline sates did not. Indeed, this exhibits the power of the persistence diagram as a tool in its own right, since the diagram can classify the three states. Classification does not however answer the more fundamental question: what shapes in the atomic configuration of amorphous solids make them rigid? In [GHO16] the authors build on [Hir+16], in an attempt to answer this question. Their idea was to start with a point cloud P corresponding to the

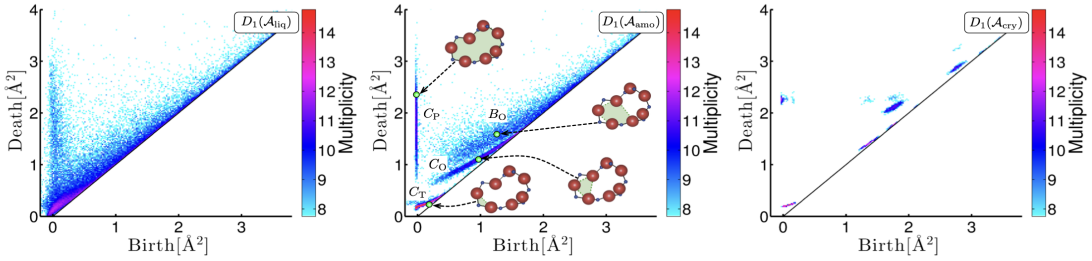


Figure 3.1: Persistence diagrams of the liquid (left), amorphous (center), and crystalline states (right) taken from [Hir+16]. The amorphous state exhibits three curves labelled C_o, C_r, and C_p, while the other states do not.

atomic configuration of an amorphous solid and compute its persistence diagram D. They noted that since the three characteristic curves of D represented the geometric features of the amorphous state, moving to a diagram D' which is constructed by moving the points along a normal direction to the curves would correspond to breaking the geometry responsible for amorphousness. Doing this in small enough steps and simultaneously deforming P to P' such that the persistence diagram of P' is D' would identify the geometric origins of rigidity. In essence, they seek to differentiate the persistence map. As we will soon see in Section 3.8 this is equivalent to optimizing the function $W_p(D, D')$, where $W(D, D')$ denotes the p-Wasserstein distance.

In general, we seek to find minima of functions of the form $\mathbb{R}^N \to \mathbb{R}$ where $\mathbb{R}^N$ is the space of point clouds composed of n points in $\mathbb{R}^d$ and the map is given by first constructing a filtration, then applying the persistence map to the filtration, and finally an appropriate $\mathbb{R}$-valued loss function to the persistence diagram. The standard method to optimize such functions is stochastic subgradient descent.

3.1 Overview

The rest of this chapter is devoted to proving that stochastic subgradient descent converges on the composition of the maps in the bottom row of Figure 3.2. The map Φ constructs a filtration of a simplicial complex K from a set $A \subset \mathbb{R}^N$. Next the associated persistence diagram is computed by the persistence map $\text{Pers} = \mu \circ \Lambda \circ \Omega \circ \Psi \circ \rho$. The discretization map ρ converts an $\mathbb{R}$-indexed filtration to a $\mathbb{N}$-indexed filtration. Next, the map Ψ applies the simplicial homology functor to the $\mathbb{N}$-indexed filtration obtained by ρ. The corresponding decomposition into interval modules is computed by Ω. We then unravel the discretization to obtain the barcode corresponding to the original $\mathbb{R}$-indexed filtration using Λ. The last step is to plot the intervals as points in $\mathbb{R}^2$ to obtain the persistence diagram. Finally, we post-compose with appropriate loss function $\mathcal{L}$.

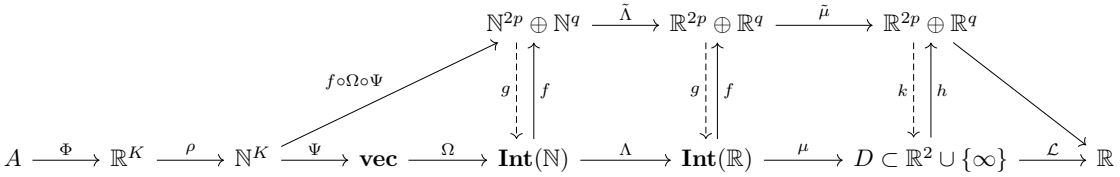


Figure 3.2: Decomposition of the persistence map.

Remark 3.1. In this chapter persistence modules refer to functors $\mathbb{R} \to \mathbf{vec}$ or $\mathbb{N} \to \mathbf{vec}$ and filtrations are always $\mathbb{R}$- or $\mathbb{N}$-indexed.

3.2 Stochastic subgradient descent and definable functions

Since the map $F = \mathcal{L} \circ \text{Pers} \circ \Phi$ is not smooth in general, we need to use results from non-smooth analysis to solve the optimization problem

$$\min_{x \in \mathbb{R}^N} F(x). \tag{3.1}$$

Although the persistence map is not smooth, it is locally Lipschitz (in fact Lipschitz by Theorem 2.1). So if we assume that Φ and $\mathcal{L}$ are locally Lipschitz, then the composition L is locally

Lipschitz. By Rademacher's theorem, locally Lipschitz functions are differentiable almost everywhere. This allows us to generalize the notion of a gradient in the following manner [Dav+20].

Definition 3.1 (Clarke subdifferential). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz and Ω be the subset of $\mathbb{R}^n$ on which f is differentiable. The Clarke subdifferential of f at x is*

$$\partial f(x) = conv\{\lim_{i\to\infty} \nabla f(x_i) : \lim_{i\to\infty} x_i = x, x_i \in \Omega\}$$

where conv denotes the convex hull.

There is always a sequence $\{x_i\} \in \Omega$ that converges to x, since the complement of Ω has Lebesgue measure zero.

Remark 3.2. Clearly if f is differentiable, the Clarke subdifferential reduces to the standard gradient.

Stochastic subgradient descent (SGD) uses the Clarke subdifferential to iterate in the direction of steepest descent

$$x_{k+1} = x_k - \alpha_k(y_k + \xi_k) \qquad y_k \in \partial f(x_k) \tag{3.2}$$

where $\{\alpha_k\}$ is a sequence of step sizes and stochasticity is added by the random sequence $\{\xi_k\}$. Solving Equation 3.1 using SGD, would entail the sequence of iterates $\{x_k\}$ converging to a critical point. In this context, x is a *critical point* if $0 \in \partial f(x)$. In their seminal work [Dav+20], Davis and coauthors proved that under standard assumptions on the step sizes $\{\alpha_k\}$ and noise $\{\xi_k\}$, SGD converges on any locally Lipschitz function that is *definable on an o-minimal structure*. This is a large class of functions that encompasses semialgebraic functions, the exponential function and semianalytic functions among others. Our approach will be to prove that the persistence map is semialgebraic and then assuming that Φ and $\mathcal{L}$ are also definable we use [Dav+20] to guarantee that SGD converges to a critical point on the composition almost surely.

Since we will primarily be working with semialgebraic functions, it is not strictly necessary to understand the definitions of definable functions in general. Nevertheless, we provide a quick overview of the relevant definitions here for completeness and critically in order to state Corollary 5.9 of [Dav+20] which is the key result we will use.

3.2.1 An excursion into o-minimal theory

The building blocks of o-minimal theory are *definable sets*. These are intended in some sense to be generalizations of semi-algebraic sets.

Definition 3.2 (Definable sets and o-minimal structures). *An* o-minimal structure *on $\mathbb{R}$ is a family $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ where each $\mathcal{D}_n$ is a set of subsets such that the following are satisfied for all $n \in \mathbb{N}$:*

1. *$\mathcal{D}_1$ is a finite union of intervals and points.*
2. *$\mathcal{D}_n$ is a boolean algebra on $\mathbb{R}^n$.*
3. *If $A \in \mathcal{D}_n$, then $A \times \mathbb{R} \in \mathcal{D}_{n+1}$.*
4. *If $A \in \mathcal{D}_{n+1}$, then its projection $\pi(A) \in \mathcal{D}_n$ where π denotes the projection onto the first n coordinates.*
5. *Every algebraic set in $\mathbb{R}^n$ is contained in $\mathcal{D}_n$.*

A set is called definable *if it belongs to $\mathcal{D}_n$ for some n. A map whose graph is a definable set is called* definable.

We will primarily work with semi-algebraic sets but will point to examples of filtrations which are definable but not semi-algebraic justifying the generality afforded by o-minimal theory.

Definition 3.3 (Semi-algebraic sets). *We say a subset $\mathbb{A} \subset \mathbb{R}^n$ is semi-algebraic if it can be written as a finite union of sets of the form $\{x \in \mathbb{R}^n | f(x) = 0,\ f \in \mathbb{R}[x_1, \ldots, x_n]\}$ and of the form $\{x \in \mathbb{R}^n | f(x) < 0,\ f \in \mathbb{R}[x_1, \ldots, x_n]\}$.*

Note that the key difference between semialgebraic and algebraic sets is that inequalities are allowed in semialgebraic sets, allowing semialgebraic sets to be closed under projections while algebraic sets are not.

Proposition 3.1 (Semialgebraic sets define an o-minimal structure). *The family $\{\mathcal{D}_n\}_{n \in \mathbb{N}}$ where $\mathcal{D}_n = \{A \subset \mathbb{R}^n | A \text{ is semialgebraic}\}$ is an o-minimal structure on $\mathbb{R}$.*

Proof. We verify the axioms listed in Definition 3.2:

1. The zero locus of a polynomial $f \in \mathbb{R}[x]$ is either a finite set of points of the entire real line. On the other hand sets of the form $\{x \in \mathbb{R} | f(x) < 0,\ f \in \mathbb{R}[x]\}$ are clearly made up of a finite union of intervals or the entire real line.
2. This is clear since unions and intersections of semialgebraic sets are semialgebraic.
3. Let $A \subset \mathbb{R}^n$ be semialgebraic, then $A \times \mathbb{R} = A \cup \{x \in \mathbb{R}^{n+1} | x_{n+1} = 0\}$ is a union of semi-algebraic sets.
4. This is the content of the famous Tariski-Seidenberg theorem - see [Dri98] for details.
5. This holds true by definition.

■

Definable maps and sets have desirable closure properties.

Proposition 3.2 (Closure properties of definable sets). *1. The interior, boundary, and closure of a definable set is definable*

2. The image of a definable set under a definable map is definable.

3. The composition of two definable maps is definable

A key property of definable sets, and indeed the reason we chose to use them in this setting, is that they admit Whitney stratifications. Intuitively, a Whitney stratification of a topological space is a partition of the space, such that the pieces of the partition glue together "nicely".

Definition 3.4 (Stratification). *A* stratification of *a set $X \subset \mathbb{R}^d$ is a finite filtration*

$$\emptyset \subset F_0 \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n = X$$

such that

1. *F_i is closed in X for all i.*
2. *The sets $S_i = F_i \setminus F_{i-1}$ are smooth manifolds of dimension i with finitely many connected components and are called the* strata *of X.*

The family $\mathcal{S} = \{S_i\}_i$ forms a partition of X and is referred to as a stratification of X. We say that a stratification is definable *if each stratum S_i is definable in an o-minimal structure.*

Arbitrary stratifications may behave very badly, in the sense that the topology might vary wildly along singularities. For this reason, Whitney imposed conditions on stratification that exclude such situations. He originally proposed two conditions (A) and (B), but Mather proved that in fact $(B) \implies (A)$, and so we will only present condition (B).

Definition 3.5 (Whitney's condition (B)). *Let X and Y be smooth manifolds.*

The pair (X, Y) satisfies the Whitney regularity condition (B) *at $x \in X$ if the following condition holds. Suppose there exist sequences $\{x_n\}_n \in X$ and $\{y_n\}_n \in Y$, such that*

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = x.$$

Then, $\lim_{n\to\infty} l_n \subset lim_{n\to\infty} T_{y_n} Y$, where $T_{y_n} Y$ denotes the tangent space of Y at y_n and l_n denotes the line spanned by $y_n - x_n$.

We say that Y is Whitney regular over X, if (X, Y) satisfies condition (B) at every $x \in X$.

A stratification is Whitney, if (S_i, S_{i-1}) satisfies condition (B) for all i .

Definition 3.6 (Stratification of maps). *Let $X \subset \mathbb{R}^d$. A map $f : X \to \mathbb{R}$ admits a Whitney stratification, if its graph admits a Whitney stratification.*

We can now finally state the key convergence result.

Theorem 3.1 (SGD converges on Whitney stratifiable maps (Corollary 5.2 of [Dav+20])). *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz map that admits a Whitney stratification. Let $\{\alpha_k\}$ be a sequence of step sizes and $\{\xi_k\}$ be a random sequence that satisfy Assumption C of [Dav+20]. Then almost surely, every sequence $\{x_k\}$ produced by Equation 3.2 converges to a critical point of f and the corresponding function values $\{f(x_k)\}$ converge.*

Remark 3.3. Assumption C imposes 3 weak requirements on the sequence of step sizes and the sequence of noise. It requires that step sizes be nonnegative, square summable, but not summable. It defines $\{\xi_k\}$ in a manner that imposes bounds on its values. It is a rather technical requirement, but easily satisfiable and not particularly relevant for this exposition.

Crucially by theorems 1.3 and 2.2 of [Tos+10], definable sets and maps admit Whitney stratifications.

Theorem 3.2 (Definable sets and maps admit Whitney stratifications). *Let $A \subset \mathbb{R}^n$ be a definable set, then there exists a definable Whitney stratification $\mathcal{S}$ of A.*

Let $A \subset \mathbb{R}^d$ be definable and $f : A \to B$ be a definable map, then there exists a definable Whitney stratification of f. Furthermore, there exists a definable Whitney stratification of A such that the restriction of f to each stratum of A is differentiable.

Theorem 3.1 and Theorem 3.2 together yield the following corollary.

Corollary 3.1 (SGD converges on definable maps). *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz map that is definable on some o-minimal structure. Let $\{\alpha_k\}$ be a sequence of step sizes and $\{\xi_k\}$ be a random sequence that satisfy Assumption C of [Dav+20]. Then almost surely, every sequence $\{x_k\}$ produced by Equation 3.2 converges to a critical point of f and the corresponding function values $\{f(x_k)\}$ converge.*

Remark 3.4. The idea of using the results in [Dav+20] to prove that SGD converges on functions based on persistence diagrams first appeared in [Car+21]. Our approach of proving that SGD converges maps based on signed barcodes is inspired by this paper. However, in contrast to [Car+21] our approach goes directly through the space of persistence modules rather than using the row reduction algorithm to go directly from filtrations to persistence diagrams. This approach has two advantages. Firstly, there is no analogue of the row reduction algorithm for signed barcodes, indeed one does not even have persistence pairs. Hence, when working in the multiparameter case, we are forced to go through the space of persistence modules. Adapting the proof in [Car+21] for the single parameter case to fit in this framework, makes the jump to multiparameter persistence a bit easier. Secondly, it allows us to work in one homology degree at a time, i.e. the persistence map sends a filtration to its persistence diagram in a fixed homology degree i. This is in contrast to [Car+21], where the persistence map sends a filtration to the union of persistence diagrams in all dimensions.

3.3 Definable parameterized families of filtrations.

We can now apply the results of the previous section to our setting. Recall from Section 2.3, that a filtation of a simplicial complex K can be interpreted as an element of the free vector space $\mathbb{R}^K$.

Definition 3.7 (Parameterized family of filtrations). *Let K be a simplicial complex and A be a set. A* family of parameterized filtrations *is a map*

$$\begin{aligned} \Phi : A &\to \mathbb{R}^K \\ x &\mapsto \Phi(x) \end{aligned}$$

such that $\tau \subset \sigma$ implies $\Phi(x)_\tau \leq \Phi(x)_\sigma$ for all $x \in A$. In other words, $\Phi(x) \in \mathbb{R}^K$ is a filtration of K. If Φ is a definable map, we call the family definable. Denote by $Filt_K \subset \mathbb{R}^K$ the subspace of vectors that define valid filtrations of K.

Example 2 (VR filtrations form a parameterized family). Consider the set of point clouds with $n > 0$ points embedded in $\mathbb{R}^d$, $d > 0$. Each such point cloud can be expressed as a vector in $R^{dn} = R^N$. Let $K = \Delta_n$ and $\Phi : \mathbb{R}^N \to \mathbb{R}^K$ be the map that send a point cloud x to its VR filtration. By Lemma 2.2, this defines a parameterized family of filtrations.

Moreover, the map Φ is definable. Indeed, given $x = (x_1, \ldots, x_n) \in \mathbb{R}^N$ and a simplex $\sigma \in K$, we have $\Phi_\sigma(x) = \max_{i,j \in \sigma} \parallel x_i - x_j \parallel$. The Euclidean norm and max are both semi-algebraic and hence definable. Consequently, their composition Φ is definable.

As promised, we now present a family of filtrations, which is not semi-algebraic but still definable.

Example 3 (Weighted Rips Filtrations). Consider again n points embedded in $\mathbb{R}^d$ and let $f : \mathbb{R}^d \to \mathbb{R}$ be a function. Then, the family of *f-weighted Rips filtrations* is a map $\Phi : A = (\mathbb{R}^{dn} \to \Delta_n)$ defined as follows: let $x \in A$ and $\sigma \in \Delta_n$, then

$$\Phi_\sigma(x) = \begin{cases} 2f(x_i) & \sigma = [i] \\ \max\{2f(x_i), 2f(x_j), \|x_i - x_j\| + f(x_i) + f(x_j)\} & \sigma = [i,j], i \neq j \\ \max\{\Phi_{[i,j]}(x) | i, j \in \sigma\} & \end{cases}$$

where $[i, j]$ denotes the full simplex spanned by the vertices $i \ldots, j$. The Euclidean distance function and max are semialgebraic and hence definable. Consequently, Φ is definable if and

only if f is definable. The o-minimal framework allows us to use a wide range of functions such as Gaussian kernels, which involves the exponential function and is thus not semialgebraic. It is however definable.

3.4 Filtrations and preorders

The key observation made in [Car+21] is that the preorder induced by a given filtration completely determines the pairs of simplices that generate the points of a persistence diagram.

Definition 3.8 (Persistence pairs). *Let $\Phi(x) \in \mathbb{R}^K$ be a filtration and $\mathcal{M} \cong \bigoplus_{I \in \mathcal{I}} k_I$ be its persistence module and interval decomposition. We call a pair of simplices (σ, τ) a* persistence pair *if $(\Phi_\sigma(x), \Phi_\tau(x)) \in \mathcal{I}$. Further, we say a simplex σ is* essential *if $(\Phi_\sigma(x), +\infty) \in \mathcal{I}$.*

Geometrically speaking, a pair (σ, τ) represents the fact that a cycle is created when σ enters the filtration and is then killed when τ enters the filtration. An essential simplex σ represents a cycle that is created when σ is added to the filtration but is never killed. An example of this is the single connected component that always survives throughout the filtration when considering 0-dimensional homology.

Remark 3.5 (The row reduction algorithm and persistence pairs).

Proposition 3.3 (Filtrations impose preorder). *Let $\Phi(x) \in \mathbb{R}^K$ be a filtration of a simplicial complex K. Then, $\Phi(x)$ imposes a preorder on the simplices of K. Moreover, if one fixes a total ordering of the vertices of K beforehand, then this preorder can be extended to a total order.*

Proof. Let σ, τ be simplices of K. We say $\sigma \leq \tau$ if $\Phi_\sigma(x) \leq \Phi_\tau(x)$. This is not a total order, since we may have $\Phi_\sigma(x) = \Phi_\tau(x)$.

Suppose we have a total order on the vertices $\{v_1, \ldots, v_n\}$ of K. This ordering can just be any arbitrary indexing. In the case of point clouds for example, the indexing is given by the data structure used to store the point cloud (a NumPy array for instance). Using this total order, [Bau21] describes a means of refining the partial order to a total order. Let $\sigma, \tau \in K$ be distinct simplices such that $\Phi_\sigma(x) = \Phi_\sigma(x)$. If $\dim \sigma \neq \dim \tau$, then we can order them according to their dimension. If their dimensions are equal, then we order them using the lexicographic order on their vertices ■

These preorders can be used to define a linear partition of Filt_K.

Lemma 3.1 (Preorders induce a partition of the filter space.). *There exists a linear (and hence semialgebraic) partition of the subspace $Filt_K \subset \mathbb{R}^K$.*

Proof. We use the imposed orders to define a partition of Filt_K as follows. Let $\leq$ be a partial order on the simplices of K. Note that since K is finite, there are only finitely many partial orders. Define a partition of Filt_K as follows

$$O_\leq = \{\Phi(x) \in \mathrm{Filt}_K | \Phi(x) \text{ imposes a partial order on } K \text{ equal to } \leq\}.$$

Two filtrations $\Phi(x), \Phi(y), \in \mathbb{R}^K$ induce the same preorder on the simplices of K if

$$v_\sigma \leq v_\tau \iff w_\sigma \leq w_\tau$$

holds for all $\sigma, \tau \in K$. The partition is completely defined by inequalities between coordinates, and thus linear. ■

The partition described above is a semi-algebraic (and hence definable) stratification of Filt_K. If two filtrations $\Phi(x), \Phi(y) \in \mathbb{R}^K$ lie in the same partition element, we say that the filtrations lie in the same *stratum*.

Crucially, we can use this stratification of Filt_K to stratify the underlying geometric space A, over which Φ is defined.

Remark 3.6 (Stratification of the underlying geometric space.). Define a partition of A by sets of the form

$$A_{\leq} = \{x \in A | \Phi(x) \text{ imposes a preorder equal to } \leq \text{ on the simplices of } K\}$$

Two elements $x, y \in A$ are in the same stratum of A if the filtrations $\Phi(x)$ and $\Phi(y)$ they induce lie in the same stratum of Filt_K. In the context of differentiability, this means a point y that lies within a small enough neighborhood of $x \in A$, will induce a filtration $\Phi(y)$ that lies in the same stratum as $\Phi(x)$.

The key fact that we will prove over the course of the next two sections is the following: if $\Phi(x), \Phi(y) \in \text{Filt}_K$ lie in the same stratum, then their persistence pairs and essential simplices are identical. In other words, the simplices that generate the bars on either barcode are the same. The only thing that changes is the length of the bars, which is determined by the birth indices of the simplices.

3.5 Discretization

Rather than working with an $\mathbb{R}$-indexed filtration, we would like to discretize the filtration to obtain one indexed by $\mathbb{N}$. The discretization map is in some sense a means of isolating solely the order imposed by a given filtration and is defined component wise as follows

$$\begin{aligned} \rho : \mathbb{R}^K &\to \mathbb{N}^K \\ v_\sigma &\mapsto \#\{\tau \in K | \tau \leq \sigma\}. \end{aligned} \tag{3.3}$$

Note that in this case $\leq$ denotes the total order on the simplices obtained by extending the partial order imposed by $\Phi(x)$ to a total order. Using the total order instead of a partial order, ensures that after discretization no two simplices have the same birth index. This fact will be important in Section 3.7.

Proposition 3.4 (Discretization is constant on similar preoders). *The map $\rho : \mathbb{R}^K \to \mathbb{N}^K$ is semialgebraic. More precisely, the restriction of ρ to each stratum of Filt_K defined in Lemma 3.1 is constant.*

Proof. Suppose now that $\Phi(x), \Phi(y) \in O_{\leq}$ for some partial order $\leq$. Although we used the total order in Equation 3.3, the definition really only depends on the partial order. Indeed, as long as the same rules are always used to extend a given partial order to a total order, for example by using a fixed indexing of the simplices as described in Proposition 3.3, then the total order obtained from two equivalent preorders will be the same.

A map whose restriction is constant on each element of a semialgebraic partition is trivially semialgebraic. In particular, this means that ρ is definable on an o-minimal structure. ■

Remark 3.7. Given a partition element $O_{\leq} \subset \text{Filt}_K$, each $\rho(O_{\leq}) \subset \mathbb{N}^K$ contains only a single element. Moreover, $\rho(O_{\leq})$ defines a partition of the subspace of vectors that define a filtration in $\mathbb{N}^K$. This is equivalent to the partition one would obtain by using the imposed preorder as in the case of $\mathbb{R}^K$ and is hence semialgebraic.

3.6 From filtrations to barcodes

The next step in the pipeline is to apply the simplicial homology functor to the $\mathbb{N}$-indexed filtration we obtained in the previous section. The map $\Psi : \mathbb{N}^K \to \mathbf{vec}$ associates to a filtration $\rho(\Phi(x)) \in \mathbb{N}^K$ the corresponding persistence module $\mathcal{M}$. Importantly, $\mathcal{M}$ is the persistence module of a discretized filtration in $\mathbb{N}^K$. Consequently, given two filtrations $\Phi(x), \Phi(y) \in \mathbb{R}^K$ that lie in the same stratum, the persistence modules after discretization $\Psi(\rho(\Phi(x)))$ and $\Psi(\rho(\Phi(y)))$ are identical.

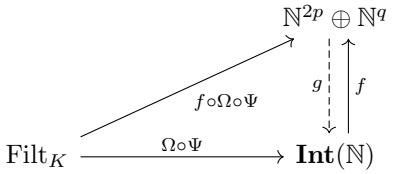
We then apply Theorem 2.1, to define $\Omega : \mathbf{vec} \to \mathbf{Int}(\mathbb{N})$ which decomposes persistence modules into interval modules. Again, since two filtrations in the same stratum are mapped to the same persistence module by $\Psi \circ \rho$, decomposing into interval modules by Ω results in the same barcode.

To prove that the composition of these two maps is definable, we first need to lift them to spaces in which it makes sense to talk about definability. Let $\mathcal{I} \subset \mathbf{Int}(\mathbb{N})$ be a barcode and denote by p the number of intervals of the form $(a, b),\ a < b < \infty$, and by q the number of intervals of the form (a, ∞). Since the barcode does not change along a single stratum of Filt_K, the numbers p and q do not vary either. So, as long as we remain in the same stratum, which is the case if we consider small enough neighborhoods, then we can lift $\mathbf{Int}(\mathbb{N})$ to the space $\mathbb{N}^{2p} \oplus \mathbb{N}^q$ by a map f. Each of the p intervals of the form $(a, b) \subset \mathbb{N}$ can be embedded in $\mathbb{R}^2$. The p intervals can then be put in non-decreasing order using the lexicographic order on $\mathbb{R}^2$, enabling us to express the p intervals as a vector in $\mathbb{R}^{2p}$. The q intervals of the form (a, ∞) can be embedded in $\mathbb{N}^q$ in the obvious manner.

$$f : \mathbf{Int}(\mathbb{N}) \to \mathbb{N}^{2p} \oplus \mathbb{N}^q$$
$$(\{(a_1, b_1), \ldots, (a_p, b_p), (c_1, \infty), \ldots, (c_q, \infty)\}) \mapsto ((a_1, b_1, a_2, b_2, \ldots, a_p, b_p), (c_1, \ldots, c_q))$$

Remark 3.8. We emphasize that due to the nature of the lift f, the codomain of $f \circ \Omega \circ \Psi \circ \rho$ varies along $\mathrm{Filt}_K \subset \mathbb{R}^K$, however; it is constant on each stratum of Filt_K. This is sufficient for our purposes, since we only need to prove that the map is differentiable on each stratum and hence differentiable almost everywhere.

Clearly this map has a left inverse g, making it a lift and leading to the following diagram.



We can now prove that the composition $f \circ \Omega \circ \Psi$ is definable by proving that its restriction to each partition element is constant as in Proposition 3.4.

Proposition 3.5. *The map $f \circ \Omega \circ \Psi$ is semialgebraic. More precisely, there exists a semialgebraic partition of the subspace of filtration vectors $\mathrm{filt}_K \subset \mathbb{N}^K$ such that the restriction of $f \circ \Omega \circ \Psi$ to each partition element is constant.*

Proof. We will work with the semialgebraic partition of filt_K defined in Remark 3.7. Two elements lie in the same stratum of $\mathrm{Filt}_K \subset \mathbb{R}^K$ if and only if they lie in the same stratum of $\mathrm{filt}_K \subset \mathbb{N}^K$.

Each stratum of filt_K has a single element and so the composition $f \circ \Omega \circ \Psi$ is trivially constant when restricted to each stratum. Consequently, the map is semialgebraic and definable on an o-minimal structure. ∎

Remark 3.9. Proposition 3.4 and Proposition 3.5 together imply that the restriction of $f \circ \Omega \circ \Psi \circ \rho$ to each stratum of $\text{Filt}_K \in \mathbb{R}^K$ is constant and hence the map is definable.

3.7 Unraveling the discretization

The barcodes computed in Section 3.6 correspond to a discretized $\rho(\Phi(x)) \in \mathbb{N}^K$ and not to $\Phi(x)$ which is the filtration that actually captures the topology of the underlying data. We now want to recover the barcodes that correspond to the actual filtration by using the preimage of ρ. In other words, given a barcode $\mathcal{I} \subset \mathbf{Int}(\mathbb{N})$ we wish to construct a barcode $\mathcal{J} \subset \mathbf{Int}(\mathbb{R})$. We do so by first finding the simplex pairs (or essential simplices) that generate each interval in $\mathcal{I}$ and then using the filtration function Φ to recover their birth indices, and hence the barcode.

It is easiest to define this map upstairs on elements of $\mathbb{N}^{2p} \oplus \mathbb{N}^q$. The first step is to recover the simplices that generate each of the intervals in the barcode.

$$\tilde{\kappa} : \mathbb{N}^{2p} \oplus \mathbb{N}^q \to K^{2p} \oplus K^q$$
$$((a_1, b_1, a_2, b_2, \ldots, a_p, b_p), (a_1, \ldots, a_q)) \mapsto ((\sigma_1, \tau_1, \sigma_2, \tau_2, \ldots, \sigma_p, \tau_p), (\omega, \ldots, \omega_q))$$

This is possible since each interval (a, b) or (a, ∞) is associated to a pair of simplices or an essential simplex. Owing to the use of a total order in Equation 3.3, no two simplices have the same birth index after discretization, ensuring that $\tilde{\kappa}$ is well-defined. We post compose $\tilde{\kappa}$ with

$$\tilde{\eta} : K^{2p} \oplus K^q \to \mathbb{R}^{2p} \oplus \mathbb{R}^q$$
$$((\sigma_1, \tau_1, \sigma_2, \tau_2, \ldots, \sigma_p, \tau_p), (\omega, \ldots, \omega_q)) \mapsto ((\Phi_{\sigma_1}(x), \Phi_{\tau_1}(x), \ldots, \Phi_{\sigma_p}(x), \Phi_{\tau_p}(x)), (\Phi_{\omega_1}(x), \ldots, \Phi_{\omega_q}(x)))$$

to obtain the map $\tilde{\Lambda}$. Due to the lift $f : \mathbf{Int}(\mathbb{N}) \to \mathbb{N}^{2p} \oplus \mathbb{N}^q$ and an identically defined lift $\mathbf{Int}(\mathbb{R}) \to \mathbb{R}^{2p} \oplus \mathbb{R}^q$, the map $\tilde{\Lambda}$ induces a map $\Lambda : \mathbf{Int}(\mathbb{N}) \to \mathbf{Int}(\mathbb{R})$.

Since the values of p and q (and hence the codomain of $f \circ \Omega \circ \Psi \circ \rho$) vary over the strata of $\text{Filt}_K \subset \mathbb{R}^K$, we will prove that $\tilde{\Lambda} \circ f \circ \Omega \circ \Psi$ is affine over each stratum of Filt_K.

Proposition 3.6 ($\tilde{\Lambda}$ is affine on each stratum). *The restriction of the map $\tilde{\Lambda} \circ f \circ \Omega \circ \Psi \circ \rho$ to each stratum of $Filt_K$ is definable.*

Proof. By Proposition 3.4 and Proposition 3.5, we know that the map $f \circ \Omega \circ \Psi \circ \rho$ is constant on each stratum. Post composing with $\tilde{\kappa}$ does not change this, since it just maps each interval of the barcode to the simplices that generates it. This depends only on the discretized barcode obtained by the map $\Omega \circ \Psi \circ \rho$, which is constant on each stratum of Filt_K. The map $\tilde{\eta}$ then simply sends each simplex σ to its birth index $\Phi_\sigma(x)$. This map is definable as long as the filtration function Φ is definable. ■

3.8 Functions based on persistence diagrams.

The final step in the pipeline is to define loss functions that optimize for certain topological features. It is easiest to define such loss functions on the space of persistence diagrams (i.e. a subset of $\mathbb{R}^2$). Mapping a collection $\mathcal{I} \subset \mathbf{Int}(\mathbb{R})$ to a persistence diagram $D \subset \mathbb{R}^2$ is simple, we just need to send an interval (a, b) to the point $(a, b) \subset \mathbb{R}^2$ and an interval (a, ∞) to the point $(a, \infty) \subset \mathbb{R}^2 \cup \{\infty\}$. This map clearly lifts to the identity on $\mathbb{R}^{2p} \oplus \mathbb{R}^q$.

Theorem 3.3 (The persistence map is definable). *The map $Pers : Filt_K \to \mathbb{R}^2 \cup \{\infty\}$ defined by the composition $\mu \circ \Lambda \circ \Omega \circ \Psi \circ \rho$ is definable. Moreover, there exists a definable stratification of $Filt_K$ such that the restriction of Pers to each stratum is differentiable.*

Proof. Over the course of the past few sections, we have proven that each of the maps in the composition is definable. By Proposition 3.2, the composition of definable maps is definable and so Pers is definable. Applying Theorem 3.2 yields the second statement. ■

We can now define loss functions on the space of persistence diagrams. In order for differentiability to continue to hold, we require these loss functions to be definable (as a map from the lifted space $\mathbb{R}^{2p} \oplus \mathbb{R}^q$) and locally Lipschitz.

One such loss function is the bottleneck or Wasserstein distance to another fixed diagram D', i.e. $\mathcal{L}(D) = d_B(D, D')$ or $\mathcal{L}(D) = W_p(D, D')$. Indeed, this is precisely the loss function used in the motivating example at the beginning of the chapter, where D' would be the persistence diagram obtained by moving the points on the curves of diagram D upward in the normal direction. These functions are clearly definable, since they involve minimums and Euclidean distance functions.

Another example of a loss function is the persistence image defined in [Ada+17]. Given a weight function $w : \mathbb{R}^2 \to \mathbb{R}$ and $\sigma \in \mathbb{R}_{>0}$, the persistence image of a diagram D is given by the function

$$\begin{aligned} I_D : \mathbb{R}^2 &\to \mathbb{R} \\ q &\mapsto \sum_{p \in D} w(p) \exp(-\frac{\|p-q\|^2}{2\sigma^2}). \end{aligned}$$

This yields a loss function on the space of persistence diagrams as follows. Fix a point $q \in \mathbb{R}^2$, a weight function w and $\sigma > 0$, then define

$$\mathcal{L}(D) = I_D(q).$$

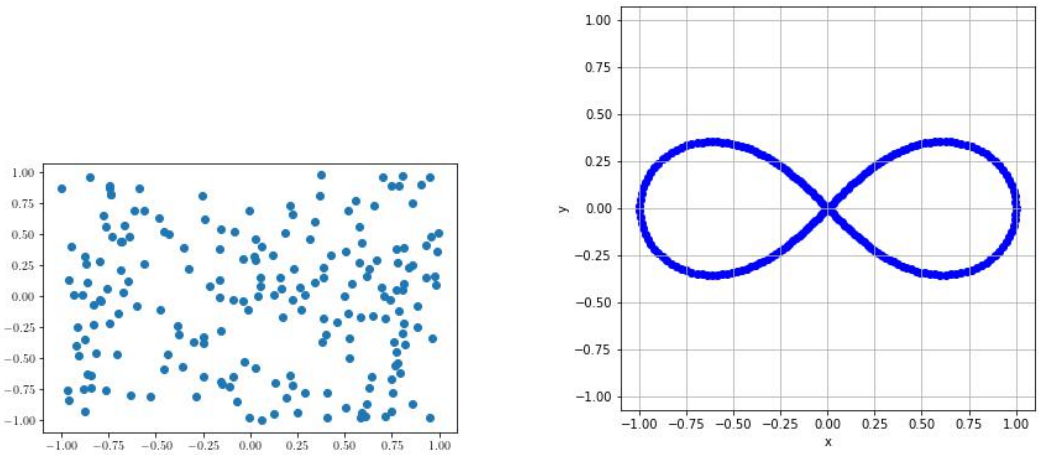
Since the exponential function, sums, multiplication, and the Euclidean norm are all definable, $\mathcal{L}$ is definable as long as the weight function w is definable. In [Ada+17], conditions for Lipschitz continuity are laid out.

Putting everything together yields the main result of this chapter.

Theorem 3.4 (SGD converges on persistent homology based maps). *Let $A \subset \mathbb{R}^K$ be definable, K be a finite simplicial complex and $\Phi : A \to \mathbb{R}^N$ be a* locally Lipschitz *family of filtrations that is* definable on an o-minimal structure. *Let $\mathcal{L} : \mathbb{R}^2 \cup \{\infty\} \to \mathbb{R}$ be a* locally Lipschitz *function that is* definable on an o-minimal structure. *Assume further, that the sequences of step sizes $\{\alpha_k\}$ and noise $\{\xi_k\}$ satisfy condition C of [Dav+20]. Then, the iterates $\{x_k\}$ produced by SGD converge almost surely to a critical point of $\mathcal{L} \circ Pers \circ \Phi$ and the function values $\mathcal{L} \circ Pers \circ \Phi(x)$ converge.*

3.9 Experiments

One of the first optimization tasks that one might attempt is point cloud optimization. Here, one aims to optimize a point cloud so that its shape moves toward some desired target shape. In [Car+21], the authors begin with a point cloud sampled uniformly from the unit square $S = [0, 1]^2$ and aim to maximize its 1-dimensional homology, i.e. create cycles. Their loss function is the Wasserstein distance to the empty diagram. By maximizing $\mathcal{L} \circ \text{Pers} \circ \Phi$ they are able to move the points of the uniform point cloud to one with many cycles. Our experiment is based on a similar idea. Our goal is to optimize a uniformly distributed point cloud, so that its shape is similar to that of $S^1 \vee S^1$. The initial point cloud X we wish to optimize, as well as the target

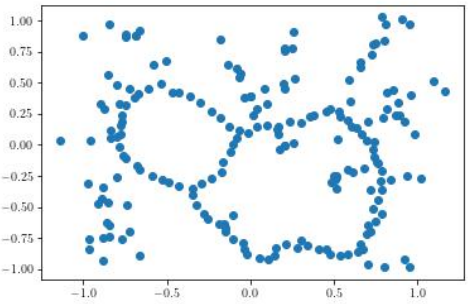


(a) Initial point cloud X comprised 200 points sampled uniformly from the unit square.

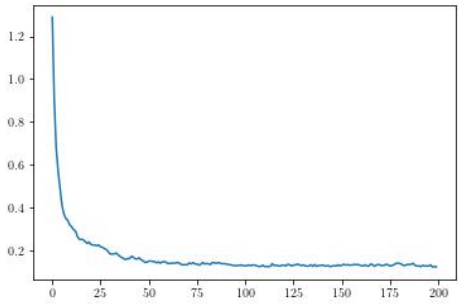
(b) Target point cloud Y comprised of 200 points sampled from the wedge product of two circles.

Figure 3.3

shape Y we wish to converge to are depicted in Figure 3.3a and Figure 3.3b. Let D' denote the 1-dimensional persistence diagram of Y. Let $\Phi : \mathbb{R}^{(2)(200)} \to \Delta_{200}$ be VR-filtrations, Pers be 1-dimensional persistence, and $\mathcal{L}$ be the Wasserstein distance to D'. We apply Adam [KB17], to optimize $\mathcal{L}(X)$ and obtain Figure 3.4a. Note that while the 1-dimensional homology of the optimized point cloud and the target are remarkably similar, there appear to be more connected components in the optimized point cloud. This is due to the fact, that we only considered 1-dimensional homology.



(a) Point cloud after 200 iterations of Adam.



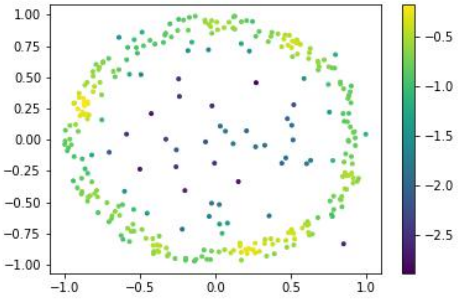
(b) The loss function plotted against number of iterations.

Figure 3.4

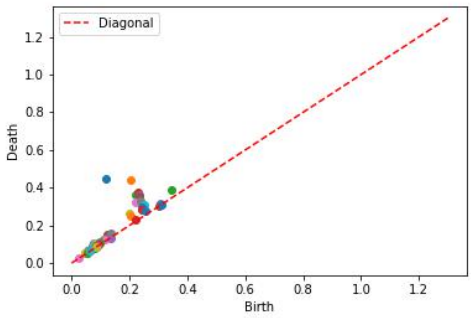
Chapter 4

A Primer on Multiparameter Persistence

Although single parameter persistence is a powerful tool that has proven successful in multiple domains, it has some drawbacks. Firstly, it is not adept at differentiating certain kinds of noise. Consider for example, the dataset depicted in Figure 4.1. Most of the points are sampled from a



(a) Points sampled from a circle with uniformly distributed background noise and a Gaussian kernel density estimate.



(b) H_1 persistence diagram of the dataset on the left.

Figure 4.1: Point cloud with a dense circle and its corresponding persistence diagram.

circle, like in Figure 2.1, but with some additional noise in the form of uniformly sampled points. However, as witnessed by the densities of the points (computed using a Gaussian kernel estimate), the circle is still the primary geometric feature of the data. Unfortunately, the persistence diagram in dimension 1, obtained using a VR complex, no longer captures this feature, and contains points close to the diagonal. It is unable to detect the fact the points along the circle have a much higher density. This makes sense, since the VR complex will include the edges spanned by the noisy points at the same time as those spanned by the denser ones, because the distances between two noisy points or a noisy point and a given point on the circle does not depend on the density of the points. One way to rectify this, could be to fix a density threshold p and only include points whose density values are above this threshold. In the above example, discarding points whose density values are lower than 0.10 would remove most of the noisy points.

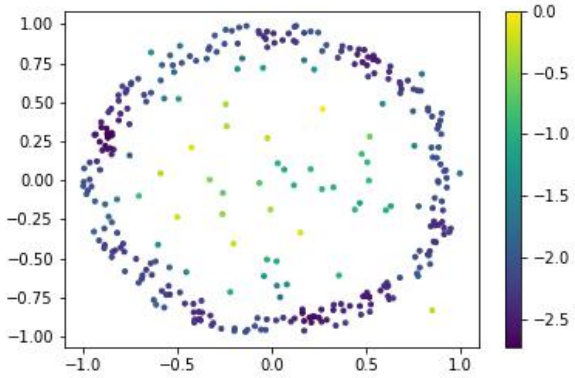


Figure 4.2: Points sampled from a circle with uniformly distributed background noise and a Gaussian kernel codensity function.

But just as in the 1-parameter case, where we glean more information by considering different radii for the VR complexes and tracking the persistence of features as the radii change, we would like to vary the density threshold and track the evolution of features along 2 parameters instead of just 1. This leads to the *function-rips* bifiltration.

Definition 4.1 (Function-Rips bifiltration). *Let X be a finite metric space and $\gamma : X \to \mathbb{R}_{\geq 0}$ be* function *on X. Then, the corresponding* superlevel function-Rips bifiltration *of X is given by $VR(\gamma^{-1}(a, \infty), r)$ for each $(a, r) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$. Analogously, the corresponding* sublevel function-Rips bifiltration *of X is given by $VR(\gamma^{-1}(0, a), r)$ for each $(a, r) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$.*

Common choices for γ include a Gaussian kernel, nearest neighbor functions or disk kernels. The corresponding function-Rips bifiltrations are called *density-Rips* bifiltrations, since they take high values on points in regions of high density and low values on sparse regions. In practice, one often uses a codensity function, i.e. a function whose values are low on regions of high density and high on low density regions. Points of high density would then appear first in the corresponding sublevel filtration. A codensity function on the noisy circle is depicted in Figure 4.2 and the corresponding sublevel density-Rips bifiltration is depicted in Figure 4.3. This bifiltration isolates the main cycle, since the noisy points appear only toward the top as the density threshold is lowered. Note that the single parameter VR complex is simply the subcomplex on the top row of the figure, which is clearly unable to detect the cycle.

Often, the underlying data already comes equipped with a function of interest. For example, persistent homology has been applied to drug discovery, where one aims to use topology to identify drug candidates that are most likely to bind to a particular molecular target. Here there are multiple other parameters that one might want to use in a filtration such as atomic mass, partial charge or bond type. In [Dem+22], the authors were able to use multiparameter persistence to significantly improve the performance in comparison to single parameter persistence.

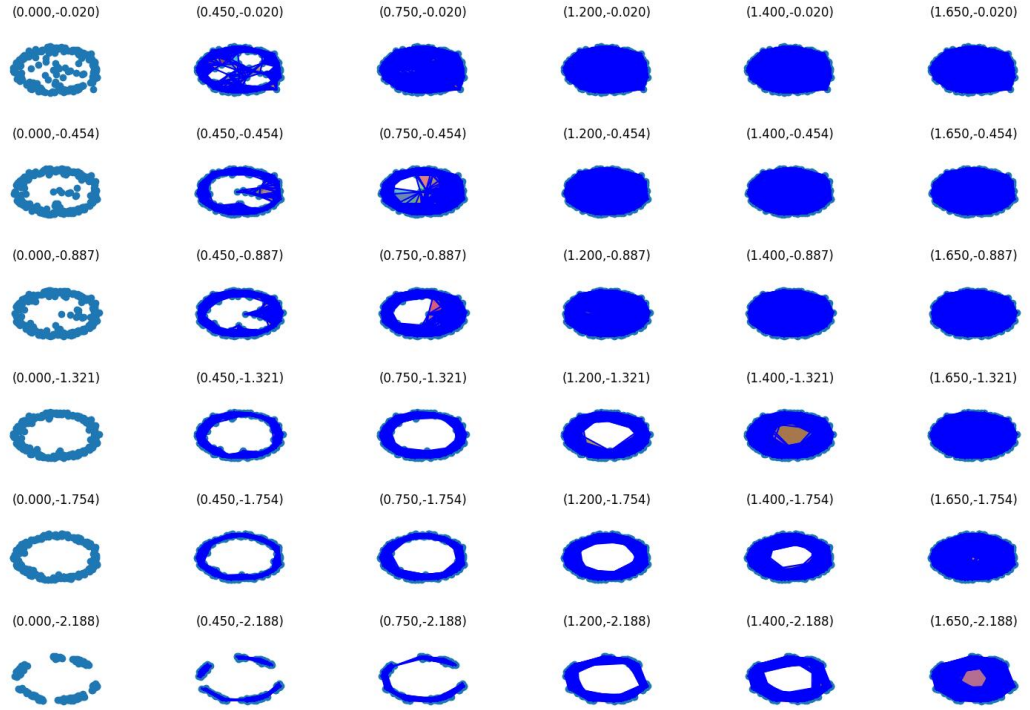


Figure 4.3: Density-Rips bifiltration of points sampled from a circle with noise in the center.

4.1 The non-existence of barcodes for multiparameter persistence modules

As alluded to earlier, there is no direct analogue to the barcodes one obtains from interval decompositions of 1-parameter persistence modules. Indeed, a classical result from representation theory known as Gabriel's theorem states that persistence modules of the form $\mathbb{R}^n \to \mathbf{vec}$ or even $\mathbb{Z}^n \to \mathbf{vec}$ cannot be decomposed into interval modules for $n > 1$. We point the reader to [Oud15] for details on the representation theory of quivers arising from persistence modules and a proof of Gabriel's theorem in this setting. We instead present an example of a simple 2-parameter persistence module that cannot be decomposed into interval modules.

Example 4 (Indecomposable 2-parameter persistence module). The main obstacle to decomposing the k^2 in the center of the persistence module depicted in Figure 4.4 is the fact that the rank of the horizontal map that passes through it is 1, while the rank of the vertical map passing through it is 0. TODO

4.2 Rank decompositions and signed barcodes

Although interval decompositions in the classical sense are impossible in this setting, allowing for negative signs in the decomposition offers us a way out. We do this by expressing invariants of persistence modules as signed sums of invariants of interval modules. One such invariant is the rank invariant. It is instructive to briefly revisit the 1-parameter case to better understand

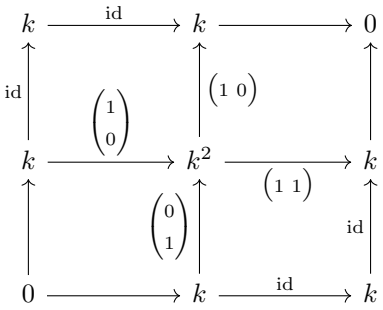


Figure 4.4: An example of an indecomposable $\mathbb{Z}^2$-indexed persistence module

the rank invariant.

Definition 4.2 (The one parameter rank invariant). *Let $\boldsymbol{P} \in \{\mathbb{R}, \mathbb{Z}\}$, $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ be a persistence module. Then, the rank invariant of $\mathcal{M}$ is a map*

$$\begin{aligned} Rk\mathcal{M} : \boldsymbol{Int}(\boldsymbol{P}) &\to \mathbb{Z} \\ I &\mapsto Rk_I\mathcal{M} = rank(\mathcal{M}(a) \to \mathcal{M}(b)) \end{aligned}$$

where $rank(\mathcal{M}(a) \to \mathcal{M}(b))$ is the rank of the linear map between the vector spaces $\mathcal{M}(a)$ and $\mathcal{M}(b)$.

In the one parameter case the rank invariant is a complete invariant of persistence modules. In particular, one can use the barcode to decompose the rank invariant of $\mathcal{M}$ into the sum of rank invariants of interval modules.

Proposition 4.1 (Rank invariant and barcodes). *Let $\boldsymbol{P} \in \{\mathbb{R}, \mathbb{Z}\}$, $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ be a persistence module. Then there exists a collection of intervals $\mathcal{I} \subset \boldsymbol{Int}(\boldsymbol{P})$ such that*

$$Rk\mathcal{M} = \sum_{I \in \mathcal{I}} Rkk_I.$$

Proof. By Theorem 2.1, we can write $\mathcal{M} \cong \bigoplus_{I \in \mathcal{I}} k_I$. Thus given an interval J, we have

$$\begin{aligned} \mathrm{Rk}_J M &= \mathrm{Rk}_J(\bigoplus_{I \in \mathcal{I}} k_I) \\ &= \sum_{I \in \mathcal{I}} \mathrm{Rk}_J k_I \end{aligned}$$

■

One can go in the other direction as well.

Proposition 4.2. *The barcode $barc(\mathcal{M})$ of a persistence module $\mathcal{M}$ can be recovered from the rank invariant $Rk\mathcal{M}$.*

Proof. Let $I = (a, b) \subset \mathbb{N}$ be an interval. We provide an explicit formula for the number of copies of I in $\mathrm{barc}(\mathcal{M})$ in terms of the rank invariant.

Denote by $C(I)$ the number of intervals in $\mathrm{barc}(\mathcal{M})$ *containing* I. Intuitively,

$$\mathrm{Rk}_I\mathcal{M} = \mathrm{rank}(\mathcal{M}(a) \to \mathcal{M}(b))$$

counts the number of features alive between times a and b. However, a feature alive between a and b might have been born before a or die after b. The bar corresponding to a particular feature represents its entire lifespan, which is an interval containing (a, b). Hence,

$$C(I) = \mathrm{Rk}_I \mathcal{M}\,.$$

A standard result from combinatorics, the inclusion-exclusion principle (see [All10]), yields:

$$\#(a,b) = \begin{cases} C((a,b)) - C((a,b+1)) - C((a-1,b)) + C((a-1,b+1)) & 0 \le a < b < \infty \\ C((a,\infty)) - C(a-1,\infty) & 0 \le a < b = \infty \end{cases}.$$

Hence,

$$\#(a,b) = \begin{cases} \mathrm{Rk}_{(a,b)}\mathcal{M} - \mathrm{Rk}_{(a,b+1)}\mathcal{M} - \mathrm{Rk}_{(a-1,b)}\mathcal{M} + \mathrm{Rk}_{(a-1,b+1)}\mathcal{M} & 0 \le a < b < \infty \\ \mathrm{Rk}_{(a,\infty)}\mathcal{M} - \mathrm{Rk}_{(a-1,\infty)}\mathcal{M} & 0 \le a < b = \infty \end{cases}. \quad (4.1)$$

■

By allowing negative signs in the sum of the rank invariant in Proposition 4.1, [BOO22] extends the result to multiparameter persistence modules. The intervals in the decomposition are then called the *signed barcode* of the persistence module.

4.3 Rank decompositions

In their paper [BOO22] introducing signed barcodes, the authors begin by proving that under mild finiteness conditions, arbitrary functions of the form $r : \mathcal{I} \to \mathbb{Z}$ can be decomposed as a signed sum of rank invariants of interval modules. We first need to extend the notion of the rank invariant to intervals of arbitrary posets.

Definition 4.3 (Generalized rank invariant). *Let $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ be a persistence module and let $\mathcal{I} \subset \boldsymbol{P}$ be a collection of intervals of $\boldsymbol{P}$. Then the* generalized rank invariant *of $\mathcal{M}$ over $\mathcal{I}$ is the map*

$$\begin{aligned} &Rk_{\mathcal{I}}\mathcal{M} : \mathcal{I} \to \mathbb{Z} \\ &\qquad I \mapsto Rk_{\mathcal{I}}\mathcal{M}(I) = rank(\varprojlim \mathcal{M}\big|_I \to \varinjlim \mathcal{M}\big|_I) \end{aligned}$$

where $\mathcal{M}\big|_I$ denotes the restriction of $\mathcal{M}$ to the interval I.

The map $\varprojlim \mathcal{M}\big|_I \to \varinjlim \mathcal{M}\big|_I$ exists, and is in fact canonical, since I is connected and convex (see Theorem 2.1 of [Set]). Furthermore, since $\mathcal{M}$ is assumed to be pointwise finite dimensional and the morphism $\varprojlim \mathcal{M}\big|_I \to \varinjlim \mathcal{M}\big|_I$ factors through the finite dimensional internal spaces of $\mathcal{M}\big|_I$, $\mathrm{rank}(\varprojlim \mathcal{M}\big|_I \to \varinjlim \mathcal{M}\big|_I)$ is finite. An important - and easy to visualize - collection of intervals are rectangles.

Definition 4.4 (Rectangles). *A* rectangle *or* segment *of a poset $\boldsymbol{P}$ is an interval of the form*

$$\langle s, t\rangle = \{u \in \boldsymbol{P} | s \le u \le t\}.$$

Example 5 (The usual rank invariant). Let $\mathcal{I}$ be the collection of rectangles in a poset $\mathbf{P}$. We call the generalized rank invariant over $\mathcal{I}$ the *usual rank invariant*. This is due to the fact, that for a rectangle $\langle s, t\rangle$, we have $\varprojlim I = s$ and $\varinjlim I = t$, making the usual rank invariant an intuitive extension of the one parameter rank invariant. Just like the single parameter rank invariant, the usual rank invariant at a rectangle I can be interpreted as the number of features that are alive between s and t.

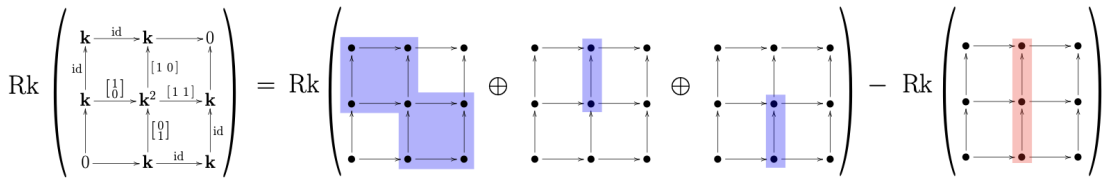


Figure 4.5: An invalid rank decomposition of the usual rank invariant with blue intervals belonging to $\mathcal{R}$ and red intervals belonging to $\mathcal{S}$ (from [BOO22]).

Definition 4.5 (Rank decomposition). *Let $\mathcal{I}$ be a collection of intervals in $\boldsymbol{P}$. A rank decomposition of a function $r : \mathcal{I} \to \mathbb{Z}$ is a pair of multisets $(\mathcal{R}, \mathcal{S})$ composed of intervals in $\mathcal{I}$ such that*

$$r = Rk_{\mathcal{I}}k_{\mathcal{R}} - Rk_{\mathcal{I}}k_{\mathcal{S}} \tag{4.2}$$

where $k_{\mathcal{R}} = \bigoplus_{R \in \mathcal{R}} k_R$ and similarly for $k_{\mathcal{S}}$. The decomposition is minimal if $\mathcal{R} \cap \mathcal{S} = \emptyset$.

Note that the equation above is an equation relating functions $\mathcal{I} \to \mathbb{Z}$ which means it holds if and only if

$$r(I) = \mathrm{Rk}_{\mathcal{I}}k_{\mathcal{R}}(I) - \mathrm{Rk}_{\mathcal{I}}k_{\mathcal{S}}(I) \qquad \forall I \in \mathcal{I}.$$

We are interested in the special case where $r = \mathrm{Rk}_{\mathcal{I}}\mathcal{M}$, where $\mathcal{M} : \mathbf{P} \to \mathbf{vec}$. Before proceeding to proving the existence of rank decompositions, let us look at an example to illustrate the importance of the collection $\mathcal{I}$ in Definition 4.5.

Example 6. Consider again the indecomposable module on the 3×3 grid G depicted in Figure 4.4. Let $\mathcal{I}$ denote the collection of rectangles in G and let $\mathrm{Rk}\mathcal{M}$ be the usual rank invariant. A rank decomposition $(\mathcal{R}, \mathcal{S})$ of $\mathrm{Rk}\mathcal{M}$ must consist only of rectangles. This excludes the decomposition depicted in Figure 4.5. This is due to the fact that the large blue interval on the left is clearly not a rectangle. This decomposition is perhaps the first one that might come to mind, if we ignore the condition that intervals must come from the underlying family. Intuitively, one begins by adding the large non-rectangular interval J to $\mathcal{R}$, because it is precisely the interval on which $\mathcal{M}$ is supported. However, we have $\mathrm{rank}(k \xrightarrow{\begin{pmatrix}0 & 1\end{pmatrix}^T} k^2 \xrightarrow{\begin{pmatrix}1 & 0\end{pmatrix}} k) = 0$ but on the same interval in J the rank is 1 since we have identity maps throughout. To rectify this we add the vertical interval to $\mathcal{S}$, to ensure the sum is 0. But this results in 0 spaces throughout the vertical interval and causes the ranks of both $k \xrightarrow{\begin{pmatrix}0 & 1\end{pmatrix}^T} k^2$ and $\xrightarrow{\begin{pmatrix}1 & 0\end{pmatrix}} k$ to be zero, while both of these ranks are 1 in the $\mathcal{M}$. To correct for this we add the these two maps individually to $\mathcal{R}$. This maintains the rank of the composition at 0, while ensuring that the ranks of each of the smaller intervals is 1. Hence, the decomposition satisfies Equation 4.2 but since not all intervals of $\mathcal{R}$ are rectangles, it is not a rank decomposition of the usual rank invariant. A valid decomposition is depicted in Figure 4.6. We will later compute it explicitly using Corollary 4.2. Note that the condition requiring the decomposition to be over the same family of intervals as the rank invariant itself is essential, as it ensures uniqueness.

If a rank decomposition exists, then it can be made minimal. Moreover, such a minimal decomposition is unique. To prove this we need the following important fact about the rank invariant, which we will use again when proving existence.

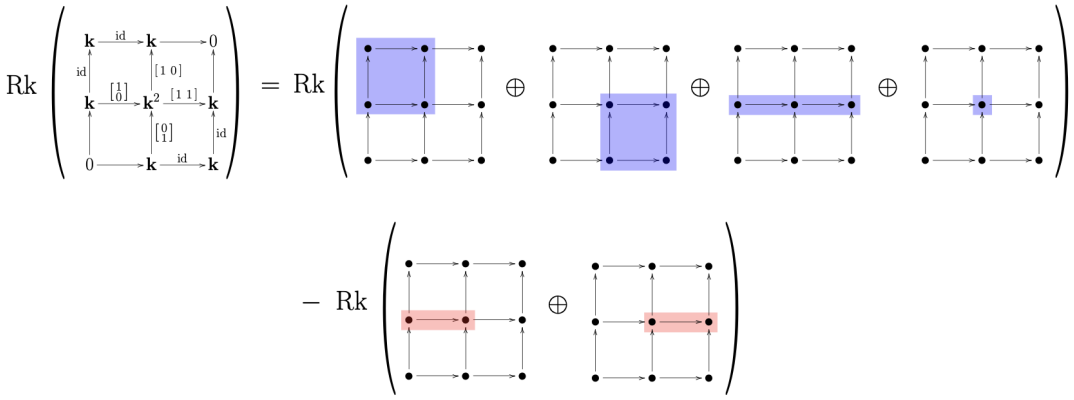


Figure 4.6: A rank decomposition of the usual rank invariant with blue rectangles belonging to $\mathcal{R}$ and red rectangles belonging to $\mathcal{S}$ (from [BOO22]).

Proposition 4.3 (Ranks and multiplicities). *Let $\mathcal{I}$ be a collection of intervals of a poset $\boldsymbol{P}$ and $\mathcal{R}$ be a multiset of intervals taken from $\mathcal{I}$. Then,*

$$Rk_{\mathcal{I}} k_{\mathcal{R}}(J) = \#\{R \in \mathcal{R} | J \subseteq R\} \qquad \forall J \in \mathcal{I}.$$

Proof. Let $J \in \mathcal{I}$ and $R \in \mathcal{R}$ be arbitrary, then $k_R\big|_J$ is simply the indicator module $k_{R\cap J}$. Keeping this in mind, we write the following decomposition of $k_{\mathcal{R}}$

$$k_{\mathcal{R}}\big|_J = \bigoplus_{R\in\mathcal{R}} k_{R\cap J} = \bigoplus_{\substack{R\in\mathcal{R}\\ J\subseteq R}} k_J \oplus \bigoplus_{R\in\tilde{\mathcal{R}}} k_R,$$

where $\tilde{\mathcal{R}} = \{R \in \mathcal{R} | R \subsetneq J\}$. Since $\mathrm{Rk}_{\mathcal{I}} k_R(J) = \mathrm{Rk}_{\mathcal{I}} k_J(J) = 1$ for all $J \subseteq R$ and the rank commutes with finite sums, we have that $\mathrm{Rk}_{\mathcal{I}} \bigoplus_{\substack{R\in\mathcal{R}\\ J\subseteq R}} k_R(J) = \#\{R \in \mathcal{R} | J \subseteq R\}$. Consequently, we need to show that $\mathrm{Rk}_{\mathcal{I}} \bigoplus_{R\in\tilde{\mathcal{R}}} k_R(J) = 0$. It suffices to show that $\mathrm{Rk}_{\mathcal{I}} k_R(J) = 0$ for all $R \subsetneq J$. In this case either $\varprojlim k_R\big|_J = 0$ or $\varprojlim k_R\big|_J = 0$ (this is best understood using the Example 7). We can write $\tilde{\mathcal{R}} = \mathcal{R}' \cup \mathcal{R}''$, where $\mathcal{R}'$ consists only of subintervals where the limit $\varprojlim k_R\big|_J = 0$ vanishes and $\mathcal{R}''$ where the colimit $\varinjlim k_R\big|_J = 0$ vanishes. Of course, this decomposition may not be unique, since there might exist subintervals where both the limit and colimit vanish. This does not matter since all we simply need to show that both $\mathrm{Rk}_{\mathcal{I}} \bigoplus_{R\in\mathcal{R}'} k_R(J) = 0$ and $\mathrm{Rk}_{\mathcal{I}} \bigoplus_{R\in\mathcal{R}''} k_R(J) = 0$. Using the fact that the direct sum commutes with the colimit, we have

$$\begin{aligned}
\mathrm{Rk}_{\mathcal{I}} \bigoplus_{R\in\mathcal{R}''} k_R(J) &= \mathrm{rank}(\varprojlim \bigoplus_{R\in\mathcal{R}''} k_R\big|_J \to \varinjlim \bigoplus_{R\in\mathcal{R}''} k_R\big|_J) \\
&= \mathrm{rank}(\varprojlim \bigoplus_{R\in\mathcal{R}''} k_R\big|_J \to \bigoplus_{R\in\mathcal{R}''} \varinjlim k_R\big|_J) \\
&= \mathrm{rank}(\varprojlim \bigoplus_{R\in\mathcal{R}''} k_R\big|_J \to 0) \\
&= 0.
\end{aligned}$$

Although the direct sum does not commute with limits, the direct product does. Expressing the direct sum as a subrepresentation of the direct product allows us to use a similar method to prove

$\mathrm{Rk}_{\mathcal{I}} \bigoplus_{R \in \mathcal{R}'} k_R(J) = 0$. Since the rank commutes with finite direct sums, these two statements together imply that $\mathrm{Rk}_{\mathcal{I}} k_R(J) = 0$ for all $R \subsetneq J$, concluding the proof. ■

Example 7. Let $\mathbf{P} = \mathbb{Z}$, $\mathcal{I}$ be the standard collection of intervals, $J = (a, b)$, and $R = (c, d)$ with $c = a \leq d < b$. The restriction $k_R\big|_J$ is depicted below

$$\begin{array}{ccccccccccc} k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 \\ a = c & & & & d & & & & & & b \end{array}$$

In this case, $\varinjlim k_R\big|_J = 0$ and since maps to 0 always have rank 0, $\mathrm{Rk}_{\mathcal{I}} k_R(J) = 0$. In category theory language, we say R *is closed under predecessors in* I. On the other hand, if $a < c \leq d = b$,we obtain

$$\begin{array}{ccccccccccc} 0 & \longrightarrow & 0 & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k \\ a & & & & c & & & & & & d = b \end{array}$$

and consequently $\varprojlim k_R\big|_J = 0$. In this case R is *closed under successors* in I. For $R \subsetneq J$ to hold it cannot be closed under both successors and predecessors in I, and hence either the limit or colimit will always be zero.

Proposition 4.4 (Minimal rank decompositions are unique). *Let $\mathcal{I}$ be a collection of intervals in a poset $\boldsymbol{P}$ and $r : \mathcal{I} \to \mathbb{Z}$ be a function. Assume that there exists a rank decomposition $(\mathcal{R}, \mathcal{S})$ of r. Then, we can obtain a unique minimal rank decomposition r.*

Proof. A rank decomposition $(\mathcal{R}, \mathcal{S})$ is unique if $\mathcal{R} \cap \mathcal{S} = \emptyset$. So, given $(\mathcal{R}, \mathcal{S})$, we can simply remove common elements in the intersection to obtain a disjoint pair $(\mathcal{R}^*, \mathcal{S}^*)$. Note that we need to remove elements from the intersection with the appropriate multiplicity, i.e. if I appears in $\mathcal{R}$ with multiplicity n and in $\mathcal{S}$ with multiplicity m with $n \geq m$, we include I with multiplicity $n - m$ in $\mathcal{R}^*$ and with multiplicity 0 in $\mathcal{S}$. The resultant pair $(\mathcal{R}^*, \mathcal{S}^*)$ is still a rank decomposition, since removing common elements in this manner does not alter the difference of their ranks.

Suppose now that $(\mathcal{R}', \mathcal{S}')$ and $(\mathcal{R}, \mathcal{S})$ are two distinct rank decompositions of r. We can think of a rank decomposition as a $\mathbb{Z}$-linear combination of the rank invariants of interval modules $\mathrm{Rk}_{\mathcal{I}} k_I$ with $I \in \mathcal{I}$. Indeed, given an interval $I \in \mathcal{R}$, set $\alpha_I = \mathrm{mult}_I \mathcal{R} - \mathrm{mult}_I \mathcal{S}$. This yields $r = \sum_{I \in \mathcal{I}} \alpha_I \mathrm{Rk}_{\mathcal{I}} k_I$. Two distinct minimal rank decompositions of r result in two such $\mathbb{Z}$-linear realizations of r. Let $r = \sum_{I \in \mathcal{I}} \alpha_I \mathrm{Rk}_{\mathcal{I}} k_I$ and $r = \sum_{I \in \mathcal{I}} \alpha'_I \mathrm{Rk}_{\mathcal{I}} k_I$ be the $\mathbb{Z}$-linear combinations corresponding to $(\mathcal{R}, \mathcal{S})$ and $(\mathcal{R}', \mathcal{S}')$ respectively. Distinctness means that the set $\{I \in \mathcal{I} | \alpha_I \neq \alpha'_I\}$ is nonempty. Let J be the maximal element (with respect to inclusion) of this set. We have

$$\sum_{I \in \mathcal{I}} \alpha_I \mathrm{Rk}_{\mathcal{I}} k_I - \sum_{I \in \mathcal{I}} \alpha'_I \mathrm{Rk}_{\mathcal{I}} k_I = r - r = 0$$

$$\sum_{I \in \mathcal{I}} (\alpha_I - \alpha'_I) \mathrm{Rk}_{\mathcal{I}} k_I = 0$$

Evaluating the sum at the maximal element J we obtained above yields

$$\sum_{I \in \mathcal{I}} (\alpha_I - \alpha'_I) \mathrm{Rk}_{\mathcal{I}} k_I(J)$$

But by Proposition 4.3,

$$\mathrm{Rk}_{\mathcal{I}} k_I(J) = \begin{cases} \mathrm{Rk}_J(k_I) = 1 & I \subseteq J \\ 0 & \text{otherwise.} \end{cases}$$

By maximality of J,

$$\sum_{I \in \mathcal{I}} (\alpha_I - \alpha'_I) \mathrm{Rk}_{\mathcal{I}} k_I(J) = (\alpha_J - \alpha'_J) \neq 0$$

which is a contradiction. ■

We now follow sections 3, and 5.1 of [BOO22] to prove that functions $r : \mathcal{I} \to \mathbb{Z}$ with locally finite support admit unique minimal rank decompositions, that they are easily computed using Möbius inversions, and finally discuss the case we are interested in, namely when r is the rank invariant of a finitely presentable persistence module.

4.3.1 Möbius inversions

One of the most elegant ways to prove the existence of rank decompositions is to use Möbius inversions. Under certain finiteness assumptions, these inversions also provide concrete formulas to compute rank decompositions. Recall first that a collection of intervals can be interpreted as a poset with inclusion $\subset$ as the order.

Definition 4.6 (Incidence algebra). *Let $\mathcal{I}$ be a locally finite collection of intervals, that is $\langle I, J \rangle = \{K \in \mathcal{I} : I \subseteq K \subseteq J\}$ is finite for all $I \subseteq J \in \mathcal{I}$. The incidence algebra over $\mathcal{I}$ consists of all functions of the form*

$$\begin{aligned} f : \{\subseteq\} &\to \mathbb{Z} \\ (J, I) &\mapsto f(J, I) \qquad\qquad J \supseteq I \end{aligned}$$

with multiplication given by convolution

$$(f \star g)(J, I) = \sum_{J \supseteq K \supseteq I} f(J, K) g(K, I).$$

The multiplicative unit is $\mathbf{1}_{J=I}$ which sends identical pairs to 1 and all other pairs to 0.

Incidence algebras are useful, because functions $r : \mathcal{I} \to \mathbb{Z}$ with locally finite support define a right module over the incidence algebra over $\mathcal{I}$ with multiplication being given by

$$(r \star f)(J, I) = \sum_{J \supseteq I} r(J) f(J, I).$$

Definition 4.7 (Möbius Inversion). *The Möbius inversion μ is defined as the multiplicative inverse of*

$$\zeta(J, I) = \begin{cases} 1 & J \supseteq I \\ 0 & \textit{otherwise} \end{cases}.$$

It can be computed recursively using

$$\mu(J, I) = \begin{cases} 1 & J = I \\ -\sum_{J \supseteq K \supsetneq I} \mu(J, K) & \textit{otherwise.} \end{cases}$$

We now want to express $r : \mathcal{I} \to \mathbb{Z}$ with locally finite support as the Möbius inversion of an easily calculable function. To that end consider a multiset of intervals $\mathcal{R}$ and the corresponding multiplicity function

$$\begin{aligned} \mathrm{mult}\mathcal{R} : \mathcal{I} &\to \mathbb{Z} \\ I &\mapsto \mathrm{mult}_I\mathcal{R} = \text{ multiplicity of } I \text{ in } \mathcal{R}. \end{aligned}$$

Rephrasing Proposition 4.3 using $\mathrm{mult}_{\mathcal{R}}$ leads to the following lemma.

Lemma 4.1 (Ranks and multiplicities). *Let $\mathcal{R}$ be a multiset of intervals of a poset P and $\mathcal{I}$ be a collection of intervals. Then,*

$$Rk_{\mathcal{I}}k_{\mathcal{R}} = mult\mathcal{R} \star \zeta.$$

Theorem 4.1 (Existence of rank decompositions). *Let $\mathcal{I}$ be a locally finite collection of intervals, and let $r : \mathcal{I} \to \mathbb{Z}$ have locally finite support. A pair $(\mathcal{R}, \mathcal{S})$ of intervals of I is a rank decomposition of r if and only if*

$$r \star \mu = mult\mathcal{R} - mult\mathcal{S}.$$

In particular $\mathcal{R}$ and $\mathcal{S}$ are given by

$$\mathcal{R} = \bigsqcup_{\substack{I \in \mathcal{I} \\ (r\star\mu)(I) > 0}} I^{(r\star\mu)(I)}, \qquad \mathcal{S} = \bigsqcup_{\substack{I \in \mathcal{I} \\ (r\star\mu)(I) < 0}} I^{(r\star\mu)(I)}$$

where $I^{(r\star\mu)(I)}$ denotes I being included with multiplicity $(r \star \mu)(I)$.

Proof. We have

$$\begin{aligned} (\mathcal{R}, \mathcal{S}) \text{ is a rank decomposition of } r &\iff r = \mathrm{Rk}_{\mathcal{I}}k_{\mathcal{R}} - \mathrm{Rk}_{\mathcal{I}}k_{\mathcal{S}} \\ &\iff r = \mathrm{mult}\mathcal{R} \star \zeta - \mathrm{mult}\mathcal{S} \star \zeta \\ &\iff r = (\mathrm{mult}\mathcal{R} - \mathrm{mult}\mathcal{S}) \star \zeta \\ &\iff r \star \mu = \mathrm{mult}\mathcal{R} - \mathrm{mult}\mathcal{S}. \end{aligned}$$

■

Imposing extra conditions on the collection of intervals $\mathcal{I}$ allows us to derive a more explicit formula for μ and consequently for the multiplicities of intervals in the rank decomposition of a given locally finite function. These conditions are usually satisfied for collections that appear in applications. In order to state the relevant proposition, we need the following definition.

Definition 4.8 (Join). *Let $\boldsymbol{P}$ be a poset and $a, b \in \boldsymbol{P}$. A* join *of a and b is an element $j \in \boldsymbol{P}$ such that*

1. *$x \leq j$ and $y \leq j$.*
2. *For any $w \in \boldsymbol{P}$ with $x \leq w$ and $y \leq w$, we have $j \leq w$.*

In other words, j is the least upper bound of x and y. A join need not always exist, but is unique when it does and is denoted by $x \vee y$.

Proposition 4.5 (Inclusion-exclusion formula for Möbius inversions). *Let $\mathcal{I}$ be a locally finite collection of intervals and assume that for each $I \in \mathcal{I}$, there exists a finite set $I^+ \subseteq \mathcal{I}$ such that*

$$\{J \in \mathcal{I} | J \supsetneq I\} = \{J \in \mathcal{I} | \exists K \in I^+ : J \supseteq K\}$$

and *that any subset I^+ has a join in $\mathcal{I}$. Then*

$$\mu(J, I) = \begin{cases} 1 & J = I \\ \sum_{\substack{x \subset I^+ \\ \vee x = J}} (-1)^{\#x} & \text{otherwise.} \end{cases}$$

The first condition of the assumption states that any element in the strict upset of I is in the upset of some element of I^+. The second assumption ensures that least upper bound of any subset of I^+ is also an interval, allowing us to replace elements of the upset of I with joins of elements from I^+. We will see explicit examples of I^+ in the next section.

Proof. ■

Applying the formula for μ obtained in Proposition 4.5 to the equation in Theorem 4.1 yields the following corollary

Corollary 4.1. *Let $\mathcal{I}$ be a locally finite collection of intervals and assume that for each $I \in \mathcal{I}$, there exists a finite set $I^+ \subseteq \mathcal{I}$ such that*

$$\{J \in \mathcal{I} | J \supsetneq I\} = \{J \in \mathcal{I} | \exists K \in I^+ : J \supseteq K\}$$

and *that any subset I^+ has a join in $\mathcal{I}$. Then, a pair $(\mathcal{R}, \mathcal{S})$ is a rank decomposition of a function $r : \mathcal{I} \to \mathbb{Z}$ with locally finite support if and only if*

$$\mathit{mult}_I \mathcal{R} - \mathit{mult}_I \mathcal{S} = r(I) - \sum_{\emptyset \neq x \subseteq I^+} (-1)^{\#x} r(\vee x) \qquad \forall I \in \mathcal{I}.$$

Proof. By Theorem 4.1,

$$r \star \mu = \mathrm{mult}\mathcal{R} - \mathrm{mult}\mathcal{S}.$$

Consequently,

$$\begin{aligned} \mathrm{mult}_I \mathcal{R} - \mathrm{mult}_I \mathcal{S} &= (r \star \mu)(I) && \forall I \in \mathcal{I} \\ &= \sum_{J \supseteq I} r(J)\mu(J, I) && \forall I \in \mathcal{I} \\ &= r(I)\mu(I, I) + \sum_{J \supsetneq I} r(J)\mu(J, I) && \forall I \in \mathcal{I} \\ &= r(I) + \sum_{J \supsetneq I} r(J) \sum_{\substack{x \subseteq I^+ \\ \vee x = J}} (-1)^{\#x} && \forall I \in \mathcal{I} \\ &= r(I) + \sum_{\emptyset \neq x \subseteq I^+} r(\vee x)(-1)^{\#x} && \forall I \in \mathcal{I} \end{aligned}$$

■

Remark 4.1. In practice, we use Corollary 4.1 to compute rank decompositions. Given an interval I, $\text{mult}_I\mathcal{R} - \text{mult}_I\mathcal{S}$ is the multiplicity of I in the decomposition. If it is positive, we include it with the appropriate multiplicity in $\mathcal{R}$ and if it is negative we include it in $\mathcal{S}$. Hence, computing the rank decomposition of r amounts to evaluating the right-hand side for every interval $I \in \mathcal{I}$. There are certain tricks one might use to reduce the number of times $r(I)$ is actually computed, but it remains highly inefficient. The choice of I^+ plays an important role, since a small I^+ could substantially reduce the sum and hence make computing the decomposition more efficient.

4.4 Applying rank decompositions to rank invariants of multiparameter persistence modules.

We can now apply the results we obtained to the specific case when r is the rank invariant of a multiparameter persistence module. In applications, we almost always deal with persistence modules indexed over finite posets such as a finite grid $G = \prod_{i=1}^{d}[1, n_i] \subset \mathbb{R}^d$. The collection $\mathcal{I}$ is most often taken to be the collection of rectangles in G, but other options such as hook modules or even the collection of all intervals is possible. Since G is finite, these collections are trivially locally finite and the rank invariant indexed over any of these collections has locally finite support. Consequently, Theorem 4.1 yields:

Corollary 4.2. *Let $\mathcal{I}$ be a collection of intervals in a finite grid $G = \prod_{i=1}^{d}[1, n_i] \subset \mathbb{R}^d$ and $\mathcal{M} : G \to \textbf{vec}$ be a persistence module. The rank invariant $Rk_{\mathcal{I}}\mathcal{M}$ admits a unique rank decomposition $(\mathcal{R}, \mathcal{S})$ over $\mathcal{I}$.*

If $\mathcal{I}$ is the collection of rectangles we can determine I^+ and apply Proposition 4.5 to obtain an explicit formula to determine $(\mathcal{R}, \mathcal{S})$.

Corollary 4.3. *Let $\mathcal{I}$ be a collection of rectangles in a finite grid $G = \prod_{i=1}^{d}[1, n_i] \subset \mathbb{R}^d$ and $\mathcal{M} : G \to \textbf{vec}$ be a persistence module. The usual rank invariant $Rk_{\mathcal{I}}\mathcal{M}$ admits a unique rank decomposition $(\mathcal{R}, \mathcal{S})$ comprised of rectangles. Moreover, the $(\mathcal{R}, \mathcal{S})$ is given by*

$$\mathcal{R} = \bigsqcup_{\substack{\langle s,t\rangle, s\leq t \\ \alpha_{\langle s,t\rangle}>0}} \langle s,t\rangle^{\alpha_{\langle s,t\rangle}}, \qquad \mathcal{S} = \bigsqcup_{\substack{\langle s,t\rangle, s\leq t \\ \alpha_{\langle s,t\rangle}<0}} \langle s,t\rangle^{\alpha_{\langle s,t\rangle}}$$

where

$$\alpha_{\langle s,t\rangle} = \sum_{\substack{s'\leq s \\ \|s-s'\|_\infty \leq 1}} \sum_{\substack{t'\geq t \\ \|t-t'\|_\infty \leq 1}} (-1)^{\|s-s'\|_1+\|t-t'\|_1} rank(M(s') \to M(t')). \tag{4.3}$$

Proof. Let $I = \langle s, t\rangle$ with $s \leq t$ be a rectangle. Then, set

$$I^+ = \{\langle s', t'\rangle | s' \leq s, t' \geq t : \|s' - s\|_\infty \leq 1, \|t' - t\|_\infty \leq 1\}.$$

Clearly, any rectangle that contains $\langle s, t\rangle$ must contain at least one of the rectangles in I^+ and the join of any subset of I^+ is trivially a rectangle. Applying Proposition 4.5 using I^+ yields the desired result. ∎

Figure 4.6 is not very compact and for other modules, the decomposition might involve many more rectangles. Moreover, many of the rectangles intersect, so plotting the entire decomposition on a single grid is not particularly useful. Instead, for each rectangle $\langle a, b\rangle$ in $\mathcal{R}$ we only plot the

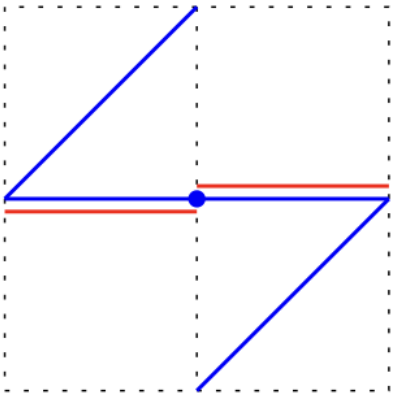


Figure 4.7: Signed barcode of the decomposition in Example 8 (from [BOO22]).

diagonal connecting a and b. Since the rectangle is completely characterized by this diagonal, we do not lose any information. We plot the diagonals arising from rectangles in $\mathcal{R}$ blue and those arising from rectangles in $\mathcal{S}$ in red to differentiate between the two. This motivates the term *signed barcode*.

Definition 4.9 (Signed barcode). *Let $\mathcal{I}$ be a collection of rectangles in a poset $\boldsymbol{P}$ and $(\mathcal{R}, \mathcal{S})$ be a rank decomposition of a function $r : \mathcal{I} \to \mathbb{Z}$. Then, the corresponding* signed barcode *is the pair (B, R), where B is the collection of diagonals of rectangles in $\mathcal{R}$ and R is the collection of diagonals of rectangles in $\mathcal{S}$.*

Example 8 (Computing the rank decomposition). As promised, we now compute the rank decomposition of the usual rank invariant of $\mathcal{M}$ depicted in Figure 4.4 using Equation 4.3. We index the underlying $G = 3 \times 3$ grid in the natural way. Begin with rectangles of the form $\langle (0,0), b \rangle$, where $b \geq (0,0)$. Any rectangle in $\langle (0,0), b \rangle^+$ would also be of this form, since there is no $s < (0,0)$ in G. Since $\mathcal{M}((0,0)) = 0$, $\mathrm{Rk}_{\mathcal{I}}\mathcal{M}(\langle (0,0), b \rangle) = 0$ for all $b \geq 0$. Consequently, $\alpha_{\langle (0,0), b \rangle = 0}$. A similar argument can be used to show $\alpha_{\langle a, (2,2) \rangle} = 0$ for all $a \leq (2,2)$, since the maps would always end in 0. For the remaining rectangles we compute the multiplicities using the formula. We exclude rectangles that start at $(0,0)$ or end at $(2,2)$ from all the I^+'s in the

computation below for the reasons mentioned above.

$$
\begin{aligned}
\alpha_{\langle(1,0),(2,0)\rangle} &= \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,0)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) = 1 - 1 = 0 \\
\alpha_{\langle(1,0),(2,1)\rangle} &= \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) = 1 \\
\alpha_{\langle(1,0),(1,1)\rangle} &= \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(1,1)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(1,2)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) \\
&= 1 - 0 - 1 = 0 \\
\alpha_{\langle(1,0),(1,2)\rangle} &= \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) = 0 \\
\alpha_{\langle(2,0),(2,1)} &= \operatorname{rank}(\mathcal{M}(2,0), \mathcal{M}(2,1)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) = 1 - 1 = 0 \\
\alpha_{\langle(0,1),(1,1)\rangle} &= \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,1)) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(2,1)) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,2)) \\
&= 1 - 1 - 1 = -1 \\
\alpha_{\langle(0,1),(2,1)\rangle} &= \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(2,1)) = 1 \\
\alpha_{\langle(0,1),(0,2)\rangle} &= \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(0,2)) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,2)) = 1 - 1 = 0 \\
\alpha_{\langle(0,1),(1,2)\rangle} &= \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,2)) = 1 \\
\alpha_{\langle(1,1),(2,1)\rangle} &= \operatorname{rank}(\mathcal{M}(1,1) \to \mathcal{M}(2,1)) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(2,1)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) \\
&= 1 - 1 - 1 = -1 \\
\alpha_{\langle(1,1),(1,2)\rangle} &= \operatorname{rank}(\mathcal{M}(1,1) \to \mathcal{M}(1,2)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(1,2)) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,2)) \\
&= 1 - 0 - 1 = 0 \\
\alpha_{\langle(0,2),(1,2)\rangle} &= \operatorname{rank}(\mathcal{M}(0,2) \to \mathcal{M}(1,2)) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,2)) = 1 - 1 = 0
\end{aligned}
$$

The only thing remaining to check is the multiplicity of the singletons $\alpha_{a,a}$ for each $a = (a_x, a_y) \in G$.

$$
\begin{aligned}
\alpha_{(1,0)} &= \operatorname{rank}(\mathrm{id}_{\mathcal{M}(1,0)}) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,0)) - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(1,1)) + \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,0)) \\
&= 1 - 1 - 1 + 1
\end{aligned}
$$

One can verify that the multiplicity of every singleton is 0 except $(1,1)$. In that case we have

$$
\begin{aligned}
\alpha_{(1,1)} &= \operatorname{rank}(\mathrm{id}_{\mathcal{M}(1,1)}) - \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,1)) - \operatorname{rank}(\mathcal{M}(1,1) \to \mathcal{M}(2,1)) \\
&\quad - \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(1,1)) - \operatorname{rank}(\mathcal{M}(1,1) \to \mathcal{M}(1,2)) + \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(2,1)) \\
&\quad + \operatorname{rank}(\mathcal{M}(0,1) \to \mathcal{M}(1,2)) + \operatorname{rank}(\mathcal{M}(1,0) \to \mathcal{M}(2,1)) \\
&= 2 - 1 - 1 - 1 - 1 + 1 + 1 + 1 = 1
\end{aligned}
$$

Consequently, we have

$$
\begin{aligned}
\mathcal{R} &= \{\langle(1,0),(2,1)\rangle, \langle(0,1),(2,1)\rangle, \langle(0,1),(1,2)\rangle, \langle(1,1),(1,1)\rangle\} \\
\mathcal{S} &= \{\langle(0,1),(1,1)\rangle, \langle(1,1),(2,1)\rangle\}.
\end{aligned}
$$

This is precisely the decomposition depicted in Figure 4.6.

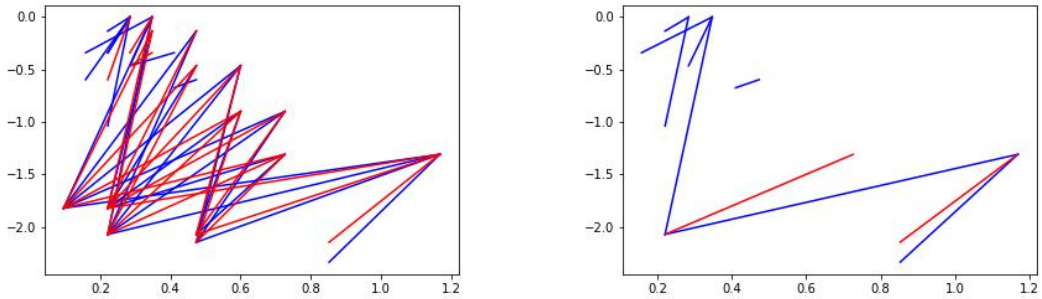


Figure 4.8: Signed barcode in dimension 1 of the circle with noise in the center. Left: raw signed barcode with positive bars colored blue and negative bars colored red. Right: Cleaned up signed barcode, where positive and negative bars that are extremely close to each other and hence cancel one another out have been removed for better readability.

The 1-dimensional signed barcode of the noisy circle depicted in Figure 4.2 is in Figure 4.8. The figure on the left is a bit hard to read due to the large number of bars that almost overlap one another. Since a blue bar and a red bar that represent the same rectangle "cancel" each other out, we can remove these bars from the picture to obtain a more interpretable figure like the one on the right. Concretely this is done as follows. Let $\epsilon > 0$ and let $\langle b_1, b_2\rangle$, $\langle r_1, r_2\rangle$ be rectangles in $\mathcal{R}$ and $\mathcal{S}$ respectively. If $\|r_1 - b_1\|^2 \leq \epsilon$ and $\|r_2 - b_2\|^2 \leq \epsilon$, then we remove the corresponding bars from the figure. In other words, if the both end points of a red and blue bar are ϵ-close to one another we remove the bar from the figure. Now that we have a cleaner figure, observe how the bars in Figure 4.8 capture the evolution of the cycles through the bifiltration of the circle in Figure 4.3.

4.5 Hilbert decompositions

Over the past few sections, we explored how the rank invariant can be used to construct a signed decomposition of persistence modules. Such a decomposition can also be constructed using the Hilbert function instead of the rank invariant.

Definition 4.10 (Hilbert function). *Let $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ be a persistence module. Then the associated Hilbert function is*

$$\begin{aligned} Hil\mathcal{M} : \boldsymbol{P} &\to \mathbb{Z} \\ i &\mapsto \dim(\mathcal{M}(i)). \end{aligned}$$

Remark 4.2 ($\mathrm{Rk}\mathcal{M}$ is strictly finer than $\mathrm{Hil}(\mathcal{M})$). The Hilbert function can be recovered from the rank invariant, since $\dim(\mathcal{M}(i)) = \mathrm{rank}(\mathcal{M}(i) \to \mathcal{M}(i))$.

Definition 4.11 (Free persistence module). *A persistence module $\mathcal{P} : \boldsymbol{P} \to \boldsymbol{vec}$ is* free *if it there exists a set $\boldsymbol{barc}(\mathcal{P}) \subset \boldsymbol{P}$ such that*

$$\mathcal{P} \cong \bigoplus_{i \in \boldsymbol{barc}(\mathcal{P})} k_{i^+}$$

where $i^+ = \{j \in \boldsymbol{P} | j \geq i\}$. We say that $\mathcal{P}$ is of finite rank *if $\boldsymbol{barc}(\mathcal{P})$ is finite.*

In other words, free persistence modules are isomorphic to direct sums of indicator modules supported on upsets of elements of the poset.

Remark 4.3. The *rank* of a free persistence module is unrelated to its *rank invariant*. This is an unfortunate coincidence.

Remark 4.4. Free persistence modules are interval decomposable with the intervals being the upsets of the elements in their barcode.

We can now define the notion of a Hilbert decomposition.

Definition 4.12 (Hilbert decomposition). *The* Hilbert decomposition *of a function $\eta : P \to \mathbb{Z}$ consists of a pair of free persistence modules $(\mathcal{P}, \mathcal{Q})$ of finite rank such that $\eta = Hil\mathcal{P} - Hil\mathcal{Q}$. The decomposition is* minimal *if $\mathcal{P} \cap \mathcal{Q} = \emptyset$ The* Betti signed barcode *of η is $(\boldsymbol{barc}(\mathcal{P}), \boldsymbol{barc}(\mathcal{Q}))$.*

We are interested in the case $\eta = \text{Hil}\mathcal{M}$, where $\mathcal{M} : \mathbf{P} \to \mathbf{vec}$ is finitely presented. We prove that $\text{Hil}\mathcal{M}$ admits a Hilbert decomposition by using projective resolutions to "approximate" $\mathcal{M}$ and then applying a Möbius inversion. Free persistence modules are projective in the sense of homological algebra, i.e. they are the projective objects in the category of persistence modules. The following proof is adapted from [BBH23].

Proposition 4.6 (Free persistence modules are projective). *Let $i \in \boldsymbol{P}$. Then, $M \cong k_{i^+}$ is a projective in the category of persistence modules.*

Proof. Assume $M \cong k_{i^+}$, then we need to show that the functor $\hom(k_{i^+}, -)$ is exact. It suffices to construct a natural isomorphism $\hom(k_{i^+}, -) \simeq (-)(i)$. To that end consider

$$\begin{aligned} \Psi : \hom(k_{i^+}, -) &\to (-)(i) \\ (k_{i^+} \xrightarrow{\varphi} \mathcal{N}) &\mapsto \varphi_i(k_{i^+}) \end{aligned}$$

Recall, that $(k_{i^+} \xrightarrow{\varphi} \mathcal{N})$ is itself a natural transformation and φ_i is the map $k_{i^+}(i) \to \mathcal{N}(i)$. This map is clearly surjective. Injectivity: $\varphi \in \ker \Psi$ implies $\Psi(\varphi) = \varphi_i(k) = 0$. Let $j \geq i$, then we have the following commutative diagram

$$\begin{array}{ccc} k_{i^+}(i) = k & \xrightarrow{\quad\text{id}\quad} & k_{i^+}(j) = k \\ \Big\downarrow{\scriptstyle \varphi_i(k)=0} & & \Big\downarrow{\scriptstyle \varphi_j(k)} \\ \mathcal{N}(i) & \longrightarrow & \mathcal{N}(j) \end{array}$$

Since φ is a natural transformation, the square must commute, forcing $\varphi_j(k) = 0$ for all $j \geq i$. If $j \notin i^+$, then φ_j is trivially 0, since $k_{i^+}(j) = 0$. Hence, $\varphi = 0$. ∎

In light of Proposition 4.6, we will use the terms projective and free interchangeably.

Definition 4.13 (Finitely presentable modules). *A persistence module $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ is* finitely presentable *if it is the cokernel of a map $\varphi : \mathcal{F} \to \mathcal{G}$, where $\mathcal{F}$ and $\mathcal{G}$ are free persistence modules of finite rank. Denote by $\boldsymbol{vec}_{fp}(\boldsymbol{P})$ the category of finitely presentable persistence modules over $\boldsymbol{P}$.*

Finitely presented modules trivially admit a projective resolution.

Definition 4.14 (Projective resolutions). *A* projective resolution *of a persistence module $\mathcal{M}$ is an exact sequence*

$$\cdots \to \mathcal{P}_n \to \mathcal{P}_{n-1} \to \dots \mathcal{P}_1 \to \mathcal{P}_0 \to \mathcal{M} \to 0$$

where each $\mathcal{P}_i$ is a projective module. A projective resolution $\mathcal{P}_\bullet \to \mathcal{M}$ is minimal *if the rank of $\mathcal{P}_k$ is minimal among all possible k^{th} terms of a projective resolution of $\mathcal{M}$ for each $k \in \mathbb{N}$. In other words, if $\mathcal{Q}_\bullet \to \mathcal{M}$ is another projective resolution of $\mathcal{M}$, the rank of $\mathcal{Q}_k$ is at most the rank of $\mathcal{P}_k$ for all $k \in \mathbb{N}$.*

Proposition 4.7 (Finitely presented modules admit projective resolutions). *Let $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ be a finitely presented. Then $\mathcal{M}$ admits a minimal projective resolution of finite length.*

Proof. We can construct a minimal projective resolution inductively. Begin with the exact sequence $0 \to \mathcal{P}_1 \xrightarrow{f} \mathcal{P}_0 \to \mathcal{M} \to 0$ obtained from the fact that $\mathcal{M}$ is finitely presented. If $\ker f = 0$, then we already have a minimal resolution. If not, we can extend the sequence to $0 \to \ker f \to \mathcal{P}_1 \xrightarrow{f} \mathcal{P}_0 \to \mathcal{M} \to 0$. We can only repeat this process finitely many times, since $\mathcal{P}_1$ is of finite rank and the ranks get smaller as we progress. The resultant resolution is minimal by construction. ■

Theorem 4.2 (Existence of projective resolutions). *Let $\mathcal{M} : \boldsymbol{P} \to \boldsymbol{vec}$ be a finitely presented. Then Hil$\mathcal{M}$ admits a minimal Hilbert decomposition.*

Proof. By Proposition 4.7, there exists a minimal projective resolution $\mathcal{P}_\bullet \to \mathcal{M}$ of finite length n. Let $i \in \mathbf{P}$ be arbitrary, then by the exactness of

$$\mathcal{P}_n \to \mathcal{P}_{n-1} \to \dots \mathcal{P}_1 \to \mathcal{P}_0 \to \mathcal{M} \to 0$$

and the rank nullity theorem we have

$$\dim \mathcal{M}(i) = \sum_{k=0}^{n} (-1)^k \dim \mathcal{P}_k.$$

The Hilbert decomposition is thus given by $(\bigoplus_{k\in 2\mathbb{N}} \mathcal{P}_k, \bigoplus_{k\in 2\mathbb{N}+1} \mathcal{P}_k)$. ■

Example 9 (Hilbert decomposition). Let $\mathcal{M} : G \to \mathbf{vec}$ be the indecomposable in Figure 4.9, where G is 5×5 grid.

k	k	0	0	0
k	k	k	0	0
k	k	k	0	0
k	k	k	k	k
k	k	k	k	k

Figure 4.9: Indecomposable "staircase" module indexed over a 5×5 grid. The generator is colored blue, the cogenerators are red, and relations between cogenerators are orange.

There is a single generator at $a = (0, 0)$, two cogenerators at $b = (1, 3)$ and $d = (2, 1)$, and a relation between cogenerators at $c = (2, 3)$. The minimal projective resolution is then

$$0 \to k_{c^+} \to k_{b^+} \oplus k_{d^+} \to k_{a^+} \to \mathcal{M}$$

and $(k_{a^+} \oplus k_{c^+}, k_{b^+} \oplus k_{d_+})$ is the Hilbert decomposition. The Betti signed barcode is $(\{a, c\}, \{b, d\})$.

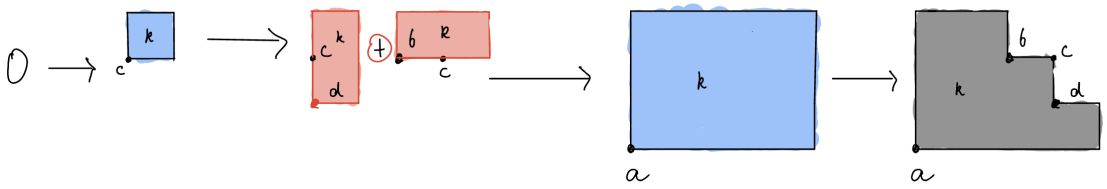


Figure 4.10: Minimal projective resolution of the module depicted in Figure 4.9.

4.6 Stability of rank and Hilbert decompositions

Just as in the 1-parameter case, convergence of SGD requires that the maps involved are locally Lipschitz. Theorem 2.2 guaranteed that 1-parameter persistence is Lipschitz with respect to the bottleneck distance. Theorem 2.2 directly related the bottleneck distance between two persistence diagrams to the infinity norm on the space of filtration functions, however there is often an intermediate step involved. One first proves that the map sending a filtration to its persistence module is stable and then proving that the map sending a persistence module to its diagram is stable. In order to do this, one needs to define a metric on the space of persistence modules. The interleaving distance [Cha+09] was introduced as a psuedometric for single parameter persistence modules $\mathcal{M} : \mathbb{R} \to \mathbf{vec}$, and it has been proven to be Lipschitz with respect to the infinity norm on filtration functions in a manner similar to Theorem 2.2. [Les15] extended the interleaving distance to modules of the form $\mathcal{M} : \mathbb{R}^n \to \mathbf{vec}$.

Definition 4.15 (Interleaving distance). *Let $\mathcal{M} : \mathbb{R}^n \to$ **vec** and $\epsilon \geq 0 \in \mathbb{R}^n$. The* ϵ-shift *is a persistence module $\mathcal{M}[\epsilon] : \mathbb{R}^n \to$ **vec**, with objects*

$$\mathcal{M}[\epsilon](r) = \mathcal{M}(r + \epsilon) \qquad\qquad r \in \mathbb{R}^n$$

and morphisms

$$(\mathcal{M}[\epsilon](r) \to \mathcal{M}[\epsilon](s)) = (\mathcal{M}(r + \epsilon) \to \mathcal{M}(s + \epsilon)) \qquad\qquad r \leq s \in \mathbb{R}^n.$$

The ϵ-shifts define an endofunctor on the category of $\mathbb{R}^n$-indexed modules. It sends an object $\mathcal{M}$ to $\mathcal{M}[\epsilon]$. A morphism $\eta : \mathcal{M} \to \mathcal{N}$ is mapped to $\eta[\epsilon] : \mathcal{M}[\epsilon] \to \mathcal{N}[\epsilon]$ that associates each $\eta_r : \mathcal{M}(r) \to \mathcal{N}(r)$ to $\eta[\epsilon]_r : \mathcal{M}(r + \epsilon) \to \mathcal{N}(r + \epsilon)$.

*We say that $\mathcal{M}, \mathcal{N} : \mathbb{R}^n \to$ **vec** are* ϵ-interleaved *if there exist morphisms $f : \mathcal{M} \to \mathcal{N}[\epsilon]$ and $g : \mathcal{N} \to \mathcal{M}[\epsilon]$ such that*

$$g[\epsilon] \circ f = \phi^{2\epsilon}_{\mathcal{M}} \qquad\qquad f[\epsilon] \circ g = \phi^{2\epsilon}_{\mathcal{N}}$$

where $\phi^{2\epsilon}_{\mathcal{M}} : \mathcal{M} \to \mathcal{M}[\epsilon]$ is the morphism that sends each $\mathcal{M}(a)$ to $\mathcal{M}(a + \epsilon)$.

The interleaving distance *between $\mathcal{M}$ and $\mathcal{N}$ is given by*

$$d_I(\mathcal{M}, \mathcal{N}) = \inf\{\epsilon \geq 0 \in \mathbb{R}^n | \mathcal{M} \text{ and } \mathcal{N} \text{ are } \epsilon \text{ interleaved } \}.$$

The following stability theorem was proven for 1-parameter persistence modules in [Cha+09] and for $\mathbb{R}^n$-indexed modules in theorem 5.3 of [Les15].

Theorem 4.3 (Stability of the interleaving distance). *Let X, Y be topological spaces and $f : X \to \mathbb{R}^n$, $g : X \to \mathbb{R}^n$ be functions. Let $\mathcal{M} : \mathbb{R}^n \to \boldsymbol{vec}$ be the persistence module obtained from obtaining homology in dimension i to the sublevel sets of f. Similarly, let $\mathcal{N} : \mathbb{R}^n \to \boldsymbol{vec}$ be the persistence module obtained from sublevel sets of g. Then,*

$$d_I(\mathcal{M}, \mathcal{N}) \leq \|f - g\|_\infty$$

In [OS21], the authors extend the bottleneck distance to a psuedometric on the space of Betti signed barcodes. They do this by interpreting classical persistence diagrams as point measures and the bottleneck distance as an optimal transport distance between point measures. Betti signed barcodes can be interpreted as signed measures and there are examples from measure theory such as the Kantorovich norm, which extend distances on positive measures to dissimilarities on signed measures. We refer to [OS21] for details and simply present the definition and the relevant stability result.

Definition 4.16 (Bottleneck dissimilarity). *Let $B = (\boldsymbol{barc}(\mathcal{P}), \boldsymbol{barc}(\mathcal{Q}))$ and $C = (\boldsymbol{barc}(\mathcal{R}), \boldsymbol{barc}(\mathcal{S}))$ be Betti signed barcodes, then their* bottleneck dissimilarity *is*

$$\widehat{d_B}(B, C) = d_b(\boldsymbol{barc}(\mathcal{P}) \cup \boldsymbol{barc}(\mathcal{S}), \boldsymbol{barc}(\mathcal{Q}) \cup \boldsymbol{barc}(\mathcal{R}))$$

where d_b is the standard bottleneck distance from Definition 2.8.

Theorem 4.4 (Stability of Hilbert decompositions). *Let $n \geq 2$ and $\mathcal{M}, \mathcal{N} : \mathbb{R}^n \to \boldsymbol{vec}$ be finitely presentable persistence modules. Let $B = (\boldsymbol{barc}(\mathcal{P}), \boldsymbol{barc}(\mathcal{Q}))$ and $C = (\boldsymbol{barc}(\mathcal{R}), \boldsymbol{barc}(\mathcal{S}))$ be the Betti signed barcodes of $\mathcal{M}$ and $\mathcal{N}$ respectively. Then,*

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

In particular, Theorem 4.3 and Theorem 4.4 prove that the map that assigns a filtration to its Hilbert decomposition is Lipschitz. The proof of Theorem 4.4 is based primarily on the fact that Hilbert decompositions arise from projective resolutions of persistence modules. An interleaving between two persistence modules can be used to establish a relationship between their projective resolutions and consequently their Hilbert decomposition. It turns out that rank decompositions of the rank invariant can also be obtained by projective resolutions, where the projective objects are defined relative to a certain exact structure. We refer to [BOO22] for details on rank exact sequences and the relative projective resolutions that arise. This interpretation of the rank decomposition, allows the results in [OS21] to be adapted to rank decompositions in [Bot+22]. In particular, theorem 6.1 of [Bot+22] states that signed barcodes of two finitely presented $\mathbb{R}^n$-indexed modules are Lipschitz with respect to bottleneck dissimilarity and the interleaving distance analogous to Theorem 4.4.

Chapter 5

Optimizing Functions Based on Signed Barcodes

We now come to the main objective of this thesis, the differentiability and optimization of functions based on signed barcodes. Just as in the one parameter case, we prove that the map that assigns to multifiltration its signed barcode is definable in an o-minimal structure. We begin by proving this for Betti signed barcodes arising from Hilbert decompositions, and then adapt that proof to signed barcodes arising from rank decompositions. The reason for doing so, is that the lifts involved in the first case are much more intuitive and easier to understand. The jump from Hilbert decompositions to rank decompositions is then not very big. In either case the first few steps of the pipeline, namely the ones that construct discretized multifiltrations and persistence modules are the same. Analogous to Chapter 3, we want to prove that the bottom row of the following diagram is definable on an o-minimal structure.

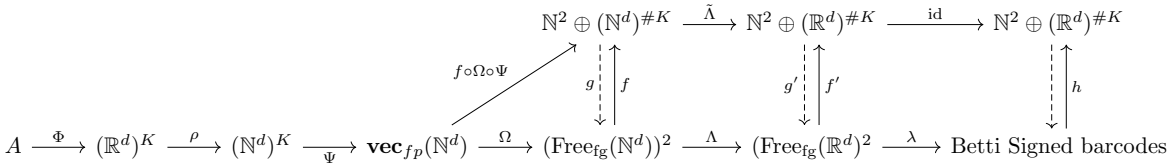


Figure 5.1: Decomposing the map that assigns to each multifiltered simplicial complex its Betti signed barcode.

The map Φ constructs from each element of A, a d-parameter filtration of a finite simplicial complex K. Next, the map ρ discretizes this filtration to construct a grid indexed over $\mathbb{N}^d$. Applying the simplicial homology functor in a given dimension over a fixed field, yields a finitely presented persistence module indexed over the poset $\mathbb{N}^d$. Note that the persistence module is finitely presented, since K and hence any multifiltration of K is finite. We then use Theorem 4.2, to obtain the minimal Hilbert decomposition of the discretized persistence module. Here $\mathrm{Free}_{\mathrm{fg}}((\mathbf{P})$ denotes finitely generated, free persistence modules indexed over $\mathbf{P}$. The map Λ unravels this discretization and delivers the minimal Hilbert decomposition of the original persistence module. Finally, we map each of the components in the Hilbert decomposition to their generators to obtain the Betti signed barcode.

5.1 Multifiltrations and the standard grid

5.1.1 Multiparameter filtrations

Definition 5.1 (Multifiltered simplicial complex). *Let K be a finite simplicial complex. A* d-filtration *of K is an increasing sequence $(K_r)_{r\in\mathbb{R}^d}$ of subcomplexes of K with respect to inclusion, i.e. $K_r \subset K_s$ for all $r \leq s$ and $K = \bigcup_{r\in\mathbb{R}^d} K_r$. A d-filtration is 1-*critical *if for each $\sigma \in K$, there is a unique minimal $\Phi_\sigma \in \mathbb{R}^d$ such that $\sigma \in K_r$. We refer to Φ_σ as the* birth index *of σ.*

Equivalently, one can view a 1-critical d-filtration of K as a vector Φ in the free vector space $(\mathbb{R}^d)^K$, where Φ_σ is the birth index of $\sigma \in K$. This makes use of the fact that the basis vectors of $(\mathbb{R}^d)^K$ are simply the simplices themselves. Note that if $\tau \subset \sigma$, then $\Phi_\tau \leq \Phi_\sigma$, so that the vector does indeed define a filtration.

Remark 5.1. A filtration Φ induces a *preorder* on the elements of K as follows: $\tau \leq \sigma$ if $\Phi_\tau \leq \Phi_\sigma$, where we use the product order on $\mathbb{R}^d$ to compare Φ_τ and Φ_σ.

Definition 5.2 (Parameterized family of filtrations). *Let K be a finite simplicial complex, and A a set. A map $\Phi : A \to (\mathbb{R}^d)^K$ is called a* parametrized family of (d)-filtrations *if for any $x \in A$ and $\sigma, \tau \in K$ with $\tau \subset \sigma$, we have $\Phi_\tau(x) \leq \Phi_\sigma(x)$.*

Remark 5.2. This induced preorder can be extended to a total order. Begin by fixing an indexing of the simplices of K in a manner consistent with face relations, i.e. if $\sigma_i \subset \sigma_j$, then $i \leq j$. Such an ordering of simplices of K exists, because K is finite. Note that each simplex can be uniquely identified by its index. Now, if τ and σ are distinct simplices, such that $\Phi_\sigma(x) = \Phi_\tau(x)$ or if they are incomparable with respect to Φ, use the total order induced by the indexing described above to order them. Henceforth, the relation $\leq$ on the simplices of K will denote the total order on K obtained by extending the preorder induced by a filtration Φ using the fixed indexing of simplices.

We will refer to the set of vectors in $(\mathbb{R}^d)^K$ that define a 1-critical d-filtration on K by Filt_K. Since K is finite, any filtration $\Phi \in (\mathbb{R}^d)^K$ can be discretized and represented as a vector in $(\mathbb{N}^d)^K$. This is due to the fact that the filtration is finite, and so the new simplices are added to the filtration only finitely many times. We define the discretization ρ on the basis vectors of $(\mathbb{R}^d)^K$ as follows:

$$\begin{aligned} \rho :\ & (\mathbb{R}^d)^K \to (\mathbb{N}^d)^K \\ & (v_\sigma)_i \mapsto \#\{\tau \in K : (v_\tau)_i \leq (v_\sigma)_i\} \qquad\qquad i \in \{1, \ldots, d\} \end{aligned}$$

Lemma 5.1. *Let $v \in (\mathbb{R}^d)K$ be a 1-critical, d-filtration of K and $\rho(v) \in (\mathbb{N}^d)K$ be its discretization. Then the discretized birth indices $\rho(v)_\sigma$ and $\rho(v)_\tau$ are distinct for all distinct $\sigma, \tau \in K$.*

Proof. The lemma hinges on the fact that the relation $\leq$ is a total order on K, obtained by extending the partial order. In fact, the statement does not hold if only used the partial order. Let $v \in (\mathbb{R}^d)^K$ be a 1-critical, d-filtration of K and $\tau, \sigma \in K$, such that $\rho(v_\sigma)_i = \rho(\tau)_i$ for all $i \in \{1, \ldots, d\}$. Then,

$$\#\{\nu \in K : (v_\nu)_i \leq (v_\sigma)_i\} = \#\{\nu \in K : (v_\nu)_i \leq (v_\tau)_i\} \text{ for all } i.$$

This means that $v_\sigma = v_\tau$ in the total order, which forces $\sigma = \tau$ since the total order extends the induced partial order in a manner that uniquely identifies each simplex. ∎

Implicitly, ρ projects the birth indices $(v_\sigma)_{\sigma\in K}$ onto each parameter $i \in \{1,\dots,d\}$ and reparameterizes each projected set by the integers $\{1,\dots,\#K\}$. The map ρ depends solely on the preorder induced by the filtration on the simplices of K, yielding the following result.

Proposition 5.1. *The map $\rho : Filt_K \to (\mathbb{N}^d)^K$ is semi-algebraic, and thus definable on an o-minimal structure. More precisely, there exists a linear partition of $Filt_k$ such that the restriction of ρ to each element of this partition is a constant map.*

Proof. We begin by defining a semi-algebraic partition of Filt_K in a manner similar to that of Proposition 3.2 of [Car+21]. Let O denote the set of all possible preorders on the simplices of K. Since K is finite, O is also finite. The set $F_{\geq} = \{\Phi \in \mathrm{Filt}_K : \Phi \text{ induces a preorder on } K \text{ equal to } \geq\}$, where $\geq \in O$ is a preorder on K, is defined by linear equalities and hence linear.

By construction ρ uses only information contained in the induced preorder, and so the restriction of ρ to each partition element is constant. A map, whose restriction is constant on each element of a semi-algebraic partition, is semi-algebraic. ∎

Visually, the fact that ρ is constant on each partition element, simply means that two filtrations that have the same preorder induce the same grid.

5.2 From multifiltrations to persistence modules

Just as with the standard barcodes in one-parameter persistence, computing Betti signed barcodes of multifiltrations can be done without explicitly going through the space of persistence modules. However, in order to prove definability, we work directly with the persistence modules.

$$\begin{aligned}\Psi : (\mathbb{N}^d)^K &\to \mathbf{vec_{fp}}(\mathbb{N}^d)\\ v &\mapsto \text{the } n^{\text{th}} \text{ persistence module } M \text{ of the multifiltration of } K \text{ given by } v\end{aligned}$$

Remark 5.3. Since Ψ acts on a filtration indexed by $\mathbb{N}^d$, the persistence modules we consider are defined on $\mathbb{N}^d$ rather than on $\mathbb{R}^d$. In particular, M is not the persistence module that corresponds to a filtration $\Phi \in (\mathbb{R}^d)^K$, but rather to its discretization $\rho(\Phi)$.

This map is Lipschitz with respect to the supremum norm on $(\mathbb{N}^d)^K$ and the interleaving distance on persistence modules, by the weak stability theorem [Cha+16].

5.3 The discretized minimal Hilbert decomposition

Given a persistence module $M : \mathbb{N}^d \to \mathbf{vec}$, Theorem 4.2 guarantees the existence of a minimal Hilbert decomposition (P, Q) of $\mathrm{Hil}(M)$. Define

$$\begin{aligned}\Omega : \mathbf{vec_{fp}}(\mathbb{N}^d) &\to (\mathrm{Free}_{\mathrm{fg}}(\mathbb{N}^d))^2\\ M &\mapsto (P, Q) \qquad \text{where } P \text{ and } Q \text{ form the minimal Hilbert decomposition.}\end{aligned}$$

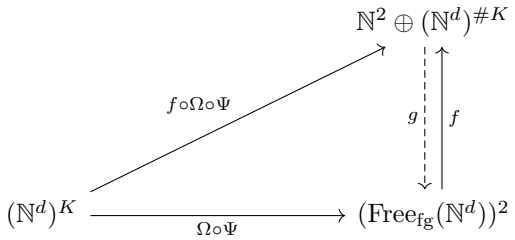
We prove that $\Omega \circ \Psi$ is semialgebraic by lifting this map to $(\mathbb{N}^d)^K \to \mathbb{N}^2 \oplus (\mathbb{N}^d)^K$ and showing that this map is semialgebraic.

Observe first that $\mathbf{barc}(P)$ can be represented as a vector in $(\mathbb{N}^d)^{\mathrm{rank}P}$ by using the lexicographic order on $\mathbb{N}^d$ to order the elements of $\mathbf{barc}(P)$ in non-decreasing order. An analogous representation exists for $\mathbf{barc}(Q)$ as well, leading to the map f, where $\mathbf{barc}(P)$ and $\mathbf{barc}(Q)$ denote the vectorization described above.

$$\begin{aligned}f : (\mathrm{Free}_{\mathrm{fg}}(\mathbb{N}^d))^2 &\to (\mathbb{N}^d)^K \to \mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K}\\ (P, Q) &\mapsto (\mathrm{rank}P, \mathrm{rank}Q, \mathbf{barc}(P), \mathbf{barc}(Q), 0)\end{aligned}$$

Remark 5.4. Since $\text{rank}P + \text{rank}Q \leq \#K$, we pack any remaining indices with zeros. We map into $\mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K}$ rather than into $\mathbb{N}^2 \oplus (\mathbb{N}^d)^{\text{rank}P+\text{rank}Q}$, since the latter would vary as the preorder induced by filtrations in $(\mathbb{N}^d)^K$ varies.

This map clearly has a left inverse g, since P and Q are completely characterized by their barcodes and the first two coordinates allow us to discern the two barcodes in the vector $(\text{rank}P, \text{rank}Q, \mathbf{barc}(P), \mathbf{barc}(Q), 0)$. This yields the following commutative diagram:



The left inverse of f makes $f \circ \Omega \circ \Psi$ a lift of $\Omega \circ \Psi$.

Theorem 5.1. *The map $f \circ \Omega \circ \Psi$ is semialgebraic. More precisely, there exists a semialgebraic partition of $(\mathbb{N}^d)^K$ such that $f \circ \Omega \circ \Psi$ restricted to each element of the partition is a constant map.*

Proof. Let $\text{filt}_K \subset (\mathbb{N}^d)^K$ denote the image $\rho(\text{Filt}_K)$. We can define a partition filt_K semi-algebraically in a manner similar to that of Proposition 5.1. Further, by Proposition 5.1 each partition element consists of a single vector. Consequently, Ψ is constant on each partition element of filt_K. Since Ω depends solely on the persistence module given by Ψ, which is constant across a given partition element, $\Omega \circ \Psi$ is also constant along each partition element. Post composing with f does not affect this constancy across the element. ■

5.4 Recovering the Betti signed barcode

In order to recover the minimal Hilbert decomposition of the persistence module that corresponds to our original filtration $v \in (\mathbb{R}^d)^K$, and hence its Betti signed barcode, we need to unravel the discretization in ρ to find the actual barcodes. This map will be affine on each stratum and hence definable.

Since free persistence modules are completely characterized by their barcodes, we may define a map $\Lambda : (\text{Free}_{\text{fg}}(\mathbb{N}^d)^2 \to (\text{Free}_{\text{fg}}(\mathbb{R}^d))^2$ by defining a map on each element of $(\mathbf{barc}(P), \mathbf{barc}(Q))$. To make proving the semi-algebraic nature of the lift of Λ clearer and to compute derivatives, we further decompose $\Lambda = \eta \circ \kappa$, where $\kappa : (\text{Free}_{\text{fg}}(\mathbb{N}^d))^2 \to (K^d)^{\text{rank}P} \oplus (K^d)^{\text{rank}Q}$. Let $i = (i_j)_{1 \leq j \leq d} \in \mathbf{barc}(P)$. Let $i = (i_j)_{1 \leq j \leq d} \in \mathbf{barc}(P)$, then κ sends i_j to $\sigma \in K$, such that $\rho(v_\sigma)_j = i_j$ and $\sigma \leq \tau$ for all $\tau \in K$ that fulfill $\rho(v_\tau)_j = i_j$. The map is well-defined, because the total order assigns each simplex a unique identifier and so the second minimality condition guarantees uniqueness.

The map $\eta : (K^d)^{\text{rank}P} \oplus (K^d)^{\text{rank}Q} \to \mathbb{N}^2 \oplus \text{Free}_{\text{fg}}(\mathbb{R}^d)$ is also defined component wise. We construct $P' \in \text{Free}_{\text{fg}}(\mathbb{R}^d)$ using the elements of $(K^d)^{\text{rank}P}$. Each element $(\sigma_1, \ldots, \sigma_d) \in K^d$ defines an element $(\Phi_{\sigma_1}(a))_1, (\Phi_{\sigma_2}(a))_2, \ldots, (\Phi_{\sigma_d}(a))_d \in \mathbb{R}^d$ of $\mathbf{barc}(P')$, where $(\Phi_\sigma(a)_j$ denotes the j^{th} component of $\Phi_\sigma(a) \in \mathbb{R}^d$. We define $Q' \in \text{Free}_{\text{fg}}(\mathbb{R}^d)$ analogously using elements of $(K^d)^{\text{rank}Q}$.

Remark 5.5. Clearly, $\text{rank}P' = \text{rank}P$ and $\text{rank}Q' = \text{rank}Q$. The minimality condition makes κ well-defined, however; the map $\eta \circ \kappa$ does not depend on this and one could in principle choose

any simplex that fulfills the first condition. Indeed, suppose there exist $\sigma, \sigma' \in K$ such that $\rho(\sigma)_j = i_j$ for some $i \in \mathbf{barc}(P)$ or $\mathbf{barc}(Q)$ and $1 \leq j \leq d$, then per construction of ρ, we must have $(\Phi_\sigma(a))_j = (\Phi'_\sigma(a))_j$. In other words, $\eta(\sigma') = \eta(\sigma)$.

The definition of Λ on the elements of the barcodes allows us to easily lift the map to $\tilde{\Lambda} = \tilde{\eta} \circ \tilde{\kappa} : \mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K} \to \mathbb{N}^2 \oplus (\mathbb{R}^d)^{\#K}$, where $\tilde{\kappa} : \mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K} \to \mathbb{N}^2 \oplus (K^d)^{\#K}$ and $\tilde{\eta} : \mathbb{N}^2 \oplus (K^d)^{\#K} \to \mathbb{N}^2 \oplus (\mathbb{R}^d)^{\#K}$. The map is the identity on the first two elements, i.e. $(\text{rank}P, \text{rank}Q) \mapsto (\text{rank}P, \text{rank}Q)$. It then maps each of the first $p+q$ vectors in $(\mathbb{N}^d)^{\#K}$ to $(\sigma_i \in K)_i \in K^d$, in the manner defined by κ and all of leftover zeros used in the vectorization to $\sigma_1 \in K$. Next, $\tilde{\eta}$ remains the identity on the first two elements and maps each of the first $p+q$ elements of $(K^d)^{\#K}$ to an element in $\mathbb{R}^d$ in the manner defined by η, and any remaining components arising from the image of zeros under $\tilde{\kappa}$ to zero. In particular, we have the following commutative diagram:

$$\begin{array}{ccc} \mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K} & \xrightarrow{\quad\tilde{\Lambda}\quad} & \mathbb{N}^2 \oplus (\mathbb{R}^d)^{\#K} \\ \Big\uparrow f \;\Big\downarrow g & & \Big\uparrow f' \;\Big\downarrow g' \\ (\text{Free}_{\text{fg}}(\mathbb{N}^d))^2 & \xrightarrow{\quad\Lambda\quad} & (\text{Free}_{\text{fg}}(\mathbb{R}^d))^2 \end{array}$$

The map f' is defined in the same manner as f and possesses a right inverse g' for the same reason as f.

Proposition 5.2. *The map $\tilde{\Lambda}$ is definable. More precisely, there exists a semialgebraic partition of $Im(f) \subset \mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K}$, such that the restriction of $\tilde{\kappa}$ to each partition element is a constant map and the restriction of $\tilde{\eta}$ to the image of each partition element under $\tilde{\kappa}$ is definable.*

Proof. The sets $O_{p,q} = \{f(P,Q) : \text{rank}P = p, \text{rank}Q = q\} \subset \mathbb{N}^2 \oplus (\mathbb{N}^d)^{\#K}$ are given by equalities and define a semialgebraic partition of $\text{Im}(f)$. Observe, that this partition is the pushforward of the partition defined in Proposition 5.1. On each partition element the map $\tilde{\kappa}$ is constant. It is trivially constant on the first two components. Consider $i \in \mathbb{N}^d$, which corresponds to an element of $\mathbf{barc}(P)$ or $\mathbf{barc}(Q)$, then each component of $\tilde{\kappa}(i)$ is completely determined by the map ρ which in turn is characterized by the preorder induced by $\Phi(a)$ on K. Hence, the restriction of $\tilde{\kappa}$ to each partition element is a constant map.

The map $\tilde{\eta}$ is constant on the fist two coordinates and assigns to each vector $v \in K^d$ a vector comprised of coordinates of the birth times of each simplex in v. Just as in the one parameter case, $\tilde{\eta}$ simply assigns to each simplex in v its birth index. Since Φ is definable, this map is definable.

Consequently, the restriction of $\tilde{\Lambda}$ to each component of the partition is definable. ∎

Finally, the map

$$\begin{aligned} \lambda : (\text{Free}_{\text{fg}}((\mathbb{R}^d))^2 &\to \text{Betti Signed barcodes} \\ (P,Q) &\mapsto (\mathbf{barc}(P), \mathbf{barc}(Q)) \end{aligned}$$

yields the Betti signed barcode. This map is clearly semialgebraic because it lifts to the identity on $\mathbb{N}^2 \oplus (\mathbb{R}^d)^{\#K}$. We have now proven that each map in the composition on the bottom row of Figure 5.1 is definable and thus obtain the following theorem.

Theorem 5.2. *Let Φ be a family of 1-critical, d-parameter filtrations that is definable on an o-minimal structure. Then, A admits a finite Whitney stratification, and $\lambda \circ \Lambda \circ \Omega \circ \Psi \circ \rho \circ \Phi$ is differentiable on each stratum.*

Proof. We have proven that each map involved in the composition are definable. Consequently, by Proposition 3.2 the composition is definable. The second statement follows from Theorem 3.2. ■

Finally, using the fact that $\lambda \circ \Lambda \circ \Omega \circ \Psi \circ \rho$ is Lipschitz (by Theorem 4.4) we can prove that SGD converges on loss functions based on Betti signed barcodes

Theorem 5.3 (SGD converges on maps based on Betti signed barcodes). *Let $A \subset \mathbb{R}^N$ be definable and K be a finite simplicial complex. Suppose that $\Phi : A \to (\mathbb{R}^d)^K, d \geq 2$ is a* locally Lipschitz *family of filtrations that is* definable on an o-minimal structure *and that $\mathcal{L} : \mathbb{N}^2 \oplus (\mathbb{R}^d)^{\#K}$ is a locally Lipschitz loss function that is* definable on an o-minimal structure. *Assume further, that the sequences of step sizes $\{\alpha_k\}$ and noise $\{\xi_k\}$ satisfy condition C of [Dav+20]. Then, the iterates $\{x_k\}$ produced by SGD converge almost surely to a critical point of $\mathcal{L} \circ \lambda \circ \Lambda \circ \Omega \circ \Psi \circ \rho \circ \Phi$ and the corresponding function values converge.*

5.5 Differentiability of signed barcodes

One can proceed similarly to prove the differentiability of signed barcodes arising from minimal rank decompositions of the rank invariant.

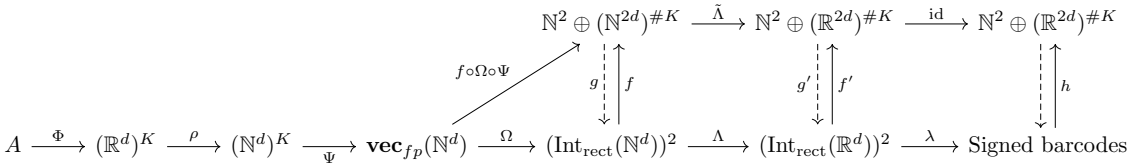


Figure 5.2: Decomposing the map that assigns to each multifiltered simplicial complex its signed barcode.

The maps Φ, ρ, and Ψ are the same as those in Figure 5.1. Instead of computing the Hilbert decomposition, Ω computes the minimal rank decomposition of a persistence module. As before, we then unravel the discretization and construct the signed barcodes.

5.5.1 The discretized minimal rank decomposition

We then apply Corollary 5.2 of [BOO22] to obtain a minimal rank decomposition $(\mathcal{R}, \mathcal{S})$ of the usual rank invariant of $M \in \mathbf{vec}_{fp}(\mathbb{N}^d)$, where $\mathcal{R}$ and $\mathcal{S}$ are multisets of rectangles contained in $\prod_{i=1}^{d}[1, \#K]$. The map Ω constructs the indicator module supported on the elements of $\mathcal{R}$ and $\mathcal{S}$.

$$\begin{aligned} \Omega : \mathbf{vec}_{fp}(\mathbb{N}^d) &\to (\mathrm{Int}_{\mathrm{rect}}(\mathbb{N}^d))^2 \\ M &\mapsto (P, Q) \\ P = \bigoplus_{I \in \mathcal{R}} k_I \text{ and } Q &= \bigoplus_{I \in \mathcal{S}} k_I \qquad \text{(with appropriate multiplicities)} \end{aligned}$$

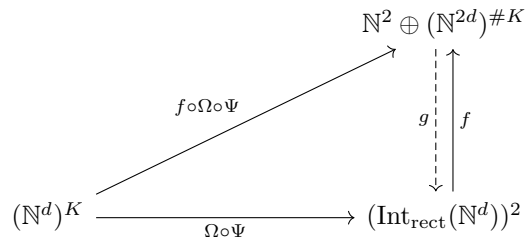
As discussed in section 6.1 of [BOO22], one can represent the minimal rank decomposition $(\mathcal{R}, \mathcal{S})$ by associating each rectangle in $\mathcal{R}$ with its positive slope diagonal and with positive sign, while rectangles in $\mathcal{S}$ are represented by their diagonals with positive slope and with a negative sign. In particular, a rectangle $\langle s, t\rangle \in \mathcal{R} \subset \mathbb{N}^d$ can be completely characterized by the pair $(s, t) \in (\mathbb{N}^d)^2$. Define

$$\mathbf{barc}(\mathcal{R}) = \{(s, t) \in (\mathbb{N}^d)^2 : \langle s, t\rangle \in \mathcal{R}\} \text{ with multiplicity}$$
$$\mathbf{barc}(\mathcal{S}) = \{(s, t) \in (\mathbb{N}^d)^2 : \langle s, t\rangle \in \mathcal{S}\} \text{ with multiplicity.}$$

Observe that there is a lexicographic order on the multisets $\mathbf{barc}(\mathcal{R})$ and $\mathbf{barc}(\mathcal{S})$, allowing for a natural embedding into $(\mathbb{N}^{2d})^{\#\mathcal{R}}$ and $(\mathbb{N}^{2d})^{\#\mathcal{S}}$. This allows us to adapt the lifting map f from earlier to this setting.

$$f : (\mathrm{Int}_{\mathrm{rect}}(\mathbb{N}^d))^2 \to \mathbb{N}^2 \oplus (\mathbb{N}^{2d})^{\#K}$$
$$M \mapsto \Big(\sum_{\langle s,t\rangle \in \mathcal{R}} \mathrm{Rk}k_{\langle s,t\rangle}(s, t), \sum_{\langle s,t\rangle \in \mathcal{S}} \mathrm{Rk}k_{\langle s,t\rangle}(s, t), \mathbf{barc}(\mathcal{R}), \mathbf{barc}(\mathcal{S}), 0\Big)$$

The expression $\sum_{\langle s,t\rangle \in \mathcal{R}} \mathrm{Rk}k_{\langle s,t\rangle}(s, t)$ is just $\#\mathcal{R}$ and similarly for $\#S$. We map into $\mathbb{N}^2 \oplus (\mathbb{N}^{2d})^{\#K}$ and pack any remaining indices with zeros for the same reason discussed in Remark 5.4. Clearly, this map has a left inverse g leading to the following commutative diagram:



The left inverse of f makes $f \circ \Omega \circ \Psi$ a lift of $\Omega \circ \Psi$. The following theorem is almost identical to Theorem 5.1.

Theorem 5.4. *The map $f \circ \Omega \circ \Psi$ is semialgebraic. More precisely, there exists a semialgebraic partition of $(\mathbb{N}^d)^K$ such that $f \circ \Omega \circ \Psi$ restricted to each element of the partition is a constant map.*

Proof. Since the map Ψ remains unchanged, its restriction is constant over each partition element defined in Theorem 5.1. The minimal rank decomposition depends solely on the persistence module obtained by Ψ. Indeed, Corollary 4.2 provides an explicit Möbius inversion formula to compute the minimal rank decomposition in this case. This formula is simply an alternating sum of the usual rank invariant of the underlying persistence module, which is invariant over each partition element. Finally, postcomposing with the lift f preserves this property. ■

5.5.2 Unraveling the discretization

The signed barcodes obtained in the previous section are the ones associated with the discretization of the original filtration. Just as with the Hilbert decomposition, we define a map whose restriction is affine on each partition element, to unravel the discretization. We define a map $\Lambda : (\mathrm{Int}_{\mathrm{rect}}(\mathbb{N}^d))^2 \to (\mathrm{Int}_{\mathrm{rect}}\mathbb{R}^d))^2$. Keeping with our earlier construction, Λ is defined on the elements of $\mathbf{barc}(\mathcal{R}), \mathbf{barc}(\mathcal{S})$, and decomposes into $\Lambda : \eta \circ \kappa$. The map

$$\kappa : (\mathrm{Int}_{\mathrm{rect}}(\mathbb{N}^d))^2 \to (K^{2d})^{\#\mathcal{R}} \oplus (K^{2d})^{\#\mathcal{S}}$$

is defined in the same way as earlier, with each element $(s,t) \in \mathbf{barc}(\mathcal{R})$ being assigned a pair $((\sigma_j)_{1\leq j\leq d}, (\tau_j)_{1\leq j\leq d})$, with the simplices being assigned using the left inverse of ρ (and the total order if needed) as earlier. Next define the map

$$\eta : (K^{2d})^{\#\mathcal{R}} \oplus (K^{2d})^{\#\mathcal{S}} \to (\mathrm{Int}_{\mathrm{rect}}(\mathbb{R}^d))^2$$

by using the elements of $(K^{2d})^{\#\mathcal{R}}$ to construct the barcode for the positive module in the decomposition. Each pair $((\sigma_j)_{1\leq j\leq d}, (\tau_j)_{1\leq j\leq d}) \in K^{2d}$ is sent to $(s,t) = ((\Phi_{\sigma_j})_j, (\Phi_P \tau_j))$ defining a rectangle in $\mathbb{R}^d$. Setting P' to be the indicator module supported on these rectangles and S' to be the module supported on the rectangles arising from the birth indices of $(K^2 d)^{\#\mathcal{S}}$ completes the definition of Λ. Once again, Λ can be lifted to a map

$$\tilde{\Lambda} = \tilde{\eta} \circ \tilde{\kappa} : \mathbb{N}^2 \oplus (\mathbb{N}^{2d})^{\#K} \xrightarrow{\tilde{\kappa}} \mathbb{N}^2 \oplus (K^{2d})^{\#K} \xrightarrow{\tilde{\eta}} \mathbb{N}^2 \oplus (\mathbb{R}^{2d})^{\#K}$$

for the same reason as earlier. This map is the identity on the first two components. Using reasoning identical to that used in Proposition 5.2, we can prove the identical statement for this map.

Proposition 5.3. *The map $\tilde{\Lambda}$ is definable. More precisely, there exists a semialgebraic partition of $Im(f) \subset \mathbb{N}^2 \oplus (\mathbb{N}^{2d})^{\#K}$, such that the restriction of $\tilde{\kappa}$ to each partition element is a constant map and the restriction of $\tilde{\eta}$ to the image of each partition element under $\tilde{\kappa}$ is definable.*

Consequently, we have:

Theorem 5.5. *Let Φ be a family of 1-critical, d-parameter filtrations that is definable on an o-minimal structure. Then, A admits a finite Whitney stratification, and $\lambda \circ \Lambda \circ \Omega \circ \Psi \circ \rho \circ \Phi$ is differentiable on each stratum.*

As with Hilbert decompositions, $\lambda \circ \Lambda \circ \Omega \circ \Psi \circ \rho$ is Lipschitz, and so the results above together deliver the following theorem.

Theorem 5.6 (SGD converges on maps based on signed barcodes). *Let $A \subset \mathbb{R}^N$ be definable and K be a finite simplicial complex. Suppose that $\Phi : A \to (\mathbb{R}^d)^K, d \geq 2$ is a* locally Lipschitz *family of filtrations that is* definable on an o-minimal structure *and that $\mathcal{L} : \mathbb{N}^2 \oplus (\mathbb{R}^{2d})^{\#K}$ is a locally Lipschitz loss function that is* definable on an o-minimal structure. *Assume further, that the sequences of step sizes $\{\alpha_k\}$ and noise $\{\xi_k\}$ satisfy condition C of [Dav+20]. Then, the iterates $\{x_k\}$ produced by SGD converge almost surely to a critical point of $\mathcal{L} \circ \lambda \circ \Lambda \circ \Omega \circ \Psi \circ \rho \circ \Phi$ and the corresponding function values converge.*

5.6 Conclusion and ongoing work

Theorem 5.3 and Theorem 5.6 tell us that SGD converges definable, locally Lipschitz functions defined on Betti signed barcodes and signed barcodes. In ongoing work we aim to provide a few examples for such loss functions and demonstrate the power of both methods using a few toy experiments. Given the fact that multiparameter persistence carries far more information than single parameter persistence, we hope that optimizing functions using signed barcodes instead of standard persistence diagrams will deliver significant improvements in performance. The main challenges at the moment are to make the computation of signed barcodes more efficient as well as to design interesting loss functions.

Bibliography

[Ada+17] Henry Adams et al. "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *J. Mach. Learn. Res.* 18.1 (2017), 218–252. ISSN: 1532-4435.

[All10] R B J T Allenby. *How to Count : An Introduction to Combinatorics*. 2nd ed. 2010. ISBN: 9780429113123.

[Bau21] Ulrich Bauer. "Ripser: efficient computation of Vietoris–Rips persistence barcodes". en. In: *Journal of Applied and Computational Topology* 5.3 (Sept. 2021), pp. 391–423. ISSN: 2367-1726, 2367-1734. DOI: 10.1007/s41468-021-00071-5. URL: https://link.springer.com/10.1007/s41468-021-00071-5 (visited on 07/04/2023).

[BBH23] Benjamin Blanchette, Thomas Brüstle, and Eric J. Hanson. *Exact structures for persistence modules*. arXiv:2308.01790 [cs, math]. Aug. 2023. URL: http://arxiv.org/abs/2308.01790 (visited on 08/07/2023).

[BL22] Magnus Bakke Botnan and Michael Lesnick. "An Introduction to Multiparameter Persistence". In: (2022). DOI: 10.48550/ARXIV.2203.14289. URL: https://arxiv.org/abs/2203.14289.

[BOO22] Magnus Bakke Botnan, Steffen Oppermann, and Steve Oudot. "Signed Barcodes for Multi-Parameter Persistence via Rank Decompositions". In: *38th International Symposium on Computational Geometry (SoCG 2022)*. Ed. by Xavier Goaoc and Michael Kerber. Vol. 224. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 19:1–19:18. ISBN: 978-3-95977-227-3. DOI: 10.4230/LIPIcs.SoCG.2022.19. URL: https://drops.dagstuhl.de/opus/volltexte/2022/16027.

[Bot+22] Magnus Bakke Botnan et al. *On the bottleneck stability of rank decompositions of multi-parameter persistence modules*. July 30, 2022. arXiv: 2208.00300[cs,math]. URL: http://arxiv.org/abs/2208.00300 (visited on 08/15/2023).

[Car09] Gunnar Carlsson. "Topology and data". In: *Bulletin of the American Mathematical Society* 46.2 (Jan. 29, 2009), pp. 255–308. ISSN: 0273-0979. DOI: 10.1090/S0273-0979-09-01249-X. URL: http://www.ams.org/journal-getitem?pii=S0273-0979-09-01249-X (visited on 08/25/2023).

[CVJ21] Gunnar Carlsson and Mikael Vejdemo-Johansson. *Topological Data Analysis with Applications*. Cambridge University Press, 2021. DOI: 10.1017/9781108975704.

[CZ09] Gunnar Carlsson and Afra Zomorodian. "The Theory of Multidimensional Persistence". In: *Discrete & Computational Geometry* 42.1 (July 2009), pp. 71–93. ISSN: 0179-5376, 1432-0444. DOI: 10.1007/s00454-009-9176-0. URL: http://link.springer.com/10.1007/s00454-009-9176-0.

[Car+21] Mathieu Carriere et al. “Optimizing persistent homology based functions”. In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, 2021, pp. 1294–1303. URL: https://proceedings.mlr.press/v139/carriere21a.html.

[Cha+09] Frédéric Chazal et al. “Proximity of Persistence Modules and Their Diagrams”. In: *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*. SCG ’09. Aarhus, Denmark: Association for Computing Machinery, 2009, 237–246. ISBN: 9781605585017. DOI: 10.1145/1542362.1542407. URL: https://doi.org/10.1145/1542362.1542407.

[Cha+16] Frédéric Chazal et al. *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics. Cham: Springer International Publishing, 2016. ISBN: 9783319425436 9783319425450. DOI: 10.1007/978-3-319-42545-0. URL: http://link.springer.com/10.1007/978-3-319-42545-0.

[Dav+20] Damek Davis et al. “Stochastic Subgradient Method Converges on Tame Functions”. In: *Foundations of Computational Mathematics* 20.1 (Feb. 2020), pp. 119–154. ISSN: 1615-3375, 1615-3383. DOI: 10.1007/s10208-018-09409-5. URL: http://link.springer.com/10.1007/s10208-018-09409-5.

[Dem+22] Andaç Demir et al. “ToDD: Topological Compound Fingerprinting in Computer-Aided Drug Discovery”. In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo et al. Vol. 35. Curran Associates, Inc., 2022, pp. 27978–27993. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/b31f6d65f2584b3c4347148db36fe07f-Paper-Conference.pdf.

[DCL22] Zhetong Dong, Jinhao Chen, and Hongwei Lin. “Topology-controllable Implicit Surface Reconstruction Based on Persistent Homology”. In: *Computer-Aided Design* 150 (2022), p. 103308. ISSN: 0010-4485. DOI: https://doi.org/10.1016/j.cad.2022.103308. URL: https://www.sciencedirect.com/science/article/pii/S001044852200077X.

[Dri98] L. P. D. Van Den Dries. *Tame Topology and O-minimal Structures*. 1st ed. Cambridge University Press, May 1998. ISBN: 9780521598385 9780511525919. DOI: 10.1017/CBO9780511525919. URL: https://www.cambridge.org/core/product/identifier/9780511525919/type/book (visited on 07/03/2023).

[EM13] Herbert Edelsbrunner and Dmitriy Morozov. “Persistent Homology: Theory and Practice”. In: *European Congress of Mathematics Kraków, 2 – 7 July, 2012*. Ed. by Rafał Latała et al. Zuerich, Switzerland: European Mathematical Society Publishing House, Nov. 30, 2013, pp. 31–50. ISBN: 9783037191200. DOI: 10.4171/120-1/3. URL: https://ems.press/doi/10.4171/120-1/3 (visited on 08/25/2023).

[EB] Gudhi Editorial Board. *Gudhi*. Version 3.7.1. URL: https://gudhi.inria.fr.

[GHO16] Marcio Gameiro, Yasuaki Hiraoka, and Ippei Obayashi. “Continuation of point clouds via persistence diagrams”. en. In: *Physica D: Nonlinear Phenomena* 334 (Nov. 2016), pp. 118–132. ISSN: 01672789. DOI: 10.1016/j.physd.2015.11.011. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167278915002626 (visited on 06/07/2023).

[Hat02] Allen Hatcher. *Algebraic topology*. Cambridge ; New York: Cambridge University Press, 2002. 544 pp. ISBN: 9780521791601 9780521795401.

[Hir+16] Yasuaki Hiraoka et al. "Hierarchical structures of amorphous solids characterized by persistent homology". In: *Proceedings of the National Academy of Sciences* 113.26 (June 28, 2016), pp. 7035–7040. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1520877113. URL: https://pnas.org/doi/full/10.1073/pnas.1520877113 (visited on 07/03/2023).

[KB17] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. arXiv:1412.6980 [cs]. Jan. 2017. URL: http://arxiv.org/abs/1412.6980 (visited on 08/09/2023).

[Les15] Michael Lesnick. "The Theory of the Interleaving Distance on Multidimensional Persistence Modules". In: *Foundations of Computational Mathematics* 15.3 (June 2015), pp. 613–650. ISSN: 1615-3375, 1615-3383. DOI: 10.1007/s10208-015-9255-y. URL: http://link.springer.com/10.1007/s10208-015-9255-y (visited on 08/16/2023).

[LOT22] Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. "A Framework for Differential Calculus on Persistence Barcodes". In: *Foundations of Computational Mathematics* 22.4 (Aug. 2022), pp. 1069–1131. ISSN: 1615-3375, 1615-3383. DOI: 10.1007/s10208-021-09522-y. URL: https://link.springer.com/10.1007/s10208-021-09522-y (visited on 08/25/2023).

[ML78] Saunders Mac Lane. *Categories for the Working Mathematician*. Vol. 5. Graduate Texts in Mathematics. New York, NY: Springer New York, 1978. ISBN: 9781441931238 9781475747218. DOI: 10.1007/978-1-4757-4721-8. URL: http://link.springer.com/10.1007/978-1-4757-4721-8 (visited on 07/02/2023).

[Moo+20] Michael Moor et al. "Topological Autoencoders". In: *Proceedings of the 37th International Conference on Machine Learning*. ICML'20. JMLR.org, 2020.

[OS21] Steve Oudot and Luis Scoccola. "On the Stability of Multigraded Betti Numbers and Hilbert Functions". In: (2021). DOI: 10.48550/ARXIV.2112.11901. URL: https://arxiv.org/abs/2112.11901.

[Oud15] Steve Y. Oudot. *Persistence theory: from quiver representations to data analysis*. eng. Mathematical Surveys and Monographs volume 209. Providence, Rhode Island: American Mathematical Society, 2015. ISBN: 9781470434434 9781470425456.

[PSO18a] Adrien Poulenard, Primoz Skraba, and Maks Ovsjanikov. "Topological Function Optimization for Continuous Shape Matching". en. In: *Computer Graphics Forum* 37.5 (Aug. 2018), pp. 13–25. ISSN: 01677055. DOI: 10.1111/cgf.13487. URL: https://onlinelibrary.wiley.com/doi/10.1111/cgf.13487 (visited on 06/07/2023).

[PSO18b] Adrien Poulenard, Primoz Skraba, and Maks Ovsjanikov. "Topological function optimization for continuous shape matching". In: *Computer Graphics Forum*. Vol. 37. 5. Wiley Online Library. 2018, pp. 13–25.

[San+23] Ainkaran Santhirasekaram et al. "Topology Preserving Compositionality for Robust Medical Image Segmentation". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. 2023, pp. 543–552.

[Set] Siddharth Setlur. "Examining Persistent Homology from Three Different Theoretical Perspectives." Bachelor Thesis. ETH Zürich. URL: https://n.ethz.ch/~ssetlur/download/bachelorarbeit.pdf.

[Tos+10] Fukui Toshizumi et al. *The Japanese-Australian Workshop on Real and Complex Singularities: JARCS III*. OCLC: 956674174. Australia: Centre for Mathematics and its Applications Mathematical Sciences Institute The Australian National University, 2010. ISBN: 9780731552078. URL: https://maths.anu.edu.au/research/cma-proceedings/japanese-australian-workshop-real-complex-singularities.

[Wya05] M. Wyart. "On the rigidity of amorphous solids". In: *Annales de Physique* 30.3 (2005), pp. 1–96. ISSN: 0003-4169, 1286-4838. DOI: 10.1051/anphys:2006003. URL: http://www.annphys.org/10.1051/anphys:2006003 (visited on 09/01/2023).

[ZC05] Afra Zomorodian and Gunnar Carlsson. "Computing Persistent Homology". In: *Discrete & Computational Geometry* 33.2 (Feb. 2005), pp. 249–274. ISSN: 0179-5376, 1432-0444. DOI: 10.1007/s00454-004-1146-y. URL: http://link.springer.com/10.1007/s00454-004-1146-y (visited on 01/02/2023).