Jacob Leygonie

Differential & Fiber of Persistent Homology

Mathematical Institute

University of Oxford

À Claire, Nana et Dede, et Claudia

Contents

Introduction

This thesis is concerned with the following two naive questions about the widely used *Persistent Homology*[1] (PH) descriptor for data analysis:

(**dPH**) Can we do Machine Learning with Persistent Homology?

(**PH^{-1}**) What are the objects giving rise to the same Persistent Homology?

These interrogations stem from the numerous applications of PH to data science problems[2] where geometry and topology are prominent, yet they are themselves of theoretical nature. Namely when it comes to solving such data science problems:

(**dPH**) Machine Learning and Optimisation pipelines are the modern favorite candidates. Since gradient descent is the unavoidable paradigm of large-scale optimisation, including PH in these pipelines requires a notion of **differentiability** for PH; and

(**PH^{-1}**) Deciding *a priori* whether PH is a relevant descriptor depends on its discriminative power, i.e. if it is able to distinguish between the important groups that constitute the data science problem at hand. This calls for an analysis of the **fiber** (pre-image) of PH.

It came as a pleasant surprise that the processes of addressing these two relatively independent problems cross-fertilized in multiple ways: the structure developed to describe PH, e.g. the stratifications of its domain and co-domain, have been fundamental in the course of tackling both (**dPH**) and (**PH^{-1}**).

This thesis is organised around six chapters that compile the six papers written during the last three years, the first three of which investigate (**dPH**) while the remaining ones are devoted to (**PH^{-1}**). Next we give an overview of these contributions.

[1] Persistent Homology [EH08, ZC05] is a central tool in *Topological Data Analysis* (TDA) whose construction originates from Algebraic Topology and provides a stable, computable and topologically meaningful invariant for geometric datasets such as graphs and point clouds. In its classical form Persistent Homology encodes the topological variations of the sublevel-sets $f^{-1}((-\infty;t])$ of a function $f : X \to \mathbb{R}$ on a topological space X (such as a graph or a point cloud).

[2] Fields of applications include Neuroscience [DBF14, BMM+16, GGB16, LEF+16, YKAY16, Cur17], Cancer Diagnosis [DCCW+10, NLC11, SCM+14], Cellular Differentiation [RCK+17], DNA Folding [ESR16], Proteins [XW14, GHI+15, KNBNH16, XLM18], Phylogenetics [CLR16], Contagion Spreading [CCR13, TKH+15, LP18], Material Science [KGKM13, KGKM14, HNH+16], Robotics [VAB13, BGK15, PHR16], Sensor Networks [DSG07], Finance [Gid17, LPRS08], Collaboration and Social Networks [CH13, BDP+15, PMRS17], Computer Vision [CIDSZ08, CBK09, BGKP13], Natural Language Processing [Zhu13], Time Series and Dynamical Systems [PDHH15, MZR16].

Chapter 1: Differential Calculus on Persistence Barcodes

In this chapter we adopt the point of view on Persistent Homology, which will carry over the entire thesis, as a map PH from the space of real-valued functions over a fixed topological space X (often taken to be a simplicial complex or a smooth manifold in applications) to the space **Bar** of *barcodes*.[3]

The space **Bar** is infinite-dimensional and has singularities, yet inspired by Diffeology Theory[4] we manage to make sense of the differentiability and derivatives for maps from and to **Bar**. These notions are convenient enough that we have techniques such as the chain rule to check whether maps are differentiable and compute their derivatives in concrete scenarios.

In particular we show that PH is differentiable on a generic (open and dense) subset of its domain when the latter consists of functions on a simplicial complex $X = K$ or Morse functions on a smooth manifold $X = M$.

Moreover we can study the singularities of PH by means of a stratification[5] of its domain: on each lower-dimensional stratum PH may not be differentiable but it admits directional derivatives.

We also consider the commonly used maps out of barcodes, namely the *vectorizations* and the metrics on **Bar**, and show that they are generically differentiable as well.

With the understanding of the differentiability properties of these classes of maps in and out **Bar**, our framework legitimates the existing, emerging and incoming applications that apply gradient descent to optimise objective functions involving Persistent Homology.[6]

[3] The *barcode* or *persistence diagram* $\mathrm{PH}(f)$ of a function f is a multi-set of points (b, d) called *intervals* that encode the scales at which the topology of the sublevel-sets $f^{-1}((-\infty; t])$ change in a given dimension. For instance in dimension 0 the interval (b, d) indicates the appeareance of a novel connected component in $f^{-1}((-\infty; b])$ that is further merged with an older component in $f^{-1}((-\infty; d])$. In dimension 1 we have the analog interpretation with loops, in dimension 2 with voids and similarly in higher dimensions. These notions are properly formalised e.g. in section 1.2.

[4] Patrick Iglesias-Zemmour. *Diffeology*. Mathematical Surveys and Monographs, volume 185. American Mathematical Society, Providence, RI, 2013

[5] Roughly speaking, a stratification of a space is a partition by submanifolds of varying dimensions that glue together nicely.

[6] Examples of applications are point cloud inference [GHO16], shape matching [PSO18], classifiers' regularization [CNBW19], image segmentation [HLSC19], graph classification [HGR+20a, YL21], surface reconstruction [BGGSSG20a] and generative modelling [GNDS20a].

Contributions, Funding and Acknowledgements

This is joint work with Ulrike Tillmann and Steve Oudot. There is a public version and a version published at *Foundations of Computational Mathematics*. Besides there is a recorded online talk given on this subject at the ATMCS conference. The authors are indebted to the anonymous reviewers for their valuable insights in the final revisions of the manuscript. The authors thank Vidit Nanda and Oliver Vipond for useful conversations. The author of this manuscript also thanks Heather Harrington for general guidance and Yixuan Wang for sharing knowledge on Morse theory and differential geometry. His research has been supported by ESPRC grant EP/R018472/1.

Chapter 2: Optimisation of Spectral Wavelets for Persistence-based Graph Classification

The first chapter guaranteed the (generic) differentiability of PH and enabled the computation of its derivatives, thereby answering (**dPH**) by the affirmative: it is theoretically possible to include Persistent Homology in Machine Learning and Optimisation pipelines.

Next we consider an actual computational application of (**dPH**): the implementation of a Machine Learning methodology based on PH and the Laplacian operator for classifying data sets of graphs.

Instead of ordinary Persistent Homology we consider the *Extended* Persistent Homology (EPH), which is more expressive[7] and more convenient to deal with computationally.[8] In turn we adapt the necessary results of the previous chapter to this context.

The resulting optimisation pipeline can be applied to arbitrary data sets of graphs, for example we demonstrate its performance on datasets of molecules, proteins and movie reviews.

The implementation makes use of the Gudhi and PyTorch libraries for TDA and Deep Learning respectively, and is available in a public GitHub repository.

[7] Informally speaking Extended Persistent Homology $\mathrm{EPH}(f)$ records the additional information of the topological variations in the super-level sets $f^{-1}([t;+\infty))$ of the function $f : X \to \mathbb{R}$.

[8] In the ordinary barcode $\mathrm{PH}(f)$ we may have unbounded intervals $(b,+\infty)$ that are not easy to handle, and those are absent in $\mathrm{EPH}(f)$.

Online resources, Contributions and Acknowledgements

This is joint work with Ka Ma Yim. There is an open-access version of this work published at *Frontiers of Artificial Intelligence* and a public repository containing the associated code base.

The authors thank Ulrike Tillmann and Heather Harrington for their close guidance and thoughtful advice on this project. In addition, the authors would like to thank Vidit Nanda, Peter Grindrod CBE, Steve Oudot, Mathieu Carrière, and Theo Lacombe for fruitful discussions. Finally they are indebted to the reviewers for their thoughtful and constructive comments. The author of this manuscript has been funded by the EPSRC grant EP/R513295/1.

Chapter 3: A Gradient Sampling Algorithm for Stratified Maps with Applications to Topological Data Analysis

In this final investigation of (**dPH**) we face the following situation: the map PH is generically differentiable by the 1st chapter, hence it is possible to optimise it using standard gradient descent (GD) as done in the 2nd chapter, however the set where PH is not differentiable hinders the convergence guarantees of GD.

From a theoretical standpoint the situation is not *as bad* as the general context of non-smooth optimisation, that is minimising locally Lipschitz objective functions, which by Rademacher Theorem are differentiable on a full-measure subset. Here the irregularities of Persistent Homology are strongly organised: there is a stratification of the domain of PH such that on lower-dimensional strata PH is not differentiable but admits as many directional derivatives as there are incident top-dimensional strata.

Therefore PH is part of the class of *stratifiably smooth functions*, for which we design an algorithm specifically tailored to put strata at work. This methodology takes great inspiration from the *Gradient Sampling* (GS) algorithm[9] for general locally Lipschitz objective functions, hence we refer to it as the *Stratified Gradient Sampling* (SGS) approach.

[9] James V Burke, Adrian S Lewis, and Michael L Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM Journal on Optimization*, 15(3):751–779, 2005

The SGS benefits from asymptotic convergence towards (generalised) critical points and a very useful stopping criterion on the norm of the (generalised) gradient which is absent from GD. Unlike both GD and GS, there is a (super-linear) convergence rate for SGS.

We apply the SGS methodology to objective functions based on Persistent Homology: for this we provide an efficient method to explore the stratification associated to PH. Through experiments we provide empirical evidence that SGS outperforms the other approaches when minimising stratifiably smooth functions.

Online resources, Contributions and Acknowledgements

This is joint work with Mathieu Carrière, Théo Lacombe and Steve Oudot. There is a public version, a version published at *Mathematical Programming*, and a public repository containing the associated codebase. The authors are indebted to Charles Arnal for his thorough proofreading and in-depth feedback.

The author of this manuscript has been funded by the EPSRC grant EP/R513295/1.

Chapter 4: The Fiber of Persistent Homology for Simplicial Complexes

In this chapter the following materialization of the problem ($\mathbf{PH^{-1}}$) is at stake: we consider the space Filt_K of *filter functions* on a fixed simplicial complex[10] K and study the properties of the fiber of the persistence map:

$$\mathrm{PH} : \mathrm{Filt}_K \longrightarrow \mathbf{Bar}.$$

We introduce two natural actions of increasing homeomorphisms of $\mathbb{R}$ on Filt_K and **Bar** turning PH into an equivariant map. In fact Filt_K and **Bar** are Whitney stratified and homotopically stratified respectively by the orbits of these actions.[11]

Over each barcode stratum $\mathcal{B} \subseteq \mathbf{Bar}$ the persistence map PH is a trivial fiber bundle: in particular all its fibers $\mathrm{PH}^{-1}(D)$ for $D \in \mathcal{B}$ are homeomorphic. More precisely, we find that the fiber enjoys the structure of a *polyhedral complex*.[12] From this structure together with upper-bounds on the dimension of the polyhedra we derive some geometric consequences about $\mathrm{PH}^{-1}(D)$.

Between barcodes D, D' taken from adjacent strata $\mathcal{B} \subseteq \mathcal{B}'$ we find a natural notion of morphism, making **Bar** into a topological category. Each morphism $\phi : D \to D'$ functorially gives rise to a map $\mathcal{L}_\phi : \mathrm{PH}^{-1}(D) \to \mathrm{PH}^{-1}(D')$ relating fibers and sending polyhedra to polyhedra, hence PH^{-1} is best viewed as a functor from barcodes to polyhedral complexes.

We illustrate the analysis by computing the 34 distinct fibers when K is the seemingly elementary example of the triangle. For some barcodes D the fiber $\mathrm{PH}^{-1}(D)$ are homeomorphic to disjoint union of circles or to a Möbius strip.

Online resources, Contributions and Acknowledgements

This is joint work with Ulrike Tillmann. There is a public version, a version published at *Journal of Pure and Applied Algebra*. Besides there is a recorded online talk given at the GEOTOP-A seminar, which reviews some contributions of the current chapter to ($\mathbf{PH^{-1}}$) as well as some contributions from the two subsequent chapters.

The author of this manuscript has been funded by the EPSRC grant EP/R513295/1.

[10] The notion of a simplicial complex K is a natural generalisation of undirected graphs to higher dimensions: $K \subseteq \mathcal{P}([n])$ is a subset of the power set of a finite set $[n]$, closed under inclusion. A filter function is a function $f : K \to \mathbb{R}$ that respects inclusions: $\sigma \subseteq \tau \Rightarrow f(\sigma) \leqslant f(\tau)$.

[11] There are various categories of stratified spaces that differ in the quality of the gluing of adjacent strata. In particular the gluing condition of Whitney stratified spaces [Tho69] is very strong compared to that of homotopically stratified space [Qui88].

[12] A polyhedral complex is a set of polyhedra that intersect at common faces only, and which alike simplicial complexes is closed by taking faces.

Chapter 5: Algorithmic Reconstruction of the Fiber of Persistent Homology on Cell Complexes

The previous chapter has established various properties of the fiber $\mathrm{PH}^{-1}(D)$ such as its structure of polyhedral complex, but also made explicit the fact that computing the homotopy type of $\mathrm{PH}^{-1}(D)$ is a particularly intricate exercise for general simplicial complexes K.

Even intermediate ambitions like determining whether $\mathrm{PH}^{-1}(D)$ is a combinatorial manifold or is non-empty are challenging. In fact for a well-chosen barcode D we show that $\mathrm{PH}^{-1}(D) \neq \varnothing$ amounts to the collapsibility[13] of K. Thus this is a quite deep property since for instance the collapsibility of (a finite iterated barycentric subdivision of) $K \times [0;1]$ for all contractible 2-complexes K is the open Zeeman's conjecture[14] that entails the 3-dimensional Poincaré conjecture.

Only when K is a finite subdivion of the line complex or of the circle the fibers have been determined explicitly[15]. In this chapter we construct an algorithm that computes the polyhedral complex forming the fiber for arbitrary K in order to generate new examples and intuition.

The fiber being a polyhedral complex then facilitates the calculation of intersections of polyhedra, which allows us to compute homology groups and other interesting statistics about the fiber $\mathrm{PH}^{-1}(D)$.

Online resources, Contributions and Acknowledgements

This is joint work with Gregory Henselman-Petrusek. There is a public version of the associated paper, which is still under review at the first journal to which it was submitted. The authors thank Heather Harrington and Ulrike Tillmann for closed guidance and regular discussions.

The author of this manuscript has been funded by the EPSRC grant EP/R513295/1.

[13] An *elementary collapse* of K is the removal of a maximal simplex $\sigma \in K$ together with one of its face $\tau \subseteq \sigma$ that has no other co-face. Then K is *collapsible* if there exists a finite sequence of elementary collapses down to a point. We refer to [Coh12] for a proper introduction to Simple Homotopy Theory.

[14] E Christopher Zeeman. On the dunce hat. *Topology*, 2(4):341–358, 1963

[15] Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020; and Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021

Chapter 6: Fiber of Persistent Homology on Morse functions

This last chapter takes a completely independent angle on ($\mathbf{PH^{-1}}$): here the Morse functions $f : M \to \mathbb{R}$ on a compact smooth finite-dimensional manifold M with boundary replace the filter functions on a simplicial complex K of the previous chapters.[16]

Diffeomorphisms of M (and in particular isotopies[17]) naturally act on Morse functions by pre-composition. Let D be the barcode of a Morse function f and let $\mathrm{PH}_f^{-1}(D) \subseteq \mathrm{PH}^{-1}(D)$ the path connected component of the fiber containing f. Our main observation[18] is that $\mathrm{PH}_f^{-1}(D)$ equals the orbit $\mathcal{O}_{\mathrm{Id}}(f)$ of f under isotopies:

$$\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f).$$

This bridge enables the use of extensive treatments of the homotopy type of diffeomorphism groups in order to analyse the fiber $\mathrm{PH}_f^{-1}(D)$. In particular we rely on the Maksymenko's study of the orbit $\mathcal{O}_{\mathrm{Id}}(f)$ in order to determine the homotopy type of the fiber when M is a connected compact surface without boundary.[19]

In the 1-dimensional settings where M is the unit interval or the circle we are able to extend the analysis to continuous functions (instead of Morse functions), and show that the fibers $\mathrm{PH}^{-1}(D)$ are made of contractible and circular components respectively.

[16] We impose that on the boundary ∂M a Morse function f has no critical point and is constant on each boundary component ∂M_j.

[17] Isotopies are diffeomorphisms in the path connected component of the identity map.

[18] This observation essentially relies on a Mather's fibration theorem for stable smooth mappings [Mat69], which we slightly adapt to the case of Morse functions with equal critical values combining results of Cerf [Cer70].

[19] Sergiy Maksymenko. Homotopy types of stabilizers and orbits of Morse functions on surfaces. *Annals of Global Analysis and Geometry*, 29(3):241–285, 2006

Online resources, Contributions and Acknowledgements

This is joint work with David Beers. There is a public version, a version published at *Journal of Applied and Computational Topology* . The authors are indebted to Ulrike Tillmann whose help has been decisive in overcoming several technical obstacles.

The author of this manuscript has been funded by the EPSRC grant EP/R513295/1.

Part I

Differential of Persistent Homology

1
Differential Calculus on Persistence Barcodes

Abstract

We define notions of differentiability for maps from and to the space of persistence barcodes. Inspired by the theory of diffeological spaces, the proposed framework uses lifts to the space of ordered barcodes, from which derivatives can be computed. The two derived notions of differentiability (respectively from and to the space of barcodes) combine together naturally to produce a chain rule that enables the use of gradient descent for objective functions factoring through the space of barcodes. We illustrate the versatility of this framework by showing how it can be used to analyze the smoothness of various parametrized families of filtrations arising in topological data analysis.

1.1 Introduction

1.1.1 Motivation

Barcodes have been introduced in topological data analysis (TDA) as a means to encode the topological structure of spaces and real-valued functions. They have been shown to provide complementary information compared to classical geometric or statistical methods, which explains their interest for applications. However, so far they have been essentially used as an alternative representation of the input, engineered by the user, as opposed to optimized to fit the problem best.

Optimizing barcodes using e.g. gradient descent requires to differentiate objective functions that factor through the space *Bar* of barcodes:

$$\mathcal{M} \longrightarrow Bar \longrightarrow \mathbb{R}, \tag{1.1}$$

where $\mathcal{M}$ is a parameter space equipped with a differential structure, typically a smooth finite-dimensional manifold. A compelling example arises in the context of supervised learning, where the barcodes can be used as features for data, generated by using some *filter function* $f : K \to \mathbb{R}$ on a fixed graph or simplicial complex K. Instead of considering f as a hyperparameter, it can be beneficial to optimize it among a family $\{f_\theta : K \to \mathbb{R}\}_{\theta \in \mathcal{M}}$ parametrized by a smooth map which we call the *parametrization*:

$$F : \theta \in \mathcal{M} \longmapsto f_\theta \in \mathbb{R}^K.$$

Post-composing F with the *persistent homology* operator PH_p in homology degree p yields a map $\mathrm{PH}_p \circ F : \mathcal{M} \to Bar$. Given a loss function $\mathcal{L} : Bar \to \mathbb{R}$, the goal is then to minimize the functional

$$\mathcal{M} \xrightarrow{\mathrm{PH}_p \circ F} Bar \xrightarrow{\mathcal{L}} \mathbb{R} \tag{1.2}$$

using variational approaches, which are standard in large-scale learning applications. In order to do so, we need to put a sensible smooth structure on *Bar* and to derive an analogue of the chain rule, so that we can compute the differential of $\mathcal{L} \circ \mathrm{PH}_p \circ F$ as the composition of the differentials of $\mathcal{L}$ and $\mathrm{PH}_p \circ F$. The difficulty arises as *Bar* is not a manifold and so far has not been given a structure in which the above makes sense.

Beyond optimization, we want to be able to address other types of applications where differential calculus is involved. For this, a variety of potential scenarios must be considered, e.g. when the filter functions are defined on a fixed smooth manifold, or when the second arrow in (1.1) takes its values in $\mathbb{R}^n$ or more generally in some smooth finite-dimensional manifold. The goal of our study is to provide a unified framework that accounts for all these scenarios.

1.1.2 *Related work*

Despite the lack of a smooth structure on the space *Bar*, developing heuristic methods to differentiate the composition in Equation (1.2) has been an active direction of research lately, leading to innovative computational applications. In Table 1.1, we specify, for each of these contributions, the choice of parametrization F and of loss function $\mathcal{L}$, the optimization problem under consideration, and the sufficient conditions worked out to guarantee the differentiability of the composition in (1.2).

In the context of point cloud inference considered by [GHO16], the positions of points in a fixed Euclidean space form the parameter space $\mathcal{M}$, and the resulting Rips filtration (resp. Alpha filtration) of the total complex on the point cloud is the parametrization F. The loss function $\mathcal{L}$ is given by the least-squares approximation of a fixed barcode. By developing a clear functional point of view on the connection between the barcode of the Rips or Alpha filtration and the positions of the points in the cloud, based on lifts to Euclidean space, the authors show that $\mathcal{L}$ is differentiable wherever the pairwise distances between points in the cloud are distinct. The approach is further refined by [DC19], where it is observed that the parametrization F is a subanalytic map, which implies that the barcode-valued map admits subanalytic (hence generically differentiable) lifts. In turn, this fact is leveraged to show that any probability measure with a density w.r.t. the Hausdorff measure on $\mathcal{M}$ induces an expected persistence diagram (viewed as a measure in the plane) with a density w.r.t. the Lebesgue measure.

In many applications, F parametrizes *lower-star filtrations*, i.e. filter functions induced by their restrictions to the vertices of K [BGGSSG20a, CNBW19, GNDS20a, HGR$^+$20a, HLSC19, PSO18]. In [PSO18], the problem of shape matching is cast into an optimization problem involving the barcodes of the shapes. [CNBW19] uses the degree-0 persistent homology as a regularizer for classifiers. Similarly, [HLSC19] proposes a persistence based regularization as an additional loss for deep learning models in the context of image segmentation. In [HGR$^+$20a], a dataset of graphs is seen as part of a bigger common simplicial complex, which allows to learn a filter function which is shared across the whole dataset. These contributions require the differentiability of (1.2), and they show that it holds whenever the filter function f_θ is injective over the vertex set.

Functions on a grid are used in [BGGSSG20a] to tackle the problem of surface reconstruction. These functions are sums of gaussians whose means and variances are parameters one wants to optimize according to an objective/loss that depends on the degree-1 persistent homology of the functions. [GNDS20a] considers optimization problems involving

Targeted Application	**Loss function $\mathcal{L}$**	**Parametrization F**	**Conditions for Differentiability of (1.2)**
Topological regulariser for classification [CNBW19]	Penalty on the degree-0 persistent homology	Parametrizations of lower star filtrations	Injectivity of the filter function over the vertices
Topological regulariser for image segmentation [HLSC19]	Penalty on the degrees 0 and 1 persistent homology		
Graph classification [HGR+20a]	Any loss of supervised learning		
Shape matching [PSO18]	Bottleneck or Wasserstein distance		
Surface reconstruction [BGGSSG20a]	Sum of persistences		Local correspondence between barcodes interval endpoints and vertices in the simplicial complex
Generative modelling, Adversarial attacks, and Regularization [GNDS20a]	Weighted sum of persistences		
Point cloud continuation [GHO16]	Least-squares approximation of a fixed barcode	Point clouds determining a Rips filtration	Distinct pairwise distances between points

Table 1.1: Some of the current frameworks for differentiating the composition in (1.2). The first column lists the targeted applications. The second and third columns show the choices of loss function $\mathcal{L}$ and parametrization F. The differentiability of $\mathcal{L} \circ \mathrm{PH}_p \circ F$ is guaranteed under the conditions listed in the fourth column.

persistence with many useful applications as in generative modelling, classification robustness, and adversarial attacks. Both contributions need to take the derivative of (1.2), and to do so, they require the existence of an inverse map taking interval endpoints in the persistence diagram $\mathrm{PH}_p(f_\theta)$ to the corresponding vertices of K. This is a strictly weaker requirement than the injectivity of f_θ, as used in the previous contributions, because an inverse map always exists (provided for instance by the standard reduction algorithm for persistent homology). However, per se, it does not guarantee the differentiability of the composition—see e.g. [HGR+20a] for a counter-example.

This variety of applications motivates the search for a unified framework for expressing the differentiability of the arrows in diagrams of

the form:

$$\mathcal{M} \longrightarrow Bar \longrightarrow \mathbb{R}. \tag{1.3}$$

Since the first appearance of this paper as a preprint, there have been novel applications of persistence differentiability in optimisation. For instance, the first author has developed a graph classification framework based on the Laplacian operator [YL21], applying the differentiability of the persistence map (Theorem 1.4.9) to the case of extended persistence. In addition, new heuristics to smooth and regularise loss functions as in Eq. (1.3) improved the optimisation procedure for specific data science problems [CD20, SWB21]. Another strong guarantee is provided when the loss in Eq. (1.3) is semi-algebraic (and more generally subanalytic or definable in some o-minimal structure), as then the classic stochastic gradient descent (SGD) algorithm converges to critical points [DDKL20]. The bridge between this result in non-smooth analysis and persistence based optimisation problems is made in [CCG$^+$21], where sufficient conditions for loss functions as in Eq. (1.3) to be semi-algebraic are given. The main results of [CCG$^+$21] also derive from our general framework, see Remark 1.4.25.

1.1.3 *Contributions and outline of contents*

Ultimately, our framework should make it possible to determine when and how maps between smooth manifolds $\mathcal{M}$ and $\mathcal{N}$ that factor through the space of barcodes can be differentiated:

$$\mathcal{M} \xrightarrow{B} Bar \xrightarrow{V} \mathcal{N}\ .$$

To achieve this goal, in Section 1.3 we define differentiability via lifts in full generality, thereby extending the approach initially proposed by [GHO16] for the specific case of parametrizations by Rips filtrations. Here we provide some of the details. As a space of multi-sets (assumed by default to have finitely many off-diagonal points), *Bar* does not naturally come equipped with a differential structure. However, it is covered by maps of the form:

$$\begin{array}{c} \mathbb{R}^{2m} \times \mathbb{R}^n \\ \Big\downarrow Q_{m,n} \\ Bar \end{array}$$

where $\mathbb{R}^{2m} \times \mathbb{R}^n$ can be thought of as the space of ordered barcodes with fixed number m (resp. n) of finite (resp. infinite) intervals, and where $Q_{m,n}$ is the quotient map modulo the order—turning vectors into multisets (Definition 1.3.1). Then, the map $B : \mathcal{M} \to Bar$ is said

to be *r-differentiable* at parameter $\theta \in \mathcal{M}$ if it admits a local C^r lift $\tilde{B}$ into $\mathbb{R}^{2m} \times \mathbb{R}^n$ for some $m, n \in \mathbb{N}$:

$$\begin{array}{ccc} & & \mathbb{R}^{2m} \times \mathbb{R}^n \\ & \overset{\exists \tilde{B}}{\nearrow} & \downarrow Q_{m,n} \\ \theta \in U \subset \mathcal{M} & \xrightarrow[B]{} & Bar \end{array} \tag{1.4}$$

This means that the map $\tilde{B}$ tracks smoothly and consistently the points in the barcodes $B(\theta')$, for θ' ranging over some open neighborhood U of θ. Dually, the map $V : Bar \to \mathcal{N}$ is r-differentiable at $D \in Bar$ if for every possible choice of m, n, the composition $V \circ Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to Bar$ is C^r on an open neighborhood of every pre-image $\tilde{D}$ of D:

$$\begin{array}{ccc} \mathbb{R}^{2m} \times \mathbb{R}^n & & \\ \downarrow Q_{m,n} & & \\ Bar & \xrightarrow[V]{} & \mathcal{N} \end{array} \tag{1.5}$$

The choice of m, n and pre-image $\tilde{D}$ of D should be thought of as the type of perturbation we allow around D. Thus, essentially, V is asked to be smooth with respect to any finite perturbation of D. In section 1.3.5 we connect these definitions to the theory of diffeological spaces, showing that our two definitions of differentiability for maps B and V are dual to each other and make the barcode space Bar a *diffeological space*.

We then define the differentials of the maps B and V, given simply by the differentials of the lift $\tilde{B} : \mathcal{M} \to \mathbb{R}^{2m} \times \mathbb{R}^n$ (for B) and of the composition $V \circ Q_{m,n}$ on the pre-image $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^n$ (for V). Although these differentials taken individually are not defined uniquely, their corresponding diagrams (1.4) and (1.5) combine together as follows:

$$\begin{array}{ccccc} & & \mathbb{R}^{2m} \times \mathbb{R}^n & & \\ & \overset{\exists \tilde{B}}{\nearrow} & \downarrow Q_{m,n} & & \\ \theta \in U \subset \mathcal{M} & \xrightarrow[B]{} & Bar & \xrightarrow[V]{} & \mathcal{N} \end{array}$$

implying that the composition $V \circ B = (V \circ Q_{m,n}) \circ \tilde{B}$ is a C^r map between smooth manifolds, whose derivative is obtained by composing the differentials of B and V, and this regardless of the choice of lift and pre-image. This is our analogue of the chain rule in ordinary differential calculus (Proposition 1.3.14).

In Sections 1.4 and 1.6, we focus on barcode-valued maps $B : \mathcal{M} \to Bar$ arising from filter functions on fixed smooth manifolds or simplicial

complexes. These maps are usually not differentiable everywhere on their domain. However, motivated by the aforementioned applications, we seek conditions under which B is differentiable almost everywhere on $\mathcal{M}$. A natural approach for this would be to use Rademacher's theorem [Fed69, Thm. 3.1.6], as we know that B is Lispchitz continuous by the Stability Theorem of persistent homology [BL15, CdSGO16, CSEH07]. However, this approach has several important shortcomings:

- it depends on a choice of measure on $\mathcal{M}$;
- it calls for a generalization of Rademacher's theorem to maps taking values in arbitrary metric spaces, and to the best of our knowledge, existing generalizations only provide directional metric differentials (see e.g. [Pan89]);
- more fundamentally, it is not constructive and therefore does not provide formulae for the differentials;
- finally, in the context of optimization, it is important to guarantee the existence of differentials/gradients in an open neighborhood of the considered parameter θ, and not just in a full-measure subset.

We therefore propose to follow a different approach, seeking conditions that ensure the differentiability of B on a generic (i.e. open and dense) subset of $\mathcal{M}$, with explicit differential.

Our first scenario (Section 1.4) considers a parametrization $F : \mathcal{M} \longrightarrow \mathbb{R}^K$ of filter functions on a fixed simplicial complex K. Given a homology degree $p \leqslant d$, where d is the maximal simplex dimension in K, the barcode-valued map B decomposes as $B = \mathrm{PH}_p \circ F$, and in Theorem 1.4.9 we show that B is r-differentiable on a generic subset of $\mathcal{M}$ whenever F is C^r over $\mathcal{M}$ or a generic subset thereof. The proof relies on the fact that the pre-order on the simplices of K induced by the values assigned by the filter function $F(\theta)$ is generically constant around θ in $\mathcal{M}$. We then relate the diffential of B to those of F in Proposition 1.4.14, yielding a closed formula that can be leveraged in practical implementations. Finally, we study the behavior of B at singular points by means of a stratification of the parameter space $\mathcal{M}$, whereby the top-dimensional strata are the locations where B is differentiable, and the lower-dimensional strata characterize the defect of differentiability of B. We show in Theorem 1.4.19 that we can define *directional derivatives* along each incident stratum at any given point $\theta \in \mathcal{M}$. We also show that the barcode valued map can be globally lifted and expressed as a permutation map on each stratum (Corollary 1.4.24).

In Section 1.5 we illustrate the impact of our framework on a series of examples of parametrizations coming from earlier work, including lower-star filtrations, Rips filtrations and some of their generalizations.

For each example, we examine the differentiability of the barcode-valued map and, whenever readily computable, we give the expressions of its differential. This allows us to recover the differentiability results from earlier work in a principled way.

Our second scenario (Section 1.6) considers a parametrization $F : \mathcal{M} \longrightarrow C^\infty(\mathcal{X}, \mathbb{R})$ of smooth filter functions on a fixed smooth compact d-dimensional manifold $\mathcal{X}$. In this scenario, given a parameter $\theta \in \mathcal{M}$, the barcode-valued map B computes all the barcodes of f_θ at once, and collates them in a vector of barcodes:

$$B : \theta \in \mathcal{M} \longmapsto (\mathrm{PH}_0(f_\theta), ..., \mathrm{PH}_d(f_\theta)) \in Bar^{d+1}.$$

We show that B is ∞-differentiable at any parameter θ such that f_θ is Morse with distinct critical values (Theorem 1.6.1). The key insights are: on the one hand, that at any such parameter θ the implicit function theorem allows us to smoothly track the critical points of $f_{\theta'}$ as θ' ranges over a small enough open neighborhood around θ; on the other hand, that the Stability Theorem provides a consistent correspondence between the critical points of $f_{\theta'}$ and the interval endpoints in its barcodes.

In Section 1.7 we look at examples of classes of maps $V : Bar \to \mathcal{N}$. We first consider persistence images [AEK+17] and more generally linear representations of barcodes, as an illustration of our framework on barcode vectorizations. We show that persistence images and linear representations are ∞-differentiable under suitable choices of weighting function (Propositions 1.7.3 and 1.7.5). We then consider the case where $V : Bar \to \mathbb{R}$ is the bottleneck or Wasserstein distance to a fixed barcode, and show it is semi-algebraic in a suitable sense (Proposition 1.7.7), which is useful in a context of optimisation. We then focus on the bottleneck distance to a fixed barcode D_0, which we believe can be of interest in the context of inverse problems. We show that this distance is differentiable on a generic subset of Bar (Propositions 1.7.9 and 1.7.10).

Finally, throughout the paper we sprinkle our exposition with examples of parametrizations and loss functions that illustrate our results and demonstrate their potential for applications.

1.2 *Preliminary notions*

Throughout the paper, vector spaces and homology groups are taken over a fixed field $\mathbb{k}$, omitted in our notations whenever clear from the context. As much as possible, we keep separate terminologies for different notions of differentiability, for instance: maps from or to the space of barcodes are called *r-differentiable* when maps between manifolds are simply called C^r. The only exception to this rule is

the term *smooth* for maps, which has a versatile meaning that should nonetheless always be clear from the context.

1.2.1 Persistence modules and persistent homology

Definition 1.2.1. A *persistence module* $\mathbf{V}$ is a functor from the poset $(\mathbb{R}, \leqslant)$ to the category $\mathbf{Vect}_{\Bbbk}$ of vector spaces over $\Bbbk$.

In other words, a persistence module is a collection $\mathbf{V} = \{V_t, v_{s,t} : V_s \to V_t\}_{(s,t)\in\mathbb{R}^2, s\leqslant t}$ of vector spaces V_t and linear maps $v_{s,t}$, such that $v_{t,t} = \mathrm{id}_{V_t}$ for all $t \in \mathbb{R}$ and $v_{s,t} \circ v_{r,s} = v_{r,t}$ for all $r \leqslant s \leqslant t \in \mathbb{R}$. We say that $\mathbf{V}$ is *pointwise finite-dimensional* (or *pfd* for short) if every V_t is finite-dimensional. Unless otherwise stated, persistence modules in the following will be pfd.

Definition 1.2.2. A *morphism* $\eta : \mathbf{V} \to \mathbf{W}$ between two persistence modules is a natural transformation between functors.

In other words, writing $\mathbf{V} = \{V_t, v_{s,t}\}_{s\leqslant t}$ and $\mathbf{W} = \{W_t, w_{s,t}\}_{s\leqslant t}$, a morphism $\eta : \mathbf{V} \to \mathbf{W}$ is a collection of linear maps $\{\eta_t : V_t \to W_t\}_{t\in\mathbb{R}}$ such that the following diagram commutes for all $s \leqslant t$:

$$\begin{array}{ccc} V_s & \xrightarrow{v_{s,t}} & V_t \\ {\scriptstyle \eta_s}\downarrow & & \downarrow{\scriptstyle \eta_t} \\ W_s & \xrightarrow{w_{s,t}} & W_t \end{array}$$

We say that η is an *isomorphism* of persistence modules if all the η_t are isomorphisms of vector spaces. We denote by **Pers** the category of persistence modules. **Pers** is an abelian category, so it admits kernels, cokernels, images and direct sums, which are defined pointwise. By Crawley-Bovey's Theorem[1], we know that persistence modules essentially uniquely decompose as direct sums of elementary modules called *interval modules*. The interval module $\mathbf{I}_J$ associated to an interval J of $\mathbb{R}$ is defined as the module with copies of the field $\Bbbk$ over J and zero spaces elsewhere, the copies of $\Bbbk$ being connected by identity maps.

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

Theorem 1.2.3. *For any persistence module* $\mathbf{V}$*, there is a unique multi-set* $\mathcal{J}$ *of intervals of* $\mathbb{R}$ *such that*

$$\mathbf{V} \cong \oplus_{J\in\mathcal{J}} \mathbf{I}_J, \tag{1.6}$$

Persistence modules of particular interest are the ones induced by the sub-level sets of real-valued functions.

Definition 1.2.4. Let $f : \mathcal{X} \to \mathbb{R}$ be a real-valued function on a topological space. Write $\mathcal{X}^t := f^{-1}((-\infty, t])$ for the closed sublevel set of f at level $t \in \mathbb{R}$. Given $p \in \mathbf{N}$, the *sublevel set persistent homology* of f in degree p is the (non-necessarily pfd) persistence module $\mathbf{H}_p(f)$ defined by:

- the vector spaces $\{H_p(\mathcal{X}^t)\}_{t\in\mathbb{R}}$, where H_p is the singular homology functor in degree p with coefficients in $\Bbbk$;
- the linear maps $\{v_{s,t} : H_p(\mathcal{X}^s) \to H_p(\mathcal{X}^t)\}_{s\leqslant t}$ induced by inclusions $\mathcal{X}^s \hookrightarrow \mathcal{X}^t$.

In the following we restrict our focus to *finite-type* persistence modules induced by *tame* functions, defined as follows:

Definition 1.2.5. A persistence module $\mathbf{V}$ is of *finite type* if it admits a decomposition into finitely many interval modules.

Definition 1.2.6. A function $f : \mathcal{X} \to \mathbb{R}$ is *tame* if its persistent homology modules in any degree are of finite type.

In particular, *filter functions* on a finite simplicial complex (see below) and *Morse functions* on a smooth manifold (see Section 1.2.3) are tame.

Definition 1.2.7. Let K be a finite simplicial complex. A *filter function* $f : K \to \mathbb{R}$ is a function that is monotonous with respect to inclusions of faces in K, i.e. $f(\sigma) \leqslant f(\sigma')$ for all $\sigma \subseteq \sigma' \in K$. This implies in particular that every sublevel set $K^t := \{\sigma \in K \mid f(\sigma) \leqslant t\}$ is a sub-complex of K.

1.2.2 *Persistence barcodes / diagrams*

Given a decomposition of a finite-type persistence module $\mathbf{V}$ as in (1.6), the (finite) multi-set $\mathcal{J}$ is called the *barcode* of $\mathbf{V}$. An alternate representation is as a (finite) multiset B of points in the plane, where each interval $J \in \mathcal{J}$ is mapped to the point $(\inf J, \sup J)$. To this multiset of points we add Δ^∞, that is the multiset containing countably many copies of the diagonal $\Delta := \{(b, b) \mid b \in \mathbb{R}\}$, to obtain the so-called *persistence diagram* of $\mathbf{V}$. When $\mathbf{V}$ is the sublevel set persistent homology of a tame function f in degree p, we denote by $\mathrm{PH}_p(f)$ its persistence diagram. Persistence diagrams can also be defined independently of persistence modules as follows:

Definition 1.2.8. A *persistence diagram* is the union $B \cup \Delta^\infty$ of a finite multiset B of elements in $\mathbb{R} \times \bar{\mathbb{R}}$, where $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$, with countably many copies of the diagonal Δ. The set of persistence diagrams is denoted by *Bar*.

From now on we also use the terminology *barcodes* for persistence diagrams. Following this terminology, we also call *intervals* the points in a persistence diagram. Points lying on the diagonal Δ are qualified as *diagonal*, the others are qualified as *off-diagonal*.

Remark 1.2.9. In the above definitions we follow the literature on extended persistence, in which persistence diagrams can have points

everywhere in the extended plane $\mathbb{R} \times \bar{\mathbb{R}}$. This is because our framework extends naturally to that setting. Note also that, in the literature, the diagonal is sometimes not included in the diagrams. Here we are including it with infinite multiplicity. This is in the spirit of taking the quotient category of observable persistence modules, as defined by [CCBdS16].

Frédéric Chazal, William Crawley-Boevey, and Vin de Silva. The observable structure of persistence modules. *Homology Homotopy Appl.*, 18(2):247–265, 2016

Definition 1.2.10. Given two barcodes $D, D' \in Bar$, viewed as multisets, a *matching* is a bijection $\gamma : D \to D'$. The *cost* of γ is the quantity

$$c(\gamma) := \sup_{x \in D} \|x - \gamma(x)\|_\infty \in \bar{\mathbb{R}}.$$

We denote by $\Gamma(D, D')$ the set of all matchings between D and D'.

Definition 1.2.11. The *bottleneck distance* between two barcodes $D, D' \in Bar$ is

$$d_\infty(D, D') := \inf_{\gamma \in \Gamma(D,D')} c(\gamma)$$

Given $q \in \mathbb{R}_+^*$, a slight modification of the matching cost yields the q-th *Wasserstein distance* on barcodes as introduced in [CSEHM10]:

$$d_q(D, D') := \inf_{\gamma \in \Gamma(D,D')} \Big(\sum_{x \in D} \|x - \gamma(x)\|_\infty^q\Big)^{\frac{1}{q}} \tag{1.7}$$

David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko. Lipschitz functions have l p-stable persistence. *Foundations of Computational Mathematics*, 10(2):127–139, 2010

Since we include all points in the diagonal with infinite multiplicity in our definition of barcodes, d_∞ is a true metric[2] and not just a pseudo-metric. Indeed, for any $D, D' \in Bar$, we have $d_\infty(D, D') = 0 \Rightarrow D = D'$. We call *bottleneck topology* the topology induced by d_∞, which by the previous observation makes *Bar* a Hausdorff space.

[2] In fact an extended metric as it can take infinite values.

A key fact is the Lipschitz continuity of the barcode function, known as the *Stability Theorem*[3]:

[3] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *Journal of Computational Geometry*, 6(1):162–191, 2015; Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The structure and stability of persistence modules*. SpringerBriefs in Mathematics. Springer, 2016; and David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007

Theorem 1.2.12. *Let $f, g : \mathcal{X} \to \mathbb{R}$ be two real-valued functions with well-defined barcodes. Then,*

$$d_\infty(\mathrm{PH}_p(f), \mathrm{PH}_p(g)) \leqslant \|f - g\|_\infty.$$

Note that the assumptions in the theorem are quite general and hold in our cases of interest: tame functions on a compact manifold, and filter functions on a simplicial complex.

1.2.3 *Morse functions*

Morse functions are a special type of tame functions, for which there is a bijective correspondence between critical points in the domain and interval endpoints in the barcode. This correspondence, detailed in Proposition 1.2.14, will be instrumental in the analysis of Section 1.6. For a proper introduction to Morse theory, we refer the reader to [Mil63].

John Milnor. Morse theory. *Princeton University Press, Princeton, N.J.*, pages vi+153, 1963

Definition 1.2.13. Given a smooth d-dimensional manifold $\mathcal{X}$, a smooth function $f : \mathcal{X} \to \mathbb{R}$ is called *Morse* if its Hessian at critical points (i.e. points where the gradient of f vanishes) is nondegenerate.

Note that we do not assume a priori that the values of f at critical points (called *critical values*) are all distinct. For such a value a, we call *multiplicity* of a the number of critical points in the level-set $f^{-1}(a)$. We also introduce the notation $\mathrm{Crit}(f)$ to refer to the set of critical points, which is discrete in $\mathcal{X}$. In particular, if $\mathcal{X}$ is compact, which will be the case in this paper, $\mathrm{Crit}(f)$ is finite. The number of negative eigenvalues of Δf at a critical point x is called the *index* of x.

Proposition 1.2.14. *Assume $\mathcal{X}$ is compact and all the critical values of f have multiplicity 1. Denote by $E(f)$ the multiset of finite endpoints of off-diagonal intervals (including the left endpoints of infinite intervals) of $\mathrm{PH}_0(f) \sqcup ... \sqcup \mathrm{PH}_d(f)$. Then, f induces a bijection $\mathrm{Crit}(f) \to E(f)$.*

This result is folklore, and we give a proof only for completeness.

Proof. Let $a \leqslant b$ be real numbers. Write $\mathcal{X}^a$ for the sublevel set $f^{-1}((-\infty, a])$. If $[a, b]$ contains a unique critical value c of f, then $\mathcal{X}^b$ has the homotopy type of $\mathcal{X}^a$ glued together with a cell e_p of dimension p, where p is the index of the unique critical point x associated to c. Therefore, $H_*(\mathcal{X}^b, \mathcal{X}^a)$ is trivial except for $* = p$ where it is spanned by the homology class of e_p. This does not depend on the choice of a, b surrounding c and sufficiently close to it. Then, using the long exact sequence in homology, we deduce that either there is one birth in degree p at value c in the persistent homology module, or there is one death in degree $p-1$. Hence, c is either a left endpoint of an interval of $\mathrm{PH}_p(f)$, or a right endpoint of an interval of $\mathrm{PH}_{p-1}(f)$. In either case, we can define the map $x \mapsto f(x)$ for any $x \in \mathrm{Crit}(f)$, and we have just shown that its codomain is indeed $E(f)$. The map is injective because the critical values of f have multiplicity 1 by assumption. We now show it is onto. Let $a \in \mathbb{R}$ be a non-critical value of f. For any (small enough) $\varepsilon, \eta > 0$, the interval $[a - \eta, a + \varepsilon]$ contains no critical value of f, therefore $\mathcal{X}^{a+\varepsilon}$ deform retracts onto $\mathcal{X}^{a-\eta}$, thus implying that the inclusions $H_p(\mathcal{X}^{a-\eta}) \to H_p(\mathcal{X}^{a+\varepsilon})$ are identity maps for any homology degree p. By the decomposition Theorem 1.2.3, this implies that a cannot be an endpoint of an interval summand, i.e. $a \notin E(f)$. □

The assumption that each critical value of f has multiplicity 1 is superfluous in Proposition 1.2.14, if we allow the correspondence map to match trivial intervals. Let $[a, b]$ be an interval containing a unique critical value c. One can still use Morse theory and glue as many critical cells e_p to $\mathcal{X}^a$ as there are critical points in $f^{-1}(c)$ in order to obtain a CW structure on $\mathcal{X}^b$ from the one of $\mathcal{X}^a$. Considering the

different critical cells, we know exactly the ranks of the morphisms $H_p(\mathcal{X}^a) \to H_p(\mathcal{X}^b)$ induced by inclusions in each homology degree p.

1.2.4 *Diffeology theory*

Diffeology theory provides a principled approach to equip a set with a smooth structure. We use some concepts of the theory in Section 1.3.5, where we equip the set *Bar* of barcodes with a diffeology and identify the resulting smooth maps. We refer the reader to [IZ13] for a detailed introduction to the material presented below. In the following, we call *domain* any open set in any arbitrary Euclidean space.

Patrick Iglesias-Zemmour. *Diffeology*. Mathematical Surveys and Monographs, volume 185. American Mathematical Society, Providence, RI, 2013

Definition 1.2.15. Given a non-empty set S, a *diffeology* is a collection $\mathcal{D}$ of pairs (U, P), called *plots*, where U is a domain and $P : U \to S$ is a map from U to S, satisfying the following axioms:

(**Covering**) For any element $s \in S$ and any integer $n \in \mathbf{N}$, the constant map $x \in \mathbb{R}^n \mapsto s \in S$ is a plot.

(**Locality**) If for a pair (U, P) we have that, for any $x \in U$ there exists an open neighborhood $U' \subseteq U$ of x such that the restriction $(U', P_{|U'})$ is a plot, then (U, P) itself is a plot.

ess compatibility) For any plot (U, P) and any smooth map $F : W \to U$ where W is a domain, the composition $(W, P \circ F)$ is a plot.

If a set S comes equipped with a diffeology $\mathcal{D}$, then it is called a *diffeological space*. We think of a diffeological space S as a space where we impose which functions, the plots, from a manifold to S, are smooth. Notice that any set can be made a diffeological space by taking all possible maps as plots. This is the *coarsest* diffeology on S, where $\mathcal{D}$ is said to be *finer* than the diffeology $\mathcal{D}'$ if $\mathcal{D} \subset \mathcal{D}'$, and *coarser* if the converse inclusion holds.[4] The prototypical diffeological space is the Euclidean space $\mathbb{R}^n$ with the usual smooth maps from domains to $\mathbb{R}^n$ as plots.

[4] This terminology is the opposite to the one used when comparing topologies.

Definition 1.2.16. A *morphism* $f : S \to S'$, or *smooth map*, between two diffeological spaces S and S', is a map such that for each plot P of S, $f \circ P$ is a plot of S'. f is called a *diffeomorphism* if it is a bijection and $f^{-1} : S' \to S$ is smooth. A map $f : A \to S'$, where $A \subseteq S$, is *locally smooth* if for any plot P of S, $f \circ P_{|P^{-1}(A)}$ is a plot of S'. f is a *local diffeomorphism* if it is a bijection onto its image and if f^{-1} is locally smooth as a map $S' \supseteq f(A) \to S$.

Obviously, identities are smooth, and smooth maps compose together into smooth maps, therefore we can consider the category **Diffeo** of diffeological spaces. Finite dimensional smooth manifolds with or without boundaries and corners, Fréchet manifolds and Frölicher

spaces, viewed as diffeological spaces with their usual smooth maps, form strict subcategories of **Diffeo**. In fact, finite dimensional smooth manifolds can be defined in the context of diffeology as follows:

Definition 1.2.17. A diffeological space $\mathcal{M}$ is a *n-dimensional diffeological manifold* if it is locally diffeomorphic to $\mathbb{R}^n$ at every point in $\mathcal{M}$.

Theorem 1.2.18 (Section 4.3 in [IZ13])**.** *Every n-dimensional smooth manifold $\mathcal{M}$ is an n-dimensional diffeological manifold once equipped with the diffeology given by the smooth maps $U \to \mathcal{M}$ from arbitrary domains U. Conversely, every n-dimensional diffeological manifold is an n-dimensional smooth manifold.*

Patrick Iglesias-Zemmour. *Diffeology*. Mathematical Surveys and Monographs, volume 185. American Mathematical Society, Providence, RI, 2013

One appealing feature of **Diffeo**, compared to the category of smooth manifolds for instance, is that it is closed under usual set operations—here we only consider coproducts and quotients:

Definition 1.2.19. For an arbitrary family of diffeological spaces $\{(S_j, D_j)\}_{j\in\mathcal{J}}$, the *sum diffeology* on $\bigsqcup_{j\in\mathcal{J}} S_j$ is the finest diffeology making the injections $S_i \to \bigsqcup_{j\in\mathcal{J}} S_j$ smooth.

Definition 1.2.20. For a diffeological space $(S, \mathcal{D})$ and an equivalence relation $\sim$ on S, the *quotient diffeology* on $S/\!\sim$ is the finest diffeology making the quotient map $S \to S/\!\sim$ smooth.

1.2.5 *Stratified manifolds*

Stratified manifolds play a role in Section 1.4.3 of this paper. For background material on the subject, see e.g. [Mat12].

John Mather. Notes on topological stability. *Bull. Amer. Math. Soc. (N.S.)*, 49(4):475–506, 2012

Definition 1.2.21. Let $\mathcal{M}$ be a smooth d-dimensional manifold. A *Whitney stratification* $\mathcal{S}_\mathcal{M}$ of $\mathcal{M}$ is a collection of connected smooth submanifolds (not necessarily closed) of $\mathcal{M}$, called *strata*, satisfying the following axioms:

(Partition) The strata partition $\mathcal{M}$.

(Locally finite) Each point of $\mathcal{M}$ has an open neighborhood meeting with finitely many strata.

(Frontier) For each stratum $\mathcal{M}' \in \mathcal{S}_\mathcal{M}$, the set $\overline{\mathcal{M}'}\backslash\mathcal{M}'$ is a union of strata, where $\overline{\mathcal{M}'}$ is the closure of $\mathcal{M}'$ in $\mathcal{M}$.

(Condition b) Consider a pair of strata $(\mathcal{M}', \mathcal{M}'')$ and an element $\theta \in \mathcal{M}'$. If there are sequences of points $(\theta'_k)_{k\in\mathbf{N}}$ and $(\theta''_k)_{k\in\mathbf{N}}$ lying in $\mathcal{M}'$ and $\mathcal{M}''$ respectively, both converging to θ, such that the line (θ'_k, θ''_k) (defined in some local coordinate system around θ) converges to some line l and $T_{\theta''_k}\mathcal{M}''$ converges to some flat, then this flat contains l.

Stratified maps are those that behave nicely with respect to stratifications. Here we only use a subset of the axioms they satisfy, hence we talk about *weakly stratified maps*.

Definition 1.2.22. Let $\mathcal{M}, \mathcal{N}$ be stratified manifolds. A map $f : \mathcal{M} \to \mathcal{N}$ is *weakly stratified* if the pre-images $f^{-1}(\mathcal{N}')$, for any stratum $\mathcal{N}' \in \mathcal{S}_{\mathcal{N}}$, is a union of strata in $\mathcal{S}_{\mathcal{M}}$.

1.3 *Differentiability for maps from or to the space of barcodes*

In Section 1.3.1 we provide a general framework for studying the differentiability of maps from a smooth manifold to *Bar*. Then in Section 1.3.2 we provide the analogue for maps with *Bar* as domain and a smooth manifold as co-domain. Both frameworks are in some sense dual to each other, and inspired by the theory of diffeological spaces—we develop this connection in Section 1.3.5. We then derive a chain rule in Section 1.3.3: if a map between manifolds factors through *Bar*, then it is smooth whenever both terms in the factorization are smooth according to our definitions, and in this case its differential can be computed explicitly.

1.3.1 *Differentiability of barcode valued maps*

Throughout this section, $\mathcal{M}$ denotes a smooth finite-dimensional manifold without boundary, which may or may not be compact. Our approach to characterizing the smoothness of a barcode valued map is to factor it through the bundle of *ordered barcodes*:

Definition 1.3.1. For each choice of non-negative integers m, n, the space of *ordered barcodes* with m finite bars and n infinite ones is $\mathbb{R}^{2m} \times \mathbb{R}^n$, equipped with the Euclidean norm and the resulting smooth structure. The corresponding *quotient map* $Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to Bar$ quotients the space by the action[5] of the product of symmetric groups $\Sigma_m \times \Sigma_n$, that is: for any ordered barcode $\tilde{D} = (b_1, d_1, ..., b_m, d_m, v_1, ..., v_n) \in \mathbb{R}^{2m} \times \mathbb{R}^n$,

$$Q_{m,n}(\tilde{D}) := \{(b_i, d_i)\}_{i=1}^{m} \cup \{(v_j, +\infty)\}_{j=1}^{n} \cup \Delta^{\infty}.$$

[5] Σ_m acts on $\mathbb{R}^{2m}$ by permutation of pairs of adjacent coordinates while Σ_n acts on $\mathbb{R}^n$ by permutation of coordinates.

One can think of an ordered barcode $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^n$ as a vector describing a persistence diagram with at most m bounded off-diagonal points and exactly n unbounded points. The former have their coordinates encoded in the adjacent pairs of the $2m$ first components in $\tilde{D}$, while the latter have the abscissa of their left endpoint encoded in the last n components of $\tilde{D}$. The quotient map $Q_{m,n}$ forgets about the ordering of the bars in the barcodes. So far $Q_{m,n}$ is merely a map between sets, and it is natural to ask whether it is regular in some reasonable sense:

Proposition 1.3.2. *For any $m, n \in \mathbb{N}^2$, $Q_{m,n}$ is 1-Lipschitz when Bar is equipped with the bottleneck topology.*

Proof. For any two elements $\tilde{D}_1, \tilde{D}_2 \in \mathbb{R}^{2m} \times \mathbb{R}^n$, there is an obvious matching γ on their images $Q_{m,n}(\tilde{D}_1), Q_{m,n}(\tilde{D}_2)$ given by matching the components of the vectors $\tilde{D}_1$ and $\tilde{D}_2$ entry-wise. The cost of this matching is then bounded above by the supremum norm of $\tilde{D}_1 - \tilde{D}_2$, by the definition of the matching cost $c(\gamma)$. In turn, the supremum norm is bounded above by the ℓ^2 norm. □

We then say that a barcode valued map is smooth if it admits a smooth lift into the space of ordered barcodes for some choice of m, n:

Definition 1.3.3. Let $B : \mathcal{M} \to Bar$ be a barcode valued map. Let $x \in \mathcal{M}$ and $r \in \mathbb{N} \cup \{+\infty\}$. We say that B *is r-differentiable at* x if there exists an open neighborhood U of x, integers $m, n \in \mathbb{N}$ and a map $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ of class C^r such that $B = Q_{m,n} \circ \tilde{B}$ on U. For an integer $d \in \mathbb{N}$, a function $\mathcal{B} : \mathcal{M} \to Bar^{d+1}$ is *r-differentiable at* $x \in \mathcal{M}$ if each of its $d+1$ components is. We call $\tilde{B}$ a *local lift* of B.

Remark 1.3.4 (Locally finite number of off-diagonal points)**.** If a function B as above is r-differentiable at $x \in \mathcal{M}$, then locally for any x' around x we can upper-bound the number of off-diagonal points arising in $B(x')$ by $m + n$. Notice that off-diagonal points can possibly appear in $B(x')$ and become part of the diagonal Δ in $B(x)$, which is to say that Defnition 1.3.3 does not restrict the function B to locally consist in a fixed number of off-diagonal points. Informally, in analogy with the fact that a barcode has finitely many off-diagonal points, our definition of smoothness allows finitely many appearances or disappearances of off-diagonal points in the neighborhood of a barcode.

Remark 1.3.5 (0-differentiability is stronger than bottleneck continuity)**.** If $B : \mathcal{M} \to Bar$ is 0-differentiable, then B is continuous when Bar is given the bottleneck topology. This comes from the Lipschitz continuity of $Q_{m,n}$ (Proposition 1.3.2) and the fact that continuity is stable under composition. The converse is false, because, on the one hand, if B is 0-differentiable then locally the number of off-diagonal points in the image of B is uniformly bounded (see the previous remark), while on the other hand, the number of off-diagonal points appearing in barcodes in any given open bottleneck ball is arbitrarily large.

Definition 1.3.6. Let $B : \mathcal{M} \to Bar$ be 1-differentiable at some x, and $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ be a C^1 lift of B defined on an open neighborhood U of x. The *differential* (or *derivative*) $d_{x,\tilde{B}}B$ of B at x with respect to $\tilde{B}$ is defined to be the differential of $\tilde{B}$ at x:

$$T_x\mathcal{M} \xrightarrow[d_x\tilde{B}]{} \mathbb{R}^{2m} \times \mathbb{R}^n.$$

Post-composing with the quotient map, we can see $Q_{m,n} \circ d_{x,\tilde{B}} B : T_x\mathcal{M} \to Bar$ as a multi-set of co-vectors, one above each off-diagonal point of $B(x)$ (plus some distinguished diagonal points), describing linear changes in the coordinates of the points of $B(x)$ under infinitesimal perturbations of x. In this respect, the spaces of ordered barcodes $\mathbb{R}^{2m+n}$ play the role of tangent spaces over Bar. For practical computations, it can be convenient to work with an alternate yet equivalent notion of differentiability, based on point trackings:

Definition 1.3.7. Let $B : \mathcal{M} \to Bar$ be a barcode valued map. Let $x \in \mathcal{M}$ and $r \in \mathbb{N} \cup \{+\infty\}$. A C^r *local coordinate system* for B at x is a collection of maps $\{b_i, d_i : U \to \mathbb{R}\}_{i\in I}$ and $\{v_j : U \to \mathbb{R}\}_{j\in J}$ for finite sets I, J defined on an open neighborhood U of x, such that:

(**Smooth**) The maps b_i, d_i, v_j are of class C^r;

(**Tracking**) For any $x' \in U$ we have the multi-set equality

$$B(x') = \{(b_i(x'), d_i(x'))\}_{i\in I} \cup \{(v_j(x'), +\infty)\}_{j\in J} \cup \Delta^\infty.$$

Thus, in a local coordinate system, we have maps b_i, d_i (resp. v_j) that track the endpoints of bounded (resp. unbounded) intervals in the image barcode through B. We will often abbreviate the data of a local coordinate system of B at x by $\mathcal{T} = (U, \{b_i, d_i\}_{i\in I}, \{v_j\}_{j\in J})$.

Our two notions of differentiability are indeed equivalent:

Proposition 1.3.8. *Let $B : \mathcal{M} \to Bar$ be a barcode valued map and $x \in \mathcal{M}$. Then B is r-differentiable at x if and only if it admits a C^r local coordinate system at x. Specifically, post-composing a C^r local lift $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ around x with the quotient map $Q_{m,n}$ yields a C^r local coordinate system, and conversely, fixing an order on the functions of a C^r local coordinate system yields a C^r local lift.*

Proof. ($\Rightarrow$) Let $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ be a C^r local lift of B at x. Extract the components of $(b_1(x'), d_1(x'), ..., b_m(x'), d_m(x'), v_1(x'), ..., v_n(x')) := \tilde{B}(x')$ to get a local coordinate system, which is C^r over U as $\tilde{B}$ is. ($\Leftarrow$) Let $\mathcal{T} = (U, \{b_i, d_i\}_{i\in I}, \{v_j\}_{j\in J})$ be a C^r local coordinate system for B at x. Set $m = |I|$ and $n = |J|$, and fix two arbitrary bijections $s : \{1, ..., m\} \to I$ and $t : \{1, ..., n\} \to J$. Then the map $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ defined as:

$$\tilde{B}(x') := [b_{s(1)}(x'), d_{s(1)}(x'), ..., b_{s(m)}(x'), d_{s(m)}(x'), v_{t(1)}(x'), ..., v_{t(n)}(x')]$$

is a lift of B. As a map valued in a Euclidean space, $\tilde{B}$ is C^r because all its coordinate functions are. □

Remark 1.3.9 (Non uniqueness of differentials)**.** It is important to keep in mind that the differential of B at x is not uniquely defined, as it depends on the choice of local lift. Indeed, for two distinct lifts $\tilde{B}, \tilde{B}'$ of

B at x, we usually get distinct differentials $dB_{x,\tilde{B}}$, $dB_{x,\tilde{B}'}$. For instance, if $\tilde{B}'$ is obtained from $\tilde{B}$ by appending an extra pair of coordinates of the form (f, f), where f is a smooth real function, then $dB_{x,\tilde{B}'}$ takes its values in a different codomain than that of $dB_{x,\tilde{B}}$. Note that this will not be an issue in the rest of the paper, as any choice of differential will yield a valid chain rule (Section 1.3.3).

1.3.2 *Differentiability of maps defined on barcodes*

Let $\mathcal{N}$ be a smooth finite-dimensional manifold without boundary. Our notion of differentiability for maps $V : Bar \to \mathcal{N}$ is in some sense dual to the one for maps $B : \mathcal{M} \to Bar$, as will be justified formally in the next section.

Definition 1.3.10. Let $V : Bar \to \mathcal{N}$ be a map on barcodes. Let $D \in Bar$ and $r \in \mathbb{N} \cup \{+\infty\}$. V is said to be *r-differentiable* at D, if for all integers m, n and all vectors $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^n$ such that $Q_{m,n}(\tilde{D}) = D$, the map $V \circ Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to \mathcal{N}$ is C^r on an open neighborhood of $\tilde{D}$.

Notice that for each choice of m, n we have a unique map $V \circ Q_{m,n}$, and we must check its differentiability at all the (possibly many) distinct pre-images $\tilde{D}$ of D and for all m, n. One can think of a choice of m, n and pre-image $\tilde{D}$ of D as a choice of tangent space of Bar at D.

Example 1.3.11 (Total persistence function)**.** Let $V : Bar \to \mathbb{R}$ be defined as the sum, over bounded intervals (b, d) in a barcode D, of the length $(d - b)$. Given $D \in Bar$ and an ordered barcode $\tilde{D} \in \mathbb{R}^{2m+n}$ such that $Q_{m,n}(\tilde{D}) = D$, the map $V \circ Q_{m,n}$ is a linear form and in particular is of class C^∞ at $\tilde{D}$. Explicitly, we have

$$V \circ Q_{m,n} : (b_1, d_1, ..., b_m, d_m, v_1, ..., v_n) \in \mathbb{R}^{2m+n} \mapsto \sum_{i=1}^{m} d_i - b_i \in \mathbb{R}$$

Therefore, V is ∞-differentiable everywhere on Bar.

The relationship between 0-differentiability and the bottleneck continuity for maps V is the opposite to the one that holds for maps B (recall Remark 1.3.5):

Remark 1.3.12 (Bottleneck continuity is stronger than 0-differentiability)**.** If $V : Bar \to \mathcal{N}$ is continuous when Bar is equipped with the bottleneck topology, then V is 0-differentiable. This is because the quotient map $Q_{m,n}$ is continuous (Proposition 1.3.2) and the composition of continuous maps is continuous. The converse is false, as seen for instance when taking V to be the total persistence function: although 0-differentiable (because ∞-differentiable) on Bar, V is not continuous in the bottleneck topology as it is unbounded in any open bottleneck ball.

Definition 1.3.13. Let $V : Bar \to \mathcal{N}$ be 1-differentiable at $D \in Bar$, and $\tilde{D} \in \mathbb{R}^{2m+n}$ be a pre-image of D via $Q_{m,n}$. The *differential* (or *derivative*) of V at D with respect to $\tilde{D}$ is the map

$$d_{D,\tilde{D}}V : \mathbb{R}^{2m+n} \xrightarrow[d_{\tilde{D}}(V \circ Q_{m,n})]{} T_{V(D)}\mathcal{N}.$$

1.3.3 Chain rule

We now combine the previous definitions to produce a chain rule.

Proposition 1.3.14. *Let $B : \mathcal{M} \to Bar$ be r-differentiable at $x \in \mathcal{M}$, and $V : Bar \to \mathcal{N}$ be r-differentiable at $B(x)$. Then:*

(i) $V \circ B : \mathcal{M} \to \mathcal{N}$ is C^r at x as a map between smooth manifolds;

(ii) If $r \geqslant 1$, then for any local C^1 lift $\tilde{B} : U \to \mathbb{R}^{2m+n}$ of B around x we have:

$$d_x(V \circ B) = d_{B(x),\tilde{B}(x)}V \circ d_{x,\tilde{B}}B.$$

The meaning of this formula is that, even though the differentials of B and of V may depend on the choice of lift $\tilde{B} : \mathcal{M} \to \mathbb{R}^{2m+n}$, their composition does not, and in fact it matches with the usual differential of $V \circ B$ as a map between smooth manifolds.

Proof. Since B is r-differentiable at x, there exists an open neighborhood U of x and a local C^r lift $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ for some integers m, n, such that $B|_U = Q_{m,n} \circ \tilde{B}$. Meanwhile, since V is r-differentiable at $B(x)$, the map $V \circ Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to \mathcal{N}$ is C^r at $\tilde{B}(x)$. This implies that the composition $V \circ B|_U = (V \circ Q_{m,n}) \circ \tilde{B}$ is C^r at x, and therefore that $V \circ B$ itself is C^r at x since U is open. This proves (i). The formula of (ii) follows then from applying the usual chain rule to $(V \circ Q_{m,n})$ and $\tilde{B}$, which are C^1 maps between smooth manifolds without boundary. □

Example 1.3.15. In [HGR+20a], given a C^∞ neural network architecture $F_0 : \mathbb{R}^N \to \mathbb{R}^{K_0}$ valued in the set of functions over the vertices of a fixed graph K, the optimization pipeline requires taking the gradient of the following loss function:

$$\mathcal{L} : \theta \in \mathbb{R}^N \longmapsto \sum_{(b,d) \in \mathrm{PH}_p(F_0(\theta)) \setminus \Delta \text{ bounded}} s(b,d) \in \mathbb{R},$$

where $s : \mathbb{R}^2 \to \mathbb{R}$ is a fixed smooth map, and $\mathrm{PH}_p(F_0(\theta))$ is the degree-p persistence diagram associated to the lower star filtration induced by $F_0(\theta)$ on K (see Section 1.5.1 dedicated to the full analysis of lower star filtrations). We may see $\mathcal{L}$ as the composition:

$$\mathcal{L} : \ \theta \in \mathbb{R}^N \xrightarrow{B} \mathrm{PH}_p(F_0(\theta)) \in Bar \xrightarrow{V} V(\mathrm{PH}_p(F_0(\theta))) \in \mathbb{R} \ ,$$

Christoph Hofer, Florian Graf, Bastian Rieck, Marc Niethammer, and Roland Kwitt. Graph filtration learning. In *International Conference on Machine Learning*, pages 4314–4323. PMLR, 2020

where $V : D \in Bar \mapsto \sum_{(b,d)\in D\setminus\Delta \text{ bounded}} s(b,d) \in \mathbb{R}$. On the one hand, B is ∞-differentiable at every θ where $F_0(\theta)$ is injective, as will be detailed in Section 1.5.1. On the other hand, V is ∞-differentiable everywhere on Bar, a fact obtained exactly as in the case of the total persistence function of Example 1.3.11. By the chain rule (Proposition 1.3.14), we deduce that the loss $\mathcal{L}$ is smooth at every θ where $F_0(\theta)$ is injective. Thus we recover the differentiability result of [HGR+20a]. In fact, the upcoming Theorem 1.4.9 ensures that B is ∞-differentiable over an open dense subset of $\mathbb{R}^N$, and therefore so is $\mathcal{L}$ by the chain rule.

1.3.4 *Higher-order derivatives*

The notions of derivatives introduced in Definitions 1.3.6 and 1.3.13 extend naturally to higher orders. For simplicity, we place ourselves in the Euclidean setting, letting $\mathcal{M} = \mathbb{R}^N$ and $\mathcal{N} = \mathbb{R}^{N'}$ for some $N, N' \in \mathbb{N}$.

Definition 1.3.16. Let $B : \mathbb{R}^N \to Bar$ be r-differentiable at some x, and $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ be a C^r lift of B defined on an open neighborhood U of x. The *r-th differential* (or *derivative*) of B at x with respect to $\tilde{B}$ is defined to be the r-th Fréchet differential of $\tilde{B}$ at x:

$$d^r_x\tilde{B} : (\mathbb{R}^N)^r \to \mathbb{R}^{2m} \times \mathbb{R}^n.$$

Dually:

Definition 1.3.17. Let $V : Bar \to \mathbb{R}^{N'}$ be r-differentiable at $D \in Bar$, and $\tilde{D} \in \mathbb{R}^{2m+n}$ be a pre-image of D via $Q_{m,n}$. The *r-th differential* (or *derivative*) of V at D with respect to $\tilde{D}$ is the r-th Fréchet differential of $V \circ Q_{m,n}$ at $\tilde{D}$:

$$d^r_{\tilde{D}}(V \circ Q_{m,n}) : (\mathbb{R}^{2m+n})^r \to \mathbb{R}^{N'}.$$

Note that, given maps $B : \mathbb{R}^N \to Bar$ and $V : Bar \to \mathbb{R}^{N'}$ that are r-differentiable at x and $B(x)$ respectively, the chain rule of Section 1.3.3 adapts readily to higher-order derivatives of $B \circ V$ at x.

Meanwhile, we get a natural *Taylor expansion* of B at x with respect to $\tilde{B}$:

$$T^r_{x,\tilde{B}}B : h \in \mathbb{R}^N \longmapsto \tilde{B}(x) + d_x\tilde{B}(h) + \cdots + \frac{1}{r!}d^r_x\tilde{B}(h, \cdots, h) \in \mathbb{R}^{2m} \times \mathbb{R}^n.$$

Proposition 1.3.18. *Let $B : \mathbb{R}^N \to Bar$ be r-differentiable at some x, and $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ be a C^r lift of B defined on an open neighborhood U of x. Then,*

$$d_\infty(B(x+h), (Q_{m,n} \circ T^r_{x,\tilde{B}}B)(h)) = o(\|h\|^r).$$

Proof. This follows from applying the standard Taylor-Young theorem to $\tilde{B}$, then post-composing by $Q_{m,n}$—which is 1-Lipschitz by Proposition 1.3.2. □

To our knowledge, there is in general no equivalent of this result for the map V, due to the lack of a Lipschitz-continuous section of $Q_{m,n}$.

1.3.5 *The space of barcodes as a diffeological space*

In this subsection, we detail how *Bar*, when viewed as the quotient of a disjoint union of Euclidean spaces, is canonically made into a diffeological space, as defined in Section 1.2.4. We then show that the resulting notions of diffeological smooth maps from and to *Bar* coincide with the definitions 1.3.3 and 1.3.10 of differentiability we chose for maps from and to *Bar* in the previous sections, thus making these two definitions dual to each other.

As a set, *Bar* is isomorphic to $\left(\bigsqcup_{m,n\in\mathbb{N}} \mathbb{R}^{2m+n}\right)/\sim$, where $\sim$ is the transitive closure of the following relations for m, n ranging over $\mathbb{N}$:

- For any permutations π, τ of $\{1, ..., m\}$ and $\{1, ..., n\}$ respectively,

$$[(b_i, d_i)_{i=1}^m, (v_j)_{j=1}^n] \sim [(b_{\pi(i)}, d_{\pi(i)})_{i=1}^m, (v_{\tau(j)})_{j=1}^n],$$

  which indicates that persistence diagrams are multisets (i.e. intervals are not ordered);

- Any element $[(b_i, d_i)_{i=1}^m, (v_j)_{j=1}^n] \in \mathbb{R}^{2m+n}$ such that one of the first m adjacent pairs (b_i, d_i) satisfies $b_i = d_i$ is equivalent to the element of $\mathbb{R}^{2(m-1)+n}$ obtained by removing (b_i, d_i). These identifications correspond to quotienting multisets by the diagonal Δ.

Since the Euclidean spaces $\mathbb{R}^{2m+n}$ are equipped with their Euclidean diffeologies, we obtain a canonical diffeology $\mathcal{D}(Bar)$ over *Bar* from Definitions 1.2.19 and 1.2.20. The plots of $\mathcal{D}(Bar)$ can be concretely characterized as follows:

Proposition 1.3.19. *Let $U \subseteq \mathbb{R}^d$ be open and $B : U \to Bar$. Then B is a plot in $\mathcal{D}(Bar)$ if and only if, for every $x \in U$, there exists an open neighborhood $V \subseteq U$ of x and a C^∞ lift $\tilde{B} : V \to \mathbb{R}^{2m+n}$ such that $B_{|V} = Q_{m,n} \circ \tilde{B}$.*

In other words, a plot in $\mathcal{D}(Bar)$ is an ∞-differentiable map from a domain U to *Bar*.

Proof. Note that the characterization of the quotient diffeology, as given in Definition 1.2.20, is in fact the characterization of the so-called *push-forward diffeology* induced by the quotient map—see [IZ13, § 1.43][6]. According to that characterization, $B : U \to Bar$ is a plot if and only if, for every element $z \in U$, there exists an open neighborhood $W \subseteq U$ of z such that the restriction $B_{|W}$ admits a lift[7] $\tilde{B} : W \to \bigsqcup_{m,n\in\mathbb{N}} \mathbb{R}^{2m+n}$, i.e. a plot $\tilde{B}$ of $\bigsqcup_{m,n\in\mathbb{N}} \mathbb{R}^{2m+n}$ that matches with $B_{|W}$ once post-composed with the quotient map modulo $\sim$. In turn, by the characterization of the sum diffeology in [IZ13, § 1.39], $\tilde{B}$ is a plot of $\bigsqcup_{m,n\in\mathbb{N}} \mathbb{R}^{2m+n}$ if and only

[6] Patrick Iglesias-Zemmour. *Diffeology*. Mathematical Surveys and Monographs, volume 185. American Mathematical Society, Providence, RI, 2013

[7] Strictly speaking, according to [IZ13, § 1.43] there is also the alternative that the restriction $B_{|W}$ be constant, but in this case it also admits a lift to $\bigsqcup_{m,n\in\mathbb{N}} \mathbb{R}^{2m+n}$. Indeed, calling D the unique barcode in the image of $B_{|W}$, we can choose one pre-image $\tilde{D}$ of D in one of the spaces of ordered barcodes $\mathbb{R}^{2m+n}$, then take $\tilde{B}$ to be the constant map $W \to \{\tilde{D}\}$.

if, for any $x \in W$, there is an open neighborhood $V \subseteq W$ of x and a pair of indices (m, n) such that the restriction $\tilde{B}_{|V}$ maps into $\mathbb{R}^{2m+n}$ and is in fact a plot of $\mathbb{R}^{2m+n}$. Equivalently, we have $B_{|V} = Q_{m,n} \circ \tilde{B}_{|V}$, where $\tilde{B}_{|V}$ is of class C^∞ (since the spaces of ordered barcodes are equipped with their canonical Euclidean diffeologies). □

Corollary 1.3.20. *The smooth maps in **Diffeo** from a smooth manifold $\mathcal{M}$ without boundary (equipped with the diffeology from Theorem 1.2.18) to the diffeological space Bar are exactly the ∞-differentiable maps from $\mathcal{M}$ to Bar.*

Proof. Let $B : \mathcal{M} \to Bar$ be a smooth map in **Diffeo**. For any plot $\phi : U \to \mathcal{M}$, the composition $B \circ \phi$ is a plot in $\mathcal{D}(Bar)$, therefore it locally rewrites as $Q_{m,n} \circ \tilde{B}$ for some C^∞ lift $\tilde{B}$, by Proposition 1.3.19. Choosing ϕ to be a local coordinate chart, we then locally have $B = Q_{m,n} \circ \tilde{B} \circ \phi^{-1}$, which means that B is ∞-differentiable. Conversely, if B is ∞-differentiable, it locally rewrites as $B = Q_{m,n} \circ \tilde{B}$, hence for any plot $\phi : U \to \mathcal{M}$ the composition $B \circ \phi$ locally rewrites as $Q_{m,n} \circ \tilde{B} \circ \phi$ and therefore is a plot in $\mathcal{D}(Bar)$ by Proposition 1.3.19. □

Dually:

Corollary 1.3.21. *The smooth maps in **Diffeo** from the diffeological space Bar to a smooth manifold $\mathcal{N}$ without boundary (equipped with the diffeology from Theorem 1.2.18) are exactly the ∞-differentiable maps from Bar to $\mathcal{N}$.*

Proof. Let $V : Bar \to \mathcal{N}$ be a smooth map in **Diffeo**. By Proposition 1.3.19, any ∞-differentiable map $B : U \to Bar$ defined on a domain U is a plot, therefore the composition $V \circ B : U \to \mathcal{N}$ is a plot hence C^∞. In particular, the map $Q_{m,n} = Q_{m,n} \circ \mathrm{Id}_{\mathbb{R}^{2m+n}} : \mathbb{R}^{2m+n} \to Bar$ is ∞-differentiable, therefore $V \circ Q_{m,n}$ is C^∞. This shows that V is ∞-differentiable. Conversely, if V is ∞-differentiable, the maps $V \circ Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to \mathcal{N}$, for varying integers m, n, are C^∞. By Proposition 1.3.19, if $B : U \to Bar$ is a plot, then it locally rewrites as $Q_{m,n} \circ \tilde{B}$ for some C^∞ lift $\tilde{B}$, therefore $V \circ B$ is locally of the form $(V \circ Q_{m,n}) \circ \tilde{B}$, which is of class C^∞ as a map between manifolds by the chain rule. Thus, $V \circ B$ is a plot, and therefore V is smooth in **Diffeo**. □

Conceptually, we have made *Bar* into a diffeological space by viewing it as the quotient of the direct limit of the spaces of ordered barcode. Then, ∞-differentiable maps are simply morphisms in **Diffeo** from or to smooth manifolds, rather than maps satisfying the a priori unrelated definitions 1.3.3 and 1.3.10. More generally, by seeing *Bar* as one object in **Diffeo** where morphisms can come in or out, we have notions of smooth maps from or to *Bar* with respect to any other diffeological space. For instance, a map $f : Bar \to Bar$ is smooth if and only if all the maps $f \circ Q_{m,n}$, for varying integers m, n, are ∞-differentiable (the proof is left as an exercise to the reader). Note however that diffeology

does not characterize the r-differentiable maps for finite r nor the maps that are differentiable only locally, two concepts that are prominent in our analysis.

1.4 *The case of barcode valued maps derived from real functions on a simplicial complex*

In this section we consider barcode valued maps $B_p : \mathcal{M} \to Bar$ that factor through the space $\mathbb{R}^K$ of real functions on a fixed finite abstract simplicial complex K:

$$B_p : \quad \mathcal{M} \xrightarrow{F} \mathbb{R}^K \xrightarrow{\mathrm{PH}_p} Bar$$

In other words, we consider barcodes derived from real functions on K. Note that PH_p, the barcode map in degree p, is only defined on the subspace of filter functions, i.e. functions $K \to \mathbb{R}$ that are monotonous with respect to inclusions of faces in K. This subspace is a convex polytope bounded by the hyperplanes of equations $f(\sigma) = f(\sigma')$ for $\sigma \subsetneq \sigma' \in K$. From now on, we consistently assume that F takes its values in this polytope.

Example 1.4.1 (Height filters)**.** Given an embedded simplicial comple $K \subseteq \mathbb{R}^d$, let $\mathcal{M} = \mathbb{S}^{d-1}$ and $F : \theta \mapsto (\sigma \in K \mapsto \max_{x \in \sigma} \langle \theta, x \rangle)$. The filter functions considered here are the height functions on K, parametrized on the unit sphere $\mathbb{S}^{d-1}$ by the map F.

By analogy with the previous example, we generally call F the *parametrization* associated to B, although it may not always be a topological embedding of $\mathcal{M}$ into $\mathbb{R}^K$ (it may not even be injective). We also call $\mathcal{M}$ the *parameter space*, and use the generic notation θ to refer to an element in $\mathcal{M}$.

As we shall see in Section 1.4.1, a local coordinate system for the map B_p at $\theta \in \mathcal{M}$ can be derived when the order of the values of the filter function $F(\theta)$ remains constant locally around θ. For this purpose we introduce the following equivalence relation on filter functions $K \to \mathbb{R}$:

Definition 1.4.2. Given a filter function $f : K \to \mathbb{R}$, the increasing order of its values induce a pre-order on the simplices of K. Two filter functions f, g are said to be *ordering equivalent*, written $f \sim g$, if they induce the same pre-order on K. This relation is an equivalence relation on filter functions, and we denote by $[f]$ the equivalence class of f. The (finite) set of equivalence classes is denoted by $\Omega(\mathbb{R}^K)$.

In order to compare barcodes across an entire equivalence class of functions, we introduce *barcode templates* as follows:

Definition 1.4.3. Given a filter function $f \in \mathbb{R}^K$ and a homology degree $0 \leqslant p \leqslant d$, a *barcode template* (P_p, U_p) is composed of a multiset P_p of pairs of simplices in K, together with a multiset U_p of simplices in K, such that:

$$\mathrm{PH}_p(f) = \left\{(f(\sigma), f(\sigma'))\right\}_{(\sigma,\sigma')\in P_p} \cup \left\{(f(\sigma), +\infty)\right\}_{\sigma\in U_p} \cup \Delta^\infty \tag{1.8}$$

Note that we do not require a priori that $\dim\sigma = p$ and $\dim\sigma' = p+1$.

Proposition 1.4.4. *For any filter function $f \in \mathbb{R}^K$ and homology degree $0 \leqslant p \leqslant d$, there exists a barcode template (P_p, U_p) of f.*

Proof. Consider the interval decomposition $\mathbf{H}_p(f) \cong \oplus_{J\in\mathcal{J}} \mathbb{I}_J$ of the p-th persistent homology module of f. Note that every interval endpoint in the decomposition corresponds to the f-value of some simplex of K (since the persistent homology module has internal isomorphisms in-between these values). For every bounded interval J with endpoints $b, d \in \mathbb{R}$ choose an element (σ_J, σ'_J) in $f^{-1}(b) \times f^{-1}(d) \subseteq K \times K$, then form the multiset $P_p := \{(\sigma_J, \sigma'_J) \mid J \in \mathcal{J} \text{ bounded}\}$. Meanwhile, for every unbounded interval J with finite endpoint $v \in \mathbb{R}$ choose an element σ_J in $f^{-1}(v)$, then form the multiset $U_p := \{\sigma_J \mid J \in \mathcal{J} \text{ unbounded}\}$. □

Barcode templates get their name from the fact that they are an invariant of the ordering equivalence relation $\sim$:

Proposition 1.4.5. *If f, f' are ordering equivalent filter functions, then any barcode template of f is also a barcode template of f' and vice-versa.*

The proof, detailed hereafter, relies on the following elementary lemma.

Lemma 1.4.6. *Let $\mathbb{V}$ be a persistence module, and $h : \mathbb{R} \to \mathbb{R}$ be a continuous increasing function. Denote by $\mathbb{V}_h$ the shift of $\mathbb{V}$ by h, i.e for any $s \leqslant t$, $\mathbb{V}_{h,t} := \mathbb{V}_{h(t)}$ and $v^{\mathbb{V}_h}_{s,t} := v^{\mathbb{V}}_{h(s),h(t)}$. If $\mathbb{V}$ decomposes as $\mathbb{V} \cong \oplus_{J\in\mathcal{J}} \mathbb{I}_J$, then $\mathbb{V}_h \cong \oplus_{J\in\mathcal{J}} \mathbb{I}_{h^{-1}(J)}$.*

Proof. The operation that takes a persistence module to its shift by h is an endofunctor of **Pers** which commutes with direct sums. In particular it preserves isomorphisms. □

Proof of Proposition 1.4.5. Let f, f' be two ordering equivalent filter functions. Since $f \sim f'$, we have $f(\sigma) = f(\sigma') \Rightarrow f'(\sigma) = f'(\sigma')$ for any pair of simplices $\sigma, \sigma' \in K$. Therefore the map $h : f(\sigma) \in f(K) \mapsto f'(\sigma) \in f'(K)$ is well-defined. Furthermore, h is an increasing function and we extend it monotonously and continuously over all $\mathbb{R}$. Then, by the reparametrization Lemma 1.4.6, any barcode template of f is also a barcode template of f'. □

1.4.1 *Generic smoothness of the barcode valued map*

We now state our first significant results (one local and the other global) about the differentiability of the map B_p in the context of this section. Equipping $\mathbb{R}^K$ with the usual Euclidean norm, we assume that the parametrization F is of class C^r as a map $\mathcal{M} \to \mathbb{R}^K$. Under this hypothesis, we show that B_p is r-differentiable in the sense of Definition 1.3.3 on a generic (open and dense) subset of $\mathcal{M}$. The intuition behind these results is that, whenever the filter functions $F(\theta')$ are all ordering equivalent in a neighborhood of θ, we can pick a barcode template that is consistent across all filter functions $F(\theta')$ in this neighborhood (by Propositions 1.4.4 and 1.4.5) and the Equation (1.8) then behaves like a local coordinate system for B at θ.

Here is our local result:

Theorem 1.4.7 (Local discrete smoothness)**.** *Let $\theta \in \mathcal{M}$. Suppose the parametrization $F : \mathcal{M} \to \mathbb{R}^K$ is of class C^r ($r \geqslant 0$) on some open neighborhood U of θ, and that $F(\theta') \sim F(\theta)$ for all $\theta' \in U$. Then, B_p is r-differentiable at θ.*

Proof. Note that, as an open set, U is an open submanifold of $\mathcal{M}$ of same dimension. By Proposition 1.4.4, we can pick a barcode template (P_p, U_p) for $F(\theta)$. By Proposition 1.4.5, this barcode template is consistent for all $F(\theta')$ where $\theta' \in U$. Therefore, we can locally write:

$$\forall \theta' \in U,\ B_p(\theta') = \left\{(F(\theta')(\sigma), F(\theta')(\sigma'))\right\}_{(\sigma,\sigma') \in P_p} \cup \left\{(F(\theta')(\sigma), +\infty)\right\}_{\sigma \in U_p} \cup \Delta^\infty$$

which is a local coordinate system for B_p at θ. This local coordinate system is C^r because F itself is C^r over U. As a result, B_p is r-differentiable at θ, by Proposition 1.3.8. □

Corollary 1.4.8. *Let $\theta \in \mathcal{M}$. Suppose that the parametrization F is of class C^r ($r \geqslant 0$) on some open neighborhood of θ, and that the filter function $F(\theta)$ is injective. Then, B_p is r-differentiable at θ.*

Proof. For such a θ, all the quantities $F(\theta)(\sigma) - F(\theta)(\sigma')$ for $\sigma \neq \sigma' \in K$ are either strictly positive or strictly negative. Therefore, by continuity they keep their sign in an open neighborhood of θ, over which all filter functions are thus ordering equivalent. The result follows then from Theorem 1.4.7. □

Here is our global result:

Theorem 1.4.9 (Global discrete smoothness)**.** *Suppose the parametrization $F : \mathcal{M} \to \mathbb{R}^K$ is continuous over $\mathcal{M}$ and of class C^r ($r \geqslant 0$) on some open subset U of $\mathcal{M}$. Then, B_p is r-differentiable on the set $U \cap \tilde{\mathcal{M}}$, where*

$$\tilde{\mathcal{M}} := \left\{\theta \in \mathcal{M} \mid \exists \text{ open neighborhood } U_\theta \text{ of } \theta \text{ s.t. } F(\theta') \sim F(\theta) \text{ for all } \theta' \in U_\theta\right\}, \tag{1.9}$$

which is generic (i.e. open and dense) in $\mathcal{M}$. In particular, if F is C^r on some generic subset of $\mathcal{M}$ in the first place, then so is B_p (on some possibly smaller generic subset).

Proof. Observe that $\tilde{\mathcal{M}}$ is open in $\mathcal{M}$. As a consequence, for every $\theta \in U \cap \tilde{\mathcal{M}}$ there is some open neighborhood on which F is C^r and all the filter functions $F(\theta')$ are ordering equivalent, which by Theorem 1.4.7 implies that B_p is r-differentiable at θ. Thus, all that remains to be shown is that $\tilde{\mathcal{M}}$ is dense in $\mathcal{M}$, which is the subject of Lemma 1.4.10 below. □

Lemma 1.4.10. *If a parametrization $F : \mathcal{M} \to \mathbb{R}^K$ is continuous, then the set $\tilde{\mathcal{M}}$ (as defined in Eq. (1.9)) is dense in $\mathcal{M}$.*

Proof. Let $h : \mathcal{M} \to \mathbb{R}$ be a continuous function. Consider the boundary of the zero-level set $h^{-1}(0)$:

$$\partial h^{-1}(0) = \overline{h^{-1}(0)} \setminus (h^{-1}(0))^{\circ}.$$

Since h is continuous, $h^{-1}(0)$ is closed in $\mathcal{M}$, therefore $\partial h^{-1}(0)$ is closed with empty interior, i.e. its complement $(\partial h^{-1}(0))^c$ in $\mathcal{M}$ is open and dense.

Consider now the case of function $h_{\sigma,\sigma'} : \theta \in \mathcal{M} \mapsto F(\theta)(\sigma) - F(\theta)(\sigma') \in \mathbb{R}$ for some fixed simplices $\sigma \neq \sigma'$ of K. The map $h_{\sigma,\sigma'}$ is continuous by continuity of the parametrization F, therefore the previous paragraph implies that $(\partial h^{-1}_{\sigma,\sigma'}(0))^c$ is generic in $\mathcal{M}$. Hence, the finite intersection

$$\hat{\mathcal{M}} := \bigcap_{\sigma \neq \sigma' \in K} (\partial h^{-1}_{\sigma,\sigma'}(0))^c$$

is also generic in $\mathcal{M}$. We now show that $\hat{\mathcal{M}}$ is a subspace of $\tilde{\mathcal{M}}$.

Let $\theta \in \hat{\mathcal{M}}$ and $\sigma \neq \sigma' \in K$. If $h_{\sigma,\sigma'}(\theta) > 0$, then by continuity we have $h_{\sigma,\sigma'} > 0$ over some open neighborhood $V_{\sigma,\sigma'}$ of θ. Similarly if $h_{\sigma,\sigma'}(\theta) < 0$. And if $h_{\sigma,\sigma'}(\theta) = 0$, then, since $\theta \in \hat{\mathcal{M}}$, θ lies in the interior of the level set $h^{-1}_{\sigma,\sigma'}(0)$, and therefore there is also an open neighborhood $V_{\sigma,\sigma'}$ of θ over which $h_{\sigma,\sigma'} = 0$. Let V be the finite intersection $\bigcap_{\sigma \neq \sigma' \in K} V_{\sigma,\sigma'}$, which is open and non-empty in $\mathcal{M}$. For every $\sigma \neq \sigma' \in K$, the sign $F(\theta')(\sigma) - F(\theta')(\sigma')$ is constant over all $\theta' \in V$, where by sign we really distinguish between three possibilities: negative, positive, null. Therefore, the pre-order on the simplices of K induced by $F(\theta')$ is constant over the $\theta' \in V$. In other words, all the $F(\theta')$ are ordering equivalent. Therefore, $\theta \in \tilde{\mathcal{M}}$. Since this is true for any $\theta \in \hat{\mathcal{M}}$, we conclude that $\hat{\mathcal{M}} \subseteq \tilde{\mathcal{M}}$, and so the latter is also dense in $\mathcal{M}$. □

Example 1.4.11 (Height functions again)**.** Let us reconsider the scenario of Example 1.4.1. The parametrization F of height filters is C^0 on the

entire sphere $\mathbb{S}^{d-1}$. Moreover, F is smooth at every direction $\theta \in \mathbb{S}^{d-1}$ that is not orthogonal to some difference $v - v'$ of vertices $v \neq v' \in K_0$ in $\mathbb{R}^d$. The set U of such directions is generic in $\mathbb{S}^{d-1}$, therefore B_p is ∞-differentiable over the generic subset $U \cap \tilde{\mathbb{S}}^{d-1}$ by Theorem 1.4.9, with $\tilde{\mathbb{S}}^{d-1}$ defined as in Eq. (1.9). In fact, we have $U \cap \tilde{\mathbb{S}}^{d-1} = U$ in this case. Indeed, for any direction $\theta \in U$, the values of the height function h_θ at the vertices of K are pairwise distinct, and by continuity this remains true in a neighborhood of θ. The pre-order on the simplices of K induced by the height function is then constant over this neighborhood.

In Theorems 1.4.7 and 1.4.9, one cannot avoid the condition that filter functions are locally ordering equivalent. Indeed, in the next examples, we highlight that there is generally no hope for the barcode valued map B_p to be differentiable everywhere, even if the parametrization F is. This is because, essentially, the time of appearance of a simplex is a maximum of smooth functions, which can be non-smooth at a point where two functions achieve the maximum. The condition that the induced pre-order is locally constant around θ is only a sufficient condition though, because a maximum of two smooth functions can still be smooth at a point where the maximum is attained by the two functions. We provide a second example to illustrate this fact.

Example 1.4.12 (Singular parameter)**.** Let us consider the following geometric simplicial complex K on the real line:

That is, K has vertices $K_0 = \{a, b\}$ with respective coordinates $\{0, 1\}$, and edges $K_1 = \{ab\}$. Consider the parametrization that filters the complex according to the squared euclidean distance to a point, i.e $F : \theta \in \mathbb{R} \mapsto (\sigma \in K \mapsto \max_{x\in\sigma}(x - \theta)^2)$. The map B_0 is then essentially a real function that tracks the squared euclidean distance of the vertex closest to θ, specifically:

$$B_0(\theta) = \{(\min(\theta^2, (1-\theta)^2), +\infty)\} \cup \Delta^\infty.$$

Hence, B_0 is not differentiable at $\theta = \frac{1}{2}$ since $\frac{1}{2}$ is a singular point of the map $\theta \mapsto \min(\theta^2, (1-\theta)^2)$. Meanwhile, for $\theta < \frac{1}{2}$, we have $F(\theta)(a) < F(\theta)(b)$, whereas whenever $\theta > \frac{1}{2}$, we have $F(\theta)(a) > F(\theta)(b)$. In particular, the pre-order induced by the filter functions $F(\theta)$ is not constant around $\theta = \frac{1}{2}$, and so $\frac{1}{2} \notin \tilde{\mathbb{R}}$.

Example 1.4.13 (Only sufficient condition)**.** We remove the edge ab from the geometric complex K in the previous example, and we see the points a and b as lying on the x-axis of $\mathbb{R}^2$. Consider the parametrization of height filters $F : \theta \in \mathbb{S}^1 \mapsto (\sigma \in K \mapsto \max_{x\in\sigma}\langle\theta, x\rangle)$. The map B_p is then trivial for each degree p except 0, where it writes as follows:

$$B_0(\theta) = \{(\langle\theta, a\rangle, +\infty), (\langle\theta, b\rangle, +\infty)\} \cup \Delta = \{(0, +\infty), (\langle\theta, (1,0)\rangle, +\infty)\} \cup \Delta^\infty.$$

We see that we have a valid local coordinate system given by the two smooth maps $\theta \mapsto 0$ and $\theta \mapsto \langle \theta, (0,1) \rangle$, so the map B_0 is ∞-differentiable everywhere on $\mathbb{S}^1$ by Proposition 1.3.8. Meanwhile, we have $F(\theta)(a) < F(\theta)(b)$ whenever $\langle \theta, (1,0) \rangle > 0$, and $F(\theta)(a) > F(\theta)(b)$ whenever $\langle \theta, (1,0) \rangle < 0$, therefore the pre-order induced by the filter functions $F(\theta)$ is not constant around $\theta = (0,1)$ and $v = (0,-1)$, hence $(0,1), (0,-1) \notin \tilde{\mathbb{R}}$.

1.4.2 *Differential of the barcode valued map*

Given a continuous parametrization $F : \mathcal{M} \to \mathbb{R}^K$ of class C^1 on some open set $U \subseteq \mathcal{M}$, Theorem 1.4.9 guarantees that a barcode template, through Equation (1.8), provides a C^1 local coordinate system for B_p around each point $\theta \in U \cap \tilde{\mathcal{M}}$. In turn, by Proposition 1.3.8, any arbitrary ordering on the functions of this local coordinate system induces a C^1 local lift of B_p. Hence we have the following formula for the corresponding differential:

Proposition 1.4.14. *Given $\theta \in U \cap \tilde{\mathcal{M}}$ and a barcode template (P_p, U_p) of $F(\theta)$, for any choice of ordering $(\sigma_1, \sigma_1'), \cdots, (\sigma_m, \sigma_m'), \tau_1, \cdots, \tau_n$ of (P_p, U_p), the map*

$$\tilde{B}_p : \theta' \mapsto \left[(F(\theta')(\sigma_i), F(\theta')(\sigma_i'))_{i=1}^m, (F(\theta')(\sigma_j))_{j=1}^n \right]$$

is a local C^1 lift of B_p around θ, and the corresponding differential for B_p at θ is:

$$d_{\theta, \tilde{B}_p} B_p(.) = \left[(d_\theta F(\cdot)(\sigma_i), d_\theta F(\cdot)(\sigma_i'))_{i=1}^m, (d_\theta F(\cdot)(\tau_j))_{j=1}^n \right].$$

Remark 1.4.15 (Algorithm for computing derivatives)**.** Suppose we are given a parametrization F whose differential we can compute. Let $\theta \in \mathcal{M}$. If the barcode of $F(\theta)$ is given to us, then the proof of Proposition 1.4.4 provides an algorithm to build a barcode template (P_p, U_p) for $F(\theta)$. If the barcode of $F(\theta)$ is not given in the first place, then the matrix reduction algorithm for computing persistence[8] outputs both the barcode and a barcode template. In both scenarios, Proposition 1.4.14 gives a formula to compute a differential of B_p at θ from the barcode template (P_p, U_p). The optimization pipelines mentioned in the introduction [BGGSSG20a, CNBW19, GHO16, HGR+20a, PSO18] apply this strategy to compute differentials.

[8] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002; and Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005

1.4.3 *Directional differentiability of the barcode valued map along strata*

In this section we define directional derivatives for the barcode valued map $B_p : \mathcal{M} \to \mathit{Bar}$ at points where it may not be differentiable in the sense of Definition 1.3.3. For this we stratify the parameter space $\mathcal{M}$ in

such a way that B_p is differentiable on the top-dimensional strata, then we define its derivatives on lower-dimensional strata via directional lifts. Intuitively, the strata in $\mathcal{M}$ are prescribed by the ordering equivalence classes in $\mathbb{R}^K$, as we know from Theorem 1.4.7 that the pre-order on simplices plays a key role in the differentiability of B_p.

Formally, consider the stratification of $\mathbb{R}^K$ formed by the collection $\Omega(\mathbb{R}^K)$ of ordering equivalence classes. This is a Whitney stratification, obtained by cutting $\mathbb{R}^K$ with the hyperplanes $\{f(\sigma) = f(\sigma')\}$ for varying simplices $\sigma \neq \sigma' \in K$. We look for stratifications of $\mathcal{M}$ that make the parametrization F weakly stratified (in the sense of Definition 1.2.22) and smooth on each stratum. Here are typical scenarios where such stratifications exist:

Proposition 1.4.16. *Let $F : \mathcal{M} \to \mathbb{R}^K$ be a continuous parametrization. Suppose that, either*

(i) $\mathcal{M}$ is a semi-algebraic set in $\mathbb{R}^N$ and F is a semi-algrebraic map, or

(ii) $\mathcal{M}$ is a compact subanalytic set in a real analytic manifold and F is a subanalytic map.

Then, there is a Whitney stratification of $\mathcal{M}$, made of semi-algebraic (resp. subanalytic) strata, such that F is weakly stratified with C^∞ restrictions to each stratum.

Proof. This is Section I.1.7 of [GM88], after observing that the stratification $\Omega(\mathbb{R}^K)$ is made of semi-algebraic strata. □

Mark Goresky and Robert MacPherson. *Stratified Morse theory*. Ergebnisse der Mathematik, volume 14. Springer-Verlag, Berlin, 1988

Example 1.4.17. We consider the parametrization F of height filters on the sphere $\mathbb{S}^{d-1}$ from Example 1.4.11. By Proposition 1.4.16, there is a stratification of $\mathbb{S}^{d-1}$ that makes F weakly stratified and C^∞ on each stratum. To be more specific, such a stratification is obtained by taking the pre-images[9] of the strata of $\Omega(\mathbb{R}^K)$ via F. Figure 1.1 illustrates the result in the case $d = 3$, where the obtained stratification of $\mathbb{S}^2$ is made of an arrangement of great circles, each circle being the pre-image of a set $\{F(\theta)(v) = F(\theta)(v')\}$ for vertices $v \neq v'$.

[9] This is called the *pull-back stratification*. In fact, for any smooth map $F : \mathcal{M} \to \mathbb{R}^K$ that is transverse with respect to $\Omega(\mathbb{R}^K)$ and to any stratification of $\mathcal{M}$ (e.g. the trivial one), the pull-back of $\Omega(\mathbb{R}^K)$ via F makes the latter weakly stratified and C^∞ on each stratum—see e.g. the section I.1.3 in [GM88].

Once a stratification $\mathcal{S}_\mathcal{M}$ of $\mathcal{M}$ is given, we can introduce a notion of derivative for B_p at $\theta \in \mathcal{M}$ in the direction of an *incident* stratum $\mathcal{M}'$, i.e. a stratum whose closure in $\mathcal{M}$ contains θ.

Definition 1.4.18. Let $B : \mathcal{M} \to \mathit{Bar}$ be a map defined on a stratified space $(\mathcal{M}, \mathcal{S}_\mathcal{M})$. Let $\theta \in \mathcal{M}$, and let $\mathcal{M}' \in \mathcal{S}_\mathcal{M}$ be a stratum incident to θ. The map B is *r-differentiable* at θ *along* $\mathcal{M}'$ if there is an open neighborhood U of θ in $\mathcal{M}$ and a C^r map $\tilde{B} : U \to \mathbb{R}^{2m} \times \mathbb{R}^n$ for some integers m, n such that $B = Q_{m,n} \circ \tilde{B}$ on $U \cap \mathcal{M}'$. The differential $d_\theta \tilde{B}$ is called a *directional derivative* of B at θ along $\mathcal{M}'$.

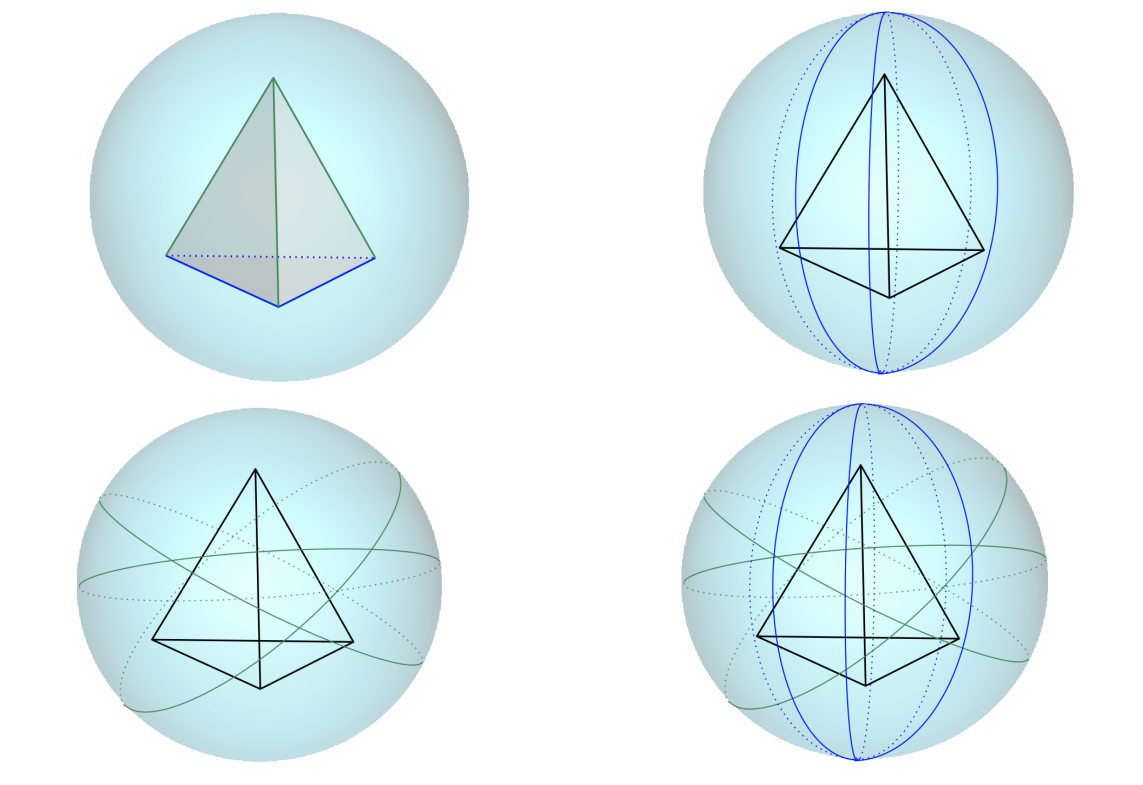


Figure 1.1: The regular tetrahedron K embedded in $\mathbb{R}^3$ (top left) induces the stratification of S^2 for the parametrization of height filters (bottom right). The intersections of great circles are 0-dimensional strata, the parts of the great circles that do not intersect with each other are 1-dimensional strata, the rest of S^2 forms the 2-dimensional strata. Each great circle corresponds to unit vectors that are orthogonal to a given edge of the simplicial complex. The edges joining the vertices in the base of the tetrahedron produce the blue circles (top right), while the edges joining the apex of the tetrahedron to the base face produce the green circles (bottom left).

This definition agrees with the notions of r-differentiability and derivatives introduced in Section 1.3 when $\mathcal{M}'$ contains an open neighborhood around θ, i.e. for θ located in a top-dimensional stratum $\mathcal{M}'$. When θ is located in some lower-dimensional stratum, it admits finitely many incident strata $\mathcal{M}'$ (possibly not top-dimensional), each one of which yields a specific directional derivative at θ. The definition of each derivative involves a local C^r lift $\tilde{B}$ of B near θ in $\mathcal{M}'$. This lift is required to extend smoothly over an open neighborhood U in $\mathcal{M}$, to ensure that $\tilde{B}$ and its derivatives have well-defined limits at θ.

Theorem 1.4.19 (Discrete smoothness along strata)**.** *Let $r \in \mathbb{N}$ and $F : \mathcal{M} \to \mathbb{R}^K$. Suppose $\mathcal{S}_{\mathcal{M}}$ is a Whitney stratification of $\mathcal{M}$ such that:*

(i) *F is a weakly stratified map with respect to $\mathcal{S}_{\mathcal{M}}$ and $\Omega(\mathbb{R}^K)$, and*

(ii) *the restriction of F to each stratum of $\mathcal{S}_{\mathcal{M}}$ is C^r, and*

(iii) *for every $\theta \in \mathcal{M}$ and every incident stratum $\mathcal{M}' \in \mathcal{S}_{\mathcal{M}}$, there is an open neighborhood U of θ in $\mathcal{M}$ such that $F_{|\mathcal{M}' \cap U}$ extends to a C^r map $U \to \mathbb{R}^K$.*

Then, at every $\theta \in \mathcal{M}$, the barcode valued map $B_p : \mathcal{M} \to Bar$ is r-differentiable along each stratum incident to θ. In particular, B_p is r-differentiable in the sense of Definition 1.3.3 inside each top-dimensional stratum.

Proof. Let $\theta \in \mathcal{M}$ and $\mathcal{M}'$ a stratum incident to θ. By **(i)**, combined with Propositions 1.4.4 and 1.4.5, there exists a barcode template (P_p, U_p) that is consistent across all $F(\theta')$ for $\theta' \in \mathcal{M}'$. Therefore, for all $\theta' \in \mathcal{M}'$:

$$B_p(\theta') = \left\{(F(\theta')(\sigma), F(\theta')(\sigma'))\right\}_{(\sigma,\sigma')\in P_p} \cup \left\{(F(\theta')(\sigma), +\infty)\right\}_{\sigma\in U_p} \cup \Delta^\infty, \tag{1.10}$$

which by **(ii)** provides a C^r local coordinate system for $B_{p|\mathcal{M}'}$. Then by Proposition 1.3.8, there is a C^r lift of $B_{p|\mathcal{M}'}$, whose coordinate functions are of the form $\theta' \mapsto F(\theta')(\sigma)$. Using **(iii)**, we extend each coordinate function of this lift (hence the lift itself) to an open neighborhood U of θ in $\mathcal{M}$. □

Combining Proposition 1.4.16 with Theorem 1.4.19 yields the following:

Corollary 1.4.20. *Under the hypotheses of Proposition 1.4.16, there is a Whitney stratification of $\mathcal{M}$, made of semi-algebraic (resp. subanalytic) strata, such that B_p is ∞-differentiable on the top-dimensional strata (whose union is generic in $\mathcal{M}$). If furthermore F is globally C^r, then B_p is everywhere r-differentiable along incident strata.*

Example 1.4.21. Consider again the setup of Example 1.4.12. We stratify $\mathbb{R}$ by the point $\{\frac{1}{2}\}$ and the half-lines $(-\infty; \frac{1}{2})$ and $(\frac{1}{2}; +\infty)$. The parametrization F is C^∞ and sends strata into strata, therefore by Theorem 1.4.19 the barcode valued map B_0 admits directional derivatives everywhere on $\mathbb{R}$. More precisely, recall that we have a lift $\tilde{B_0} : \theta \mapsto \min(\theta^2, (1-\theta)^2)$, which is smooth in the top-dimensional strata, while at $\theta = \frac{1}{2}$ it admits directional derivatives along the two half-lines, whose values are 1 and -1 respectively and thus do not agree.

Example 1.4.22. Consider again the stratification $\mathcal{S}_{\mathbb{S}^{d-1}}$ by the great circles of the parameter space $\mathbb{S}^{d-1}$ associated to the parametrization of height filters (Example 1.4.17). By Corollary 1.4.20, we know that there exists a refinement $\mathcal{S}'_{\mathbb{S}^{d-1}}$ of $\mathcal{S}_{\mathbb{S}^{d-1}}$ such that B_p admits directional derivatives along incident strata of $\mathcal{S}'_{\mathbb{S}^{d-1}}$ at every point $\theta \in \mathbb{S}^{d-1}$. In fact, we can even take $\mathcal{S}'_{\mathbb{S}^{d-1}}$ to be $\mathcal{S}_{\mathbb{S}^{d-1}}$ itself. Indeed, all the directions in a given stratum $\mathcal{M}' \in \mathcal{S}_{\mathbb{S}^{d-1}}$ induce the same pre-order on the simplices of K, therefore

- the restriction $F_{|\mathcal{M}'}$ is valued in a stratum of $\Omega(\mathbb{R}^K)$, and
- for every simplex $\sigma \in K$, there is a vertex $\bar{v}(\sigma)$ such that $F_{|\mathcal{M}'}(.)(\sigma) = \langle ., \bar{v}(\sigma)\rangle$.

Consequently, the assumptions of Theorem 1.4.19 hold, and the barcode valued map B_p admits directional derivatives along incident strata of $\mathcal{S}_{\mathbb{S}^{d-1}}$ at every point $\theta \in \mathbb{S}^{d-1}$.

1.4.4 *The barcode valued map as a permutation map*

In this section, we work out a global lift of the barcode valued map, which restricts nicely to each stratum of a stratification of $\mathcal{M}$. To do so, we first focus on the map PH which, given a filter function $f \in \mathbb{R}^K$ on a fixed simplicial complex K of dimension d, returns the vector of all its barcodes $(\mathrm{PH}_p(f))_{p=0}^d$. We observe that PH admits a global Euclidean lift, and furthermore, that this lift is essentially a permutation map on each stratum of $\Omega(\mathbb{R}^K)$. Throughout, we fix an ordering of the simplices of K, so that the canonical basis of $\mathbb{R}^K$ turns into a basis of $\mathbb{R}^{\#K}$, and we let $\phi : \mathbb{R}^K \to \mathbb{R}^{\#K}$ be the corresponding isomorphism.

Proposition 1.4.23. *There exist integers m_p, n_p for $0 \leqslant p \leqslant d$ such that $\sum_{p=0}^d (2m_p + n_p) = \#K$, and a map* $\mathrm{Perm} : \mathbb{R}^K \to \prod_{p=0}^d \mathbb{R}^{2m_p} \times \mathbb{R}^{n_p} \cong \mathbb{R}^{\#K}$ *whose restriction* $\mathrm{Perm}_{|S}$ *to each ordering equivalence class $S \in \Omega(\mathbb{R}^K)$ is a permutation matrix, and such that the following diagram commutes:*[10]

[10] Strictly speaking, like PH, this diagram applies only to the set of filter functions in $\mathbb{R}^K$.

$$\begin{array}{ccc} \mathbb{R}^{\#K} & \xrightarrow{\quad\mathrm{Perm}\quad} & \prod_{p=0}^d \mathbb{R}^{2m_p} \times \mathbb{R}^{n_p} \\ \phi \uparrow & & \downarrow Q := \prod_{p=0}^d Q_{m_p,n_p} \\ \mathbb{R}^K & \xrightarrow{\quad\mathrm{PH}\quad} & Bar^{d+1} \end{array} \tag{1.11}$$

For simplicity, from now on we identify $f \in \mathbb{R}^K$ with its image in $\mathbb{R}^{\#K}$ without explicitly mentioning the map ϕ.

Proof. Given a filter function $f \in \mathbb{R}^K$, we define a total barcode template (P, U) for f to be the data of $d+1$ barcode templates (P_p, U_p) for f in each homology degree, such that each simplex of K appears exactly once, in a unique P_p or U_p. We further require that the pairs (σ, σ') appearing in P_p consist of a p-dimensional simplex σ and a $(p+1)$-dimensional simplex σ', while the unpaired simplices appearing in U_p must be p-dimensional. A simplex σ is then labelled *positive* if it appears as the first component of a pair in some P_p or U_p, and *negative* otherwise.

Note that total barcode templates always exist, by an argument similar to (yet somewhat more involved than) the one used in the proof of Proposition 1.4.4. Alternatively, applying the matrix reduction algorithm for computing persistence[11] to the sublevel-sets filtration of f produces a total barcode template. By Proposition 1.4.5, total barcode templates are invariant under ordering equivalences. We therefore fix a unique total barcode template $(P(S), U(S))$ per ordering equivalence class $S \in \Omega(\mathbb{R}^K)$ (there are only finitely many such classes), and we denote by $m_p(S) := \#P_p(S)$, $n_p(S) := \#U_p(S)$ their sizes in each homology degree p.

[11] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002; and Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005

Since the barcode templates $(P(S), U(S))$ are total, we have $\sum_{p=0}^{d}(2m_p(S) + n_p(S)) = \#K$. Besides, since the number of infinite intervals in the barcode of a filter function is given by the Betti numbers of the simplicial complex K, an easy induction on the homology degree shows that the number of positive (resp. negative) simplices in each homology degree is independent of the choice of filter function and of total barcode template. Therefore, the integers $m_p(S), n_p(S)$ do not depend on the stratum S.

For each stratum $S \in \Omega(\mathbb{R}^K)$ and homology degree p, we pick arbitrary orderings $(\sigma_{k,S}, \sigma'_{k,S})_{k=0}^{m_p}$ of $P_p(S)$ and $(\tau_{k,S})_{k=0}^{n_p}$ of $U_p(S)$. Any filter function $f \in S$ admits $(P(S), U(S))$ as total barcode template, therefore we get that $\mathrm{PH}_p(f) = Q_{m_p,n_p}((f(\sigma_{k,S}), f(\sigma'_{k,S}))_{k=0}^{m_p}, (f(\tau_{k,S}))_{k=0}^{n_p})$ in every homology degree p. We simply set $\mathrm{Perm}(f) := [(f(\sigma_{k,S}), f(\sigma'_{k,S}))_{k=0}^{m_p}, (f(\tau_{k,S}))_{k=0}^{n_p}]_{p=0}^{d} \in \prod_{p=0}^{d} \mathbb{R}^{2m_p} \times \mathbb{R}^{n_p}$, which ensures the commutativity of (1.11). Since each simplex of K appears exactly once in $(P(S), U(S))$, the vector $\mathrm{Perm}(f)$ is a re-ordering of the coordinates of f (i.e. of its values on the simplices) and therefore $\mathrm{Perm}_{|S}$ is a permutation matrix. □

We now turn to the parametrized barcode valued map

$$B : \theta \in \mathcal{M} \underset{F}{\longrightarrow} F(\theta) \in \mathbb{R}^K \underset{\mathrm{PH}}{\longrightarrow} [\mathrm{PH}_p(F(\theta))]_{p=0}^{d} \in Bar^{d+1}$$

determined by a parametrization $F : \mathcal{M} \to \mathbb{R}^K$ of filter functions. We show that if $\mathcal{M}$ admits a Whitney stratification $\mathcal{S}_{\mathcal{M}}$ satisfying the assumptions of Theorem 1.4.19, then B admits a global lift $\tilde{B}$ that acts as a permutation of F-values on each stratum.

Corollary 1.4.24. *Using the same notations as in Proposition 1.4.23, the map*

$$\tilde{B} : \theta \in \mathcal{M} \longmapsto \mathrm{Perm}(F(\theta)) \in \prod_{p=0}^{d} \mathbb{R}^{2m_p} \times \mathbb{R}^{n_p}$$

is a global lift of B, i.e $Q \circ \tilde{B} = B$ everywhere on $\mathcal{M}$. If moreover $\mathcal{M}$ admits a Whitney stratification $\mathcal{S}_{\mathcal{M}}$ satisfying the assumptions of Theorem 1.4.19, then $\tilde{B} = \mathrm{Perm}_{\mathcal{M}'} \circ F$ for some permutation matrix $\mathrm{Perm}_{\mathcal{M}'}$ over each stratum $\mathcal{M}' \in \mathcal{S}_{\mathcal{M}}$. Consequently, B is r-differentiable along incident strata everywhere on $\mathcal{M}$, with directional derivatives given by the ones of $\tilde{B}$.

The last part of the statement expresses the fact that directional derivatives of B are simply given by permuting the directional derivatives of the coordinate functions of F.

Proof. The first part of the statement is a straight consequence of Proposition 1.4.23. Let $\mathcal{S}_{\mathcal{M}}$ be a stratification satisfying the assumptions of Theorem 1.4.19. As F is weakly stratified with respect to $\mathcal{S}_{\mathcal{M}}$ and $\Omega(\mathbb{R}^K)$, it sends strata into strata and therefore by Proposition 1.4.23

we have $\tilde{B} = \mathrm{Perm}_{\mathcal{M}'} \circ F$ for some permutation matrix $\mathrm{Perm}_{\mathcal{M}'}$ over each stratum $\mathcal{M}' \in \mathcal{S}_{\mathcal{M}}$. Then, since F admits local smooth extensions over each stratum $\mathcal{M}'$ of $\mathcal{S}_{\mathcal{M}}$, so do its coordinate functions and in turn so does $\tilde{B} = \mathrm{Perm}_{\mathcal{M}'} \circ F$. These local extensions of $\tilde{B}$ yield directional derivatives for B along incident strata. □

Remark 1.4.25. Recall that the map Perm is a linear map when restricted to the strata of $\Omega(\mathbb{R}^K)$, which are simply polyhedra in $\mathbb{R}^K$. Therefore, if $\mathcal{M}$ is a semi-algebraic set (resp. subanalytic set or definable set in an o-minimal structure) and F is a semi-algebraic (resp. subanalytic or definable) map, then the global lift $\tilde{B} = \mathrm{Perm} \circ F$ of Corollary 1.4.24 is itself a semi-algebraic (resp. subanalytic or definable) map. Thus we recover Proposition 3.2 and Corollary 3.3 of [CCG+21]. Meanwhile, the differentiability of $\tilde{B}$ on top-dimensional strata (as per Corollary 1.4.20) recovers their Proposition 3.4.

Mathieu Carriere, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hariprasad Kannan, and Yuhei Umeda. Optimizing persistent homology based functions. In *International Conference on Machine Learning*, pages 1294–1303. PMLR, 2021

We conclude this section with a side result whose proof relies on Proposition 1.4.23. This result states that PH is locally an isometry on top-dimensional strata of $\Omega(\mathbb{R}^K)$. It involves the distance $d_0(f)$ of any filter function $f \in \mathbb{R}^K$ to the union of strata of $\Omega(\mathbb{R}^K)$ of codimension at least 1:

$$d_0(f) = \frac{1}{2} \min_{\sigma \neq \sigma'} |f(\sigma) - f(\sigma')|.$$

Proposition 1.4.26. *Let $f, g \in \mathbb{R}^K$ be two filter functions that are located in the closure of a common top-dimensional stratum $S \in \Omega(\mathbb{R}^K)$. Then:*

$$\max_{0 \leqslant p \leqslant d} d_\infty(\mathrm{PH}_p(f), \mathrm{PH}_p(g)) \geqslant \min(\|f - g\|_\infty, \max(d_0(f), d_0(g))). \quad (1.12)$$

In particular, for any filter function $f \in \mathbb{R}^K$ located in a top-dimensional stratum, the map PH *is a local isometry in a closed ball of radius $d_0(f)$ around f, specifically:*

$$\forall g \in \mathbb{R}^K,\ \|f - g\|_\infty \leqslant d_0(f) \implies \max_{0 \leqslant p \leqslant d} d_\infty(\mathrm{PH}_p(f), \mathrm{PH}_p(g)) = \|f - g\|_\infty \quad (1.13)$$

$$\forall g, h \in \mathbb{R}^K,\ \max(\|f - g\|_\infty, \|f - h\|_\infty) \leqslant \frac{d_0(f)}{3} \implies \max_{0 \leqslant p \leqslant d} d_\infty(\mathrm{PH}_p(g), \mathrm{PH}_p(h)) = \|g - h\|_\infty. \quad (1.14)$$

Proof. We have several persistence diagrams to compare, so we first simplify the problem as follows. Given two vectors $D = (D_0, ..., D_d) \in Bar^{d+1}$ and $D' = (D'_0, ..., D'_d)^{d+1}$ of $d+1$ barcodes, let $\Gamma(D, D')$ be the set of *multi-matchings* between D and D', where a multi-matching is a bijection $\gamma : \bigsqcup_{p=0}^{d} D_i \to \bigsqcup_{p=0}^{d} D'_i$ such that $\gamma(D_p) = D'_p$ for all $0 \leqslant p \leqslant d$. The notions of cost $c(\gamma)$ and optimality are the same as for

ordinary matchings. Specifically, for an optimal γ in $\Gamma(\mathrm{PH}(f), \mathrm{PH}(g))$, we have $c(\gamma) = \max_{0 \leqslant p \leqslant d} d_\infty(\mathrm{PH}_p(f), \mathrm{PH}_p(g))$.

We fix an ordering $\sigma_1 < ... < \sigma_{\#K}$ of the simplices of K, which yields an isomorphism $\phi : \mathbb{R}^K \to \mathbb{R}^{\#K}$. We denote by f_i the i-th component of f through this isomorphism, i.e $f_i = f(\sigma_i)$. Let us assume that $f, g \in \mathbb{R}^K$ are two filter functions in a common top dimensional stratum S. If we can prove (1.12) in this case then it will hold for f, g in the closure of S by a continuity argument. Since f and g are both in S, they induce the same strict order on the simplices, and without loss of generality we can assume that $f_1 < ... < f_{\#K}$ and $g_1 < ... < g_{\#K}$. By Proposition 1.4.23, we can write $\mathrm{PH}(f) = Q(P(S)f)$ and $\mathrm{PH}(g) = Q(P(S)g)$ for a fixed permutation matrix $P(S)$, which implies that:

$$\forall 1 \leqslant i \neq j \leqslant \#K, \forall 0 \leqslant p \leqslant d,\ (f_i, f_j) \in \mathrm{PH}_p(f) \Leftrightarrow (g_i, g_j) \in \mathrm{PH}_p(g) \tag{1.15}$$

$$\forall 1 \leqslant k \leqslant \#K, \forall 0 \leqslant p \leqslant d,\ (f_k, +\infty) \in \mathrm{PH}_p(f) \Leftrightarrow (g_k, +\infty) \in \mathrm{PH}_p(g)$$

Let $\gamma \in \Gamma(\mathrm{PH}(f), \mathrm{PH}(g))$ be optimal. Consider the case where γ sends an off-diagonal point (b, d) of $\mathrm{PH}(f)$ onto the diagonal Δ. As (b, d) is of the form (f_i, f_j) (or $(f_i, +\infty)$), this implies that $c(\gamma) \geqslant \frac{|f_i - f_j|}{2} \geqslant d_0(f)$. In addition, $\mathrm{PH}(f)$ and $\mathrm{PH}(g)$ have exactly the same number of bounded and unbounded intervals in each degree, which implies that there exists an off-diagonal interval (b', d') of $\mathrm{PH}(g)$ which has pre-image in the diagonal Δ. Again, (b', d') must be of the form (g_k, g_l) (or $(g_k, +\infty)$), so that $c(\gamma) \geqslant \frac{|g_k - g_l|}{2} \geqslant d_0(g)$. Therefore, $c(\gamma) \geqslant \max(d_0(f), d_0(g))$ and we are done.

We now treat the case where all off-diagonal intervals are sent to off-diagonal intervals by γ. We denote by $O(f, g) \subset \Gamma(\mathrm{PH}(f), \mathrm{PH}(g))$ the set of multi-matchings that send off-diagonal intervals to off-diagonal intervals. By the decomposition $\mathrm{PH}(f) = Q(P(S)f)$ (resp. $\mathrm{PH}(g) = Q(P(S)g)$) and from the fact that no two values of f (resp. of g) are equal, the bounded end-points of off-diagonal intervals of $\mathrm{PH}(f)$ (resp. $\mathrm{PH}(g)$) are in bijection with the set $\{f_1, ..., f_{\#K}\}$ (resp. $\{g_1, ..., g_{\#K}\}$). Therefore, any multi-matching $\nu \in O(f, g)$ induces a permutation $\pi(\nu)$ of $\{1, ..., \#K\}$. Let us denote by $c(\pi) := \max_i(|f_i - g_{\pi(i)}|)$ the *cost* of a permutation π of $\{1, ..., \#K\}$. In this formulation, we have:

$$\max_{0 \leqslant p \leqslant d} d_\infty(\mathrm{PH}_p(f), \mathrm{PH}_p(g)) = c(\gamma) = \min_{\nu \in O(f,g)} c(\nu) = \min_{\nu \in O(f,g)} c(\pi(\nu)) \tag{1.16}$$

Consider the following relaxed optimization problem, in which the pairing of coordinates in (1.15) is ignored:

$$\min_{\pi \text{ permutation of } \{1,...,\#K\}} c(\pi) \tag{1.17}$$

From the fact that $f_1 < ... < f_{\#K}$ and $g_1 < ... < g_{\#K}$, the monotonicity of the optimal transport map[12] for the ∞-Wasserstein distance in $\mathbb{R}$

[12] Julien Rabin, Gabriel Peyré, Julie Delon, and Marc Bernot. Wasserstein barycenter and its application to texture mixing. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 435–446. Springer, 2011

guarantees that $\pi = \mathrm{Id}$ is a minimizer[13] of (1.17). Therefore,

$$\max_{0\leqslant p\leqslant d} d_\infty(\mathrm{PH}_p(f),\mathrm{PH}_p(g)) = \min_{\nu\in O(f,g)} c(\nu) \geqslant \min_{\pi \text{ permutation}} c(\pi) = c(\mathrm{Id}) = \|f-g\|_\infty$$

and we are done.

We now address the second part of the proposition. Let $f \in \mathbb{R}^K$ be a filter function. By the stability Theorem 1.2.12, showing that Eq. (1.13) holds amounts to showing that $\max_{0\leqslant p\leqslant d} d_\infty(\mathrm{PH}_p(f),\mathrm{PH}_p(g)) \geqslant \|f-g\|_\infty$. We denote by S the top dimensional stratum S that contains f, and let $g \in \mathbb{R}^K$ be another filter such that $\|f-g\|_\infty \leqslant d_0(f)$. This implies that g is also in the (closure of the) stratum S. We can then apply (1.12), and since by assumption $\|f-g\|_\infty \leqslant d_0(f) \leqslant \max(d_0(f),d_0(g))$, we have the desired result.

[13] Indeed, for a permutation π, let $\mathrm{inv}(\pi)$ denote its number of inversions. Let π be a minimizer (1.17) with minimal $\mathrm{inv}(\pi)$. Assuming by contradiction that $\mathrm{inv}(\pi) > 0$, there exist $i < j$ such that $\pi(i) > \pi(j)$. Let τ be the transposition that swaps $\pi(i)$ and $\pi(j)$. Since $f_1 < ... < f_{\#K}$ and $g_1 < ... < g_{\#K}$, a simple case analysis shows that $c(\tau\circ\pi) \leqslant c(\pi)$, thus raising a contradiction with the minimality of $\mathrm{inv}(\pi)$. Therefore $\pi = \mathrm{Id}$ is a minimizer of (1.17).

Using similar arguments, we finally prove Eq. (1.14). Let $f \in \mathbb{R}^K$ be a filter function in some top dimensional stratum S, and $g,h \in \mathbb{R}^K$ be such that $\|f-g\| \leqslant \frac{d_0(f)}{3}$ and $\|f-h\| \leqslant \frac{d_0(f)}{3}$. By the stability Theorem 1.2.12, showing that Eq. (1.14) holds amounts to showing that $\max_{0\leqslant p\leqslant d} d_\infty(\mathrm{PH}_p(g),\mathrm{PH}_p(h)) \geqslant \|g-h\|_\infty$. For every $i \neq j \in \{1,...,\#K\}$,

$$|g_i-g_j| = |(f_i-f_j)-[(f_i-g_i)+(g_j-f_j)]| \geqslant |f_i-f_j| - 2\|f-g\|_\infty \geqslant 2d_0(f) - 2\|f-g\|_\infty \geqslant \frac{4d_0(f)}{3},$$

so that $d_0(g) \geqslant \frac{2d_0(f)}{3}$. Similarly, $d_0(h) \geqslant \frac{2d_0(f)}{3}$. Meanwhile,

$$\|g-h\|_\infty \leqslant \|f-g\|_\infty + \|f-h\|_\infty \leqslant \frac{2d_0(f)}{3}.$$

Therefore, $\|g-h\|_\infty \leqslant \max(d_0(g),d_0(h))$, and since both g,h are in (the closure of) S, we conclude by using Eq. (1.12). □

1.5 *Application to common simplicial filtrations*

In this section we leverage Theorems 1.4.7 and 1.4.9 in the case of a few important classes of parametrizations of filter functions on a simplicial complex K of dimension d. In each case, we derive a characterization of the parameter values where B_p is differentiable, and whenever possible we provide an explicit differential of B_p using Proposition 1.4.14. In the following we fix a homology degree $0 \leqslant p \leqslant d$.

1.5.1 *Lower star filtrations*

Parametrizations of lower star filtrations are involved in most practical scenarios [BGGSSG20a, CNBW19, GNDS20a, HGR+20a, PSO18], here we provide a common analysis of their differentiability.

Definition 1.5.1. Given a function $f : K_0 \to \mathbb{R}$ defined on the vertices of K, we extend it to each simplex σ of K by its highest value on the vertices of σ. The sub-level sets of this function together form the *lower-star filtration* of K induced by f.

One interest of lower-star filtrations is that any parametrization $\mathcal{M} \to \mathbb{R}^{K_0}$ on the vertex set of K induces a valid parametrization $\mathcal{M} \to \mathbb{R}^K$ on K itself. Sufficient conditions for the differentiablity of such parametrizations are easy to work out thanks to the following observation:

Proposition 1.5.2. *Let $F_0 : \mathcal{M} \to \mathbb{R}^{K_0}$ be a C^r parametrization of filter functions on the vertices of K. Then, the induced parametrization $F : \mathcal{M} \to \mathbb{R}^K$ is C^r at each $\theta \notin \mathrm{Sing}(F_0)$, where $\mathrm{Sing}(F_0)$ is the boundary of the set:*

$$\{\theta \in \mathcal{M},\ \exists (v, v') \in K_0, F_0(\theta)(v) = F_0(\theta)(v')\}.$$

Specifically, for every $\theta \notin \mathrm{Sing}(F_0)$, letting

$$\bar{v} : \sigma \in K \mapsto \mathrm{argmax}_{v \text{ vertex in } \sigma} F_0(\theta)(v) \in K_0$$

by breaking ties wherever necessary, there is an open neighborhood U of θ such that $F(\theta')(\sigma) = F_0(\theta')(\bar{v}(\sigma))$ for every $\theta' \in U$ and $\sigma \in K$, from which follows that F is C^r at θ.

Proof. The continuity of F comes from the continuity of F_0 and of the max function. If $\theta \in \mathcal{M} \backslash \mathrm{Sing}(F_0)$, then the pre-order on K_0 induced by $F_0(.)$ is constant in an open neighborhood U of θ. We want to check that F is C^r at θ, i.e. that all maps $\theta' \mapsto F(\theta')(\sigma)$ are C^r at θ, for a fixed simplex $\sigma \in K$. For σ a vertex of K, this is true by assumption because $F(.)(\sigma) = F_0(.)(\sigma)$. For an arbitrary simplex σ, $F(.)(\sigma) = \max_{v \text{ vertex in } \sigma} F_0(.)(v)$. Since the pre-order induced on K_0 by F_0 is constant over U, the maximum above is attained at vertex $\bar{v}(\sigma)$, and this fact holds for all θ' in U. Thus, $F(.)(\sigma)_{|U} = F_0(.)(\bar{v}(\sigma))_{|U}$, which allows us to conclude. □

Remark 1.5.3. Recall that $\mathrm{Sing}(F_0)$ is by definition the boundary of $\{\theta \in \mathcal{M},\ \exists (v, v') \in K_0, F_0(\theta)(v) = F_0(\theta)(v')\}$, whose complement may not be generic (in fact it may even be empty, e.g. when $F_0 = 0$). This shows the interest of working with locally constant pre-orders on vertices, and not just with locally injective parametrizations as in [BGGSSG20a, CNBW19, GNDS20a, HGR+20a, PSO18].

Defining $\mathrm{Sing}(F_0)$ and $\bar{v}$ as in Proposition 1.5.2, and combining this result with Proposition 1.4.14, we deduce the following result on the differentiability of B_p, which only relies on the differentiability of F_0:

Corollary 1.5.4. *For any C^r parametrization $F_0 : \mathcal{M} \to \mathbb{R}^{K_0}$ on the vertices of K, the induced barcode valued map $B_p : \theta \in \mathcal{M} \mapsto \mathrm{PH}_p(F(\theta)) \in Bar$ is r-differentiable outside $\mathrm{Sing}(F_0)$. Moreover, at $\theta \in \mathcal{M} \backslash \mathrm{Sing}(F_0)$, for any barcode template (P_p, U_p) of $F(\theta)$ and any choice of ordering $(\sigma_1, \sigma_1'), \cdots, (\sigma_m, \sigma_m')$, $\tau_1, \cdots, \tau_n$ of (P_p, U_p), the map $\tilde{B}_p : \mathcal{M} \to \mathbb{R}^m \times \mathbb{R}^n$ defined by:*

$$\tilde{B}_p : \theta' \longmapsto \left[(F_0(\theta')(\bar{v}(\sigma_i)), F_0(\theta')(\bar{v}(\sigma_i')))_{i=1}^m, (F_0(\theta')(\bar{v}(\sigma_j')))_{j=1}^n \right]$$

is a local C^r lift of B_p around θ. The corresponding differential for B_p at θ is:

$$d_{\theta,\tilde{B}_p} B_p(\cdot) = \left[(d_\theta F_0(\cdot)(\bar{v}(\sigma_i)), d_\theta F_0(\cdot)(\bar{v}(\sigma'_i)))_{i=1}^m, (d_\theta F_0(\cdot)(\bar{v}(\sigma'_j)))_{j=1}^n \right].$$

Proof. For $\theta \in \mathcal{M} \backslash \mathrm{Sing}(F_0)$, the pre-order on the vertices K_0 induced by F_0 is constant in an open neighborhood U of θ. By Proposition 1.5.2, each $F(\theta')(\sigma)$ rewrites as $F_0(\theta')(\bar{v}(\sigma))$ for $\theta' \in U$, which implies that the pre-order on the simplices of K induced by F is also constant over U. The fact that B_p is r-differentiable at θ follows then from Theorem 1.4.7, since F itself is C^r on an open neighborhood of θ (again by Proposition 1.5.2, and by the fact that $\mathrm{Sing}(F_0)$ is closed). The rest of the corollary is an immediate consequence of Proposition 1.4.14. □

Example 1.5.5. Consider our running example of parametrization of height filtrations $F_0(\theta) = h_\theta : v \in K_0 \mapsto \langle v, \theta \rangle \in \mathbb{R}$, where K is a fixed geometric simplicial complex in $\mathbb{R}^d$ and $\theta \in \mathbb{S}^{d-1}$. In this case, we know from Example 1.4.11 that B_p is generically ∞-differentiable. Corollary 1.5.4 provides another proof of this fact: since F_0 is C^∞, B_p is ∞-differentiable outside $\mathrm{Sing}(F_0)$, which has generic complement in $\mathbb{S}^{d-1}$. Moreover, the components of the differential of B_p at $\theta \in \mathbb{S}^{d-1} \backslash \mathrm{Sing}(F_0)$ are the $d_\theta F_0(\cdot)(v)$, whose corresponding gradients (in the tangent space $T_\theta \mathbb{S}^{d-1}$ equipped with the Riemannian structure inherited from $\mathbb{R}^d$) are $v - \langle v, \theta \rangle \theta$.

1.5.2 *Rips filtrations of point clouds*

Given a finite point cloud $P = (p_1, \cdots, p_n) \in \mathbb{R}^{nd}$, the *Rips* filtration of P is a filtration of the total complex $K := 2^{\{1,\cdots,n\}} \backslash \{\varnothing\}$ with $n := \#P$ vertices, where the time of appearance of a simplex $\sigma \subseteq \{1, \cdots, n\}$ is $\max_{i,j \in \sigma} \|p_i - p_j\|_2$. [GHO16] optimize the positions of the points of P in $\mathbb{R}^d$ so that the barcode of the Rips filtration reaches some target barcode. Here we see $\mathbb{R}^{nd}$ as our parameter space $\mathcal{M}$, and we consider the parametrization

Marcio Gameiro, Yasuaki Hiraoka, and Ippei Obayashi. Continuation of point clouds via persistence diagrams. *Physica D: Nonlinear Phenomena*, 334:118–132, 2016

$$F(P)(\sigma) := \max_{i,j \in \sigma} \|p_i - p_j\|_2.$$

The differentiability result of [GHO16] can be expressed as a result on the differentiability of the barcode-valued map $B_p = \mathrm{PH}_p \circ F$ using our framework. We require that the points of P lie in general position as defined hereafter:

Definition 1.5.6 ([GHO16])**.** P is in *general position* if the following two conditions hold:

(i) $\forall i \neq j \in \{1, ..., n\}$, $p_i \neq p_j$;

(ii) $\forall \{i, j\} \neq \{k, l\}$, where $i, j, k, l \in \{1, ..., n\}$, $\|p_i - p_j\|_2 \neq \|p_k - p_l\|_2$.

We denote by $\tilde{\mathcal{P}} \subseteq \mathbb{R}^{nd}$ the subspace of point clouds in general position.

Proposition 1.5.7. *$\tilde{\mathcal{P}}$ is generic in $\mathbb{R}^{nd}$.*

Proof. The set of point clouds P such that $p_i \neq p_j$ for all $1 \leqslant i \neq j \leqslant n$ is clearly generic in $\mathbb{R}^{nd}$. Moreover, the maps $P = (p_1, ..., p_n) \mapsto \|p_i - p_j\|_2^2 - \|p_k - p_l\|_2^2$ are smooth everywhere and are submersions on a generic subset of $\mathbb{R}^{nd}$, therefore their 0-sets have generic complements, whose (finite) intersection is also generic. □

We next observe that the parametrization F is C^∞ at point clouds P in general position.

Proposition 1.5.8. *The parametrization $F : \mathbb{R}^{nd} \to \mathbb{R}^K$ is C^∞ over $\tilde{\mathcal{P}}$. Specifically, given $P \in \tilde{\mathcal{P}}$, letting $\{\bar{v}(\sigma), \bar{w}(\sigma)\} = \operatorname{argmax}_{i,j\in\sigma} \|p_i - p_j\|_2$ for every $\sigma \in K$, there is an open neighborhood U of P such that $F(P')(\sigma) = \|p'_{\bar{v}(\sigma)} - p'_{\bar{w}(\sigma)}\|_2$ for every $P' = (p'_1, \cdots, p'_n) \in U$ and $\sigma \in K$, from which follows that F is C^∞ at P.*

Proof. The continuity of F follows from the continuity of the Euclidean norm and max function. Assuming P is in general position, the distances $\|p_i - p_j\|_2$, for $i \neq j$ ranging in $\{1, \cdots, n\}$, are strictly ordered. By continuity of F, this order remains the same over an open neighborhood U of P in $\mathbb{R}^{nd}$. Therefore, every $P' = (p'_1, \cdots, p'_n) \in U$ is also in general position, and $F(P')(\sigma) = \|p'_{\bar{v}(\sigma)} - p'_{\bar{w}(\sigma)}\|_2$ for all $\sigma \in K$. Now, the map $P' \mapsto \|p'_{\bar{v}(\sigma)} - p'_{\bar{w}(\sigma)}\|_2$ is C^∞ at P for each σ because $p'_{\bar{v}(\sigma)} \neq p'_{\bar{w}(\sigma)}$. This implies that F is C^∞ at P. □

Defining $\bar{v}, \bar{w}$ as in Proposition 1.5.8, and combining this result with Proposition 1.4.14, we deduce the following differential of B_p, which only relies on derivatives of the Euclidean distance between points:

Corollary 1.5.9. *The barcode valued map $B_p : P \in \mathbb{R}^{nd} \mapsto \mathrm{PH}_p(F(P)) \in Bar$ is ∞-differentiable in $\tilde{\mathcal{P}}$. Moreover, at $P \in \tilde{\mathcal{P}}$, for any barcode template (P_p, U_p) o $F(P)$ and any choice of ordering $(\sigma_1, \sigma'_1), \cdots, (\sigma_m, \sigma'_m), \tau_1, \cdots, \tau_n$ of (P_p, U_p), the map $\tilde{B}_p$ defined on a point cloud $P' = (p'_1, ..., p'_n)$ by:*

$$\tilde{B}_p(P') = \left[\left(\|p'_{\bar{v}(\sigma_i)} - p'_{\bar{w}(\sigma_i)}\|_2, \|p'_{\bar{v}(\sigma'_i)} - p'_{\bar{w}(\sigma'_i)}\|_2 \right)_{i=1}^m, \left(\|p'_{\bar{v}(\tau_j)} - p'_{\bar{w}(\tau_j)}\|_2 \right)_{j=1}^n \right]$$

is a local C^∞ lift of B_p around P. The corresponding differential $d_{P,\tilde{B}_p} B_p : \mathbb{R}^{nd} \to \mathbb{R}^{2m} \times \mathbb{R}^n$ is defined on a tangent vector $u \in \mathbb{R}^{nd}$ by:

$$d_{P,\tilde{B}_p} B_p(u) = \left[\left(\langle \mathbf{P}_{\bar{v}(\sigma_i),\bar{w}(\sigma_i)}, u \rangle, \langle \mathbf{P}_{\bar{v}(\sigma'_i),\bar{w}(\sigma'_i)}, u \rangle \right)_{i=1}^m, \left(\langle \mathbf{P}_{\bar{v}(\tau_j),\bar{w}(\tau_j)}, u \rangle \right)_{j=1}^n \right],$$

where $\mathbf{P}_{i,j}$ denotes the vector with $\frac{p_i - p_j}{\|p_i - p_j\|_2}$ as i-th component (resp. $\frac{p_j - p_i}{\|p_i - p_j\|_2}$ as j-th component) and 0 as other components.

This result implies in particular that B_p is generically ∞-differentiable, since by Proposition 1.5.7 the set of point clouds in general position is generic in $\mathbb{R}^{nd}$.

Proof. By Proposition 1.5.8, F is C^∞ in $\tilde{\mathcal{P}}$, which is open by Proposition 1.5.7. Given P in general position, the distances $\|p_i - p_j\|_2$, for $i \neq j$ ranging in $\{1, \cdots, n\}$, are strictly ordered, and this order remains the same over an open neighborhood U of P in $\mathbb{R}^{nd}$ by continuity. By Proposition 1.5.8 again, we have $F(P')(\sigma) = \|p'_{\vec{v}(\sigma)} - p'_{\vec{w}(\sigma)}\|_2$ for every $P' = (p'_1, ..., p'_n) \in U$ and $\sigma \in K$. Therefore, the pre-order induced by F on the simplices of K is constant over U. Consequently, B_p is ∞-differentiable at P by Theorem 1.4.7. The rest of the statement is an immediate consequence of Proposition 1.4.14. □

We conclude this section by considering a parametrization that constraints the points $p_1, ..., p_n$ to evolve along smooth submanifolds $\mathcal{M}_1, ..., \mathcal{M}_n$ of $\mathbb{R}^d$:

Proposition 1.5.10. *Let $\mathcal{M}_1, ..., \mathcal{M}_n$ be smooth submanifolds of $\mathbb{R}^d$. Denoting by $\iota : \mathcal{M}_1 \times ... \times \mathcal{M}_n \hookrightarrow \mathbb{R}^{nd}$ the inclusion map, the barcode valued map $B_p = \mathrm{PH}_p \circ F \circ \iota$ is generically ∞-differentiable.*

Proof. Let $\mathcal{M} := \mathcal{M}_1 \times ... \times \mathcal{M}_n$. The parametrization $F \circ \iota$ is C^∞ at parameters $\theta \in \mathcal{M}$ such that, locally, for all θ' in a sufficiently small open neighborhood around θ, the following quantities are constant:

(i) the indices of the points at distance 0 of each other in $\iota(\theta')$, and

(ii) the pre-order on the pairwise distances in $\iota(\theta')$.

Note that, in this case, the point clouds $\iota(\theta')$ are not necessarily in general position, but the way they violate conditions (i) and (ii) of Definition 1.5.6 is constant. Let U' (resp. U) denote the set of points in $\mathcal{M}$ where (i) (resp. (ii)) is satisfied. From the above, $F \circ \iota$ is C^∞ over $U \cap U'$. We now show that $U \cap U'$ is generic in $\mathcal{M}$.

Calling U_{ijkl} the quadric $\{P \in \mathbb{R}^{nd} \mid \|p_i - p_j\|_2 = \|p_k - p_l\|_2\}$, and U'_{ij} the hyperplane $\{P \in \mathbb{R}^{nd} \mid p_i = p_j\}$, for i, j, k, l ranging in $\{1, \cdots, n\}$, we have:

$$U = \bigcap_{\{i,j\} \neq \{k,l\}} \mathcal{M} \backslash \partial \iota^{-1}(U_{ijkl})$$
$$U' = \bigcap_{i \neq j} \mathcal{M} \backslash \partial \iota^{-1}(U'_{ij}).$$

Indeed, for any $\{i, j\} \neq \{k, l\}$, the order between $\|p_i - p_j\|_2$ and $\|p_k - p_l\|_2$ in $\iota(\theta)$ is strict when θ is in the (open) complement of $\iota^{-1}(U_{ijkl})$, constantly an equality when θ is inside the (open) interior $\iota^{-1}(U_{ijkl})^\circ$,

and not locally constant when θ lies on the boundary $\partial\iota^{-1}(U_{ijkl})$. Hence the formula for U. The formula for U' follows from the same argument.

The sets $\partial\iota^{-1}(U_{ijkl})$ and $\partial\iota^{-1}(U'_{ij})$ are boundaries of closed sets, and thus their complements in $\mathcal{M}$ are generic. As finite intersections of generic sets, U and U' are themselves generic. Theorem 1.4.9 allows us to conclude. □

1.5.3 *Rips filtrations of clouds of ellipsoids*

As pointed out in [BKSW18], in some cases, growing isotropic balls around the points of $P = (p_1, ..., p_n) \in \mathbb{R}^{nd}$ may result in a loss of geometric information. It is then advised to grow rather ellipsoids with distinct covariance matrices around each point, to account for the local anistropy of the problem. Formally, the *Ellipsoid-Rips* filtration of P with respect to the vector of covariance matrices $A = (A_1, ..., A_n) \in (S_{d,+}(\mathbb{R}))^n$ is a filtration of the total complex $K := 2^{\{1,...,n\}} \setminus \{\varnothing\}$ with $n := \#P$ vertices, in which the time of appearance of a simplex $\sigma \subseteq \{1, ..., n\}$ is given by:

Paul Breiding, Sara Kališnik, Bernd Sturmfels, and Madeleine Weinstein. Learning algebraic varieties from samples. *Revista Matemática Complutense*, 31(3):545–593, 2018

$$\max_{i,j\in\sigma} r_{i,j}(A) \quad \text{where} \quad r_{i,j}(A) := \left\| \frac{p_i - p_j}{\frac{1}{2}\left(\sqrt{q_i\left(\frac{p_i-p_j}{\|p_i-p_j\|_2}\right)} + \sqrt{q_j\left(\frac{p_j-p_i}{\|p_i-p_j\|_2}\right)}\right)} \right\|_2,$$

where the $q_i : x \in \mathbb{R}^d \mapsto \langle A_i x, x\rangle$ are the quadrics determined by the positive definite matrices A_i.[14] Here we see the space $(S_{d,+}(\mathbb{R}))^n$ as our parameter space $\mathcal{M}$, whose smooth structure is inherited from that of $\left(\mathbb{R}^{\frac{d(d+1)}{2}}\right)^n$, and we consider the parametrization:

$$F(A)(\sigma) := \max_{i,j\in\sigma} r_{i,j}(A).$$

[14] The quantity $r_{i,j}(A)$ serves as a proxy for the intersection of the two ellipsoids with covariance matrices A_i and A_j centered at p_i and p_j suggested by [BKSW18], as the problem of computing intersections of quadrics is in general NP-hard.

We are then interested in the differentiability of the barcode valued map $B_p = \mathrm{PH}_p \circ F$. Inspired by the case of isotropic Rips filtrations, we require that the covariance matrices in A lie in general position as defined hereafter:

Definition 1.5.11. The pair (A, P) is in *general position* if the two following conditions hold:

- all points in P are distinct, i.e, $p_i \neq p_j$ whenever $1 \leqslant i \neq j \leqslant n$;
- all pairwise "ellipsoidal" distances are distinct, i.e, $r_{i,j}(A) \neq r_{k,l}(A)$ whenever $\{i, j\} \neq \{k, l\} \subseteq \{1, ..., n\}$.

Proposition 1.5.12. *Assume the points of P to be pairwise distinct. Then, the set of vectors of covariance matrices A such that (A, P) is in general position is generic in $S_{d,+}(\mathbb{R}^d)^n$.*

Proof. First, we claim that the sets $O_{ijkl} := \{A \in S_{d,+}(\mathbb{R}^d)^n \mid r_{i,j}(A) = r_{k,l}(A)\}$, for $\{i,j\} \neq \{k,l\}$, are level-sets of some smooth real valued functions on $S_{d,+}(\mathbb{R}^d)^n$ whose gradients are nowhere zero. To prove this fact, we introduce the quantities $C := \frac{\|p_i - p_j\|_2}{\|p_k - p_l\|_2}$ and $(x,y) := (\frac{p_i - p_j}{\|p_i - p_j\|_2}, \frac{p_k - p_l}{\|p_k - p_l\|_2})$. Then:

$$\begin{aligned} A = (A_1, ..., A_n) \in O_{ijkl} &\Leftrightarrow r_{i,j}(A) = r_{k,l}(A) \\ &\Leftrightarrow \frac{\sqrt{q_i(x)} + \sqrt{q_j(x)}}{\sqrt{q_k(y)} + \sqrt{q_l(y)}} = C \\ &\Leftrightarrow \frac{\sqrt{< A_i x, x >} + \sqrt{< A_j x, x >}}{\sqrt{< A_k y, y >} + \sqrt{< A_l y, y >}} = C. \end{aligned}$$

Note that x, y are non zero because points in P are distinct. Therefore, the map $f_{ijkl} := A \in S_{d,+}(\mathbb{R})^d \mapsto \frac{\sqrt{<A_i x,x>}+\sqrt{<A_j x,x>}}{\sqrt{<A_k y,y>}+\sqrt{<A_l y,y>}} \in \mathbb{R}$ is well-defined and smooth on $S_{d,+}(\mathbb{R})^n$ as the two inner products in the denominator are always strictly positive. We want to compute $\nabla f_{ijkl} = (\nabla_{A_1} f_{ijkl}, ..., \nabla_{A_n} f_{ijkl})$ where $\nabla_{A_t} f_{ijkl}$ is the gradient of f_{ijkl} with respect to the t-th component of A. For $t = i$:

$$\nabla_{A_i} f_{ijkl} = \frac{1}{\sqrt{< A_k y, y >} + \sqrt{< A_l y, y >}} \times \frac{1}{2\sqrt{< A_i x, x >}} \times \nabla_{A_i} < A_i x, x > .$$

The first two factors are strictly positive scalars for any $A \in S_{d,+}(\mathbb{R})^d$. The last factor is the gradient of a non-zero linear map, so it is non-zero. As a consequence, the gradient $\nabla_A f_{ijkl}$ is nowhere zero, which proves our claim.

Then, by the constant rank theorem, each O_{ijkl} is a smooth submanifold of $S_{d,+}(\mathbb{R}^d)^n$ of dimension strictly lower than that of $S_{d,+}(\mathbb{R}^d)^n$. Taking their (finite) union allows us to conclude. □

From this point, the same chain of arguments as in the isotropic case allows us to show that the parametrization F is C^∞ at vectors of covariance matrices A in general position, and to express the differential of B_p at A. Assume the points of P to be pairwise distinct, and denote by $\tilde{\mathcal{A}} \subseteq S_{d,+}(\mathbb{R}^d)^n$ the subspace of covariance matrices A such that (A, P) is in general position.

Proposition 1.5.13. *The parametrization $F : S_{d,+}(\mathbb{R}^d)^n \to \mathbb{R}^K$ is C^∞ over $\tilde{\mathcal{A}}$. Specifically, given $A \in \tilde{\mathcal{A}}$, letting $\{\bar{v}(\sigma), \bar{w}(\sigma)\} = \operatorname{argmax}_{i,j \in \sigma} r_{i,j}(A)$ for every $\sigma \in K$, there is an open neighborhood U of A such that $F(A')(\sigma) = r_{\bar{v}(\sigma),\bar{w}(\sigma)}(A')$ for every $A' = (A'_1, ..., A'_n) \in U$ and $\sigma \in K$, from which follows that F is C^∞ at A.*

Proof. Let $A \in \tilde{\mathcal{A}}$. Then, the maps $r_{i,j}$ are C^∞ because the points of P are pairwise distinct, and furthermore the quantities $r_{i,j}(A)$, for $i \neq j$ ranging in $\{1, \cdots, n\}$, are strictly ordered. By continuity, this order remains

the same over an open neighborhood U of A in $S_{d,+}(\mathbb{R}^d)^n$. Therefore, for every $A' \in U$, for all $\sigma \in K$, we have $F(A')(\sigma) = r_{\bar{v}(\sigma),\bar{w}(\sigma)}(A')$. This implies that F is C^∞ at A. □

Defining $\bar{v}, \bar{w}$ as in Proposition 1.5.13, and combining this result with Proposition 1.4.14, we deduce the following formula for the differential of B_p, which only rely on derivatives of the maps $r_{i,j}$:

Corollary 1.5.14. *The barcode valued map $B_p : A \in S_{d,+}(\mathbb{R}^d)^n \mapsto \mathrm{PH}_p(F(A)) \in$ Bar is ∞-differentiable over $\tilde{\mathcal{A}}$. Moreover, at $A \in \tilde{\mathcal{A}}$, for any barcode template (P_p, U_p) of $F(A)$ and any choice of ordering $(\sigma_1, \sigma'_1), \cdots, (\sigma_m, \sigma'_m), \tau_1, \cdots, \tau_n$ of (P_p, U_p), the map $\tilde{B}_p$ defined by:*

$$A' = (A'_1, ..., A'_n) \longmapsto \left[\left(r_{\bar{v}(\sigma_i),\bar{w}(\sigma_i)}(A'), r_{\bar{v}(\sigma'_i),\bar{w}(\sigma'_i)}(A') \right)_{i=1}^{m}, \left(r_{\bar{v}(\tau_j),\bar{w}(\tau_j)}(A') \right)_{j=1}^{n} \right]$$

is a local C^∞ lift of B_p around P, whose differential provides a closed formula for $d_{A,\tilde{B}_p} B_p$.

This result implies in particular that B_p is generically ∞-differentiable, since by Proposition 1.5.12 the set of vectors of covariance matrices in general position is generic in $S_{d,+}(\mathbb{R}^d)^n$ (provided the points of P are pairwise distinct).

Proof. By Proposition 1.5.13, F is C^∞ in $\tilde{\mathcal{A}}$, which is open by Proposition 1.5.12. Given $A \in \tilde{\mathcal{A}}$, the quantities $r_{i,j}(A)$, for $i \neq j$ ranging in $\{1, \cdots, n\}$, are strictly ordered, and this order remains the same over an open neighborhood U of A in $S_{d,+}(\mathbb{R}^d)^n$ by continuity. By Proposition 1.5.13 again, we have $F(A')(\sigma) = r_{\bar{v}(\sigma),\bar{w}(\sigma)}(A')$ for every $A' = (A'_1, ..., A'_n) \in U$ and $\sigma \in K$. Therefore, the pre-order induced by F on the simplices of K is constant over U. Consequently, B_p is ∞-differentiable at A by Theorem 1.4.7. The rest of the statement is an immediate consequence of Proposition 1.4.14. □

Remark 1.5.15. Corollaries 1.5.9 and 1.5.14 can be combined together to generically differentiate the barcode valued map B_p with respect to both the point positions and the covariance matrices. The corresponding parameter space is $\mathbb{R}^{nd} \times S_{d,+}(\mathbb{R}^d)^n$.

1.5.4 *Arbitrary filtrations of a simplicial complex*

In certain scenarios, the optimization takes place in the entire space of filter functions $\mathrm{Filt}(K)$ on a fixed simplicial complex K. For instance, in the context of topological simplification of a filter function f_0, as described by [AGH+09a, ELZ02], one looks for a filter function $f \in \mathbb{R}^K$ which is ε-close to f_0 in supremum norm and whose diagram $\mathrm{PH}_p(f)$ equals $\mathrm{PH}_p(f_0) \backslash \Delta_\epsilon$, where Δ_ϵ is the set of intervals of $\mathrm{PH}_p(f_0)$ that

Dominique Attali, Marc Glisse, Samuel Hornus, Francis Lazarus, and Dmitriy Morozov. Persistence-sensitive simplification of functions on surfaces in linear time. In *Topological Methods in Data Analysis and Visualization (TopoInVis 2009)*. Springer, 2009; and Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002

are ε-close to the diagonal. One way to formalize this question is as a soft-constrained optimization problem, whereby the bottleneck distance to the simplified barcode is to be minimized in tandem with the supremum-norm distance to the original function:

$$\min_{f\in \mathrm{Filt}(K)} d_\infty(\mathrm{PH}_p(f), \mathrm{PH}_p(f_0)\backslash\Delta_\epsilon) + \lambda \left\| f - f_0 \right\|_\infty,$$

for some fixed mixing parameter λ. This optimization problem can be tackled using a variational approach, for which it is more convenient to work in the manifold $\mathbb{R}^K$ containing $\mathrm{Filt}(K)$. However, in order to avoid leaving $\mathrm{Filt}(K)$, we consider the parametrization of $\mathbb{R}^K$ given by the indicator function of $\mathrm{Filt}(K)$:

$$\begin{array}{rccl} F := \mathbf{1}_{\mathrm{Filt}(K)} : & \mathbb{R}^K & \to & \mathbb{R}^K \\ & f & \mapsto & \begin{cases} f \text{ if } f \in \mathrm{Filt}(K) \\ 0 \text{ otherwise,} \end{cases} \end{array}$$

which is smooth generically. The optimisation becomes then:

$$\min_{f\in \mathbb{R}^K} d_\infty(\mathrm{PH}_p(F(f)), \mathrm{PH}_p(f_0)\backslash\Delta_\epsilon) + \lambda \left\| F(f) - f_0 \right\|_\infty. \tag{1.18}$$

Implementing a variational approach such as gradient descent requires both terms in (1.18) to be differentiable. The second term is generically differentiable, as the parametrization F and the norm $\|\cdot\|_\infty$ are. The first term is the composition

$$f \in \mathbb{R}^K \longrightarrow \mathrm{PH}_p(F(f)) \in Bar \longrightarrow d_\infty(\mathrm{PH}_p(F(f)), \mathrm{PH}_p(f_0)\backslash\Delta_\epsilon) \in \bar{\mathbb{R}}, \tag{1.19}$$

which by the chain rule (Proposition 1.3.14) is differentiable as long as both arrows are. Since F is generically differentiable, so is the first arrow by Theorem 1.4.9. The second arrow is the bottleneck distance to a fixed diagram and therefore also generically differentiable, as will be argued in Section 1.7. There, we also view Eq. (1.18) as an instance of semi-algebraic loss function, which can be minimised via Stochastic Gradient Descent (SGD).

1.6 *The case of barcode valued maps derived from real functions on a manifold*

In this section we consider barcode valued maps that factor through the space $\mathbb{R}^{\mathcal{X}}$ of real functions on a fixed smooth compact d-manifold $\mathcal{X}$ without boundary. Since we seek statements about the differentiability of B, we restrict the focus to maps that factor through $C^\infty(\mathcal{X}, \mathbb{R})$ equipped with the standard Whitney C^∞ topology:[15]

[15] This topology coincides with all usual topologies on $C^\infty(\mathcal{X}, \mathbb{R})$ because $\mathcal{X}$ is compact.

$$B : \quad \mathcal{M} \xrightarrow{F} C^\infty(\mathcal{X}, \mathbb{R}) \xrightarrow{\mathrm{PH}} Bar^{d+1} .$$

Here, PH is the map that takes a function $f \in C^\infty(\mathcal{X}, \mathbb{R})$ to the vector of its barcodes $(\mathrm{PH}_p(f))_{p=0}^d$. It is well-defined on $C^\infty(\mathcal{X}, \mathbb{R})$, as continuous functions on triangulable spaces have well-defined persistence diagrams[16]. However, as in the previous sections, we want to work only with barcodes that have finitely many off-diagonal points, therefore we further assume that F takes its values in the subset $\mathrm{Tame}(\mathcal{X})$ of tame C^∞ functions—note that $\mathrm{Tame}(\mathcal{X})$ contains the generic subset of Morse functions[17] on $\mathcal{X}$. Hence the factorization:

$$B : \quad \mathcal{M} \xrightarrow{F} \mathrm{Tame}(\mathcal{X}) \xrightarrow{\mathrm{PH}} Bar^{d+1} .$$

As before, we call F the *parametrization* associated to B, and $\mathcal{M}$ the *parameter space*, whose elements are generally refered to as θ. We also denote $F(\theta)$ by f_θ to emphasize the fact that F is valued in a function space. The map PH takes f_θ to the vector of its barcodes $(\mathrm{PH}_p(f_\theta))_{p=0}^d$, so we can take advantage of the bijective correspondence between the critical points of f_θ (provided f_θ is Morse) and the interval endpoints in this vector (Proposition 1.2.14).

As in the case of a parametrization valued in the space of filter functions on a simplicial complex, we need F to be smooth in some reasonable sense to ensure that the composite B is ∞-differentiable. For this, we define a curve $c : \mathbb{R} \to C^\infty(\mathcal{X}, \mathbb{R})$ to be *differentiable* if the limit $\lim_{h\to 0} \frac{c(t+h)-c(t)}{h}$ exists for all $t \in \mathbb{R}$. The limit can be viewed as a curve, and when iterated limits exist, we say that c is a *smooth* curve. We then say that the parametrization F is *smooth*[18] if it sends every smooth curve $\theta(t)$ in $\mathcal{M}$ to a smooth curve $F(\theta(t))$ in $C^\infty(\mathcal{X}, \mathbb{R})$. By Corollary 11.9 in [Mic80][19], if F is smooth, then its uncurrified version

$$\tilde{F} : (\theta, x) \in \mathcal{M} \times \mathcal{X} \longmapsto F(\theta)(x) \in \mathbb{R} \tag{1.20}$$

is a smooth map in the usual sense, to which we can therefore apply standard results from differential calculus, typically the implicit function theorem. This will be instrumental in the proof of our main result (Theorem 1.6.1).

[16] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The structure and stability of persistence modules*. SpringerBriefs in Mathematics. Springer, 2016

[17] John Milnor. Morse theory. *Princeton University Press, Princeton, N.J.*, pages vi+153, 1963

[18] This notion of smooth map is key in the theory developped in [FK88, KM97] for adapting the concepts of differential geometry to a wide category of infinite-dimensional vector spaces, including Banach and Fréchet spaces.

[19] Peter W. Michor. *Manifolds of differentiable mappings*. Shiva Mathematics Series, volume 3. Shiva Publishing Ltd., Nantwich, 1980

1.6.1 *Smoothness of the barcode valued map*

Theorem 1.6.1 (Continuous smoothness)**.** *Let $F : \mathcal{M} \to C^\infty(\mathcal{X}, \mathbb{R})$ be a parametrization of class C_c^∞ valued in $\mathrm{Tame}(\mathcal{X})$. Let $\theta \in \mathcal{M}$ be a parameter such that f_θ is Morse with critical values of multiplicity 1. Then, B is ∞-differentiable at θ.*

Proof. Since f_θ is a Morse function on a compact manifold, $\mathrm{Crit}(f_\theta)$ is a finite set whose cardinality we denote by N_θ. We will proceed by proving the following statements in sequence:

(i) There exist an open neighborhood U of θ and smooth maps $\pi_l : U \to \mathcal{X}$ for $1 \leqslant l \leqslant N_\theta$ that track the critical points, that is:

$$\forall \theta' \in U, \mathrm{Crit}(f_{\theta'}) = \{\pi_l(\theta')\}_{1 \leqslant l \leqslant N_\theta} \tag{1.21}$$

(ii) Shrinking U if necessary, we further have that for any $\theta' \in U$, $f_{\theta'}$ is Morse with critical values of multiplicity 1.

(iii) Let $\theta' \in U$ and $(b, d) \in \mathrm{PH}_p(\mathcal{X}, f_{\theta'}) \backslash \Delta$ for some homology degree p. Then, either $d = +\infty$, in which case there exists a unique $1 \leqslant l \leqslant N_\theta$ such that $b = f_{\theta'}(\pi_l(\theta'))$, or $d < +\infty$, in which case there exist unique $1 \leqslant l \neq l' \leqslant N_\theta$ such that $(b, d) = (f_{\theta'}(\pi_l(\theta')), f_{\theta'}(\pi_{l'}(\theta')))$.

(iv) For all $\theta_1, \theta_2 \in U$, $1 \leqslant l \neq l' \leqslant N_\theta$, and $0 \leqslant p \leqslant d$, we have: $(f_{\theta_1}(\pi_l(\theta_1)), f_{\theta_1}(\pi_{l'}(\theta_1))) \in \mathrm{PH}_p(f_{\theta_1})$ (resp. $(f_{\theta_1}(\pi_l(\theta_1)), +\infty) \in \mathrm{PH}_p(f_{\theta_1})$) if and only if $(f_{\theta_2}(\pi_l(\theta_2)), f_{\theta_2}(\pi_{l'}(\theta_2))) \in \mathrm{PH}_p(f_{\theta_2})$ (resp. $(f_{\theta_2}(\pi_l(\theta_2)), +\infty) \in \mathrm{PH}_p(f_{\theta_2})$).

(v) There exist smooth local coordinate systems for B_p at θ for every $0 \leqslant p \leqslant d$. Therefore, by Proposition 1.3.8, the barcode valued map B is ∞-differentiable at θ.

The proofs of assertions (i) and (ii) use differential geometry: we show that we can smoothly track the critical points of $f_{\theta'}$ as θ' varies in a neighborhood of θ. The proof of assertion (iii) simply exploits the fact that the endpoints in the barcodes of a Morse function are its critical values (Propostion 1.2.14). Assertion (iv) means that the critical points do not exchange their contributions to the persistence diagrams when the parameter is varying. This will be shown using standard tools in persistence theory. Assertion (v) is obtained by re-indexing the set $\{1, ..., N_\theta\}$ such that, through this re-indexation, the maps $\theta' \mapsto f_{\theta'}(\pi_l(\theta'))$ provide local coordinate systems as defined in Definition 1.3.6.

Proof of assertion (i): The tangent bundle $T\mathcal{X} = \bigsqcup_{x \in \mathcal{X}} \{x\} \times T_x\mathcal{X}$ is a smooth manifold of dimension $2d$. Let $x_1, ..., x_{N_\theta}$ be the critical points of f_θ. Locally, in an open neighborhood $\mathcal{V}$ of these critical points, the tangent bundle is parallelizable, i.e. we have a diffeomorphism $T\mathcal{V} \cong \mathcal{V} \times \mathbb{R}^d$ and the projection onto the second component provides a smooth map to $\mathbb{R}^d$. Consider the map:

$$\partial F : (\theta', x) \in \mathcal{M} \times \mathcal{V} \mapsto \nabla f_{\theta'}(x) \in T_x\mathbb{V} \cong \mathbb{R}^d,$$

which is smooth due to the smoothness of $\tilde{F}$, see Eq. (1.20). Then, at the critical points we have $\partial F(\theta, x_l) = \nabla f_\theta(x_l) = 0$. Moreover, because f_θ is Morse, $\nabla_x \partial F(\theta, x_l) = \nabla^2 f_\theta(x_l)$ is invertible, where $\nabla_x \partial F$ denotes the first derivative of ∂F with respect to its second argument. We can

then apply the implicit function theorem to ∂F: there exist an open neighborhood U_l of θ, an open neighborhood V_l of x_l (contained in $\mathcal{V}$) and a smooth diffeomorphism $\pi_l : U_l \to V_l$ such that

$$\forall (\theta', x) \in U_l \times V_l,\ \partial F(\theta', x) = 0 \iff x = \pi_l(\theta'). \tag{1.22}$$

Let $U = \bigcap_{l=1}^{N_\theta} U_l$. After shrinking each V_l so that it equals $\pi_l(U)$, we obtain that (1.22) holds over $U \times V_l$ for every $1 \leqslant l \leqslant N_\theta$. Now, by definition of ∂F and the ($\Leftarrow$) of (1.22), we have

$$\forall \theta' \in U, \{\pi_l(\theta')\}_{1 \leqslant l \leqslant N_\theta} \subseteq \mathrm{Crit}(f_{\theta'}).$$

We now show the converse inclusion. From the ($\Rightarrow$) in Equation (1.22), it is sufficient to prove that no critical points of $f_{\theta'}$ can be found in the compact set $W := \mathcal{X} \backslash (\bigcup_{l=1}^{N_\theta} V_l)$ when θ' ranges over U. We equip $\mathcal{X}$ with an arbitrary Riemannian metric g, and we consider the smooth map:

$$\partial G : (\theta', x) \in U \times \mathcal{X} \longmapsto g(\nabla f_{\theta'}(x), \nabla f_{\theta'}(x)) \in \mathbb{R},$$

where $\nabla f_{\theta'}(x) \in T_x\mathcal{X}$. In particular, $\partial G(\theta', x)$ is zero if and only if x is a critical point of $f_{\theta'}$. As a result, ∂G does not vanish on $\{\theta\} \times W$ since W includes no critical point of $f_{\theta'}$. By the compactness of W and the continuity of ∂G, there exists an open neighborhood U' of θ such that $\partial G_{|U' \times W}$ does not vanish either. Assertion (i) follows after shrinking U to $U \cap U'$.

Proof of assertion (ii): Let U be as in assertion (i). Since f_θ is Morse, $\nabla_x \partial F(\theta, x_l) = \nabla^2 f_\theta(x_l)$ is invertible for each $l \in \{1, ..., N_\theta\}$. ∂F is of class C^1 as it is of class C^∞, so we get open neighborhoods U_l' of θ and V_l' of x_l such that $\nabla_x \partial F$ is invertible over $U_l' \times V_l'$. We shrink U to $U \cap (\bigcap_{l=1}^{N_\theta} U_l')$ and each V_l to $V_l \cap V_l'$, so that the critical points of $f_{\theta'}$ are non-degenerate for $\theta' \in U$. Shrinking U further if necessary, a similar argument ensures that the critical values of $f_{\theta'}$ have multiplicity 1 for all $\theta' \in U$. This concludes the proof of assertion (ii).

Proof of assertion (iii): Let $\theta' \in U$. Let $(b, d) \in \mathrm{PH}_p(f_{\theta'}) \backslash \Delta$ for some homology degree $0 \leqslant p \leqslant d$. We assume that $d < +\infty$. From assertion (ii), $f_{\theta'}$ is Morse with critical values of multiplicity 1. Therefore, by Proposition 1.2.14, $f_{\theta'}$ induces a bijection between the multisets $\mathrm{Crit}(f_{\theta'})$ and $E(f_{\theta'})$. Meanwhile, assertion (i) provides the equality $\mathrm{Crit}(f_{\theta'}) = \{\pi_l(\theta')\}_{1 \leqslant l \leqslant N_\theta}$, so $f_{\theta'}$ induces a bijection $\{\pi_l(\theta')\}_{1 \leqslant l \leqslant N_\theta} \to E(f_{\theta'})$. By taking pre-images of b and d which are in $E(f_{\theta'})$, there exist some unique indices $1 \leqslant l \neq l' \leqslant N_\theta$ such that $(b, d) = (f_{\theta'}(\pi_l(\theta')), f_{\theta'}(\pi_{l'}(\theta')))$. The case $d = +\infty$ is proven the same way.

Proof of assertion (iv): The maps $(\theta_1, \theta_2) \in U^2 \mapsto |f_{\theta_1}(\pi_l(\theta_1)) - f_{\theta_2}(\pi_{l'}(\theta_2))| \in \mathbb{R}_+$, for varying $1 \leqslant l \neq l' \leqslant N_\theta$, are continuous. They are strictly positive at (θ, θ) because f_θ has critical values of multiplicity 1, so

$$\inf_{1 \leqslant l \neq l' \leqslant N_\theta} |f_\theta(\pi_l(\theta)) - f_\theta(\pi_{l'}(\theta))| > 0.$$

By continuity, shrinking U further if necessary, we have

$$\inf_{1 \leqslant l \neq l' \leqslant N_\theta,\, (\theta_1,\theta_2) \in U^2} |f_{\theta_1}(\pi_l(\theta_1)) - f_{\theta_2}(\pi_{l'}(\theta_2))| > 0.$$

Let ε be a real number such that:

$$0 < \varepsilon < \inf_{1 \leqslant l \neq l' \leqslant N_\theta,\, (\theta_1,\theta_2) \in U^2} |f_{\theta_1}(\pi_l(\theta_1)) - f_{\theta_2}(\pi_{l'}(\theta_2))|. \tag{1.23}$$

By continuity of $\tilde{F}$ and compactness of $\mathcal{X}$, we can shrink U further[20] so that $\|f_{\theta_1} - f_{\theta_2}\|_\infty \leqslant \frac{\varepsilon}{2}$ for any $\theta_1, \theta_2 \in U$. From the Stability Theorem 1.2.12 we then have:

$$\forall \theta_1, \theta_2 \in U, \forall 0 \leqslant p \leqslant d, \quad d_\infty(\mathrm{PH}_p(f_{\theta_1}), \mathrm{PH}_p(f_{\theta_2})) \leqslant \frac{\varepsilon}{2}. \tag{1.24}$$

[20] This does not prevent our choice of ε from satisfying (1.23), because shrinking U increases the right-hand side of this equation.

Let us fix two parameters $\theta_1, \theta_2 \in U$ and a homology degree p. Let $1 \leqslant l_1 \neq l_1' \leqslant N_\theta$ be such that $(f_{\theta_1}(\pi_{l_1}(\theta_1)), f_\theta(\pi_{l_1'}(\theta_1))) \in \mathrm{PH}_p(f_{\theta_1})$. From Equation (1.24), there exists a matching $\gamma : \mathrm{PH}_p(f_{\theta_1}) \to \mathrm{PH}_p(f_{\theta_2})$ with cost $c(\gamma) \leqslant \frac{\varepsilon}{2}$. In particular, if we denote $(b, d) := \gamma(f_{\theta_1}(\pi_{l_1}(\theta_1)), f_{\theta_1}(\pi_{l_1'}(\theta_1))) \in \mathbb{R}^2$, then

$$|f_{\theta_1}(\pi_{l_1}(\theta_1)) - b| \leqslant \frac{\varepsilon}{2} \quad \text{and} \quad |f_{\theta_1}(\pi_{l_1'}(\theta_1)) - d| \leqslant \frac{\varepsilon}{2}. \tag{1.25}$$

Of course we cannot have $d = +\infty$. Also, we cannot have $(b, d) \in \Delta$, i.e $b = d$, because then the triangle inequality would imply that $|f_{\theta_1}(\pi_{l_1}(\theta_1)) - f_{\theta_1}(\pi_{l_1'}(\theta_1))| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, which contradicts (1.23). Thus, (b, d) is a bounded off-diagonal point of $\mathrm{PH}_p(f_{\theta_2})$. By assertion (iii), there exist indices $1 \leqslant l_2 \neq l_2' \leqslant N_\theta$ such that $b = f_{\theta_2}(\pi_{l_2}(\theta_2))$ and $d = f_{\theta_2}(\pi_{l_2'}(\theta_2))$. Equations (1.25) and (1.23) together force $l_2 = l_1$ and $l_2' = l_1'$. Hence, $(f_{\theta_2}(\pi_{l_1}(\theta_2)), f_{\theta_2}(\pi_{l_1'}(\theta_2))) = (b, d) \in \mathrm{PH}_p(f_{\theta_2})$, which proves the result. The case of an index $1 \leqslant l \leqslant N_\theta$ such that $(f_{\theta_1}(\pi_l(\theta_1)), +\infty) \in \mathrm{PH}_p(f_{\theta_1})$ is treated in the same way.

Proof of assertion (v): For any homology degree $0 \leqslant p \leqslant d$, by assertion (iii), each bounded off-diagonal interval (b, d) in $\mathrm{PH}_p(f_\theta) \backslash \Delta$ can be rewritten as $(f_\theta(\pi_{l_{b,p}}(\theta)), f_\theta(\pi_{l_{d,p}}(\theta)))$ for some indices $l_{b,p} \neq l_{d,p}$. Similarly, each interval $(v, +\infty)$ can be rewritten as $(f_\theta(\pi_{l_{v,p}}(\theta)), +\infty)$ for some index $l_{v,p}$. By assertion (iv), for any parameter $\theta' \in U$, $B_p(\theta')$ equals

$$\left\{(f_{\theta'}(\pi_{l_{b,p}}(\theta')), f_{\theta'}(\pi_{l_{d,p}}(\theta')))\right\}_{(b,d) \in \mathrm{PH}_p(f_\theta) \backslash \Delta} \cup \left\{(f_{\theta'}(\pi_{l_{v,p}}(\theta')), +\infty)\right\}_{(v,+\infty) \in \mathrm{PH}_p(f_\theta)} \cup \Delta^\infty.$$

This provides a smooth local coordinate system (see Definition 1.3.7) for B_p at θ, therefore B_p is ∞-differentiable at θ by Proposition 1.3.8. Since this is true for every $0 \leqslant p \leqslant d$, B itself is ∞-differentiable at θ. □

Remark 1.6.2 (Multiplicity one)**.** The upcoming Figure 1.5 shows how important the assumption that f_θ has critical values of multiplicity 1 is for the conclusion of Theorem 1.6.1 to hold. Roughly speaking, the assumption implies that the critical points do not exchange their contributions to the persistence diagrams of f_θ under perturbations of θ. We proved this fact using the Stability Theorem for persistence diagrams (see the proof of assertion (iv) above), however it is also a consequence of the so-called *structural stability theorem* for dynamical systems[21]. This result implies that the gradient vector field induced by a Morse function f_θ with distinct critical values is *structurally stable*, and as an immediate consequence, that the Morse-Smale complex of f_θ does not change as we smoothly perturb f_θ. The Morse-Smale complex allows us to recover the persistence module completely and, in turn, the barcode of f_θ.

[21] Jacob Palis and Stephen Smale. Structural stability theorems. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 223–231. Amer. Math. Soc., Providence, R.I., 1970

1.6.2 *Discussion: generic differentiability*

Theorem 1.6.1 guarantees that B is ∞-differentiable at parameters θ that produce Morse functions with critical values of multiplicity 1. The set of such functions is a generic subspace of $C^\infty(\mathcal{X}, \mathbb{R})$, see [GG73]. We can also argue that, under some extra conditions on the parametrization F, the set $D(\mathcal{M}, \mathcal{X})$ of parameters $\theta \in \mathcal{M}$ that produce Morse functions f_θ with critical values of multiplicity 1 is generic in $\mathcal{M}$:

M. Golubitsky and V. Guillemin. *Stable mappings and their singularities*. Springer-Verlag, New York-Heidelberg, 1973. Graduate Texts in Mathematics, Vol. 14

Proposition 1.6.3 ([Nic11])**.** *If F is smooth and* generically large*, i.e. for generic $x \in \mathcal{X}$ the map $\theta \in \mathcal{M} \mapsto df_\theta(x) \in T_x\mathcal{X}^*$ is a submersion, then $D(\mathcal{M}, \mathcal{X})$ is generic in $\mathcal{M}$.*

Liviu Nicolaescu. *An invitation to Morse theory*. Universitext. Springer, New York, second edition, 2011

There are important examples where this result applies, such as for instance:

Example 1.6.4 ([Nic11])**.** Assume $\mathcal{X}$ is embedded in $\mathbb{R}^d$ and translated so as not to contain the origin. Then, each of the following parametrizations F is smooth and generically large:

$$v \in \mathbb{R}^d \mapsto (x \in \mathcal{X} \mapsto \langle v, x\rangle \in \mathbb{R})$$
$$p \in \mathbb{R}^d \mapsto (x \in \mathcal{X} \mapsto |x - p|^2 \in \mathbb{R})$$
$$A \in S_+(\mathbb{R}^d) \mapsto (x \in \mathcal{X} \mapsto \frac{1}{2}\langle Ax, x\rangle \in \mathbb{R})$$

1.6.3 *A simple example*

Take the ground space $\mathcal{X}$ to be the torus $\mathbb{S}^1 \times \mathbb{S}^1$ embedded in $\mathbb{R}^3$, the parameter space $\mathcal{M}$ to be the 2-sphere $\mathbb{S}^2$, and the parametrization F to

be the family of height filtrations, i.e $F : \theta \in \mathbb{S}^2 \mapsto (x \in \mathcal{X} \mapsto \langle \theta, x \rangle \in \mathbb{R})$. For a generic direction $\theta \in \mathbb{S}^2$, the induced height function, which we denote by h_θ, will be Morse and no two critical points are in the same level set. In this case we can track the critical points smoothly as we vary θ, and the barcodes $\mathrm{PH}_p(h_\theta)$ also evolve smoothly. An example of this situation is given in Figure 1.2.

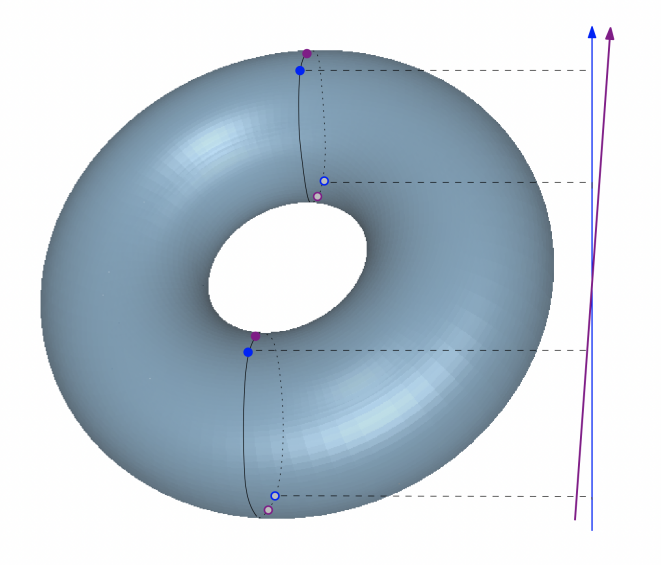


Figure 1.2: A torus filtered by a generic height function. The blue arrow indicates the direction θ. By the correspondence of Proposition 1.2.14, the 4 critical points (blue dots) correspond from bottom to top to an infinite bar in degree 0, an infinite bar in degree 1, another infinite bar in degree 1, and an infinite bar in degree 2 of the resulting barcode $\mathrm{PH}(h_\theta)$. The implicit function theorem applied to these critical points allows us to track them smoothly when perturbing the height function (purple arrow). The correspondence to points in the barcode remains unchanged.

Even in this elementary situation, the singular parameters $\theta \in \mathbb{S}^2$ can exhibit pathological behaviors. There are two specific heights, on opposite sides of the sphere $\mathbb{S}^2$, that produce *Morse-Bott* functions. We show one of them in Figure 1.3. At such a parameter θ, the critical sets are codimension-1 submanifolds of $\mathcal{X}$, and smooth perturbations of θ may result in discontinuous changes in the critical set.

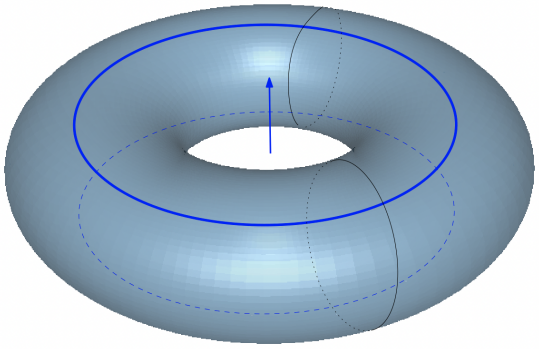


Figure 1.3: Horizontal torus filtered by the vertical height function h_θ. The critical sets are the two blue circles, one of which corresponds to both a birth of a connected component and a loop, while the other corresponds to the births of a loop and a 2-cycle. Observe that any slight perturbation of θ results in a valid Morse function with 4 critical points, however, the locations of these points do not vary smoothly, and not even continuously, at θ.

There are other directions θ at which the assumptions of Theo-

rem 1.6.1 are not met, yet the interval endpoints in the barcode can still be tracked smoothly. Such a case is shown in Figure 1.4, where the height function h_θ is Morse but with a critical value of multiplicity 2. In this specific case, the implicit function theorem still applies to both critical points and provides a smooth local coordinate system for the barcode of h_θ. However, in the general case, such a Morse

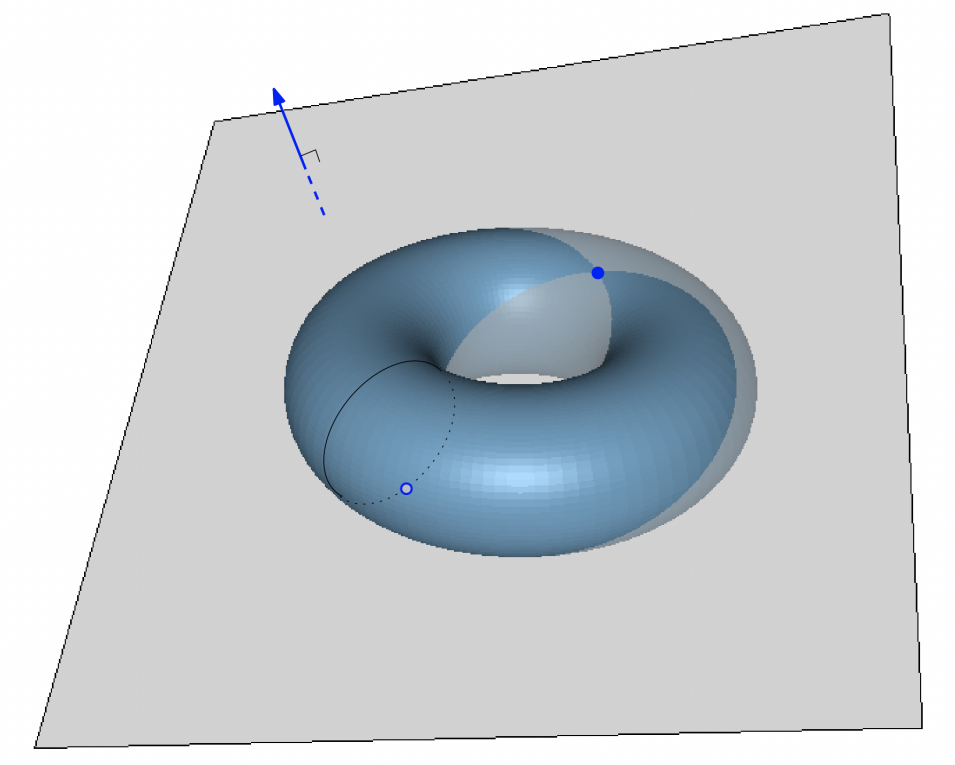


Figure 1.4: A height function h_θ (blue arrow) that is Morse with two critical points (blue dots) in the same level set (the hyperplane), producing two distinct loops in the persistence diagram. The critical points can still be tracked smoothly around θ, and no change in the pairing occurs in the barcode $\mathrm{PH}(h_\theta)$.

function with two critical points sitting in the same level-set can induce a change in the correspondence with interval endpoints in the barcode, potentially resulting in non-smooth behavior of the barcode valued map B. An example is given in Figure 1.5.

1.7 *The case of maps on barcodes derived from vectorizations and loss functions*

We continue on with examples of differentiable maps, this time focusing on maps $V : \mathit{Bar} \to \mathcal{N}$ defined on barcodes and valued in a smooth finite-dimensional manifold. There is a plethora of examples of such maps V in the literature on topological data analysis [AEK+17, Bub15, COO15, CCO17, DFF15, Kal19]. Most of them take $\mathcal{N}$ to be a Euclidean or Hilbert space, and they were designed to provide meaningful (e.g. stable, discriminative) representations of barcodes that can be fed to

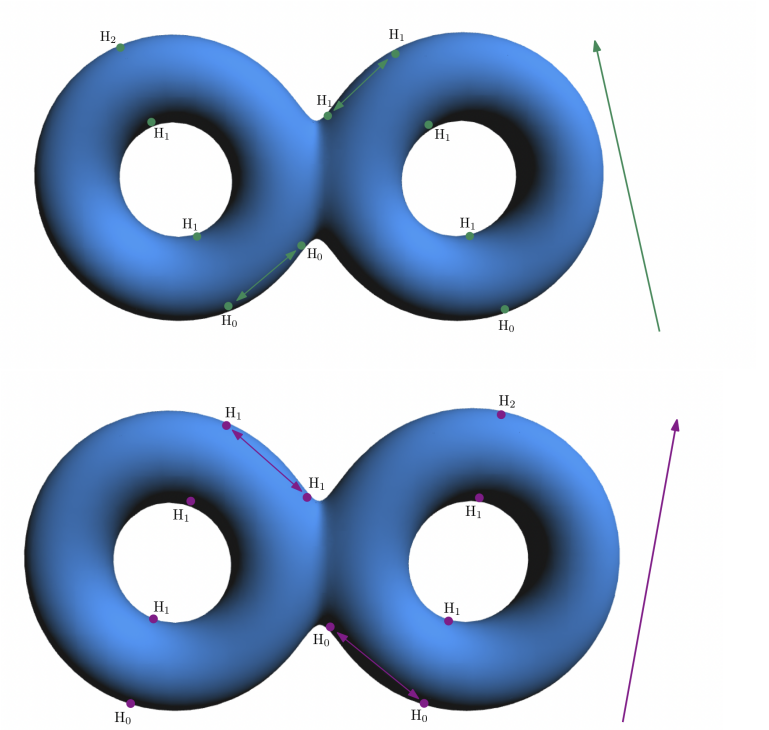


Figure 1.5: A 2-torus filtered by two infinitesimal perturbations of the vertical height function, together with their critical points and labels to indicate the dimension of the homology group that they affect. Paired critical points correspond to bounded intervals in the associated barcodes. Here, the vertical height function is Morse and the critical points evolve smoothly. However, the pairing between critical points is not constant, nor their homological dimensions. Therefore, the barcode valued map is not smooth at the vertical direction.

machine learning algorithms. A prototypical example of such a map is the persistence image of [AEK+17][22], which we study in Section 1.7.1. Other maps even have $\mathcal{N} = \mathbb{R}$ as codomain, and they are meant to be used as loss terms in optimization tasks [CNBW19, BGGSSG20a, HLSC19]. Many examples of such vectorizations and loss functions are part of the wide class of *linear representations*, which we study in Section 1.7.2. In Section 1.7.4, we study an important example of non-linear loss, namely the bottleneck distance to a fixed barcode, which we believe can be of interest in the context of inverse problems. The machinery developed in this section is likely to be adaptable to other examples of maps on barcodes, however the purpose of the section is to provide a proof of concept rather than an exhaustive treatment.

[22] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research*, 18(1):218–252, 2017

1.7.1 *The differentiability of persistence images*

Recall that Bar is equipped with the bottleneck topology. Let Bar_n be the subset of Bar containing the barcodes with n infinite intervals. In particular, Bar_0 is the set of barcodes whose intervals are bounded.

Proposition 1.7.1. *The set of path connected components of Bar is enumerable. More precisely,* $\pi_0(Bar) = \{Bar_n\}_{n=0}^{+\infty}$.

Proof. Since $Bar = \bigsqcup_{n=0}^{+\infty} Bar_n$, we only need to prove that each Bar_n is a

maximal connected subset of Bar. First note that Bar_n is path connected, as we can always move n infinite intervals to n other ones continuously, and similarly move the bounded off-diagonal intervals to the diagonal. We now prove the maximality of Bar_n. Let $A \subseteq Bar \backslash Bar_n$ be non-empty. Any element in A has infinite bottleneck distance to any element in Bar_n, since their numbers of infinite intervals are different. Therefore, $A \cup Bar_n$ cannot be path-connected, and so Bar_n is maximal. □

We view the persistence image as a map $V : Bar_0 \to \mathbb{R}^{n^2}$ for some *discretization step* $n \in \mathbb{N}$:

Definition 1.7.2. Let $D \in Bar_0$. We fix a *weighting function* $\omega : \mathbb{R} \to \mathbb{R}$ that is zero at the origin. For $(b, d) \in \mathbb{R}^2$, consider the Gaussian

$$g_{b,d} : (x, y) \in \mathbb{R}^2 \mapsto \frac{1}{2\pi\sigma^2} e^{-[(x-b)^2+(y-(d-b))^2]/2\sigma^2}$$

for some fixed variance $\sigma > 0$. The *persistence surface* associated to D is the map

$$\rho_D : (x, y) \in \mathbb{R}^2 \mapsto \sum_{(b,d)\in D} \omega(d - b) g_{b,d}(x, y).$$

Given a square $B \subset \mathbb{R}^2$, we subdivide it into n^2 regular squares $B_{k,l}$ for $1 \leqslant k, l \leqslant n$. Then we define the *persistence image* of D to be the histogram

$$V_{B,n} : D \in Bar_0 \mapsto \left(\int_{(x,y)\in B_{k,l}} \rho_D(x, y) dx dy \right)_{1\leqslant k,l\leqslant n} \in \mathbb{R}^{n^2}$$

Proposition 1.7.3. *If ω is C^r over $\mathbb{R}^2$ for some integer $r \in \mathbb{N}$, then $V_{B,n}$ is r-differentiable everywhere in Bar_0.*

Proof. The maps $(b, d) \in \mathbb{R}^2 \mapsto \int_{(x,y)\in B_{k,l}} g_{b,d}(x, y) dx dy \in \mathbb{R}$ are C^∞ for any fixed box $B_{k,l}$. For any space of ordered barcodes $\mathbb{R}^{2m} \times \mathbb{R}^0$ and any $\tilde{D} = (b_1, d_1, ..., b_m, d_m) \in \mathbb{R}^{2m} \times \mathbb{R}^0$,

$$V_{B,n}(Q_{m,0}(\tilde{D})) = \left(\sum_{1\leqslant i\leqslant m} \omega(d_i - b_i) \int_{(x,y)\in B_{k,l}} g_{b_i,d_i}(x, y) dx dy \right)_{1\leqslant k,l\leqslant n} \in \mathbb{R}^{n^2},$$

which is C^r at every $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^0$. □

In [AEK+17], the weighting function ω is chosen to be the ramp function $\omega_t : \mathbb{R} \mapsto \mathbb{R}$ defined as

$$\omega_t(u) = \begin{cases} 0 \text{ if } u \leqslant 0 \\ \dfrac{u}{t} \text{ if } 0 \leqslant u \leqslant t \\ 1 \text{ if } t \leqslant u \end{cases} \tag{1.26}$$

for some parameter $t > 0$. Thus, the ramp function is differentiable everywhere except at 0 and t. This implies that the persistence image $V_{B,n}$ is nowhere differentiable, as every neighborhood of a barcode always contains some neighborhood of the diagonal Δ. Thanks to Proposition 1.7.3, this issue can be resolved by taking any C^r approximation of the ramp function, which makes the persistence image r-differentiable over Bar_0.

1.7.2 *Linear representations of barcodes*

The analysis of persistence images in the previous section can be generalized to the following wide class of vectorizations:

Definition 1.7.4. Let $\phi : \mathbb{R}^2 \to \mathbb{R}^k$, $\psi : \mathbb{R} \to \mathbb{R}^k$ and $\omega : \mathbb{R} \to \mathbb{R}$ be continuous maps such that $\omega(0) = 0$. The associated *linear representation* is the map

$$V : D \in Bar \longmapsto \sum_{(b,d)\in D \text{ bounded}} \omega(d-b)\phi(b,d) + \sum_{(v,+\infty)\in D} \psi(v) \in \mathbb{R}^k.$$

Properties of linear representations valued in Banach spaces such as continuity, lipschitzness and stochastic convergence are analyzed in [DC19, DL20][23]. Many vectorizations in the litterature are linear representations, e.g. persistence images [AEK+17] and its variations [CWRW15, KHF16, RHBK15], persistence silhouettes [CFL+14] and weighted Betti curves [Ume17].

When $k = 1$, a linear representation may be viewed as a loss function on persistence diagrams. The total persistence in Example 1.3.11, and more generally the q-Wasserstein distance to the empty diagram, are such loss functions. In addition, the structure elements of [HKN19, Definition 9] form a wide class of parametrized linear losses and linear representations that can be optimised.

In all these examples, the maps ϕ, ψ and ω are not necessarily smooth by design, see e.g. the ramp function in Eq. (1.26) for persistence images, but one can always replace them with smooth approximations. We then get r-differentiable maps on barcodes, as expressed in the following result.

Proposition 1.7.5. *If the maps ϕ, ψ are C^r on generic subsets of $\mathbb{R}^2$ containing the diagonal Δ, and if ω is C^r on a generic subset of $\mathbb{R}$ containing the origin, then the associated linear representation V is generically r-differentiable. Whenever ϕ, ψ and ω are in fact C^r everywhere, then V is r-differentiable everywhere.*

Proof. The subspace of barcodes whose intervals avoid the set of non-differentiability of ϕ, ψ and ω is clearly generic in Bar. Let D be a

[23] Vincent Divol and Frédéric Chazal. The density of expected persistence diagrams and its kernel based estimation. *Journal of Computational Geometry*, 10(2):127–153, 2019; and Vincent Divol and Théo Lacombe. Understanding the topology and the geometry of the space of persistence diagrams via optimal partial transport. *Journal of Applied and Computational Topology*, pages 1–53, 2020

barcode therein. For any space of ordered barcodes $\mathbb{R}^{2m} \times \mathbb{R}^n$ and pre-image $\tilde{D} = [(b_i, d_i)_{i=1}^m, (v_j)_{j=1}^n] \in \mathbb{R}^{2m} \times \mathbb{R}^n$ of D, we have

$$V(Q_{m,n}(\tilde{D})) = \sum_{1 \leqslant i \leqslant m} \omega(d_i - b_i)\phi(b_i, d_i) + \sum_{1 \leqslant j \leqslant n} \psi(v_j),$$

which is C^r in a neighborhood of $\tilde{D}$. □

Let us consider an everywhere r-differentiable linear representation V, and a barcode valued map B on a simplicial complex, which is (generically) differentiable (Theorem 1.4.9). Using the chain rule 1.3.14, the composition $V \circ B$ is then itself (generically) differentiable, hence amenable to gradient descent based optimisation.

1.7.3 *Semi-algebraic and subanalytic functions on barcodes*

We consider another important class of examples arising from loss functions on barcodes that restrict to semi-algebraic maps on the spaces of ordered barcodes. The subanalytic and definable counterparts are analogously defined and the results of this section are valid in these situations as well. See also [CCG+21] for a full treatment of semi-algebraic loss functions in persistence.

Mathieu Carriere, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hariprasad Kannan, and Yuhei Umeda. Optimizing persistent homology based functions. In *International Conference on Machine Learning*, pages 1294–1303. PMLR, 2021

Definition 1.7.6. We say that a map $V : Bar \to \mathbb{R}$ is *semi-algebraic* if all the precompositions $V \circ Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to \mathbb{R}$ are semi-algebraic.

A prototypical example of semi-algebraic loss on barcodes is the distance to a target barcode D_0:

$$d_q(D_0, .) : D \in Bar \mapsto d_q(D_0, D) \in \mathbb{R} \cup \{+\infty\}.$$

Here, d_q is the q-th Wasserstein distance on barcodes for any $q \in \mathbb{R}_+^*$ as defined in Eq. (1.7), and d_∞ is the bottleneck distance.

Proposition 1.7.7. *For any target barcode D_0 and non-negative number $q \in \mathbb{R}_+^*$, the map $d_q(D_0, .) : Bar \to \mathbb{R}$ is semi-algebraic.*

Proof. We consider the case where $q = \infty$, as the same line of arguments works for arbitrary Wasserstein metrics, and rewrite $d_q(D_0, .)$ as d_{D_0} for simplicity. Let $m, n \in \mathbb{N}$. We assume that n is the number of infinite intervals in D_0, as otherwise the map $d_{D_0} \circ Q_{m,n} : \mathbb{R}^{2m} \times \mathbb{R}^n \to \mathbb{R}$ takes infinite value everywhere. Then, $d_{D_0} \circ Q_{m,n}$ can be expressed as a minimum of finitely many cost functions, $\min c(\gamma_{m,n})(.)$, each of which is defined in terms of a fixed partial matching $\gamma_{m,n}$ of coordinates in $\mathbb{R}^{2m} \times \mathbb{R}^n$ with interval endpoints of D_0. As a point-wise maximum of finitely many absolute values, each cost function $c(\gamma_{m,n})(.)$ is semi-algebraic, and so $d_{D_0} \circ Q_{m,n}$ is semi-algebraic. □

Semi-algebraic functions V on barcodes are particularly useful in the context of optimisation when pre-composed with a semi-algebraic parametrization of filter functions $F : \mathcal{M} \to \mathbb{R}^K$ on a fixed simplicial complex K. Indeed, composition preserves semi-algebraicity, and so from Remark 1.4.25 the loss function given by the composition

$$\mathcal{L} : \mathcal{M} \xrightarrow{F} \mathbb{R}^K \xrightarrow{\mathrm{PH}_p} Bar \xrightarrow{V} \mathbb{R} \tag{1.27}$$

is a semi-algebraic map. Then [DDKL20, Corollary 5.9][24] guarantees that the well-known stochastic gradient descent (SGD) algorithm converges almost surely to critical points of $\mathcal{L}$.[25]

This guarantee can be applied to various optimisation problems. When choosing the Rips parametrization F of point clouds as in Section 1.5.2, minimizing the loss $\mathcal{L} = d_q(D_0, .) \circ \mathrm{PH}_p \circ F$ amounts to solving the problem of point cloud inference originally proposed in [GHO16][26], see [GNDS20a][27] for implementations. Besides, from Section 1.5.4, for F the parametrization of all filter functions on a fixed simplicial complex and an adequate target barcode D_0, the minimisation of $\mathcal{L}$ yields an approach to function simplification. However, when F is not semi-algebraic, typically in the continuous setting developped in Section 1.6, and more generally for an arbitrary barcode valued map $B : \mathcal{M} \to Bar$, it is unclear how to perform full-fledged continuous gradient descent to minimize

$$\mathcal{L} : \mathcal{M} \xrightarrow{B} Bar \xrightarrow{d_q(D_0,.)} \mathbb{R} \,. \tag{1.28}$$

While implementing a solution to this problem is beyond the scope of this paper, it serves as a motivation for the next section where we show that the bottleneck distance to D_0 is generically ∞-differentiable, as then the chain rule of Proposition 1.3.14 enables the use of gradient descent.

1.7.4 *The bottleneck distance to a diagram*

For simplicity, we denote the bottleneck distance to a fixed barcode D_0 by:

$$d_{D_0} : D \in Bar \longmapsto d_\infty(D, D_0) \in \mathbb{R}.$$

For ease of exposition, we start with the special case where $D_0 = \Delta^\infty$ is the empty diagram (the diagonal Δ with infinite multiplicity). We then provide the details of the analysis for the general, technically more involved case of an arbitrary fixed barcode D_0.

Recall that $d_{\Delta^\infty}(D) = +\infty$ for any diagram $D \in Bar$ with infinite bars. Consequently, we consider the restriction of d_{Δ^∞} to the subset Bar_0 introduced in Section 1.7.1. This restriction is valued in the real line:

[24] Damek Davis, Dmitriy Drusvyatskiy, Sham Kakade, and Jason D Lee. Stochastic subgradient method converges on tame functions. *Foundations of Computational Mathematics*, 20(1):119–154, 2020

[25] The loss $\mathcal{L}$ must also be locally Lipschitz for this result to hold. By the Stability Theorem [CSEH07] , PH_p is Lipschitz continuous, hence this additional mild requirement is met whenever F and V are locally Lipschitz (for instance when $V = d_q(D_0, .)$ is the distance to a fixed barcode).

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

[26] Marcio Gameiro, Yasuaki Hiraoka, and Ippei Obayashi. Continuation of point clouds via persistence diagrams. *Physica D: Nonlinear Phenomena*, 334:118–132, 2016

[27] Rickard Brüel Gabrielsson, Bradley J Nelson, Anjan Dwaraknath, and Primoz Skraba. A topology layer for machine learning. In *International Conference on Artificial Intelligence and Statistics*, pages 1553–1563. PMLR, 2020

$d_{\Delta^\infty} : Bar_0 \to \mathbb{R}$. Consider the set Bar_Δ of barcodes which admit a unique point at maximal distance to the diagonal Δ:

$$Bar_\Delta := \left\{ D \in Bar_0 \;\middle|\; \#\mathrm{argmax}_{(b,d)\in D} \frac{|d-b|}{2} = 1 \right\}. \tag{1.29}$$

For $D \in Bar_\Delta$, we let $(\bar{b}_D, \bar{d}_D) \in D$ be the unique interval in the set $\mathrm{argmax}_{(b,d)\in D} \frac{|d-b|}{2}$.

Proposition 1.7.8. *Bar_Δ is generic in Bar_0. Moreover, given $D \in Bar_\Delta$, for $\varepsilon > 0$ small enough, any D' at bottleneck distance less than ε from D satisfies $d_{\Delta^\infty}(D') = \frac{|\bar{d}_{D'} - \bar{b}_{D'}|}{2}$ and $\|(\bar{b}_{D'}, \bar{d}_{D'}) - (\bar{b}_D, \bar{d}_D)\|_\infty < \varepsilon$.*

Proof. Given $D \in Bar_0$, consider the set $\mathrm{argmax}_{(b,d)\in D} \frac{|d-b|}{2}$. If this set is not a singleton, then we can move infinitesimally one of its elements away from the diagonal, so as to get a diagram in Bar_Δ. Thus, Bar_Δ is dense in Bar_0. Let now $D \in Bar_\Delta$, and let δ be the second maximal distance to the diagonal:

$$\delta := \max_{(b,d)\in D\setminus\{(\bar{b}_D,\bar{d}_D)\}} \frac{|d-b|}{2}$$

and $\alpha := \frac{|\bar{d}_D - \bar{b}_D|}{2} - \delta > 0$. Take $\varepsilon \in \left(0, \frac{\alpha}{4}\right)$. If D' is at bottleneck distance less than ε from D, all the points of D' are within distance less than ε either from the diagonal or from an off-diagonal point of D. As we have picked $\varepsilon < \frac{\alpha}{4}$, there is a unique off-diagonal point $(\bar{b}', \bar{d}')$ of D' that is within distance less than ε from $(\bar{b}_D, \bar{d}_D)$, and it must be the unique furthest point from Δ in D'. So indeed $D' \in Bar_\Delta$ and $(\bar{b}_{D'}, \bar{d}_{D'}) = (\bar{b}', \bar{d}')$. Therefore, Bar_Δ is open, which concludes the proof. □

Not surprisingly, d_{Δ^∞} is smooth at every $D \in Bar_\Delta$, with partial derivatives related to the ones of the map $(\bar{b}_D, \bar{d}_D) \mapsto \frac{|\bar{d}_D - \bar{b}_D|}{2}$.

Proposition 1.7.9. *For any $D \in Bar_\Delta$,*

(i) d_{Δ^∞} is ∞-differentiable at D, and

(ii) for any $m \in \mathbb{N}$ and $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^0$ such that $Q_{m,0}(\tilde{D}) = D$, there are exactly two non-zero components in the gradient $\nabla_{\tilde{D}}(d_{\Delta^\infty} \circ Q_{m,0})$, one with value $\frac{1}{2}$ and the other with value $-\frac{1}{2}$.

Proof. Let $m \in \mathbb{N}$ and $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^0$ be such that $Q_{m,0}(\tilde{D}) = D$. Without loss of generality, we can write $\tilde{D} = (\bar{b}_D, \bar{d}_D, b_2, d_2, ..., b_m, d_m)$ where (b_i, d_i) is distinct from $(\bar{b}_D, \bar{d}_D)$ for all $2 \leqslant i \leqslant m$. By Proposition 1.3.2, $Q_{m,0}$ is continuous. Therefore, by Proposition 1.7.8, there is an open neighborhood U of $\tilde{D}$, such that for any $\tilde{D}' = (\bar{b}_{D'}, \bar{d}_{D'}, b'_2, d'_2, ..., b'_m, d'_m) \in U$, $Q_{m,0}(\tilde{D}')$ is in Bar_Δ and $d_{\Delta^\infty}(Q_{m,0}(\tilde{D}')) = \frac{|\bar{d}_{D'} - \bar{b}_{D'}|}{2} > 0$. Assertions (i) and (ii) follow. □

Next we generalize Proposition 1.7.9, namely we show the generic differentiability of the bottleneck distance $d_{D_0} : Bar \to \mathbb{R} \cup \{+\infty\}$ to an arbitrary fixed diagram $D_0 \in Bar$.

Throughout, we denote by Δ_ϵ the set of elements in $\mathbb{R}^2$ that are at distance less than $\epsilon > 0$ to the diagonal Δ. We equip, for the rest of this section only, the spaces of ordered barcodes with the supremum norm $\|.\|_\infty$ rather than the Euclidean norm. Note that (the proof of) Proposition 1.3.2 ensures that the quotient maps $Q_{m,n}$ are 1-Lipchitz with respect to the metrics in place. We denote by $\mathcal{B}(.,*)$ the ball centered at . with radius $*$ with respect to the supremum norm or bottleneck metric depending on the context.

Proposition 1.7.10. *Let $D_0 \in Bar$ and n be the number of infinite bars in D_0. For generic $D \in Bar_n$, d_{D_0} is ∞-differentiable at D. Moreover, for any $m \in \mathbb{N}$ and $\tilde{D} \in \mathbb{R}^{2m} \times \mathbb{R}^n$ such that $Q_{m,n}(\tilde{D}) = D$, exactly one of the following possibilities holds:*

(i) either the gradient $\nabla_{\tilde{D}}(d_{D_0} \circ Q_{m,n})$ has exactly two non-zero components, one with value $\frac{1}{2}$ and the other with value $-\frac{1}{2}$; or

(ii) the gradient $\nabla_{\tilde{D}}(d_{D_0} \circ Q_{m,n})$ has a unique non-zero component with value 1 or -1.

Proposition 1.7.10 states the generic smoothness of d_{D_0}. We first observe that all the compositions $d_{D_0} \circ Q_{m,n}$ are smooth on a generic subset of $\mathbb{R}^{2m+n}$.

Lemma 1.7.11. *For every $m \in \mathbb{N}$, the map*

$$d_{D_0} \circ Q_{m,n} : \mathbb{R}^{2m+n} \to \mathbb{R}$$

is generically smooth, with gradients that are either 0 or as in (i) or (ii) of Proposition 1.7.10.

Proof. Let $m \in \mathbb{N}$. Define an *ordered matching* $\tilde{\gamma} : \mathbb{R}^{2m+n} \to \mathbb{R}^{2m+n}$ to be an affine map whose first m pairs of coordinate functions (resp. last n coordinate functions) are of the form $\tilde{D} := [(b_i, d_i)_{i=1}^m, (v_j)_{j=1}^n] \mapsto (b_i, d_i) - (b_{0,i}, d_{0,i})$ where $(b_{0,i}, d_{0,i})$ is either an off-diagonal point in D_0 or $(b_{0,i}, d_{0,i}) = (\frac{b_i+d_i}{2}, \frac{b_i+d_i}{2})$ is the orthogonal projection of (b_i, d_i) onto Δ (resp. are of the form $\tilde{D} \mapsto v_j - v_{0,j}$ for some infinite interval $(v_{0,j}, +\infty)$ in D_0). We further require that the collection of intervals $(b_{0,i}, d_{0,i})$ (resp. $(v_{0,j}, +\infty)$) involved in this way are distinct elements in D_0. We denote by $D_0(\tilde{\gamma})$ the set of bounded off-diagonal intervals $(b_0, d_0) \in D_0$ that are not in the collection $\{b_{0,i}, d_{0,i}\}_{i=1}^m$.

Since the maximum of smooth functions over $\mathbb{R}^{2m+n}$ is smooth[28] on a generic subset of $\mathbb{R}^{2m+n}$, the map

$$\tilde{c}(\tilde{\gamma}) : \tilde{D} \in \mathbb{R}^{2m+n} \mapsto \max(\|\tilde{\gamma}(\tilde{D})\|_\infty, \{\frac{|d_0 - b_0|}{2}\}_{(b_0,d_0)\in D_0(\tilde{\gamma})}) \in \mathbb{R} \quad (1.30)$$

[28] This is the same argument as in the proof of Proposition 1.4.9. Namely, the set where at least two of the smooth functions involved in the maximum are equal is closed, and therefore the boundary of this set has generic complement. On this complement the maximum of the smooth functions locally equals a unique smooth function.

is itself C^∞ on a generic subset of $\mathbb{R}^{2m+n}$, with gradients either equal to 0 or as in (i) or (ii) of Proposition 1.7.10. Let $\tilde{\Gamma}_m$ be the set of ordered matchings $\tilde{\gamma} : \mathbb{R}^{2m+n} \to \mathbb{R}^{2m+n}$, which is non-empty and finite. Then the map

$$\tilde{d}_{D_0,m} : \tilde{D} \in \mathbb{R}^{2m+n} \mapsto \min_{\tilde{\gamma} \in \tilde{\Gamma}_m} \tilde{c}(\tilde{\gamma})(\tilde{D}) \in \mathbb{R}$$

is C^∞ on a generic subset of $\mathbb{R}^{2m+n}$, with gradients either equal to 0 or as in (i) or (ii) of Proposition 1.7.10.

We will be done if we can show that the two maps $d_{D_0} \circ Q_{m,n}$ and $\tilde{d}_{D_0,m}$ are equal over $\mathbb{R}^{2m+n}$. Fix an ordered barcode $\tilde{D} \in \mathbb{R}^{2m+n}$ and let $D := Q_{m,n}(\tilde{D})$. Let $\tilde{\gamma} : \mathbb{R}^{2m+n} \to \mathbb{R}^{2m+n}$ be an ordered matching. The components of $\tilde{\gamma}$ determine a matching γ between D and D_0, sending (b_i, d_i) onto $(b_{0,i}, d_{0,i})$ and $(v_j, +\infty)$ onto $(v_{0,j}, +\infty)$. By definition of the cost of a matching 1.2.10 and Equation (1.30), we have $c(\gamma) = \tilde{c}(\tilde{\gamma})(\tilde{D})$. This yields $\tilde{d}_{D_0,m}(\tilde{D}) \geqslant d_{D_0}(D) = d_{D_0} \circ Q_{m,n}(\tilde{D})$. Conversely, among the optimal matchings from D to D_0, it is always possible to find one that sends off-diagonal points of D (and D_0) on the diagonal only by orthogonal projection. This allows us to lift γ at the level of $\tilde{D}$ and to define an ordered matching $\tilde{\gamma}$ such that $\tilde{c}(\tilde{\gamma})(\tilde{D}) = c(\gamma)$. This yields $\tilde{d}_{D_0,m}(\tilde{D}) \leqslant d_{D_0}(D) = d_{D_0} \circ Q_{m,n}(\tilde{D})$ and therefore $d_{D_0} \circ Q_{m,n} = \tilde{d}_{D_0,m}$ on $\mathbb{R}^{2m+n}$. □

We cannot directly use Lemma 1.7.11 to prove Proposition 1.7.10. As a matter of fact, by the definition of ∞-differentiability (Definition 1.3.10), Proposition 1.7.10 is asking that for generic $D \in Bar_n$ *all* the maps $d_{D_0} \circ Q_{m,n}$, for varying $m \in \mathbb{N}$, should be smooth at pre-images of D. However, Lemma 1.7.11 only guarantees that the maps $d_{D_0} \circ Q_{m,n}$, taken individually, are smooth over generic subsets of $\mathbb{R}^{2m+n}$, and it is not clear a priori how to glue at the level of barcodes these generic subsets lying in different spaces of ordered barcodes $\mathbb{R}^{2m+n}$. In order to leverage Lemma 1.7.11, we devise intermediate results that infer the smoothness of the maps $d_{D_0} \circ Q_{m',n}$ from the knowledge of the smoothness of a well-chosen map $d_{D_0} \circ Q_{m,n}$. The high-level intuition of each of these intermediate steps is as follows:

1. Infinitesimal perturbations of a given diagram D can be understood as infinitesimal moves of the off-diagonal points of D, together with appearances of small intervals from the diagonal. In Lemma 1.7.12, we devise a generic condition on D ensuring that these new small off-diagonal intervals appearing when perturbing D do not play any role in the bottleneck distance to D_0.

2. Given a barcode D, we take a pre-image $\tilde{D}_m \in Q_{m,n}^{-1}(D)$ of D which is minimal in the sense that its pairs of adjacent components are not trivial, i.e not of the form (b, b). In other words, $\tilde{D}_m$ is an ordering of the endpoints of off-diagonal intervals appearing in D

without extra pairs (b, b) lying on the diagonal. Up to an infinitesimal perturbation of $\tilde{D}_m$, Lemma 1.7.11 ensures that $d_{D_0} \circ Q_{m,n}$ is smooth in an open neighborhood of $\tilde{D}_m$. It is easy to observe that for any other pre-image $\tilde{D}_{m'}$ of D, the components of the ordered barcode $\tilde{D}_{m'}$ only differ with those of $\tilde{D}_m$ by the addition of trivial pairs of the form (b, b). According to the previous item, those trivial pairs do not play any role when computing the bottleneck distance to D_0. Therefore, since $d_{D_0} \circ Q_{m,n}$ is smooth in a neighborhood of $\tilde{D}_m$, the map $d_{D_0} \circ Q_{m',n}$ is itself smooth in an open neighborhood of $\tilde{D}_{m'}$. We make these intuitions rigorous in Lemma 1.7.13.

3. The previous arguments allow us to construct open balls $\mathcal{B}(\tilde{D}_{m'}, \epsilon)$ of the same radius $\epsilon > 0$ around all pre-images $\tilde{D}_{m'} \in \mathbb{R}^{2m'+n}$ of a generic diagram $D \in Bar$ over which all maps $d_{D_0} \circ Q_{m',n}$ are smooth. To conclude that d_{D_0} itself is ∞-differentiable in a neighborhood of D, we show in Lemma 1.7.14 that the ϵ-bottleneck ball around D is covered by the union of the images of the balls $\mathcal{B}(\tilde{D}_{m'}, \epsilon)$.

Let $\hat{Bar}$ be the set of barcodes $D \in Bar_n$ such that no intervals of D_0 is at distance $d_\infty(D, D_0)$ to its diagonal projection. It is easy to check that $\hat{Bar}$ is generic in Bar_n. When perturbing a given barcode D infinitesimally, an arbitrary number of new off-diagonal points may appear from the diagonal. We show that, for $D \in \hat{Bar}$, these new off-diagonal intervals can be disregarded when computing the bottleneck distance to D_0.

Lemma 1.7.12. *Let $D \in \hat{Bar}$. There exists $\epsilon > 0$ such that for any barcode D' which is ϵ-close to D we have that $d_\infty(D', D_0) > \epsilon$, and there exists an optimal matching from D' to D_0 sending $D' \cap \Delta_\epsilon$ (i.e. those points of D' that are ϵ-close to the diagonal), onto the diagonal Δ.*

Proof. Let $D \in \hat{Bar}$. Denote by α the minimal gap $|\frac{|d_0 - b_0|}{2} - d_\infty(D, D_0)|$ between the distance of off-diagonal intervals (b_0, d_0) of D_0 from their diagonal projections and $d_\infty(D, D_0)$. Since $D \in \hat{Bar}$, α is strictly positive.

We have $d_\infty(D, D_0) > 0$ as otherwise $d_\infty(D, D_0) = 0$ would imply that $D \notin \hat{Bar}$ (as the distance from a diagonal element of D_0 to its diagonal projection would also be 0), so we can pick $\epsilon > 0$ such that

$$\epsilon < \min(\frac{d_\infty(D, D_0)}{2}, \frac{\alpha}{2}).$$

We now prove that the conclusion of the Lemma holds in the bottleneck ball $\mathcal{B}(D, \epsilon)$. Let $D' \in \mathcal{B}(D, \epsilon)$. Since $\epsilon < \frac{d_\infty(D,D_0)}{2}$, we have $d_\infty(D', D_0) > \epsilon$.We assume, seeking contradiction, that there is no optimal matching from D' to D_0 that sends all points of $D' \cap \Delta_\epsilon$ onto Δ.

We restrict our attention to the set $\Gamma^*(D', D_0)$ of optimal matchings from D' to D_0 that are allowed to send off-diagonal points of D' and

D_0 to the diagonal only by orthogonal projections. This set is finite and non-empty. We define the Δ-*degree* of a matching $\gamma \in \Gamma^*(D', D_0)$ to be the number of off-diagonal points of D' and D_0 that are sent to their diagonal projections, and take γ with maximal Δ-degree. By assumption, there exists an off-diagonal point $(b', d') \in D' \cap \Delta_\epsilon$ sent to some off-diagonal point $(b_0, d_0) \in D_0$. Recall that $|\frac{|d_0-b_0|}{2} - d_\infty(D, D_0)| \geqslant \alpha$. We divide the analysis into two cases: either $\frac{|d_0-b_0|}{2} \geqslant d_\infty(D, D_0) + \alpha$, or $\frac{|d_0-b_0|}{2} \leqslant d_\infty(D, D_0) - \alpha$.

In the case where $\frac{|d_0-b_0|}{2} \geqslant d_\infty(D, D_0) + \alpha$, we have:

$$\begin{aligned}
\|(b_0, d_0) - (b', d')\|_\infty &= \|[(b_0, d_0) - (\frac{b'+d'}{2}, \frac{b'+d'}{2})] - [(b', d') - (\frac{b'+d'}{2}, \frac{b'+d'}{2})]\|_\infty \\
&\geqslant \|(b_0, d_0) - (\frac{b'+d'}{2}, \frac{b'+d'}{2})\|_\infty - \|(b', d') - (\frac{b'+d'}{2}, \frac{b'+d'}{2})\|_\infty \\
&\geqslant \frac{|d_0 - b_0|}{2} - \frac{|d' - b'|}{2} \\
&\geqslant d_\infty(D, D_0) + \alpha - \epsilon \\
&\geqslant d_\infty(D', D_0) + \alpha - 2\epsilon \\
&> d_\infty(D', D_0)
\end{aligned}$$

where the first inequality holds by the triangle inequality, the second from the fact that a minimizer of the distance from (b_0, d_0) to the diagonal is the orthogonal projection of (b_0, d_0) onto Δ, the third by assumption on (b_0, d_0) and (b', d'), the fourth by the triangle inequality and the last one by $\epsilon < \frac{\alpha}{2}$. This yields a contradiction as γ is optimal and its cost may not exceed $d_\infty(D', D_0)$.

Consider now the case where $\frac{|d_0-b_0|}{2} \leqslant d_\infty(D, D_0) - \alpha$. On the one hand, by the triangle inequality and by the fact that $\epsilon < \alpha$:

$$\frac{|d_0 - b_0|}{2} \leqslant d_\infty(D, D_0) - \alpha \leqslant d_\infty(D', D_0) + \epsilon - \alpha < d_\infty(D', D_0)$$

On the other hand, since $\epsilon < \frac{d_\infty(D,D_0)}{2}$,

$$\frac{|d' - b'|}{2} \leqslant \epsilon < \frac{d_\infty(D, D_0)}{2} \leqslant \frac{d_\infty(D', D_0) + \epsilon}{2} < d_\infty(D', D_0),$$

where the last inequality comes from the fact that $\epsilon < d_\infty(D', D_0)$ because $\epsilon < \frac{d_\infty(D', D_0)+\epsilon}{2}$. To sum up, both quantities $\frac{|d_0-b_0|}{2}$ and $\frac{|d'-b'|}{2}$ are upper-bounded by $d_\infty(D', D_0)$. Modifying γ by sending (b_0, d_0) and (b', d') to their diagonal projections, we obtain a matching in $\Gamma^*(D', D_0)$ with Δ-degree strictly higher than that of γ, which contradicts the maximality of the Δ-degree of γ. □

We say that an ordered barcode $\tilde{D}_m = [(b_i, d_i)_{i=1}^m, (v_j)_{j=1}^n] \in \mathbb{R}^{2m+n}$ is *minimal* if $b_i \neq d_i$ for $1 \leqslant i \leqslant m$. This terminology is justified by the fact that the image $D := Q_{m,n}(\tilde{D}_m) \in Bar_n$ contains exactly m bounded-off

diagonal intervals and n unbounded ones, and therefore any other pre-image $\tilde{D}_{m'} \in \mathbb{R}^{2m'+n}$ of D must lie in a space of ordered barcodes of dimension at least $2m+n$ (i.e. $m' \geqslant m$). We show that under suitable assumptions, the differentiability of all the maps $d_{D_0} \circ Q_{m',n}$ at pre-images $\tilde{D}_{m'}$ of D can be inferred from the differentiability of $d_{D_0} \circ Q_{m,n}$ at the minimal pre-image $\tilde{D}_m$.

Lemma 1.7.13. *For every $m \in \mathbb{N}$, the set of minimal ordered barcodes in $\mathbb{R}^{2m+n}$ is open. Moreover, given a minimal $\tilde{D}_m \in \mathbb{R}^{2m+n}$ with $D := Q_{m,n}(\tilde{D}_m) \in \hat{Bar}$, if $d_{D_0} \circ Q_{m,n}$ is C^∞ in an open neighborhood of $\tilde{D}_m$, then there is an $\epsilon > 0$ such that for all other pre-images $\tilde{D}_{m'}$ of D, the map $d_{D_0} \circ Q_{m',n}$ is C^∞ in $\mathcal{B}(\tilde{D}_{m'}, \epsilon)$, with gradients as in (i) or (ii) of Proposition 1.7.10.*

Proof. Is is clear that the set of minimal ordered barcodes in $\mathbb{R}^{2m+n}$ is open. We address the second part of the Lemma. Let $\tilde{D}_m \in \mathbb{R}^{2m+n}$ be a minimal ordered barcode such that $D := Q_{m,n}(\tilde{D}_m) \in \hat{Bar}$, and assume there is an open neighborhood U of $\tilde{D}_m$ within which $d_{D_0} \circ Q_{m,n}$ is C^∞. By continuity of the quotient map and from the fact that $\hat{Bar}$ is open, we can assume without loss of generality that $Q_{m,n}(U)$ is contained in $\hat{Bar}$.

For any other pre-image $\tilde{D}_{m'} \in \mathbb{R}^{2m'+n}$ of D, i.e. an ordered barcode such that $Q_{m',n}(\tilde{D}_{m'}) = D = Q_{m,n}(\tilde{D}_m)$, the first m' adjacent pairs of components of $\tilde{D}_{m'}$ must describe in an arbitrary order the m bounded off-diagonal points of D together with $m' - m$ trivial pairs of the form (b, b). The last n components of $\tilde{D}_{m'}$ must be in correspondance with the left endpoints of infinite intervals in D. In other words, the first $2m'$ components of $\tilde{D}_{m'}$ consist of a re-ordering of the first $2m$ components of $\tilde{D}_m$, together with $m' - m$ trivial pairs of the form (b, b). The last n components of $\tilde{D}_{m'}$ consist of a re-ordering of those of $\tilde{D}_m$.

To every pre-image $\tilde{D}_{m'}$ of D as above, we associate the linear projection $L_{m',m} : \mathbb{R}^{2m'+n} \to \mathbb{R}^{2m+n}$ that sends $\tilde{D}_{m'}$ to $\tilde{D}_m$ by re-arranging the m non trivial pairs of components and the n last components, and forgetting the $m' - m$ trivial pairs. Since $D \in \hat{Bar}$, Lemma 1.7.12 provides an $\epsilon > 0$ such that for any $D' \in \mathcal{B}(D, \epsilon)$, the points of D' that are in Δ_ϵ may be sent onto the diagonal when computing the bottleneck distance from D' to D_0, and furthermore they can be disregarded when computing $d_\infty(D', D_0)$. Therefore, using that the quotient map $Q_{m',n}$ is 1-Lipschitz, we know that for any pre-image $\tilde{D}_{m'}$ of D and $\tilde{D}'_{m'} \in \mathcal{B}(\tilde{D}_{m'}, \epsilon)$, the $m' - m$ pairs of components $\tilde{D}'_{m'}$ with persistence less than ϵ can be disregarded when computing $d_{D_0} \circ Q_{m',n}(\tilde{D}'_{m'})$. Formally, for every $m' \in \mathbb{N}$,

$$\forall \tilde{D}_{m'} \in Q^{-1}_{m',n}(D),\ \forall \tilde{D}'_{m'} \in \mathcal{B}(\tilde{D}_{m'}, \epsilon),\ d_{D_0} \circ Q_{m',n}(\tilde{D}'_{m'}) = d_{D_0} \circ Q_{m,n} \circ L_{m',m}(\tilde{D}'_{m'}). \tag{1.31}$$

Note that the maps $L_{m',m}$ are 1-Lipschitz. Therefore, we can reduce ϵ in order to ensure that $L_{m',m}(\mathcal{B}(\tilde{D}_{m'}, \epsilon)) \subset U$ for every pre-image $\tilde{D}_{m'}$ of D. Applying the chain rule on $d_{D_0} \circ Q_{m,n}$ and $L_{m',m}$ —which is an affine map hence C^∞— in Equation (1.31), we obtain that all the maps $d_{D_0} \circ Q_{m',n}$ are C^∞ in $\mathcal{B}(\tilde{D}_{m'}, \epsilon)$. Also by the chain rule, by definition of $L_{m',m}$, the components of the gradients of the maps $d_{D_0} \circ Q_{m',n}$ are a re-ordering of the components of the gradient of $d_{D_0} \circ Q_{m,n}$. By Lemma 1.7.11, the gradient of the latter is either 0 or as in (i) or (ii) of Proposition 1.7.10. However, the gradient of $d_{D_0} \circ Q_{m,n}$ being 0 at some elements $\tilde{D}'_m \in U$ would mean that the bottleneck distance $d_\infty(Q_{m',n}(\tilde{D}'_m), D_0)$ equals the distance of some off-diagonal interval (b_0, d_0) to its diagonal projection, which is impossible since $Q_{m',n}(\tilde{D}'_m) \in \hat{Bar}$. □

By means of Lemma 1.7.13, we can deduce at once the differentiability of all the maps $d_{D_0} \circ Q_{m',n}$ over balls of the same radius. We need a last result that connects these balls to an actual open neighborhood of D in Bar_n.

Lemma 1.7.14. *For any $D \in Bar_n$, there exists an $\epsilon > 0$ such that for every $m' \in \mathbb{N}$,*

$$Q^{-1}_{m',n}(\mathcal{B}(D, \epsilon)) \subseteq \bigcup_{\tilde{D}_{m'} \in R^{2m'+n}, Q_{m',n}(\tilde{D}_{m'}) = D} \mathcal{B}(\tilde{D}_{m'}, \epsilon).$$

Proof. Let $D \in Bar_n$, and $\eta > 0$ be less than all the pairwise distances between geometrically distinct off-diagonal points in D, and less than all the distances from off-diagonal points in D to the diagonal. We take $\epsilon > 0$ such that $\epsilon < \frac{\eta}{2}$. Let $D' \in \mathcal{B}(D, \epsilon)$. Then, for every off-diagonal point (b, d) of D, the number of (off-diagonal) points of D' lying in $\mathcal{B}((b, d), \epsilon)$ equals the multiplicity of (b, d) in D. Let us say that these points in D' are of *type* **(a)**. The points of D' that are not in the balls $\mathcal{B}((b, d), \epsilon)$, for (b, d) ranging over off-diagonal intervals of D, must be ϵ-close to the diagonal, and we say that these points are of type **(b)**. Note that we can accordingly characterize the components of a pre-image $\tilde{D}'_{m'} \in Q^{-1}_{m',n}(D')$: the pairs of components in $\tilde{D}'_{m'}$ must either be trivial (i.e of the form (b, b)), or equal to some off-diagonal point of type **(a)** or **(b)**. All off-diagonal points of D', of type **(a)** or **(b)**, counted with multiplicity, must appear as a pair in $\tilde{D}'_{m'}$.

Given such a pre-image $\tilde{D}'_{m'} \in Q^{-1}_{m',n}(\mathcal{B}(D, \epsilon))$ of D', we construct another ordered barcode $\tilde{D}_{m'} \in \mathbb{R}^{2m'+n}$ by modifying the components of $\tilde{D}'_{m'}$ at cost less than ϵ (i.e such that $\|\tilde{D}'_{m'} - \tilde{D}_{m'}\|_\infty < \epsilon$) as follows:

- The last n components of $\tilde{D}'_{m'}$ parametrize the left endpoints of infinite intervals in D'. We change them at cost less than ϵ into the left endpoints of infinite intervals in D.

- If a pair (b', d') among the first m' pairs of components of $\tilde{D}'_{m'}$ is of type **(a)**, it is ϵ-close to a unique off-diagonal point (b, d) of D. We change it into (b, d).
- If a pair (b', d') among the first m' pairs of components of $\tilde{D}'_{m'}$ is of type **(b)**, it is ϵ-close to the diagonal. We transform it into $(\frac{b'+d'}{2}, \frac{b'+d'}{2})$.
- The remaining pairs in the first m' pairs of components of $\tilde{D}'_{m'}$ must be trivial, and we leave them unchanged.

In this way, we have constructed an ordered barcode $\tilde{D}_{m'}$ such that $\tilde{D}'_{m'} \in \mathcal{B}(\tilde{D}_{m'}, \epsilon)$ and also, by construction, $\tilde{D}_{m'}$ is a pre-image of D, i.e $Q_{m',n}(\tilde{D}_{m'}) = D$. □

We are now ready to prove Proposition 1.7.10.

Proof of Proposition 1.7.10. Consider the set of barcodes $D \in Bar_n$ that admit an open neighborhood within which d_{D_0} is ∞-differentiable. By definition, this set is open in Bar_n, and we are left to show that it is also dense. Given an arbitrary $D \in Bar^n$, we will perform a series of infinitesimal perturbations of D, so that there exists a (small) open neighborhood U of D over which d_{D_0} is ∞-differentiable.

Since $\hat{Bar}$ is generic in Bar_n, up to an infinitesimal perturbation, we can assume that D lies in $\hat{Bar}$. Let $\tilde{D}_m \in \mathbb{R}^{2m+n}$ be a minimal pre-image of D. By Lemma 1.7.13, the set of minimal ordered barcodes in $\mathbb{R}^{2m+n}$ is open. Moreover, $d_{D_0} \circ Q_{m,n}$ is smooth on a generic subset of $\mathbb{R}^{2m+n}$ by Lemma 1.7.11. Therefore, up to an infinitesimal perturbation of $\tilde{D}_m$ (which results in an infinitesimal perturbation of D by continuity of $Q_{m,n}$), we can further assume that $d_{D_0} \circ Q_{m,n}$ is smooth on a ball $\mathcal{B}(\tilde{D}_m, \epsilon)$ for some $\epsilon > 0$, while $\tilde{D}_m$ remains minimal and D stays in $\hat{Bar}$.

Reducing ϵ if necessary, by Lemma 1.7.13 all the maps $d_{D_0} \circ Q_{m',n}$ are smooth over $\mathcal{B}(\tilde{D}_{m'}, \epsilon)$, with gradients as in (i) or (ii) of Proposition 1.7.10, where $\tilde{D}_{m'}$ ranges over the pre-images of D. Reducing ϵ further if necessary, we conclude that d_{D_0} is ∞-differentiable over $\mathcal{B}(D, \epsilon)$ by Lemma 1.7.14. □

1.8 Questions

Our goal in this paper was to deepen our understanding of the differential properties of maps factorizing through the space of barcodes Bar. Generalizing the lifting argument present implicitly in [GHO16], we defined differentiability for maps from and to Bar via lifts and pre-images in spaces of ordered barcodes, and we showed that this approach is, in essence, connected to the theory of diffeological spaces. We then worked out sufficient conditions on a parametrization of filter functions,

Marcio Gameiro, Yasuaki Hiraoka, and Ippei Obayashi. Continuation of point clouds via persistence diagrams. *Physica D: Nonlinear Phenomena*, 334:118–132, 2016

for its induced barcode-valued map to be differentiable on a generic subset of the parameter space, and we also studied the behavior of the barcode-valued map at singularities (in the case where the domain of the filter functions is a simplicial complex). Then, we leveraged our results to recover the differentiability properties derived in earlier work in a principled way. Note that our analysis was carried out on barcodes of ordinary persistence, and that it can be adapted to extended persistence and zigzags in a fairly straightforward way.

This account is but a first step towards laying down the foundations of differential calculus on barcodes, and many important questions remain open, such as:

- In the case of filter functions on a compact smooth manifold $\mathcal{X}$ (Section 1.6), we do not have a clear understanding of the behavior of the barcode-valued map B at its singular points. We would like to mimick the analysis of the discrete setting (proposition 1.4.19), which would involve analyzing the behavior of critical points in Whitney neighborhoods of functions that are not Morse. Cerf theory[29] seems to be the appropriate tool for this.
- So far, our filter functions have been defined on a fixed domain, we do not know what happens if we allow for smooth variations of this domain. As an analogy with continuity considerations, the Stability Theorem as recorded here is about variations of the filter function, but there are also versions about variations (in the Gromov-Hausdorff distance) of the domain[30]. To make sense of what *smooth* variations of the domain could be, one can for instance consider the space of embeddings between given manifolds.
- The Stability Theorem and its converse [31] yield an isometry between the spaces of persistence modules and of barcodes. This isometry can be leveraged to extend our framework to the space of persistence modules. In multi-persistence however, no such isometry (and even no notion of barcode) exists in general. It would be desirable to have an intrinsic version of our framework, defining differentiability from and to the space of persistence modules.

[29] Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970

[30] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the twenty-fifth annual Symposium on Computational Geometry*, pages 237–246. ACM, 2009

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

2

Optimisation of Spectral Wavelets for Persistence-based Graph Classification

Abstract

A graph's spectral wavelet signature determines a filtration, and consequently an associated set of extended persistence diagrams. We propose a framework that optimises the choice of wavelet for a dataset of graphs, such that their associated persistence diagrams capture features of the graphs that are best suited to a given data science problem. Since the spectral wavelet signature of a graph is derived from its Laplacian, our framework encodes geometric properties of graphs in their associated persistence diagrams and can be applied to graphs without a priori node attributes. We apply our framework to graph classification problems and obtain performances competitive with other persistence-based architectures. To provide the underlying theoretical foundations, we extend the differentiability result for ordinary persistent homology to extended persistent homology.

2.1 Introduction

2.1.1 Background

Graph classification is a challenging problem in machine learning. Unlike data represented in Euclidean space, there is no easily computable notion of distance or similarity between graphs. As such, graph classification requires techniques that lie beyond mainstream machine learning techniques focused on Euclidean data. Much research has been conducted on methods such as graph neural networks (GNNs) [XSC+19] and graph kernels [VSKB10, SVP+09] that embed graphs in Euclidean space in a consistent manner.

Recently, *persistent homology* [ZC05, EH08] has been applied as a feature map that explicitly represents topological and geometric features of a graph as a set of *persistence diagrams* (a.k.a. *barcodes*). In the context of our discussion, the persistent homology of a graph $G = (V, E)$ depends on a vertex function $f : V \to \mathbb{R}$. In the case where a vertex function is not given with the data, several schemes have been proposed in the literature to assign vertex functions to graphs in a consistent way. For example, vertex functions can be constructed using local geometric descriptions of vertex neighbourhoods, such as discrete curvature [ZW19], heat kernel signatures [CCI+20] and Weisfeiler–Lehman graph kernels [RBB19].

However, it is often difficult to know *a priori* whether a heuristic vertex assignment scheme will perform well in addressing different data science problems. For a single graph, we can optimise the vertex function over $|V|$ many degrees of freedom in $\mathbb{R}^V$. In recent years, there have been many other examples of persistence optimisation in data science applications. The first two examples of persistence optimisation are the computation of Fréchet mean of barcodes using gradients on Alexandrov spaces [TMMH14], and that of point cloud inference [GHO16], where a point cloud is optimised so that its barcode fits a target fixed barcode. The latter is an instance of topological inverse problems (see [OS20a] for a recent overview of such). Another inverse problem is that of surface reconstruction [BGGSSG20b]. Besides, in the context of shape matching [PSO18], persistence optimisation is used in order to learn an adequate function between shapes. Finally, there are also many recent applications of persistence optimisation in Machine Learning, such as the incorporation of topological information in Generative Modelling [MHRB20a, HKND19, GNDS20b] or in Image Segmentation [HLSC19, COB+19], the design of topological losses for Regularization in supervised learning [CNBW19] or for dimension reduction [Kac20].

Each of these applications can be thought of as minimising a certain

loss function over a manifold M of parameters:

$$\min_{\theta \in M} \mathcal{L}(\theta),$$

where $\mathcal{L} : M \to \mathbf{Bar}^N \to \mathbb{R}$ factors through the space $\mathbf{Bar}^N$ of N-tuples of barcodes. The aim is to find the parameter θ that best fits the application at hand. Gradient descent is a very popular approach in minimisation, but it requires the ability to differentiate the loss function. In fact, [LOT21] provide notions of differentiability for maps in and out **Bar** that are compatible with smooth calculus, and show that the loss functions $\mathcal{L}$ corresponding the applications cited in the above paragraph are generically differentiable. The use of (stochastic) gradient descent is further legitimated by [CCG+21], where convergence guarantees on persistence optimisation problems are devised, using a recent study of stratified non-smooth optimisation problems [DDKL20]. In practice, the minimisation of $\mathcal{L}$ can be unstable due to its non-convexity and partial non-differentiability. Some research has been conducted in order to smooth and regularise the optimisation procedure [SWB21, CD20].

In a supervised learning setting, we want to optimise our vertex function assignment scheme over many individual graphs in a dataset. Since graphs may not share the same vertex set and come in different sizes, optimising over the $|V|$ degrees of freedom of any one graph is not conducive to learning a vertex function assignment scheme that can generalise to another graph. The degrees of freedom in any practical vertex assignment scheme should be independent of the number of vertices of a graph. However, a framework for parametrising and optimising the vertex functions of many graphs over a common parameter space M is not immediately apparent.

The first instance of a graph persistence optimisation framework (`GFL`) [HGR+20b] uses a one layer graph isomorphism network (GIN) [XSC+19] to parametrise vertex functions. The GIN learns a vertex function by exploiting the local topology around each vertex. In this paper, we propose a different framework for assigning and parametrising vertex functions, based on a graph's Laplacian operator.

2.1.2 *Outline and Contributions*

We address the issue of vertex function parametrisation and optimisation using *wavelet signatures*. Wavelet signatures are vertex functions derived from the eigenvalues and eigenvectors of the graph Laplacian and encode multiscale geometric information about the graph [LH13]. The wavelet signature of a graph is dependent on a choice of wavelet $g : \mathbb{R} \to \mathbb{R}$, a function on the eigenvalues of the graph's Laplacian matrix. We can thus obtain a parametrisation of vertex functions for

any graph $F : M \to \mathbb{R}^V$ by parametrising g. Consequently, the *extended persistence* of a graph – which has only four non-trivial persistence diagrams – can be varied over the parameter space M. If we have a function Out : $\mathbf{Bar}^4 \to \mathbb{R}$ on persistence diagrams that we wish to minimise, we can optimise over M to minimise the loss function

$$\mathcal{L} : M \xrightarrow{F} \mathbb{R}^V \xrightarrow{\text{EPH}} \mathbf{Bar}^4 \xrightarrow{\text{Out}} \mathbb{R}. \tag{2.1}$$

If $\mathcal{L}$ is generically differentiable, we can optimise the wavelet signature parameters $\theta \in M$ using gradient descent methods. We illustrate an application of this framework to a graph classification problem in fig. 2.1, where the loss function $\mathcal{L}$ is the classification error of a graph classification prediction model based on the graph's extended persistence diagrams.

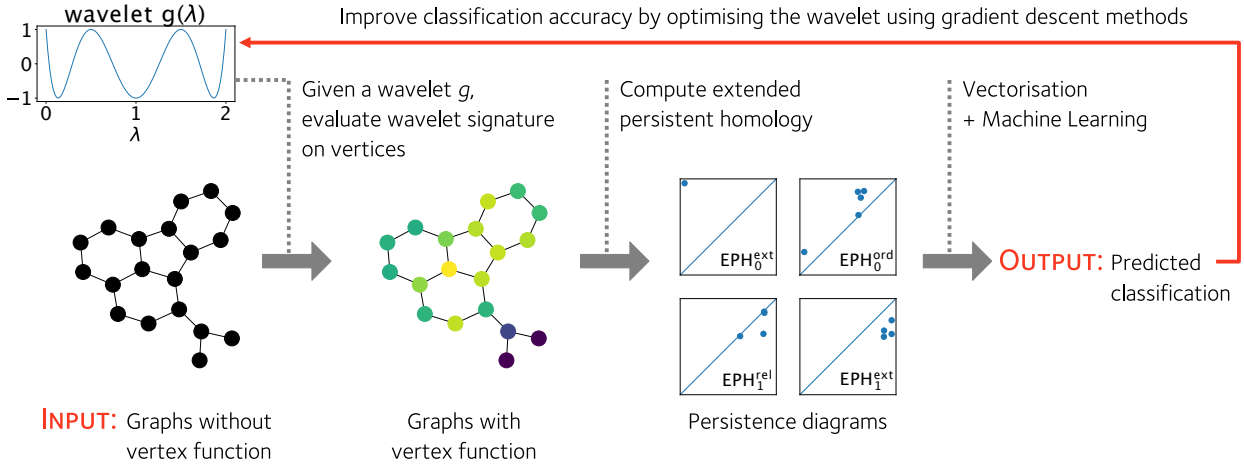


Figure 2.1: Given a wavelet $g : \mathbb{R} \to \mathbb{R}$, we can equip any graph with a non-trivial vertex function. This allows us to compute the extended persistence diagrams of a graph and use the diagrams as features of the graph to predict a graph's classification in some real world setting. The wavelet g can be optimised to improve the classification accuracy of a graph classification pipeline based on the extended persistence diagrams of a graph's vertex function.

In section 2.2, we describe the assignment of vertex functions $F : M \to \mathbb{R}^V$ by reviewing the definition of wavelet signatures. While spectral wavelets have been used in graph neural network architectures that predict vertex features [XSC+19] and compress vertex functions [RG19], they have not been considered in a persistent homology framework for graph classification. We describe several ways to parametrise wavelets. We also show in Proposition 2.2.2 that wavelet signatures are independent of the choice of eigenbasis of the graph Laplacian from which it is derived, ensuring that it is well-defined.

In section 2.3, we describe the theoretical basis for optimising the *extended* persistent homology of a vertex function EPH : $\mathbb{R}^V \to \mathbf{Bar}^4$ and elucidate what it means for $\mathcal{L}$ to be differentiable. In Proposition 2.3.3, we generalise the differentiability formalism of ordinary persistence[1] to extended persistence. We prove this result in Proposition 2.3.3

[1] Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021

Finally, in section 2.4, we apply our framework to graph classification problems on several benchmark datasets. We show that our model is competitive with state-of-the-art persistence-based models. In particular, optimising the vertex function appreciably improves the prediction accuracy on some datasets.

2.2 *Filter Function Parametrization*

We describe our recipe for assigning vertex functions to any simplicial graph $G = (V, E)$ based on a parametrised spectral wavelet, the first part F of the loss function

$$\mathcal{L} : M \xrightarrow{F} \mathbb{R}^V \xrightarrow{\text{EPH}} \mathbf{Bar}^4 \xrightarrow{\text{Out}} \mathbb{R} \ . \qquad \text{((2.1) recalled)}$$

Our recipe is based on a graph's wavelet signature, a vertex function derived from the graph's Laplacian. The wavelet signature also depends on a so-called 'wavelet function' in $g : \mathbb{R} \to \mathbb{R}$, which is independent of the graph. By modulating the wavelet, we can jointly vary the wavelet signature across many graphs. We parametrise the wavelet using a finite linear combination of basis functions, such that the wavelet signature can be manipulated in a computationally tractable way. In the following section, we define the wavelet signature and describe our linear approach to wavelet parametrisation.

2.2.1 *Wavelet Signatures*

The wavelet signature is a vertex function initially derived from wavelet transforms of vertex functions on graphs [HVG11], a generalisation of wavelet transforms for square integrable functions on Euclidean space [Gra95, Chu16] for signal analysis [AHHH01]. Wavelet signatures for graphs have been applied to encode geometric information about meshes of 3D shapes [AHHH01, LH13]. Special cases of wavelets signatures, such as the heat kernel signature [SOG09] and wave kernel signature [ASC11], have also been applied to describe graphs and 3D shapes [BK10, HRG14].

The wavelet signature of a graph is constructed from the graph's Laplacian operator. A graph's normalised Laplacian $L \in \mathbb{R}^{V \times V}$ is a symmetric positive semi-definite matrix, whose entries are given by

$$L_{uv} = \begin{cases} 1 & u = v \\ -\frac{1}{\sqrt{k_u k_v}} & (u, v) \in E \\ 0 & \text{otherwise} \end{cases} \qquad (2.2)$$

where k_u is the degree of vertex u. The Laplacian's eigenvalues λ and eigenvectors $\boldsymbol{\phi}$ are known to encode various topological and geomet-

ric information about the graph[2]; for example, the number of zero eigenvalues corresponds to the number of connected components of the graph. The spectrum of the normalised Laplacian have real eigenvalues[3] in $[0,2]$. As such, any function $g : \mathbb{R} \to \mathbb{R}$ evaluated on the eigenvalues need only be defined on $[0,2]$. Moreover, functions on a compact domain are easily parametrised using convenient bases.

[2] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997; and Türker Biyikoglu, Josef Leydold, and Peter F Stadler. *Laplacian eigenvectors of graphs: Perron-Frobenius and Faber-Krahn type theorems*. Springer, 2007

[3] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997

Definition 2.2.1. [Wavelet Signature[4]] Let $L \in \mathbb{R}^{V \times V}$ be the normalised Laplacian of a simplical graph $G = (V, E)$. Let $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{|V|}$ be an orthonormal eigenbasis for L and $\lambda_1, \ldots, \lambda_{|V|}$ be their corresponding eigenvalues. The wavelet signature $W : \mathbb{R}^{[0,2]} \to \mathbb{R}^V$ maps a function $g : [0,2] \to \mathbb{R}$, which we refer to as a *wavelet*, to a vertex function $W(g) \in \mathbb{R}^V$ linearly, where the value of $W(g)$ on vertex v is given by

[4] Chunyuan Li and A Ben Hamza. A multiresolution descriptor for deformable 3d shape retrieval. *The Visual Computer*, 29(6-8):513–524, 2013

$$W(g)_v = \sum_{i=1}^{|V|} g(\lambda_i) \left(\boldsymbol{\phi}_i\right)_v^2, \tag{2.3}$$

and $\left(\boldsymbol{\phi}_i\right)_v$ denotes the component of eigenvector $\boldsymbol{\phi}_i$ corresponding to vertex v.

If the eigenvalues of L have geometric multiplicity one (i.e. their eigenspaces are one dimensional), then the orthonormal eigenvectors are uniquely defined up to a choice of sign. It is then apparent from 2.3 that the wavelet signature is independent of the choice of sign. However, if some eigenvalues have geometric multiplicity greater than one, then the orthonormal eigenvectors of L are uniquely defined up to orthonormal transformations in the individual eigenspaces. However we show that the wavelet signature is well-defined even when the multiplicities of eigenvalues are greater than one.

Proposition 2.2.2. *The wavelet signature of a graph is independent of the choice of orthonormal eigenbasis for the Laplacian.*

Remark 2.2.3. In addition to the traditional view of wavelets from a spectral signal processing perspective[5], we can also relate the wavelet signature of a vertex v to the degrees of vertices in some neighbourhood of v prescribed by g. Consider a wavelet $g : [0,2] \to \mathbb{R}$. On a finite graph G, the normalised Laplacian L has at most $|V|$ many distinct eigenvalues. As such, there exists a polynomial $\hat{g}(x) = \sum_{n=0}^{p} a_n x^n$ of finite order that interpolates g at the eigenvalues $g(\lambda_i) = \hat{g}(\lambda_i)$. Therefore, $W(g) = W(\hat{g})$. Moreover, the vertex values assigned by $W(\hat{g})$ are the diagonal entries of the matrix polynomial $\hat{g}(L)$:

[5] David K Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011

$$\hat{g}(L)_{vv} = \sum_{n=0}^{p} a_n (L^n)_{vv} = \sum_{i=1}^{|V|} \hat{g}(\lambda_i) \left(\boldsymbol{\phi}_i\right)_v^2 = \sum_{i=1}^{|V|} g(\lambda_i) \left(\boldsymbol{\phi}_i\right)_v^2 = W(g)_{vv}. \tag{2.4}$$

Furthermore, we can also write the matrix polynomial $\hat{g}(L)$ as a matrix polynomial in $A = I - L$, the *normalised adjacency matrix*. From the definition of L, we can compute the diagonal entry of a monomial A^r corresponding to vertex v as an inverse degree weighted count of paths[6] $[v_0, v_1, \ldots, v_r]$ on the graph which begin and end on vertex $v = v_0 = v_r$, see [New18][7]:

$$(A^r)_{vv} = \frac{1}{k_v} \sum_{[v,v_1,\ldots,v_{r-1},v]} \left(\prod_{l=1}^{r-1} \frac{1}{k_{v_l}} \right). \tag{2.5}$$

[6] Here a path refers to a sequences of vertices that are connected to the next vertex in the sequence by an edge.

[7] Mark Newman. *Networks*. Oxford university press, 2018

By expressing the wavelet signature as a matrix polynomial in A, we see that g controls how information at different length scales of the graph contribute to the wavelet signature. For instance, if g were an order p polynomial, then $W(g)_v$ only takes the degrees of vertices that are $\lfloor p/2 \rfloor$ away from v into account. As a corollary, since $W(g)$ can be specified by replacing g with a polynomial $\hat{g}$ of order at most $|V| - 1$, the wavelet signature at a vertex is only dependent on the subgraph of G that is within $\lfloor |V| - 1 \rfloor / 2$ steps away from v.

Proof. Let $\mathrm{Spec}(L) \subset \mathbb{R}$ denote the spectrum of L and $\boldsymbol{\phi}_1, \ldots \boldsymbol{\phi}_{|V|}$ be a set of orthonormal eigenvectors of L. Let us denote $\Phi(\lambda)$ to be a $|V| \times m$ matrix where m corresponds to the geometric multiplicity of λ, and the m column vectors of $\Phi(\lambda)$ correspond to eigenvectors $\boldsymbol{\phi}_{i_1}, \ldots, \boldsymbol{\phi}_{i_m}$ with eigenvalue λ. Then we can rewrite the wavelet signature eq. (2.3) as

$$W(g)_v = \sum_{i=1}^{|V|} g(\lambda_i) \left(\boldsymbol{\phi}_i\right)_v^2 = \sum_{\lambda \in \mathrm{Spec}(L)} g(\lambda) (\Phi(\lambda)\Phi(\lambda)^\intercal)_{vv} \tag{2.6}$$

Suppose we have another choice of eigenbasis of L. Without loss of generality for $\lambda \in \mathrm{Spec}(L)$, the new basis $\boldsymbol{\phi}'_{i_1}, \ldots, \boldsymbol{\phi}'_{i_m}$ for $\mathrm{eig}(\lambda)$ is related to the previous eigenbasis $\boldsymbol{\phi}_{i_1}, \ldots, \boldsymbol{\phi}_{i_m}$ by an orthonormal transformation transformation $U(\lambda) \in \mathbb{R}^{m \times m}$ on $\Phi(\lambda)$:

$$\Phi'(\lambda) = \begin{bmatrix} \boldsymbol{\phi}'_1 & \cdots & \boldsymbol{\phi}'_m \end{bmatrix} = \Phi(\lambda) U(\lambda).$$

As $U(\lambda)$ is an orthonormal transformation with $U(\lambda)U(\lambda)^\intercal = 1$,

$$\begin{aligned} \Phi'(\lambda)\Phi'(\lambda)^\intercal &= \Phi(\lambda)U(\lambda)(\Phi(\lambda)U(\lambda))^\intercal \\ &= \Phi(\lambda)U(\lambda)U(\lambda)^\intercal\Phi(\lambda)^\intercal \\ &= \Phi(\lambda)\Phi(\lambda)^\intercal. \end{aligned}$$

Since the $V \times V$ matrix $\Phi(\lambda)\Phi(\lambda)^\intercal$ is independent of the choice of eigenbasis, the wavelet signature given on the right hand side of eq. (2.6) must also be independent of the choice of eigenbasis. □

2.2.2 *Parametrising the Wavelet*

We see from remark 2.2.3 that the choice of wavelet g determines how the topology and geometry of the graph is reflected in the vertex function. Though the space of wavelets is potentially infinite dimensional, here we only consider wavelets $g_\theta(x)$ that are parametrised by parameters θ in a finite dimensional manifold, so that we can easily optimise them using computational methods. In particular, we focus on wavelets written as a linear combination of m *basis functions* $h_1, \dots, h_m : [0,2] \to \mathbb{R}$

$$g_\theta(x) := \sum_{j=1}^{m} \theta_j h_j(x) \tag{2.7}$$

This parametrisation of wavelets in turn defines a parametrisation of vertex functions $F : \mathbb{R}^m \to \mathbb{R}^V$ for our optimisation pipeline in (2.1)

$$F: \quad \theta \in \mathbb{R}^m \quad \longmapsto \quad F(\theta) := W(g_\theta) \quad \in \mathbb{R}^V. \tag{2.8}$$

Since $W(g)$ is a linear function of the wavelet g, F is a linear transformation:

$$F(\theta) = W\left(\sum_{j=1}^{m} \theta_j h_j(x)\right) = \sum_{j=1}^{m} \theta_j W\big(h_j\big). \tag{2.9}$$

We can write F as a $|V| \times m$ matrix acting on a vector $[\theta_1, \dots \theta_m]^\intercal \in \mathbb{R}^m$, whose columns are the vertex functions $W\big(h_j\big)$.

Example 2.2.4 (Chebyshev Polynomials)**.** Any Lipschitz continuous function on an interval can be well approximated by truncating its Chebyshev series at some finite order.[8] The Chebyshev polynomials $T_n : [-1,1] \to \mathbb{R}$

$$T_n(x) = \cos(n \arccos(x)) \quad n \in \mathbb{N}_{\geqslant 0}. \tag{2.10}$$

form an orthonormal set of functions. We can thus consider $h_j(\lambda) = T_j(\lambda - 1)$, $j = 0, 2, \dots, m$ as a naïve basis for wavelets. We exclude $T_1(x) = x$ in the linear combination as $W(T_1(1-x)) = 0$ for graphs without self loops.

[8] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*, volume 50. Siam, 1997

Example 2.2.5 (Radial Basis Functions)**.** In the machine learning community, a *radial function* refers loosely to a continuous monotonically decreasing function $\rho : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$. There are many possible choices for ρ, for example, the inverse multiquadric

$$\rho(r) = \left(\left(\frac{r}{\epsilon}\right)^2 + 1\right)^{-\frac{1}{2}} \tag{2.11}$$

where $\epsilon \neq 0$ is a width parameter. We can obtain a naïve wavelet basis $h_j(x) = \rho\big(\|x - x_j\|\big)$ using copies of ρ offset by a collection of centroids

$x_j \in \mathbb{R}$ along $\mathbb{R}$. In general, the centroids are parameters that could be optimised, but we fix them in this study. This parametrisation can be considered as a *radial basis function neural network*. RBNNs are well-studied in function approximation and subsequently machine learning; we refer readers to [CCG91, PS91] for further details.

Sheng Chen, Colin FN Cowan, and Peter M Grant. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on neural networks*, 2(2):302–309, 1991; and Jooyoung Park and Irwin W Sandberg. Universal approximation using radial-basis-function networks. *Neural computation*, 3(2):246–257, 1991

2.2.3 *The Choice of Wavelet Basis*

The choice of basis functions determines the space of wavelet signatures and also the numerical stability of the basis function coefficients which serve as the wavelet signature parameters. The stability of the parametrisation depends on the graphs as much as the choice of wavelet basis $h_1, \ldots, h_m$. We can analyse the stability of a parametrisation F by its the singular value decomposition

$$F = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k^{\mathsf{T}} \tag{2.12}$$

where $\sigma_1, \ldots, \sigma_r$ are the non-zero singular values of the matrix, and $\mathbf{u}_k \in \mathbb{R}^{|V|}$ and $\mathbf{v}_k \in \mathbb{R}^m$ are orthonormal sets of vectors respectively. If the distribution of singular values span many orders of magnitude, we say the parametrisation is *ill-conditioned*. An ill-conditioned parametrisation interferes with the convergence of gradient descent algorithms on a loss function evaluated on wavelet signatures. We discuss the relationship between the conditioning of F and the stability of gradient descent in detail in remark 2.2.7.

We empirically observe that the coefficients of a naïve choice of basis functions, such as Chebyshev polynomials or radial basis functions, are numerically ill-conditioned. In figure 2.2, we can see that the singular values of radial basis function and Chebyshev polynomial parametrisations respectively are distributed across a large range on the logarithmic scale for some datasets of graphs in machine learning. We address this problem by picking out a new wavelet basis

$$h'_k(x) = \frac{1}{\sigma_k} \sum_{j=1}^{m} (\mathbf{v}_k)_j h_j(x), \quad k = 1, \ldots, r, \tag{2.13}$$

where σ_k are the singular values of F and $\mathbf{v}_k$ are the associated vectors in $\mathbb{R}^m$ from the singular value decomposition of matrix F in (2.12). Then the parametrisation $F' : \mathbb{R}^r \to \mathbb{R}^V$

$$F'(\theta') = \sum_{k=1}^{r} \theta'_k W(h'_k). \tag{2.14}$$

have singular values equal to one, since this is a linear combination of

orthonormal vectors $\mathbf{u}_k \in \mathbb{R}^V$:

$$W(h'_k) = \sum_{j=1}^{m} \frac{1}{\sigma_k} (\mathbf{v}_k)_j W(h_j) = \frac{1}{\sigma_k} F\mathbf{v}_k = \mathbf{u}_k. \tag{2.15}$$

As an example, we plot the new wavelet basis h'_k derived from a twelve parameter radial basis function parametrisation for the MUTAG dataset in fig. 2.3.

Remark 2.2.6 (Learning a Wavelet Basis for Wavelet Signatures on Multiple Graphs)**.** In the case where the wavelet coefficients parametrise the wavelet signatures over graphs $G_1, \ldots, G_N$, we can view the maps $F_1, \ldots, F_N$ that map wavelet basis coefficients to vertex functions of graphs $G_1, \ldots, G_N$ respectively as a parametrisation for the disjoint union $\bigsqcup_i G_i$:

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_N \end{bmatrix} = \begin{bmatrix} F_1 \\ \vdots \\ F_N \end{bmatrix} \theta =: F\theta. \tag{2.16}$$

We can then perform a singular value decomposition of the parametrisation F on $\bigsqcup_i G_i$ and derive a new, well-conditioned basis.

Remark 2.2.7 (Why the Conditioning of F Matters)**.** Let us optimise a loss function $\mathcal{L}$ on the parameter space of wavelet coefficients θ using a gradient descent algorithm. In a gradient descent step of step size s, the wavelet coefficients are updated to $\theta \mapsto \theta - s\nabla_\theta \mathcal{L}$. Using the singular value decomposition of F ((2.12)), we can write

$$\nabla_\theta \mathcal{L} = \nabla_\theta f^\intercal \nabla_f \mathcal{L} = F^\intercal \nabla_f \mathcal{L} = \sum_{k=1}^{r} \sigma_k \left\langle \nabla_f \mathcal{L}, \mathbf{u}_k \right\rangle \mathbf{v}_k. \tag{2.17}$$

The change in the vertex function is simply the matrix F applied to the change in wavelet parameters. Hence the vertex function is updated to $f \mapsto f - sF\nabla_\theta \mathcal{L}$, where

$$F\nabla_\theta \mathcal{L} = \sum_{k=1}^{r} \sigma_k \left\langle \nabla_f \mathcal{L}, \mathbf{u}_k \right\rangle F\mathbf{v}_k = \sum_{k=1}^{r} \sigma_k^2 \left\langle \nabla_f \mathcal{L}, \mathbf{u}_k \right\rangle \mathbf{u}_k. \tag{2.18}$$

If the loss function $\mathcal{L}$ has large second derivatives– for example, due to nonlinearities in the function on persistence diagrams $\mathrm{Out} : \mathbf{Bar}^4 \to \mathbb{R}$ – the projections $\left\langle \nabla_f \mathcal{L},\ \mathbf{u}_k \right\rangle$ in (2.17) and (2.18) may change dramatically from one gradient descent update to another. If the smallest singular value is much smaller than the largest, then updates to the wavelet signature can be especially unstable throughout the optimisation process. This source of instability can be removed if we choose a parametrisation with uniform singular values $\sigma_k = 1$. In this case, the update to f is simply the projection of $\nabla_f \mathcal{L}$ onto the space of wavelet

signatures spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_r$, without any distortion introduced by non-uniform singular values:

$$f \mapsto f - s \sum_{k=1}^{r} \left\langle \mathbf{u}_k, \nabla_f \mathcal{L} \right\rangle \mathbf{u}_k. \tag{2.19}$$

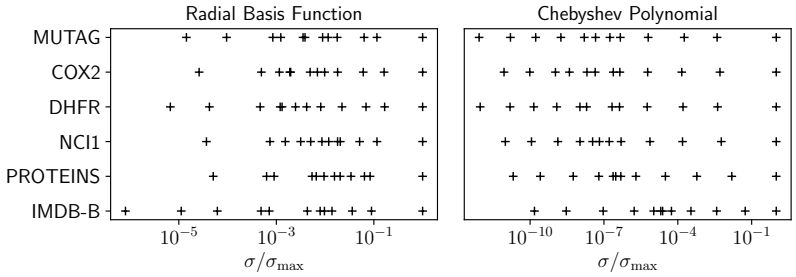


Figure 2.2: We consider the parametrisations of wavelet signatures on some datasets of graphs in machine learning, namely `MUTAG`, `COX2`, `DHFR`, `NCI1`, `PROTEINS` and `IMDB-B`, using coefficients of 12 radial basis functions (see eq. (2.30)) and a degree 13 Chebyshev polynomial respectively. For each dataset, we plot the distribution of the singular values σ of the map F in eq. (2.16) from the basis function coefficients $\theta \in \mathbb{R}^{12}$ to the wavelet signature on the whole dataset of graphs, as a fraction of the largest singular value $\sigma_{\max}$ of F. We can observe that for both parametrisations, the singular values span many orders of magnitudes across different datasets. Note that the singular values of F not only depend on the choice of basis but also on the dataset of graphs.

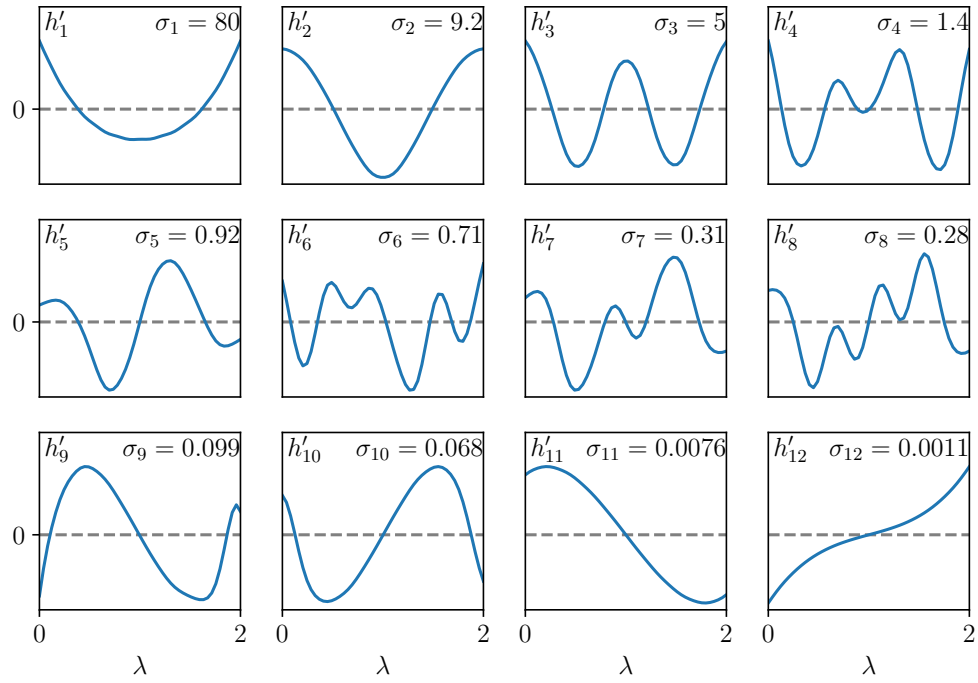


Figure 2.3: The functions shown are the new, stable wavelet basis $h'_1, \ldots, h'_{12}$ (eq. (2.13)) for the `MUTAG` dataset, derived from an initial numerically unstable parametrisation using twelve inverse multiquadric radial basis functions (eq. (2.30)). We parametrise the wavelet as a linear combination of these basis functions.

2.3 *Extended Persistent Homology*

The homology of a given graph is a computable vector space whose dimension counts the number of connected components or cycles in

the graph. Finer information can be retained by filtering the graph and analysing the evolution of the homology throughout the filtration. This evolution is described by a set of *extended persistence diagrams* (a.k.a. *extended barcodes*), a multiset of points $\langle b, d \rangle$ that record the birth b and death of homological features in the filtration. In this section, we begin by summarising these constructions. We refer the reader to [ZC05, EH08, CSEH07] for full treatments of the theory of Persistence.

Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005; Herbert Edelsbrunner and John Harer. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008; and David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007

Compared to *ordinary persistence*, extended persistence is a more informative and convenient feature map for graphs. Extended persistence encodes strictly more information than ordinary persistence. For instance, the cycles of a graph are represented as points with $d = \infty$ in ordinary persistence. Thus, only the birth coordinate b of such points contain useful information about the cycles. In contrast, the corresponding points in extended persistence are each endowed with a finite death time d, thus associating extra information to the cycles. The points at infinity in ordinary persistence also introduce obstacles to vectorisation procedures, as often arbitrary finite cutoffs are needed to 'tame' the persistence diagrams before vectorisation.

2.3.1 *Extended Persistent Homology*

Let $G = (V, E)$ be a finite graph without double edges and self-loops. For the purposes of this paper, the associated *extended persistent homology* is a map

$$\mathrm{EPH} : \mathbb{R}^V \to \mathbf{Bar}^4$$

from functions $f \in \mathbb{R}^V$ on its vertices to the space of four *persistence diagrams* or *barcodes*, which we define below. The map arises from a *filtration* of the graph, a sequential attachment of vertices and edges in ascending or descending order of f. We extend f on each edge $e = (v, v')$ by the maximal value of f over the vertices v and v', and we then let $G_t \subset G$ be the sub graph induced by vertices taking value less than t. Then we have the following sequence of inclusions:

$$\varnothing \longrightarrow \cdots \longrightarrow G_s \xrightarrow{s \leqslant t} G_t \longrightarrow \cdots \longrightarrow G. \tag{2.20}$$

Similarly, the sub graphs $G^t \subset G$ induced by vertices taking value greater than t assemble into a sequence of inclusions:

$$G \longleftarrow \cdots \longleftarrow G^s \xleftarrow{s \leqslant t} G^t \longleftarrow \cdots \longleftarrow \varnothing\ . \tag{2.21}$$

The changes in the topology of the graph along the filtration in ascending and descending order of f can be detected by its *extended*

persistence module, indexed over the poset $\mathbb{R} \cup \{\infty\} \cup \mathbb{R}[op]$:

$$\mathrm{V}_p(f): \quad \begin{array}{ccccccccc} 0 = H_p(\varnothing) & \longrightarrow \cdots \longrightarrow & H_p(G_s) & \xrightarrow{s\leqslant t} & H_p(G_t) & \longrightarrow \cdots \longrightarrow & H_p(G) \\ & & & & & & \downarrow\cong \\ 0 = H_p(G,G) & \longleftarrow \cdots \longleftarrow & H_p(G,G^s) & \xleftarrow{s\leqslant t} & H_p(G,G^t) & \longleftarrow \cdots \longleftarrow & H_p(G,\varnothing) \end{array}, \tag{2.22}$$

where H_p is the singular (relative) homology functor in degre $p \in 0, 1$ with coefficients in a fixed field, chosen to be $\mathbb{Z}/2\mathbb{Z}$ in pratice. In general terms, the modules $\mathrm{V}_0(f)$ and $\mathrm{V}_1(f)$ together capture the evolution of the connected components and loops in the sub graphs of G induced by the function f.

Each module $\mathrm{V}_p(f)$ is completely characterised by a finite multiset $\mathrm{EPH}_p(f)$ of pairs of real numbers $\langle b, d\rangle$ called *intervals* representing the birth and death of homological features. Following [CSEH09], the intervals in $\mathrm{EPH}_p(f)$ are further partitioned according to the type of homological feature they represent:

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

$$\mathrm{EPH}_p(f) = \underbrace{\{\langle b,d\rangle \mid b < d < \infty\}}_{=\mathrm{EPH}_p^{\mathrm{ord}}(f)} \sqcup \underbrace{\{\langle b,d\rangle \mid b < \infty < d\}}_{=\mathrm{EPH}_p^{\mathrm{ext}}(f)} \sqcup \underbrace{\{\langle b,d\rangle \mid \infty < b < d\}}_{=\mathrm{EPH}_p^{\mathrm{rel}}(f)}. \tag{2.23}$$

Each of the three finite multiset $\mathrm{EPH}_p^{\mathrm{k}}(f)$, for $k \in \{\mathrm{ord}, \mathrm{ext}, \mathrm{rel}\}$, is an element in the space **Bar** of so-called *barcodes* or *persistence diagrams*. However, $\mathrm{EPH}_0^{\mathrm{rel}}(f)$ and $\mathrm{EPH}_1^{\mathrm{ord}}(f)$ being trivial for graphs, we refer to the collection of four remaining persistence diagrams

$$\mathrm{EPH}(f) = \left[\mathrm{EPH}_0^{\mathrm{ord}}(f), \mathrm{EPH}_0^{\mathrm{ext}}(f), \mathrm{EPH}_1^{\mathrm{ext}}(f), \mathrm{EPH}_1^{\mathrm{rel}}(f)\right] \in \mathbf{Bar}^4 \tag{2.24}$$

as the *extended barcode* or *extended persistence diagram* of f. We have thus defined the *extended persistence map*

$$\mathrm{EPH} : \mathbb{R}^V \to \mathbf{Bar}^4.$$

Remark 2.3.1. If we only apply homology to the filtration of Eq. (2.20), we get an *ordinary persistence module* indexed over the real line, which is essentially the first row in Eq. (2.22). This module is characterised by a unique barcode $\mathrm{PH}_p(f) \in \mathbf{Bar}$. We refer to the map

$$\mathrm{PH} : f \in \mathbb{R}^V \longmapsto [\mathrm{PH}_0(f), \mathrm{PH}_1(f)] \in \mathbf{Bar}^2 \tag{2.25}$$

as the *ordinary persistence map*.

2.3.2 *Differentiability of Extended Persistence*

The extended persistence map can be shown to be locally Lipschitz by the Stability theorem[9]. The Rademacher theorem states that any

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

real-valued function that is locally Lipschitz is differentiable on a full measure set. Thus, so is our loss function

$$\mathcal{L} : M \xrightarrow{F} \mathbb{R}^V \xrightarrow{\text{EPH}} \mathbf{Bar}^4 \xrightarrow{\text{Out}} \mathbb{R} \ . \qquad \text{((2.1) recalled)}$$

as long as Out and F are smooth or locally Lipschitz.[10] If a loss function $\mathcal{L}$ is locally Lipschitz, we can use stochastic gradient descent as a paradigm for optimisation. Nonetheless, the theorem above does not rule out dense sets of non differentiability in general.

[10] In practice, a locally Lipschitz Out can be constructed out of Lipschitz stable vectorisation methods, such as Persistence Landscapes [Bub15] and Persistence Images [AEK+17].

In this section, we show that the set where EPH is not differentiable is not pathological. Namely, we show that EPH is *generically* differentiable, i.e. differentiable on an open dense subset. This property guarantees that local gradients yield reliable descent directions in a neighbourhood of the current iterate. We recall from [LOT21][11] the definition of differentiability for maps to barcodes.

[11] Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021

We call a map $F : M \to \mathbb{R}^V$ a *parametrisation*, as it corresponds to a selection of filter functions over G parametrised by the manifold M. Then $B := \text{EPH} \circ F$ is the barcode valued map whose differentiability properties are of interest in applications.

Definition 2.3.2. A map $B : M \to \mathbf{Bar}$ on a smooth manifold M is said to be differentiable at $\theta \in M$ if for some neighbourhood U of θ, there exists a finite collection of differentiable maps[12] $b_i, d_i : U \to \mathbb{R} \cup \{\infty\}$, called a *local coordinate system* for B at θ, such that

$$\forall \theta' \in U,\ B(\theta') = \big\{\big\langle b_i(\theta'), d_i(\theta')\big\rangle \mid b_i(\theta') \neq d_i(\theta')\big\}.$$

[12] By convention, a differentiable map that takes the value ∞ is constant.

For $N \in \mathbb{N}$, we say that a map $B : M \to \mathbf{Bar}^N$ is differentiable at θ if all its components are so.

In [LOT21] it is proven that the composition $\text{PH} \circ F$ is generically differentiable as long as F is so. It is possible to show that $\text{EPH} \circ F$ is generically differentiable along the same lines, but we rather provide an alternative argument in the next section. Namely, we rely on the fact that the extended persistence of G can be decoded from the ordinary persistence of the cone complex $\mathrm{C}(G)$, a connection first noted in [CSEH09][13] for computational purposes.

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

Proposition 2.3.3. *Let $F : M \to \mathbb{R}^V$ be a generically differentiable parametrisation. Then the composition $\text{EPH} \circ F$ is generically differentiable.*

For completeness, the proof is provided in the next section and treats the general case of a finite simplicial complex K of arbitrary dimension.

2.3.3 *Differentiability of the extended persistence map*

Let K be a finite simplicial complex with vertex set V and dimension $d \in \mathbb{N}$. A vertex function $f \in \mathbb{R}^V$ extends to the whole complex via $f(\sigma) :=$

$\max_{v\in\sigma} f(v)$. Filtrations, persistence modules and barcodes are then defined analogously to the case of a graph. The extended barcode of a function f now consists of $3(d+1)$ barcodes:

$$\mathrm{EPH}(f) = \left[\{\mathrm{EPH}_p^{\mathrm{k}}(f)\}_{k\in\{\mathrm{ord,ext,rel}\}}\right]_{p=0}^{d} \in \mathbf{Bar}^{3(d+1)}. \tag{2.26}$$

We then have the extended persistence map

$$\mathrm{EPH} : \mathbb{R}^V \to \mathbf{Bar}^{3(d+1)},$$

and the ordinary persistence map as in remark 2.3.1

$$\mathrm{PH} : f \in \mathbb{R}^K \longmapsto \left[\mathrm{PH}_p(f)\right]_{p=0}^{d} \in \mathbf{Bar}^{d+1}.$$

Proposition 2.3.4. *Let K be a finite simplicial complex, and let $F : M \to \mathbb{R}^V$ be a generically differentiable parametrisation. Then the composition* $\mathrm{EPH} \circ F$ *is generically differentiable.*

In particular, taking the parameter space M to be the space $\mathbb{R}^V$ of vertex functions, we obtain the generic differentiablility of the extended persistence map EPH itself. Note that, however, we could not have directly deduced the generic differentiability of any composition of the form $\mathrm{EPH} \circ F$ from the generic differentiability of EPH. This is due to the fact that the image of a parametrisation F might lie in the set where EPH is not differentiable.

The idea of our proof is to view the extended persistence of a vertex function $f \in \mathbb{R}^V$ as the ordinary persistence of an extension of f over the cone complex $\mathrm{C}(K)$. We note that this point of view has proven to be particularly useful for computing extended persistence in practice [CSEH09]. The relationship between EPH and PH can be described by a commutative diagram:

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

$$\begin{array}{ccc} \mathbb{R}^V & \xrightarrow{\quad\mathrm{EPH}\quad} & \mathbf{Bar}^3 \\ \downarrow & & \uparrow \\ \mathbb{R}^{\mathrm{C}(K)} & \xrightarrow{\quad\mathrm{PH}\quad} & \mathbf{Bar} \end{array}$$

whose vertical maps are differentiable. Thus, we can deduce the differentiability of the extended persistence map EPH from the results of[LOT21] about the ordinary persistence map PH.

Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021

Proof of Proposition 2.3.4. Let $f \in \mathbb{R}^V$ be a vertex function. Let K^t (resp. K_t) be the maximal sub complexes of K induced by vertices taking values greater (resp less) than t. For $0 \leqslant p \leqslant d$, the associated p-th extended persistent homology module $\mathrm{V}_p(f)$ is:

$$\begin{array}{ccccccccccc} 0 = H_p(\varnothing) & \longrightarrow & \cdots & \longrightarrow & H_p(K_s) & \xrightarrow{s\leqslant t} & H_p(K_t) & \longrightarrow & \cdots & \longrightarrow & H_p(K) \\ & & & & & & & & & & \downarrow\cong \\ 0 = H_p(K,K) & \longleftarrow & \cdots & \longleftarrow & H_p(K,K^s) & \xleftarrow{s\leqslant t} & H_p(K,K^t) & \longleftarrow & \cdots & \longleftarrow & H_p(K,\varnothing) \end{array} \quad . \tag{2.27}$$

As such, $\mathrm{V}_p(f)$ is a module indexed over the extended real line $\mathbb{R} \sqcup \{\infty\} \sqcup \mathbb{R}^{\mathrm{op}}$. We construct an equivalent module $\mathrm{V}_{p,R}(f)$ over the simpler, compact poset $[-R;3R]$, where $R > 0$ is a large enough constant chosen hereafter. For this, we consider the poset map that collapses $\mathbb{R} \sqcup \{\infty\} \sqcup \mathbb{R}^{\mathrm{op}}$ onto $[-R;3R]$ as in fig. 2.4. Formally, the poset map is defined on $\mathbb{R}$ as the canonical retraction onto $[-R;R]$, on $\mathbb{R}^{\mathrm{op}}$ as the symmetry $t \mapsto 2R - t$ followed by the canonical retraction onto $[-R;R]$, and sends the point ∞ to R.

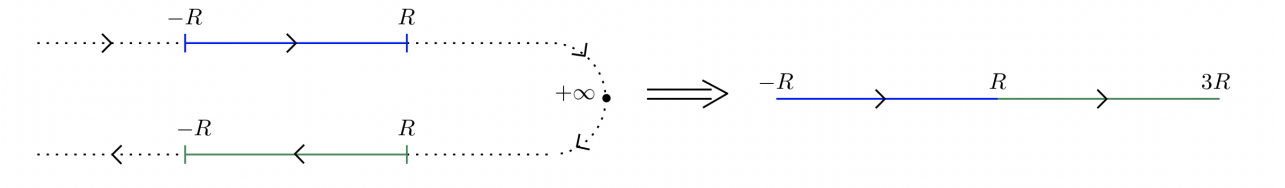


Figure 2.4: Collapsing the dotted part of the left poset yields the compact poset on the right.

If we choose $R > \sup_\sigma |f(\sigma)|$, then no new simplex enters the subcomplexes K_t or K^t for $t \notin [R;R]$ and $t \notin [R;R]^{\mathrm{op}}$. Hence, the module $\mathrm{V}_p(f)$ is locally constant outside of $[R;R]$ and $[R;R]^{\mathrm{op}}$. By applying the (inverse of the) poset map described above, we thus get a $[-R;3R]$-indexed module $\mathrm{V}_{p,R}(f)$:

$$\begin{array}{ccccccccccc} 0 = H_p(K_{-R}) & \longrightarrow & \cdots & \longrightarrow & H_p(K_s) & \xrightarrow{s\leqslant t} & H_p(K_t) & \longrightarrow & \cdots & \longrightarrow & H_p(K_R) \\ & & & & & & & & & & \downarrow\cong \\ 0 = H_p(K,K^{-R}) & \longleftarrow & \cdots & \longleftarrow & H_p(K,K^{2R-t}) & \xleftarrow{t\geqslant s} & H_p(K,K^{2R-s}) & \longleftarrow & \cdots & \longleftarrow & H_p(K;K^R) \end{array} \quad . \tag{2.28}$$

The extended module module $\mathrm{V}_p(f)$ is essentially equivalent to the ordinary module $\mathrm{V}_{p,R}(f)$, since we can retrieve the extended barcode $\mathrm{EPH}_p(f)$ of $\mathrm{V}_p(f)$ from the barcode of $\mathrm{V}_{p,R}(f)$ as follows:

- Each interval $\langle b,d\rangle$ in the barcode of $\mathrm{V}_{p,R}(f)$ such that $b \leqslant d < R$ yields an interval $\langle b,d\rangle$ in $\mathrm{EPH}_p^{\mathrm{ord}}(f)$;
- Each interval $\langle b,d\rangle$ in the barcode of $\mathrm{V}_{p,R}(f)$ such that $b < R < d$ yields an interval $\langle b,2R-d\rangle$ in $\mathrm{EPH}_p^{\mathrm{ext}}(f)$;
- Each interval $\langle b,d\rangle$ in the barcode of $\mathrm{V}_{p,R}(f)$ such that $R < b \leqslant d$ yields an interval $\langle 2R-b,2R-d\rangle$ in $\mathrm{EPH}_p^{\mathrm{rel}}(f)$.

We denote this decoding map by $\mathrm{Dec}_R : \mathbf{Bar} \to \mathbf{Bar}^3$. We next take advantage of working with the ordinary module $\mathrm{V}_{p,R}(f)$ by viewing it as the sub level set persistent homology module of a function defined on the cone $\mathrm{C}(K)$.

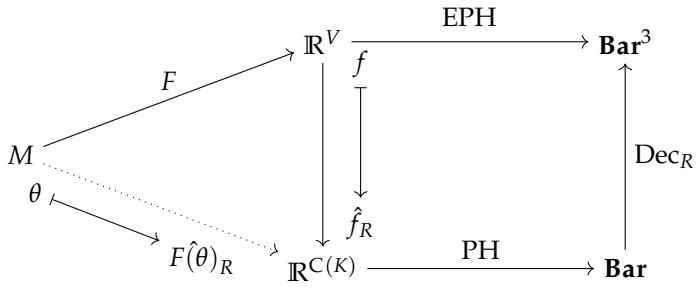
Note that the relative homology groups of $\mathrm{V}_{p,R}(f)$ in the second row of Eq. (2.28) may be replaced with ordinary (reduced) homology groups of the cones $\mathrm{C}\left(K^{2R-t}\right)$ using the functorial isomorphism:

$$H_p(K, K^{2R-t}) \cong \tilde{H}_p(K/K^{2R-t}) \cong \tilde{H}_p(K \cup \mathrm{C}\left(K^{2R-t}\right)).$$

We denote by ω the distinguished vertex of such cones. It is then clear that $\mathrm{V}_{p,R}(f)$ equals the ordinary p-th sub level set persistent (reduced) homology module of the function $\hat{f}_R : \mathrm{C}(K) \to \mathbb{R}$ defined by

$$\hat{f}_R(\sigma) := f(\sigma) \text{ and } \hat{f}_R(\sigma \sqcup \{\omega\}) := 2R - \min_{v \text{ vertex in } \sigma} f(v)$$

for any simplex $\sigma \in K$, and $\hat{f}_R(\omega) := -R$ by convention. Plugging these constructions together, we connect the ordinary and extended maps in the commmutative diagram:[14]



(2.29)

Note that this diagram only makes sense for parameters θ such that $F(\theta)_R$ is a function whose sub level sets are sub complexes of $\mathrm{C}(K)$, as $\mathrm{PH}\left(F(\theta)_R\right)$ is undefined otherwise. This requirement is satisfied whenever the inequality $\sup_\sigma |F(\theta)(\sigma)| < R$ holds. For simplicity, we assume that R can be chosen large enough for the inequality to hold for all parameters θ, hence the diagram (2.29) makes sense globally on M. One can always avoid this restriction by working locally on compact neighbourhoods in M.

From Theorem 4.9 in [LOT21], the subset $M' \subseteq M$ where the parametrisation F is differentiable and induces a locally constant preorder on simplices of K is a generic sub manifold. In turn, all the maps $\theta \mapsto \min_{v\in\sigma} F(\theta)(v)$ and $\theta \mapsto \max_{v\in\sigma} F(\theta)(v)$, for $\sigma \in K$ a simplex, are differentiable over M'. Therefore $\hat{F}_R : M \to \mathbb{R}^{\mathrm{C}(K)}$ is differentiable over the generic submanifold M'.

Since $\hat{F}_R$ is generically differentiable, so is $\mathrm{PH} \circ \hat{F}_R$ [LOT21, Theorem 4.9], i.e. we generically have local coordinate systems as in Def. 2.3.2 . Since the decoding map Dec_R in diagram (2.29) merely

[14] Strictly speaking, the decoding map should furthermore forget the unique unbounded interval $\langle b, +\infty \rangle$ in the barcode $\mathrm{PH}\left(\hat{f}_R\right)$, since the ordinary persistence map PH computes the barcode of a module made of non-reduced homology groups.

Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021

applies an affine transformation to the local coordinate systems and then splits them into three parts (the splitting is constant), we obtain local coordinate systems for EPH $\circ$ F. Therefore, EPH $\circ$ F is generically differentiable. □

2.4 *Binary graph classification*

We investigate whether optimising the extended persistence of wavelet signatures can be usefully applied to graph classification problems, where persistence diagrams are used as features to predict discrete, real life attributes of networks. In this setting, we aim to learn $\theta \in M$ that minimise the classification error of graphs over a training dataset.

We apply our wavelet optimisation framework to classification problems on the graph datasets `MUTAG` [DLdCD+91, KM12], `COX2` [SOW03], `DHFR` [SOW03], `NCI1` [WWK08, SSVL+11], `PROTEINS` [BOS+05, DD03] and `IMDB-B` [YV15]. The former five datasets are biochemical molecules and `IMDB-B` is a collection of social ego networks. In our models, we use persistence images [15] as a fixed vectorisation method and use a feed forward neural network to map the persistence images to classification labels. We also include the eigenvalues of the graph Laplacian as additional features; model particulars are described in the sections below.

To illustrate the effect of wavelet optimisation on different classification problems, we also perform a set of *control experiments* where for the same model architecture, we fix the wavelet and only optimise the parameters of the neural network. The control experiment functions as a baseline against which we assess the efficacy of wavelet optimisation.

We benchmark our results with two existing persistence based architectures, `PersLay`[16] and `GFL`[17]. `Perslay` optimises the vectorisation parameters and use two heat kernel signatures as fixed rather than optimisable vertex functions for computing extended persistence. `GFL` optimises and parametrises vertex functions using a graph isomorphism network[18], and computes *ordinary* sublevel and superlevel set persistence instead of extended persistence.

2.4.1 *Model Architecture*

We give a high level description of our model and relegate details and hyperparameter choices of the vectorisation method and neural network architecture to subsection 2.4.4. In our setting, the extended persistence diagrams of the optimisable wavelet signatures for each graph are vectorised as persistence images. We also include the static persistence images of a *fixed* heat kernel signature, $W(e^{-0.1x})$, as an ad-

[15] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research*, 18(1):218–252, 2017

[16] Mathieu Carrière, Frédéric Chazal, Yuichi Ike, Théo Lacombe, Martin Royer, and Yuhei Umeda. Perslay: A neural network layer for persistence diagrams and new graph topological signatures. In *International Conference on Artificial Intelligence and Statistics*, pages 2786–2796. PMLR, 2020

[17] Christoph Hofer, Florian Graf, Bastian Rieck, Marc Niethammer, and Roland Kwitt. Graph filtration learning. In *International Conference on Machine Learning*, pages 4314–4323. PMLR, 2020

[18] Bingbing Xu, Huawei Shen, Qi Cao, Yunqi Qiu, and Xueqi Cheng. Graph wavelet neural network. *arXiv preprint arXiv:1904.07785*, 2019

ditional set of features, alongside some non-persistence features. Both the optimised and static persistence diagrams are transformed into the persistence images using identical hyperparameters. We feed the optimisable and static persistence images into two separate convolutional neural networks (CNNs) with the same architecture. Similarly, we feed the non-persistence features as a vector into a separate multilayer perceptron. The outputs of the CNNs are concatenated with the outputs of the multi-layer perceptron. Finally, an affine transformation sends the concatenated vector to a real number whose sign determines the binary classification.

Wavelet Parametrisation We choose a space of wavelets spanned by 12 inverse multiquadric radial basis functions

$$h_j(x) = \left(\left(\frac{x - x_j}{\epsilon} \right)^2 + 1 \right)^{-\frac{1}{2}} \tag{2.30}$$

whose centroids x_j are located at $x_j = 2(j-1)/9,\ j = 0, \ldots, 11$. The width parameter is chosen to be the distance between the centroids, $\epsilon = 2/9$. On each dataset, we derive a numerically stable parametrisation using the procedure described in section 2.2.2; the parameters we optimise are the coefficients of the new basis given by (2.13). We initialise the parameters by fitting them via least squares to the heat kernel signature $W(e^{-10x})$ on the whole dataset of graphs.

Non-Persistence Features We also incorporate the eigenvalues of the normalised Laplacian as additional, fixed features of the graph. Since the number of eigenvalues for a given graph is equal to the number of vertices, it differs between graphs in the same dataset. To encode the information represented in the eigenvalues as a fixed length vector, we first sort the eigenvalues into a time-series; we then compute the log path signature of the time series up to level four, which is a fixed length vector in $\mathbb{R}^8$. The log-signature captures the geometric features of the path; we refer readers to [CK16] for details about path signatures. For `IMDB-B` in particular, we also include the maxima and minima of the heat kernel signatures $W(e^{-10x})$ and $W(e^{-0.1x})$ respectively of each graph.

Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788*, 2016

2.4.2 *Experimental set up*

We employ a *10 ten-fold* test-train split scheme on each dataset to measure the accuracy of our model. Each ten-fold is a set of ten

experiments, corresponding to a random partition of the dataset into ten portions. In each experiment, a different portion is selected as the test set while the model is trained on the remaining nine portions. We perform 10 ten-folds to obtain a total of 10×10 experiments, and report the accuracy of the classifier as the average accuracy over 100 such experiments. The epochs at which the accuracies were measured are specified in table 2.3.

Across all experiments, we use binary cross entropy as the loss function. We use the `Adam` optimiser[19] with learning rate `lr = 1e-3` to optimise the parameters of the neural network. The wavelet parameters are updated using stochastic gradient descent with learning rate `lr = 1e-2`, for all datasets except for `IMDB-B`, where the learning rate is set to `lr = 1e-1`. The batch sizes for each experiment are shown in table 2.4. In all experiments, we stop the optimisation of wavelet parameters at epoch 50 while the neural network parameters continue to be optimised.

[19] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014

We use the `GUDHI` library to compute persistence, and make use of the optimisation and machine learning library `PyTorch` for the construction of the graph classifications models.

2.4.3 *Results and Discussion*

In table 2.2, we present the classification accuracies of our model. For each dataset, we perform four experiments using our model, varying whether the wavelet parameter is optimised and whether additional features are included. In 2.1, we show the test accuracy of our model alongside two persistence-based graph classification architectures, `Perslay` and `GFL`, as well as other state-of-the-art graph classification architectures.

We first compare the performances of our model between cases where we optimise and fix the wavelets. In table 2.2, we see that on `MUTAG` and `DHFR`, optimising the wavelet improves the classification accuracy regardless of whether extra features are included. On `NCI1`, wavelet optimisation improves the classification accuracy only persistence features are included. When we include non-persistence features in the model, the performances of the optimised and control models are statistically indistinguishable for `NCI1`, suggesting that the non-persistence features play a more significant role in the classification. As for `COX2`, `PROTEINS`, and `IMDB-B`, optimising the wavelet coefficients do not bring about statistically significant improvements. This indicates that the initial wavelet signature – the heat kernel signature $W(e^{-10x})$ – is a locally optimal choice of wavelet for our neural network classifier.

We now compare our architecture to other persistence based architec-

tures, Perslay and GFL, where node attributes are excluded from their vertex function models. Except on PROTEINS, our wavelet optimised model matches or exceeds Perslay. While our model architecture and choice of wavelet initialisation is similar to that of Perslay, we differ in two important respects. Perslay fixes the vertex functions but optimises the weights assigned to points on the persistence diagrams, as well as the parameters of the persistence images. Our improvements on Perslay for MUTAG, DHFR, and IMDB-B indicate that vertex function optimisation yields improvements that cannot be obtained through vectorisation optimisation alone on some datasets of graphs.

Compared to GFL (without node attributes), both Perslay and our architecture achieves similar or higher classification accuracies on PROTEINS and NCI1. This supports wavelet signatures being viable models for vertex functions on those datasets. On the other hand, both Perslay and our model lag behind GFL on IMDB-B. We attribute this to the fact that IMDB-B, unlike the other bionformatics datasets, consists of densely connected graphs. The graphs in IMDB-B have diameter at most two and 14% of the graphs are cliques. This fact has two consequences. First, we expect the one-layer GIN used in GFL – a local topology summary – to be highly effective in optimising for the salient features of a graph with small diameter. Second, the extended persistence modules for cliques have zero persistence, since all vertices are assigned the same function value due to symmetry. In contrast, ordinary persistence used in GFL is able to capture the cycles in a complete graph as points with infinite persistence.

Compared to non-persistence state-of-the-art architectures in table 2.2, our model achieves competitive accuracies on MUTAG, COX2, and DHFR. For NCI1 and PROTEINS, all persistence architectures listed that exclude additional node attributes perform poorly in comparison, though PWL was able to achieve leading results with node attributes.

All in all, we observe that wavelet signatures can be an effective parametrisation of vertex functions when we use extended persistence as features for graph classification. In particular, on some bioinformatics datasets, we show that optimising the wavelet signature can lead to improvements in classification accuracy. The wavelet signature approach is complementary to the GFL approach to vertex function parametrisation as they show strengths on different datasets.

2.4.4 *Experimental Details*

Persistence Images Parameters We vectorised each of the three persistence diagrams EPH_0, $\mathrm{EPH}_1^{\mathrm{rel}}$ and $\mathrm{EPH}_1^{\mathrm{ext}}$ as a persistence image. Prior to vectorising the persistence diagrams, we apply a fixed and identical affine transformation to the values of the vertex functions across all

graphs in the dataset concerned, such that the maximum and minimum values taken across all graphs in the dataset of the initial vertex function prior to optimisation are scaled to 1 and 0 respectively. The persistence image is sampled on a 20×20 grid, whose grid points are equidistantly placed $\sigma = 1/17$ apart on the square $[-\sigma, 1+\sigma]^2$ of the persistence diagrams, where σ is the width of the Gaussian. The Gaussian centred on the birth and persistence coordinates $\langle b, p \rangle$ of each point is weighted according to its persistence

$$\omega(p) = \sin^2\Big(\frac{\pi}{2} \min\Big(\frac{p}{\sigma}, 1\Big)\Big).$$

Points with persistence $p \geqslant \sigma$ are assigned a uniform weight $\omega = 1$, else assigned a weight that diminishes to zero as $p \to 0$.

Convolutional Neural Network Architecture for Persistence Images We feed each set of three persistence images belonging to either the optimisable or static persistence diagrams as a three channel image into the following convolutional neural network to obtain a 22×22 image:

$$\mathrm{CNN} : \mathbb{R}^{3\times 20\times 20} \xrightarrow{\text{ReLU} \circ \text{Conv} \circ \text{BN2D}} \mathbb{R}^{20\times 21\times 21} \xrightarrow{\text{ReLU} \circ \text{Conv} \circ \text{DO} \circ \text{BN2D}} \mathbb{R}^{22\times 22}. \tag{2.31}$$

The function Conv denotes a convolutional layer with kernel size 2, stride 1, padding 1; BN2D denotes a 2D batch normalisation layer; and DO denotes a dropout layer with dropout probability $p = 0.5$.

Multilayer Perceptron for Non-Persistence Features We feed non-persistence features as a vector of length $n = \#\text{features}$ into the following multilayer perceptron:

$$\mathrm{MLP} : \mathbb{R}^n \xrightarrow{\text{BN} \circ \text{ReLU} \circ \text{Aff} \circ \text{BN}} \mathbb{R}^n \tag{2.32}$$

where $\mathrm{Aff} : \mathbb{R}^n \to \mathbb{R}^n$ denotes an affine transformation, and BN denotes a batch normalisation layer.

Path Encoding of Laplacian Eigenvalues For `MUTAG`, `COX2`, `DHFR`, and `NCI1`, we sort the Laplacian eigenvalues in ascending order and transform the one-dimensional sequence into a two-dimensional time series via a *delay embedding*

$$\begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_{N-1} & \lambda_N \end{bmatrix} \mapsto \left[\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \quad \begin{bmatrix} \lambda_2 \\ \lambda_3 \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \lambda_{N-2} \\ \lambda_{N-1} \end{bmatrix} \quad \begin{bmatrix} \lambda_{N-1} \\ \lambda_N \end{bmatrix} \right]. \tag{2.33}$$

For IMDB-B, we incorporate a fictitious time coordinate $t_j = 2(j-1)/(N-1)$ for $j = 1, \dots N$ as the second coordinate instead of a 'delayed' eigenvalue:

$$\begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_{N-1} & \lambda_N \end{bmatrix} \mapsto \left[\begin{bmatrix} \lambda_1 \\ t_1 \end{bmatrix} \quad \begin{bmatrix} \lambda_2 \\ t_2 \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \lambda_{N-1} \\ t_{N-1} \end{bmatrix} \quad \begin{bmatrix} \lambda_N \\ t_N \end{bmatrix} \right]. \tag{2.34}$$

	Non-Persistence State-of-the-Art				Persistence Based					
	P-SAN [NAK16]	RetGK [ZWX+18]	GIN [XSC+19]	FGSD [VZ17]	PWL [RBB19]	GFL [HGR+20b]		Perslay [CCI+20]	Control This paper	Wavelet Opt. This paper
Node attr.	Yes		No		Yes	No				
MUTAG	92.6	90.3	89.4	92.1	90.5	–	–	89.8	$89.0_{\pm 0.6}$	$90.4_{\pm 1.3}$
COX2	–	80.1	–	–	–	–	–	80.9	$80.8_{\pm 0.4}$	$80.8_{\pm 1.0}$
DHFR	–	81.5	–	–	–	–	–	80.3	$80.0_{\pm 0.4}$	$81.0_{\pm 0.9}$
NCI1	78.6	84.5	82.7	79.8	85.6	77.2	71.2	73.5	$74.3_{\pm 0.3}$	$74.4_{\pm 0.3}$
PROTEINS	75.9	75.8	76.2	73.4	75.9	73.4	74.1	74.8	$74.5_{\pm 0.4}$	$74.6_{\pm 0.6}$
IMDB-B	71.0	71.9	75.1	73.6	73.0	–	74.5	71.2	$71.6_{\pm 0.9}$	$72.0_{\pm 0.7}$
# Ten-folds	10	10	1	1	10	1	1	10	10	10

Table 2.1: Binary classification accuracy on datasets of graphs. The performance of our model is reported in the column Wavelet Opt. on the right hand side. The accuracies of the control model, where the wavelet parameters are fixed to the initial values, are shown in the column Control. Both these models use additional features (see 2.4.1). The accuracies of our model are the means over 10 ten-folds, recorded at epochs reported in table 2.3. We also provide standard deviations of the 10 mean accuracies of each ten-fold. For other architectures, we indicate whether their accuracies were reported as averages over 1 ten-fold or 10 ten-fold in the bottom row of the table.

	Persistence Only		Non-Persistence Features incl.	
	Control	Wavelet Opt.	Control	Wavelet Opt.
MUTAG	$89.2_{\pm 0.6}$	$89.8_{\pm 0.8}$	$89.0_{\pm 0.6}$	$90.4_{\pm 0.4}$
COX2	$79.6_{\pm 1.0}$	$79.4_{\pm 0.7}$	$80.8_{\pm 1.0}$	$80.8_{\pm 1.0}$
DHFR	$79.9_{\pm 0.4}$	$80.4_{\pm 0.4}$	$80.3_{\pm 0.9}$	$81.0_{\pm 0.9}$
NCI1	$73.7_{\pm 0.2}$	$74.3_{\pm 0.5}$	$74.3_{\pm 0.3}$	$74.4_{\pm 0.3}$
PROTEINS	$72.9_{\pm 0.3}$	$73.0_{\pm 0.4}$	$74.5_{\pm 0.4}$	$74.6_{\pm 0.6}$
IMDB-B	$68.3_{\pm 0.5}$	$68.6_{\pm 0.7}$	$71.6_{\pm 0.9}$	$72.0_{\pm 0.7}$

Table 2.2: Binary classification accuracy of our model where we vary whether non-Persistence features are included and whether the wavelet is optimised. The reported accuracies are the mean over 10 ten-folds, recorded at epochs reported in table 2.3. We also provide standard deviations of the 10 mean accuracies of each ten-fold. See section 2.4.1 for the particulars about the non-persistence features.

	Persistence Only		Non-Persistencee Features incl.	
	Control	Wavelet Opt.	Control	Wavelet Opt.
MUTAG	25	125	25	75
COX2	50	50	25	25
DHFR	125	250	125	45
NCI1	270	270	500	370
PROTEINS	50	50	125	125
IMDB-B	100	25	75	50

Table 2.3: Epochs at which accuracies in table 2.2 are recorded.

2.5 *Conclusion*

We have presented a framework for equipping any graph G with a set of extended persistence diagrams $\mathrm{EPH} \circ F : M \to \mathbf{Bar}^4$ parametrised over a manifold M, a parameter space for the graph's wavelet signature. We

	MUTAG	COX2	DHFR	NCI1	IMDB-B
# graphs	188	467	756	4110	1000
batch size	10	9	11	20	50

Table 2.4: Batch sizes in the graph classification experiments for different datasets described in 2.4.

described how wavelet signatures can be parametrised and interpreted. Given a function on extended persistence diagrams Out : $\mathbf{Bar}^4 \to \mathbb{R}$ that is differentiable, we have shown how a loss function $\mathcal{L} = \text{Out} \circ \text{EPH} \circ F$ can be generically differentiable with respect to $\theta \in M$ as $\mathcal{L}$. Thus, we can apply gradient descent methods to optimise the extended persistence diagrams of a graph to minimise $\mathcal{L}$.

We applied this framework to a graph classification architecture where the wavelet signature is optimised for classification accuracy. We are able to demonstrate an increase in accuracy on several benchmark datasets where the wavelet is optimised, and perform competitively with state-of-the-art persistence based graph classification architectures.

3
A Gradient Sampling Algorithm for Stratified Maps with Applications to Topological Data Analysis

Abstract

We introduce a novel gradient descent algorithm extending the well-known Gradient Sampling methodology to the class of stratifiably smooth objective functions, which are defined as locally Lipschitz functions that are smooth on some regular pieces—called the strata—of the ambient Euclidean space. For this class of functions, our algorithm achieves a sub-linear convergence rate. We then apply our method to objective functions based on the (extended) persistent homology map computed over lower-star filters, which is a central tool of Topological Data Analysis. For this, we propose an efficient exploration of the corresponding stratification by using the Cayley graph of the permutation group. Finally, we provide benchmark and novel topological optimization problems, in order to demonstrate the utility and applicability of our framework.

3.1 *Introduction*

In its most general instance nonsmooth, non convex, optimization seek to minimize a locally Lipschitz objective or loss function $f : \mathbb{R}^n \to \mathbb{R}$. Without further regularity assumptions on f, most algorithms—such as the usual Gradient Descent with learning rate decay, or the Gradient Sampling method—are only guaranteed to produce iterates whose subsequences are asymptotically stationary, without explicit convergence rates. Meanwhile, when restricted to the class of min-max functions (like the maximum of finitely many smooth maps), stronger guarantees such as convergence rates can be obtained [HSS17]. This illustrates the common paradigm in nonsmooth optimization: the richer the structure in the irregularities of f, the better the guarantees we can expect from an optimization algorithm. Note that there are algorithms specifically tailored to deal with min-max functions, e.g. [Ber75].

Another example are bundle methods [FGG04a, FGG04b, HMM07]. They consist, roughly, in constructing successive linear approximations of f as a proxy for minimization. Their convergence guarantees are strong, especially with an additional *semi-smoothness* property of f [Bih84, Mif77]. Other types of methods, like the variable metric ones, can also benefit from the semi-smoothness hypothesis [VL01]. We observe that it is not merely the structure on f that impacts the convergence properties in the algorithm, but also the amount of information about f that we assume to be accessible in practical computations. For instance, the bundle method [LV98] assumes that the Hessian matrix, when defined locally, can be computed. For great accounts of the theory and practice in non smooth optimization, we refer to [BKM14, Kiw06, Sho12].

The ability to cut $\mathbb{R}^n$ in well-behaved pieces where f is regular, is another type of important structure. Examples, in increasing order of generality, are semi-algebraic, (sub)analytic, definable and tame (w.r.t. an o-minimal structure), and Whitney stratified maps. For such objective functions, the traditional gradient descent (GD), or a stochastic version of it, converges to stationary points [DDKL20]. In order to obtain further theoretical guarantees such as convergence rates, it is natural to design optimization algorithms specifically taylored to regular maps. Indeed, they have strong properties, e.g. tame maps are semi-smooth [Iof09], and the generalized gradients of Whitney stratified maps are closely related to (restricted) gradients of f along strata [BDLS07]. Besides, strong convergence guarantees under (mainly) the Kurdyka–Łojasiewicz assumption [ABS13, Nol14] exist, including the class of semi-algebraic maps. Our method is related to this line of work, in that we exploit the strata of $\mathbb{R}^n$ in which f is smooth.

The motivation of this work stems from Topological Data Analysis

(TDA), where geometric objects such as graphs are described by means of computable and topological descriptors. Persistent Homology is one such descriptor, and has been successfully applied in various data science problems. Persistent Homology describes graphs and more generally simplicial complexes over n nodes by means of a signature called the barcode/persistence diagram $\mathrm{PH}(x)$. Here x is a function on the nodes, which we view as a vector in $\mathbb{R}^n$. Loosely speaking, $\mathrm{PH}(x)$ is a finite multi-set of points in the half-plane $\{(b,d) \in \mathbb{R}^2,\ d \geqslant b\}$ that encode topological and geometrical information about the graph and the function x.

Barcodes form a metric space **Bar** when equipped with the standard metrics of TDA, the so-called *bottleneck* and *Wasserstein* distances, and the persistence map $\mathrm{PH} : \mathbb{R}^n \to \mathbf{Bar}$ is locally Lipschitz [CSEH07, CSEHM10]. However **Bar** is not Euclidean nor a smooth manifold, thus hindering the use of these topological descriptors in standard statistical or machine learning pipelines. Still, there exist natural notions of differentiability for maps in and out of **Bar** [LOT21]. In particular, the persistence map $\mathrm{PH} : \mathbb{R}^n \to \mathbf{Bar}$ restricts to a locally Lipschitz, smooth map on a stratification of $\mathbb{R}^n$ by polyhedra. If we compose the persistence map with a smooth and Lipschitz map $V : \mathbf{Bar} \to \mathbb{R}$, the resulting objective (or loss) function

$$f : \mathbb{R}^n \xrightarrow{\ \mathrm{PH}\ } \mathbf{Bar} \xrightarrow{\ V\ } \mathbb{R}$$

is itself Lipschitz and smooth on the various strata. From [DDKL20], and as recalled in [CCG+21], classical (Stochastic) Gradient Descent on f asymptotically converges to stationary points. Similarly, the Gradient Sampling (GS) method asymptotically converges. See [SWB21] for an application of GS to topological optimization.

Nonetheless, we want to design algorithms that take advantage of the structure in the irregularities of the persistence PH, in order to get better theoretical guarantees. For instance, it is possible to locally integrate gradients of PH—whenever defined—to stabilise the iterates [SWB21], or to add a regularization term to f acting as a smoothing operator [CD20]. In this work, we rather exploit the stratification of $\mathbb{R}^n$ induced by PH, as it is very easy to manipulate. Indeed, we can efficiently access elements x' in neighboring strata as well as estimate the distance to these strata.

For this reason, we believe that persistence based objective functions f form a rich playground for non smooth optimisation. However, it is also per se an emerging field in TDA with applications in point cloud inference [GHO16], surface reconstruction [BGGSSG20b], shape matching [PSO18], graph classification [HGR+20b, YL21], topological regularisation for Generative Modelling [MHRB20b, HKND19, GNDS20b], Image Segmentation [HLSC19, COB+19], supervised learn-

ing [CNBW19] and dimension reduction [Kac20] to name a few.

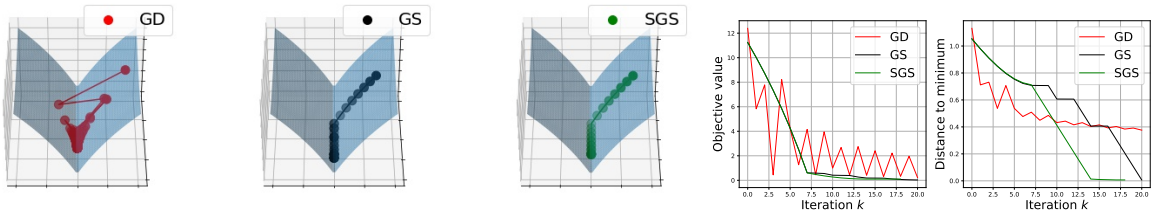


Figure 3.1: A proof-of-concept comparison between different gradient descent techniques. The objective function $f : (z_1, z_2) \in \mathbb{R}^2 \to 10\log(1 + |z_1|) + z_2^2 \in \mathbb{R}$ (blue surface) attains its minimum at $x_* = (0, 0)$ and is not smooth along the line $\{z_1 = 0\}$. In particular $\|\nabla f\| > 1$ around x_*, thus the gradient norm cannot be used as a stopping criterion. The traditional GD, for which updates are given by $x_{k+1} = x_k - \frac{\lambda_0}{k+1}\nabla f(x_k)$, oscillates around $\{z_1 = 0\}$ due to the non-smoothness of f and asymptotically converges toward x_* because of the decaying learning rate $\frac{\lambda_0}{k+1}$. In the meantime, non-smooth optimisation methods that sample points around x_k to produce reliable descent directions converge in finite time. Namely, the classical Gradient Sampling method randomly samples 3 points and manages to reach an (ϵ, η)-stationary point of f in $\sim 20.6 \pm 3.9$ iterations (averaged over 100 experiments), while our stratified approach improves on this by leveraging the fact that we explicitly have access to the two strata $\{z_1 < 0\}$ and $\{z_1 > 0\}$ where f is differentiable. In particular, we only sample additional points when x_k is ϵ-near the line $\{z_1 = 0\}$, and reach an (ϵ, η)-stationary point in 18 iterations. Right plots showcase the evolution of the objective value $f(x_k)$ and the distance to the minimum $\|x_k - x_*\|$ across iterations. Parameters: $x_0 = (0.8, 0.8)$, $\lambda_0 = 10^{-1}$, $\epsilon = 10^{-1}$, $\eta = 10^{-2}$.

3.1.1 *Contributions and outline of contents*

Our new method, called *Stratified Gradient Sampling* (SGS), is a variation of the established GS algorithm, whose main steps for updating the current iterate $x_k \in \mathbb{R}^n$ we recall in alg. 1 below. In the closing remarks

Algorithm 1 An update step with the Gradient Sampling algorithm

1: Sample $m \geqslant n + 1$ points $x_k^1, \cdots, x_k^m$ in a ball $B(x_k, \epsilon)$
2: Compute *approximate subgradient* $G_k := \{\nabla f(x_k), \nabla f(x_k^1), \cdots, \nabla f(x_k^m)\}$
3: Compute *descent direction* $g_k := \operatorname{argmin}\{|g|^2,\ g \text{ in convex hull of } G_k\}$
4: Find *step size* $t_k \geqslant 0$ so that $f(x_k - t_k g_k) \leqslant f(x_k) - \beta t_k |g_k|^2$ ($\beta \in (0, 1)$ hyperparameter)
5: Ensure that f is differentiable at $x_{k+1} := x_k - t_k g_k$ by small perturbations

of a recent overview of the GS methodology [BCLO19], the authors suggest that the GS theory and practice could be enhanced assuming some extra structure on the non differentiability of f.

In this work, we deal with *stratifiably smooth maps*, for which the non differentiability is organized around smooth submanifolds that together form a stratification of $\mathbb{R}^n$. In section 3.2, we review some background material in nonsmooth analysis and define stratifiably smooth maps $f : \mathbb{R}^n \to \mathbb{R}$, a variant of the Whitney stratifiable functions from [BDLS07] for which we do not impose any Whitney regularity on the gluing between adjacent strata of $\mathbb{R}^n$, but rather enforce that there exist local C^2-extensions of the restrictions of f to top-dimensional strata.

In order to update the current iterate x_k when minimizing a stratifiably smooth objective function f, we introduce a new descent direction g_k. As in GS, g_k is obtained in our new SGS algorithm by collecting the gradients of samples around x_k in an approximate subgradient G_k, and then by taking the element with minimal norm in the convex set

generated by G_k. A key difference with GS is that we only need to sample as many points around x_k as there are distinct strata close by, compare with the $m \geqslant n+1$ samples of alg. 1. In section 3.2.7, we show that we indeed obtain a descent direction, i.e., that we have the descent criterion $f(x_k - t_k g_k) \leqslant f(x_k) - \beta t_k \|g_k\|^2$ (as in Line 4 of alg. 1) for a suitable choice of step size t_k.

Our SGS algorithm is detailed in section 3.3.1 and its analysis in section 3.3.2. The convergence of the original GS methodology crucially relies on the sample size $m \geqslant n+1$ in order to apply the Carathéodory Theorem to subgradients. Differently, our convergence analysis relies on the fact that the gradients of f, when restricted to neighboring strata, are locally Lipschitz. Hence, our proof of asymptotic convergence to stationary points (Theorem 3.3.3) is substantially different. In Theorem 3.3.5, we determine a convergence rate of our algorithm that holds for any proper stratifiably smooth map, which is an improvement over the guarantees of GS for general locally Lipschitz maps. Finally, in section 3.3.3, we adapt our method and results to the case where only estimated distances to nearby strata are available.

In section 3.4, we introduce the persistence map PH over a simplicial complex K, which gives rise to a wide class of stratifiably smooth objective functions with rich applications in TDA. We characterize strata around the current iterate (i.e., filter function) x_k by means of the permutation group over n elements, where n is the number of vertices in K. Then, the Cayley graph associated to the permutation group allows us to efficiently explore the set of neighboring strata by increasing order of distances to x_k, that are needed to compute descent directions.

Section 3.5 is devoted to the implementation of the SGS algorithm for the optimization of persistent homology-based objective functions f. In 3.5.1, we provide empirical evidence that SGS behaves better than GD and GS with a simple experiment about minimization of total persistence. In 3.5.3 and 3.5.2, we consider two novel topological optimization problems which we believe are of interest in real-world applications. On the one hand, the *Topological Template Registration* of a filter function x defined on a complex K, is the task of finding a filter function x' over a smaller complex K' that preserves the barcode of x. On the other hand, given a Mapper graph G, which is a standard visualization tool for arbitrary data sets [1], we can bootstrap the data set in order to produce multiple bootstrapped graphs G_i. The *Topological Mean* is then the task of finding a new graph G^* whose barcode is as close as possible to the mean of the barcodes associated to the graphs G_i. As a result we obtain a smoothed version G^* of the Mapper graph G in which spurious and non-relevant graph attributes are removed.

[1] Gurjeet Singh, Facundo Mémoli, and Gunnar Carlsson. Topological methods for the analysis of high dimensional data sets and 3D object recognition. In *4th Eurographics Symposium on Point-Based Graphics (SPBG 2007)*, pages 91–100. The Eurographics Association, 2007

3.2 *A direction of descent for stratifiably smooth maps*

In this section we define the class of stratifiably smooth maps whose optimisation is at stake in this work. For such maps we can define an approximate sub gradient and the resulting descent direction, which is the key ingredient of our algorithm. We first collect some useful background in nonsmooth analysis, essentially from [Cla90]. Throughout, $f : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz (non necessary smooth, nor convex) and proper (i.e., compact sublevel sets) function, which we want to minimize.

Frank H Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990

3.2.1 *Nonsmooth Analysis*

First-order information of f at $x \in \mathbb{R}^n$ in a direction $v \in \mathbb{R}^n$ is captured by its *generalized directional derivative*:

$$f^\circ(x; v) = \limsup_{y \to x, t \downarrow 0^+} \frac{f(y + tv) - f(y)}{t}, \tag{3.1}$$

Besides, the *generalized gradient* is the following set of linear subapproximations:

$$\partial f(x) := \{\zeta \in \mathbb{R}^n,\ f^\circ(x; v) \geqslant \langle \zeta, v \rangle \text{ for all } v \in \mathbb{R}^n\}. \tag{3.2}$$

Given an arbitrary set $S \subset \mathbb{R}^n$ of Lebesgue measure 0 around x, we have an alternative description of the generalized gradient in terms of limits of surrounding gradients, whenever defined:

$$\partial f(x) = \overline{\text{co}}\{\lim \nabla f(x_i) \,|\, x_i \to x,\ \nabla f(x_i) \text{ is defined, } \lim \nabla f(x_i) \text{ exists, } x_i \notin S\}, \tag{3.3}$$

where $\overline{\text{co}}$ is the operation of taking the closure of the convex hull.[2] The duality between generalized directional derivatives and gradients is captured by the equality:

[2] Here the equality holds as well (by some compactness argument), when taking the convex hull without closure. As we take closed convex hulls later on, we chose not to introduce this subtlety explicitly.

$$f^\circ(x; v) = \max\{\langle \zeta, v \rangle, \zeta \in \partial f(x)\}. \tag{3.4}$$

The *Goldstein subgradient* [3] is an ϵ-relaxation of the generalized gradient:

$$\partial_\epsilon f(x) = \overline{\text{co}}\{\lim \nabla f(x_i) \,|\, x_i \to x',\ \nabla f(x_i) \text{ is defined , } \lim \nabla f(x_i) \text{ exists , } |x - x'| \leqslant \epsilon\}. \tag{3.5}$$

[3] AA Goldstein. Optimization of lipschitz continuous functions. *Mathematical Programming*, 13(1):14–22, 1977

Given $x \in \mathbb{R}^n$, we say that:

$$x \text{ is a } \textit{stationary point} \text{ (for } f) \text{ if } 0 \in \partial f(x).$$

Any local minimum is stationary, and conversely if f is convex. We also have weaker notions. Namely, given $\epsilon, \eta > 0$,

$$\begin{aligned} x \text{ is } \epsilon\textit{-stationary} \text{ if } 0 \in \partial_\epsilon f(x); \text{ and} \\ x \text{ is } (\epsilon, \eta)\textit{-stationary} \text{ if } d(0, \partial_\epsilon f(x)) \leqslant \eta. \end{aligned}$$

3.2.2 *Stratifiably smooth maps*

Desirable properties for an optimization algorithm is that it produces iterates $(x_k)_k$ that either converge to an (ϵ, η)-stationary point in finitely many steps, or whose subsequences (or some of them) converge to an ϵ-stationary point. For this, we work in the setting of objective functions that are smooth when restricted to submanifolds, that together partition $\mathbb{R}^n$.

Definition 3.2.1. A *stratification* $\mathcal{X} = \{X_i\}_{i \in I}$ of a closed subset $\mathbb{X} \subseteq \mathbb{R}^n$ is a locally finite partition of $\mathbb{X}$ by smooth submanifolds X_i -called *strata*- such that for $i \neq j \in I$:

$$\overline{X_i} \cap X_j \neq \varnothing \Rightarrow X_j \subseteq \overline{X_i} \backslash X_i.$$

This makes $(\mathbb{X}, \mathcal{X})$ into a *stratified space*.

Note that we do not impose any (usually needed) gluing conditions between adjacent strata, as we do not require them in the analysis. In particular, semi-algebraic, subanalytic, or definable subsets of $\mathbb{R}^n$, together with Whitney stratified sets are stratified in the above weak sense. We next define the class of maps f with smooth restrictions $f_{|X_i}$ to strata X_i of some stratification $\mathcal{X}$, inspired by the Whitney stratifiable maps of [BDLS07] (there $\mathcal{X}$ is required to be Whitney) and the *stratifiable functions* of [DIL15], however we further require that the restrictions $f_{|X_i}$ admit local extensions of class C^2.

Jérôme Bolte, Aris Daniilidis, Adrian Lewis, and Masahiro Shiota. Clarke subgradients of stratifiable functions. *SIAM Journal on Optimization*, 18(2):556–572, 2007

Dmitriy Drusvyatskiy, Alexander D Ioffe, and Adrian S Lewis. Clarke subgradients for directionally lipschitzian stratifiable functions. *Mathematics of Operations Research*, 40(2):328–349, 2015

Definition 3.2.2. The map $f : \mathbb{R}^n \to \mathbb{R}$ is *stratifiably smooth* if there exists a stratification $\mathcal{X}$ of $\mathbb{R}^n$, such that for each top-dimensional stratum $X_i \in \mathcal{X}$, the restriction $f_{|X_i}$ admits an extension f_i of class C^2 in a neighborhood of X_i.

Remark 3.2.3. The slightly weaker assumption that the extension f_i is continuously differentiable with locally Lipschitz gradient would have also been sufficient for our purpose.

We denote by X_x the stratum containing x, and by $\mathcal{X}_x \subseteq \mathcal{X}$ the set of strata containing x in their closures. More generally, for $\epsilon > 0$, we let $\mathcal{X}_{x,\epsilon} \subseteq \mathcal{X}$ be the set of strata X_i such that the closure of the ball $B(x, \epsilon)$ has non-empty intersection with the closure of X_i. Local finiteness in the definition of a stratification implies that $\mathcal{X}_{x,\epsilon}$ (and $\mathcal{X}_x$) is finite.

If f is stratifiably smooth and $X_i \in \mathcal{X}_x$ is a top-dimensional stratum, there is a well-defined limit gradient $\nabla_{X_i} f(x)$, which is the unique limit of the gradients $\nabla f_{|X_i}(x_n)$ where $x_n \in X_i$ is any sequence converging to x. Indeed, this limit exists and does not depend on the choice of sequence since $f_{|X_i}$ admits a local C^2 extension f_i. The following result states that the generalized gradient at x can be retrieved from these finitely many limit gradients along the various adjacent top-dimensional strata.

Proposition 3.2.4. *If f is stratifiably smooth, then for any $x \in \mathbb{R}^n$ we have:*

$$\partial f(x) = \mathrm{co}\{\nabla_{X_i} f(x),\ X_i \in \mathcal{X}_x \text{ is of dimension } n\}.$$

More generally, for $\epsilon > 0$:

$$\partial_\epsilon f(x) = \overline{\mathrm{co}}\{\nabla_{X_i} f(x') \,|\, |x' - x| \leqslant \epsilon,\ X_i \in \mathcal{X}_{x'} \subseteq \mathcal{X}_{x,\epsilon} \text{ is of dimension } n\}.$$

Proof. We show the first equality only, as the second can be proven along the same lines. We use the description of $\partial f(x)$ in terms of limiting gradients from (3.3), which clearly implies the inclusion of the right-hand side in $\partial f(x)$. Conversely, let S be the union of strata in $\mathcal{X}_x$ with positive codimension, which is of measure 0. Let x_i be a sequence avoiding S, converging to x, such that $\nabla f(x_i)$ converges as well. Since $\mathcal{X}_x$ is finite, up to extracting a subsequence, we can assume that all x_i are in the same top-dimensional stratum $X_i \in \mathcal{X}_x$. Consequently, $\nabla f(x_i)$ converges to $\nabla_{X_i} f(x)$. □

3.2.3 *Direction of descent*

Thinking of x as a current position, we look for a direction of *(steepest) descent*, in the sense that a perturbation of x in this direction produces a (maximal) decrease of f. Given $\epsilon \geqslant 0$, we let $g(x, \epsilon)$ be the projection of the origin on the convex set $\partial_\epsilon f(x)$. Equivalently, $g(x, \epsilon)$ solves the minimization problem:

$$g(x, \epsilon) = \mathrm{argmin}\{\|g\|,\ g \in \partial_\epsilon f(x)\}. \tag{3.6}$$

Introduced in [Gol77], the direction $-g(x, \epsilon)$ is a good candidate of direction for descent, as we explain now. Since $g(x, \epsilon)$ is the projection of the origin on the convex closed set $\partial_\epsilon f(x)$, we have the classical inequality $\langle g(x, \epsilon), g(x, \epsilon) - g\rangle \leqslant 0$ that holds for any g in the Goldstein subgradient at x. Equivalently,

$$\forall g \in \partial_\epsilon f(x),\ \langle -g(x, \epsilon), g\rangle \leqslant -\|g(x, \epsilon)\|^2. \tag{3.7}$$

AA Goldstein. Optimization of lipschitz continuous functions. *Mathematical Programming*, 13(1):14–22, 1977

Informally, if we think of a small perturbation $x - tg(x, \epsilon)$ of x along this direction, for $t > 0$ small enough, then $f(x - tg(x, \epsilon)) \approx f(x) - t\langle \nabla f(x), g(x, \epsilon)\rangle$. Using (3.7), since $\nabla f(x) \in \partial_\epsilon f(x)$, we deduce that $f(x - tg(x, \epsilon)) \leqslant f(x) - t\|g(x, \epsilon)\|^2$. So f locally decreases at linear rate in the direction $-g(x, \epsilon)$. This intuition relies on the fact that $\nabla f(x)$ is well-defined so as to provide a first order approximation of f around x, and that t is chosen small enough. In order to make this reasoning completely rigorous, we need the following well-known result (see for instance [Cla90]):

Frank H Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990

Theorem 3.2.5 (Lebourg Mean value Theorem)**.** *Let $x, x' \in \mathbb{R}^n$. Then there exists some $y \in [x, x']$ and some $w \in \partial f(y)$ such that:*

$$f(x') - f(x) = \langle w, x' - x\rangle.$$

Let $t > 0$ be lesser than $\frac{\epsilon}{\|g(x,\epsilon)\|}$, in order to ensure that $x' := x - tg(x,\epsilon)$ is ϵ-close to x. Then by the mean value theorem (and Proposition 3.2.4), we have that

$$f(x - tg(x,\epsilon)) - f(x) = -t\langle w, g(x,\epsilon)\rangle$$

for some $w \in \partial_\epsilon f(x)$. (3.7) yields

$$\forall t \leqslant \frac{\epsilon}{\|g(x,\epsilon)\|}, \qquad f(x - tg(x,\epsilon)) \leqslant f(x) - t\|g(x,\epsilon)\|^2, \tag{3.8}$$

as desired.

In practical scenarios however, it is unlikely that the exact descent direction $-g(x,\epsilon)$ could be determined. Indeed, from (3.6), it would require the knowledge of the set $\partial_\epsilon f(x)$, which consists of infinitely many (limits of) gradients in an ϵ-neighborhood of x. We now build, provided f is stratifiably smooth, a faithful approximation $\tilde{\partial}_\epsilon f(x)$ of $\partial_\epsilon f(x)$, by collecting gradient information in the strata that are ϵ-close to x.

For each top-dimensional stratum $X_i \in \mathcal{X}_{x,\epsilon}$, let x_i be an arbitrary point in $\overline{X}_i \cap \overline{B}(x,\epsilon)$. Define

$$\tilde{\partial}_\epsilon f(x) := \overline{\text{co}}\{\nabla_{X_i} f(x_i),\ X_i \in \mathcal{X}_{x,\epsilon}\}. \tag{3.9}$$

Of course, $\tilde{\partial}_\epsilon f(x)$ depends on the choice of each $x_i \in X_i$. But this will not matter for the rest of the analysis, as we will only rely on the following approximation result which holds for arbitrary choices of points x_i:

Proposition 3.2.6. *Let $x \in \mathbb{R}^n$ and $\epsilon > 0$. Assume that f is stratifiably smooth. Let L be a Lipschitz constant of the gradients ∇f_i restricted to $\overline{B}(x,\epsilon) \cap X_i$, where f_i is some local C^2 extension of $f_{|X_i}$, and $X_i \in \mathcal{X}_{x,\epsilon}$ is top dimensional. Then we have:*

$$\tilde{\partial}_\epsilon f(x) \subseteq \partial_\epsilon f(x) \subseteq \tilde{\partial}_\epsilon f(x) + \overline{B}(0, 2L\epsilon).$$

In particular, $d_H(\tilde{\partial}_\epsilon f(x), \partial_\epsilon f(x)) \leqslant 2L\epsilon$.

Note that, since the f_i are of class C^2, their gradients are locally Lipschitz, hence by compactness of $\overline{B}(x,\epsilon)$, the existence of the Lipschitz constant L above is always guaranteed.

Proof. From Proposition 3.2.4, we have

$$\partial_\epsilon f(x) = \overline{\text{co}}\{\nabla_X f(x') \,|\, |x' - x| \leqslant \epsilon,\ X \in \mathcal{X}_{x'} \subseteq \mathcal{X}_{x,\epsilon} \text{ is of dimension } n\ \}.$$

This yields the inclusion $\tilde{\partial}_\epsilon f(x) \subseteq \partial_\epsilon f(x)$. Now, let $x' \in \mathbb{R}^n$, $|x' - x| \leqslant \epsilon$, and let $X_i \in \mathcal{X}_{x'} \subseteq \mathcal{X}_{x,\epsilon}$ be a top-dimensional stratum touching x'. Based on how x_i is defined in (3.9), we have that x' and x_i both

belong to $\overline{B}(x,\epsilon)$, and they both belong to the stratum X_i. Therefore, $|\nabla_{X_i} f(x') - \nabla_{X_i} f(x_i)| \leqslant 2L\epsilon$, and so $\nabla_{X_i} f(x') \in \tilde{\partial}_\epsilon f(x) + \overline{B}(0, 2L\epsilon)$. The result follows from the fact that $\tilde{\partial}_\epsilon f(x) + \overline{B}(0, 2L\epsilon)$ is convex and closed. □

Recall from (3.8) that the (opposite to the) descent direction $-g(x,\epsilon)$ is built as the projection of the origin onto $\partial_\epsilon f(x)$. Similarly, we define our approximate descent direction as $-\tilde{g}(x,\epsilon)$, where $\tilde{g}(x,\epsilon)$ is the projection of the origin onto the convex closed set $\tilde{\partial}_\epsilon f(x)$:

$$\tilde{g}(x,\epsilon) = \operatorname{argmin}\{\|\tilde{g}\|,\ \tilde{g} \in \tilde{\partial}_\epsilon f(x)\}. \tag{3.10}$$

We show that this choice yields a direction of decrease of f, in a sense similar to (3.8).

Proposition 3.2.7. *Let f be stratifiably smooth, and let x be a non-stationary point. Let $0 < \beta < 1$, and $\epsilon_0 > 0$. Denote by L a Lipschitz constant for all gradients of the restrictions f_i to the ball $\overline{B}(x,\epsilon_0)$ (as in Proposition 3.2.6). Then:*

(i) *For $0 < \epsilon \leqslant \epsilon_0$ small enough we have $\epsilon \leqslant \frac{1-\beta}{2L}\|\tilde{g}(x,\epsilon)\|$; and*

(ii) *For such ϵ, we have $\forall t \leqslant \frac{\epsilon}{\|\tilde{g}(x,\epsilon)\|}$, $f(x - t\tilde{g}(x,\epsilon)) \leqslant f(x) - \beta t\|\tilde{g}(x,\epsilon)\|^2$.*

Proof. Saying that x is non-stationary is equivalent to the inequality $\|g(x,0)\| > 0$. We show that the map $\epsilon \in \mathbb{R}^+ \mapsto \|g(x,\epsilon)\| \in \mathbb{R}^+$, which is non-increasing, is continuous at 0^+. Let ϵ be small enough such that the sets of strata incident to x are the same that meet with the ϵ-ball around x, i.e., $\mathcal{X}_{x,\epsilon} = \mathcal{X}_x$. Such an ϵ exists since there are finitely many strata that meet with a sufficiently small neighborhood of x. Of course, all smaller values of ϵ enjoy the same property. By Proposition 3.2.4, we then have the nesting

$$\partial f(x) \subseteq \partial_\epsilon f(x) \subseteq \partial f(x) + \overline{B}(0, 2L\epsilon),$$

where L is a Lipschitz constant for the gradients in neighboring strata. In turn, $0 \leqslant \|g(x,0)\| - \|g(x,\epsilon)\| \leqslant 2L\epsilon$. In particular, $\|g(x,\epsilon)\| \to \|g(x,0)\| > 0$ as ϵ goes to 0, hence $\epsilon = o(\|g(x,\epsilon)\|)$. Besides, the inclusion $\tilde{\partial}_\epsilon f(x) \subseteq \partial_\epsilon f(x)$ (Proposition 3.2.6) implies that $\|\tilde{g}(x,\epsilon)\| \geqslant \|g(x,\epsilon)\| > 0$. This yields $\epsilon = o(\|\tilde{g}(x,\epsilon)\|)$ and so item **(i)** is proved.

We now assume that ϵ satisfies the inequality of item **(i)**, and let $0 \leqslant t \leqslant \frac{\epsilon}{\|\tilde{g}(x,\epsilon)\|}$. By the Lebourg mean value theorem, there exists a $y \in [x, x - t\tilde{g}(x,\epsilon)]$ and some $w \in \partial f(y)$ such that:

$$f(x - t\tilde{g}(x,\epsilon)) - f(x) = t\langle w, -\tilde{g}(x,\epsilon)\rangle.$$

Since $t \leqslant \frac{\epsilon}{\|\tilde{g}(x,\epsilon)\|}$, y is at distance no greater than ϵ from x. In particular, w belongs to $\partial_\epsilon f(x)$. From Proposition 3.2.6, there exists some $\tilde{w} \in \tilde{\partial}_\epsilon f(x)$ at distance no greater than $2L\epsilon$ from w. We then rewrite:

$$f(x - t\tilde{g}(x,\epsilon)) - f(x) = t\langle w - \tilde{w}, -\tilde{g}(x,\epsilon)\rangle + t\langle \tilde{w}, -\tilde{g}(x,\epsilon)\rangle. \tag{3.11}$$

On the one hand, by the Cauchy-Schwarz inequality:

$$\langle w - \tilde{w}, -\tilde{g}(x,\epsilon)\rangle \leqslant |w - \tilde{w}| \cdot \|\tilde{g}(x,\epsilon)\| \leqslant 2L\epsilon\|\tilde{g}(x,\epsilon)\| \leqslant (1-\beta)\|\tilde{g}(x,\epsilon)\|^2, \tag{3.12}$$

where the last inequality relies on the assumption that $\epsilon \leqslant \frac{1-\beta}{2L}\|\tilde{g}(x,\epsilon)\|$. On the other hand, since $\tilde{g}(x,\epsilon)$ is the projection of the origin onto $\tilde{\partial}_\epsilon f(x)$, we obtain $\langle \tilde{g}(x,\epsilon) - \tilde{w}, \tilde{g}(x,\epsilon)\rangle \leqslant 0$, or equivalently:

$$\langle \tilde{w}, -\tilde{g}(x,\epsilon)\rangle \leqslant -\|\tilde{g}(x,\epsilon)\|^2. \tag{3.13}$$

Plugging the inequalities of eqs. (3.12) and (3.13) into eq. (3.11) proves item **(ii)**. □

3.3 *Stratified Gradient Sampling (SGS)*

In this section we develop a gradient descent algorithm for the optimisation of stratifiably smooth functions and then detail its convergence properties. We require that the objective function $f : \mathbb{R}^n \to \mathbb{R}$ has the following properties:

- **(Proper)**: *f has compact sublevel sets.*
- **(Stratifiably smooth)**: *f is stratifiably smooth, and for each iterate x and $\epsilon \geqslant 0$ we have an oracle* **Sample**(x,ϵ) *that samples one ϵ-close element x' in each ϵ-close top-dimensional stratum X'.*
- **(Differentiability check)** *We have an oracle* DiffOracle(x) *checking whether an iterate $x \in \mathbb{R}^n$ belongs to the set $\mathcal{D} \subset \mathbb{R}^n$ over which f is differentiable.*

3.3.1 *The algorithm*

That f is a proper map is also needed in the original GS algorithm[4], but is a condition that can be omitted [5] to allow the values $f(x_k)$ to decrease to $-\infty$. In our case we stick to this assumption because we need the gradient of f (whenever defined) to be Lipschitz on sub-level sets.

Similarly, the ability to check that an iterate x_k belongs to $\mathcal{D}$ is standard in the GS methodology. We use it to make sure that f is differentiable at each iterate x_k. For this we call a subroutine **MakeDifferentiable** which slightly perturbs the iterate x_k to achieve differentiability and maintain a descent condition. Note that these considerations are mainly theoretical because generically the iterates x_k are points of differentiability, hence **MakeDifferentiable** is unlikely to do anything.

Finally, the requirement that f is stratifiably smooth replaces the classical weaker assumption used in the GS algorithm that f is locally Lipschitz and that the set $\mathcal{D}$ where f is differentiable is open and

[4] James V Burke, Adrian S Lewis, and Michael L Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM Journal on Optimization*, 15(3):751–779, 2005

[5] Krzysztof C Kiwiel. Convergence of the gradient sampling algorithm for nonsmooth nonconvex optimization. *SIAM Journal on Optimization*, 18(2):379–388, 2007

dense. There are many possible ways to design the oracle **Sample**(x, ϵ): for instance the sampling could depend upon arbitrary probability measures on each stratum, or it could be set by deterministic rules depending on the input (x, ϵ) as will be the case for the persistence map in section 3.4. However our algorithm and its convergence properties are oblivious to these degrees of freedom, as from section 3.2.3 any sampling allows us to approximate the Goldstein subgradient $\partial_\epsilon f(x_k)$ using finitely many neighbouring points to compute $\tilde{\partial}_\epsilon f(x_k)$. In turn we have an approximate descent direction g_k which can be used to produce the subsequent iterate $x_{k+1} := x_k - t_k g_k$ as in the classical smooth gradient descent.

The details of the main algorithm **SGS** are given in alg. 2.

The algorithm **SGS** calls the method **UpdateStep** of 3 as a subroutine to compute the right descent direction g_k and the right step size t_k. Essentially this method progressively reduces the exploration radius ϵ_k of the ball where we compute the descent direction $g_k := \tilde{g}(x_k, \epsilon_k)$ until the criteria of 3.2.7 ensuring that the loss sufficiently decreases along g_k are met.

Given the iterate x_k and the radius ϵ_k, the calculation of $g_k := \tilde{g}(x_k, \epsilon_k)$ is done by the subroutine **ApproxGradient** in 4: points x' in neighboring strata that intersect the ball $B(x_k, \epsilon_k)$ are sampled using **Sample**(x_k, ϵ_k) to compute the approximate Goldstein gradient and in turn the descent direction g_k.

Much like the classical GS algorithm, our method behaves like the well-known smooth gradient descent where the gradient is replaced with a descent direction computed from gradients in neighboring strata. A key difference however is that, in order to find the right exploration radius ϵ_k and step size t_k, the **UpdateStep** needs to maintain a constant C_k to approximate the ratio $\frac{1-\beta}{2L}$ of Proposition 3.2.7, as no Lipschitz constant L may be explicitly available.

To this effect, **UpdateStep** maintains a relative balance between the exploration radius ϵ_k and the norm of the descent direction g_k, controlled by C_k, i.e., $\epsilon_k \simeq C_k \|g_k\|$. As we further maintain $C_k \simeq \frac{1-\beta}{2L}$, we know that the convergence properties of ϵ_k and g_k are closely related. Thus, the utility of this controlling constant is mainly theoretical, to ensure convergence of the iterates x_k towards stationary points in Theorem 3.3.3. In practice, we start with a large initial constant C_0, and decrease it on line 11 of alg. 3 whenever it violates a property of the target constant $\frac{1-\beta}{2L}$ given by Proposition 3.2.7.

Remark 3.3.1. Assume that we know a common Lipschitz constant L for all gradients ∇f_i in the ϵ-neighborhood of the current iterate x_k, recall that f_i is any C^2 extension of the restriction $f_{|X_i}$ to the neighboring top-dimensional stratum $X_i \in \mathcal{X}_{x,\epsilon}$. Then we can simplify alg. 3 by decreas-

Algorithm 2 SGS$(f, x_0, \epsilon, \eta, C_0, \beta, \gamma)$

Require: Loss function f, initial iterate $x_0 \in \mathcal{D}$, exploration radius $\epsilon > 0$, initial constant $C_0 > 0$ controlling exploration radius, critical distance to origin $\eta \geqslant 0$, descent rate $0 < \beta < 1$, step size decay rate $0 < \gamma < 1$

$k \leftarrow 0$
repeat
  $(t_k, g_k, C_{k+1}) \leftarrow$ **UpdateStep**$(f, x_k, \epsilon, \eta, C_k, \beta, \gamma)$ via alg. 3
  $x_{k+1} \leftarrow x_k - t_k g_k$
  $x_{k+1} \leftarrow$ **MakeDifferentiable**(x_{k+1}, x_k, t_k, g_k)
  $k \leftarrow k + 1$
until $\|g_k\| \leqslant \eta$
return x_k

Algorithm 3 UpdateStep$(f, x_k, \epsilon, \eta, C_k, \beta, \gamma)$

$\epsilon_k \leftarrow \epsilon$ and $C_{k+1} \leftarrow C_k$
repeat
  $g_k \leftarrow$ **ApproxGradient**(x_k, ϵ_k) via alg. 4
  if $\|g_k\| \leqslant \eta$ **then**
    Break, **return** $t_k = 0$, g_k and C_{k+1} # *Set $\eta = 0$ to reach an ϵ-stationary point*
  end if
  $t_k \leftarrow \frac{\epsilon_k}{\|g_k\|}$ # *Candidate of update step*
  while $f(x_k - t_k g_k) > f(x_k) - \beta t_k \|g_k\|^2$ and $\epsilon_k \leqslant C_{k+1}\|g_k\|$ **do**
    $C_{k+1} \leftarrow \gamma C_{k+1}$ # *Once $C_{k+1} \leqslant \frac{1-\beta}{2L}$, this loop never occurs by (ii) of Proposition 3.2.7*
  end while
  if $f(x_k - t_k g_k) > f(x_k) - \beta t_k \|g_k\|^2$ or $\epsilon_k > C_{k+1}\|g_k\|$ **then**
    $\epsilon_k \leftarrow \gamma \epsilon_k$ # *Reduce ϵ_k to satisfy criterion (i) of Proposition 3.2.7*
  end if
until $f(x_k - t_k g_k) < f(x_k) - \beta t_k \|g_k\|^2$ and $\epsilon_k < C_{k+1}\|g_k\|$
return t_k, g_k and C_{k+1}

Algorithm 4 ApproxGradient(x_k, ϵ_k)

$G_k \leftarrow \{\nabla f(x_k)\}$ # *Eventually G_k will be some approximate Goldstein subgradient $\tilde{\partial} f_{\epsilon_k}(x_k)$*
$\{x_k^1, \cdots, x_k^m\} \leftarrow$ **Sample**(x_k, ϵ_k) # *ϵ-close samples from ϵ-close top dim strata*
for $1 \leqslant l \leqslant m$ **do**
  $G_k \leftarrow G_k \cup \{\nabla f(x_k^l)\}$ # *Add gradients from remote strata*
end for
Solve the quadratic minimization problem $g_k = \operatorname{argmin}\{\|g\|^2,\ g \in \overline{\operatorname{co}}(G_k)\}$
return g_k # *$g_k = \tilde{g}(x_k, \epsilon_k)$ is the approximate steepest descent direction*

Algorithm 5 MakeDifferentiable(x_{k+1}, x_k, t_k, g_k)

$r \leftarrow t_k \|g_k\|$
while $x_{k+1} \notin \mathcal{D}$ or $f(x_{k+1}) > f(x_k) - \beta t_k \|g_k\|^2$ **do**
  Replace x_{k+1} with a sample in $B(x_k - t_k g_k, r)$
  $r \leftarrow \frac{r}{2}$
end while
return x_{k+1}

ing the exploration radius ϵ_k progressively until $\epsilon_k \leqslant \frac{(1-\beta)}{2L}\|\tilde{g}(x_k, \epsilon_k)\|$ as done in alg. 6: This ensures by Proposition 3.2.7 that the resulting update step satisfies the descent criterion $f(x_k - t_k \tilde{g}(x_k, \epsilon_k)) < f(x_k) - \beta t_k \|\tilde{g}(x_k, \epsilon_k)\|^2$. In particular the parameter C_k is no longer needed, and the theoretical guarantees of the simplified algorithm are unchanged. Note that for objective functions from TDA (see section 3.4), the stability theorems (e.g. from [CSEH07]) often provide global Lipschitz constants, enabling the use of the simplified update step described in alg. 6.

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

Algorithm 6 SimpleUpdateStep$(f, x_k, \epsilon, \eta, \beta, \gamma)$

$\epsilon_k \leftarrow \epsilon$ and $g_k \leftarrow$ **ApproxGradient**(x_k, ϵ_k) via alg. 4
repeat
  if $\|g_k\| \leqslant \eta$ **then**
    Break, **return** $t_k = 0$ and g_k # *Set $\eta = 0$ to reach an ϵ-stationary point*
  end if
  $\epsilon_k \leftarrow \gamma \epsilon_k$ and $g_k \leftarrow$ **ApproxGradient**(x_k, ϵ_k) via alg. 4
until $\epsilon_k \leqslant \frac{1-\beta}{2L}\|g_k\|$
return $t_k := \frac{\epsilon_k}{\|g_k\|}$ and g_k

Remark 3.3.2. In the situation of Remark 3.3.1, let us further assume that $\epsilon \in \mathbb{R}_+ \mapsto \|\tilde{g}(x, \epsilon)\| \in \mathbb{R}_+$ is non-increasing. This mono-

tonicity property mimics the fact that $\epsilon \in \mathbb{R}_+ \mapsto \|g(x,\epsilon)\| \in \mathbb{R}_+$ is non-increasing, since increasing ϵ grows the Goldstein generalized gradient $\partial_\epsilon f(x)$, of which $g(x,\epsilon)$ is the element with minimal norm. If the initial exploration radius ϵ does not satisfy the termination criterion (Line 8), $\epsilon \leqslant \frac{1-\beta}{2L}\|\tilde{g}(x_k,\epsilon)\|$, then one can set $\epsilon_k := \frac{1-\beta}{2L}\|\tilde{g}(x_k,\epsilon)\| \leqslant \epsilon$, yielding $\epsilon_k \leqslant \frac{1-\beta}{2L}\|\tilde{g}(x_k,\epsilon_k)\|$. This way alg. 6 is further simplified: in constant time we find a ϵ_k that yields the descent criterion $f(x_k - t_k\tilde{g}(x_k,\epsilon_k)) < f(x_k) - \beta t_k\|\tilde{g}(x_k,\epsilon_k)\|^2$, and the parameter γ is no longer needed. A careful reading of the proofs provided in the following section yields that the convergence rate (Theorem 3.3.5) of the resulting algorithm is unchanged, however the asymptotic convergence (Theorem 3.3.3), case (**b**), needs to be weakened: converging subsequences converge to ϵ-stationary points instead of stationary points. We omit details for the sake of concision.

3.3.2 *Convergence*

We show convergence of the algorithm towards stationary points in 3.3.3. Finally, 3.3.5 provides a non-asymptotic sub linear convergence rate, which is by no mean tight yet gives a first estimate of the number of iterations required in order to reach an approximate stationary point.

Theorem 3.3.3. *If $\eta = 0$, then almost surely the algorithm either (**a**) converges in finitely many iterations to an ϵ-stationary point, or (**b**) produces a bounded sequence of iterates $(x_k)_k$ whose converging subsequences all converge to stationary points.*

As an intermediate result, we first show that the update step computed in 3 is obtained after finitely many iterations and estimate its magnitude relatively to the norm of the descent direction.

Lemma 3.3.4. **UpdateStep**$(f, x_k, \epsilon, \eta, C_k, \beta, \gamma)$ *terminates in finitely many iterations. In addition, let L be a Lipschitz constant for the restricted gradients ∇f_i (as in Proposition 3.2.6) in the ϵ-neighborhood of x_k. Assume that $\frac{1-\beta}{2L} \leqslant C_k$. If x_k is not an (ϵ,η)-stationary point, then the returned exploration radius ϵ_k satisfies:*

$$\min\left(\frac{\gamma^2(1-\beta)}{2L}\|\tilde{g}(x_k, \frac{1}{\gamma}\epsilon_k)\|, \epsilon\right) \leqslant \epsilon_k \leqslant \min(C_k\|\tilde{g}(x_k,\epsilon_k)\|, \epsilon).$$

Moreover the returned controlling constant C_{k+1} satisfies:

$$C_{k+1} \geqslant \frac{\gamma(1-\beta)}{2L}.$$

Proof. If x_k is a stationary point, then **ApproxGradient** returns a trivial descent direction $g_k = 0$ because the approximate gradient G_k contains $\nabla f(x_k)$ (Line 1). In turn, **UpdateStep** terminates at Line 4.

Henceforth we assume that x_k is not a stationary point and that the breaking condition of Line 4 in alg. 3 is never reached (otherwise the algorithm terminates). Therefore, at each iteration of the main loop, either C_{k+1} is replaced by γC_{k+1} (line 9), or ϵ_k is replaced by $\gamma\epsilon_k$ (line 12), until both the following inequalities hold (line 14):

$$\textbf{(A)}\quad \epsilon_k < C_{k+1}\|\tilde{g}(x_k, \epsilon_k)| \quad \text{and}$$

$$\textbf{(B)}\quad f(x_k - t_k\tilde{g}(x_k, \epsilon_k)) < f(x_k) - \beta t_k\|\tilde{g}(x_k, \epsilon_k)\|^2.$$

Once C_{k+1} becomes lower than $\frac{1-\beta}{2L}$, inequality **(A)** implies inequality **(B)** by Proposition 3.2.7 (**ii**). It then takes finitely many replacements $\epsilon_k \leftarrow \gamma\epsilon_k$ to reach inequality **(A)**, by Proposition 3.2.7 **(i)**. At that point (or sooner), alg. 3 terminates. This concludes the first part of the statement, namely **UpdateStep** terminates in finitely many iterations.

Next we assume that x_k is not an (ϵ, η)-stationary point, which ensures that the main loop of **UpdateStep** cannot break at Line 5. We have the invariant $C_{k+1} \geqslant \gamma\frac{1-\beta}{2L}$: this is true at initialization ($C_{k+1} = C_k$) by assumption, and in later iterations C_{k+1} is only replaced by γC_{k+1} whenever **(A)** holds but not **(B)**, which forces $C_{k+1} \geqslant \frac{1-\beta}{2L}$ by Proposition 3.2.7 **(ii)**.

At the end of the algorithm, $\epsilon_k \leqslant C_{k+1}\|\tilde{g}(x_k, \epsilon_k)\|$ by inequality **(A)**, and so we deduce the right inequality $\epsilon_k \leqslant \min(C_k\|\tilde{g}(x_k, \epsilon_k)\|, \epsilon)$.

Besides, if both **(A)** and **(B)** hold when entering the main loop (line 11) for the first time, then $\epsilon_k = \epsilon$. Otherwise, let us consider the penultimate iteration of the main loop for which the update step is $\frac{1}{\gamma}\epsilon_k$. Then, either condition **(A)** does not hold, namely $\frac{1}{\gamma}\epsilon_k > C_{k+1}\|\tilde{g}(x_k, \frac{1}{\gamma}\epsilon_k)\| \geqslant \gamma\frac{1-\beta}{2L}\|\tilde{g}(x_k, \frac{1}{\gamma}\epsilon_k)\|$, or condition **(B)** does not hold, which by the assertion **(ii)** of Proposition 3.2.7 implies $\frac{1}{\gamma}\epsilon_k \geqslant \frac{1-\beta}{2L}\|\tilde{g}(x_k, \frac{1}{\gamma}\epsilon_k)\|$. In any case, we deduce that

$$\epsilon_k \geqslant \min\left(\frac{\gamma^2(1-\beta)}{2L}\left\|\tilde{g}\left(x_k, \frac{1}{\gamma}\epsilon_k\right)\right\|, \epsilon\right).$$

□

Proof of Theorem 3.3.3. We assume that alternative (**a**) does not happen. By Lemma 3.3.4, alg. 3 terminates in finitely many iteration and by Line 14 we have the guarantee:

$$\forall k \geqslant 0,\ f(x_k - t_k\tilde{g}(x_k, \epsilon_k)) < f(x_k) - \beta t_k\|\tilde{g}(x_k, \epsilon_k)\|^2. \tag{3.14}$$

The subsequent iterate x_{k+1} is initialized at $x_k - t_k\tilde{g}(x_k, \epsilon_k)$ by **MakeDifferentiable** (see alg. 5) and replaced by a sample in a progressively shrinking ball $B(x_k - t_k\tilde{g}(x_k, \epsilon_k), r)$ until two conditions are reached. The first condition is that f is differentiable at x_{k+1}, and since $\mathcal{D}$ has full measure

by Rademacher's Theorem, this requirement is almost surely satisfied in finitely many iterations. The second condition is that

$$\forall k \geqslant 0,\ f(x_{k+1}) < f(x_k) - \beta t_k \|\tilde{g}(x_k, \epsilon_k)\|^2, \tag{3.15}$$

which by (3.14) and continuity of f is satisfied in finitely many iterations as well. Therefore **MakeDifferentiable** terminates in finitely many iterations almost surely. In particular, the sequence of iterates $(x_k)_k$ is infinite.

By (3.15) the sequence of iterates' values $(f(x_k))_k$ is decreasing and it must converge by compactness of the sublevel sets below f. Using (3.15), we obtain:

$$\epsilon_k \|\tilde{g}(x_k, \epsilon_k)\| = t_k \|\tilde{g}(x_k, \epsilon_k)\|^2 \leqslant \frac{1}{\beta}(f(x_k) - f(x_{k+1})) \longrightarrow 0^+, \tag{3.16}$$

Let also L be Lipschitz constant for the restricted gradients ∇f_i (as in Proposition 3.2.6) in the ϵ-offset of the sublevel set $\{x, f(x) \leqslant f(x_0)\}$. Up to taking L large enough, there is another Lipschitz constant $L' < L$ ensuring that

$$\frac{1}{\gamma}\frac{1-\beta}{2L} \leqslant \frac{1-\beta}{2L'} \leqslant C_0.$$

By theorem 3.3.4, upon termination of algorithm 3, $C_1 \geqslant \gamma\frac{1-\beta}{2L'} \geqslant \frac{1-\beta}{2L}$. If $C_1 \leqslant \frac{1-\beta}{2L'}$, the **(ii)** of Proposition 3.2.7 prevents Line 9 in alg. 3 from ever occurring again, i.e., $C_k = C_1$ is constant in later iterations. Otherwise, C_1 satisfies $C_1 \geqslant \frac{1-\beta}{2L'}$ just like C_0. A quick induction then yields:

$$\forall k \geqslant 0,\ C_k \geqslant \frac{1-\beta}{2L}.$$

Therefore, by Lemma 3.3.4:

$$\forall k \geqslant 0,\ \min\left(\frac{\gamma^2(1-\beta)}{2L}\left\|\tilde{g}\left(x_k, \frac{1}{\gamma}\epsilon_k\right)\right\|, \epsilon\right) \leqslant \epsilon_k \leqslant \min(C_0\|\tilde{g}(x_k, \epsilon_k)\|, \epsilon). \tag{3.17}$$

In particular, using the rightmost inequality and (3.16), we get $\epsilon_k \to 0^+$. In turn, using the leftmost inequality, we get that

$$\left\|\tilde{g}\left(x_k, \frac{1}{\gamma}\epsilon_k\right)\right\| \to 0^+. \tag{3.18}$$

The sequence of iterates $(x_k)_k$ is bounded; up to extracting a converging subsequence, we assume that it converges to some x_*. Let $\epsilon' > 0$. We claim that $0 \in \partial_{\epsilon'} f(x_*)$, that is x_* is ϵ'-stationary. As $x_k \to x_*$ and $\epsilon_k \to 0$, we have that for k large enough $B(x_k, \frac{1}{\gamma}\epsilon_k) \subseteq B(x_*, \epsilon')$, which implies that:

$$\partial_{\frac{1}{\gamma}\epsilon_k} f(x_k) \subseteq \partial_{\epsilon'} f(x_*).$$

Besides, from Proposition 3.2.6, we have $\tilde{\partial}_{\frac{1}{\gamma}\epsilon_k} f(x_k) \subseteq \partial_{\frac{1}{\gamma}\epsilon_k} f(x_k)$, so that $\tilde{g}(x_k, \frac{1}{\gamma}\epsilon_k) \in \partial_{\frac{1}{\gamma}\epsilon_k} f(x_k)$. Hence $\tilde{g}(x_k, \frac{1}{\gamma}\epsilon_k) \in \partial_{\epsilon'} f(x_*)$. In the limit, (3.18) implies $0 \in \partial_{\epsilon'} f(x_*)$, as desired.

Following [BLO05], the intersection of the Goldstein subgradients $\partial_{\epsilon'} f(x_*)$, over $\epsilon' > 0$, yields $\partial f(x_*)$. Hence, $0 \in \partial f(x_*)$ and x_* is a stationary point. □

James V Burke, Adrian S Lewis, and Michael L Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM Journal on Optimization*, 15(3):751–779, 2005

Theorem 3.3.5. *If $\eta > 0$, then alg. 2 produces an (ϵ, η)-stationary point using at most $O\left(\frac{1}{\eta \min(\eta, \epsilon)}\right)$ iterations.*

Proof. Assume that alg. 2 has run over k iterations without producing an (ϵ, η)-stationary point. From alg. 3 (line 14), alg. 5 (Line 2) and the choice $t_j = \frac{\epsilon_j}{\|\tilde{g}(x_j, \epsilon_j)\|}$ of update step for $j \leqslant k$, it holds that $\beta \epsilon_j \|\tilde{g}(x_j, \epsilon_j)\| \leqslant f(x_j) - f(x_{j+1})$, and in turn

$$\sum_{j=0}^{k-1} \epsilon_j \|\tilde{g}(x_j, \epsilon_j)\| \leqslant \frac{f_0 - f^*}{\beta},$$

where $f_0 := f(x_0)$ and f^* is a minimal value of f. Besides, using Lemma 3.3.4, ϵ_j is either bigger than $\frac{\gamma^2(1-\beta)}{2L} \left\|\tilde{g}\left(x_j, \frac{1}{\gamma}\epsilon_j\right)\right\|$ or than ϵ, hence

$$\sum_{j=0}^{k-1} \epsilon_j \|\tilde{g}(x_j, \epsilon_j)\| \geqslant k \min_{j<k} \|\tilde{g}(x_j, \epsilon_j)\| \times \min_{j<k} \epsilon_j > k \times \eta \times \min\left(\epsilon, \frac{\gamma^2(1-\beta)}{2L}\eta\right).$$

The two equations cannot simultaneously hold whenever

$$k \geqslant \frac{f_0 - f^*}{\beta} \times \frac{1}{\eta \min\left(\epsilon, \frac{\gamma^2(1-\beta)}{2L}\eta\right)},$$

which allows us to conclude. □

3.3.3 *Approximate distance to strata*

The algorithm and its convergence assume that strata X that are ϵ-close to an iterate x can be detected by the oracle $\mathbf{Sample}(x, \epsilon)$. However in practice computing distances $d(x, X)$ to sub manifolds may be expansive or even impossible. Instead we can hope for approximate distances $\hat{d}(x, X)$. Typically when we have an assignment

$$(x, X) \in \mathbb{R}^n \times \mathcal{X} \longmapsto \tilde{x}_X \in X \subseteq \mathbb{R}^n, \text{ at our disposal, we can define } \hat{d}(x, X) := d(x, \tilde{x}_X),$$

and replace the accurate oracle $\mathbf{Sample}(x, \epsilon)$ with the following approximate oracle:

$$\mathbf{ApproxSample}(x, \epsilon) := \{\tilde{x}_X \mid X \in \mathcal{X}, d(x, \tilde{x}_X) \leqslant \epsilon\} = \{\tilde{x}_X \mid X \in \mathcal{X}, \hat{d}(x, X) \leqslant \epsilon\}.$$

Therefore for the purpose of this section we make the following assumption: *To every iterate $x \in \mathbb{R}^n$ and stratum X we can associate an element $\tilde{x}_X$ that belongs to X, in particular we have the corresponding oracle* **ApproxSample**. *Moreover there exists a constant $a \geqslant 1$ such that the resulting approximate distance to strata $\hat{d}(x, X) := d(x, \tilde{x}_X)$ satisfies:*

$$\forall x \in \mathbb{R}^n, \forall X \in \mathcal{X},\ \hat{d}(x, X) \leqslant a d(x, X).$$

Note that we always have a reverse inequality $\hat{d}(x, X) \geqslant d(x, X)$ since $\tilde{x}_X \in X$. In the case of the persistence map this will specialize to $d(x, X) \leqslant \hat{d}(x, X) \leqslant 2d(x, X)$, that is $a = 2$, see Proposition 3.4.6.

We then replace the approximate Goldstein subgradient $\tilde{\partial}_\epsilon f(x)$ with $\hat{\partial}_\epsilon f(x)$, defined in the exact same way except that it is computed from strata satisfying $\hat{d}(x, X) \leqslant \epsilon$, that is, $\hat{\partial}_\epsilon f(x)$ contains $\nabla f(\tilde{x}_X)$ for each such stratum. The proof of Proposition 3.2.6 adapts straightforwardly to the following statement:

Proposition 3.3.6. *Let $x \in \mathbb{R}^n$ and $\epsilon > 0$. Assume that f is stratifiably smooth. Let L be a Lipschitz constant of the gradients ∇f_i restricted to $\overline{B}(x, a\epsilon) \cap X_i$, where f_i is some local C^2 extension of $f_{|X_i}$, and $X_i \in \mathcal{X}_{x,\epsilon}$ is top dimensional. Then we have:*

$$\hat{\partial}_\epsilon f(x) \subseteq \partial_\epsilon f(x) \text{ and } \partial_\epsilon f(x) \subseteq \hat{\partial}_{a\epsilon} f(x) + \overline{B}(0, (a+1)L\epsilon).$$

Proof. The inclusion $\hat{\partial}_\epsilon f(x) \subseteq \partial_\epsilon f(x)$ is clear. Conversely, let $\nabla_X f(x') \in \partial_\epsilon f(x)$, where X is a top-dimensional stratum incident to x' and $|x' - x| \leqslant \epsilon$. We then have $\hat{d}(x, X) \leqslant a\epsilon$ and hence $\tilde{x}_X$ is a point in $\hat{\partial}_{a\epsilon} f(x)$ which is $(a+1)\epsilon$-close to x'. Therefore $\nabla_X f(x') \in \hat{\partial}_{a\epsilon} f(x) + \overline{B}(0, (a+1)L\epsilon)$. The rest of the proof is then conducted as in 3.2.6. □

The vector $\hat{g}(x, \epsilon)$ with minimal norm in $\hat{\partial}_\epsilon f(x)$ can then serve as a new descent direction in place of $\tilde{g}(x, \epsilon)$:

Proposition 3.3.7. *Let f be stratifiably smooth, and x be a non stationnary point. Let $0 < \beta < 1$, and $\epsilon_0 > 0$. Denote by L a Lipschitz constant for all gradients of the restriction f_i in the ball $\overline{B}(x, a\epsilon_0)$ (as in Proposition 3.3.6).*

(i) *For $0 < \epsilon \leqslant \epsilon_0$ small enough we have $\epsilon \leqslant \frac{1-\beta}{2L}|\hat{g}(x, \epsilon)|$; and*

(ii) *For such ϵ, we have $\forall t \leqslant \frac{\epsilon}{a|\hat{g}(x,\epsilon)|}$, $f(x - t\hat{g}(x, \epsilon)) \leqslant f(x) - \beta t|\hat{g}(x, \epsilon)|^2$.*

Proof. The proof of Proposition 3.2.7 can be replicated by replacing ϵ with $\frac{\epsilon}{a}$ and using Proposition 3.3.6 instead of Proposition 3.2.6. □

In light of this result, we can use $g_k = \hat{g}(x_k, \epsilon_k)$ as a descent direction, which in practice simply amounts to replace the accurate oracle **Sample**(x_k, ϵ_k) in alg. 4 with the approximate oracle **ApproxSample**(x_k, ϵ_k).

The only difference is that the assignment of update step in alg. 3 (Line 7) should take the constant a into account, namely:

$$\text{(Line 7')} \qquad t_k \leftarrow \frac{\epsilon}{a\|g_k\|}.$$

The convergence analysis of section 3.3.2 holds as well for this algorithm, and the proofs of Theorem 3.3.3 and Theorem 3.3.5 are unchanged.

3.4 *Application to Topological Data Analysis*

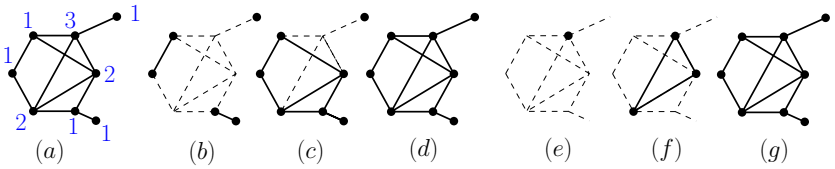


Figure 3.2: Sublevel sets and superlevel sets filtrations illustrated on graphs. (a) Input graph (V,E) along with the values of a function $x : V \to \mathbb{R}$ (blue). (b, c, d) Sublevel sets for t = 1,2,3 respectively. (e, f, g) Superlevel sets for t = 3,2,1 respectively.

In this section, we define the persistence map $\mathrm{PH} : \mathbb{R}^n \to \mathbf{Bar}$ which is a central descriptor in TDA that gives rise to prototypical stratifiably smooth objective functions f in this work. We refer the reader to [EH08, Oud15, ZC05] for full treatments of the theory of Persistence. We then introduce the stratification that makes PH a stratifiably smooth map, by means of the permutation group. Using the associated Cayley graph we give a way to implement the oracle $\mathbf{Sample}(x, \epsilon)$ that samples points in nearby top dimensional strata, which is the key ingredient of alg. 4 for computing descent directions.

Persistent Homology and Barcodes Let $n \in \mathbb{N}$, and let $\{v_1, \dots, v_n\}$ be a (finite) set of vertices. A *simplicial complex* K is a subset of the power set $\mathcal{P}(\{v_1, \dots, v_n\})$ whose elements are called *simplices*, and which is closed under inclusion: if $\sigma \in K$ is a simplex and $\sigma' \subseteq \sigma$, then $\sigma' \in K$. The *dimension* of the complex is the maximal cardinality of its simplices minus one. In particular a 1-dimensional simplicial complex is simply an undirected graph.

A *filter function* is a function on the vertices of K, which we equivalently view as a vector $x \in \mathbb{R}^n$. Given $t \in \mathbb{R}$, we have the sub complexes $K_{\leqslant t} = \{\sigma \in K,\ \forall v \in \sigma,\ x(v) \leqslant t\}$. This yields a nested sequence of sub complexes called the *sublevel sets* filtration of K by x:

$$\varnothing \longrightarrow \cdots \longrightarrow K_{\leqslant s} \xrightarrow{s \leqslant t} K_{\leqslant t} \longrightarrow \cdots \longrightarrow K, \tag{3.19}$$

See Fig. 3.2 for an illustration on graphs. The *(Ordinary) Persistent Homology* of x in degree $p \in \{0, \cdots, \dim K\}$ records topological changes in (3.19) by means of points $(b, d) \in \mathbb{R}^2$, here $b < d$, called *intervals*. For

instance, in degree $p = 0$, an interval (b, d) corresponds to a connected component that appears in $K_{\leqslant b}$ and that is merged with an older component in $K_{\leqslant d}$. In degree $p = 1$ and $p = 2$, intervals track loops and cavities respectively, and more generally an interval (b, d) in degree p tracks a p-dimensional sphere that appears in $K_{\leqslant b}$ and persists up to $K_{\leqslant d}$.

Note that there are possibly infinite intervals (b, ∞) for p-dimensional cycles that persist forever in the filtration (3.19). Such intervals are not easy to handle in applications, and it is common to consider the *(Extended) Persistent Homology* of x, for which they do not occur, i.e. we append the following sequence of pairs involving superlevel sets $K_{\geqslant t} := \{\sigma \in K |\ \forall v \in \sigma,\ x(v) \geqslant t\}$ to (3.19):

$$K \cong (K, \varnothing) \longrightarrow \cdots \longrightarrow (K, K_{\geqslant s}) \xrightarrow{s \geqslant t} (K, K_{\geqslant t}) \longrightarrow \cdots \longrightarrow (K, K). \tag{3.20}$$

Together intervals (b, d) form the (extended) *barcode* $\mathrm{PH}_p(x)$ of x in degree p, which we simply denote by $\mathrm{PH}(x)$ when the degree is clear from the context.

Definition 3.4.1. A *barcode* is a finite multi-set of pairs $(b, d) \in \mathbb{R}^2$ called *intervals*, with $b \leqslant d$. Two barcodes differing by intervals of the form (b, b) are identified. We denote by **Bar** the set of barcodes.

The set **Bar** of barcodes can be made into a metric space as follows. Given two barcodes $D := \{(b, d)\}$ and $D' := \{(b', d')\}$, a *partial matching* $\gamma : D \to D'$ is a bijective map from some subset $A \subseteq D$ to some $B \subseteq D'$. For $q \geqslant 1$ the *q-th diagram distance* $W_q(D, D')$ is the following cost of transferring intervals (b, d) to intervals (b', d'), minimized over partial matchings γ between D and D':

$$W_q(D, D') := \inf_{\gamma} \Bigg(\sum_{(b,d) \in A} \|\gamma(b, d) - (b, d)\|_2^q + \sum_{(b,d) \in D \setminus A} \left(\frac{d - b}{\sqrt{2}} \right)^q + \sum_{(b',d') \in D' \setminus B} \left(\frac{d' - b'}{\sqrt{2}} \right)^q \Bigg)^{\frac{1}{q}}. \tag{3.21}$$

In particular the intervals that are not in the domain A and image B of γ contribute to the total cost relative to their distances to the diagonal $\{b = d\} \subset \mathbb{R}^2$.

The Stability Theorem[6] implies that the map $\mathrm{PH} : \mathbb{R}^n \to \mathbf{Bar}$, which we refer to as the *persistence map* in what follows, is Lipschitz continuous.

Differentiability of Persistent Homology Next we recall from [LOT21] the notions of differentiability for maps in and out of **Bar** and the differentiability properties of PH. Note that the results of [LOT21] focus

[6] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007; and David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko. Lipschitz functions have l p-stable persistence. *Foundations of Computational Mathematics*, 10(2):127–139, 2010

Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021

on ordinary persistence, yet they easily adapt to extended persistence, see e.g. [YL21].

Ka Man Yim and Jacob Leygonie. Optimization of spectral wavelets for persistence-based graph classification. *Frontiers in Applied Mathematics and Statistics*, 7:16, 2021

Given $r \in \mathbb{N}$, we define a *local $\mathcal{C}^r$-coordinate system* as a collection of C^r real-valued maps coming in pairs $b_i, d_i : U \to \mathbb{R}$ defined on some open Euclidean set $U \subseteq \mathbb{R}^n$, indexed by a finite set I, and satisfying $b_i(x) \leqslant d_i(x)$ for all $x \in U$ and $i \in I$. A local $\mathcal{C}^r$-coordinate system is thus equally represented as a map valued in barcodes

$$\tilde{B} : x \in U \mapsto \left\{b_i(x), d_i(x)\right\}_{i \in I} \in \mathbf{Bar},$$

where each interval $(b_i(x), d_i(x))$ is identified and tracked in a $\mathcal{C}^r$ manner.

Definition 3.4.2. A map $B : \mathbb{R}^n \to \mathbf{Bar}$ is *r-differentiable* at $x \in \mathbb{R}^n$ if $B_{|U} = \tilde{B}_{|U}$ for some local $\mathcal{C}^r$-coordinate system $\tilde{B}$ defined in a neighborhood U of x.

Similarly,

Definition 3.4.3. A map $V : \mathbf{Bar} \to \mathbb{R}^m$ is *r-differentiable* at $D \in \mathbf{Bar}$ if $V \circ \tilde{B} : \mathbb{R}^n \to \mathbb{R}^m$ is of class $\mathcal{C}^r$ in a neighborhood of the origin for all $n \in \mathbb{N}$ and local $\mathcal{C}^r$-coordinate system $\tilde{B}$ defined around the origin such that $\tilde{B}(0) = D$.

These notions compose together via the chain rule, so for instance an objective function $f = V \circ B : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable in the usual sense as soon as B and V are so.

We now define the stratification $\mathcal{X}$ of $\mathbb{R}^n$ such that the persistence map $B = \mathrm{PH}$ is r-differentiable (for any r) over each stratum. Denote by Σ the group of permutations on $\{1, \cdots, n\}$. Each permutation $\pi \in \Sigma$ gives rise to a closed polyhedron

$$\mathcal{S}_\pi := \left\{ x \in \mathbb{R}^n \,|\, \forall 1 \leqslant i < n,\, x_{\pi(i)} \leqslant x_{\pi(i+1)} \right\}, \tag{3.22}$$

which is a *cell* in the sense that its (relative) interior is a top-dimensional stratum of our stratification $\mathcal{X}$. The (relative) interiors of the various faces of the cells $\mathcal{S}_\pi$ form the lower dimensional strata. In terms of filter functions, a stratum is simply a maximal subset whose functions induce the same pre-order on vertices of K. We then have that any persistence based loss is stratifiably smooth w.r.t. this stratification.

Proposition 3.4.4. *Let $V : \mathbf{Bar} \to \mathbb{R}$ be a 2-differentiable map. Then the objective function $f := V \circ \mathrm{PH}$ is stratifiably smooth for the stratification $\mathcal{X}$ induced by the permutation group Σ.*

Proof. From Proposition 4.23 and Corollary 4.24 in [LOT21], on each cell $\mathcal{S}_\pi$ we can define a local C^2 coordinate system adapted to PH that consists of linear maps $b_i, d_i : \mathcal{S}_\pi \to \mathbb{R}$, in particular it admits a C^2

extension on a neighborhood of $\mathcal{S}_\pi$. Since V is globally 2-differentiable, by the chain rule, we incidentally obtain a local C^2 extension f_i of $f_{|\mathcal{S}_\pi} = (V \circ \mathrm{PH})_{|\mathcal{S}_\pi}$. □

Remark 3.4.5. Note that the condition that $f = V \circ \mathrm{PH}$ is a proper map, as required for the analysis of alg. 2, sometimes fails because PH may not have compact level-sets. The intuitive reason for this is that a filter function x can have an arbitrarily large value on two distinct entries—one simplex creates a homological cycle that the other destroys immediately—that may not be reflected in the barcode $\mathrm{PH}(x)$. Hence the fiber of PH is not bounded. However, when the simplicial complex K is (homeomorphic to) a compact oriented manifold, any filter function must reach its maximum at the simplex that generates the fundamental class of the manifold (or one of its components), hence PH has compact level-sets in this case. Finally, we note that it is always possible to turn a loss function f based on PH into a proper map by adding a regularization term that controls the norm of the filter function x.

3.4.1 *Exploring the space of strata*

In the setting of Proposition 3.4.4, the objective function $f = V \circ \mathrm{PH} : \mathbb{R}^n \to \mathbb{R}$ is a stratifiably smooth map, where the stratification involved is induced by the group Σ of permutations on $\{1, \cdots, n\}$, with cells $\mathcal{S}_\pi$ as in (3.22). In order to calculate the approximate subgradient $\partial_\epsilon f(x)$, we need to compute the set $\mathcal{X}_{x,\epsilon}$ of cells $\mathcal{S}_\pi$ that are at Euclidean distance no greater than ϵ from x:

$$d_2^2(x, \mathcal{S}_\pi) := \min_{p \in \mathcal{S}_\pi} \|x - p\|_2^2 \leqslant \epsilon^2. \tag{3.23}$$

Estimating distances to strata In practice however, solving the quadratic problem of (3.23) to compute $d_2(x, \mathcal{S}_\pi)$ can be done in $O(n \log n)$ time using solvers for isotonic regression [7]. Since we want to approximate many such distances to neighboring cells, we rather propose the following estimate which boils down to $O(1)$ computations to estimate $d_2(x, \mathcal{S}_\pi)$. For any $\pi \in \Sigma$, we consider the *mirror of x in $\mathcal{S}_\pi$*, denoted by $x^\pi \in \mathbb{R}^n$ and obtained by permuting the coordinates of x according to π:

$$\forall 1 \leqslant i \leqslant n, \qquad x^\pi_{\pi(i)} := x_i. \tag{3.24}$$

In the rest of this section, we assume that the point x is fixed and has increasing coordinates, $x_i \leqslant x_{i+1}$, which can always be achieved after a suitable re-ordering of these coordinates. The proxy $d_2(x, x^\pi)$ then yields a good estimate of $d_2(x, \mathcal{S}_\pi)$, as expressed by the following result.

Proposition 3.4.6. *For any permutation $\pi \in \Sigma$, we have:*

$$d_2(x, \mathcal{S}_\pi) \leqslant d_2(x, x^\pi) \leqslant 2 d_2(x, \mathcal{S}_\pi).$$

[7] Michael J Best and Nilotpal Chakravarti. Active set algorithms for isotonic regression; a unifying framework. *Mathematical Programming*, 47(1):425–439, 1990

Proof. The left inequality is clear from the fact that x^π belongs to the cell $\mathcal{S}_\pi$. To derive the right inequality, let $\hat{x}^\pi$ be the projection of x onto $\mathcal{S}_\pi$. It is a well-known fact in the discrete optimal transport literature, or alternatively a consequence of Lemma 3.4.9 below, that

$$d_2(x^\pi, \hat{x}^\pi) = \min_{\tau \in \Sigma} d_2(x^\tau, \hat{x}^\pi),$$

so that in particular $d_2(x^\pi, \hat{x}^\pi) \leqslant d_2(x, \hat{x}^\pi)$. Consequently,

$$d_2(x, x^\pi) \leqslant d_2(x, \hat{x}^\pi) + d_2(x^\pi, \hat{x}^\pi) \leqslant 2d_2(x, \hat{x}^\pi) = 2d_2(x, \mathcal{S}_\pi).$$

□

Our approximate oracle for estimating the Goldstein subgradient, see section 3.3.3, computes the set of mirrors x^π that are at most ϵ-away from the current iterate $x := x_k$, that is:

$$\textbf{ApproxSample}(x, \epsilon) := \left\{ x^\pi \,|\, d_2(x, x^\pi) \leqslant \epsilon, \pi \in \Sigma \right\}.$$

Remark 3.4.7. Recall that the oracle is called several times in alg. 3 when updating the current iterate x_k with a decreasing exploration radius ϵ_k. However, for the oracle above we have

$$\textbf{ApproxSample}(x, \epsilon') \subseteq \textbf{ApproxSample}(x, \epsilon) \text{ whenever } \epsilon' \leqslant \epsilon,$$

so that once we have computed $\textbf{ApproxSample}(x_k, \epsilon_k)$ for an initial value ϵ_k and the current x_k, we can retrieve $\textbf{ApproxSample}(x_k, \epsilon')$ for any $\epsilon' < \epsilon_k$ in a straightforward way, avoiding re-sampling neighboring points around x_k and computing the corresponding gradient each time ϵ_k decreases, saving an important amount of computational resources.

Sampling in nearby strata In order to implement the oracle $\textbf{ApproxSample}(x, \epsilon)$, we consider the Cayley graph with permutations Σ as vertices and edges between permutations that differ by elementary transpositions (those that swap consecutive elements). In other words, the Cayley graph is dual to the stratification of filter functions: a node corresponds uniquely to a cell $\mathcal{S}_\pi$ and an edge corresponds to a pair of cells sharing a common $(n-1)$-dimensional face.

We explore this graph by maintaining a list $\mathcal{V}$ of visited nodes and a list $\mathcal{N}$ of nodes to visit next. Initially, $\mathcal{V}$ reduces to the source node, the identity permutation, and $\mathcal{N}$ consists of its immediate neighbors, the elementary transpositions. At each step of the exploration procedure, we add a node that has not yet been visited, i.e. a permutation $\pi \in \mathcal{N}$, to the list of $\mathcal{V}$ of visited nodes. If $d_2(x, x^\pi) \leqslant \epsilon$, then we record the mirror point x^π in $\textbf{ApproxSample}(x, \epsilon)$, and we add in $\mathcal{N}$ all the neighbors π' of π that are not yet in $\mathcal{V} \cup \mathcal{N}$. Note that by storing x^π and $d_2(x, x^\pi)$,

we will later be able to retrieve $x^{\pi'}$ and $d_2(x, x^{\pi'})$ in $O(1)$ time for any neighbor π'. The following result entails that this exploration procedure indeed computes the set $\mathbf{ApproxSample}(x, \epsilon)$ of all mirrors x^π that satisfy $d_2(x, x^\pi) \leqslant \epsilon$.

Proposition 3.4.8. *Let $\pi' \in \Sigma$ be a permutation such that $d_2(x, x^{\pi'}) \leqslant \epsilon$. Then there exists a path in the Cayley graph from the identity to π' whose nodes π also satisfy $d_2(x, x^\pi) \leqslant \epsilon$.*

Proof. In fact, from the well-known Lemma 3.4.9 below, any shortest path from the identity to π' does the job, because along such a path the set $\mathrm{inv}(\pi)$ of inversions, i.e. pairs (i, j) satisfying $(j - i)(\pi(j) - \pi(i)) < 0$, grows incrementally. □

Lemma 3.4.9. *Let $x, y \in \mathbb{R}^n$ be two vectors whose coordinates are sorted in the same order, namely $x_i \leqslant x_j \Leftrightarrow y_i \leqslant y_j$. Given $\pi \in \Sigma$ a permutation, let $\mathrm{inv}(\pi)$ be the set of inversions, i.e. pairs (i, j) satisfying $(j - i)(\pi(j) - \pi(i)) < 0$. Then*

$$\mathrm{inv}(\pi) \subseteq \mathrm{inv}(\pi') \Rightarrow \sum (x_i - y_{\pi(i)})^2 \leqslant \sum (x_i - y_{\pi'(i)})^2.$$

Remark 3.4.10. For an arbitrary filter function x, the computation of the barcode $\mathrm{PH}(x)$ has complexity $O(\#K^3)$, here $\#K$ is the number of vertices and edges in the graph K (or the number of simplices if K is a simplicial complex). In the SGS algorithm we need to compute $\mathrm{PH}(x_\pi)$ for each cell $\mathcal{S}_\pi$ near the current iterate x_k, which can quickly become too expansive. Below we describe two heuristics that we implemented in some of our experiments (see section 3.5.3) to reduce time complexity.

The first method bounds the number of strata that can be explored with a hyper-parameter $N \in \mathbb{N}$, enabling a precise control of the memory footprint of the algorithm. In this case exploring the Cayley graph of Σ using Dijkstra's algorithm is desirable, since it allows to retrieve the N strata that are the closest to the current iterate x_k. Note that in the original Dijkstra's algorithm for computing shortest-path distances to a source node, each node is reinserted in the priority queue each time one of its neighbors is visited. However in our case we know the exact distances $d(x, x_\pi)$ to the source each time we encounter a new node, permutation π, so we can simplify Dijkstra's algorithm by treating each node of the graph at most once. The second approach is *memoization*: inside a cell $\mathcal{S}_\pi$, all the filter functions induce the same pre-order on the n vertices of K, hence the knowledge of the barcode $\mathrm{PH}(x_\pi)$ of one of its filter functions allows to compute $\mathrm{PH}(x'_\pi)$ for any other $x'_\pi \in \mathcal{S}_\pi$ in $O(\#K)$ time. We can take advantage of this fact by recording the cells $\mathcal{S}_\pi$ (and the barcode $\mathrm{PH}(x_\pi)$ of one filter function x_π therein) that are met by the SGS (or GS) algorithm during the optimization, thereby avoiding redundant computations whenever the cell $\mathcal{S}_\pi$ is met for a second time.

3.5 *Experiments*

In this section we apply our approach to the optimisation of objective functions based on the persistence map PH, and compare it with other methods. There are two natural classes of objective functions that we can build on top of the barcode $\mathrm{PH}(x)$. One consists in turning $\mathrm{PH}(x)$ into a vector using one of the many existing vectorisation techniques for barcodes [Bub15, AEK+17, CL19, CCI+20] and then to apply any standard objective function defined on Euclidean vector space. In this work we focus on the second type of objective functions which are based on direct comparisons of barcodes by means of metrics W_q on **Bar** as introduced in section 3.4.

We consider three experimental settings in increasing level of complexity. Section 3.5.1 is dedicated to the optimization of an elementary objective function in TDA that allows for explicit comparisons of SGS with other optimization techniques. Section 3.5.2 and 3.5.3 introduce two novel topological optimization tasks: that of *topological registration* for translating filter functions between two distinct simplicial complexes, and that of *topological Fréchet mean* for smoothing the *Mapper graphs* built on top of a data set.

Implementation Our implementation is done in `Python 3` and relies on `TensorFlow`[8] for automatic-differentiation, `Gudhi`[9] for TDA-related computations (barcodes, distances W_q, Mapper graphs), `cvxpy`[10] for solving the quadratic minimization problem involved in alg. 4. Our implementation handles both ordinary and extended persistence, complexes of arbitrary dimension, and can easily be tuned to enable general objective functions (assuming those are provided in an automatic differentiation framework). Our code is publicly available at `https://github.com/tlacombe/topt`.

[8] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pages 265–283, 2016

[9] Clément Maria, Jean-Daniel Boissonnat, Marc Glisse, and Mariette Yvinec. The gudhi library: Simplicial complexes and persistent homology. In *International congress on mathematical software*, pages 167–174. Springer, 2014

[10] Steven Diamond and Stephen Boyd. Cvxpy: A python-embedded modeling language for convex optimization. *The Journal of Machine Learning Research*, 17(1):2909–2913, 2016

3.5.1 *Proof-of-concept: Minimising total extended persistence*

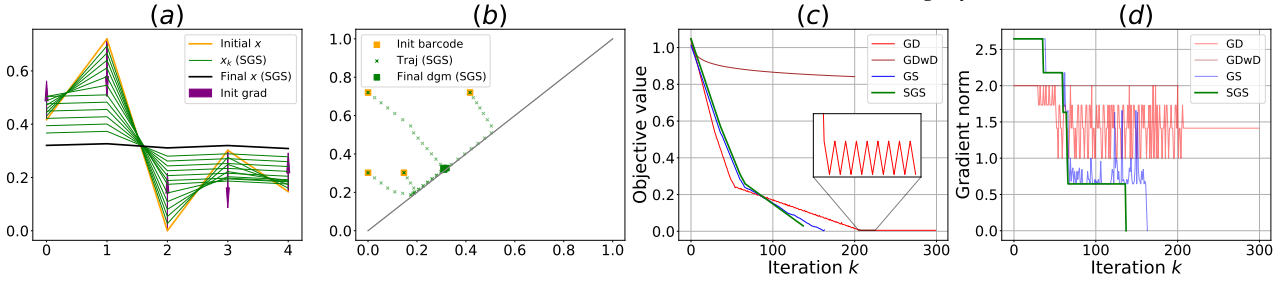


Figure 3.3: Comparison of vanilla Gradient Descent (GD), Gradient Descent with decay (GDwD), Gradient Sampling (GS) and our Stratified Gradient Sampling (SGS) on a toy example. *(a)* Evolution of filter functions $(x_k)_k$ as the total extended persistence Pers is minimised with the SGS method. Purple arrows indicate descent direction at $k = 0$. As expected, the minimization tends to make $(x_k)_k$ topologically as trivial as possible, that is flat in this framework (b) Barcodes $\mathrm{PH}(x_k)$, represented as *persistence diagrams* extended with the point $(\min(x), \max(x))$. *(c)* Objective function's value $\mathrm{Pers}(x_k)$ across iterations k. *(d)* Corresponding gradient norms $(\|g_k\|)_k$. Only GS and SGS reach the stopping criterion $\|g_k\| < \eta$ (slightly faster for SGS than GS).

The goal of this experiment is to provide a simple yet instructive

framework where one can clearly compare different optimization methods. Here we consider the vanilla Gradient Descent (GD), its variant with learning-rate Decay (GDwD), the Gradient Sampling (GS) methodology and our Stratified Gradient Sampling (SGS) approach. Recall that GD is very well-suited to smooth optimization problems, while GS deals with objective functions that are merely locally Lipschitz with a dense subset of differentiability. To some extent, SGS is specifically tailored to functions with an intermediate degree of regularity since their restrictions to strata are assumed to be smooth, and this type of functions arise naturally in TDA.

We consider the elementary example of filter functions x on the graph obtained from subdividing the unit interval with n vertices and the associated (extended) barcodes $\mathrm{PH}(x) = \mathrm{PH}_0(x)$ in degree 0.[11] When the target diagram is empty, $D = \varnothing$, the objective $x \mapsto W_1(\mathrm{PH}(x), \varnothing)$ to minimize is also known in the TDA literature as the *total extended persistence* of $\mathrm{PH}(x)$:

$$\mathrm{Pers} : x \in \mathbb{R}^n \longmapsto \sum_{(b,d)\in \mathrm{PH}(x)} |d-b| \in \mathbb{R}.$$

[11] In this setting the extended barcode can be derived from the ordinary barcode by adding the interval $(\min(x), \max(x))$.

Throughout the minimization, the sublevel sets of x are simplified until they become topologically trivial: $\mathrm{Pers}(x)$ is minimal if and only if x is constant. This elementary optimization problem enables a clear comparison of the GD, GS and SGS methods.

For each mode $\in \{\mathrm{GD}, \mathrm{GDwD}, \mathrm{GS}, \mathrm{SGS}\}$ we get a gradient g_k^{mode} and thus build a sequence of iterates

$$x_{k+1} := x_k - \epsilon_k g_k^{\mathrm{mode}},\ k \geqslant 0.$$

For GD, the update step $\epsilon_k = \epsilon$ is constant, for GDwD it is set to be $\epsilon_k = \epsilon/(1+k)$, and for mode $\in \{\mathrm{GS}, \mathrm{SGS}\}$ it is reduced until $\mathrm{Pers}(x_k - \epsilon_k g_k^{\mathrm{mode}}) < \mathrm{Pers}(x_k) - \beta\epsilon_k \|g_k^{\mathrm{mode}}\|^2$ (and in addition $\epsilon_k < C_k \|g_k^{\mathrm{SGS}}\|$ for SGS). In each case the condition $\|g_k^{\mathrm{mode}}\| \leqslant \eta$ is used as a stopping criterion.

For the experiments, the graph consists of $n = 5$ vertices, $x_0 = (0.4, 0.72, 0, 0.3, 0.14)$, $\epsilon = \eta = 0.01$, and we also have the hyper-parameters $\gamma = 0.5$ and $\beta = 0.5$ for the GS and SGS algorithm. The results are illustrated in Fig. 3.3.

Whenever differentiable, the objective Pers has gradient norm greater than 1, so in particular it is *not* differentiable at its minima, which consists of constant functions. Therefore GD oscillates around its optimal value: the stopping criterion $\|g_k^{\mathrm{GD}}\| \leqslant \eta$ is never met which prevents from detecting convergence. Setting ϵ_k to decay at each iteration in GDwD theoretically ensures the convergence of the sequence $(x_k)_k$, but comes at the expense of a dramatic decrease of the convergence rate.

In contrast, the GS and SGS methods use a fixed step-size ϵ_k yet they converge since they compute a descent direction by minimizing $\|g\|$

over the convex hull of the surrounding gradients $\{\nabla \mathrm{Pers}(x_k), \nabla \mathrm{Pers}(x^{(1)}), \ldots, \nabla \mathrm{Pers}(x^{(m)})\}$, as described in alg. 1 and alg. 4. Here $x^{(1)}, \ldots, x^{(m)}$ are either sampled randomly around the current iterate x_k (with $m = n + 1$) for GS or in the strata around x_k (if any) for SGS. We observe that it takes less iterations for SGS to converge: 137 iterations versus $\sim$ 165 iterations for GS (averaged over 10 runs). This is because in GS the convex hull of the random sampling $\{x^{(1)}, \ldots, x^{(m)}\}$ may be far from the actual generalized gradient $\partial_\epsilon f$, incidentally producing sub-optimal descent directions and missing local minima, while in SGS the sampling takes all nearby strata into account which guarantees a reliable direction (as in Proposition 3.2.7), and in fact since the objective Pers restricts to a linear map on each stratum the approximate gradient $\tilde{\partial}_\epsilon f(x_k)$ equals $\partial_\epsilon f(x_k)$.

Another difference is that GS samples $n + 1 = 6$ nearby points at each iteration k, while SGS samples as many points as there are nearby strata, and for early iterations there is just one such stratum. In practice, this results in a total running time of $\sim$ 2.7s for GS vs. 2.4s for SGS to reach convergence.[12]

[12] Experiment run on a `Intel(R) Core(TM) i5-8350U @ 1.70GHz` CPU

3.5.2 *Topological Registration*

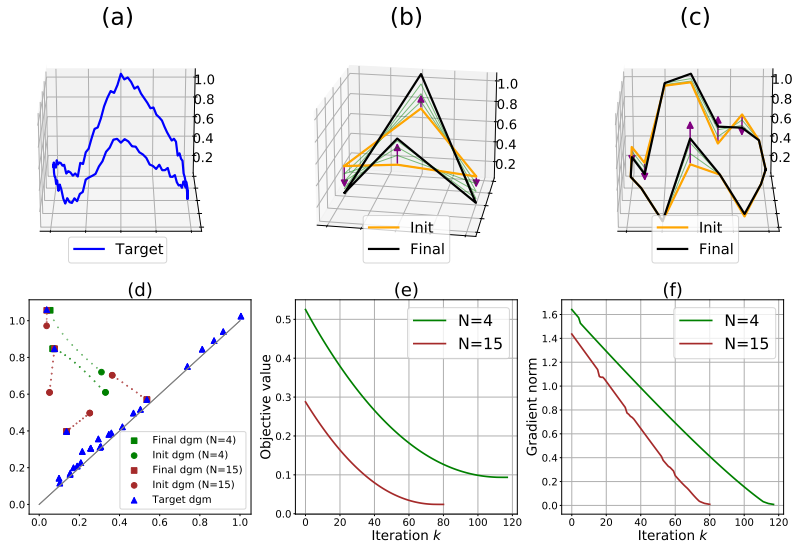


Figure 3.4: Illustration of topological registration *(a)* The target function defined on a (circular) simplicial complex of which we want to reproduce the topology. *(b)* The registration obtained when using a template with $N = 4$ vertices. Purple arrows indicate descent direction at $k = 0$. *(c)* The registration obtained when using a template with $N = 15$ vertices. *(d)* The target persistence diagram (blue triangles) along with the diagram trajectories through iterations for both cases. (green and brown, respectively). *(e)* The objective function's value through iteration. Using a larger template naturally allows to retrieve a better (lower) objective value. *(f)* Corresponding gradient norms, both reaching the stopping criterion $\|g_k\| \leqslant \eta$.

We now present the more sophisticated optimization task of *topological registration*. This problem takes inspiration from registration experiments in shapes and medical images analysis [JM00, DGS01, DFP+11], where we want to translate noisy real-world data (e.g. MRI images of a brain) into a simpler and unified format (e.g. a given template of the brain).

Problem formulation In a topological analog of this problem the observation consists of a filter function F defined on a simplicial complex K which may have, for instance, a large number of vertices, and the template consists of a simplicial complex K' simpler than K (e.g. with fewer vertices). The goal is then to find a filter function x on K' such that (K', x) faithfully recovers the topology of the observation (K, F). Formally we minimise the q-th distance ($q \in [1, +\infty]$) between their barcodes

$$x \mapsto W_q(\mathrm{PH}(x, K'), \mathrm{PH}(F, K)), \tag{3.25}$$

where we include the complexes in the notations $\mathrm{PH}(F, K)$ of the barcodes to make a clear distinction between filter functions defined on K and K'.

Experiment We minimise (3.25) using our SGS approach. The observed simplicial complex K is taken to be the subdivision of the unit circle with 120 vertices, see Fig. 3.4. Let $F = F_0 + \zeta$ where $F_0 \in \mathbb{R}^{120}$ is a piecewise linear map satisfying $F_0[0] = 0, F_0[30] = 1, F_0[45] = 0.05, F_0[60] = 0.35, F_0[75] = 0.1, F_0[90] = 0.8$ and ζ is a uniform random noise in $[0, 0.1]^{120}$. The (extended) barcode of F_0 consists of two long intervals $(0, 1), (0.05, 0.9)$ and one smaller interval $(0.1, 0.35)$ that corresponds to the small variation of F_0 (Fig. 3.4 (a, left of the plot)). The stability of persistent homology implies that the barcode of F, which is a noisy version of F_0, contains a perturbation of these three intervals along with a collection of small intervals of the form (b, d) with $(d - b) < 0.2$, since 0.1 is the amplitude of the noise ζ. The persistence diagram representation of this barcode can be seen on Fig. 3.4 (d, blue triangles): the three intervals are represented by the points away from the diagonal $\{x = y\} \subset \mathbb{R}^2$ and the topological noise is accounted by the points close to the diagonal.

We propose to compute a topological registration x of (K, F) for two simpler circular complexes with $n = 4$ and $n = 15$ vertices respectively (Fig. 3.4, (b,c)). We initialize the vertex values x_0 randomly (uniformly in $[0, 1]^n$), and minimize (3.25) via SGS. We use $q = 2$, and the parameters of alg. 4 are set to $\epsilon = 0.01$, $\eta = 0.01$, $\beta = 0.5$, $\gamma = 0.5$.

With $n = 4$ vertices, the final filter function x returned by alg. 4 reproduces the two main peaks of F that correspond to the long intervals $(0, 1), (0.05, 0.9)$, but it fails to reproduce the small bump corresponding to $(0.1, 0.35)$ as it lacks the degrees of freedom to do so. *A fortiori* the noise appearing in F is completely absent in x, as observed in Fig. 3.4 (d) where the two points appearing in the barcode of x_0 are pushed towards the two points of the target barcode of F as it is the best way to reduce the distance W_q. Using $n = 15$ vertices the barcode $\mathrm{PH}(x)$ retrieves the third interval $(0.1, 0.35)$ as well and thus the final filter function x reaches a lower objective value. However x

also fits some of the noise, as one of the interval in the diagram of x_k is pushed toward a noisy interval close to the diagonal (see Fig. 3.4 (d)).

3.5.3 *Topological Mean of Mapper graphs*

In our last experiment, we provide an application of our SGS algorithm to the *Mapper* data visualization technique[13]. Intuitively, given a data set X, Mapper produces a graph $\mathrm{Map}(X)$, whose attributes, such as its connected components and loops, reflect similar structures in X. For instance, the branches of the Mapper graph in [RCK+17] correspond to the differentiation of stem cells into specialized cells. Besides its potential for applications, Mapper enjoys strong statistical and topological properties [BBMW21, CO17, CMO18, CM19, MW16, BBBN20]. In this last experiment, we propose an optimization problem to overcome one of the main Mapper limitations, i.e., the fact that Mapper sometimes contains irrelevant features, and solve it with the SGS algorithm. We refer the reader to the last paragraph of this section for a more detailed introduction to Mapper and its parameters.

[13] Gurjeet Singh, Facundo Mémoli, and Gunnar Carlsson. Topological methods for the analysis of high dimensional data sets and 3D object recognition. In *4th Eurographics Symposium on Point-Based Graphics (SPBG 2007)*, pages 91–100. The Eurographics Association, 2007

Problem formulation It is a well-known fact that the Mapper graph is not robust to certain changes of parameters[14] which may introduce artificial graph attributes, see [AGH+09b] for an approach to curate $\mathrm{Map}(X)$ from its irrelevant attributes. In our case we assume that $\mathrm{Map}(X)$ is a graph embedded in some Euclidean space $\mathbb{R}^d$ ($d = 2$ in our experiments), which is typically the case when the data set X is itself in $\mathbb{R}^d$, and we modify the embedding of the nodes in order to cancel geometric outliers. For notational clarity we distinguish between the embedded graph $\mathrm{Map}(X) \subseteq \mathbb{R}^d$ and its underlying abstract graph K. Let n be the number of vertices of K.

[14] More details about these parameters are given in the last paragraph of this section

Dominique Attali, Marc Glisse, Samuel Hornus, Francis Lazarus, and Dmitriy Morozov. Persistence-sensitive simplification of functions on surfaces in linear time. In *Topological Methods in Data Analysis and Visualization (TopoInVis 2009)*. Springer, 2009

We propose an elementary scheme inspired from [CMO18] in order to produce a simplified embedding of $\mathrm{Map}(X)$. For this, we consider a family of bootstrapped data sets $\hat{X}_1, \ldots, \hat{X}_k$ obtained by sampling the data set $\mathrm{card}(X)$ times with replacements, from which we derive new mapper graphs $K_1, \cdots, K_k$, whose embeddings $\mathrm{Map}(\hat{X}_1), \ldots, \mathrm{Map}(\hat{X}_k)$ in $\mathbb{R}^d$ are fixed during the experiment. In particular, given a fixed unit vector $\mathbf{e}$ in $\mathbb{R}^d$, the projection $F_{\mathbf{e}}$ onto the line parametrized by $\mathbf{e}$ induces filter functions for each K_i, hence barcodes $\mathrm{PH}(F_{\mathbf{e}}, K_i)$.

Mathieu Carrière, Bertrand Michel, and Steve Oudot. Statistical analysis and parameter selection for Mapper. *Journal of Machine Learning Research*, 19(12):1–39, 2018

We minimize the following objective over filter functions $\tilde{F}_{\mathbf{e}} \in \mathbb{R}^n$:

$$\tilde{F}_{\mathbf{e}} \in \mathbb{R}^n \mapsto \sum_{i=1}^{k} W_2(\mathrm{PH}(\tilde{F}_{\mathbf{e}}, K), \mathrm{PH}(F_{\mathbf{e}}, K_i))^2 \in \mathbb{R}. \tag{3.26}$$

By viewing the optimized filter function $\tilde{F}_{\mathbf{e}}$ as the coordinates of the vertices of K along the $\mathbf{e}$-axis, we obtain a novel embedding of the

mapper graph $\mathrm{Map}(X)$ in $\mathbb{R}^d$ that is the *topological barycenter* of the family $(F_{\mathbf{e}}, \mathrm{Map}(\hat{X}_i))$.

To further improve the embedding $\mathrm{Map}(X)$, we jointly optimize Eq. (3.26) over a family $\{\mathbf{e}_j\}_j$ of directions. Intuitively, irrelevant graph attributes do not appear in most of the subgraphs $\mathrm{Map}(\hat{X}_i)$ and thus are removed in the optimized embedding of $\mathrm{Map}(X)$.

Remark 3.5.1. In some sense, the minimization (3.26) corresponds to pulling back to filter functions the well-known minimization problem $\mathbf{Bar} \ni D \mapsto \sum_{i=1}^{k} W_2(D, D_i)^2$ that defines the *barycenter* or *Fréchet mean*[15] of barcodes $D_1, \ldots, D_k$. Indeed, a *topological mean* of a set of filter functions $x^1, \ldots, x^k$ on simplicial complexes $K_1, \ldots, K_k$ can be defined as a minimizer of $x \in \mathbb{R}^n \mapsto \sum W_2(\mathrm{PH}(x, K), \mathrm{PH}(x^i, K_i))^2$. In our experiment, x is interpreted as a radial projection onto the $\mathbf{e}$-axis, and in fact when considering several directions $\{\mathbf{e}_j\}_j$ the mean resulting from the optimization is actually that of the so-called Persistent Homology Transform[16].

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

[16] Katharine Turner, Sayan Mukherjee, and Doug M. Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014

Experiment To illustrate this new method for Mapper, we consider a data set X of single cells characterized by chromatin folding[17]. Each cell is encoded by the squared distance matrix M of its DNA fragments. This data set was previously studied in [CR20], in which it was shown that the Mapper graph could successfully capture the cell cycle, represented as a big loop in the graph. However, this attribute could only be observed by carefully tuning the parameters. Here we start with a Mapper graph computed out of arbitrary parameters, and then curate the graph using bootstrap iterates as explained in the previous paragraphs.

[17] Takashi Nagano, Yaniv Lubling, Csilla Várnai, Carmel Dudley, Wing Leung, Yael Baran, Netta Mendelson-Cohen, Steven Wingett, Peter Fraser, and Amos Tanay. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*, 547:61–67, 2017

Specifically, we processed the data set X with the stratum-adjusted correlation coefficient (SCC)[18], with 500kb and convolution parameter $h = 1$ on all chromosomes. Then, we run a kernel PCA on the SCC matrix to obtain two lenses. From these lenses we construct a Mapper graph, with a fixed set of parameters for the Mapper algorithm (resolution 15, gain 0.4 on both lenses, and hierarchical clustering with threshold 2 on Euclidean distance). The resulting Mapper graph $\mathrm{Map}(X)$ displayed in fig. 3.5 (upper left) contains the expected main loop associated to the cell cycle, but it also contains many spurious branches. However computing the Mapper graph with same parameters on a bootstrap iterate results in less branches but also in a coarser version of the graph (fig. 3.5, upper middle).

[18] Tao Yang, Feipeng Zhang, Galip Yardımcı, Fan Song, Ross Hardison, William Noble, Feng Yue, and Qunhua Li. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Research*, 27(11):1939–1949, 2017

After using the SGS algorithm capped at 150 strata (see Remark 3.4.10), $\epsilon = 0.01$, $\eta = 0.01$, $\beta = 0.5$, $\gamma = 0.5$, initialized with $\mathrm{Map}(X)$, and with loss computed out of 10 bootstrap iterates and 4 directions with angles $\{0, \pi/2, \pi/4, -\pi/4\}$, the resulting Mapper, shown in fig. 3.5 (upper right), offers a good compromise: its resolution remains high and it is

curated from irrelevant and artifactual attributes.

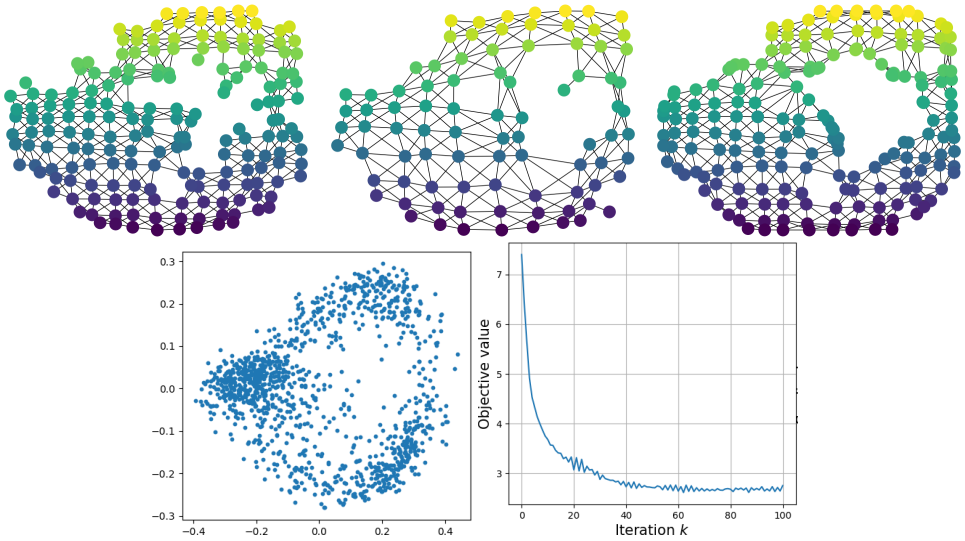


Figure 3.5: Different Mapper graphs colored with the first kernel PCA component. *(Top row)* Left: original Mapper graph computed with a set of arbitrary parameters with many spurious branches. Middle: Mapper graph obtained from bootstrap with very low resolution. Right: curated Mapper graph obtained as the Fréchet mean of the bootstrap iterates. *(Bottom row)* Left: visualization of the data set with kernel PCA. Right: the evolution of the loss (3.26).

A brief review of Mapper and its parameters Mapper is a visualization tool that allows to represent any data set X equipped with a metric and a continuous function $f : X \to \mathbb{R}$ with a graph. It is based on the Nerve Theorem, which essentially states that, under certain conditions, the nerve of a cover of a space has the same topology of the original space, where a cover is a family of subspaces whose union is the space itself, and the (1-skeleton of a) nerve is a graph whose nodes are the cover elements and whose edges are determined by the intersections of cover elements. The whole idea of Mapper is that since covering a space is not always simple, an easier way is to cover the image of a continuous function defined on the space with regular intervals, and then pull back this cover to obtain a cover of the original space.

More formally, Mapper has three parameters: the *resolution* $r \in \mathbb{N}^*$, the *gain* $g \in [0,1]$, and a *clustering method* $\mathcal{C}$. Essentially, the Mapper is defined as $\mathrm{Map}(X) = \mathcal{N}(\mathcal{C}(f^{-1}(\mathcal{I}(r,g))))$, where $\mathcal{I}(r,g)$ stands for a cover of $\mathrm{im}(f)$ with r intervals with g% overlap, and $\mathcal{N}$ stands for the nerve operation, which is applied on the cover $\mathcal{C}(f^{-1}(\mathcal{I}(r,g)))$ of X. This cover is made of the connected components (assessed by $\mathcal{C}$) of the subspaces $f^{-1}(I)$, $I \in \mathcal{I}(r,g)$. See Figure 3.6.

The influence of the parameters $r, g, \mathcal{C}, f$ on the Mapper shape is still an active research area. For instance, the number of Mapper nodes increases with the resolution, and the number of edges increases with the gain, but these parameters, as well as the function f and the clustering method $\mathcal{C}$, can also have more subtle effects on the Mapper shape. We refer the interested reader to the references mentioned in this article for a more detailed introduction to Mapper.

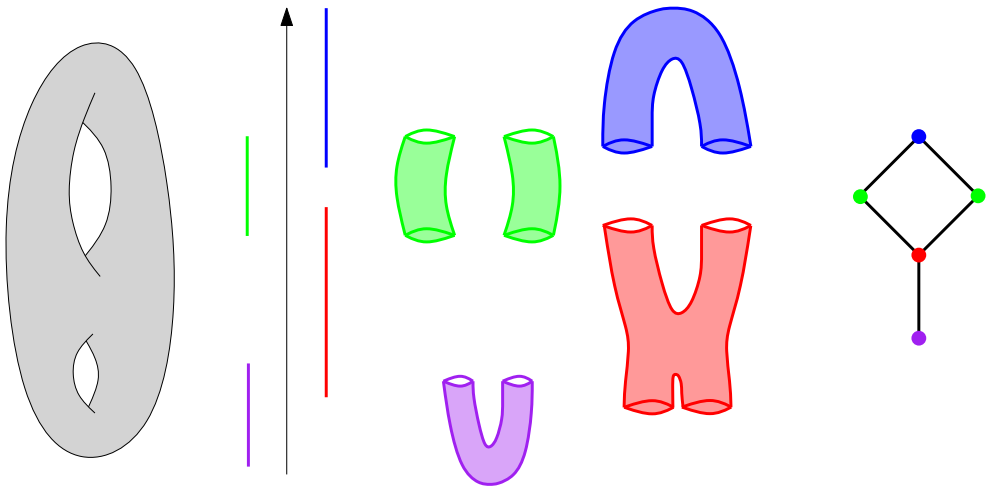


Figure 3.6: Example of Mapper computation on a double torus with height function covered by 4 intervals. Each interval is pulled back in the original space through f^{-1} and then separated into its connected components using $\mathcal{C}$. For example, the cover element obtained with the pre-image of the green interval is separated into its two connected components. The nerve of this cover is then computed to obtain the Mapper.

Part II

Fiber of Persistent Homology

4
The Fiber of Persistent Homology for Simplicial Complexes

Abstract

We study the inverse problem for persistent homology: For a fixed simplicial complex K, we analyse the fiber of the continuous map PH on the space of filters that assigns to a filter $f : K \to \mathbb{R}$ the total barcode of its associated sublevel set filtration of K. We find that PH is best understood as a map of stratified spaces. Over each stratum of the barcode space the map PH restricts to a (trivial) fiber bundle with fiber a polyhedral complex. Amongst other we derive a bound for the dimension of the fiber depending on the number of distinct endpoints in the barcode. Furthermore, taking the inverse image PH^{-1} can be extended to a monodromy functor on the (entrance path) category of barcodes. We demonstrate our theory on the example of the simplicial triangle giving a complete description of all fibers and monodromy maps. This example is rich enough to have a Möbius band as one of its fibers.

4.1 Introduction

Topological Data Analysis (TDA) is a rapidly expanding, new area [CZ09, EH10b, Oud15] which has been applied to a large variety of data science problems. Its best-known tool, persistent homology, provides a non-linear dimension reduction method which is computable [OPT+17, ZC05] and robust with respect to small perturbations of the underlying data [CSEH07]. A growing number of vectorisation methods [AEK+17, Bub15] enable statistical studies of the outcome of persistent homology and combining it with machine learning methods.

It is thus natural to ask how much information can be recovered from persistent homology: Given a particular instance of a persistence module, what can we say about the data set it has been derived from? Any qualitative or quantitative understanding of the information loss would be of great value for future applications, and several approaches to variations of this question have recently appeared [Cur18, CMT18, CMW20, TMB14b]. We refer to [OS20b] for a survey of inverse problems for persistent homology.

In this paper we analyse this foundational problem in a general form: For a fixed simplicial complex K, we study the persistent homology map PH (and its inverse) from the real-valued functions on K to the space of barcodes. We are naturally led to study PH as a map of stratified spaces and to extend PH^{-1} to a functor defined on a natural category of barcodes. This rich structure we expect may also be of interest outside the data science community. Indeed, persistent homology has attracted much recent interest from other branches of mathematics, in particular symplectic topology, stemming from its connection to Morse theory and its close cousin Floer homology [Bar94a, PS16, UZ16], and may yet find uses in other areas, see for example [MM20].

4.1.1 Content and results

Given a fixed finite simplical complex K, let Filt_K be the space of its filters. By definition these are functions $f : K \to \mathrm{I} = [0,1] \subset \mathbb{R}$ that are monotonic with respect to face inclusions,

$$\sigma \subseteq \sigma' \Rightarrow f(\sigma) \leqslant f(\sigma') \qquad \text{for all } \sigma, \sigma' \in K.$$

Thus each sublevel set $f^{-1}((-\infty, t])$ defines a simplicial subcomplex of K and every f gives rise to a filtration. Persistent homology then defines a continuous map

$$\mathrm{PH} : \mathrm{Filt}_K \longrightarrow \mathbf{Bar}^\infty,$$

where $\mathbf{Bar}^\infty$ is the space of total barcodes and PH assigns the union of barcodes in all homological degrees. As K is fixed we restrict our

attention to the image

$$\mathbf{Bar}_K := \mathrm{PH}(\mathrm{Filt}_K).$$

In this notation, to understand the information loss of persistent homology is to understand the fiber $\mathrm{PH}^{-1}(D)$ at a barcode $D \in \mathbf{Bar}_K$. This naturally leads us to a closer analysis of the spaces Filt_K and $\mathbf{Bar}_K$ themselves. We will endow them with monoid actions and stratifications, and show that PH is compatible with these extra structures.

Let $\mathrm{End}(\mathrm{I}, \leqslant)$ be the monoid of order preserving continuous maps of the unit interval I that fix the endpoints, and let $\mathrm{Aut}(\mathrm{I}, \leqslant)$ be its subgroup of homeomorphisms. The monoid $\mathrm{End}(\mathrm{I}, \leqslant)$ and hence $\mathrm{Aut}(\mathrm{I}, \leqslant)$ act continuously on Filt_K by post-composition and on $\mathbf{Bar}_K$ by moving the endpoints of the bars. As the endpoints of the bars in $\mathrm{PH}(f)$ are a subset of the values of f, the map PH is readily seen to be equivariant with respect to these actions (Lemma 4.2.5).

For our further analysis it is important that both Filt_K and $\mathbf{Bar}_K$ have a natural stratification where each i-dimensional stratum is identified with an open simplex

$$\mathring{\Delta}^i = \{(x_1, \cdots, x_i) \mid 0 < x_1 < \cdots < x_i < 1\}$$

such that the coordinates are given by the distinct values in the image of f, and respectively, the distinct endpoints of the bars in D. We identify each such stratum as an $\mathrm{Aut}(\mathrm{I}, \leqslant)$-orbit, and thus PH, by its equivariance, is a strongly stratified map taking a stratum of filters surjectively onto a stratum of barcodes (Proposition 4.2.15). In particular, the inverse image $\mathrm{PH}^{-1}(\mathcal{B})$ of any barcode stratum $\mathcal{B}$ is a finite union of filter strata. We then show that PH over a stratum $\mathcal{B}$ is a fiber bundle with fiber a polyhedral complex (Theorem 4.3.2), and derive some general properties of this fiber. Thus, for example, we show that the dimension of the fiber over a barcode D is bounded by half the difference between the number $\sharp K$ of simplices and the number $\sharp D$ of endpoints in the barcode (Proposition 4.3.5):

$$\dim \mathrm{PH}^{-1}(D) \leqslant \frac{\sharp K - \sharp D}{2}.$$

Unlike Filt_K, $\mathbf{Bar}_K$ is not (the realisation of) a simplical complex: On the boundary of a barcode stratum, viewed as an open simplex, the 1-dimensional subspaces corresponding to bars (x_j, x_j) of length zero are collapsed. Nevertheless, we are able to describe the attaching (or monodromy) maps of the fiber over $\mathcal{B}$ to the fiber of a lower dimensional stratum $\mathcal{B}' \subseteq \bar{\mathcal{B}}$ in its closure. These attaching maps are homotopic to maps of polyhedra but are not in general homotopic to each other (Proposition 4.4.12). We find that this structure is most naturally described in terms of the category $\mathbf{Bar}_\mathbf{K}$: Its objects are the barcodes

in $\mathbf{Bar}_K$ and its space of morphisms from D to D' is the subspace of maps $\phi \in \mathrm{End}(\mathrm{I}, \leqslant)$ that send the endpoints of the bars in D surjectively to those of D'. Each of these morphism spaces is discrete up to homotopy (Theorem 4.4.2). Taking the inverse image then extends to a functor from $\mathbf{Bar_K}$ to the category of topological spaces and continuous maps

$$\mathrm{PH}^{-1} : \mathbf{Bar_K} \longrightarrow \mathbf{Top}$$

taking a morphism $\phi \in \mathbf{Bar_K}(D, D')$ to the continuous map $\mathcal{L}_\phi : f \in \mathrm{PH}^{-1}(D) \longmapsto \phi \circ f \in \mathrm{PH}^{-1}(D')$, which is indeed well-defined by the equivariance of PH under the action of $\mathrm{End}(\mathrm{I}, \leqslant)$.

The category $\mathbf{Bar_K}$, which we were naturally led to consider, is closely related to the entrance path category $\mathbf{Ent}(\mathbf{Bar}_K)$ of the space $\mathbf{Bar}_K$ which we prove to be homotopically stratified (Proposition 4.5.4) in the sense of Quinn.[1] Indeed, we show that descending to the homotopy category, i.e. replacing morphism spaces with the set of their connected components, induces an isomorphism of categories (Proposition 4.5.7)

$$h\mathbf{Bar_K} \simeq \mathbf{Ent}(\mathbf{Bar}_K).$$

[1] Frank Quinn. Homotopically stratified sets. *Journal of the American Mathematical Society*, 1(2):441–499, 1988

Recall that the entrance path category is the analogue for stratified spaces of the fundamental groupoid, and that functors from the entrance category to the category of sets are in correspondence with branched covers. Taking this analogy one step further by replacing the category of sets with the homotopy category of spaces $h\mathbf{Top}$, we may most naturally think of PH as a stratified fiber bundle with polyhedral fibers:

$$\mathrm{PH}^{-1} : \mathbf{Ent}(\mathbf{Bar}_K) \longrightarrow h\mathbf{Top}.$$

Finally we remark in section 4.6.1 that there is no loss of generality by restricting to filters that take value in the bounded interval $\mathrm{I} \subset \mathbb{R}$. We also note in section 4.6.2 that a similar analysis as above holds if instead of all filter functions we consider only lower star filters. Indeed, the space Low_K of lower star filters is a union of strata in Filt_K. Thus the fiber of PH restricted to the lower star filters is again a polyhedral complex.

In Section 4.7, illustrating our theory, we describe in complete detail the case when K is a triangle (with 6 simplices). For each of the 34 barcode strata in $\mathbf{Bar}_K$ we describe the fiber with the action of the symmetry group of K and their monodromy maps. While most fibers consist of a set of discrete points, three fibers are homeomorphic to a circle, one is homeomorphic to two copies of the circle, and one is homeomorphic to the Möbius band. As far as we are aware this is the first non-contractible simplicial complex for which the fibers of PH have been studied and also the first example where the fibers are not homotopy discrete.

4.1.2 *Related previous work*

Our set-up here is most closely related to that in the work of Cyranka, Mischaikow and Weibel [CMW20] where the authors consider lower star filters on the n-fold subdivided interval and show that the fibers of PH are homotopy discrete. Previously, Curry in [Cur18] considers the interval with the set of continuous maps. In particular he bounds the connected components of the fiber in terms of the nestings of the intervals in the barcode. Curry with coauthors also studies the higher dimensional example of a sphere in [CCF+20] with the set of functions that arise as compositions of an embedding of $\mathbb{S}^2$ into $\mathbb{R}^3$ followed by a projection onto the last coordinate. To the best of our knowledge, these are all the examples where the fibers of persistent homology have been analysed.

In general, the persistent homology associated to a single filter cannot determine the underlying simplicial complex or its homotopy type, no more than homology can determine the homotopy type of the underlying space. However, under some conditions a family of such functions might suffice. To understand this question Turner, Mukherjee and Boyer introduced the persistent homology transform (PHT) [TMB14b], and proved that indeed under certain circumstances PHT is injective on shapes embedded in $\mathbb{R}^3$, see also [CP16, GLM18] for a generalisation to higher dimensions. It has even been possible to find algorithmically a left inverse for PHT for some specific classes of sets [BFM+20, Bet18, FMM+19, Mic20].

There are other persistence based invariants of spaces. One such (stable and computable) invariant for metric graphs has been proposed by Dey, Shi, and Wang [DSW15]. In [OS17] Oudot and Solomon show that the fiber of this intrinsic transform is generically globally and always locally injective.

4.2 *Stratifications of the spaces of filters and barcodes*

We first recall some background theory and define the persistence map PH the fibers of which are the object of interest. In section 4.2.2, we introduce the topological monoid $\mathrm{End}(I, \leqslant)$ of non-decreasing maps of the interval which provides essential structure: it acts continuously on the spaces of filters and barcodes, and the persistence map PH is equivariant with respect to this action. In section 4.2.3, we show furthermore that the orbits of the subgroup of homeomorphisms, $\mathrm{Aut}(I, \leqslant)$, provide stratifications for the space of filters and the space of barcodes and that, due to its equivariance, PH is a strongly stratified map between them. It follows now easily that the fibers over barcodes from the same stratum are pairwise homeomorphic, thus turning the identi-

fication of the fiber into a finite problem. In section 1.4 we also show that the image $\mathbf{Bar}_K$ of the space of filters under PH has the quotient topology.

4.2.1 *The definition of the persistence map*

Let **SCpx** denote the category of finite (abstract) simplicial complexes and inclusions, and let $K \in \mathbf{SCpx}$ be an arbitrary simplicial complex of dimension $d \in \mathbb{N}$, which is fixed throughout the paper. We consider K as a subset of the power set on its vertices. Recall, if $\sigma \in K$ then all its non-empty subsets $\sigma' \subset \sigma$, i.e. its *faces*, are also in K. We write $\sharp K$ for the total number of simplices in the complex K.

We denote by $\mathrm{I} := [0,1]$ the closed unit interval. A typical function on K valued in I is denoted by $f \in \mathrm{I}^K$.

Definition 4.2.1. A filter function, or *filter* for short, on K is a map $f : K \to \mathrm{I}$ that is monotonic with respect to face inclusions: For all simplices $\sigma', \sigma \in K$

$$\sigma' \subseteq \sigma \Rightarrow f(\sigma') \leqslant f(\sigma).$$

The set of all filters on K is denoted by Filt_K.

The monotonicity condition on filters is equivalent to the property that their sublevel sets are simplicial subcomplexes of K. Thus a filter f gives rise to a filtration $K(f) = \{f^{-1}((-\infty, t])\}_{t\in\mathbb{R}}$, of K which we may think of as a functor from $\mathbb{R}$ (as an ordered set) to the category of simplicial complexes

$$K(f) : (\mathbb{R}, \leqslant) \longrightarrow \mathbf{SCpx}.$$

We can then compose this with the functor H_p which takes a simplicial complex to its p-th simplicial homology with coefficients in a fixed field $\Bbbk$. This defines the *pth persistent homology functor*

$$\mathrm{H}_p(K(f)) : (\mathbb{R}, \leqslant) \longrightarrow \mathbf{Vect}_\Bbbk,$$

which is an instance of a one-parameter, pointwise finite dimensional, finite persistence module, or *persistence module* for short. We denote by **Pers** the category of such persistence modules and natural transformations between them.

Given an interval $J \subseteq \mathbb{R}$, the associated *interval module* $\mathbb{I}_J \in \mathbf{Pers}$ has copies of the field $\Bbbk$ over J and zero elsewhere, the copies of $\Bbbk$ being connected by identity maps. Given a persistence module $\mathbb{V} \in \mathbf{Pers}$, by the Decomposition Theorem[2], there exists a unique finite set $\mathcal{J}$ of intervals such that we have an isomorphism:

$$\mathbb{V} \cong \bigoplus_{J\in\mathcal{J}} \mathbb{I}_J.$$

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

The finite multiset $\mathbf{Bar}(\mathbb{V})$ of pairs $(\inf J, \sup J) \in (\mathbb{R} \sqcup \{-\infty\}) \times (\mathbb{R} \sqcup \{\infty\})$ for intervals $J \in \mathcal{J}$ appearing in the above decomposition is the so-called *barcode* of the module $\mathbb{V}$. If $\mathbb{V} = \mathrm{H}_p(K(f))$ is a persistent homology module, then the intervals that occur are all half-open intervals of the form $J = [b, d)$ with restricted values $(b, d) \in \mathrm{I} \times (\mathrm{I} \sqcup \{\infty\})$. Consequently, we formally define barcodes as follows.

Definition 4.2.2. A barcode D is a finite multi-set of pairs (b, d) in $\mathrm{I} \times (\mathrm{I} \sqcup \{\infty\})$, with $b < d$, called the *intervals* or *bars* of D. An interval (b, d) is *bounded* (resp. *infinite*) if $d < \infty$ (resp. $d = \infty$). The multiplicity of an interval $(b, d) \in D$ is denoted by $D(b, d) \in \mathbb{N}$. The set of all barcodes is denoted by $\mathbf{Bar}$.

We can now define the *degree p persistence map* as the composition

$$\mathrm{PH}_p := \mathbf{Bar}(\mathrm{H}_p \circ K(.)) : \mathrm{Filt}_K \longrightarrow \mathbf{Bar}$$

and the (total) *persistence map* as the product

$$\mathrm{PH} := (\mathrm{PH}_0, \ldots, \mathrm{PH}_d) : \mathrm{Filt}_K \longrightarrow \mathbf{Bar}^{d+1}.$$

We will refer to elements $D = (D_0, \ldots, D_d) \in \mathbf{Bar}^{d+1}$ simply as barcodes. We will mainly be interested in barcodes in the image of PH:

$$\mathbf{Bar}_K := \mathrm{PH}(\mathrm{Filt}_K) \subseteq \mathbf{Bar}^{d+1}.$$

The set of filters Filt_K is naturally topologised as a subset of the finite dimensional Euclidean space $\mathbb{R}^K$.

The standard topology on $\mathbf{Bar}$, and hence on $\mathbf{Bar}_K$, is induced by an (extended) metric which we now recall. A *matching* γ between two barcodes $D, D' \in \mathbf{Bar}$ is a partial injective map γ from intervals of D to those of D'. The *cost* $c(\gamma)$ of a matching is the maximum of the following three quantities: (i) the maximum $\|(b, d) - \gamma(b, d)\|_\infty$ over intervals $(b, d) \in D$ where γ is defined, (ii) the maximal length $\frac{d-b}{2}$ over intervals $(b, d) \in D$ where γ is not defined, and (iii) the maximal length $\frac{d'-b'}{2}$ over intervals $(b', d') \in D'$ that are not in the image of γ. Here we allow ∞ as a possible value for d, d', and hence also for the maxima; and $\|.\|_\infty$ denotes the (extended) supremum norm on $\mathbb{R} \times (\mathbb{R} \cup \{\infty\})$. The *bottleneck distance*, d_b, between D and D' is then defined as:

$$d_b(D, D') := \inf_{\gamma \text{ matching}} c(\gamma).$$

Since our barcodes are finite multisets, d_b defines a true (extended) metric on $\mathbf{Bar}$. We endow $\mathbf{Bar}$ with the induced *bottleneck topology*.

We thus have the following instance of the Stability Theorem[3] in our context.

Theorem 4.2.3. *The degree p persistence map* PH_p *is Lipschitz continuous, and thus so is the persistence map* PH*.*

[3] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *Journal of Computational Geometry*, 6(1):162–191, 2015; Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The structure and stability of persistence modules*. SpringerBriefs in Mathematics. Springer, 2016; and David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007

For later reference, we record the following elementary fact.

Proposition 4.2.4. *Let $D \in \mathbf{Bar}^{d+1}$ be a non-empty barcode. Then for small enough ϵ, another barcode D' is ϵ-close to D if and only if its intervals satisfy the following:*

- *for each integer $0 \leqslant p \leqslant d$ and interval $(b,d) \in D_p$, the intervals (b',d') in D'_p satisfying $\|(b',d') - (b,d)\|_\infty < \epsilon$ have multiplicities summing up to $D_p(b,d)$, the multiplicity of (b,d) in D_p;*
- *the other intervals $(b',d') \in D'$, that is those that are not ϵ-close to intervals in D, are ϵ-*small*, i.e. $|d'-b'| < \epsilon$.*

Proof. Take $\epsilon \leqslant \frac{\alpha}{2}$ where α is the minimum of (a) the lengths $d-b$ of intervals $(b,d) \in D$ and (b) all pairwise distances $\|(b,d) - (\overline{b},\overline{d})\|_\infty$ for any two geometrically distinct intervals (b,d) and $(\overline{b},\overline{d})$ in D. Note that $\alpha > 0$ as by definition all our barcodes are finite, i.e. D has finite support. □

4.2.2 *Actions on filters and barcodes, and equivariance of the persistence map*

Let $\mathrm{Aut}(\mathrm{I},\leqslant)$ be the space of orientation preserving homeomorphisms of I, and $\mathrm{End}(\mathrm{I},\leqslant)$ be the space of continuous non-decreasing maps that fix the boundary points 0 and 1. We consider them as subspaces of the space of all continuous maps of I to itself with the compact open (or equivalently $||.||_\infty$-metric) topology. In this topology $\mathrm{End}(\mathrm{I},\leqslant)$ is the closure of $\mathrm{Aut}(\mathrm{I},\leqslant)$. For future reference we note that the straight line interpolation

$$\phi_t := (1-t)\phi + t\phi'$$

between maps $\phi, \phi' \in \mathrm{End}(\mathrm{I},\leqslant)$ defines a continuous path in $\mathrm{End}(\mathrm{I},\leqslant)$.

Since the boundary points are fixed by elements in $\mathrm{Aut}(\mathrm{I},\leqslant)$ and $\mathrm{End}(\mathrm{I},\leqslant)$, they extend by the identity to automorphisms and endomorphisms of the real line $\mathbb{R}$ and the extended real line $\mathbb{R} \cup \{\pm\infty\}$. When the context requires it, we will tacitly extend our maps without changing notation.

The monoid $\mathrm{End}(\mathrm{I},\leqslant)$ acts from the left on Filt_K by post-composition:

$$\phi.f := \phi \circ f.$$

It also acts from the left on **Bar**, and hence diagonally on $\mathbf{Bar}^{d+1}$, by applying ϕ to all the endpoints of the bars in D with the convention that $\phi(\infty) = \infty$ and bars of length zero are suppressed:

$$\phi.D := \phi(D) = \{(\phi(b),\phi(d)) \mid (b,d) \in D \text{ and } \phi(b) \neq \phi(d)\}.$$

Thus $\phi(D)$ contains (with multiplicities) an interval $(\phi(b),\phi(d))$ for each interval $(b,d) \in D$ as long as $\phi(b) \neq \phi(d)$, and an interval $(\phi(b),\infty)$ for each interval $(b,\infty) \in D$.

A key result used in this work is that the persistence map is equivariant with respect to these actions.

Lemma 4.2.5 (Equivariance)**.** *The persistence map* PH *is* $\mathrm{End}(I,\leqslant)$*-equivariant: For all* $\phi \in \mathrm{End}(I,\leqslant)$ *and* $f \in \mathrm{Filt}_K$

$$\mathrm{PH}(\phi \circ f) = \phi(\mathrm{PH}(f)).$$

Proof. We fix a filter $f \in \mathrm{Filt}_K$ and a map $\phi \in \mathrm{End}(\mathrm{I},\leqslant)$.

Recall that $\mathrm{PH}(f)$ is the union of $\mathrm{PH}_p(f)$ for $p = 0,\dots,d$, and $\mathrm{PH}_p(f)$ is given by the composition $\mathbf{Bar}(\mathrm{H}_p \circ K(f))$. By definition of $K(\phi \circ f)$, for $t \in \mathbb{R}$ we have

$$K(\phi\circ f)(t) = (\phi\circ f)^{-1}\big((-\infty;t]\big) = f^{-1}\big((-\infty;\max(\phi^{-1}(\{t\}))]\big) = K(f)\big(\max(\phi^{-1}(\{t\}))\big).$$

Note that ϕ is non-decreasing and continuous. Thus the inverse image $\phi^{-1}(\{t\})$ of the point t is a closed, bounded interval and hence contains its maximum. On composition with the singular homology functor H_p this yields

$$\mathrm{H}_p \circ K(\phi\circ f)(t) = \mathrm{H}_p \circ K(f)\big(\max(\phi^{-1}(\{t\}))\big).$$

Hence the barcode $\mathrm{PH}_p(\phi \circ f)$ is the barcode of the persistence module $t \mapsto \mathrm{H}_p \circ K(f)\big(\max(\phi^{-1}(\{t\}))\big)$, which rewrites uniquely as a sum of interval modules:

$$\begin{aligned}
\mathrm{H}_p \circ K(f)\big(\max(\phi^{-1}(\{t\}))\big) &\simeq \Big[\bigoplus_{(b,d)\in \mathrm{PH}_p(f)} \mathbb{I}_{[b,d)}\Big]\big(\max(\phi^{-1}(\{t\}))\big)\\
&= \bigoplus_{(b,d)\in \mathrm{PH}_p(f)} \mathbb{I}_{[b,d)}\big(\max(\phi^{-1}(\{t\}))\big)\\
&= \bigoplus_{(b,d)\in \mathrm{PH}_p(f)} \mathbb{I}_{[\phi(b),\phi(d))}(t).
\end{aligned}$$

The first equality follows from the definition of the barcode $\mathrm{PH}_p(f)$, i.e. $\mathrm{H}_p \circ K(f)$ decomposes as $\bigoplus_{(b,d)\in \mathrm{PH}_p(f)} \mathbb{I}_{[b,d)}$. The second equality holds because pre-composition by the map $t \mapsto \max(\phi^{-1}(\{t\}))$ induces an additive endofunctor on persistence modules. The third one is a consequence of ϕ being non-decreasing, as then $\max \phi^{-1}(\{t\}) \in [b,d)$ is equivalent to $t \in [\phi(b),\phi(d))$.

This yields $\mathrm{PH}_p(\phi\circ f) = \phi(\mathrm{PH}_p(f))$, and hence $\mathrm{PH}(\phi\circ f) = \phi(\mathrm{PH}(f))$. □

Remark 4.2.6. In the above proof we indirectly made use of the following more general categorical framework where both $K(f)$ and $\mathrm{H}_p \circ K(f)$ are considered as functors from the category $(\mathbb{R},\leqslant)$ of ordered real numbers defining concrete instances of filtrations and persistence modules, that is functors

$$F : (\mathbb{R},\leqslant) \longrightarrow \mathbf{SCpx} \qquad \text{and} \qquad \mathbb{V} : (\mathbb{R},\leqslant) \longrightarrow \mathbf{Vect}_{\Bbbk}.$$

Precomposition with any endofunctor α of $(\mathbb{R}, \leqslant)$ defines a right action both on filtrations and on persistence modules. Furthermore, composition by any functor $L : \mathbf{SCpx} \to \mathbf{Vect}_{\Bbbk}$ defines a map from filtrations to persistence modules, and we have the following general equivariance result due to associativity for composition of functors:

$$L(F.\alpha) = L \circ (F \circ \alpha) = (L \circ F) \circ \alpha = L(F).\alpha.$$

The endofunctors of $(\mathbb{R}, \leqslant)$ are the (weakly) order preserving maps of $\mathbb{R}$, that is maps α satisfying: $t' < t \Rightarrow \alpha(t') \leqslant \alpha(t)$. Note that α does not have to be continuous.

For the proof of Lemma 4.2.5 we have used the functor $L = \mathrm{H}_p$ and the fact that an element $\phi \in \mathrm{End}(\mathrm{I}, \leqslant)$ gives rise to an endofunctor $\alpha := \max \phi^{-1}$ defined by $t \mapsto \max \phi^{-1}(\{t\})$. We furthermore used that $K(\phi \circ f) = K(f) \circ \max \phi^{-1}$ and $\mathbb{I}_{[\phi(b),\phi(d))} = \mathbb{I}_{[b,d)} \circ \max \phi^{-1}$ to translate the given action on the space of filters and barcodes into the functorial setting.

It is an easy exercise to show that the action of $\mathrm{End}(\mathrm{I}, \leqslant)$ on filters is continuous. We next show that the action is also continuous on barcodes.

Proposition 4.2.7. $\mathrm{End}(I, \leqslant)$ *acts continuously on* **Bar**, *i.e. it is induced by a continous map*

$$\mathrm{End}(I, \leqslant) \times \mathbf{Bar} \to \mathbf{Bar}.$$

Proof. We show that the map $\mathrm{End}(\mathrm{I}, \leqslant) \times \mathbf{Bar} \to \mathbf{Bar}$ is sequentially continuous. Let $(\phi_n, D_n) \in \mathrm{End}(\mathrm{I}, \leqslant) \times \mathbf{Bar}$ be a sequence converging to some (ϕ, D). Let $\epsilon > 0$ be small enough such that the intervals of any barcode that is ϵ-close to D satisfy the alternative of Proposition 4.2.4. As I is compact, ϕ is uniformly continuous and there exists $\eta > 0$ such that $|\phi(d) - \phi(b)| < \epsilon$ whenever $b, d \in \mathrm{I}$ satisfy $|d - b| < \eta$. Let n be large enough such that D_n is $\min(\epsilon, \eta)$-close to D, and moreover $\|\phi_n - \phi\|_\infty < \epsilon$.

If $(b_n, d_n) \in D_n$ is a small interval, i.e. $d_n - b_n < \min(\epsilon, \eta)$, then:

$$|\phi_n(d_n) - \phi_n(b_n)| < |\phi_n(d_n) - \phi(d_n)| + |\phi(d_n) - \phi(b_n)| + |\phi(b_n) - \phi_n(b_n)| < 3\epsilon.$$

Therefore $(\phi_n(b_n), \phi_n(d_n))$ is a 3ϵ-small interval in $\phi_n(D_n)$.

Else, (b_n, d_n) is $\min(\epsilon, \eta)$-close to a unique interval (b, d) of D, and then

$$\begin{aligned}\|(\phi_n(b_n), \phi_n(d_n)) - (\phi(b), \phi(d))\|_\infty &\leqslant \|(\phi_n(b_n), \phi_n(d_n)) - (\phi(b_n), \phi(d_n))\|_\infty \\ &\quad + \|(\phi(b_n), \phi(d_n)) - (\phi(b), \phi(d))\|_\infty < 2\epsilon.\end{aligned}$$

This yields a canonical matching from $\phi_n(D_n)$ to $\phi(D)$ with cost less than 3ϵ. □

4.2.3 *Stratifications of the spaces of filters and barcodes*

We introduce a weak notion of stratification in the sense that we will not require (for now) any conditions on how strata are glued together. We will return to this in Section 4.5.1.

Definition 4.2.8. A *stratification* of a topological space X is a filtration $\varnothing = X_{-1} \subseteq X_0 \subseteq X_1 \subseteq \cdots$ by a (possibly infinite) sequence of closed subspaces X_i, $i \in \mathbb{N}$, where the sets $X_i \backslash X_{i-1}$ are topological manifolds of dimension i. The path connected components of $X_i \backslash X_{i-1}$ are called *i-strata*, or strata of dimension i. A *stratified map* between two stratified spaces X and X' is a continuous, filtration preserving map of the underlying spaces. A *strongly stratified map* is a stratified map that maps any stratum of X surjectively to a stratum of X'.

We will show that Filt_K and $\mathbf{Bar}_K$ are stratified spaces with strata given by the $\mathrm{Aut}(\mathrm{I}, \leqslant)$-orbits, each homeomorphic to an open standard simplex for some i,

$$\mathring{\Delta}^i := \left\{(x_1, \cdots, x_i) \mid 0 < x_1 < \cdots < x_i < 1\right\} \subset \mathbb{R}^i.$$

Recall that $\mathrm{Aut}(\mathrm{I}, \leqslant)$ acts continuously on the space of filters Filt_K by post-composition. For $f \in \mathrm{Filt}_K$, we denote the associated orbit by

$$\mathcal{S} = \mathcal{S}_f := \{\phi.f \mid \phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)\}.$$

Two filters are in the same orbit if they induce the same pre-order on the simplices of K: $\sigma \leqslant \sigma' \iff f(\sigma) \leqslant f(\sigma')$. Inside a given orbit a filter is uniquely determined by the sequence of its values that are not equal to 0 or 1 sorted in increasing order. Varying f by an element in $\mathrm{Aut}(\mathrm{I}, \leqslant)$ varies this sequence over the whole open standard simplex. Thus for each orbit $\mathcal{S}$ the map

$$\mathcal{S} \xrightarrow{\cong} \mathring{\Delta}^{\dim \mathcal{S}}, \qquad \dim \mathcal{S} := \sharp(\mathrm{Im}(f) \cap (0,1))$$

that sends a filter to the increasing sequence of its distinct values that are not equal to 0 or 1 defines an affine homeomorphism. The inverse map μ is a coordinate chart for the stratum $\mathcal{S}$ which in fact extends to the closure,

$$\mu : \Delta^{\dim \mathcal{S}} \xrightarrow{\cong} \overline{\mathcal{S}}, \tag{4.1}$$

and $\overline{\mathcal{S}}$ is the orbit of $\mathrm{End}(\mathrm{I}, \leqslant)$. We record that the orbits define a stratification.

Proposition 4.2.9. *For each $i \geqslant 0$, let F_i be the union of $\mathrm{Aut}(I, \leqslant)$-orbits $\mathcal{S}$ with $\dim \mathcal{S} \leqslant i$. This defines a stratification of Filt_K. The i-strata are given by the orbits of dimension i.*

Remark 4.2.10. A stratum is simply an equivalence class of filters, where filters are declared equivalent if they induce the same pre-order on simplices. This point of view was already adopted in [LOT21] in the context of persistence differentiation. Equivalently, the stratification is the hyperplane arrangement generated by the equalities $f(\sigma) = f(\sigma')$. It is well-known to be a Whitney stratification, but we will not make use of this richer structure here.

Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021

Similarly we construct a stratification of barcodes $\mathbf{Bar}^{d+1}$. For $D \in \mathbf{Bar}^{d+1}$, we consider the associated $\mathrm{Aut}(\mathrm{I}, \leqslant)$-orbit

$$\mathcal{B} = \mathcal{B}_D := \{\phi.D \,|\, \phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)\}.$$

The orbits partition the space of barcodes $\mathbf{Bar}^{d+1}$. Within such an orbit, the multiplicities and the nestings of bars are constant, and it is only the consecutive values of the interval endpoints that can vary. Thus for each orbit $\mathcal{B}$ the map

$$\mathcal{B} \xrightarrow{\cong} \mathring{\Delta}^{\dim \mathcal{B}}, \qquad \dim \mathcal{B} = \dim \mathcal{B}_D = \dim D := \sharp \text{ distinct endpoints of } D \text{ that are in } (0,1)$$

that sends a barcode to the increasing sequence of its distinct values of interval endpoints that are not equal to 0, 1 or ∞ defines a homeomorphism. The inverse map ν is then a coordinate chart for the stratum:

$$\nu : \mathring{\Delta}^{\dim \mathcal{B}} \xrightarrow{\cong} \mathcal{B}. \tag{4.2}$$

Remark 4.2.11. In fact ν is even a local isometry when $\mathcal{B}$ is equipped with the bottleneck distance and $\mathring{\Delta}^{\dim \mathcal{B}}$ with the $\|.\|_\infty$-metric: This is because in a fixed stratum barcodes have a constant number and nestings of bars; hence endpoints can be matched (in an increasing order) and the $\|.\|_\infty$-metric gives us the cost of the induced matching, which will be optimal when the barcodes are close enough.

Proposition 4.2.12. *For each $i \geqslant 0$, let B_i be the union of the* $\mathrm{Aut}(I, \leqslant)$*-orbits $\mathcal{B}$ with $\dim \mathcal{B} \leqslant i$. This defines a stratification of $\mathbf{Bar}^{d+1}$. The i-strata are the orbits of dimension i.*

Proof. As orbits provide a partition, the sets B_i for $i \in \mathbb{N}$ form a filtration of $\mathbf{Bar}^{d+1}$. Each of the B_i is also closed as the complement is open: barcodes close to a given barcode D have the same number of bars with endpoints in $(0,1)$ or more as can be deduced from Proposition 4.2.4.

Furthermore, the complements $B_i \backslash B_{i-1}$ are by definition the union of finitely many (disjoint) orbits $\mathcal{B}_1, \ldots, \mathcal{B}_{l_i}$, each homeomorphic to $\mathring{\Delta}^i$. The lemma below implies that the closure of each $\mathcal{B}_j$ does not intersect any of the $\mathcal{B}_k$ for $k \neq j$. Thus a path D_t of diagrams in $B_i \backslash B_{i-1}$ with $D_0 \in \mathcal{B}_j$ cannot leave the orbit $\mathcal{B}_j$. On the other hand, each stratum $\mathcal{B}_j$ is path connected since D_0 can be connected to any other diagram $D_1 \in \mathcal{B}_j$

by a linear path $D_t := (t\phi + (1-t)\mathrm{Id}).D_0$ where $\phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ is such that $D_1 = \phi(D_0)$. Hence the path connected components of $B_i \backslash B_{i-1}$ are the orbits of dimension i, and $B_i \backslash B_{i-1}$ is a manifold of dimension i. □

Lemma 4.2.13. *Let $D, D' \in \mathbf{Bar}^{d+1}$ be two barcodes. Then the following are equivalent:*

(1) There exists a non-decreasing map $\phi \in \mathrm{End}(I, \leqslant)$ such that $D' = \phi.D$;

(2) $\mathcal{B}_{D'} \subseteq \overline{\mathcal{B}_D}$, i.e. the stratum containing D' is in the closure of that containing D.

Proof. Assume (1) and let $\phi \in \mathrm{End}(\mathrm{I}, \leqslant)$ be such that $D' = \phi.D$. Consider the paths

$$\phi_t := (1-t)\mathrm{Id} + t\phi \qquad \text{and} \qquad D_t := \phi_t.D.$$

For $t \in [0,1)$, $\phi_t \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ and hence $D_t \in \mathcal{B}_D$. By continuity of the monoid action, Proposition 4.2.7, the path of barcodes D_t is continuous in t on the whole interval $[0,1]$. Consequently, in the limit, $D' = D_1 \in \overline{\mathcal{B}_D}$. If $D'' \in \mathcal{B}_{D'}$ is another barcode from the orbit defined by D' then there exists a $\beta \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ with $D'' = \beta.D'$. Consider $\beta.D_t = (\beta \circ \phi_t).D$. By the same argument as above, this is a continuous path of barcodes from D to D'' that is contained entirely in $\mathcal{B}_D$ with the possible exception when $t = 1$. Hence $D'' \in \overline{\mathcal{B}_D}$, and more generally $\mathcal{B}_{D'} \subseteq \overline{\mathcal{B}}_D$ which is (2).

Conversely, assume (2). If $D' \in \mathcal{B}_D$ then by definition of $\mathcal{B}_D$ there exists a $\phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ with $D' = \phi.D$ and (1) is satisfied. So we may assume $D' \notin \mathcal{B}_D$ (and hence the entire orbit $\mathcal{B}_{D'}$ is contained in the boundary $\overline{\mathcal{B}_D} \backslash \mathcal{B}_D$). Let $D_n, n \geqslant 0$, be a sequence in $\mathcal{B}_D$ converging to D', and let $\phi_n \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ such that $D_n = \phi_n.D$. Then, by the characterisation of the local neighborhoods in **Bar**, Proposition 4.2.4 for small enough ϵ and large enough n, the bars in D' can be matched up (one-to-one) with bars in D_n that are ϵ-close, and furthermore any additional bar in D_n is of length less ϵ. Let γ be an optimal matching from D_n to D' which collapses the small bars. The number of intervals in D_n is the same as in D, in particular finite, so for ϵ small enough relative to the distances between consecutive endpoints x_i, x_{i+1} of D', the union of the ϵ-small intervals in D_n do not cover any segment $[x_i, x_{i+1}]$. In this case we can use γ to construct a non-decreasing map $\phi' \in \mathrm{End}(\mathrm{I}, \leqslant)$ with $\phi'.D_n = D'$. Hence, $D' = \phi'.D_n = (\phi' \circ \phi_n).D = \phi.D$ with $\phi := \phi' \circ \phi_n$. This gives (1). □

Remark 4.2.14. The monoid $\mathrm{End}(\mathrm{I}, \leqslant)$ acts coordinate-wise on any simplex Δ^i, and under this action the orbit of any point in the interior $\mathring{\Delta}^i$ is the closed simplex Δ^i. As both μ and ν are compatible with the action

restricted to $\mathrm{Aut}(\mathrm{I}, \leqslant)$, they can be extended to equivariant maps from the closed simplex:

$$\nu : \Delta^{\dim \mathcal{B}} \longrightarrow \overline{\mathcal{B}}, \quad \text{with} \quad \nu(\phi.x) := \phi.\nu(x). \tag{4.3}$$

This is easily seen to be well-defined, i.e. given $x, x' \in \mathring{\Delta}^{\dim \mathcal{B}}$ and $\phi, \phi' \in \mathrm{End}(\mathrm{I}, \leqslant)$ such that $\phi.x = \phi'.x'$ we have $\phi.\nu(x) = \phi'.\nu(x')$. From the above lemma $\overline{\mathcal{B}}$ is the monoid orbit of $\mathrm{End}(\mathrm{I}, \leqslant)$, therefore the extension ν is surjective. Hence $\overline{\mathcal{B}}$ (as a set) can be identified as a quotient of the closed standard simplex. Indeed, ν is a strongly stratified map where the stratification on the standard simplex is the usual one and $\overline{\mathcal{B}}$ is considered a sub-stratified space of $\mathbf{Bar}^{d+1}$. In general ν is not injective on the boundary, see Example 4.2.19.

With both the stratifications of Filt_K and $\mathbf{Bar}^{d+1}$ in place, the persistence map $\mathrm{PH} : \mathrm{Filt}_K \to \mathbf{Bar}^{d+1}$ is then a map of stratified spaces in the following strong way:

Proposition 4.2.15. *The persistence map is a strongly stratified map. Namely, let $\mathcal{S}$ be an i-stratum in the space of filters. Then there exists a j-stratum $\mathcal{B}$ with $j \leqslant i$ and $\mathrm{PH}(\mathcal{S}) = \mathcal{B}$.*

Proof. Strata in the spaces of filters and barcodes are the $\mathrm{Aut}(\mathrm{I}, \leqslant)$-orbits with respect to which PH is equivariant by Lemma 4.2.5. We thus have for all $f \in \mathrm{Filt}_K$ and associated stratum $\mathcal{S}_f$

$$\mathrm{PH}(\mathcal{S}_f) = \{\mathrm{PH}(\phi.f) = \phi.\mathrm{PH}(f) \,|\, \phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)\} = \mathcal{B}_{\mathrm{PH}(f)}. \qquad \square$$

Furthermore, over a fixed stratum in $\mathbf{Bar}_K$ the fibers are all homeomorphic.

Proposition 4.2.16. *Let $\mathcal{B} \subseteq \mathbf{Bar}_K$ be a barcode stratum. The pre-images of* PH *over elements in $\mathcal{B}$ are pairwise homeomorphic.*

Proof. Let $D, D' \in \mathcal{B}$ so that $D' = \phi.D$ for some $\phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)$. By the equivariance of PH, Lemma 4.2.5, the action (by post-composition) of ϕ on Filt_K restricts to a map from $\mathrm{PH}^{-1}(D)$ to $\mathrm{PH}^{-1}(D') = \mathrm{PH}^{-1}(\phi.D) = \phi.\mathrm{PH}^{-1}(D)$. $\square$

Therefore $\mathbf{Bar}_K$ is a stratified subspace of $\mathbf{Bar}^{d+1}$ consisting of the union of strata $\mathrm{PH}(\mathcal{S})$, where $\mathcal{S} \subseteq \mathrm{Filt}_K$ varies over the set of strata of the space of filters. In particular $\mathbf{Bar}_K$ is a finite union of strata, finite dimensional and compact, unlike $\mathbf{Bar}^{d+1}$ which has infinitely many strata of arbitrarily large dimensions.

4.2.4 *The space $\mathbf{Bar}_K$ as a quotient space*

We will now show that as a topological space $\mathbf{Bar}_K$ is the quotient of the space of filters Filt_K induced by the persistence map.

Proposition 4.2.17. *The quotient topology on* $\mathbf{Bar}_K = \mathrm{PH}(\mathrm{Filt}_K)$ *induced by* PH *agrees with the bottleneck topology, that is we have a homeomorphism from the quotient*

$$\overline{\mathrm{PH}} : (\mathrm{Filt}_K / \sim) \xrightarrow{\cong} \mathbf{Bar}_K,$$

where $\sim$ *is defined by* $f \sim f' \Leftrightarrow \mathrm{PH}(f) = \mathrm{PH}(f')$.

Proof. By Theorem 4.2.3, $\mathrm{PH} : \mathrm{Filt}_K \to \mathbf{Bar}_K$ is continuous, and hence, by the universal property of the quotient, it induces a continuous bijection $\overline{\mathrm{PH}}$. It remains to prove that the inverse $\overline{\mathrm{PH}}^{-1}$ is also continuous, or equivalently that $\overline{\mathrm{PH}}$ is open.

Let U be an open set in $(\mathrm{Filt}_K / \sim)$ and let $D \in U$. Then by definition of the quotient topology $\mathrm{PH}^{-1}(U)$ is open and contains $\mathrm{PH}^{-1}(D)$. Since PH is continuous, $\mathrm{PH}^{-1}(D)$ is closed, and being a subset of I^K it is in fact compact. Thus for some $\eta > 0$ we have that the η-offset of $\mathrm{PH}^{-1}(D)$ lies in $\mathrm{PH}^{-1}(U)$:

$$\mathrm{PH}^{-1}(D)_\eta := \left\{ f \in \mathrm{Filt}_K, \exists g \in \mathrm{PH}^{-1}(D), \|f - g\|_\infty < \eta \right\} \subseteq \mathrm{PH}^{-1}(U).$$

We will show that $\mathrm{PH}^{-1}(D') \subseteq \mathrm{PH}^{-1}(D)_\eta$ for D' close enough to D in the bottleneck metric. By the above this implies that $\mathrm{PH}^{-1}(D') \subseteq \mathrm{PH}^{-1}(U)$, which amounts to $D' \in U$, and hence U is an open set in $\mathbf{Bar}_K$.

By Proposition 4.2.4, for ϵ small enough, bars $(b', d') \in D'$ that are not ϵ-small are matched up with bars of D that are ϵ-close. We consider the following equivalence relation on interval endpoints b', d' of D'. First, we deem equivalent all endpoints that are ϵ-close to the same endpoint x_i of D. Then, we deem equivalent two ϵ-small intervals that overlap: $[b', d'] \cap [b'', d''] \neq \varnothing$ and take the transitive closure of that relation. Since there are at most $\sharp K$ endpoints in D', the endpoints in the same equivalence class span a range of size at most $\sharp K \times \epsilon$. Thus if ϵ has been chosen small enough to start with, then it is impossible to find in the same equivalence class two endpoints of D' that are ϵ-close to distinct endpoints $x_i \neq x_j$ of D.

This allows constructing a map $\phi : \mathrm{I} \to \mathrm{I}$ such that $\phi(D') = D$ as follows: Over the span of an equivalence class of endpoints we define ϕ as the constant map with value x_i in the case where there is an endpoint b' or d' which is ϵ-close to the endpoint x_i of D, and with an arbitrary value in the span in the case where there is no such endpoint in the equivalence class. We extend ϕ linearly on I. By design ϕ differs from the identity map by at most $\sharp K \times \epsilon$, because the span of each equivalence class has diameter bounded by $\sharp K \times \epsilon$. Hence if we take an arbitrary $f \in \mathrm{PH}^{-1}(D')$, then $g := \phi \circ f$ belongs to $\mathrm{PH}^{-1}(D)$ by equivariance of PH, and g is $(\sharp K \times \epsilon)$-close to f. Up to shrinking ϵ so that $\sharp K \times \epsilon < \eta$, we have $f \in \mathrm{PH}^{-1}(D)_\eta$. Therefore $\mathrm{PH}^{-1}(D') \subseteq \mathrm{PH}^{-1}(D)_\eta$. □

The top dimensional strata of Filt_K consists of the injective filters $f : K \to \mathrm{I}$ that do not take the values 0 or 1. Hence the dimension of the top strata is $\sharp K$. The interval endpoints of a barcode $D = \mathrm{PH}(f)$ form a subset of the values of f and in general $\dim \mathcal{B}_D \leqslant \dim \mathcal{S}_f$. However, when f is injective, each simplex enters the next sublevel set of the filtration by itself and hence induces a change in homology. Thus in particular we see that the dimension of the top dimensional barcode strata in $\mathbf{Bar}_K$ is again $\sharp K$. Let

$$\mathbf{Bar}_K^{\mathrm{top}} := \bigcup_{\dim \mathcal{B} = \sharp K} \mathcal{B} = \{\mathrm{PH}(f) \mid f \in \mathrm{Filt}_K \text{ is injective and does not take the values 0 or 1}\}.$$

Proposition 4.2.18. *The barcode space $\mathbf{Bar}_K$ is the closure of its top dimensional strata, i.e.*

$$\mathbf{Bar}_K = \overline{\mathbf{Bar}_K^{\mathrm{top}}}.$$

Proof. Let $D = \mathrm{PH}(f)$ be a barcode in the image. We can always factor f as $f = \phi \circ g$ for some injective filter $g \in \mathrm{Filt}_K$ and a non-decreasing map $\phi \in \mathrm{End}(\mathrm{I}, \leqslant)$. By the equivariance of PH (Lemma 4.2.5), we have $D = \phi(\mathrm{PH}(g))$. Up to an arbitrarily small perturbation, g does not take the values 0 and 1, and hence $\mathrm{PH}(g)$ is an element in a top dimensional stratum. The result then follows from Lemma 4.2.13. □

The space $\mathbf{Bar}_K$ can thus be built as a quotient of a finite collection of (closed) simplices corresponding to the top dimensional barcode strata where some of the faces may be identified to each other and where i-dimensional faces may be reduced to a j-dimensional simplex through collapsing $(i-j)$-dimensional affine subspaces (corresponding to bars of length zero). The following example illustrates this.

Example 4.2.19. Let K represent the unit interval with vertices a, b and 1-simplex σ. The space of filters consists of two 3-dimensional strata corresponding to the induced orderings $a < b < \sigma$ and $b < a < \sigma$, which are mapped to each other via the action on K given by the involution (a, b) on its set of vertices. All faces are included in Filt_K and the two simplices are glued together along their common face corresponding to $a = b \leqslant \sigma$. Under the map PH the two 3-simplices are identified to one 3-simplex giving a unique top dimensional stratum $\mathcal{B}_{\mathrm{top}}$ in the space of barcodes $\mathbf{Bar}_K$ parametrising barcodes of the form $\{(x_1, \infty), (x_2, x_3)\}$ with $0 < x_1 < x_2 < x_3 < 1$. The barcode space $\mathbf{Bar}_K$ can then be identified with the quotient space of the closed 3-simplex where the 2-dimensional face corresponding to $0 \leqslant x_1 < x_2 = x_3 \leqslant 1$ is collapsed to the line segment $0 \leqslant x_1 = x_2 = x_3 \leqslant 1$, i.e. we identify:

$$(x_1, x_2, x_3) \sim (x_1', x_2', x_3') \quad \text{whenever } x_2 = x_3 \text{ and } x_2' = x_3'.$$

See Figure 4.1. We thus see that the fiber $\mathrm{PH}^{-1}(D)$ of a barcode $D = \{(x_1, \infty), (x_2, x_3)\} \in \mathbf{Bar}_K$ consists of two points if $x_1 < x_2 < x_3$, of

one point if $x_1 = x_2 < x_3$, and of an interval (consisting of two intervals glued together) if $x_1 < x_2 = x_3$.

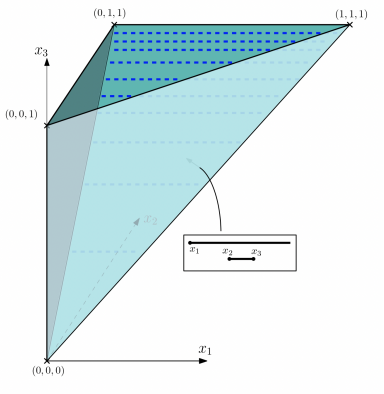


Figure 4.1: When K is the simplicial interval, the space $\mathbf{Bar}_K$ is the quotient space of the closed 3-simplex with each dotted line on the back face collapsed to one point.

4.3 *The persistence map as a polyhedral stratified fiber bundle*

Although the persistence map PH is globally not a fibration, we show in section 4.3.1 that it is a trivial fibration over each barcode stratum with a polyhedral complex as fiber. Using this structure in section 4.3.2, we derive topological properties of the fiber.

4.3.1 *Polyhedral structure on the fiber*

We strengthen Propositions 4.2.15 and 4.2.16, showing that, over each barcode stratum in the image $\mathbf{Bar}_K = \mathrm{PH}(\mathrm{Filt}_K)$, the persistence map is a trivial fiber bundle with a polyhedral complex as fiber. The intuition behind this result is that PH can be viewed as a piecewise linear projection as follows. Given a filter stratum $\mathcal{S} \subseteq \mathrm{Filt}_K$ and the barcode stratum $\mathcal{B} = \mathrm{PH}(\mathcal{S})$, the restriction $\mathrm{PH}_{|\mathcal{S}}$ can also be described as a map $\pi_{\mathcal{S}}^{\mathcal{B}} : \mathring{\Delta}^{\dim \mathcal{S}} \to \mathring{\Delta}^{\dim \mathcal{B}}$ via the coordinate charts $\mu : \mathring{\Delta}^{\dim \mathcal{S}} \xrightarrow{\cong} \mathcal{S}$ and $\nu : \mathring{\Delta}^{\dim \mathcal{B}} \xrightarrow{\cong} \mathcal{B}$ given by Eq. (4.1) and Eq. (4.2):

$$\begin{array}{ccc} \mathring{\Delta}^{\dim \mathcal{S}} & \xrightarrow{\mu} & \mathcal{S} \\ \Big\downarrow{\scriptstyle \pi_{\mathcal{S}}^{\mathcal{B}}} & & \Big\downarrow{\scriptstyle \mathrm{PH}} \\ \mathring{\Delta}^{\dim \mathcal{B}} & \xrightarrow{\nu} & \mathcal{B} \end{array} \tag{4.4}$$

The map $\pi_{\mathcal{S}}^{\mathcal{B}}$ is the linear projection from $\mathbb{R}^{\dim \mathcal{S}}$ to $\mathbb{R}^{\dim \mathcal{B}}$ that records the values of a filter f which are bounded interval endpoints in the associated barcode $\mathrm{PH}(f)$. That PH can be described by this diagram follows from the following two elementary observations: (i) $\pi_{\mathcal{S}}^{\mathcal{B}}$ is $\mathrm{Aut}(\mathrm{I}, \leqslant)$-equivariant since μ and ν are equivariant, where $\mathrm{Aut}(\mathrm{I}, \leqslant)$ acts on $\mathring{\Delta}^{\dim \mathcal{S}}$ and $\mathring{\Delta}^{\dim \mathcal{B}}$ coordinate-wise; and (ii) the set of bounded interval endpoints in the barcode PH is a subset of the values of f.

The fiber of such a projection map, restricted to the polyhedron $\mathring{\Delta}^{\dim \mathcal{S}} \cong \mathcal{S}$, is itself a polyhedron. In this section, we glue the polyhedra obtained in this way over the various filter strata $\mathcal{S}$ in order to describe the whole fiber of PH over a barcode stratum as a complex of polyhedra.

Recall that a *(bounded) polyhedron* is a bounded, finite intersection of closed half-spaces in a Euclidean space. The *dimension* of a polyhedron is the dimension of its affine hull. A *face* of a polyhedron $\mathcal{P}$ is the intersection of $\mathcal{P}$ with a supporting hyperplane, and is itself a polyhedron. In particular, a polyhedron $\mathcal{P}$ is a bounded convex Euclidean set. For such sets, we have the notions of relative interior and relative

boundary, which offer the advantage not to depend on the ambient Euclidean space, and with a slight abuse of notations we denote them by $\mathring{\mathcal{P}}$ and $\partial\mathcal{P}$ respectively.

Definition 4.3.1. A *polyhedral complex* is a finite set Π of polyhedra in some Euclidean space $\mathbb{R}^n$, such that (i) if F is a face of $\mathcal{P} \in \Pi$, then $F \in \Pi$, and (ii) for all $\mathcal{P}, \mathcal{P}' \in \Pi$, the intersection $\mathcal{P} \cap \mathcal{P}'$ is either empty or is a face of both $\mathcal{P}$ and $\mathcal{P}'$. By convention, the empty set is in Π. The *support* of Π is $\bigcup_{\mathcal{P}\in\Pi} \mathcal{P} \subseteq \mathbb{R}^n$. The *dimension* of a polyhedral complex is the maximal dimension of its polyhedra.

A map of polyhedral complexes, or *polyhedral map*, is a map that sends a polyhedron of the domain to a polyhedron of the co-domain surjectively, and whose restriction to each polyhedron $\mathcal{P}$ of the domain is affine (i.e. can be extended to an affine map to the ambient space of $\mathcal{P}$). In the case where the polyhedra are simplicial complexes, the notion of polyhedral map coïncides with that of simplicial map, in that it is induced by a map defined on the abstract simplicial complexes. More generally, polyhedral complexes can be thought of as geometric realisations of simplicial complexes. In fact, it is a standard fact that a polyhedral complex admits a finite triangulation on the same set of vertices. For the proof of this result and more about polyhedral geometry, we refer the reader to [GB09].

Joseph Gubeladze and Winfried Bruns. *Polytopes, rings, and K-theory*. Springer, 2009

Theorem 4.3.2. *Let $\mathcal{B} \subseteq \mathbf{Bar}_K$ be a barcode stratum in the image of* PH*, and $D \in \mathcal{B}$ be any barcode. For each filter stratum $\mathcal{S} \subseteq \mathrm{Filt}_K \cap \mathrm{PH}^{-1}(\mathcal{B})$, let*

$$\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) = \overline{\mathrm{PH}^{-1}(D) \cap \mathcal{S}} \subseteq \mathbb{R}^K$$

be the (closure of the) restriction to the stratum $\mathcal{S}$ of the fiber of the persistence map over D. Then:

(a) *Each $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is a polyhedron in $\mathbb{R}^K$ of dimension $\dim \mathcal{S} - \dim \mathcal{B}$, and is affinely isomorphic to the product of $\dim \mathcal{B} + 1$ standard simplices (of various dimensions):*

$$\Delta_0 \times \Delta_1 \times \Delta_2 \times \cdots \times \Delta_{\dim \mathcal{B}-1} \times \Delta_{\dim \mathcal{B}}.$$

Moreover, the relative interior of $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is $\mathrm{PH}^{-1}_{|\mathcal{S}}(D)$;

(b) *The fiber $\mathrm{PH}^{-1}(D)$ is the support of the polyhedral complex*

$$\left\{ \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) \mid \mathcal{S} \text{ filter stratum} \right\};$$

(c) *There is a homeomorphism* $\mathbf{\Phi}$ *giving the following commutative diagram:*

$$\begin{array}{ccc} \mathcal{B} \times \mathrm{PH}^{-1}(D) & \xrightarrow{\quad\mathbf{\Phi}\quad} & \mathrm{PH}^{-1}(\mathcal{B}) \\ & {\scriptstyle \pi_1} \searrow \quad \swarrow {\scriptstyle \mathrm{PH}} & \\ & \mathcal{B} & \end{array} \tag{4.5}$$

where π_1 *denotes the projection onto the first factor. Additionally, for any barcode* $D' \in \mathcal{B}$ *and filter stratum* $\mathcal{S}$*, the restricted map* $\mathbf{\Phi}(D', .)_{|\mathcal{S}}$ *is an affine isomorphism between* $\mathrm{PH}^{-1}_{|\mathcal{S}}(D)$ *and* $\mathrm{PH}^{-1}_{|\mathcal{S}}(D')$*. In particular, the polyhedral structure of* $\mathrm{PH}^{-1}(D)$ *is the same for all barcodes in* $\mathcal{B}$*.*

The situation of assertion **(a)** is depicted in Fig. 4.2.

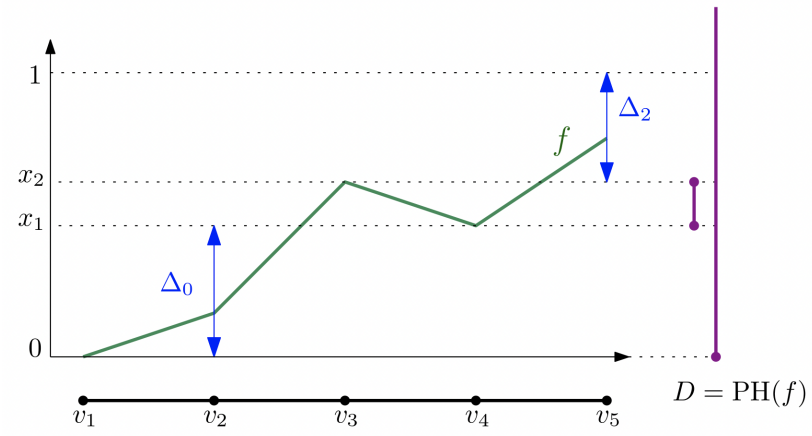


Figure 4.2: The filter f (green) on the complex obtained from subdividing the unit interval with five vertices yields the 2-dimensional barcode D (purple). The stratum $\mathcal{S}$ of f contains those filters which assign values in strictly increasing order to the vertices v_1, v_2, v_4, v_3 and v_5 and whose values over edges (v_i, v_{i+1}) are the maxima of their values over the endpoints v_i and v_{i+1}. The filters in the restricted fiber $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ must fix the equalities $f(v_1) = 0$, $f(v_3) = x_2$ and $f(v_4) = x_1$, but we may vary $f(v_2) \in [0, x_1]$ and $f(v_5) \in [x_1, x_2]$. These two degrees of freedom correspond to the two simplices Δ_0 and Δ_2, which are in this case of dimension 1, whereas Δ_1 is a singleton as no value of f lies in (x_1, x_2). We thus get the affine isomorphism between $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ and the product $\Delta_0 \times \Delta_1 \times \Delta_2$ of Theorem 4.3.2.

Proof. Recall from Eq. (4.4) how the persistence map rewrites as a projection map $\pi^{\mathcal{B}}_{\mathcal{S}} : \mathbb{R}^{\dim \mathcal{S}} \to \mathbb{R}^{\dim \mathcal{B}}$ onto the $\dim \mathcal{B}$ consecutive filter values that modify the homology groups of the sublevel sets. So the fiber of $\pi^{\mathcal{B}}_{\mathcal{S}}$ over the element $(x_1, \cdots, x_{\dim D}) := \nu^{-1}(D)$ is automatically affinely isomorphic to the product

$$\begin{aligned} \mathrm{PH}^{-1}(D) \cap \mathcal{S} &\cong (\pi^{\mathcal{B}}_{\mathcal{S}})^{-1}(x_1, \cdots, x_{\dim D}) \\ &\cong \mathring{\Delta}_0 \times \mathring{\Delta}_1 \times \mathring{\Delta}_2 \times \cdots \times \mathring{\Delta}_{\dim \mathcal{B}-1} \times \mathring{\Delta}_{\dim \mathcal{B}} \end{aligned} \tag{4.6}$$

of open simplices $\mathring{\Delta}_i$ whose dimensions sum up to $\dim \mathcal{S} - \dim \mathcal{B}$, where the open simplex $\mathring{\Delta}_i$ corresponds to values in-between x_i and x_{i+1}, the i-th and $i+1$-th endpoint values in D (recall the convention that $x_0 = 0$ and $x_{\dim D+1} = 1$). The linear isomorphism μ extends to a stratified

linear isomorphism between the closed simplex $\Delta^{\dim \mathcal{S}}$ and $\overline{\mathcal{S}}$. Therefore, the homeomorphism in Eq. (4.6) extends to the closure and yields the affine homeomorphism:

$$\overline{\mathrm{PH}^{-1}(D) \cap \mathcal{S}} \cong \Delta_0 \times \Delta_1 \times \Delta_2 \times \cdots \times \Delta_{\dim \mathcal{B}-1} \times \Delta_{\dim \mathcal{B}}. \tag{4.7}$$

This proves assertion **(a)**. We can also deduce from the previous arguments that if $\mathcal{S}' \subseteq \partial\mathcal{S}$ is a boundary stratum, then

$$\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D) \subseteq \partial\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D). \tag{4.8}$$

To see this, let us fix $f \in \mathcal{S} \cap \mathrm{PH}^{-1}(D)$ and let f' be an arbitrary filter in $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$. Then the respective coordinates $\mu^{-1}(f)$ and $\mu^{-1}(f')$ are in the fiber of $\pi^{\mathcal{B}}_{\mathcal{S}}$ over $\nu^{-1}(D)$, therefore so is the straight line $[\mu^{-1}(f); \mu^{-1}(f')]$ by linearity of the projection. Since $\mu([\mu^{-1}(f); \mu^{-1}(f'))) \subseteq \mathcal{S}$, we have $\mu([\mu^{-1}(f); \mu^{-1}(f'))) \subseteq \mathrm{PH}^{-1}_{|\mathcal{S}}(D)$ and we deduce that $f' \in \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$, and consequently $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D) \subseteq \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$. Since $\mathcal{S}' \subseteq \partial\mathcal{S}$, the polyhedron $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$ in fact lies in the relative boundary of $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$.

We now show assertion **(b)**, namely that the set of polyhedra $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is a polyhedral complex. So we first show that if $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is a polyhedron and $F \subset \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is one of its face, then F is of the form $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$. We assume that F is a proper face, i.e. has codimension 1 in $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$. The general case follows by induction on the codimension.

The restriction of the affine homeomorphism of Eq. (4.7) to the face F implies that F is affinely isomorphic to the product

$$F \cong \Delta_0 \times \Delta_1 \times \Delta_2 \times \cdots \times \Delta'_i \times \cdots \times \Delta_{\dim \mathcal{B}-1} \times \Delta_{\dim \mathcal{B}}, \tag{4.9}$$

where Δ'_i is a proper face of Δ_i. Note that Δ'_i is obtained by replacing one inequality between consecutive coordinates in Δ_i with an equality. This uniquely defines a stratum $\mathcal{S}' \subseteq \partial\mathcal{S}$ of codimension 1 in $\mathcal{S}$ such that $\mathring{F} \subseteq \mathcal{S}'$. Since $F \subseteq \mathrm{PH}^{-1}(D)$, we have $F \subseteq \overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$. However, by Eq. (4.8), the polyhedron $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$ lies inside the relative boundary of $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$, and in fact by convexity, inside a proper face of $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$. Since $F \subseteq \overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$ is itself a proper face, we deduce that $F = \overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$, as desired.

We now show that a non-empty intersection $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) \cap \overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$ of two polyhedra in the fiber is a common face of $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ and $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$. If $\mathcal{S} = \mathcal{S}'$ there is nothing to prove. Otherwise, $\mathcal{S} \cap \mathcal{S}' = \varnothing$ so that $\mathrm{PH}^{-1}_{|\mathcal{S}}(D) \cap \mathrm{PH}^{-1}_{|\mathcal{S}'}(D) = \varnothing$. We consider two cases:

1. Either $\mathcal{S}' \subseteq \partial\mathcal{S}$, and then $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$ lies in a proper face $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}''}}(D)$

of $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$, and we must have $\mathcal{S}' \subseteq \partial\mathcal{S}''$. Thus we are done by an induction on the codimension of $\mathcal{S}'$ in $\mathcal{S}$. Similarly if $\mathcal{S} \subseteq \partial\mathcal{S}'$;

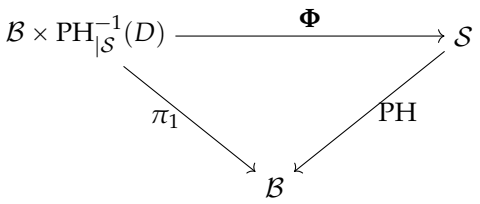
2. Or $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ and $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$ intersect only at their relative boundaries. In this case, by convexity of the two polyhedra, their intersection is the intersection $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}_2}}(D) \cap \overline{\mathrm{PH}^{-1}_{|\mathcal{S}'_2}}(D)$ of some proper faces $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}_2}}(D) \subseteq \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ and $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'_2}}(D) \subseteq \overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D)$. We are then left with the initial problem with polyhedra of smaller dimensions. We then conclude via an induction on the dimension of the polyhedra.

We now address the proof of assertion **(c)**. Given a barcode $D' \in \mathcal{B}$, define $\phi_{D'} : \mathrm{I} \to \mathrm{I}$ to be the unique piecewise linear interpolation of the increasing map that takes the (bounded) endpoints of D' to the (bounded) endpoints of D, further fixing 0 and 1. Clearly, $\phi_{D'}$ is an orientation preserving homeomorphism. From the homeomorphism $\nu : \mathring{\Delta}^{\dim \mathcal{B}} \xrightarrow{\cong} \mathcal{B}$, the intervals' endpoints in a barcode D' vary continuously with $D' \in \mathcal{B}$. Therefore $D' \mapsto \phi_{D'}$ is continuous, and in turn the map

$$\mathbf{\Phi} : (D', f) \in \mathcal{B} \times \mathrm{PH}^{-1}(D) \mapsto \phi_{D'}^{-1} \circ f \in \mathrm{PH}^{-1}(\mathcal{B}) \tag{4.10}$$

is continuous. Similarly, its inverse given by $\mathbf{\Phi}^{-1}(f') = (\mathrm{PH}(f'), \phi_{\mathrm{PH}(f')} \circ f')$ is continuous, so that $\mathbf{\Phi}$ is a homeomorphism. Using Lemma 4.2.5, we have $\mathrm{PH} \circ \mathbf{\Phi} = \pi_1$, i.e. the diagram in Eq. (4.5) commutes.

Let $\mathcal{S} \subseteq \mathrm{PH}^{-1}(\mathcal{B})$ be a filter stratum in the fiber. For $D' \in \mathcal{B}$, $\mathbf{\Phi}(D', .) = \phi_{D'}^{-1} \circ \_$ is the post-composition by $\phi_{D'}^{-1} \in \mathrm{Aut}(\mathrm{I}, \leqslant)$, see Eq (4.10), so by Lemma 4.2.5, the previous homeomorphism restricts to:



To finish the proof, we show that the homeomorphism $\mathbf{\Phi}(D', .) : \mathrm{PH}^{-1}(D) \cap \mathcal{S} \to \mathrm{PH}^{-1}(D') \cap \mathcal{S}$ is the restriction of an affine endomorphism of $\mathbb{R}^K$, for any barcode $D' \in \mathcal{B}$. Equivalently, we describe the coordinate functions $\mathbf{\Phi}(D', .)_\sigma : f \mapsto \mathbf{\Phi}(D', f)(\sigma)$ as affine forms, for each simplex $\sigma \in K$. Let $(x'_1, \cdots, x'_{\dim \mathcal{B}}) := \nu^{-1}(D')$, and let $x'_0 = 0$ and $x'_{\dim \mathcal{B}+1} = 1$ by convention. Given $\sigma \in K$ and $f \in \mathrm{PH}^{-1}(D') \cap \mathcal{S}$, there is an index $0 \leqslant i \leqslant \dim \mathcal{B}$ such that $x'_i \leqslant f(\sigma) \leqslant x'_{i+1}$. If $f' \in \mathrm{PH}^{-1}(D') \cap \mathcal{S}$ is another filter in the fiber, we have $f' = \phi \circ f$ for some $\phi \in \mathrm{Aut}(\mathrm{I}, \leqslant)$, and $\phi(D') = D'$ by the equivariance Lemma 4.2.5, so that $x'_i \leqslant f'(\sigma) \leqslant x'_{i+1}$ as well. Since $\phi_{D'}^{-1}$ is affine over $[x'_i; x'_{i+1}]$, we conclude that $\mathbf{\Phi}(D', .)_\sigma$ is the restriction of an affine map, as desired. □

4.3.2 *Topology of the fiber*

In this section, we collect a few results that restrict the topology of the fiber of PH over a barcode D in the image $\mathbf{Bar}_K = \mathrm{PH}(\mathrm{Filt}_K)$. We make use of the previous sections, and in particular we derive a bound for the dimension of the polyhedra in the fiber $\mathrm{PH}^{-1}(D)$.

With our first result we obtain finer control on the type of strata arising in $\mathbf{Bar}_K$. Denote by

$$\mathrm{rk}(\partial_p) := \dim \mathrm{Im}(\partial_p : C_p(K, \Bbbk) \to C_{p-1}(K, \Bbbk))$$

the rank of the boundary map in the simplicial chain complex, and by $\beta_p(K) := \dim \mathrm{H}_p(K, \Bbbk)$ the p-th Betti of K.

Proposition 4.3.3. *Let $D = (D_0, \ldots, D_d) \in \mathbf{Bar}_K$. Then, for any homology degree $0 \leqslant p \leqslant d$:*

(i) *The number of infinite intervals in D_p equals $\beta_p(K)$; and*

(ii) *The number of bounded intervals in D_p is smaller than or equal to $\mathrm{rk}(\partial_{p+1})$.*

Proof. Item **(i)** follows from the fact that infinite bars in D correspond to the homology of K, and item **(ii)** is a consequence of the fact that the introduction of a $(p+1)$-simplex in a complex either increases the dimension of the $(p+1)$-th homology or reduces the dimension of the p-th homology by one. □

Next we determine a bound for the dimension of the polyhedral complex $\mathrm{PH}^{-1}(D)$. Recall that barcode strata in $\mathbf{Bar}_K$ have maximal dimension $\sharp K$. Therefore by the **(a)** of Theorem 4.3.2 we have the obvious bound $\dim \mathrm{PH}^{-1}(D) \leqslant \mathrm{codim}(\mathcal{B})$, where we define the *codimension* of a barcode stratum $\mathcal{B}$ as:

$$\mathrm{codim}(\mathcal{B}) := \sharp K - \dim(\mathcal{B}) \geqslant 0.$$

To improve this bound, we introduce the following quantity, which can be thought of as the number of missing bounded intervals in the target barcode D.

Definition 4.3.4. Let $\sharp K$ be the number of simplices in K, and $\sharp D$ be the number of interval endpoints in D with finite value (counted with multiplicities). The *bounded deficit* of D is the quantity:

$$\frac{\sharp K - \sharp D}{2} \geqslant 0.$$

Note that the bounded deficit is the same for all barcodes inside a given stratum.

Proposition 4.3.5. *For any barcode $D \in \mathbf{Bar}_K$, the dimension of the fiber of* PH *over D is less than or equal to the bounded deficit:*

$$\dim \mathrm{PH}^{-1}(D) \leqslant \frac{\sharp K - \sharp D}{2} \leqslant \mathrm{codim}(D).$$

It follows in particular that $\mathrm{H}_p(\mathrm{PH}^{-1}(D)) = 0$ for $p > \frac{\sharp K - \sharp D}{2}$. We apply Proposition 4.3.5 to various fibers in the case where K is a triangle in Section 4.7. For these fibers, the bounded deficit is almost systematically a tight upper-bound on the dimension of the fiber.

Proof. The inequality $\frac{\sharp K - \sharp D}{2} \leqslant \mathrm{codim}(D)$ follows directly from the two inequalities $\dim D \leqslant \sharp K$ and $\dim D \leqslant \sharp D$. So we now investigate the left inequality $\dim \mathrm{PH}^{-1}(D) \leqslant \frac{\sharp K - \sharp D}{2}$.

Let $\mathcal{S}$ be a filter stratum and let $(x_1, \cdots, x_{\dim D}) := \nu^{-1}(D) \in \mathring{\Delta}^{\dim D}$, with $x_0 := 0$ and $x_{\dim D+1} := 1$. We set $\{D\} := \{x_i\}_{i=0}^{\dim D+1}$. Let $\sigma \in K$ be a simplex. By the equivariance Lemma 4.2.5, a filter $f \in \mathrm{PH}^{-1}(D) \cap \mathcal{S}$ satisfy $f(\sigma) \in \{D\}$ if and only if all filters $f \in \mathrm{PH}^{-1}(D) \cap \mathcal{S}$ satisfy $f(\sigma) \in \{D\}$. There are at least $\sharp D$ such simplices, since each interval endpoint in the barcode D must correspond to at least one simplex entering the filtration.

Meanwhile, there are exactly $\dim \mathcal{S} - \dim D$ distinct values $x = f(\sigma)$ that are not in $\{D\}$ for all filters $f \in \mathrm{PH}^{-1}(D) \cap \mathcal{S}$. Each such value x is attained by at least 2 simplices, as otherwise $f^{-1}(x)$ would be a singleton hence would have non-zero Euler characteristic and we would have $x \in \{D\}$. We obtain

$$2 \times (\dim \mathcal{S} - \dim D) + \sharp D \leqslant \sharp K.$$

From the item **(a)** of Theorem 4.3.2, the polyhedron $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ has dimension $\dim \mathcal{S} - \dim D$, and the above inequality yields $\dim \mathrm{PH}^{-1}(D) \leqslant \frac{\sharp K - \sharp D}{2}$. □

When D is the image of an injective filter f, then $\sharp K = \dim D = \sharp D$ and hence, by Proposition 4.3.5, the dimension of the fiber above is zero. We thus have the following immediate consequence.

Corollary 4.3.6. *If $f \in \mathrm{Filt}_K$ is injective, then the fiber $\mathrm{PH}^{-1}(\mathrm{PH}(f))$ is a finite set.*

Barcodes corresponding to injective filters are of maximal dimension. At the other extreme we have barcodes of dimension 1, i.e. barcodes D consisting simply of an infinite interval (x, ∞), possibly with multiplicity. The constant filter with value x gives rise to such a barcode. Although the fiber $\mathrm{PH}^{-1}(D)$ does not reduce to this constant filter, we nevertheless show that it retracts to it.

Proposition 4.3.7. *A barcode of dimension 1 in $\mathbf{Bar}_K$ has contractible fiber.*

Proof. Let x be the unique endpoint value in D. By assumption, there exists a filter f in $\mathrm{PH}^{-1}(D)$. Note that we must have $\min f = x$. We show that the straight line homotopy $(1-t)f + tx$ lies in $\mathrm{PH}^{-1}(D)$. For $t < 1$, the filters $(1-t)f + tx$ and f induce the same pre-order on

simplices of K, hence lie in a common stratum $\mathcal{S}$. By Proposition 4.2.15, $\mathrm{PH}(\mathcal{S}) = \mathcal{B}$ where $\mathcal{B}$ is the stratum containing D, which consist in barcodes D' obtained from D by moving the unique endpoint value x to any other value x'. Hence $\mathrm{PH}((1-t)f + tx)$ is such a barcode, and must equal D since $x' = \min(1-t)f + tx = \min f = x$. By continuity, at $t = 1$, we further get that the constant filter x is in the fiber. Then $\mathrm{PH}^{-1}(D)$ is star-shaped around x. □

Remark 4.3.8. The dimension $\dim \mathrm{PH}^{-1}(D)$ being upper-bounded, we can ask if conversely the star of every polyhedron in the fiber has dimension $\dim \mathrm{PH}^{-1}(D)$. The example section 4.7, suggests that this property holds true in the case where the complex K is a manifold. In general however, two distinct filters in the fiber may have neighborhoods of distinct dimensions, see for instance the barcode D of Fig. 4.3.

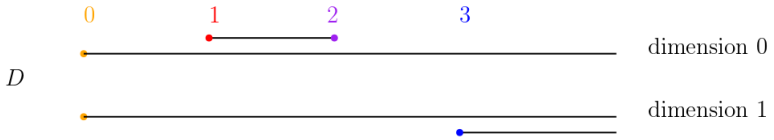


Figure 4.3: A barcode D where each endpoint value is given a different color. In Fig. 4.4 we draw the simplicial complex (left) and two components in the fiber with distinct dimension.

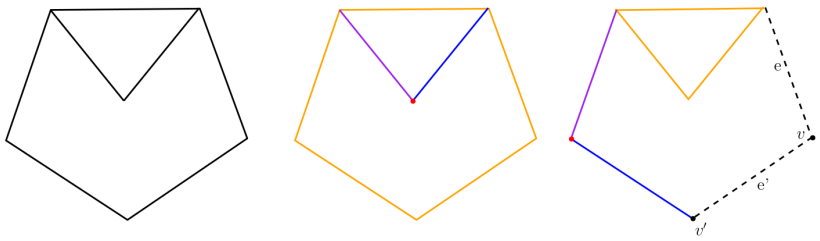


Figure 4.4: The simplicial complex (left) with two filters (middle and right) in the fiber $\mathrm{PH}^{-1}(D)$ over the barcode of Fig. 4.3, whose values on simplices are indicated by the colors. The filter f on the right has unspecified values on vertices and edges v, v', e, e' (colored in black), which means that whatever these values are, we can modify them as long as $f(v) = f(e)$ and $f(v') = f(e')$ and still get a filter whose barcode is D. Therefore, the star of f is 2-dimensional. However, the filter in the middle is alone in its neighborhood: all its values are fixed to endpoints of D and infinitesimal changes of any of these values yield filters out of $\mathrm{PH}^{-1}(D)$.

4.4 *The barcode category and the fiber functor*

In section 4.4.1, we define morphisms between barcodes. This makes the image of the persistence map into a topological category, which is homotopy discrete (Theorem 4.4.2). We further show that morphisms of barcodes can be deformed into particularly nice morphisms which we refer to as *simplicial* morphisms (Proposition 4.4.7). Such morphisms can always be described as finite compositions of morphisms between codimension 1 barcodes (Proposition 4.4.8). In section 4.4.2, morphisms of barcodes are pulled-back to provide maps of fibers, that are furthermore maps of polyhedral complexes up to homotopy (Proposition 4.4.12).

4.4.1 *The barcode category is homotopy discrete*

In this section we make the image $\mathbf{Bar}_K$ of the persistence map into a category, so that in the next section we view the fiber PH^{-1} as a functor.

Definition 4.4.1. We denote by $\mathbf{Bar_K}$ the **Top**-enriched *category of barcodes* with $D \in \mathbf{Bar}_K$ as objects and non-decreasing continuous maps $\phi \in \mathrm{End}(\mathrm{I}, \leqslant)$ such that $\phi(D) = D'$ as space of morphisms between D and D', $\mathbf{Bar_K}(D, D')$. Two morphisms $\phi_0, \phi_1 \in \mathbf{Bar_K}(D, D')$ are *homotopic as morphisms* if they belong to the same path connected component of $\mathbf{Bar_K}(D, D')$.

The first main result of this section is that the spaces of morphisms in $\mathbf{Bar_K}$ are made of contractible components.

Theorem 4.4.2. *For any two barcodes D, D' in $\mathbf{Bar}_K$, the space of morphisms $\mathbf{Bar_K}(D, D')$ has finitely many path connected components each of which is contractible. In particular, the category $\mathbf{Bar_K}$ is* homotopy discrete, *i.e.* $\mathbf{Bar_K}(D, D') \simeq h\mathbf{Bar_K}(D, D')$.

When D and D' belong to the same stratum, $\mathbf{Bar_K}(D, D')$ is contractible. Indeed, a morphism ϕ from D to D' is then simply a non-decreasing map that sends the i-th endpoint of D to the i-th endpoint of D'. Hence, given an arbitrary $\phi_0 \in \mathbf{Bar_K}(D, D')$, the straight line homotopy $(t, \phi) \mapsto (1-t)\phi + t\phi_0$ is a deformation retract of $\mathbf{Bar_K}(D, D')$ onto the point ϕ_0. In general, however, there may be more than one connected component as we will see.

As before, we use the coordinate charts $\nu : \mathring{\Delta}^{\dim D} \xrightarrow{\cong} \mathcal{B}_D$ of Eq. (4.2) in order to define the consecutive endpoint values $(x_1, \cdots, x_{\dim D}) = \nu^{-1}(D)$ and $(x'_1, \cdots, x'_{\dim D'}) = \nu^{-1}(D')$. By convention, we also set $x_0 = x'_0 = 0$ and $x_{\dim D+1} = x_{\dim D'+1} = 1$. In order to prove Theorem 4.4.2, we first associate to a morphism $\phi \in \mathbf{Bar_K}(D, D')$ the induced partial map from the endpoints of D to the endpoints of D'.

Definition 4.4.3. Let $\phi \in \mathbf{Bar_K}(D, D')$. The *index* of ϕ is the induced partial map $\Phi : \{x_1, \cdots, x_{\dim D}\} \to \{0, x'_1, \cdots, x'_{\dim D'}, 1\}$. Equivalently, the index is the collection of pre-images $\Phi := \{\Phi(j)\}_{j=0}^{\dim D'+1}$, where

$$\Phi(j) := \Phi^{-1}(x'_j) = \left\{x_i \mid \phi(x_i) = x'_j,\ 1 \leqslant i \leqslant \dim D\right\}$$

Given $\phi_0, \phi_1 \in \mathbf{Bar_K}(D, D')$, we write $\phi_0 \nearrow \phi_1$ whenever ϕ_1 is an extension of ϕ_0:

$$\forall 0 \leqslant j \leqslant \dim D' + 1,\ \Phi_0(j) \subseteq \Phi_1(j).$$

Lemma 4.4.4. *Let $\phi_0, \phi_1 \in \mathbf{Bar_K}(D, D')$. Then the following are equivalent:*

(i) The two morphisms are homotopic, i.e. $\phi_0 \sim \phi_1$ as morphisms.

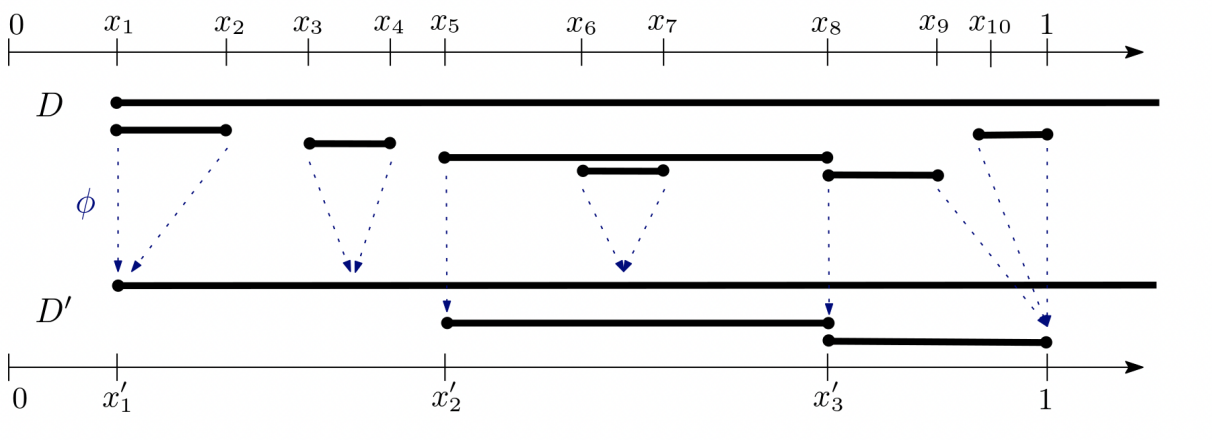


Figure 4.5: The map $\phi : D \to D'$ has index Φ consisting of $\Phi(0) = \varnothing$, $\Phi(1) = \{x_1, x_2\}$, $\Phi(2) = \{x_5\}$, $\Phi(3) = \{x_8\}$ and $\Phi(4) = \{x_9, x_{10}\}$.

(ii) *For all $(b,d) \in D$ and $(b',d') \in D'$, we have $(\phi_0(b), \phi_0(d)) = (b',d')$ if and only if $(\phi_1(b), \phi_1(d)) = (b',d')$.*

(iii) *The straight line interpolation $t \mapsto t\phi_0 + (1-t)\phi_1$ is a path in $\mathbf{Bar}_{\mathbf{K}}(D, D')$.*

(iv) *There is some $\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ such that $\phi \nearrow \phi_0$ and $\phi \nearrow \phi_1$.*

In particular, if $\phi_0 \nearrow \phi_1$, then ϕ_0 and ϕ_1 are homotopic.

Proof. $[(i) \Rightarrow (ii)]$: Let ϕ_t be a path in $\mathbf{Bar}_{\mathbf{K}}(D, D')$ joining ϕ_0 and ϕ_1. Given $(b,d) \in D$ and $(b',d') \in D'$, let

$$\mathrm{I}_{b,d}^{b',d'} := \left\{ t \in \mathrm{I},\ \phi_t(b,d) = (\phi_t(b), \phi_t(d)) = (b',d') \right\} \subseteq \mathrm{I}.$$

The sets $\mathrm{I}_{b,d}^{b',d'}$ are closed in I. There are also open since the map $\sum_{[(b,d),(b',d')] \in D \times D'} \mathbb{1}_{\mathrm{I}_{b,d}^{b',d'}} : \mathrm{I} \to \mathbb{N}$ is constant and equals the number of intervals in D'. Therefore, $\mathrm{I}_{b,d}^{b',d'} = \mathrm{I}$ or $\mathrm{I}_{b,d}^{b',d'} = \varnothing$.

$[(ii) \Rightarrow (iii)]$: For $t \in \mathrm{I}$, $\phi_t := t\phi_0 + (1-t)\phi_1$ is non-decreasing. Then, $\phi_t(b,d)$ equals a non-trivial interval $(b',d') \in D'$ if and only if $\phi_0(b,d) = \phi_1(b,d) = (b',d')$. All the other intervals $(b,d) \in D$ must then be trivialized, i.e. $\phi_0(b) = \phi_0(d)$ and $\phi_1(b) = \phi_1(d)$, so that $\phi_t(b) = \phi_t(d)$. This ensures that $\phi_t \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ since for instance $\phi_0 \in \mathbf{Bar}_{\mathbf{K}}(D, D')$.

$[(iii) \Rightarrow (i)]$: This implication is immediate. From now on, we have $(i) \Leftrightarrow (ii) \Leftrightarrow (iii)$.

$[(iii) \Rightarrow (iv)]$: Let ϕ_t be the straight line interpolation between ϕ_0 and ϕ_1. Let $0 \leqslant j \leqslant \dim D' + 1$. Let $1 \leqslant i \leqslant \dim D$ be such that $x_i \notin \Phi_0(j)$. Without loss of generality, we assume that $\phi_0(x_i) < x'_j$. Since x'_j is the endpoint of a non-trivial interval of D' and $\phi_0(D) = D'$, there must exist an endpoint $x_{i'}$, with $i' > i$, of an interval in D such that $\phi_0(x_{i'}) = x'_j$. By the item (ii), we also have $\phi_1(x_{i'}) = x'_j$. In turn, $\phi_1(x_i) \leqslant x'_j$ as ϕ_1 is non-decreasing. Therefore $\phi_{\frac{1}{2}}(x_i) < x'_j$, and in

particular $x_i \notin \Phi_{\frac{1}{2}}(j)$. We have therefore proved that $\Phi_{\frac{1}{2}}(j) \subseteq \Phi_0(j)$, so that $\phi_{\frac{1}{2}} \nearrow \phi_0$. Similarly, $\phi_{\frac{1}{2}} \nearrow \phi_1$.

$[(iv) \Rightarrow (i)]$: It is enough to show that if $\phi \nearrow \phi_0$, then ϕ and ϕ_0 are homotopic, as we then show in the exact same way that $\phi \sim \phi_1$, which implies $\phi_0 \sim \phi_1$. Let $(b,d) \in D$ be a bounded interval (the case where $d = d' = \infty$ is dealt with similarly) such that $\phi(b,d) = (b',d')$ for some $(b',d') \in D'$. Then there are indices $i < i'$ and $j < j'$ such that $(b,d) = (x_i, x_{i'})$ and $(b',d') = (x'_j, x'_{j'})$. Since $\Phi(j) \subseteq \Phi_0(j)$ and $\Phi(j') \subseteq \Phi_0(j')$, we have $\phi_0(b,d) = (b',d')$ as well. Therefore, the images $\phi_0(b,d)$ of intervals (b,d) such that $\phi(b,d) \in D'$ cover all the intervals in D'. Hence, any other interval in D is trivialized by ϕ_0, which guarantees that the assertion (ii) holds. We are done since $(ii) \Rightarrow (iii) \Rightarrow (i)$. □

Remark 4.4.5. Two morphisms $\phi_0, \phi_1 \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ with the same index are homotopic. In the case where D and D' belong to strata that differ by one dimension, the converse is true as well so that in this case the index is a (complete) homotopy invariant. To see this, observe that the index Φ associated to a morphism ϕ in the case $\dim D - \dim D' = 1$ cover the endpoints of D and must consist of singletons $\Phi(j)$ except for a unique $\Phi(k)$ which is the unique pair $\{x_i, x_{i+1}\}$ of consecutive endpoints of D collapsed by ϕ. Therefore, if $\phi_0, \phi_1 \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ are homotopic, by Lemma 4.4.4 there is a third morphism $\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ such that $\phi \nearrow \phi_0$ and $\phi \nearrow \phi_1$. By the above restriction on the index of morphisms in $\mathbf{Bar}_{\mathbf{K}}(D, D')$, this immediately implies that ϕ_0 and ϕ_1 have the same index.

We are now ready to prove the theorem.

Proof of Theorem 4.4.2. Let $D, D' \in \mathbf{Bar}_K$. By the assertion (iv) of Lemma 4.4.4, if $\phi_0, \phi_1 \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ are two morphisms such that $\phi_0 \nearrow \phi_1$, then ϕ_0 and ϕ_1 belong to the same path connected component of $\mathbf{Bar}_{\mathbf{K}}(D, D')$. Since there are finitely many possible indices, $\mathbf{Bar}_{\mathbf{K}}(D, D')$ has finitely many path connected components.

Let Ω be a path connected component of $\mathbf{Bar}_{\mathbf{K}}(D, D')$. By finiteness of all possible indices Φ, we can find a minimal ψ in Ω for the pre-order $\nearrow$, i.e. if $\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ is any another morphism satisfying $\phi \nearrow \psi$, then $\psi \nearrow \phi$ as well. In fact, the minimality of ψ together with the assertion (iv) of Lemma 4.4.4 implies that for any morphism $\phi \in \Omega$, we have $\psi \nearrow \phi$. We then have a deformation retraction

$$(t, \phi) \in [0;1] \times \Omega \longmapsto t\psi + (1-t)\phi \in \Omega$$

of Ω onto $\{\psi\}$, which is well-defined by the assertion (iii) of Lemma 4.4.4. Consequently, Ω is contractible. □

We denote by $\mathrm{I}(D)$ the simplicial complex obtained from subdividing the unit interval by the $\dim D$ endpoints $0 < x_1 < \cdots < x_{\dim D} < 1$. The

morphisms from D to D' that send endpoints to endpoints piecewise linearly are of particular interest to us.

Definition 4.4.6. A morphism $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$ is *simplicial* if it is induced from the geometric realisation of a simplicial map from $\mathrm{I}(D)$ to $\mathrm{I}(D')$.

Proposition 4.4.7. *For any two barcodes $D, D' \in \mathbf{Bar}_K$, each connected component of $\mathbf{Bar}_\mathbf{K}(D, D')$ contains at least one simplicial map.*

Proof. Let Ω be a path connected component of $\mathbf{Bar}_\mathbf{K}(D, D')$ and $\phi \in \Omega$. Denote by Φ the associated index. We can modify ϕ by sending each $x_i \notin \Phi$ to the unique endpoint x'_j satisfying $x'_j \leqslant \phi(x_i) < x'_{j+1}$.[4] The morphism ϕ is then simplicial. □

[4] Note that the morphism ϕ is then a maximal element for the pre-order $\nearrow$.

Therefore, a morphism of barcodes is (up to homotopy) a simplicial map over the unit interval. We next show that simplicial morphisms are finite compositions of simplicial morphisms between barcodes that differ by one dimension. Note that this implies that the morphisms in $h\mathbf{Bar}_\mathbf{K}$ are generated by morphisms between barcodes that differ by one dimension.

Proposition 4.4.8. *Let $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$ be a simplicial morphism. There exists a finite sequence of barcodes*

$$D =: D_1, D_2, \cdots, D_k := D'$$

satisfying $0 \leqslant \dim D_i - \dim D_{i+1} \leqslant 1$ for $1 \leqslant i \leqslant k-1$, together with simplicial morphisms $\phi_i \in \mathbf{Bar}_\mathbf{K}(D_i, D_{i+1})$ between them, such that

$$\phi = \phi_k \circ \phi_{k-1} \circ \cdots \circ \phi_2 \circ \phi_1.$$

Proof. We proceed by induction on $\dim D - \dim D'$. If $\dim D - \dim D' \in \{0, 1\}$, the statement is trivial. So we assume that $\dim D - \dim D' \geqslant 2$. We first treat the case where $\Phi(0) \neq \varnothing$, which means that ϕ sends the first endpoint x_1 of D to 0. Then, we may write ϕ as a composition $\phi' \circ \phi_1$ where ϕ_1 is the map that collapses the interval $[0; x_1]$ to 0 and sends $[x_1; 1]$ to $[0; 1]$ linearly. Clearly then, $D_1 := \phi_1(D)$ satisfies $\dim D - \dim D_1 = 1$, and the morphisms $\phi_1 \in \mathbf{Bar}_\mathbf{K}(D, D_1)$ and $\phi' \in \mathbf{Bar}_\mathbf{K}(D_1, D')$ are simplicial. The symmetric case where $\Phi(x_{\dim D'+1}) \neq \varnothing$ is dealt with similarly.

Therefore, we may assume that $\Phi(0) = \Phi(\dim D' + 1) = \varnothing$. The morphism ϕ being simplicial, there is an index $1 \leqslant j \leqslant \dim D'$ such that $|\Phi(j)| \geqslant 2$, i.e. the pre-image by ϕ of the endpoint x'_j is a sequence of at least two consecutive endpoints x_i of D. We choose an endpoint $x_i \in \Phi(j)$ for which there exists a non-trivial interval (x_i, d) (or (b, x_i)) in D such that $\phi(d) \neq x'_j$ (or $\phi(b) \neq x'_j$). Such an endpoint must exist because x'_j is the endpoint of a non-trivial interval in D',

which is the image by ϕ of an interval in D. We may assume that $x_{i+1} \in \Phi(j)$, as otherwise $x_{i-1} \in \Phi(j)$ and the rest of the proof can be conducted similarly. So $\phi(x_i) = \phi(x_{i+1}) = x'_j$, and in fact $\phi([x_i, x_{i+1}]) = x'_j$ since ϕ is non-decreasing. We may thus factor ϕ as $\phi' \circ \phi_1$, where $\phi_1 \in \mathrm{End}(\mathrm{I}, \leqslant)$ is the map that collapses the interval $[x_i; x_{i+1}]$ onto x'_j, extended linearly on I. The image $D_1 := \phi_1(D)$ then satisfies $\dim D - \dim D_1 = 1$, since $\phi_1((x_i, d)) = (x'_j, \phi(d))$ is a non-trivial interval in D_1 and ϕ acts injectively on the endpoints x_k, for $k \notin \{i, i+1\}$. Finally, the morphisms $\phi_1 \in \mathbf{Bar}_{\mathbf{K}}(D, D_1)$ and $\phi' \in \mathbf{Bar}_{\mathbf{K}}(D_1, D')$ are simplicial, which concludes the proof. □

4.4.2 *Monodromies and polyhedral maps of fibers*

We now analyse how fibers relate to each other as we cross barcode strata. We associate to each map of barcodes a map between the corresponding fibers as follows.

Definition 4.4.9. Let $D \in \mathbf{Bar}^{d+1}$ and $\phi \in \mathrm{End}(\mathrm{I}, \leqslant)$. The *monodromy* $\mathcal{L}_\phi$ associated to ϕ is the map between fibers:

$$\mathcal{L}_\phi : f \in \mathrm{PH}^{-1}(D) \longmapsto \phi \circ f \in \mathrm{PH}^{-1}(\phi(D)).$$

Note that the monodromy is well-defined since PH is $\mathrm{End}(\mathrm{I}, \leqslant)$-equivariant by Lemma 4.2.5,

$$\mathrm{PH}(\phi \circ f) = \phi(\mathrm{PH}(f)),$$

and hence $\phi \circ f \in \mathrm{PH}^{-1}(\phi(D))$. Furthermore, given another $\phi' \in \mathrm{End}(\mathrm{I}, \leqslant)$ by definition

$$\mathcal{L}_{\phi' \circ \phi} = \mathcal{L}_{\phi'} \circ \mathcal{L}_\phi,$$

and in particular, if ϕ is invertible then $\mathcal{L}_\phi$ is a homeomorphism. Furthermore, the monodromy assignment $\phi \mapsto \mathcal{L}_\phi$ is continuous. We thus see that monodromies turn the inverse image PH^{-1} into a functor as follows.

Definition 4.4.10. The *fiber functor* is a functor of **Top**-enriched categories

$$\mathrm{PH}^{-1} : \mathbf{Bar}_{\mathbf{K}} \longrightarrow \mathbf{Top}$$

that sends a barcode D to the fiber $\mathrm{PH}^{-1}(D)$ and a morphism $\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ to the monodromy $\mathcal{L}_\phi$.

Remark 4.4.11. By definition, the sets of morphisms in such a **Top**-enriched category come equipped with a topology, and so taking connected components yields the associated *homotopy category*. In our case, PH^{-1} descends to define a functor of homotopy categories:

$$\mathrm{PH}^{-1} : h\mathbf{Bar}_{\mathbf{K}} \longrightarrow h\mathbf{Top}.$$

When ϕ is not simplicial (Definition 4.4.6), it may send a bounded interval (b, d) of D "to the middle of nowhere", i.e. $\phi(b) = \phi(d)$ may not equal an interval endpoint in D'. In turn, if this is the case, the monodromy $\mathcal{L}_\phi$ is not a polyhedral map; see Fig. 4.6 below. However, simplicial maps induce polyhedral monodromies, as stated in the following result.

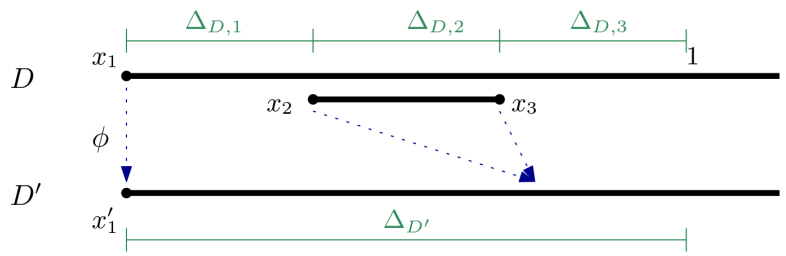


Figure 4.6: The barcode D has dimension 3. By Theorem 4.3.2, the polyhedral complex $\mathrm{PH}^{-1}(D)$ is made of polyhedra that are products $\Delta_{D,1} \times \Delta_{D,2} \times \Delta_{D,3}$ of 3 standard simplices. Each simplex $\Delta_{D,i}$ corresponds to the values of the filters in the fiber that are in-between the i-th endpoint x_i and $(i+1)$-th endpoint x_{i+1} of D. Likewise, a polyhedron in $\mathrm{PH}^{-1}(D')$ is simply a unique standard simplex $\Delta_{D'}$. The map ϕ collapses x_2 and x_3 strictly in-between x'_1 and 1. In turn, the monodromy $\mathcal{L}_\phi$ collapses the second standard simplex $\Delta_{D,2}$ of any polyhedron in $\mathrm{PH}^{-1}(D)$ in the interior of $\Delta_{D'}$, and so is not a polyhedral map.

Proposition 4.4.12. *Let $D, D' \in \mathbf{Bar}_K$ and $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$. If ϕ is simplicial, then the monodromy $\mathcal{L}_\phi : \mathrm{PH}^{-1}(D) \to \mathrm{PH}^{-1}(D')$ is a polyhedral map. In particular, for any $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$, the monodromy $\mathcal{L}_\phi$ is homotopic to a polyhedral map.*

Given a barcode D, we denote by $\mathrm{Aut}(\mathrm{I}, \leqslant)_D$ the stabilizer of D, i.e. the group of homeomorphisms of the real line that fix the endpoints of D and hence act trivially on D. The proof of Proposition 4.4.12 mainly relies on the following lemma.

Lemma 4.4.13. *Let $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$ be a simplicial morphism. Then there exists a group homomorphism*

$$\phi_{\mathrm{post}} : \mathrm{Aut}(I, \leqslant)_D \longmapsto \mathrm{Aut}(I, \leqslant)_{D'},$$

such that for all $\alpha \in \mathrm{Aut}(I, \leqslant)_D$, $\phi \circ \alpha = \phi_{\mathrm{post}}(\alpha) \circ \phi$. Similarly, there exists a group homomorphism

$$\phi_{\mathrm{pre}} : \mathrm{Aut}(I, \leqslant)_{D'} \longmapsto \mathrm{Aut}(I, \leqslant)_D,$$

such that for all $\beta \in \mathrm{Aut}(I, \leqslant)_{D'}$, $\phi \circ \phi_{\mathrm{pre}}(\beta) = \beta \circ \phi$.

Proof. Let $\alpha \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ such that $\alpha(D) = D$. Note that, since ϕ is simplicial, for any index $0 \leqslant i \leqslant \dim D$, the following alternative holds:

(a) Either $\phi([x_i, x_{i+1}]) = x'_j$ for some index $0 \leqslant j \leqslant \dim D' + 1$;

(b) Or $\phi_{|[x_i, x_{i+1}]}$ is a linear bijection onto $[x'_j, x'_{j+1}]$ for some index $0 \leqslant j \leqslant \dim D'$.

According to this alternative, we define $\beta := \phi_{\mathrm{post}}(\alpha)$ on each $\phi([x_i, x_{i+1}])$ as follows:

(a) Either $\phi([x_i, x_{i+1}]) = x'_j$, in which case we set $\beta(x'_j) := x'_j$;

(b) Or $\phi_{|[x_i,x_{i+1}]}$ is a linear bijection onto $[x'_j, x'_{j+1}]$, in which case we set $\beta_{|[x'_j,x'_{j+1}]} := \phi \circ \alpha \circ \phi^{-1}_{|[x_j,x_{j+1}]}$.

Note that in case (b), β is well-defined since $\alpha(D) = D$ implies that α restricts to a homeomorphism of each line segment $[x_i; x_{i+1}]$. Moreover, β is defined on the whole unit interval I since ϕ is surjective. It is then clear that we have $\beta \circ \phi \circ \alpha = \phi$ on each $[x_i; x_{i+1}]$, so that the equality holds on I as desired. Note that $\beta \in \mathrm{Aut}(\mathrm{I}, \leqslant)_{D'}$ since

$$\beta(D') = \beta(\phi \circ \alpha(D)) = \phi(D) = D'.$$

By construction, the association $\phi_{\mathrm{post}} : \alpha \mapsto \beta$ is a group homomorphism. Conversely, if we are rather given a map $\beta \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ such that $\beta(D') = D'$, then we can construct the map $\alpha := \phi_{\mathrm{pre}}(\beta)$ satisfying $\phi \circ \alpha = \beta \circ \phi$ as follows:

(a) Either $\phi([x_i, x_{i+1}]) = x'_j$, in which case we set $\alpha([x_i, x_{i+1}]) := \mathrm{Id}_{|[x_i,x_{i+1}]}$;

(b) Or $\phi_{|[x_i,x_{i+1}]}$ is a linear bijection onto $[x'_j, x'_{j+1}]$, in which case we set $\alpha_{|[x_i,x_{i+1}]} := \phi^{-1}_{|[x_i,x_{i+1}]} \circ \beta \circ \phi$.

This construction also yields a group homomorphism $\phi_{\mathrm{pre}} : \beta \mapsto \alpha$. □

Proof of Proposition 4.4.12. The second part of the statement follows directly from Proposition 4.4.7, which states that any morphism of barcodes is homotopic to a simplicial map. Henceforth, we fix a simplicial morphism $\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ and show that the monodromy $\mathcal{L}_\phi$ is a polyhedral map, i.e. it is affine on each polyhedron and sends polyhedra to polyhedra surjectively.

Let $\mathcal{S} \subseteq \mathrm{Filt}_K$ be a filter stratum. We have $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) \cong \Delta_0 \times \cdots \times \Delta_{\dim D}$, where the standard simplex Δ_i corresponds to filter values that are in-between the endpoints x_i and x_{i+1} of D. Since ϕ is affine over $[x_i, x_{i+1}]$, each coordinate function $\mathcal{L}_{\phi,\sigma} : f \in \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) \mapsto \mathcal{L}_\phi(f)(\sigma) \in \mathbb{R}$, for $\sigma \in K$, is affine as well. So the restriction of $\mathcal{L}_\phi$ to $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is an affine map, as desired.

It remains to show that the image $\mathcal{L}_\phi(\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D))$ equals a polyhedron $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D')$. We fix a filter $f \in \mathrm{PH}^{-1}_{|\mathcal{S}}(D)$ and denote by $\mathcal{S}'$ the stratum

containing $\phi \circ f$. Note that, if $g \in \mathrm{Filt}_K$ is a filter, then:

$$\begin{aligned} g \in \mathcal{L}_\phi(\mathrm{PH}^{-1}_{|\mathcal{S}}(D)) &\iff \exists f' \in \mathrm{PH}^{-1}_{|\mathcal{S}}(D),\ g = \phi \circ f' \\ &\iff \exists \alpha \in \mathrm{Aut}(\mathrm{I}, \leqslant)_D,\ g = \phi \circ (\alpha \circ f) \\ &\qquad\qquad \text{by the equivariance Lemma 4.2.5} \\ &\iff \exists \beta \in \mathrm{Aut}(\mathrm{I}, \leqslant)_{D'},\ g = \beta \circ (\phi \circ f) \\ &\qquad\qquad \text{by Lemma 4.4.13} \\ &\iff g \in \mathrm{PH}^{-1}_{\mathcal{S}'}(D'). \\ &\qquad\qquad \text{by the equivariance Lemma 4.2.5} \end{aligned}$$

Therefore, $\mathcal{L}_\phi(\mathrm{PH}^{-1}_{|\mathcal{S}}(D))$ equals $\mathrm{PH}^{-1}_{|\mathcal{S}'}(D')$ and in fact $\mathcal{L}_\phi(\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D))$ equals $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}'}}(D')$ since the image of a closed polyhedron via an affine map is again a closed polyhedron. □

The following is a consequence of the Propositions 4.4.7, 4.4.8 and 4.4.12:

Corollary 4.4.14. *For any $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$, the monodromy $\mathcal{L}_\phi$ is homotopic to a polyhedral map. This polyhedral map may further be chosen as a composition*

$$\mathcal{L}_{\phi_k \circ \cdots \circ \phi_1} = \mathcal{L}_{\phi_k} \circ \cdots \circ \mathcal{L}_{\phi_1}$$

of monodromies $\mathcal{L}_{\phi_k}$ that are polyhedral maps between fibers over barcodes that differ by one dimension.

Remark 4.4.15. The monodromy associated to a simplicial morphism ϕ acts as a projection map on each polyhedron $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ of the fiber. Indeed, recalling that $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ is (isomorphic to) a product $\Delta_0 \times \cdots \times \Delta_{\dim D}$ of standard simplices, we have the commutative diagram

$$\begin{array}{ccc} \Delta_0 \times \cdots \times \Delta_{\dim D} & \xrightarrow{\quad\pi\quad} & \Delta_0 \times \cdots \times \Delta_{\dim D} / \prod_{\phi(x_i)=\phi(x_{i+1})} \Delta_i \\ \cong \downarrow & & \downarrow \cong \\ \overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) & \xrightarrow{\quad\mathcal{L}_\phi\quad} & \mathcal{L}_\phi(\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)) \end{array}$$

where π is the projection map. In other words, the product of simplices describing the image polyhedron $\mathcal{L}_\phi(\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D))$ is obtained from the product describing $\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D)$ by collapsing standard simplices Δ_i whenever ϕ collapses the i-th and $i+1$-th endpoint of D.

4.5 *The space of barcodes is homotopically stratified*

In this section, we take a closer look at the stratification of barcodes and show that $\mathbf{Bar}_K$ is homotopically stratified. This naturally leads us to

introducing the entrance path category of barcodes. We observe that the entrance path category is isomorphic to the homotopy category, $h\mathbf{Bar}_{\mathbf{K}}$, of barcodes. Recall from Theorem 4.4.2 that $\mathbf{Bar}_{\mathbf{K}}$ is homotopy discrete. Similarly, we find that the space of entrance paths between fixed barcodes is contractible.

4.5.1 *Regularity of the stratification of barcodes*

The notion of stratification used so far (Definition 4.2.8) is a very weak one, as it does not impose restrictions on the neighborhoods of strata. Although the space of filters Filt_K is Whitney stratified, the regularity of the stratification of $\mathbf{Bar}_K$ is less apparent. We show that the space $\mathbf{Bar}_K$ of barcodes in the image of the persistence map is homotopically stratified in the sense of Quinn.[5] Stratified coverings over homotopically stratified spaces are classified by the entrance path category, which we introduce and analyse in the next section in the case of barcodes.

[5] Frank Quinn. Homotopically stratified sets. *Journal of the American Mathematical Society*, 1(2):441–499, 1988

The local neighborhoods in a homotopically stratified space X are defined in terms of paths that cross strata in decreasing order of dimension:

Definition 4.5.1. Let X be a stratified space. A continuous path $\gamma \in X^I$ is an *entrance path* if for any $0 \leqslant t \leqslant t' \leqslant 1$, the stratum containing $\gamma(t)$ has greater or equal dimension than that containing $\gamma(t')$.

Definition 4.5.2. Let X be a stratified space. An entrance path γ is *elementary* if it stays in a unique stratum until the very last moment, that is if $\gamma([0;1))$ belongs to a fixed stratum. Given two strata X^i and X^j, $j < i$, the *homotopy link* $\mathrm{Holink}(X^i, X^j)$ is the space of elementary paths starting in X^i and ending in X^j with the compact open topology.

Definition 4.5.3. A stratified space X is *homotopically stratified* if it satisfies the following conditions for any pair of strata X^i and X^j, where $j < i$:

1. The inclusion $X^j \hookrightarrow X^i \cup X^j$ is *tame*, which means that there is a strong deformation retraction of a neighborhood of X^j in $X^i \cup X^j$ onto X^j such that points remain in the same stratum until the very last moment during the deformation;
2. The evaluation at time $t = 1$

$$\mathrm{ev}_1 : \gamma \in \mathrm{Holink}(X^i, X^j) \longmapsto \gamma(1) \in X^j$$

is a fibration.

Note that, in the original formulation of homotopically stratified spaces [Qui88], the strata are not necessarily topological manifolds, and so we should really refer to the spaces of Definition 4.5.3 as manifold

stratified spaces, as done in [Wei94] for instance. However, this distinction is irrelevant for our purposes, since the strata in $\mathbf{Bar}_K = \mathrm{PH}(\mathrm{Filt}_K)$ are manifolds.

Shmuel Weinberger. *The topological classification of stratified spaces*. University of Chicago Press, 1994

Proposition 4.5.4. *The filtered space* $\mathbf{Bar}_K$ *is homotopically stratified.*

Proof. Let $\mathcal{B}$ and $\mathcal{B}'$ be two barcode strata with $\mathcal{B}' \subseteq \bar{\mathcal{B}}$. Recall that the coordinate chart ν extends to a continuous, surjective, stratum-preserving map $\nu : \Delta^{\dim \mathcal{B}} \to \bar{\mathcal{B}}$; see Eq. (4.3). Then, the inverse image $\nu^{-1}(\mathcal{B}')$ of $\mathcal{B}'$ is a union of faces in $\Delta^{\dim \mathcal{B}}$, and so the inclusion $\nu^{-1}(\mathcal{B}') \hookrightarrow \nu^{-1}(\mathcal{B}') \cup \mathring{\Delta}^{\dim \mathcal{B}}$ is tame, and a neighborhood deformation retracts onto $\nu^{-1}(\mathcal{B}')$. Composing with ν, we get a deformation retraction of a neighborhood of $\mathcal{B} \cup \mathcal{B}'$ onto $\mathcal{B}'$. Therefore, the inclusion $\mathcal{B}' \hookrightarrow \mathcal{B} \cup \mathcal{B}'$ is tame.

To check that the evaluation map $\mathrm{ev}_1 : \gamma \in \mathrm{Holink}(\mathcal{B}, \mathcal{B}') \mapsto \gamma(1) \in \mathcal{B}'$ is a fibration, it is enough to find a section for the map:[6]

[6] Witold Hurewicz. On the concept of fiber space. *Proceedings of the National Academy of Sciences of the United States of America*, 41(11):956, 1955

$$\tilde{\gamma}_h \in \mathrm{Holink}(\mathcal{B}, \mathcal{B}')^{[0;1]} \longmapsto (\mathrm{ev}_1 \circ \tilde{\gamma}_h, \tilde{\gamma}_0) \in \mathcal{B}'^{[0;1]} \times_{\mathcal{B}'} \mathrm{Holink}(\mathcal{B}, \mathcal{B}'),$$

where $\mathcal{B}'^{[0;1]} \times_{\mathcal{B}'} \mathrm{Holink}(\mathcal{B}, \mathcal{B}')$ is the fiber product

$$\mathcal{B}'^{[0;1]} \times_{\mathcal{B}'} \mathrm{Holink}(\mathcal{B}, \mathcal{B}') := \big\{(\gamma, \tilde{\gamma}) \in \mathcal{B}'^{[0;1]} \times \mathrm{Holink}(\mathcal{B}, \mathcal{B}') \,|\, \gamma(0) = \tilde{\gamma}(1)\big\}.$$

Using the coordinate chart $\nu : \mathring{\Delta}^{\dim \mathcal{B}'} \xrightarrow{\cong} \mathcal{B}'$, we may view a path $\gamma : h \in [0;1] \mapsto \gamma(h) \in \mathcal{B}'$ via its coordinates $0 < x_1^\gamma(h) < \cdots < x_{\dim \mathcal{B}'}^\gamma(h) < 1$. Given $0 \leqslant h \leqslant 1$, let $\phi_h^\gamma \in \mathrm{Aut}(\mathrm{I}, \leqslant)$ be the map that sends the endpoint values $x_i^\gamma(0)$ to $x_i^\gamma(h)$, and is extended linearly on each line segment $[x_i^\gamma(0), x_{i+1}^\gamma(0)]$. Clearly, the association $(\gamma, h) \mapsto \phi_h^\gamma$ is continuous, and we have $\gamma(h) = \phi_h^\gamma(\gamma(0))$. Then, the map

$$(\gamma, \tilde{\gamma}) \in \mathcal{B}'^{[0;1]} \times_{\mathcal{B}'} \mathrm{Holink}(\mathcal{B}, \mathcal{B}') \longmapsto \tilde{\gamma}_h(t) := \phi_h^\gamma(\tilde{\gamma}(t)) \in \mathrm{Holink}(\mathcal{B}, \mathcal{B}')^{[0;1]}$$

is the desired section. □

4.5.2 *The entrance path category of barcodes*

The entrance path category is the suitable generalisation of the fundamental groupoïd for stratified spaces where ordinary paths are replaced by entrance paths between points; see [Tre09] for the original constructions.

David Treumann. Exit paths and constructible stacks. *Compositio Mathematica*, 145(6):1504–1532, 2009

Definition 4.5.5. Let X be a stratified space. The *entrance path category* of X is $\mathbf{Ent}(X) := h\mathrm{Path}_{\leqslant}(X)$, where $\mathrm{Path}_{\leqslant}(X)$ is the topologically enriched category with X as the set of objects and spaces of entrance paths equipped with the compact open topology as morphisms. In other words, $\mathbf{Ent}(X)$ has the points of X as objects and the homotopy classes of entrance paths as morphisms.

Remark 4.5.6. In order to make $\mathrm{Path}_{\leqslant}(X)$ into a category where concatenation of paths defines a strictly associative composition of morphisms, paths of all positive lengths need to be allowed. This is analogous to replacing loop spaces by Moore loop spaces and the resulting morphism spaces are homotopy equivalent. In particular the definition of $\mathbf{Ent}(X)$ is not affected. We will ignore this subtlety in what follows.

Let ϕ in $\mathbf{Bar}_{\mathbf{K}}$ be a morphism between barcodes D and D'. Define the path:

$$\gamma_\phi(t) := (t\phi + (1-t)\mathrm{Id})(D).$$

For times $t < 1$, $t\phi + (1-t)\mathrm{Id}$ is a homeomorphism of the unit interval and so $\gamma_\phi(t)$ stays in the stratum containing D. Hence, $\gamma_\phi(t)$ is in fact an elementary entrance path. It is then clear that the association

$$\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D') \longmapsto \gamma_\phi \in \mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D')$$

is continuous.

Proposition 4.5.7. *The association $[\phi] \mapsto [\gamma_\phi]$ is functorial and induces an isomorphism of categories:*

$$h\mathbf{Bar}_{\mathbf{K}} \cong \mathbf{Ent}(\mathbf{Bar}_K).$$

Before proving this, we first characterize when elementary entrance paths are homotopic, in a similar fashion to Lemma 4.4.4 for morphisms of barcodes. Recall that an entrance path γ from D to D' is elementary if it stays in a unique stratum until the very last moment, that is if $\gamma([0;1))$ belongs to the stratum $\mathcal{B}_D$. Using the coordinate chart $\nu : \mathring{\Delta}^{\dim D} \xrightarrow{\cong} \mathcal{B}_D$, we may view $\gamma_{|[0;1)}$ via its coordinates $0 < x_1^\gamma(t) < \cdots < x_{\dim D}^\gamma(t) < 1$, $0 \leqslant t < 1$. Alternatively, for each interval $(b,d) \in D$, there is a continuously evolving interval $(b^\gamma(t), d^\gamma(t))$ starting at (b,d), $0 \leqslant t < 1$, and together the intervals $(b^\gamma(t), d^\gamma(t))$ form the barcode $\gamma(t)$.

Lemma 4.5.8. *Let $\gamma_0, \gamma_1 \in \mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D')$ be elementary entrance paths. Then the following are equivalent:*

(i) The two entrance paths are homotopic through elementary entrance paths.

(ii) For all $(b,d) \in D$ and $(b',d') \in D'$, we have $\lim_{t\to 1^-}(b^{\gamma_0}(t), d^{\gamma_0}(t)) = (b',d')$ if and only if $\lim_{t\to 1^-}(b^{\gamma_1}(t), d^{\gamma_1}(t)) = (b',d')$.

(iii) For all $(b,d) \in D$ and $(b',d') \in D'$, we have $\lim_{t\to 1^-}(d^{\gamma_0}(t) - b^{\gamma_0}(t)) = 0$ if and only if $\lim_{t\to 1^-}(d^{\gamma_1}(t) - b^{\gamma_1}(t)) = 0$.

Proof. [$(i) \Rightarrow (ii)$ and (iii)]: Let γ_h be a homotopy between γ_0 and γ_1 through entrance paths, and let $h \in [0;1]$. Using Proposition 4.2.4, we can partition the intervals in D into sets $D_{(a)}^{\gamma_h}$ and $D_{(b)}^{\gamma_h}$ as follows:

(a) For each interval $(b', d') \in D'$, there is a unique $(b, d) \in D$ for which $\lim_{t\to 1^-}(b^{\gamma_h}(t), d^{\gamma_h}(t)) = (b', d')$;

(b) For all other intervals $(b, d) \in D$, we have $\lim_{t\to 1^-}(b^{\gamma_h}(t) - d^{\gamma_h}(t)) = 0$.

By continuity, the classification remains constant along the homotopy $h \mapsto \gamma_h$, i.e. $D^{\gamma_h}_{(a)}$ and $D^{\gamma_h}_{(b)}$ are the same for all $0 \leqslant h \leqslant 1$.

$[(ii) \Rightarrow (iii)]$: By assumption $D^{\gamma_0}_{(a)} = D^{\gamma_1}_{(a)}$, hence $D^{\gamma_0}_{(b)} = D^{\gamma_1}_{(b)}$.

$[(iii) \Rightarrow (ii)]$: By assumption $D^{\gamma_0}_{(b)} = D^{\gamma_1}_{(b)}$, hence $D^{\gamma_0}_{(a)} = D^{\gamma_1}_{(a)} =: D_{(\mathrm{a})}$.

Then, γ_0 and γ_1 induce bijections from $D_{(\mathrm{a})}$ to the set of intervals in D'. Since γ_0 and γ_1 are entrance paths, these bijections are monotonic with respect to the endpoint values. So they are in fact the same bijections.

$[(ii)$ and $(iii) \Rightarrow (i)]$: We define a homotopy $h \mapsto \gamma_h$ between the restrictions of γ_0 and γ_1 to $[0; 1)$ by interpolating the coordinates:

$$\forall 0 \leqslant t < 1,\ \forall 1 \leqslant i \leqslant \dim D,\ x_i^{\gamma_h}(t) := hx_i^{\gamma_1}(t) + (1-h)x_i^{\gamma_0}(t).$$

For each interval $(b, d) \in D$, we then have

$$\forall 0 \leqslant t < 1,\ (b^{\gamma_h}(t), d^{\gamma_h}(t)) = (hb^{\gamma_1}(t) + (1-h)b^{\gamma_0}(t), hd^{\gamma_1}(t) + (1-h)d^{\gamma_0}(t)).$$

Since $D^{\gamma_0}_{(b)} = D^{\gamma_1}_{(b)}$ and $D^{\gamma_0}_{(a)} = D^{\gamma_1}_{(a)}$, we have $\lim_{t\to 1^-}(b^{\gamma_h}(t), d^{\gamma_h}(t)) = (b', d')$ (resp. $\lim_{t\to 1^-} d^{\gamma_h}(t) - b^{\gamma_h}(t) = 0$) if and only if $\lim_{t\to 1^-}(b^{\gamma_0}(t), d^{\gamma_0}(t)) = (b', d')$ (resp. $\lim_{t\to 1^-} d^{\gamma_0}(t) - b^{\gamma_0}(t) = 0$). Therefore,

$$\forall h \in [0; 1],\ \lim_{t\to 1^-} \gamma_h(t) = D',$$

hence the homotopy $h \mapsto \gamma_h$ extends to a homotopy between γ_0 and γ_1 on the whole unit interval. □

Proof of Proposition 4.5.7. The functoriality of $[\phi] \mapsto [\gamma_\phi]$ amounts to showing that if $\phi \in \mathbf{Bar}_{\mathbf{K}}(D, D')$ and $\psi \in \mathbf{Bar}_{\mathbf{K}}(D', D'')$ are two morphisms, then the entrance paths $\gamma_{\psi\circ\phi}$ and $\gamma_\psi.\gamma_\phi$ are homotopic. For $t \in [0; 1]$, let ϕ_t denote the interpolated map $(1-t)\mathrm{Id} + t\phi$, so that $\gamma_\phi(t) = \phi_t(D)$, $\gamma_\psi(t) = \psi_t(D')$ and $\gamma_{\psi\circ\phi}(t) = (\psi \circ \phi)_t(D)$. Besides, the concatenated path $\gamma_\psi.\gamma_\phi(t)$ equals $\phi_{2t}(D)$ for $t \leqslant \frac{1}{2}$ and $\psi_{2t-1}(D') = \psi_{2t-1}(\phi(D))$ for $t \geqslant \frac{1}{2}$. A homotopy between $\gamma_{\psi\circ\phi}$ and $\gamma_\psi.\gamma_\phi$ can then be defined as:

$$\mathrm{H}(h, t) := [h\phi_{2t} + (1-h)(\psi \circ \phi)_t](D) \text{ for } t \leqslant \frac{1}{2},$$

and

$$\mathrm{H}(h, t) := [h\psi_{2t-1} \circ \phi + (1-h)(\psi \circ \phi)_t](D) \text{ for } t \geqslant \frac{1}{2}.$$

Hence, we obtain a functor from $h\mathbf{Bar}_{\mathbf{K}}$ to $\mathbf{Ent}(\mathbf{Bar}_K)$, which is the identity on objects. Given barcodes D, D', we show that this functor gives a bijection $h\mathbf{Bar}_{\mathbf{K}}(D, D') \xrightarrow{\cong} \mathbf{Ent}(\mathbf{Bar}_K)(D, D')$.

Let γ be an elementary entrance path from D to D'. We construct a morphism ϕ from D to D' using the classification of the maps $(b^\gamma(t), d^\gamma(t))$:

(a) For each interval $(b', d') \in D'$, there is a unique $(b, d) \in D$ for which $\lim_{t\to 1^-}(b^\gamma(t), d^\gamma(t)) = (b', d')$. We then set $\phi(b) := b'$ and $\phi(d) := d'$;

(b) For all other intervals $(b, d) \in D$, we have $\lim_{t\to 1^-}(b^\gamma(t) - d^\gamma(t)) = 0$. We then set $\phi(b) = \phi(d)$ to be an arbitrary value such that ϕ remains non-decreasing.

We extend ϕ to a non-decreasing map of the unit interval arbitrarily. Since $\gamma(1) = D'$, we have $\phi(D) = D'$ by construction. Besides, the interpolated path $\gamma_\phi(t) = [t\phi + (1-t)\mathrm{Id}](D)$ is an elementary entrance path beween D and D', which satisfies the (i) and (ii) of Lemma 4.5.8 with respect to γ, hence is homotopic to γ. More generally, an arbitrary entrance path γ from D to D' is homotopic to a finite concatenation of elementary paths. In turn, γ is homotopic to an interpolated morphism γ_ϕ by applying the previous argument to each elementary path. Therefore, the map $[\phi] \in h\mathbf{Bar}_\mathbf{K}(D, D') \mapsto [\gamma_\phi] \in \mathbf{Ent}(\mathbf{Bar}_K)$ is surjective.

Besides, comparing the (ii) of Lemma 4.4.4 with the (ii) of Lemma 4.5.8, we see that if two morphisms between D and D' induce interpolated paths that are homotopic, then they must be homotopic. Consequently, the map $[\phi] \in h\mathbf{Bar}_\mathbf{K}(D, D') \mapsto [\gamma_\phi] \in \mathbf{Ent}(\mathbf{Bar}_K)$ is injective. □

Remark 4.5.9. It is well-known that coverings over a topological space X satisfying mild properties are classified by the fundamental groupoïd of X. When X is stratified, it is natural to consider stratified coverings, i.e. maps restricting to coverings over each individual stratum. If X is homotopically stratified with locally simply connected and locally connected strata, the stratified coverings which are either local homeomorphisms or *branched covers* are classified by the entrance path category of X.[7] That is, functors from $\mathbf{Ent}(X)$ to $\mathbf{Set}$ functorially give rise to such stratified coverings, in fact also to constructible cosheaves, and conversely.[8] Although in the case of barcodes the inverse image PH^{-1} is valued in $h\mathbf{Top}$, we have a natural set valued functor:

$$\mathbf{Ent}(\mathbf{Bar}_K) \xrightarrow{\mathrm{PH}^{-1}} h\mathbf{Top} \xrightarrow{\pi_0} \mathbf{Set}.$$

[7] Jonathan Woolf. The fundamental category of a stratified space. *arXiv preprint arXiv:0811.2580*, 2008

[8] See [CP16] for similar classifications of functors over $\mathbf{Ent}(X)$ when X is conically stratified, and [Tre09] for some 2-categorical equivalences.

We next prove an analogue of Theorem 4.4.2 for entrance paths.

Proposition 4.5.10. *For any two barcodes D, D' in $\mathbf{Bar}_K$, the space of morphisms $\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D')$ has finitely many path connected components each of which is contractible. In other words, the category $\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)$ is* homotopy discrete, *i.e. $\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D') \simeq h\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D')$.*

Proof. Let D and D' be two barcodes, and let $\mathrm{Holink}(D, D')$ be the space of elementary entrance paths from D to D'. For homotopically stratified metric spaces, the space of entrance paths and that of elementary entrance paths are homotopy equivalent [Milo6, Theorem 4.9][9], from which we deduce that:

[9] David A Miller. Popaths and holinks. *Journal of Homotopy and Related Structures*, 1(1):1–9, 2006

$$\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D') \simeq \mathrm{Holink}(D, D').$$

The proof of the statement then follows from Lemma 4.5.8. In more detail, let Ω be a path connected component in $\mathrm{Holink}(D, D')$, and let $\gamma_0 \in \Omega$. Recall that we can partition the intervals in D into sets $D^{\gamma_0}_{(a)}$ and $D^{\gamma_0}_{(b)}$ as follows:

(a) For each interval $(b', d') \in D'$, there is a unique $(b, d) \in D$ for which $\lim_{t\to 1^-}(b^{\gamma_0}(t), d^{\gamma_0}(t)) = (b', d')$;

(b) For all other intervals $(b, d) \in D$, we have $\lim_{t\to 1^-}(b^{\gamma_0}(t) - d^{\gamma_0}(t)) = 0$.

From Lemma 4.5.8, any other path $\gamma \in \Omega$ satisfies $D^{\gamma}_{(a)} = D^{\gamma_0}_{(a)}$ and $D^{\gamma}_{(b)} = D^{\gamma_0}_{(b)}$. Define $r : \Omega \times [0;1] \times [0;1) \to \mathbf{Bar}_K$ by:

$$r : (\gamma, h, t) \in \Omega \times [0;1] \times [0;1) \longmapsto \left\{((1-h)b^{\gamma}(t) + hb^{\gamma_0}(t), (1-h)d^{\gamma}(t) + hd^{\gamma_0}(t))\right\}_{(b,d)\in D} \in \mathcal{B}_D \subseteq \mathbf{Bar}_K.$$

We can continuously extend r at time $t = 1$ by $r(\gamma, h, 1) := D'$. We then get a deformation retraction

$$R : (\gamma, h) \in \Omega \times [0;1] \longmapsto r(\gamma, h, .) \in \Omega$$

of Ω onto $\{\gamma_0\}$. □

Combining Proposition 4.5.7 and 4.5.10 we can summarise our results in this section with the following.

Corollary 4.5.11. *For any two barcodes $D, D' \in \mathbf{Bar}_K$ we have*

$$\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D') \simeq h\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K)(D, D') = \mathbf{Ent}(\mathbf{Bar}_K)(D, D') \cong h\mathbf{Bar}_{\mathbf{K}}(D, D') \simeq \mathbf{Bar}_{\mathbf{K}}(D, D'),$$

and hence the natural weak equivalences of categories

$$\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K) \xrightarrow{\simeq} h\mathrm{Path}_{\leqslant}(\mathbf{Bar}_K) = \mathbf{Ent}(\mathbf{Bar}_K) \xleftarrow{\cong} h\mathbf{Bar}_{\mathbf{K}} \xleftarrow{\simeq} \mathbf{Bar}_{\mathbf{K}}.$$

4.6 *Variations of the fiber problem*

We adapt our analysis to two further situations of interest, namely when we remove the constraint that filters and barcodes take value in the unit interval, and when we restrict PH to the subspace of filters determined by their values on vertices. Finally, we point out that the action of the symmetries of K on the filters restricts to the fibers.

4.6.1 *The case of unbounded filters and barcodes*

In the previous sections, the values of filters and the interval endpoints of barcodes were constrained to lie in the unit interval. Here we briefly outline how our analysis can be adapted when we consider the unbounded case and replace the interval I by the real line $\mathbb{R}$. We denote by $\mathrm{Filt}(\mathbb{R})_K$ the filter functions with unrestricted real values and by $\mathbf{Bar}(\mathbb{R})$ the space of finite barcodes with unrestricted endpoints. As before, persistent homology defines a map

$$\mathrm{PH} : \mathrm{Filt}(\mathbb{R})_K \longrightarrow \mathbf{Bar}(\mathbb{R}).$$

Let $\mathrm{Aut}(\mathbb{R}, \leqslant)$ be the group of continuous automorphisms of the ordered real line that are the identity outside a compact set. Similarly let $\mathrm{End}(\mathbb{R}, \leqslant)$ be the monoid of continuous order preserving maps of the real line that are the identity outside a compact set. Both spaces then act on the extended spaces of filters and barcodes. The proof of Lemma 4.2.5 generalises to show that the persistence map above is equivariant with respect to these extended actions. The actions are continuous as in Proposition 4.2.7 when we equip $\mathrm{Aut}(\mathbb{R}, \leqslant)$ and $\mathrm{End}(\mathbb{R}, \leqslant)$ with the L^∞ topology. Using the action of $\mathrm{Aut}(\mathbb{R}, \leqslant)$ one can construct stratifications of filter and barcode spaces such that the $\mathrm{Aut}(\mathbb{R}, \leqslant)$-orbits are the strata and the analogues of Propositions 4.2.9 and 4.2.12 hold.

Indeed, most of the results and their proofs can easily be adapted with the caveat that fibers no longer have to be compact. Thus, item **(a)** of Theorem 4.3.2 needs to be reinterpreted: The polyhedron in the fiber are not necessarily products of closed standard simplices but instead we have

$$\overline{\mathrm{PH}^{-1}_{|\mathcal{S}}}(D) \cong \Delta_1 \times \cdots \times \Delta_{\dim \mathcal{B}-1} \times \Delta'_{\dim \mathcal{B}},$$

where the last term may be a simplex missing its last face, i.e. $\Delta'_{\dim \mathcal{B}}$ may be the closed i-simplex Δ^i or of the form:

$$\Delta^i_\infty := \{0 \leqslant x_1 \leqslant \cdots \leqslant x_i < \infty\}.$$

Modulo this subtlety, Theorems 4.3.2 and Proposition 4.4.12 hold also in the unbounded case we consider here, and PH is again a stratified fiber bundle whose fibers are (possibly unbounded) polyhedra. The proofs of these results in the unbounded situation do not present additional difficulties.

Next we provide a necessary and sufficient criterion for the simplicial complex K that ensures that all the fibers of PH are bounded.

Definition 4.6.1. A subset $L \subseteq K$ is *-removable*, or simply *removable*, if $K \backslash L$ is a subcomplex of K and the inclusion $K \backslash L \hookrightarrow K$ induces an isomorphism on (standard) homology with -coefficients. K is said to be *-essential*, or simply *essential*, if it has no removable subsets.

For instance, any pair (σ, σ') where σ' has σ as its only co-face provides an example of a removable subset for any . This is an elementary collapse familiar from simple homotopy theory. More elaborate examples include the wedge product $K \vee A$ of two simplicial complexes K and A, where A is -acyclic, i.e. A has trivial reduced (ordinary) homology with coefficients. Then $L = A \backslash \{*\}$ is removable. A rich source of such A are the classifying spaces of perfect groups, or the classifying spaces of finite groups when is of characteristic zero.

Proposition 4.6.2. *The fibers of the persistence map* PH *are all compact if and only if the complex K is essential.*

Proof. If K is not essential, it has a removable subset $L \subseteq K$. Given a partial filter $f : K \backslash L \to \mathbb{R}$ and a real value x with $x \geqslant \max_{\sigma \in K \backslash L} f(\sigma)$, f can be extended to a filter f_x on all of K by assigning the common value x to all the simplices in L. Thus the fiber of $D = \mathrm{PH}(f_x)$ contains the open half line $\{f_x \mid x \in [\max_{\sigma \in K \backslash L} f(\sigma); \infty)\}$ and is hence not compact.

Conversely, if PH has a non-compact fiber over some barcode D, it means there exists an $f \in \mathrm{PH}^{-1}(D)$ attaining values higher than the largest (bounded) endpoint $\max(D)$ of D. Therefore the set L of simplices on which f takes value larger than $\max(D)$ is removable. □

Next we will exhibit a family of simplicial complexes that are essential. We say that K is a *triangulated (oriented) manifold* if its geometric realisation $|K|$ is homeomorphic to a closed (orientable) manifold. Note that this manifold will necessarily be compact since K is finite.

Proposition 4.6.3. *Let K be a triangulated manifold. If either the field of coefficients $\Bbbk$ is of characteristic 2 or K is oriented, then K is essential.*

Proof. Without loss of generality, we may assume that K is connected and of dimension d. Then any filter f attains its maximum value on a top dimensional simplex, since all lower-dimensional simplices have co-faces. By our assumptions, we have $\mathrm{H}_d(K) \cong \Bbbk$ and a generator of this top dimensional homology class is the sum of all top dimensional simplices of K (with appropriate signs). At the level of barcodes, this means that there is an infinite interval in homological degree d starting at $\max_{\sigma \in K} f(\sigma)$. Therefore the fiber of the persistence map over any barcode D is bounded, and since it is closed by continuity of PH, it is also compact and hence essential by the previous result. □

The converse of Proposition 4.6.3 is false as can be seen from the following simple counterexample.

Example 4.6.4. Let K be the wedge product of two triangles. So K has five vertices and six 1-simplices. Its first homology group is of rank 2

but any subcomplex will have at most rank 1. Thus K is essential, but K is not a manifold.

The point-set topology is a little delicate when working with the unbounded real line $\mathbb{R}$ instead of the compact interval I. This is part of the reason why we chose to work with I for the main part of our paper. For example, Proposition 4.2.17 cannot be adapted to the unbounded situation: If the fibers of PH are not compact then the bottleneck topology and the quotient topology induced by PH do not necessarily agree on the image $\mathbf{Bar}_K(\mathbb{R})$ as Example 4.6.5 below shows.

Example 4.6.5. Let us consider again the Example 4.2.19 of the complex K representing the unit interval with vertices a, b and 1-simplex σ. As a set $\mathbf{Bar}_K(\mathbb{R})$ can be identified (as in Example 4.2.19) with the 3-simplex $\Delta_3' = \{(x_1, x_2, x_3), -\infty < x_1 \leqslant x_2 \leqslant x_3 < \infty\}$ with two missing faces, where in addition the 2-dimensional face corresponding to $-\infty < x_1 < x_2 = x_3 < \infty$ is collapsed to the line segment $-\infty < x_1 = x_2 = x_3 < \infty$. Consider the set

$$U := \left\{(x_1, x_2, x_3) \,|\, x_3 - x_2 < e^{-x_2}\right\} \subseteq \mathbf{Bar}_K(\mathbb{R})$$

of barcodes whose unique bounded bar (x_2, x_3) has length less than e^{-x_2}. We then have

$$\mathrm{PH}^{-1}(U) := \left\{f \in \mathrm{Filt}_K(\mathbb{R}) \,|\, f(\sigma) - \max(f(a), f(b)) < e^{-\max(f(a),f(b))}\right\} \subseteq \mathrm{Filt}_K(\mathbb{R}) \subseteq \mathbb{R}^3,$$

which is open in $\mathrm{Filt}_K(\mathbb{R})$ for the usual topology induced by the $\|.\|_\infty$-metric. Therefore U is an open set in the quotient topology. However it does not contain any bottleneck ball, hence is not an open set in the bottleneck topology.

Furthermore, the choice of topology on $\mathrm{End}(\mathbb{R}, \leqslant)$ and $\mathrm{Aut}(\mathbb{R}, \leqslant)$ matters in the unbounded situation: Replacing the L^∞ topology by the compact open topology results in the actions not being (sequentially) continuous as can be seen in the following example.

Example 4.6.6. Consider the sequence of barcodes D_n containing a single interval $(n, n + 2^{-n})$. The sequence D_n converges to the empty diagram $D_\varnothing$ in the bottleneck topology. In addition, let $\phi_n : \mathbb{R} \to \mathbb{R}$ be the map such that $\phi_n(n) = n$, $\phi_n(n + 2^{-n}) = n + 1$, $\phi_n|_{[n;n+2^{-n}]}$ and $\phi_n|_{[n+2^{-n};n+2]}$ are linear, and outside $[n; n + 2]$ ϕ_n is the identity. Then the sequence ϕ_n converges to the identity map of the real line in the compact open topology. If the action were continuous in both variables, the sequence $\phi_n(D_n)$ would converge to the empty diagram $\mathrm{Id}(D_\varnothing) = D_\varnothing$. However, each of the barcodes $\phi_n(D_n)$ contains a unique interval $(n, n + 1)$, and the sequence does therefore not converge in the bottleneck topology.

However, in our analysis, we have never needed to make full use of the continuity of the action of $\mathrm{End}(\mathrm{I}, \leqslant)$. Instead, it is enough to ensure that the action is continuous w.r.t. the choice of ϕ. Namely, fixing $D \in \mathbf{Bar}$, the map $\phi \in \mathrm{End}(\mathrm{I}, \leqslant) \mapsto \phi(D) \in \mathbf{Bar}$ is continuous. In the current unbounded situation, it can also be proven that the map $\phi \in \mathrm{End}(\mathbb{R}, \leqslant) \mapsto \phi(D) \in \mathbf{Bar}(\mathbb{R})$ is continuous, where we consider the compact open topology on $\mathrm{End}(\mathbb{R}, \leqslant)$. This is precisely what is needed to carry the analysis through in a similar fashion.

Finally, if we do not impose that maps in $\mathrm{End}(\mathbb{R}, \leqslant)$ and $\mathrm{Aut}(\mathbb{R}, \leqslant)$ equal the identity outside a compact set, then the analysis breaks down in the L^∞ topology. For instance, straight line interpolations on which our results rely, would not always give continuous paths.

4.6.2 *The case of lower star filters*

The lower star filtration form an interesting subspace of the space of all filters on K and one might want to restrict one's attention to these as for example in [CMW20]. We summarise briefly how our analysis can be adapted and compared to this case.

Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020

Let K be a finite simplicial complex with vertex set V. A *lower star filter* on K is a filter $f \in \mathrm{Filt}_K$ such that for any simplex $\sigma \in K$:

$$f(\sigma) = \max_{v \in \sigma} f(v).$$

Being determined by their values on vertices, such filters offer many advantages in practice. Any function $f : V \to \mathrm{I}$ can be extended uniquely to a lower star filter. Hence, the subspace $\mathrm{Low}_K \subseteq \mathrm{Filt}_K$ of lower star filters is canonically isomorphic to I^{V}. We denote its image under the persistence map PH by $\mathbf{Bar}_K^{\mathrm{Low}}$.

The actions of $\mathrm{Aut}(\mathrm{I}, \leqslant)$ and $\mathrm{End}(\mathrm{I}, \leqslant)$ by post-composition restrict to I^{V} and, by the equivariance of PH, also to $\mathbf{Bar}_K^{\mathrm{Low}}$. As the strata are given by $\mathrm{Aut}(\mathrm{I}, \leqslant)$-orbits, we see that both Low_K and $\mathbf{Bar}_K^{\mathrm{Low}}$ are sub-stratified spaces, each consisting of a subcollection of full strata from Filt_K and $\mathbf{Bar}_K$ respectively. Thus PH restricts to a strongly stratified map

$$\mathrm{PH}_{|\mathrm{Low}} : \mathrm{Low}_K \longrightarrow \mathbf{Bar}_K^{\mathrm{Low}}$$

and hence satisfies similar properties as PH. In particular, the fiber $\mathrm{PH}_{\mathrm{Low}}^{-1}(D)$ has again the structure of a polyhedral complex and the analogue of Theorem 4.3.2 holds.

The space $\mathbf{Bar}_K^{\mathrm{Low}}$ also gives rise to a subcategory $\mathbf{Bar_K}^{\mathrm{Low}}$ of $\mathbf{Bar_K}$. We note that this is a full subcategory. Thus Theorem 4.4.2 and Proposition 4.4.12 also hold for this subcategory. In particular, $\mathbf{Bar_K}^{\mathrm{Low}}$ is homotopy discrete. As before, we can associate to morphisms in $\mathbf{Bar_K}^{\mathrm{Low}}$ monodromies between fibers, turning the inverse image into a functor

$$\mathrm{PH}_{|\mathrm{Low}}^{-1} : \mathbf{Bar_K}^{\mathrm{Low}} \longrightarrow \mathbf{Top}.$$

Up to homotopy, the monodromies between fibers are again polyhedral maps.

Remark 4.6.7. Vice-versa, we may also consider the space of filters on K as a subspace of the space of lower star filters on its barycentric subdivision $\hat{K}$. Recall that the vertices of $\hat{K}$ are the simplices of K. Thus $f \in \mathrm{Filt}_K$ uniquely gives rise to $\hat{f} \in \mathrm{Low}_{\hat{K}}$ via

$$\hat{f}(\hat{\sigma}) := f(\sigma),$$

where $\hat{\sigma}$ is the vertex of $\hat{K}$ corresponding to the simplex σ in K. This way, we get a nested sequence of spaces:

$$\mathrm{Low}_K \subset \mathrm{Filt}_K \subset \mathrm{Low}_{\hat{K}}.$$

It is a straightforward exercise to show that $\mathrm{PH}(f) = \mathrm{PH}(\hat{f})$. Thus we also have a nested sequence of barcode spaces:

$$\mathbf{Bar}_K^{\mathrm{Low}} \subset \mathbf{Bar}_K \subset \mathbf{Bar}_{\hat{K}}^{\mathrm{Low}}.$$

All these inclusions are also $\mathrm{End}(\mathrm{I}, \leqslant)$-equivariant and PH defines an equivariant map between these nested sequences. Thus, by similar arguments as before, PH and the inclusions are compatible with the stratifications in the strongest sense giving rise to a sequence of full, homotopy discrete subcategories

$$\mathbf{Bar_K}^{\mathrm{Low}} \subset \mathbf{Bar_K} \subset \mathbf{Bar}_{\hat{\mathbf{K}}}{}^{\mathrm{Low}}.$$

It would be interesting to analyse how the fibers of PH, or more generally the fiber functors on these three categories are related.

4.6.3 Symmetries restricted to fibers

In this brief section we examine how symmetries of the simplicial complex restrict to the fibers of the persistence map. For simplicity we return to filters and barcodes in the unit interval I, but the analysis can be carried out in the unbounded situation or when restricting to lower star filters in the same way.

Let $\mathbb{G}(K)$ be the group of isomorphisms of the simplicial complex K. Then $\mathbb{G}(K)$ can be identified as the subgroup of the group of symmetries $\mathrm{Sym}(K_0)$ of the vertices K_0 which consists of all those s that map a subset $\sigma \in K$ to a subset $s(\sigma) \in K$. Pre-composition with the inverse induces a left action of $\mathbb{G}(K)$ on the space Filt_K of filters via

$$s.f := f \circ s^{-1}.$$

Thus f and $s.f$ take the same values and furthermore, if $\mathcal{S} \subseteq \mathrm{Filt}_K$ is a filter stratum then $s.\mathcal{S}$ is another stratum of the same dimension, $\dim s.\mathcal{S} = \dim \mathcal{S}$, and s maps $\mathcal{S}$ to $s(\mathcal{S})$ via an affine isomorphism.

Proposition 4.6.8. *For all $f \in \mathrm{Filt}_K$ and all $s \in \mathbb{G}(K)$ we have*

$$\mathrm{PH}(f) = \mathrm{PH}(s.f).$$

Equivalently, the action of $\mathbb{G}(K)$ on Filt_K restricts to the fiber $\mathrm{PH}^{-1}(D)$ for every $D \in \mathbf{Bar}_K$. Furthermore, $\mathbb{G}(K)$ acts through maps of polyhedra on $\mathrm{PH}^{-1}(D)$.

Proof. The symmetry s maps the sublevel-set filtration of f isomorphically to that of $s.f$. Thus the associated persistence modules are isomorphic and so the two resulting barcodes in $\mathbf{Bar}_K$ are the same. Hence, f and $s.f$ are in the same fiber. Since s defines an affine isomorphism from a stratum $\mathcal{S}$ to the stratum $s.\mathcal{S}$ and preserves fibers, it restricts to an affine isomorphism from $\mathrm{PH}^{-1}(D) \cap \mathcal{S}$ to $\mathrm{PH}^{-1}(D) \cap s(\mathcal{S})$ for any barcode $D \in \mathbf{Bar}_K$. □

Proposition 4.6.9. *Given two barcodes D, D' and a morphism $\phi \in \mathbf{Bar}_\mathbf{K}(D, D')$, the monodromy $\mathcal{L}_\phi : \mathrm{PH}^{-1}(D) \to \mathrm{PH}^{-1}(D')$ is $\mathbb{G}(K)$-equivariant.*

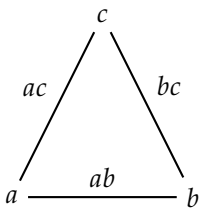
Proof. Let $f \in \mathrm{PH}^{-1}(D)$, $s \in \mathbb{G}(K)$ and $\sigma \in K$. Then the statement of the proposition follows from

$$\mathcal{L}_\phi(s.f)(\sigma) = (\phi \circ s.f)(\sigma) = \phi(f(s^{-1}(\sigma))) = \mathcal{L}_\phi(f)(s^{-1}(\sigma)) = s.\mathcal{L}_\phi(f)(\sigma).$$

□

4.7 *A detailed example: the fiber of the persistence map over the triangle*

In this section we illustrate our theory developed so far on an example. The simplest non-contractible simplicial complex K is a triangle. We denote its vertices by a, b, c and its edges by ab, ac, bc.



In this simple case, a filter is any map $f : K \to \mathrm{I}$ such that its value on an edge is greater (or equal) than the value on the endpoints of this edge. By the elder rule and since K is connected, $\min(f(a), f(b), f(c))$ is the left endpoint of the unique infinite interval in $\mathrm{PH}_0(f)$. For similar reasons, $\mathrm{PH}_1(f)$ contains a unique unbounded interval with left endpoint given by $\max(f(ab), f(bc), f(ac))$.

In addition, the example of the triangle has interesting symmetries. The symmetries of the triangle $\mathbb{G}(K) = \mathbf{D}_3$ is the dihedral group which

can also be identified with the symmetric group Σ_3, the set of bijections of the set $\{a, b, c\}$ of vertices of K.

Our goal in this section is to determine all the barcode strata and the corresponding fibers, see Figure 4.7; describe the action of $\mathbb{G}(K) \simeq \mathbf{D}_3$ on the fibers, and compute the monodromies between the non-discrete fibers.

Summary of the results:

1. In section 4.7.1 we compute all 34 barcode strata in the image $\mathbf{Bar}_K = \mathrm{PH}(\mathrm{Filt}_K)$;

2. In section 4.7.2 we compute the fibers of PH over the distinct barcode strata. We find only five strata with fibers that are not discrete. By Theorem 4.3.2, these fibers are polyhedral complexes, and we exhibit the polyhedra in $\mathbb{R}^K$ making up the complexes. The results in section 4.6.3 guarantee that $\mathbb{G}(K)$ acts on each fiber, and we describe this action on the fibers for these five barcode strata;

3. Finally, by the results in section 4.4.2, the closure containment relations between barcode strata yield monodromies between fibers, which up to homotopy are polyhedral maps. We describe the monodromies between non-discrete fibers in section 4.7.3.

This simple example of the triangle shows that the fibers of the persistence map can be topologically distinct from each other. Furthermore, the topology of the fibers can also be more complex than that of the underlying simplicial complex K, especially for low dimensional strata in the space of barcodes. Similar observations hold when restricting the fibers to the subspace Low_K of lower star filters, as we detail in section 4.7.4. We contrast this with the case studied in [CMW20] where K is a triangulation of the interval $[0, 1]$ and lower star filters are considered. In that case the fibers of the persistence map are all disjoint unions of contractible sets. In particular, our example of the triangle shows that the fiber of the persistence map does not have to be a union of contractible sets.

Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020

4.7.1 *Computation of* $\mathbf{Bar}_K$ *and its strata*

For notational convenience, we replace the unit interval with a bigger interval, $\mathrm{I} := [-10; 10]$. This allows considering barcodes with only integer valued endpoints. In addition, we restrict ourselves to strata of barcodes with endpoints strictly in $(-10, 10)$, since they completely determine strata (and their fibers) where the endpoints -10 and 10 are allowed.

The top dimensional filter strata of the space of filters Filt_K correspond to injective filters and can equivalently be thought of as orderings

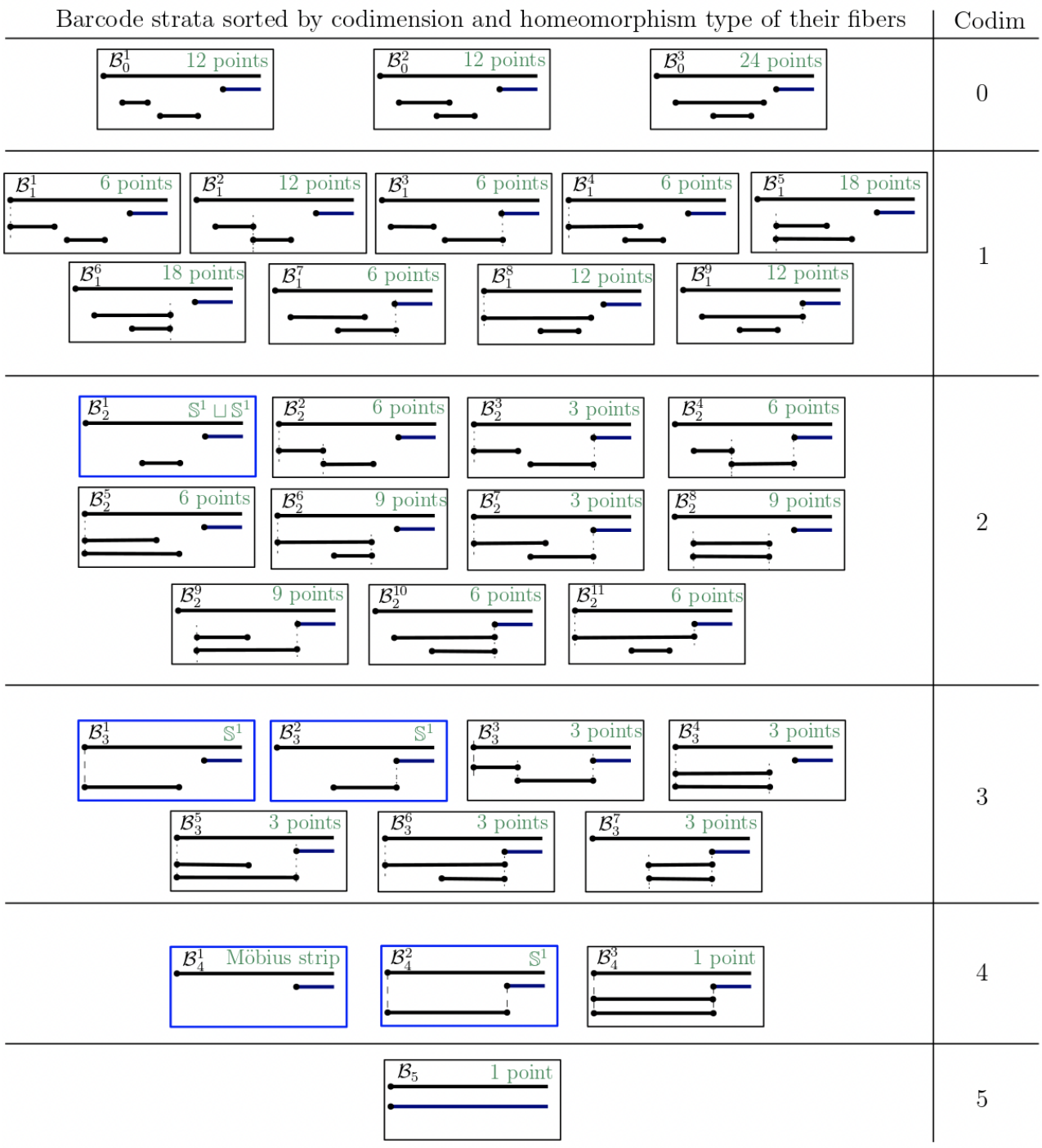


Figure 4.7: The list of all barcode strata in $\mathbf{Bar}_K$ and the associated fibers when K is the simplicial triangle.

of the simplices in the triangle, where an edge must appear after its vertices. This allows us to count these top dimensional strata. Namely, considering the case where the vertices of the triangle appear before all the edges, and separately the case where one edge appears before the last vertex, we get the following count of possible orderings and hence

$$36 + 12 = 48 \text{ top dimensional filter strata.}$$

The top dimensional barcode strata in the image $\mathbf{Bar}_K = \mathrm{PH}(\mathrm{Filt}_K)$, by Proposition 4.2.18, are given by the image of the top dimensional filter strata. We argue that there are precisely three: From section 4.6.3, the persistence map is $\mathbb{G}(K)$-equivariant, so we may restrict ourselves to those filter strata, viewed as orderings, for which the vertex a is first, $f(a) = 0$, followed by b, $f(b) = 1$. The following simplex must either be ab or c. In the first case, $f(ab) = 2$, all filters satisfying this yield the same barcode $[\{(0,\infty),(1,2),(3,4)\},\{(5,\infty)\}]$. We denote by $\mathcal{B}_0^1$ the corresponding codimension 0 barcode stratum. In the second case, $f(c) = 2$, we get two other barcodes depending on whether ab is the next simplex in the ordering or not, $[\{(0,\infty),(1,3),(2,4)\},\{(5,\infty)\}]$ and $[\{(0,\infty),(1,4),(2,3)\},\{(5,\infty)\}]$. We denote the corresponding barcode strata by $\mathcal{B}_0^2$ and $\mathcal{B}_0^3$.

The list of all barcode strata in $\mathbf{Bar}_K$, by Proposition 4.2.18, can be derived from the three top dimensional barcode strata $\mathcal{B}_0^1$, $\mathcal{B}_0^2$ and $\mathcal{B}_0^3$ by collapsing their interval endpoints. We draw all the 34 resulting strata in Figure 4.7, sorted by codimension, and give them labels that are used in the rest of the section.

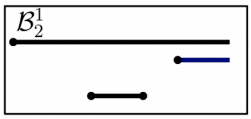
4.7.2 *Computation of fibers and the action of* $\mathbb{G}(K)$

The fibers over the various barcode strata in the image of the persistence map, whose computations are detailed in this section, are summarised in Figure 4.7 (in green in the top right corner of each box). More precisely, each stratum is represented by one of its barcodes, where the blue interval is the one corresponding to the degree 1 homology of the triangle and the others correspond to degree 0 homology. Strata come with a label, for instance $\mathcal{B}_0^2$, whose subscript gives the codimension and the superscript allows to enumerate strata of a given codimension. The homeomorphism type of the fiber of PH over each stratum is given in green. For instance, the fiber $\mathrm{PH}^{-1}(\mathcal{B}_1^2)$ is finite and consists of 12 distinct filters. The blue boxes highlight the five strata with non-discrete fibers.

Strata with discrete fibers. Most of the barcode strata have finite fibers. For instance, the unique lowest-dimensional stratum is the one labelled by $\mathcal{B}_5$: there only the two essential homological features are present,

and appear at the same time. This implies that all the simplices of the triangle must appear at a given time, and therefore the fiber of PH over $\mathcal{B}_5$ consists of a unique constant function. This agrees with the prediction of Proposition 4.3.7 that the fiber is contractible.

The barcode strata with maximal number of bounded intervals, that is 2 such intervals, have zero bounded deficit (Def. 4.2.2). There are 28 such strata. By Proposition 4.3.5, their fibers are discrete. For instance, a simple counting argument gives discrete fibers for the top dimensional barcode strata $\mathcal{B}_0^1$, $\mathcal{B}_0^2$ and $\mathcal{B}_0^3$, with 12, 12, and 24 points in their fibers respectively. More generally, Proposition 4.3.5 upper-bounds the dimension of the fibers over arbitrary strata by their bounded deficit. In this example, note that the bounded deficit in fact equals the dimension of the fiber in all cases, except in the degenerate case of the stratum $\mathcal{B}_5$. In the remainder of this section we compute explicitly the fibers over the five barcode strata that have fibers of dimension greater than 0.



Stratum $\mathcal{B}_2^1$. We take a representative barcode $D := [\{(b_1, \infty), (b_2, d_2)\}, \{(b_3, \infty)\}] \in \mathbf{Bar}^2$, where $b_1 < b_2 < d_2 < b_3$, of the codimension 2 stratum $\mathcal{B}_2^1$.

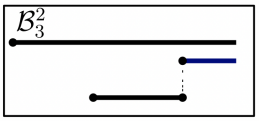
Without loss of generality, $(b_1, b_2, d_2, b_3) = (0, 1, 2, 3)$. We use the identification of a filter f with the vector $(f(a), f(b), f(c), f(ab), f(ac), f(bc)) \in \mathrm{I}^6$. If f yields barcode D, then:

- There is a vertex v (resp. an edge e) at which f attains its minimum (resp. maximum) value, equal to 0 (resp. 3).
- There is another vertex v' and edge e' such that $f(v'), f(e') = 1, 2$.
- The remaining vertex and **incident** edge have same value $0 \leqslant t \leqslant 3$.

Let us assume that the vertex a should create the first connected component, while the vertex c should create the second one. Then the vertex b must appear at the same time as an edge connecting b to either a or c. If b appears at the same time as ab, and if we wish that bc is the edge closing the loop, we get a 1-simplex $\{(0, t, 1, t, 2, 3)\}_{0 \leqslant t \leqslant 3}$ in the fiber, which we break into three 1-simplices that lie in (the closures of) different filter strata: $\{(0, t, 1, t, 2, 3)\}_{0 \leqslant t \leqslant 1}$, $\{(0, t, 1, t, 2, 3)\}_{1 \leqslant t \leqslant 2}$ and $\{(0, t, 1, t, 2, 3)\}_{2 \leqslant t \leqslant 3}$. If rather the edge ac closes the loop of the triangle, we get the two sets $\{(0, t, 1, t, 3, 2)\}_{1 \leqslant t \leqslant 2}$ and $\{(0, t, 1, t, 3, 2)\}_{0 \leqslant t \leqslant 1}$ in the fiber. Now, let us assume that b appears at the same time as bc, which imposes that b appears after time 1. Then if ab closes the loop of the triangle, we get the two sets $\{(0, t, 1, 3, 2, t)\}_{1 \leqslant t \leqslant 2}$ and $\{(0, t, 1, 3, 2, t)\}_{2 \leqslant t \leqslant 3}$ in the fiber, while if we wish that the edge ac closes this loop, we get the unique set $\{(0, t, 1, 2, 3, t)\}_{1 \leqslant t \leqslant 2}$ in the fiber. All in all, we have gathered 8 embeddings of the standard 1-simplex in I^6. By symmetry, if we vary the choice of two vertices that create the first two connected

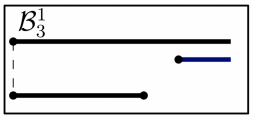
components, we get 6 analogous collections of 8 embeddings of the standard 1-simplex described above, which together cover the fiber. This provides a description of the fiber in terms of a graph, whose incidence structure is explicited in Fig. 4.10. In particular, the fiber of PH over D is homeomorphic to $\mathbb{S}^1 \sqcup \mathbb{S}^1$.

We further depict the action of $\mathbb{G}(K)$ on $\mathrm{PH}^{-1}(D)$ in Fig. 4.10. We consider the cyclic map $(a,b,c) \mapsto (b,c,a)$ and the elementary transposition $(a,b,c) \mapsto (b,a,c)$ as generators. We see that the cyclic map preserves the two connected components of the fiber $\mathbb{S}^1 \sqcup \mathbb{S}^1$, whereas the elementary transposition swaps them. More generally, even permutations $g \in \mathbb{G}(K)$ preserve the connected components of the fiber, while odd permutations exchange them.



Stratum $\mathcal{B}_3^2$. Fixing two endpoints via $d_2 = b_3$ in the stratum $\mathcal{B}_2^1$, we get the stratum $\mathcal{B}_3^2$ represented below.

As a representative of this stratum, we consider the barcode $D = [\{(0,\infty),(1,2)\},\{(2,\infty)\}] \in \mathbf{Bar}^2$. The fiber of D can be obtained similarly as in the previous case, that is by providing a cover of the fiber by embeddings of the 1-simplex in I^6. Imposing that vertices a and then c should be responsible for the appearance of the first two components, we get the sets $\{(0,t,1,t,2,2)\}_{0\leqslant t\leqslant 1}$, $\{(0,t,1,t,2,2)\}_{1\leqslant t\leqslant 2}$ and $\{(0,t,1,2,2,t)\}_{1\leqslant t\leqslant 2}$ in the fiber. By symmetry in the choice of these two vertices, we get a total of 6×3 embeddings of the standard 1-simplex that together cover the fiber. This decomposition describes the fiber as a graph, which is described in Fig 4.11 and is homeomorphic to $\mathbb{S}^1$. The action of the symmetries on the circle is described in the figure as well.

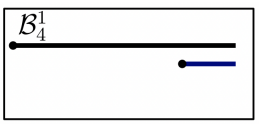


Stratum $\mathcal{B}_3^1$. A representative of the stratum $\mathcal{B}_3^1$ is $D := [\{(0,\infty),(0,1)\},\{(2,\infty)\}] \in \mathbf{Bar}^2$.

We describe the fiber with an explicit cover by embeddings of the 1-simplex in I^6. Imposing that vertices a and c should be responsible for the appearance of the first two components, we get the following embeddings:

- $\{(0,t,0,1,2,t)\}_{0\leqslant t\leqslant 1}$, $\{(0,t,0,2,1,t)\}_{0\leqslant t\leqslant 1}$, $\{(0,t,0,t,1,2)\}_{0\leqslant t\leqslant 1}$, $\{(0,t,0,t,2,1)\}_{0\leqslant t\leqslant 1}$, $\{(0,t,0,t,1,2)\}_{1\leqslant t\leqslant 2}$ and $\{(0,t,0,2,1,t)\}_{1\leqslant t\leqslant 2}$.

By symmetry in the choice of these two vertices, we get a total of 3×6 embeddings of the standard 1-simplex that together cover the fiber. By inspecting adjacency relations, this decomposition describes the fiber as a graph isomorphic to the graph of the fiber over $\mathcal{B}_3^2$. In particular, the fiber is homeomorphic to $\mathbb{S}^1$.



Stratum $\mathcal{B}_4^1$. Let $D = [\{(b_1,\infty)\},\{(b_2,\infty)\}] \in \mathbf{Bar}^2$ (where $b_1 < b_2$) be a barcode in the stratum $\mathcal{B}_4^1$.

We may set $b_1 = 0$ and $b_2 = 1$ for simplicity. If a filter $f : K \to \mathrm{I}$ yields the barcode D, then there are two pairs $(v_i, e_i)_{i=1,2}$ of vertices and edges such that $0 \leqslant f(e_i) = f(v_i) =: t_i \leqslant 1$ and the remaining vertex v_0 and edge e_0 realize the minimum 0 and maximum 1 of f respectively.

For a fixed choice of $(v_i, e_i)_{i=0,1,2}$ as above, the two parameters $t_1 \leqslant t_2$ describe a 2-simplex. This 2-simplex corresponds to one of the top dimensional polyhedron in the fiber of PH over D, which we know is a polyhedral complex from Theorem 4.3.2. To count and describe these 2-simplices, fix v_0, say $v_0 = a$. Then we distinguish between the two cases: (i) where e_0 contains v_0 and (ii) where e_0 does not contain v_0. In case (i), if we choose $e_0 = ac$, then b and ab simultaneously appear at time t_1 and finally c and bc appear at time t_2. The resulting simplex in the fiber is denoted by $S := \{(0, t_1, t_2, t_1, 1, t_2)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$. In case (ii), there is only one choice for e_0, i.e. $e_0 = bc$. The other vertex-edge pairs have to be the pairs (b, ab) and (c, ac). If we decide that the pair (b, ab) appears before (c, ac), we obtain the simplex $\bar{S} := \{(0, t_1, t_2, t_1, t_2, 1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$. All the other simplices in the fiber may be derived from the action of $\mathfrak{G}(K)$ on S and $\bar{S}$. More precisely, letting $\tau : (a, b, c) \mapsto (b, a, c)$ be the elementary transposition and $c : (a, b, c) \mapsto (b, c, a)$ the cyclic permutation, we obtain all the 2-simplices in the fiber:

(i) $S = \{(0, t_1, t_2, t_1, 1, t_2)\}$, $\tau.S = \{(t_1, 0, t_2, t_1, t_2, 1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $c.S = \{(t_2, 0, t_1, 1, t_2, t_1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $c^2.S = \{(t_1, t_2, 0, t_2, t_1, 1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $\tau c.S = \{(0, t_2, t_1, 1, t_1, t_2)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $\tau c^2.S = \{(t_2, t_1, 0, t_2, 1, t_1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$;

(ii) $\bar{S} = \{(0, t_1, t_2, t_1, t_2, 1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $\tau.\bar{S} = \{(t_1, 0, t_2, t_1, 1, t_2)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $c.\bar{S} = \{(t_2, 0, t_1, t_2, 1, t_1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $c^2.\bar{S} = \{(t_1, t_2, 0, 1, t_1, t_2)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $\tau c.\bar{S} = \{(0, t_2, t_1, t_2, t_1, 1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$, $\tau c^2.\bar{S} = \{(t_2, t_1, 0, 1, t_2, t_1)\}_{0 \leqslant t_1 \leqslant t_2 \leqslant 1}$.

All the 12 sets are embeddings of the 2-simplex Δ^2 in I^6, where we represent the simplex Δ^2 in $\mathbb{R}^2$ conveniently for our purpose as in Fig 4.8.

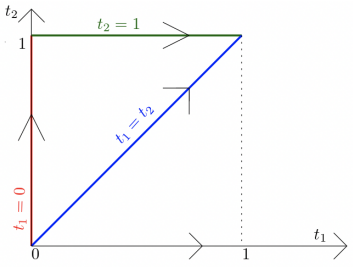
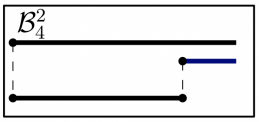


Figure 4.8: An embedding of the standard simplex Δ^2, with faces in red, green and blue.

The 2-simplices in the orbit of $\bar{S}$ meet with 3 distinct other 2-simplices at its 3 faces (obtained by setting $t_1 = 0$, $t_2 = 1$ or $t_1 = t_2$). The 2-simplices in the orbit of S only meet with two other simplices. Glued together, these simplices form a Möbius band embedded in $\mathbb{R}^6$, as explicited in Figure 4.12. Therefore the fiber of the persistence map over the barcodes in the stratum $\mathcal{B}_4^1$ is isomorphic, as a simplicial complex, to the Möbius band.

We again consider how the symmetry group Σ_3 of the triangle acts on the fiber. Note that the action on the fiber must preserve the orientations and colors described in Fig. 4.12. The elementary transposition simply rotates the Möbius band by an angle of π. The action of the cyclic permutation is slightly more involved as it reverses and translates the Möbius band.



Stratum $\mathcal{B}_4^2$. The last stratum of barcodes whose fiber we explicitely compute has representative $D := [\{(0,\infty)\},\{(0,1),(1,\infty)\}]$:

In this case, the fiber is the union of the segments:

- $\{(0,t,0,t,1,1)\}_{0\leqslant t\leqslant 1}$, $\{(0,t,0,1,1,t)\}_{0\leqslant t\leqslant 1}$, $\{(0,0,t,1,1,t)\}_{0\leqslant t\leqslant 1}$, $\{(0,0,t,1,t,1)\}_{0\leqslant t\leqslant 1}$, $\{(t,0,0,1,t,1)\}_{0\leqslant t\leqslant 1}$, $\{(t,0,0,t,1,1)\}_{0\leqslant t\leqslant 1}$;

which assemble into a regular hexagon, and therefore the fiber of PH over the stratum $\mathcal{B}_4^2$ is homeomorphic to $\mathbb{S}^1$.

4.7.3 *Computation of monodromies between fibers*

We describe the monodromies between non-discrete fibers. We focus on pairs of barcode strata that differ by one dimension, since a monodromy between an arbitrary pair of strata is a composition of such elementary monodromies by Corollary 4.4.14.

Monodromy from $\mathcal{B}_2^1$ to $\mathcal{B}_3^1$ and $\mathcal{B}_3^2$. From the previous section, the fibers over $\mathcal{B}_3^1$ and $\mathcal{B}_3^2$ are isomorphic cyclic graphs. We only describe monodromies from $\mathrm{PH}^{-1}(\mathcal{B}_2^1)$ to $\mathrm{PH}^{-1}(\mathcal{B}_3^2)$ since monodromies from $\mathrm{PH}^{-1}(\mathcal{B}_2^1)$ to $\mathrm{PH}^{-1}(\mathcal{B}_3^1)$ are identical. The stratum $\mathcal{B}_2^1$ contains the stratum $\mathcal{B}_3^2$ in its closure. From Lemma 4.2.13, this ensures that the set of morphisms $\mathbf{Bar}_{\mathbf{K}}(D_2, D_3^2)$ is non-empty for any representatives $(D_2, D_3^2) \in \mathcal{B}_2^1 \times \mathcal{B}_3^2$. For clarity, we take as representatives the barcodes from the previous section, that is $D_2 = [\{(0,\infty),(1,2)\},\{(3,\infty)\}]$ and $D_3^2 = [\{(0,\infty),(1,2)\},\{(2,\infty)\}]$.

From section 4.4.2, the number of different homotopy classes of monodromies from $\mathrm{PH}^{-1}(D_2)$ to $\mathrm{PH}^{-1}(D_3^2)$ is upper-bounded by the cardinality of $\pi_0(\mathbf{Bar}_{\mathbf{K}}(D_2, D_3^2))$. Moreover, the homotopy type of an element $\phi \in \mathbf{Bar}_{\mathbf{K}}(D_2, D_3^2)$ is completely characterized by its index by Remark 4.4.5. In the current situation, any map ϕ from D_2 to D_3^2 must collapse the third and fourth endpoint of D_2. This means that up to homotopy, there is a unique monodromy map from $\mathrm{PH}^{-1}(D_2)$ to $\mathrm{PH}^{-1}(D_3^2)$ to describe. It is then natural to choose ϕ to be any linear extension of the map:

$$\phi : \{0,1,2,3\} \mapsto \{0,1,2,2\}.$$

The resulting monodromy $\mathcal{L}_\phi$, from Proposition 4.4.12, is a simplicial map between fibers. To visualize $\mathcal{L}_\phi$, it is convenient to see how the fiber $\mathrm{PH}^{-1}(D_2)$ is progressively deformed into a subset of $\mathrm{PH}^{-1}(D_3^2)$ under the path $t \mapsto t\mathcal{L}_\phi + (1-t)\mathrm{Id}$. As depicted in Fig. 4.9, this transformation has the effect to collapse some edges and to identify some edges and nodes.

In the same figure, we see that the monodromy is a surjection onto the fiber $\mathrm{PH}^{-1}(D_3^2)$. By Proposition 4.6.9, the equivariance of the

monodromy w.r.t. the $\mathbb{G}(K) = \Sigma_3$ action on the fibers predicts that the image, through the monodromy, of the action on $\mathrm{PH}^{-1}(D_2)$ must equal the action on $\mathrm{PH}^{-1}(D_3^2)$. Let us for instance consider the cyclic permutation $(a,b,c) \mapsto (b,c,a)$ of the triangle, which acts on the two hexagons constituting $\mathrm{PH}^{-1}(D_2)$ by rotation of an angle of $\frac{2\pi}{3}$. The monodromy identifies the two hexagons in the fiber, hence the induced action of g on $\mathrm{PH}^{-1}(D_3^2)$ is the rotation by the same angle, which agrees with the direct computation of the action of g on $\mathrm{PH}^{-1}(D_3^2)$ in Fig. 4.11. The same observation can be made about the elementary transposition $\tau : (a,b,c) \mapsto (b,a,c)$ of the triangle.

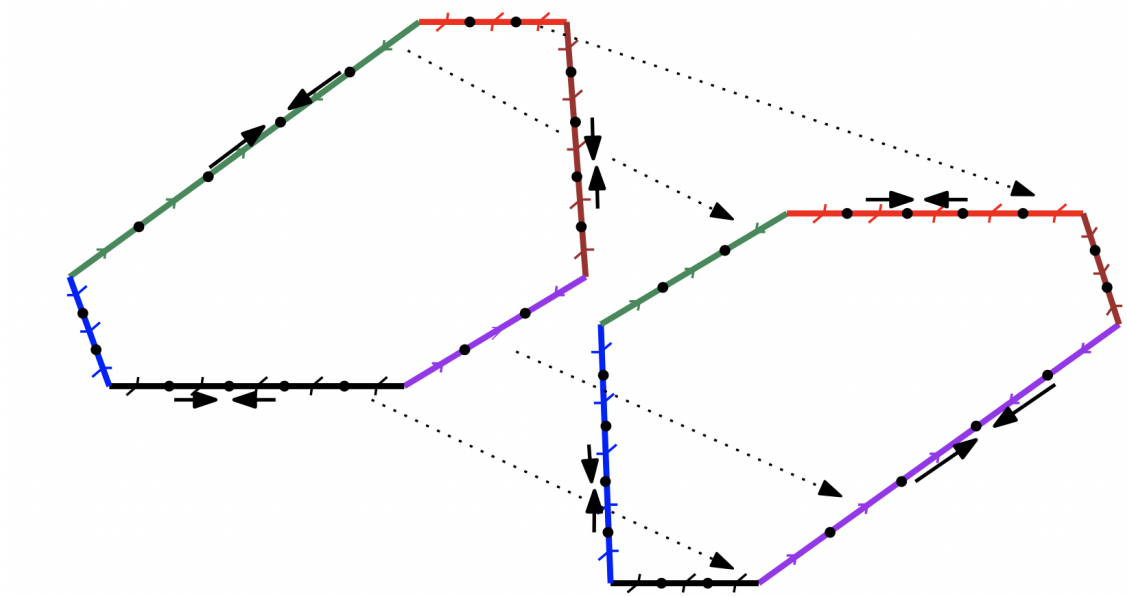


Figure 4.9: Starting with the two irregular hexagons describing the fiber of the barcode $[\{(0,\infty),(1,2)\},\{(3,\infty)\}]$ of stratum $\mathcal{B}_2^1$ (see Fig. 4.10), the arrows describe the process of continuously tracking the fiber as the interval $(3,\infty)$ gets closer to the interval $(2,\infty)$, thus ending to the barcode $[\{(0,\infty),(1,2)\},\{(2,\infty)\}]$ of stratum $\mathcal{B}_3^2$. Plain arrows show edges of the fiber that are collapsed during this process. Meanwhile, the two hexagons merge into the regular hexagon depicted in Fig. 4.11. This merging happens by identifying edges of the two components, of the same color and orientation, following the dotted arrows.

Monodromy from $\mathcal{B}_3^1$ and $\mathcal{B}_3^2$ to $\mathcal{B}_4^1$. The stratum $\mathcal{B}_3^2$ contains the stratum $\mathcal{B}_4^1$ in its closure. We let $D_3^2 := [\{(0,\infty),(1,2)\},\{(2,\infty)\}]$ be the representative barcode of the stratum $\mathcal{B}_3^2$, and $D_4^1 := [\{(0,\infty)\},\{(1,\infty)\}]$ be that of $\mathcal{B}_4^1$, both as in the previous section. The set of morphisms $\mathbf{Bar}_{\mathbf{K}}(D_3^2, D_4^1)$ is non-empty, and in fact contains a unique homotopy class by Remark 4.4.5, since all maps $\phi \in \mathbf{Bar}_{\mathbf{K}}(D_3^2, D_4^1)$ must collapse the interval $(1,2)$. It is then natural to choose ϕ to be any linear extension of the map:

$$\phi : \{0,1,2\} \mapsto \{0,1,1\}.$$

From Proposition 4.4.12, the resulting monodromy $\mathcal{L}_\phi : \mathrm{PH}^{-1}(D_3^2) \to \mathrm{PH}^{-1}(D_4^1)$ is a map of polyhedral complexes. Recall that the fiber $\mathrm{PH}^{-1}(D_3^2)$ is a polyhedral complex described in Fig. 4.11. The monodromy map $\mathcal{L}_\phi$ collapses 12 out of the 18 edges in $\mathrm{PH}^{-1}(D_3^2)$. The remaining 6 edges form a regular hexagon which is mapped onto the green circle of the Möbius strip describing the fiber of $\mathcal{B}_4^1$, see Fig 4.12. Likewise, the monodromy from the fiber of (a representative of) the stratum $\mathcal{B}_3^1$ to the fiber over $\mathcal{B}_4^1$ is unique up to homotopy, and collapses 12 out of the 18 edges in the fiber over $\mathcal{B}_3^1$, sending the remaining 6 edges onto the red circle of the Möbius strip describing the fiber of $\mathcal{B}_4^1$.

Monodromy from $\mathcal{B}_3^1$ and $\mathcal{B}_3^2$ to $\mathcal{B}_4^2$. Recall that the fiber over $\mathcal{B}_3^2$ is a cyclic graph with 18 edges depicted in Figure 4.11, while the fiber over $\mathcal{B}_4^2$ is a regular hexagon. There is again a unique (up to homotopy) monodromy $\mathcal{L}_\phi$ between these fibers. The simplicial map $\mathcal{L}_\phi$ collapses 12 out of the 18 edges in $\mathrm{PH}^{-1}(D_3^2)$. The remaining 6 edges (one for each color in Figure 4.11) form the regular hexagon $\mathrm{PH}^{-1}(\mathcal{B}_4^2)$. In particular, $\mathcal{L}_\phi$ is a fibration. The monodromy from $\mathrm{PH}^{-1}(D_3^1)$ to $\mathrm{PH}^{-1}(\mathcal{B}_4^2)$ can be described in the same way.

4.7.4 *Lower star filters*

If we consider the restriction of PH to the subspace $\mathrm{Low}_K \subseteq \mathrm{Filt}_K$ of lower star filters, there are only 2 barcode strata in the image: $\mathcal{B}_4^1$ and $\mathcal{B}_5$. The reason for this is that edges enter the sublevel set filtration of a lower star filter at the same time as one of their vertices, hence there are no bounded intervals in the resulting barcode. The fibers of the restricted persistence map over the two strata can be derived from the general case. Namely, the fiber over $\mathcal{B}_5$ is the constant filter, while the fiber over $\mathcal{B}_4^1$ is a hexagon which embeds into the Möbius strip as the green zig-zag in Fig. 4.12.

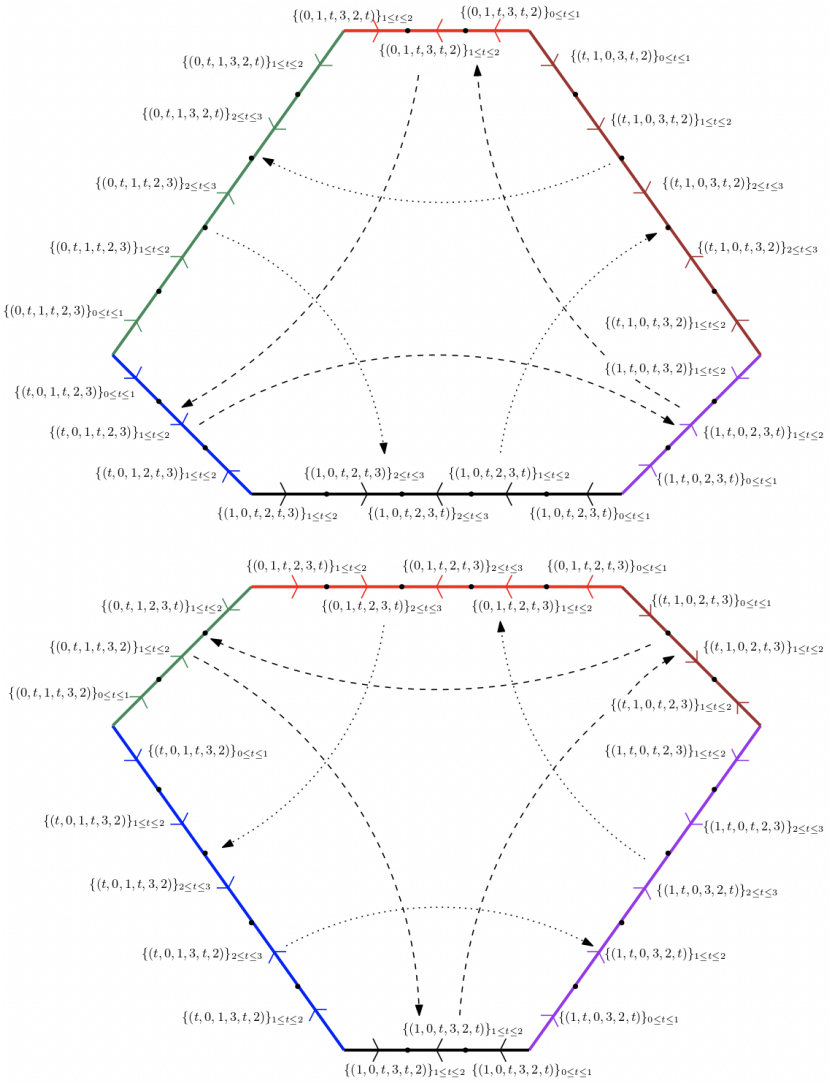


Figure 4.10: The fiber of the barcode in stratum $\mathcal{B}_2^1$. We observe two connected components. Each edge represents an embedding of the standard 1-simplex in the fiber, oriented with an arrow toward increasing values of the parameter t. The six colors correspond to the six possible choices of two vertices responsible for the appearance of the first two connected components. If a and c (resp. a and b, b and a, b and c, c and a, c and b) create these components, we color the edge in green (resp. red, black, blue, purple and brown). The cyclic symmetry $(a,b,c) \mapsto (b,c,a)$ acts on the fiber by rotating each connected component by an angle of $\frac{2\pi}{3}$, as depicted by the dotted arrows. The transposition $(a,b,c) \mapsto (b,a,c)$ acts by swapping the connected components by reflecting along the horizontal axis.

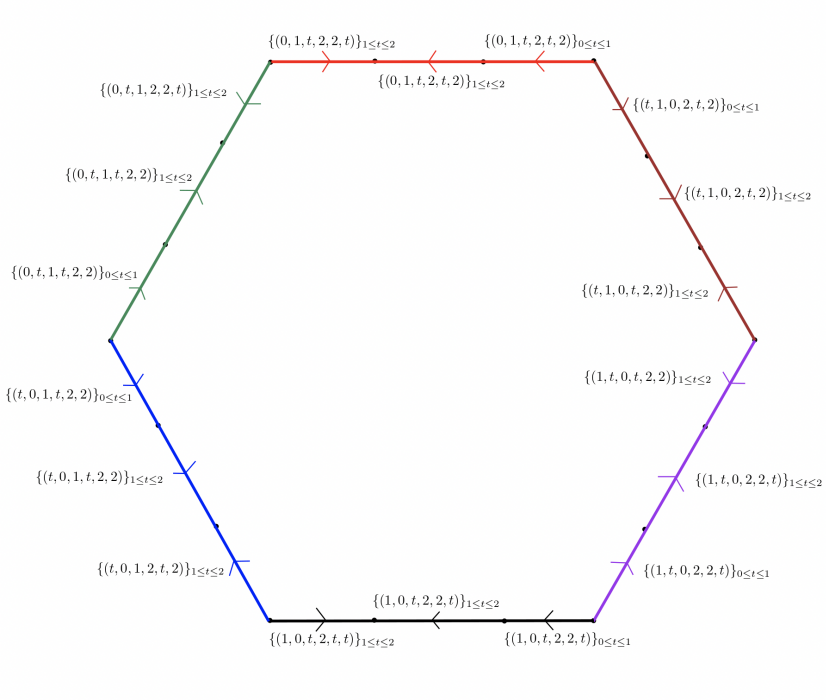


Figure 4.11: The fiber of the barcode in stratum $\mathcal{B}_3^2$. Each edge represents an embedding of the standard 1-simplex in the fiber, oriented toward increasing values of the parameter t. The six colors correspond to the six possible choices of two vertices responsible for the appearance of the first two connected components in the barcode. The coloring convention is the same as in case $\mathcal{B}_2^1$, Fig 4.10. The cyclic permutation of the triangle acts as a rotation by an angle of $\frac{2\pi}{3}$ oriented counter-clockwise, while the elementary transposition acts as the horizontal symmetry of the hexagon.

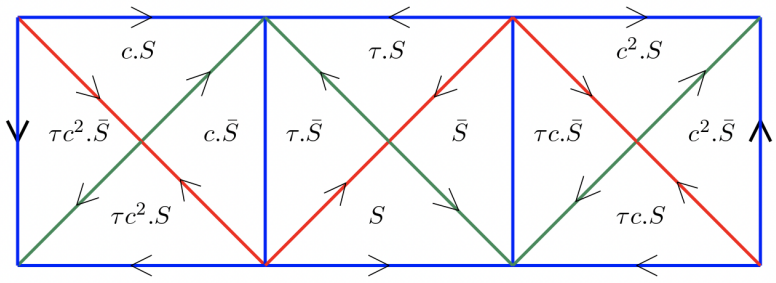


Figure 4.12: The 2-simplices in the fiber of the stratum $\mathcal{B}_4^1$ glued together in a Möbius strip. Blue (resp. red and green) edges correspond to setting $t_1 = t_2$ (resp. $t_1 = 0$ and $t_2 = 1$) in their cofaces. The left and right extreme blue oriented edges are identified. The action of the transposition $(a, b, c) \mapsto (b, a, c)$ on the fiber can be described as a rotation of the Möbius strip by an angle of π. The action of the cyclic permutation $(a, b, c) \mapsto (b, c, a)$ on the fiber can be described as the composition of (i) the symmetry of the Möbius strip around its middle horizontal line, followed by (ii) a unit translation on the left of each simplex.

5
Algorithmic Reconstruction of the Fiber of Persistent Homology on Cell Complexes

Abstract

Let K be a finite simplicial, cubical, delta or CW complex. The persistence map PH takes a filter $f : K \to \mathbb{R}$ as input and returns the barcodes $\mathrm{PH}(f)$ of the associated sublevel set persistent homology modules. We address the inverse problem: given a target barcode D, computing the fiber $\mathrm{PH}^{-1}(D)$. For this, we use the fact that $\mathrm{PH}^{-1}(D)$ decomposes as complex of polyhedra when K is a simplicial complex, and we generalise this result to arbitrary based chain complexes. We then design and implement a depth first search algorithm that recovers the polyhedra forming the fiber $\mathrm{PH}^{-1}(D)$. As an application, we solve a corpus of 120 sample problems, providing a first insight into the statistical structure of these fibers, for general CW complexes.

5.1 Introduction

Persistent Homology (PH) is a computable descriptor[1] for data science problems where topology is prominent, e.g. for analysing graphs and simplicial complexes K. Given a function $f : K \to \mathbb{R}$, one assembles the homology groups of the sub complexes $f^{-1}((-\infty, t])$ into a module which is faithfully represented by a so-called barcode $D = \mathrm{PH}(f)$. Using vectorisation methods[2] the topological information of the barcode can then be used in statistical studies and machine learning problems.

To gain an a priori understanding of the problems where PH is applied, it is key to know about the invariance of PH, i.e. the context in which it is a discriminating descriptor. Hence the general inverse problem: what are the different functions f giving rise to the same barcode $D = \mathrm{PH}(f)$? Equivalently we are interested in the properties of the fiber $\mathrm{PH}^{-1}(D)$ over a target barcode D. Prior work[3] has shown the fiber to be the geometric realization of a polyhedral complex; each polyhedron represents the restriction of the fiber to strata of the space of filters. Similarly the space of barcodes is given a stratification, and $\mathrm{PH}^{-1}(D)$ is piecewise-affinely isomorphic to $\mathrm{PH}^{-1}(D')$ for any D, D' that belong to the same barcode stratum. Furthermore, if D' lies in the closure of the stratum containing D, then there is a natural map $\mathrm{PH}^{-1}(D) \to \mathrm{PH}^{-1}(D')$ that respect the polyhedral structure of $\mathrm{PH}^{-1}(D)$ and $\mathrm{PH}^{-1}(D')$.

However, understanding the fiber $\mathrm{PH}^{-1}(D)$ remains a challenge, in general. If the underlying simplicial complex K is a subdivision of the unit interval or of the circle, then the fiber $\mathrm{PH}^{-1}(D)$ is a disjoint union of contractible sets[4] and circles[5], respectively. Outside these and a few other 1-dimensional examples, however, little is known. Even establishing that $\mathrm{PH}^{-1}(D) \neq \varnothing$ can be difficult, as we shall see in section 5.7.[6]

Algorithm development represents an important step toward addressing this challenge. Computerized calculations offer a range of new examples (intractable by hand) with which to test hypotheses and search for patterns, thus contributing to the growth of theory. They also represent an essential prerequisite to many scientific applications.

In this paper we propose an algorithmic approach to compute the fiber of the persistence map PH, for an arbitrary finite simplicial complex K, over an arbitrary barcode D. We then generalize this approach to include finite CW complexes.

Outline of contributions After introducing elementary material from Persistence theory (section 5.2), we define in section 5.3 the key data structures used in our method. The algorithm, presented in section 5.4, computes the list of polyhedra in the fiber $\mathrm{PH}^{-1}(D)$. For this, we

[1] Herbert Edelsbrunner and John Harer. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008; and Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005

[2] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research*, 18(1):218–252, 2017; and Peter Bubenik. Statistical topological data analysis using persistence landscapes. *The Journal of Machine Learning Research*, 16(1):77–102, 2015

[3] Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021

[4] Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020

[5] Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021

[6] For certain types of barcode D, we show that $\mathrm{PH}^{-1}(D) = \varnothing$ iff K is collapsible.

adopt an inductive approach that constructs filtrations of K simplex by simplex, and simultaneously, an associated barcode via the standard reduction algorithm for computing persistent homology. Through this incremental construction, we ensure that the filtrations remain compatible with the target barcode D. The resulting collection of polyhedra is the polyhedral complex that characterises the homeomorphism type of $\mathrm{PH}^{-1}(D)$.

That the fiber $\mathrm{PH}^{-1}(D)$ is a polyhedral complex is generalised to arbitrary based chain complexes in section 5.5, in particular including CW complexes, delta complexes, cubical complexes, on which we can take arbitrary filter functions, or the sub spaces of lower-star and lower-edge filter functions. In turn our algorithm adapts to these situations.

In a section 5.6 dedicated to experiments, we apply the algorithm to multiple complexes and barcodes, and report statistics about the fiber $\mathrm{PH}^{-1}(D)$, such as its number of polyhedra and its Betti numbers. Sometimes unexpectedly, most of the properties observed for the 1-dimensional complexes studied in [CMW20, LT21, MW21][7] do not hold for more general complexes. For instance, already for a triangulated 2-sphere the fiber $\mathrm{PH}^{-1}(D)$ over some barcodes D has non-trivial homology in degree 3, unlike all the existing examples of graphs, for which $\mathrm{PH}^{-1}(D)$ has vanishing homology in dimension greater than 1. We also find CW complexes, for example the real projective plane, that are topological manifolds whose fibers are not necessarily manifolds.

By contrast, the algorithm allows us to observe novel trends that hold consistently across examples: for instance to each simplicial complex K we associate a specific barcode D_K such that the fiber $\mathrm{PH}^{-1}(D_K)$ and K have the same Betti numbers (see conjecture 5.6.1). Our code is publicly available at [JHP22][8].

Related works Finally we note that for related inverse problems algorithmic approaches to reconstruct the fiber have been designed. For instance instead of taking a single function the Persistent Homology Transform (PHT) of [TMB14b][9] computes the barcodes of a family of functions over a fixed shape in $\mathbb{R}^3$, which is enough to completely characterise this shape, see [CMT18, GLM18][10] for generalisations to higher dimensions. These sorts of results motivated the design of algorithms to reconstruct the complex K from the associated family of barcodes.[11]

5.2 *Background*

We fix a finite simplicial complex K of dimension $\dim K$ and with $\sharp K$ many simplices. Throughout, (simplicial) homology is taken with coefficients in a fixed field $\Bbbk$. A *filter* over K is a map $f : K \to \mathrm{I}$ valued in the interval $\mathrm{I} := [1; \sharp K] \subset \mathbb{R}$, whose sublevel sets are subcomplexes of K.

[7] Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020; Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021; and Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021

[8] Leygonie Jacob and Gregory Henselman-Petrusek. Software Companion to Algorithmic Reconstruction of the Fiber of Persistent Homology on Cell Complexes. https://github.com/Eetion/phfibre, 2022

[9] Katharine Turner, Sayan Mukherjee, and Doug M Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014

[10] Justin Curry, Sayan Mukherjee, and Katharine Turner. How many directions determine a shape and other sufficiency results for two topological transforms. *arXiv preprint arXiv:1805.09782*, 2018; and Robert Ghrist, Rachel Levanger, and Huy Mai. Persistent homology and euler integral transforms. *Journal of Applied and Computational Topology*, 2(1-2):55–60, 2018

[11] Leo M Betthauser. *Topological Reconstruction of Grayscale Images*. PhD thesis, University of Florida, 2018; and Brittany Terese Fasy, Samuel Micka, David L Millman, Anna Schenfisch, and Lucia Williams. Persistence diagrams for efficient simplicial complex reconstruction. *arXiv preprint arXiv:1912.12759*, 2019

If we regard K as a poset partially ordered by face inclusions $\sigma \subseteq \sigma'$, then f can be regarded, equivalently, as an order-preserving function into I. The set of filters is a polyhedron contained in the Euclidean space $\mathbb{R}^K$.

For each homology degree $0 \leqslant p \leqslant \dim K$, we have a (finite) persistent homology module arising from the sub-level sets of f. The isomorphism type of each module of this form is uniquely determined by the associated barcode $\mathrm{PH}_p(f)$; concretely, the barcode is a finite multi-set of pairs (b, d), called *intervals*, each of which characterizes the appearance and annihilation of a class of p-cycles in the filtration. For further details, see [ZC05, EH10a]. We denote the set of all barcodes by **Bar**.

The persistence map PH takes a filter f and returns the $\dim K + 1$ barcodes of interest:

$$\mathrm{PH} : f \in \mathrm{Filt}(K) \longmapsto \big(\mathrm{PH}_0(f), \mathrm{PH}_1(f), \cdots, \mathrm{PH}_{\dim K}(f)\big) \in \mathbf{Bar}^{\dim K+1}.$$

Our goal is to compute the fiber $\mathrm{PH}^{-1}(D)$ over some $(d+1)$-tuple of barcodes $D = (D_0, D_1, \cdots, D_{\dim K}) \in \mathbf{Bar}^{\dim K+1}$, from now on called barcode for simplicity.

There is another, equivalent formulation of D which is sometimes better suited to formal arguments. In this approach, we replace the sequence of multisets D with a single, bona fide set X. We regard X as a set of indices x, with one index for each interval in the barcode, and write $\mathrm{birth}(x)$, $\mathrm{death}(x)$, and $\dim(x)$ for the birth time, death time, and dimension, respectively, of the corresponding interval. Thus D_p is the multiset $\{(\mathrm{birth}(x), \mathrm{death}(x)) : x \in X,\ \dim(x) = p\}$. The set X is especially useful for algorithms, e.g. when writing for-loops, however it is nonstandard as a convention.

By way of compromise, we will identify D with X. When we refer to fixing $(b, d) \in D$, we mean fixing $x \in X$ such that $\mathrm{birth}(x) = b$ and $\mathrm{death}(x) = d$. By $\dim(b, d)$, we then mean $\dim(x)$.

We also abuse notation and write $(b, d) \in D$ whenever $(b, d) \in D_p$ for some $0 \leqslant p \leqslant \dim K$.

Given a barcode D, let $\mathrm{End}(\mathrm{I}, \leqslant)(D)$ denote the set of all endpoints of intervals in D,

$$\{b : (b, d) \in D \text{ for some } d\} \cup \{d : (b, d) \in D \text{ for some } b\} \subseteq \mathbb{R} \sqcup \{+\infty\}$$

Set $\mathrm{End}(\mathrm{I}, \leqslant)^*(D) = \mathrm{End}(\mathrm{I}, \leqslant)(D) \backslash \{\infty\}$ and write $\dim D = |\mathrm{End}(\mathrm{I}, \leqslant)^*(D)|$ for the number of finite endpoints. In particular, $\dim D \leqslant \sharp K$. By normalizing, we may assume without loss of generality that $\mathrm{End}(\mathrm{I}, \leqslant)^*(D)$ has form $\{1, \ldots, \dim D\}$. For instance, the barcode D with 4 intervals (in possibly different homology degrees) $(1, 2)$, $(1, 3)$, $(2, 3)$ and $(2, +\infty)$ has dimension 3, since the endpoints values form the set $\{1, 2, 3\}$.

Fix the barcode D. For any pair of filters $f, g : K \to \mathrm{I}$, let us define the relation $f \sim g$ by either of the following two equivalent axioms:

(A1) There exists a non-decreasing map $\psi : \mathrm{I} \to \mathrm{I}$ such that $\psi(i) = i$ for each $i \in \{1, \dots, \dim D\}$, and $g = \psi \circ f$.

(A2) The function g satisfies

$$f(e) \in \mathrm{End}(\mathrm{I}, \leqslant)^*(D) \implies g(e) = f(e), \quad f(e) \leqslant f(e') \implies g(e) \leqslant g(e'). \tag{5.1}$$

Note that $\sim$ is reflexive and transitive but not symmetric. Given a filter f we write $\eta^i := f(E) \cap (i, i+1)$ and

$$\mathcal{P}(f) := \{g \in \mathrm{I}^K : \ f \sim g\} \subseteq \mathrm{I}^K.$$

Theorems 5.2.1 and 5.2.2 below are proved in [LT21, Theorem 2.2][12]. Recall that a finite set $\mathcal{P}$ of polyhedra in $\mathbb{R}^K$ is a *polyhedral complex* if for each polyhedron $X \in \mathcal{P}$, the faces of X belong to $\mathcal{P}$ as well, and if, furthermore, any two polyhedra $X, X' \in \mathcal{P}$ intersect at a common face. The underlying space of $\mathcal{P}$ is $|\mathcal{P}| = \bigcup_{X \in \mathcal{P}} X$.

[12] Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021

Theorem 5.2.1. *Let $f \in \mathrm{PH}^{-1}(D)$ be a filter in the fiber. Then $\mathcal{P}(f) \subseteq \mathrm{PH}^{-1}(D)$, and $\mathcal{P}(f)$ is a polyhedron which is affinely isomorphic to the product $\Delta_{\#\eta^1} \times \cdots \times \Delta_{\#\eta^{\dim D}}$, where Δ_k stands for the standard geometric simplex of dimension k.*

The polyhedra $\mathcal{P}(f)$ then assemble into a polyhedral complex with underlying space the fiber $\mathrm{PH}^{-1}(D)$:

Theorem 5.2.2. *The set $\{\mathcal{P}(f) \mid f \in \mathrm{PH}^{-1}(D)\}$ is a polyhedral complex.*

In section 5.5 we extend Theorems 5.2.1 and 5.2.2 to CW complexes and more generally to based chain complexes.

5.3 Data structures

5.3.1 Classifications of simplices compatible with a barcode

Let us fix a finite simplicial complex K and a barcode D.

An important part of our strategy will be to introduce the concept of D-compatible factorization. This object allows one to define clean correspondences between simplices in K and interval endpoints. This, in turn, allows one to define equivalence classes of simplices according to whether or not they contribute to the barcode D – a major benefit, in terms of overall organization.

Let $f \in \mathrm{PH}^{-1}(D)$ be given. Then $\mathrm{End}(\mathrm{I}, \leqslant)^*(D) = \{1, \cdots, \dim D\} \subseteq \mathrm{Im}(f)$. Note that containment may be proper because f can take non-integer values, in general. However, we claim $\{1, \cdots, \dim D\} = \mathrm{Im}(f)$

when f is injective; in this case simplices appear "one at a time" in the corresponding filtration, and a simple counting argument allows one to infer that $\dim D = \#K = \#\mathrm{Im}(f)$ as a result. We make use of this property in the general case.

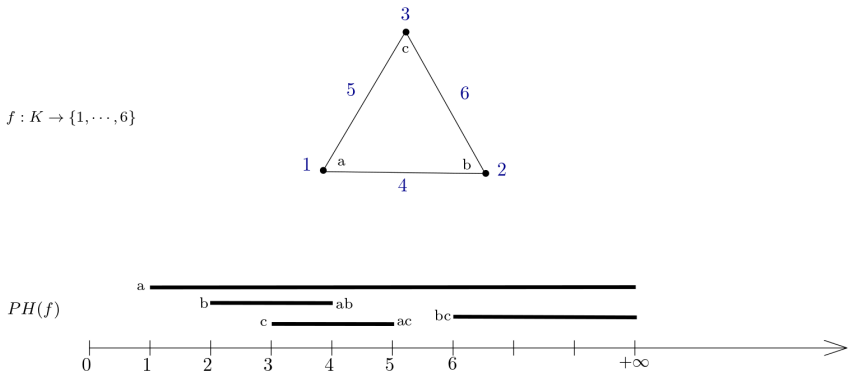


Figure 5.1: Given an injective filter f, the finite endpoints in the barcode $\mathrm{PH}(f)$ are canonically identified with simplices in K. Here, K is the triangle with vertices denoted by a, b and c. The values (in dark blue) of the bijection $f : K \to \{1, \cdots, 6\}$ are displayed along the simplices of K. No two distinct intervals in the resulting barcode (all homology degrees included) share a common end value, and these endpoints are annotated with the corresponding simplex

Definition 5.3.1. Let f be an injective filter valued in $\{1, \cdots, \sharp K\}$. The *associated total order* is the unique total order on simplices $\sigma_1 < \cdots < \sigma_{\sharp K}$ such that $f(\sigma_k) = k$ for each k. If there exists a non-decreasing continuous map $\phi : \mathbb{R} \mapsto \mathrm{I}$ such that $\mathrm{PH}(\phi \circ f) = D$, we say that (f, ϕ) is a *D-compatible factorization*.

Remark 5.3.2. Given a non-decreasing continuous map ϕ and an arbitrary barcode $D' = [D'_0, \ldots, D'_{\sharp K}]$, let $\phi(D')$ be the barcode such that $\phi(D')_p$ contains one interval of form $(\phi(b'), \phi(d'))$ for each interval (b', d') in D'_p that satisfies $\phi(b') \neq \phi(d')$. Here, by convention, we define $\phi(\infty) = \infty$. Then the persistence map is equivariant with respect to this action, in the sense that $\mathrm{PH}(\phi \circ f) = \phi(\mathrm{PH}(f))$; see for instance [LT21, Lemma 1.5][13]. This gives another perspective on the idea of D-compatibility: The barcode $\mathrm{PH}(f)$ has endpoints in bijection with the simplices of K, and ϕ 'contracts' this barcode into D. See Fig. 5.2 for illustration.

[13] Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021

We now describe how a D-compatible factorization (f, ϕ) allows one to disambiguate the homological roles of each simplex. Let $\sigma \in K$ be a given. By injectivity of f, there exists a unique $\sigma' \in K$ such that $(f(\sigma), f(\sigma'))$ or $(f(\sigma'), f(\sigma))$ lies in the barcode $\mathrm{PH}(f)$. We write $[\sigma, \sigma']_f$ for this pairing; where context leaves no room for confusion, we simply write $[\sigma, \sigma']$. There is also the case where σ is unpaired, i.e. when it corresponds to an infinite interval $(f(\sigma), \infty)$. We then write $[\sigma, \infty]$. Then (f, ϕ) gives rise to the following classification of pairs of simplices $[\sigma, \sigma']$ (with the possibility that $\sigma' = \infty$) according to whether or not $(\phi \circ f(\sigma), \phi \circ f(\sigma'))$ is an interval in the barcode D:

Critical $(\phi \circ f(\sigma), \phi \circ f(\sigma')) = (b, d)$ for some non-trivial interval $(b, d) \in D$; We say that $[\sigma, \sigma']$ is (f, ϕ)-*critical* for (b, d) and record this association via $\pi_+(b, d) := \sigma$ and $\pi_-(b, d) := \sigma'$;

Non-Critical Otherwise ϕ cancels the interval $(f(\sigma), f(\sigma'))$, i.e. $\phi \circ f(\sigma) = \phi \circ f(\sigma')$; We say that $[\sigma, \sigma']$ is (f, ϕ)-*non-critical*.

In particular we have a well-defined bijection of intervals in D with critical pairs:

$$\pi : D \to K \times (K \sqcup \{\infty\}), \quad (b, d) \longmapsto [\pi_+(b, d), \pi_-(b, d)].$$

By abuse of terminology we say that a simplex σ is *critical* if it is part of a critical pair, and that it is *non-critical* otherwise. Note that an unpaired simplex is necessarily critical for some $(b, \infty) \in D$. In Fig. 5.2, we provide an example of D-compatible (f, ϕ) and the resulting critical and non-critical simplices.

There are uncountably infinitely many D-compatible factorizations. In the two following definitions we compress the necessary information into finitely many equivalence classes of D-compatible factorizations.

Definition 5.3.3. A *classification* $\mathcal{K}$ of simplices of K relative to D, or *classification* for short, consists of:

(i) A bijective filter $f : K \to \{1, \cdots, \sharp K\}$, or equivalently a total ordering $\sigma_1 < \cdots < \sigma_{\sharp K}$;

(ii) A partition of simplices into consecutive ordered sets $\Phi_1, \cdots, \Phi_m$, which we refer to as *classes*:

$$\underbrace{\sigma_1 < \sigma_2 < \cdots < \sigma_{\sharp \Phi_1}}_{\Phi_1} < \underbrace{\cdots < \cdots < \cdots}_{\Phi_2, \cdots, \Phi_{m-1}} < \underbrace{\sigma_{(\sharp K - \sharp \Phi_m + 1)} < \cdots < \sigma_{\sharp K}}_{\Phi_m};$$

(iii) For each interval endpoint j, a class Φ_{i_j}, with $j \longmapsto i_j$ a non-decreasing assignment.

Definition 5.3.4. A D-compatible classification $\mathcal{K}$ is a classification induced by a D-compatible factorization (f, ϕ) as follows:

(i) $f : K \to \{1, \cdots, \sharp K\}$ equals the bijective filter of $\mathcal{K}$ inducing the order $\sigma_1 < \cdots < \sigma_{\sharp K}$;

(ii) The image of $\phi \circ f$ has cardinality m, i.e. we can write $\mathrm{Im}(\phi \circ f) = \{t_1 < \cdots < t_m\}$, and the associated sequence of pre-images equal the classes of $\mathcal{K}$:

$$(\phi \circ f)^{-1}(t_1) = \Phi_1, \ \cdots, \ (\phi \circ f)^{-1}(t_m) = \Phi_m.$$

(iii) For each interval endpoint $j \in \mathrm{End}(\mathrm{I}, \leqslant)^*(D)$, we have $(\phi \circ f)^{-1}(j) = \Phi_{i_j}$.

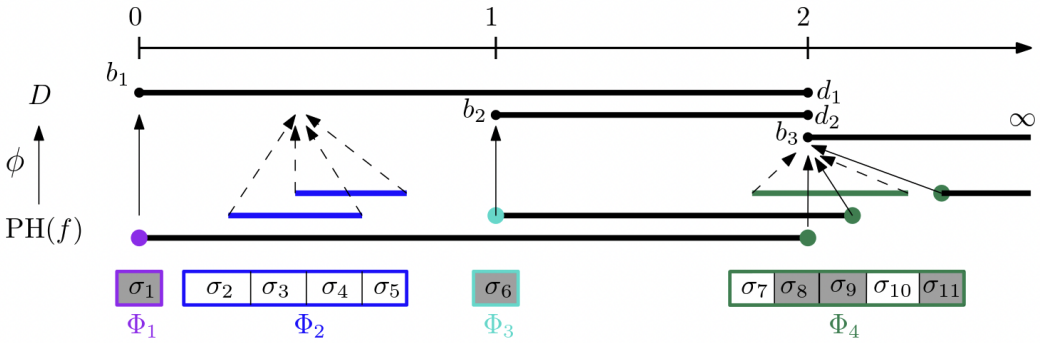


Figure 5.2: With a simplicial complex K and an injective filter f in the background, ϕ maps the barcode PH(f) to D via the arrows going up, so that (f,ϕ) is D-compatible. In the corresponding D-compatible classification, there are 4 distinct equivalence classes Φ_i of simplices, and they are distinguished by the use of different colors. Critical simplices are grayed and correspond to intervals via the upward solid arrows: Here $\pi(b_1,d_1) = [\sigma_1,\sigma_8]$, $\pi(b_2,d_2) = [\sigma_6,\sigma_9]$ and $\pi(b_3,\infty) = [\sigma_{11},\infty]$, hence the critical simplices form the set $\{\sigma_1,\sigma_6,\sigma_8,\sigma_9,\sigma_{11}\}$. The other simplices are non-critical and cancelled in pairs by the dotted arrows: $\{\sigma_2,\sigma_3,\sigma_4,\sigma_5,\sigma_7,\sigma_{10}\}$.

We depict all this data in Fig 5.2. Note that given a classification $\mathcal{K}$ the filter f induces pairings $[\sigma,\sigma']$.

Proposition 5.3.5. *A classification $\mathcal{K}$ is the D-compatible classification induced by some D-compatible factorization if and only if the following two rules are satisfied:*

Critical rule: For any $1 \leqslant j < j' \leqslant \dim D$ and $0 \leqslant p \leqslant \sharp K$, there are as many pairs $[\sigma_j,\sigma_{j'}]_f$ with $\sigma_j \in \Phi_{i_j}$ and $\sigma_{j'} \in \Phi_{i_{j'}}$, here $\dim\sigma_j = p$, as there are copies of the interval (j,j') in D_p;

Non-critical rule: For any remaining pair $[\sigma,\sigma']$, both σ and σ' belong to the same class Φ_i.

Proof. Indeed, when these two rules are satisfied, we can define ϕ directly on the classes $\Phi_1,\cdots,\Phi_m$ as any order-preserving map sending $f(\Phi_{i_j})$ to j. □

It is natural to group D-compatible factorizations (f,ϕ) according to the classification $\mathcal{K}$ they induce: it is clear that whenever factorizations induce the same classification, then they induce the same pairs $[\sigma,\sigma]$ of simplices, the same critical and non-critical pairs (and simplices) and the same bijection π from intervals in D to critical pairs. Therefore these concepts are defined as well given a D-compatible classification $\mathcal{K}$.

5.3.2 *Relations to the fiber*

We explain how to retrieve the polyhedra that compose the fiber $\mathrm{PH}^{-1}(D)$ (see Theorem 5.2.2) from the set of D-compatible classifications $\mathcal{K}$. Note that if two D-compatible factorizations (f,ϕ) and (f',ϕ') induce the same classification $\mathcal{K}$, then by Axiom **(A2)**, they determine the same polyhedron, $\mathcal{P}(\phi\circ f) = \mathcal{P}(\phi'\circ f')$. Thus the following definition is unambiguous:

$$\mathcal{P}(\mathcal{K}) := \mathcal{P}(\phi\circ f).$$

Moreover, $\mathcal{P}(\mathcal{K}) \subseteq \mathrm{PH}^{-1}(D)$ by Theorem 5.2.1, because (f, ϕ) is D-compatible. Conversely, a filter $g \in \mathrm{PH}^{-1}(D)$ can always be written as $g = \phi \circ f$ for some injective filter f and non-decreasing map $\phi : \mathrm{I} \to \mathrm{I}$, so that $\mathcal{P}(g) = \mathcal{P}(\phi \circ f)$. Therefore we have the following result:

Theorem 5.3.6. *The set* $\{\mathcal{P}(\mathcal{K}) \mid \mathcal{K}$ *is a* D*-compatible classification*$\}$ *is the polyhedral complex underlying* $\mathrm{PH}^{-1}(D)$.

The upshot is that it is enough to compute the D-compatible classifications $\mathcal{K}$ in order to cover the fiber $\mathrm{PH}^{-1}(D)$.

5.4 *Algorithm*

We now propose an algorithm to retrieve the fiber $\mathrm{PH}^{-1}(D)$, i.e. that computes all the D-compatible classifications $\mathcal{K}$ of the previous section. Our implementation is simple in spirit: Algorithm 8 builds these classifications from scratch, simplex by simplex, and tries all the possible ways to do so. Along the way the algorithm manipulates *partial* D-compatible classifications, which can be thought of as the result of cutting a D-compatible classification $\mathcal{K}$ at a given step (see Fig. 5.3).

Given an integer $1 \leqslant j \leqslant \dim D + 1$, let $D^{<j} = (D_0^{<j}, \cdots, D_{\dim K}^{<j})$ be the truncated version of D such that

$$D_p^{<j} := \left\{(b,d) \in D_p : d < j\right\} \cup \left\{(b,\infty) \in D_p : (b,d) \in D_p, \;\; b < j \leqslant d\right\}.$$

for each $0 \leqslant p \leqslant \dim K$.

Definition 5.4.1. A *partial classification* is a tuple

$$\mathcal{F} = \left(\mathcal{K}_{\mathrm{in}}, j, \Phi_{\mathrm{cur}}, \mathcal{I}_{\mathrm{cur}}, \Phi_{\mathrm{cur}}^{\mathrm{n.c.}}\right)$$

where $\mathcal{K}_{\mathrm{in}}$ is a $D^{<j}$-compatible classification, and the set Φ_{cur} is a linearly ordered subset of $K \backslash \mathcal{K}_{\mathrm{in}}$ called the *current class*. We say that the current class Φ_{cur} is *incomplete* under either of the following two non-exclusive conditions:

(**Critical rule violation**) $\mathcal{I}_{\mathrm{cur}} \neq \varnothing$; and (**Non-critical rule violation**) $\Phi_{\mathrm{cur}}^{\mathrm{n.c.}} \neq \varnothing$.

Otherwise Φ_{cur} is *complete*. The (**Critical rule violation**) occurs when the current class Φ_{cur} is a truncated version of the critical class Φ_{i_j} containing the simplices critical for the interval endpoint j, and there are still intervals (b, j) or (j, d) that are unpaired with simplices of K. These missing intervals are stored in $\mathcal{I}_{\mathrm{cur}}$. Meanwhile, the (**Non-critical rule violation**) indicates that some non-critical birth simplices $\tau \in \Phi_{\mathrm{cur}}$ are not yet paired with a death non-critical simplex, i.e. τ creates a $\dim \tau$-cycle which must be destroyed in the same class. These unpaired non-critical simplices are stored in $\Phi_{\mathrm{cur}}^{\mathrm{n.c.}}$.

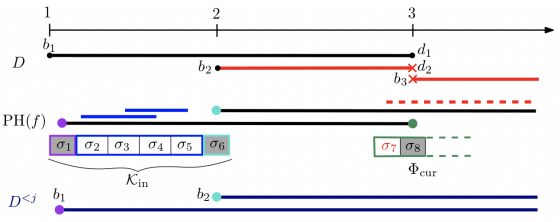


Figure 5.3: We consider the barcode D (top) of Fig. 5.2 and cut the D-compatible classification at σ_8 to get a partial classification. The first three classes Φ_1, Φ_2, Φ_3 are fully present in the resulting partial classification $\mathcal{F}$. So here $\mathcal{K}_{\text{in}}$ is a $D^{<j}$-compatible classification where $D^{<j}$ (bottom) is the barcode spanned by the interval endpoints 1 and 2. The next endpoint value to consider is then $j = 3$. The current class $\Phi_{\text{cur}} = \{\sigma_7, \sigma_8\}$ is incomplete for two reasons. On the one hand **Critical rule** is violated: The current class Φ_{cur} is intended to contain critical simplices for the endpoint $j = 3$, as already it contains the critical simplex $\sigma_8 = \pi_-(b_1, d_1)$, but the intervals (b_2, d_2) and (b_3, ∞) (in red) also have the endpoint $j = 3$ and are yet unmatched: so $\mathcal{I}_{\text{cur}} = \{(b_2, d_2), (b_3, \infty)\}$. Hence we should define the simplex $\pi_-(b_2, d_2)$ that destroys the cycle generated by $\sigma_6 = \pi_+(b_2, d_2)$, and a simplex $\pi_-(b_3, \infty)$ that creates a $\dim(b_3, \infty)$-cycle, before completion of the class. On the other hand **Non-critical rule** is violated: here $\Phi^{\text{n.c.}}_{\text{cur}} = \{\sigma_7\}$ contains the unpaired simplex σ_7 which generates an unexpected (dotted red) interval, which shall be destroyed by another simplex before completion of the class.

Algorithm 8 builds all the partial classifications starting with the empty one:

$$\mathcal{F} = (\mathcal{K}_{\text{in}}, j, \Phi_{\text{cur}}, \mathcal{I}_{\text{cur}}, \Phi^{\text{n.c.}}_{\text{cur}}) := (\varnothing, 1, \varnothing, D^{<2}, \varnothing),$$

and records complete D-compatible classifications in a list **Results**. Note that the set $\mathcal{I}_{\text{cur}}$ of intervals is initialised with intervals starting at the first endpoint 1 of D, since it is necessary that the first simplex to enter the filtration will be critical for such an interval. To check that $\mathcal{K}$ is complete is done at line 1 of Algorithm 8, and means that all the simplices of K have entered the filtration and that the target barcode D has been reached.

If the partial classification $\mathcal{F}$ is not complete but the current class Φ_{cur} is complete, i.e. the algorithm checks that **Critical rule** and **Non-critical rule** are satisfied (line 3), then Φ_{cur} is added to the classification (line 4), and the algorithm prepares the next class to build according to the following alternative. Either the next class will be the class $\Phi_{\text{cur}} = \Phi_{i_j}$ that will contain all simplices critical for intervals that have the endpoint j (lines 6 to 9), or the next class will contain only non-critical simplices (line 10). In practice either we fill $\mathcal{I}_{\text{cur}}$ with all intervals (b, j) and (j, d) in D that have j as endpoint, or we set $\mathcal{I}_{\text{cur}} = \varnothing$.

The remainder of the algorithm enumerates all ways to extend the partial classification (which is provided by the user as input) by adding one simplex to the current class Φ_{cur}. There are four possible types of extensions, which we explain and illustrate with the example depicted in Fig. 5.4.

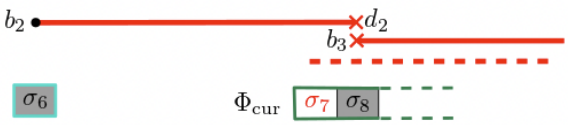


Figure 5.4: We consider the part of the partial classification from Fig. 5.3 that is relevant to the current class Φ_{cur} only. For Φ_{cur} to be completed we must at least find critical simplices $\pi_+(b_3, \infty)$ and $\pi_-(b_2, d_2)$ to account for the red intervals $(b_2, d_2), (b_3, \infty)$, and pair the non-critical simplex σ_7 with another non-critical simplex to destroy the unexpected dotted red interval.

- (Lines 14 to 19) The first type of extension consists in adding a simplex $\pi_+(j, d)$ critical for the birth of an interval $(j, d) \in \mathcal{I}_{\text{cur}}$ as in Fig. 5.5.

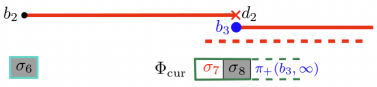


Figure 5.5: Adding a birth simplex $\pi_+(b_3, \infty)$ generating a $\dim(b_3, \infty)$-cycle, to be critical for the interval (b_3, ∞).

- (Lines 20 to 26) The second type of extension consists in adding a

critical simplex $\pi_-(b,j)$ to account for the death of an interval $(b,j) \in \mathcal{I}_{\text{cur}}$ as in Fig. 5.6.

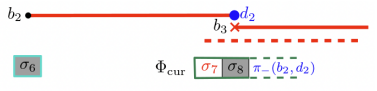


Figure 5.6: Adding a death simplex $\pi_-(b_2,d_2)$ destroying the $\dim(b_2,d_2)$-cycle generated by $\pi_+(b_2,d_2)$, to be critical for the interval (b_2,d_2).

- (Lines 27 to 31) The third type of extension consists in adding a simplex σ to pair with a non-critical unpaired simplex $\sigma' \in \Phi^{\text{n.c.}}_{\text{cur}}$ as in Fig. 5.7.

- (Lines 31 to 36) The fourth type of extension consists in adding a non-critical simplex σ to $\Phi^{\text{n.c.}}_{\text{cur}}$ as in Fig. 5.8, i.e. a simplex that creates a $\dim\sigma$-cycle that shall later be destroyed.

To figure out which simplices $\sigma \notin \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ can extend the partial D-compatible classifications $\mathcal{F}$ through the algorithm, we maintain a matrix δ which is derived from the boundary matrix ∂ via elementary row and column operations. The range of possible extensions can be deduced from the sparsity pattern of δ as follows; see [CSEM06][14] and Remark 5.4.4 for full details on the construction and interpretation of this matrix.

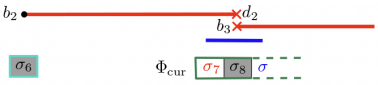


Figure 5.7: Finding a simplex σ to destroy the cycle generated by an unpaired non-critical simplex $\sigma_7 \in \Phi^{\text{n.c.}}_{\text{cur}}$, resulting in the removal of σ_7 in $\Phi^{\text{n.c.}}_{\text{cur}}$.

- Simplex σ can be added as a (critical or non-critical) simplex to Φ_{cur} if and only if $\delta\sigma = 0$ or the lowest non-zero entry of column associated to σ corresponds to another simplex σ' that already belongs to the partial classification: $\sigma' \in \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$. We then write $\text{low}(\delta\sigma) := \sigma'$;

- If a simplex σ such that $\delta\sigma = 0$ is added to Φ_{cur}, then it is a birth simplex and creates a new pair $[\sigma,\infty]$;

- If a simplex σ such that $\text{low}(\delta\sigma) = \sigma' \in \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ is added to Φ_{cur}, then it is a death simplex, which has the effect to replace a pair $[\sigma',\infty]$ in $\mathcal{F}$ with a new pair $[\sigma',\sigma]$.

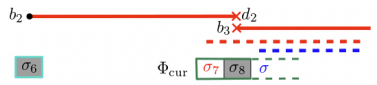


Figure 5.8: Finding a non-critical simplex σ generating a new cycle, resulting in the addition of σ in $\Phi^{\text{n.c.}}_{\text{cur}}$.

Given a complete D-compatible classification, recall that we have a correspondence π of intervals in D with critical pairs. When $\mathcal{F}$ is a (non-complete) partial D-compatible classification, the correspondence is not necessarily defined for all intervals in D: For instance there are no birth and death critical simplices for (b,d) in $\mathcal{F}$, for any interval $(b,d) \in D$ such that $j < b < d$. In this case we only have a partially defined correspondence, which we indicate by the conventions $\pi_+(b,d) := \varnothing$ and $\pi_-(b,d) := \varnothing$. Note that it is also possible to have $\pi(b,d) = (\sigma,\varnothing)$ whenever we already have a birth critical simplex $\sigma = \pi_+(b,d)$ in $\mathcal{F}$ that is not yet associated with a death critical simplex. In the algorithm, we store and update the partially defined correspondence π as we incrementally construct the partial classification $\mathcal{F}$.

Finally, to improve the time-efficiency of our algorithm, we maintain an array $\mathrm{M} \in \mathbb{N}^{\dim K+1}$ of integers that constrain the non-critical simplices that can be added to the classification in each dimension.

[14] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the twenty-second annual Symposium on Computational Geometry*, pages 119–126. ACM, 2006

In the array $\mathrm{M} = (\mathrm{M}_0, \cdots, \mathrm{M}_{\dim K})$, each M_p is the number of non-critical positive p-simplices that remain to be added in the classification. Since non-critical simplices come in pairs, M_p is also the number of non-critical negative $(p+1)$-simplices that remain to be added in the classification: At initialization this is the rank $\mathrm{rank}(\partial_{p+1})$ of the boundary matrix restricted to $(p+1)$-simplices, minus the number of bounded bars $(b,d) \in D$ in degree p (because those bars are in 1-1 correspondence with negative critical $(p+1)$-simplices). By convention $\mathrm{M}_{-1} = \mathrm{M}_{\dim K+1} = 0$.

In practice we call the main algorithm **ComputeFiltrations**(K, D) (see Alg 7) to build the initial partial classification and then call the exploration (see Alg. 8) as a subroutine **ExtendFiltration**$(\mathcal{F}, \textbf{Results}, \delta, \pi, \mathrm{M})$.

Algorithm 7 ComputeFiltrations(K, D)

Require: Finite simplicial complex K and target barcode D
Ensure: List **Results** of all D-compatible classifications
1: $\mathcal{F} = \big(\mathcal{K}_{\mathrm{in}}, j, \Phi_{\mathrm{cur}}, \mathcal{I}_{\mathrm{cur}}, \Phi^{\mathrm{n.c.}}_{\mathrm{cur}}\big) \leftarrow (\varnothing, \varnothing, 1, \varnothing, D^{<2}, \varnothing)$
2: Initialise δ as the boundary matrix of K
3: Initialise correspondence $\pi(b,d) := [\varnothing, \varnothing]$ for bounded $(b,d) \in D$ and $\pi(b,\infty) := [\varnothing, \infty]$ for infinite $(b,\infty) \in D$
4: $\mathrm{M}_p \leftarrow \mathrm{rank}(\partial_{p+1}) - \sharp\big\{(b,d) \in D \mid \dim(b,d) = p+1,\ d < \infty\big\}$, for $0 \leqslant p \leqslant \dim K$
5: **return** **ExtendFiltration**$(\mathcal{F}, \textbf{Results}, \delta, \pi, \mathrm{M})$

Theorem 5.4.2. *Algorithm* **ComputeFiltrations**(K, D) *returns the list of all D-compatible classifications.*

Proof. From Line 1 of Alg. 8, the output consists in the D-compatible classifications $\mathcal{K}_{\mathrm{in}}$ extracted from complete partial classifications $\mathcal{F}$ that are encountered by the algorithm. Thus it suffices to show that any given partial classification $\mathcal{F}$ is explored by the algorithm. We proceed by induction on the number m of classes $\Phi_1, \cdots, \Phi_m$ forming $\mathcal{K}_{\mathrm{in}}$, and on the number of simplices in $\Phi_1 \sqcup \cdots \sqcup \Phi_m \sqcup \Phi_{\mathrm{cur}}$. Note that the empty partial classification is visited at the beginning of the algorithm, so we may assume that $\Phi_1 \sqcup \cdots \sqcup \Phi_m \sqcup \Phi_{\mathrm{cur}} \neq \varnothing$. If $\Phi_{\mathrm{cur}} = \varnothing$, we have $m \geqslant 1$ and $\Phi_m \neq \varnothing$. We can then form the partial classification $\mathcal{F}'$ with $\mathcal{K}'_{\mathrm{in}} := \Phi_1 \sqcup \cdots \sqcup \Phi_{m-1}$, $\Phi'_{\mathrm{cur}} := \Phi_m$, $\mathcal{I}'_{\mathrm{cur}} := \varnothing$, $(\Phi^{\mathrm{n.c.}}_{\mathrm{cur}})' := \varnothing$, and $j' := j-1$ or $j' := j$ depending on whether $\Phi_m = \Phi_{i_{j-1}}$ or not. Clearly then, by Lines $3-10$ of Alg. 8, if $\mathcal{F}'$ is explored by the algorithm, then so is $\mathcal{F}$. Otherwise, $\Phi_{\mathrm{cur}} \neq \varnothing$ has a last simplex σ. We can then form $\mathcal{F}'$ by removing σ, i.e. $\Phi'_{\mathrm{cur}} := \Phi_{\mathrm{cur}} \backslash \{\sigma\}$, with $\mathcal{K}'_{\mathrm{in}} := \mathcal{K}_{\mathrm{in}}$ and $j' = j$ unchanged, and with the obvious changes in $\mathcal{I}'_{\mathrm{cur}}$ if σ were a critical simplex, or in $(\Phi^{\mathrm{n.c.}}_{\mathrm{cur}})'$ if it were non-critical. Then $\mathcal{F}$ is one of the four incremental extensions of $\mathcal{F}'$ depicted in Fig. 5.4, therefore by Lines $14-37$ of Alg. 8, if $\mathcal{F}'$ is explored by the algorithm, then so is $\mathcal{F}$. □

Algorithm 8 ExtendFiltration$(\mathcal{F}, \textbf{Results}, \delta, \pi, \mathrm{M})$

Require: Partial classification $\mathcal{F}$, list **Results**, matrix δ, pairing π, integers $\mathrm{M} \in \mathbb{N}^{d+1}$

if $|\mathcal{K}_{\text{in}}| = \sharp K$ and $D^{<j} = D$ and $\Phi_{\text{cur}} = \varnothing$ **then**
  return **Results** $\cup \{\mathcal{K}_{\text{in}}\}$
else if $\Phi_{\text{cur}} \neq \varnothing$ and $\mathcal{I}_{\text{cur}} = \varnothing$ and $\Phi^{\text{n.c.}}_{\text{cur}} = \varnothing$ **then**
  $\mathcal{K}'_{\text{in}} \leftarrow \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$
  if Φ_{cur} contains critical simplices for the endpoint j **then** $j' \leftarrow j+1$ **else** $j' \leftarrow j$
  if $D^{<j'} \neq D$ **then**
    $\mathcal{I}'_{\text{cur}} \leftarrow \{(b,d) \in D \mid b = j' \text{ or } d = j'\}$
    $\mathcal{F}' \leftarrow (\mathcal{K}'_{\text{in}}, j', \varnothing, \varnothing, \mathcal{I}'_{\text{cur}})$ and **Results** $\leftarrow$ **ExtendFiltration**$(\mathcal{F}', \textbf{Results}, \delta, \pi, \mathrm{M})$
  end if
  $\mathcal{F}' \leftarrow (\mathcal{K}'_{\text{in}}, j', \varnothing, \varnothing, \varnothing)$
  return **ExtendFiltration**$(\mathcal{F}', \textbf{Results}, \delta, \pi, \mathrm{M})$
else
  for interval (b,d) in $\mathcal{I}_{\text{cur}}$ **do**
    if $b = j$ **then**
      for simplex $\sigma \notin \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ such that $\delta\sigma = 0$ and $\dim\sigma = \dim(b,d)$ **do**
        $\Phi'_{\text{cur}} \leftarrow \Phi_{\text{cur}} \cup \{\sigma\}$ and $\mathcal{I}'_{\text{cur}} \leftarrow \mathcal{I}_{\text{cur}} \backslash \{(b,d)\}$ and $\mathcal{F}' \leftarrow (\mathcal{K}_{\text{in}}, j, \Phi'_{\text{cur}}, \mathcal{I}'_{\text{cur}}, \Phi^{\text{n.c.}}_{\text{cur}})$
        $\pi' \leftarrow \pi$ where π' agrees with π on all inputs, except that $\pi'_+(b,d) := \sigma$ instead of $\varnothing$
        Results $\leftarrow$ **ExtendFiltration**$(\mathcal{F}', \textbf{Results}, \delta, \pi', \mathrm{M})$
      end for
    else if $d = j$ **then**
      for simplex $\sigma \notin \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ such that $\mathrm{low}(\delta\sigma) = \pi_+(b,d)$ and $\dim\sigma = \dim(b,d)+1$ **do**
        $\Phi'_{\text{cur}} \leftarrow \Phi_{\text{cur}} \cup \{\sigma\}$ and $\mathcal{I}'_{\text{cur}} \leftarrow \mathcal{I}_{\text{cur}} \backslash \{(b,d)\}$ and $\mathcal{F}' \leftarrow (\mathcal{K}_{\text{in}}, j, \Phi'_{\text{cur}}, \mathcal{I}'_{\text{cur}}, \Phi^{\text{n.c.}}_{\text{cur}})$
        $\delta' \leftarrow$ reduce δ with respect to σ, c.f. Remark 5.4.4.
        $\pi' \leftarrow \pi$, where π' agrees with π on all inputs, except that $\pi'_-(b,d) := \sigma'$ instead of $\varnothing$
        Results $\leftarrow$ **ExtendFiltration**$(\mathcal{F}', \textbf{Results}, \delta', \pi', \mathrm{M})$
      end for
    end if
  end for
  for simplex $\tau \in \Phi^{\text{n.c.}}_{\text{cur}}$ **do**
    for simplex $\sigma \notin \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ such that $\mathrm{low}(\delta\sigma) = \tau$ and $\dim\sigma = \dim\tau + 1$ **do**
      $\mathcal{F}' \leftarrow (\mathcal{K}_{\text{in}}, j, \Phi_{\text{cur}} \cup \{\sigma\}, \mathcal{I}_{\text{cur}}, \Phi^{\text{n.c.}}_{\text{cur}} \backslash \{\tau\})$ and $\delta' \leftarrow$ reduce δ with respect to σ, c.f. Remark 5.4.4
      Results $\leftarrow$ **ExtendFiltration**$(\mathcal{F}', \textbf{Results}, \delta', \pi, \mathrm{M})$
    end for
  end for
  for simplex $\sigma \notin \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ such that $\delta\sigma = 0$ **do**
    if $\mathrm{M}_{\dim\sigma} > 0$ **then**
      $\mathrm{M}_{\dim\sigma} \leftarrow \mathrm{M}_{\dim\sigma} - 1$ and $\mathcal{F}' \leftarrow (\mathcal{K}_{\text{in}}, j, \Phi_{\text{cur}} \cup \{\sigma\}, \mathcal{I}_{\text{cur}}, \Phi^{\text{n.c.}}_{\text{cur}} \cup \{\sigma\})$
      Results $\leftarrow$ **ExtendFiltration**$(\mathcal{F}', \textbf{Results}, \delta, \pi, \mathrm{M})$
    end if
  end for
  return List **Results** enriched with D-compatible classifications that extend the classification $\mathcal{K}_{\text{in}}$
end if

Remark 5.4.3. The exploration of Algorithm 8 is equivalently viewed as a Depth-First Search (DFS) on *the tree of partial D-compatible classifications*: each node is a partial classification, whose children differ by the addition of a single simplex according to the four types of extension depicted in Fig. 5.4. The algorithm records the subset of leaves that correspond to D-compatible classifications. It would also have been possible to design a Breadth-First-Search (BFS) algorithm. However the BFS approach requires more storage, because we can forget the information stored in a node (e.g. the boundary matrix of a partial classification) only when all its children are treated. Hence in a BFS version of the algorithm we would eventually need to store the information of the entire tree, while in the DFS version at most one branch is stored at a time.

Remark 5.4.4. We implicitly maintain a matrix factorization $\delta = \partial V$ at each step of the algorithm, where $\partial \in \mathbb{k}^{K\times K}$ is the total boundary matrix of K and $\delta, V \in \mathbb{k}^{K\times K}$ are square. This factorization must satisfy two conditions, each of which is stated in terms of a sequence of form $\xi = (\sigma_1, \ldots, \sigma_p, \tau_1, \ldots, \tau_q, v_1, \ldots, v_r)$, where $\sigma_1 < \cdots < \sigma_p$ is the linear order on $\mathcal{K}_{\text{in}}$ and $v_1 < \ldots < v_q$ is the linear order on Φ_{cur}. We write $\hat{\partial}$ for the matrix obtained by permuting rows and columns of ∂ such that simplex ξ_k indexes the kth row and column of $\hat{\partial}$, for each k. Matrices $\hat{\delta}$, and $\hat{V}$ are defined similarly, by permuting rows and columns to ensure that ξ_k indexes the kth row and column of each matrix. Our two conditions can now be stated as follows:

- Matrix $\hat{V}$ must be upper unitriangular.
- Matrix $\hat{\delta}$ must be *partially reduced* in the following sense. For each birth-death pair $[\tau, \tau']$ in $\mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$ (including non-critical pairs and excluding pairs of form $[\tau, \infty]$), the entry $\hat{\delta}(\tau, \tau')$ must be nonzero, and each entry that lies either directly below $\hat{\delta}(\tau, \tau')$ in column τ' (respectively, each entry to the right of $\hat{\delta}(\tau, \tau')$ in row τ) must equal zero.

It can be shown that the low function of $\hat{\delta}$ agrees with the low function of any $R = DV$ decomposition of$\hat{\partial}$ (when restricting each of these functions to the subset $\mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$; their values may differ for $\tau \notin \mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}}$), c.f. [CSEM06][15].

To obtain such a factorization after we have added σ to Φ_{cur}, first fix a compatible sequence ξ and permute the columns of δ accordingly; then perform one further swap to ensure that the column indexed by σ appears directly to the right of the last column indexed by $\mathcal{K}_{\text{in}}$, keeping the location of all columns indexed by $\mathcal{K}_{\text{in}}$ fixed. Perform the same permutation on rows. Then add multiples of column σ to columns on its right as necessary to ensure that the unique nonzero

[15] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the twenty-second annual Symposium on Computational Geometry*, pages 119–126. ACM, 2006

entry in row $\mathrm{low}(\delta\sigma)$ appears in column σ. A routine exercise shows that the resulting matrix $\hat{\delta}'$, fits into a matrix factorization $\hat{\delta}' = \hat{\partial}\hat{V}$ of the appropriate form.

Remark 5.4.5. From Theorem 5.3.6 the polyhedra $\mathcal{P}(\mathcal{K})$ induced by the D-compatible classifications $\mathcal{K}$ describe the fiber $\mathrm{PH}^{-1}(D)$ as a polyhedral complex. In applications where it is desirable to dispose of a simplicial complex structure for $\mathrm{PH}^{-1}(D)$, we can simply form the nerve of the cover associated to the polyhedra $\mathcal{P}(\mathcal{K})$, which by the Nerve theorem for Euclidean closed convex sets is homotopy equivalent to $\mathrm{PH}^{-1}(D)$ (see e.g. [Wei52][16]). In practice, computing the nerve amounts to finding non-empty intersections $\bigcap_{l=1}^{n} \mathcal{P}(\mathcal{K}_l)$, which in a polyhedral complex boils down to intersecting vertex sets:

[16] André Weil. Sur les théoremes de de rham. *Comment. Math. Helv*, 26(1):119–145, 1952

$$\mathrm{V}\left(\bigcap_{l=1}^{n} \mathcal{P}(\mathcal{K}_l)\right) = \bigcap_{l=1}^{n} \mathrm{V}(\mathcal{P}(\mathcal{K}_l)).$$

We include the construction of this simplicial complex for describing $\mathrm{PH}^{-1}(D)$ in our implementation.

Remark 5.4.6. Since the polyhedral decomposition of the fiber realizes a regular CW complex, computing the $\mathbb{Z}_2$ linear boundary operator of this object reduces to enumerating the codimension-1 faces of each polyhedron. These may be computed from the standard formula for the boundary of a product of copies of standard geometric simplices:

$$\partial(\sigma_1 \times \cdots \sigma_m) = \bigcup_k [\sigma_1 \times \cdots \times \partial(\sigma_k) \times \cdots \times \sigma_m]$$

Computing the coefficients of the $\mathbb{Z}$-linear boundary matrix is slightly more involved, due to orientation. We deffer this problem to later work.

Remark 5.4.7 (Generalization to barcodes for persistent (relative) (co)homology)**.** The discussion this far has focused exclusively on barcodes of the homological persistence module (obtained by applying the homology functor to a nested sequence of cell complexes). However, there are several other formulae for generating persistence modules from a filtered cell complex. While each construction has distinct and useful algebraic properties, their barcodes are completely determined by that of the homological barcode[17]. Thus the procedure to compute fiber of PH also serves to compute the persistence fiber of these other constructions. For a detailed discussion, see section 5.8.

[17] Vin De Silva, Dmitriy Morozov, and Mikael Vejdemo-Johansson. Dualities in persistent (co) homology. *Inverse Problems*, 27(12):124003, 2011

5.5 *Generalisation to based chain complexes*

We generalise the fact that the fiber of the persistence map PH is a polyhedral complex to filters defined directly at the level of based chain complexes. These include filters on simplicial complexes, cubical

complexes, delta complexes and CW complexes. In turn our approach for computing $\mathrm{PH}^{-1}(D)$ adapts to these situations as well.

Definition 5.5.1. A based, finite-dimensional, $\Bbbk$-linear chain complex is a pair (C, E) such that $\sum_i \dim(C_i) < \infty$ and E is a union of bases E_i of C_i for all i. A *filter* on (C, E) is a real-valued function $f : E \to \mathrm{I}$ such that the linear span of $\{e \in E : f(e) \leqslant t\}$ forms a linear subcomplex of C, for each $t \in \mathrm{I}$.

Here are some examples of (C, E) induced by combinatorial complexes:

1. **Simplicial Complexes.** Basis E is the collection of simplices in a simplicial complex K. We recover the standard setting of filters over K.
2. **Cubical complexes.** Basis E is the collection of cubes in a cubical complex.
3. **Delta and CW Complexes.** Basis E is the collection of cells in a delta complex or CW complex K.

These variations are of interest in practice: For instance with delta and CW complexes we can decompose topological spaces with much fewer simplices, while cubical complexes appear naturally e.g. in image analysis. The main result of this section, Theorem 5.5.14, generalizes the structure theorem for simplicial complexes (Theorem 5.2.2) to these important variants. In particular, Theorem 5.5.14 implies each of the following results.

Theorem 5.5.2. *Let K be a simplicial, cubical, delta or CW complex and let D be a barcode. Then the fiber $\mathrm{PH}^{-1}(D)$ is the underlying space of a polyhedral complex whose polyhedra are products of standard simplices.*

Theorem 5.5.3. *Theorem 5.5.2 remains true if we restrict to lower-star filtrations or Vietoris-Rips filtrations.*

5.5.1 Polyhedral decomposition of the ambient cube

Let E be a finite set, and let $\Gamma = \{\gamma_0 < \cdots < \gamma_m\}$ be a finite subset of the interval I, where $\gamma_0 = 1$ and $\gamma_m \leqslant \sharp K$. For any pair of functions $f, g : E \to \mathrm{I}$, let us define the relation $f \sim_\Gamma g$ by either of the following two equivalent axioms:

Item 1 There exists an order-preserving map $\psi : \mathrm{I} \to \mathrm{I}$ such that $\psi(\gamma_i) = \gamma_i$ for each i, and $g = \psi \circ f$.

Item 2 The function g satisfies

$$f(e) \in \Gamma \implies g(e) = f(e) \qquad f(e) \leqslant f(e') \implies g(e) \leqslant g(e'). \tag{5.2}$$

Note that the relation $\sim_\Gamma$ is reflexive and transitive but not symmetric. We write

$$\mathcal{P}_\Gamma(f) := \{g \in \mathrm{I}^E : \ f \sim_\Gamma g\}.$$

Lemma 5.5.4. *The set $\mathcal{P}_\Gamma(f)$ is a compact polyhedron.*

Proof. Axiom Item 2 represents a family of logical conditions, each of which determines either a hyperplane, i.e. $\{g : g(e) = c\}$, or a closed half-space, i.e. $\{g : g(e') - g(e) \geqslant 0\}$. The intersection of these sets, $\mathcal{P}_\Gamma(f)$, is a bounded polyhedron. □

Lemma 5.5.5. *If $g \in \mathcal{P}_\Gamma(f)$ then $\mathcal{P}_\Gamma(g) \subseteq \mathcal{P}_\Gamma(f)$.*

Proof. Relation $\sim_\Gamma$ is transitive. □

Proposition 5.5.6. *The set of convex polyhedra $\mathcal{P}_\Gamma = \{\mathcal{P}_\Gamma(f) : f : E \to I\}$ is a polyhedral complex; the underlying space is $|\mathcal{P}_\Gamma| = I^E$.*

Proof. Follows from Lemmas 5.5.4 and 5.5.5. □

Next we provide the description of each polyhedron $\mathcal{P}_\Gamma(f)$ in the complex as a product of standard simplices. For convenience let $\gamma_{m+1} := \infty$, and for each $i \in \{1, \ldots, m\}$, let $\eta^i = f(E) \cap (\gamma_i, \gamma_{i+1})$.

Lemma 5.5.7. *$\mathcal{P}_\Gamma(f)$ is affinely isomorphic to $\Delta_{\#\eta^0} \times \cdots \times \Delta_{\#\eta^m}$, where Δ_k stands for the standard geometric simplex of dimension k.*

Proof. Follows from Axiom Item 1. □

5.5.2 *Polyhedral decomposition of* $\mathrm{PH}^{-1}(D)$ *for based chain complexes*

Let (C, E) be a based, finite-dimensional, $\Bbbk$-linear chain complex.

Lemma 5.5.8. *Let f be a filter on (C, E); D be the associated (total) barcode; and $\Gamma := \mathrm{End}(I, \leqslant)^*(D)$ be the set of finite endpoints of intervals in D. Then each element of $\mathcal{P}_\Gamma(f)$ is a bona fide filter on (C, E) with barcode D.*

Proof. The persistence map is equivariant in the sense that $\mathrm{PH}(\psi \circ f) = \psi.\mathrm{PH}(f)$ for $\psi : \mathrm{I} \to \mathrm{I}$ a non-decreasing map, as can be proven e.g. following [LT21, Lemma 1.5][18]. Here ψ acts point-wise on intervals of $\mathrm{PH}(f)$, i.e. each interval $(b, d) \in \mathrm{PH}(f)$ produces an interval $(\psi(b), \psi(d))$ in $\psi.\mathrm{PH}(f)$ whenever $\psi(b) \neq \psi(d)$. The result follows from Axiom Item 1. □

[18] Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021

Given S a union of polyhedra in a polyhedral complex $\mathcal{P}$, then the subcomplex induced by S is defined as

$$\mathcal{P}[S] = \{X \in \mathcal{P} : X \subseteq S\}. \tag{5.3}$$

Theorem 5.5.9. *Let (C, E) be a point-wise finite dimensional based chain complex; D be a barcode; $\Gamma := \mathrm{End}(I, \leqslant)^*(D)$; and $\mathrm{PH}^{-1}(D)$ be the set of filters on (C, E) with barcode D. Then $\mathrm{PH}^{-1}(D)$ is a union of polyhedra in $\mathcal{P}_\Gamma$, hence there exists a well-defined polyhedral subcomplex*

$$\mathcal{P}_\Gamma[\mathrm{PH}^{-1}(D)] \subseteq \mathcal{P}_\Gamma$$

with underlying space $|\mathcal{P}_\Gamma[\mathrm{PH}^{-1}(D)]| = \mathrm{PH}^{-1}(D)$.

Proof. Lemma 5.5.8 implies that $\mathrm{PH}^{-1}(D)$ is a union of the polyhedra in $\{\mathcal{P}_\Gamma(f) : f \in \mathrm{PH}^{-1}(D)\}$, and by continuity it is a closed subset of the polyhedral complex $\mathcal{P}_\Gamma = \{\mathcal{P}_\Gamma(f) : f : E \to \mathrm{I}\}$ (Proposition 5.5.6). Therefore $\mathrm{PH}^{-1}(D)$ is a sub-polyhedral complex. □

5.5.3 *Polyhedral decomposition of* $\mathrm{PH}^{-1}(D)$ *for CW complexes*

The polyhedral decomposition of $\mathrm{PH}^{-1}(D)$ for based chain complexes (Theorem 5.5.9) does not carry over directly to arbitrary CW complexes because given a CW complex K, there may exist filters on the associated based chain complex (C, E) that do not correspond to valid filters of K. Here the notion of filter on a CW complex (respectively, delta or cubical complex) naturally generalises the simplicial situation: it qualifies any function $f : K \to \mathrm{I}$ whose sub-level sets are sub-CW complexes (respectively, sub-delta or sub-cubical complexes).

Example 5.5.10. Let $K = \{v, e\}$ be the CW decomposition of S^1 with one vertex, v, and one edge, e. Since all boundary maps are 0, the filtration $\{e\} \subseteq \{v, e\}$ is perfectly valid for (C, E), but not for K.

Fortunately this problem is simple to address. Lemma 5.5.11 represents the only technical observation needed to extend our polyhedral characterization of the persistence fiber from regular finite CW complexes to arbitrary finite CW complexes. The proof is vacuous.

Lemma 5.5.11. *Suppose that a function $P : I^E \to \{\mathrm{True}, \mathrm{False}\}$ satisfies the condition that*

$$P(f) = \mathrm{True} \implies P \text{ evaluates to } \mathrm{True} \text{ on each element of } \mathcal{P}_\Gamma(f). \quad (5.4)$$

Then $P^{-1}(\mathrm{True})$ is a union of polyhedra in $\mathcal{P}_\Gamma$, hence $\mathcal{P}_\Gamma[P^{-1}(\mathrm{True})]$ is a polyhedral subcomplex.

Theorem 5.5.12. *Let K be a simplicial, cubical, delta or CW complex. Let D be a barcode, and let (C, E) be the induced based chain complex. Then the fiber $\mathrm{PH}^{-1}(D)$ is a polyhedral complex whose polyhedra are products of standard simplices.*

Proof. Let K be a finite CW complex, and let $P(f) = \mathrm{True}$ iff $f^{-1}(-\infty, t]$ is a CW subcomplex of K for each $t \in \mathrm{I}$. Then $P(f) \implies P(g)$ for each

g such that $f \sim_\Gamma g$, by Axiom Item 1. Lemma 5.5.11 therefore implies that $\mathrm{PH}^{-1}(D)$ is the underlying space of a polyhedral complex. The polyhedra in this complex are isomorphic to products of simplices, by Lemma 5.5.7. □

In practice, it is simple to adapt the **ExtendFiltration** algorithm to compute $\mathrm{PH}^{-1}(D)$ for a CW complex. One must simply add the condition that $\mathcal{K}_{\mathrm{in}} \cup \{\sigma\}$ be a bona fide subcomplex to each of the lists of criteria found on lines 15, 21, 28, and 32.

5.5.4 *Polyhedral decomposition of* $\mathrm{PH}^{-1}(D) \cap S$*, for a restricted family* S

It is common in practice to restrict one's attention to a restricted family of filters, S. A noteworthy family of examples come from the *lower-p filtrations*. Given a CW complex K with associated based chain complex (C, E), we define the space of *lower-p filtrations* on K as

$$\mathrm{Low}_p(K) = \{f : E \to \mathrm{I} : \text{ for each cell } \tau \in E, \text{ one has } f(\tau) = \max\{f(\sigma) : \sigma \subseteq \mathrm{cl}(\tau), \ \dim(\sigma) \leqslant p\}\}$$

Examples include

1. The space of lower-star filtration, $\mathrm{Low}_0(K)$.
2. The space of lower-edge filtrations, $\mathrm{Low}_1(K)$. These include the so-called Vietoris-Rips filtrations of a combinatorial simplicial complex.

The important fact about these spaces is that they are unions of polyhedra in the polyhedral decomposition of the ambient cube:

Lemma 5.5.13. *For each p and each finite subset $\Gamma \subset I$, the space $\mathrm{Low}_p(K)$ is a union of polyhedra in $\mathcal{P}_\Gamma$.*

Proof. Fix $f \in \mathrm{Low}_p(K)$, and suppose that $f \sim_\Gamma g$ for some g. Then it follows from Axiom Item 1 that $g \in \mathrm{Low}_p(K)$. In particular, if $\mathrm{Low}_p(K)$ contains an element of a polyhedron in $\mathcal{P}_\Gamma$, then it contains the entire polyhedron. □

Then we have the following.

Theorem 5.5.14. *Suppose that (C, E) is the based chain complex associated with a finite simplicial complex, a delta complex or a CW complex K. Let D be a barcode and $\Gamma = \mathrm{End}(I, \leqslant)^*(D)$ be the associated set of finite endpoints. Then*

$$S = \mathrm{PH}^{-1}(D) \cap \mathrm{Low}_p(K)$$

is a union of polyhedra in $\mathcal{P}_\Gamma$. Consequently there is a nested sequence of polyhedral subcomplexes

$$\mathcal{P}_\Gamma[S] \subseteq \mathcal{P}_\Gamma[\mathrm{PH}^{-1}(D)] \subseteq \mathcal{P}_\Gamma. \tag{5.5}$$

When $p = 0$, the set S is the space of lower-star filters with barcode D. When $p = 1$, the set S is the space of lower-edge filters with barcode D.

Proof. Follows from Lemma 5.5.13. □

Remark 5.5.15 (Adapting the **ExtendFiltration** algorithm)**.** If we wish only to compute a polyhedral decomposition of D-compatible lower-p filters, then the exploration of Alg. 8 can be considerably reduced. For this we maintain a list L of simplices $\sigma \notin \mathcal{K}_{\text{in}}$ whose faces τ with $\dim(\tau) \leqslant p$ satisfy $\tau \in \mathcal{K}_{\text{in}}$; we deem the current class Φ_{cur} incomplete as long as L is non-empty. In practice this simply amounts to modifying Line 3 to:

(Line 3′) **else if** $\mathrm{L} = \varnothing$ and $\Phi_{\text{cur}} \neq \varnothing$ and $\mathcal{I}_{\text{cur}} = \varnothing$ and $\Phi^{\text{n.c.}}_{\text{cur}} = \varnothing$ **then**

In addition we need to update the list L each time we add a new vertex τ in the classification (which can happen at lines 16, 22, 29 and 35) simply by adding the following line:

Add to list L all simplices σ such that $\mathcal{K}_{\text{in}} \cup \Phi_{\text{cur}} \cup \{\tau\}$ contains the k-dimensional faces of σ for each $k \leqslant p$.

5.6 Experiments

Using our algorithm we compute the number of polyhedra in $\mathrm{PH}^{-1}(D)$, binned by dimension, and the non-zero Betti numbers $\beta_p(\mathrm{PH}^{-1}(D))$. We also record the special cases where the facets, i.e. polyhedra in $\mathrm{PH}^{-1}(D)$ that are maximal w.r.t. inclusions, do not consist only in the top-dimensional polyhedra.

Implementation Algorithm 3.3.1 was implemented in the programming language Rust, and is publicly available [19]. This implementation accommodates user-defined coefficient fields, based complexes, and restricted families of filters, e.g. lower-star. The implementation uses several dependencies from the ExHACT library[20] for low-level functions, including reduction of boundary matrices and implementation of common coefficient fields.

[19] Leygonie Jacob and Gregory Henselman-Petrusek. Software Companion to Algorithmic Reconstruction of the Fiber of Persistent Homology on Cell Complexes. https://github.com/Eetion/phfibre, 2022

[20] Haibin Hang and Gregory Henselman-Petrusek. Exact homological algebra for computational topology (ExHACT). https://github.com/ExHACT, 2021

Reading the results The outputs of the algorithm are reported in figures. Each figure corresponds to a specific simplicial, or CW complex and provides statistics about the fiber $\mathrm{PH}^{-1}(D)$ for various barcodes in a table. By convention, black intervals in the target barcode D are of dimension 0, blue intervals are of dimension 1, while green intervals are of dimension 2. In all cases the number of polyhedra in $\mathrm{PH}^{-1}(D)$ is binned by dimension in the form of an array, and the Betti numbers are computed with coefficients in $\mathbb{Z}_2$. Unless explicitly stated otherwise:

- The facets are the top dimensional polyhedra. Otherwise, we explicitly report in red the facets binned by dimension in an array;
- The persistence modules associated to barcodes D are computed with coefficients in $\mathbb{Z}_2$. Otherwise, in special cases where we also compute persistence modules with coefficients in $\mathbb{Q}$, the coefficient field is indicated in blue by a specific mention;
- The fiber is computed inside the space of all filters. Otherwise, we provide as many columns for statistics about $\mathrm{PH}^{-1}(D)$ as there are categories of filters to consider.

5.6.1 *Simplicial Complexes*

In all the examples of this section K is a simplicial complex. When K is a tree (Figure 5.9), we report these statistics both when the domain of PH consists of all filters and when it is restricted to lower star filters. For lower star filters on the interval the fiber is shown by [CMW20][21] to consist of contractible components. Our computations indicate that this property holds as well for general filters on the interval, however it breaks for other trees where the fiber has loops as indicated by non-trivial Betti numbers β_1.

[21] Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020

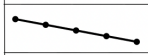
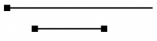
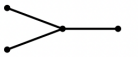
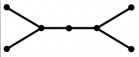
Complex K	Barcode D	Fiber	
		All filtrations	Lower star filtrations
		Polyhedra: [21, 105, 231, 258, 144, 32] $\beta_0 = 1$	Polyhedra: [21, 105, 231, 258, 144, 32] $\beta_0 = 1$
		Polyhedra: [990, 4674, 8214, 6352, 1824] $\beta_0 = 2$	Polyhedra: [330, 1064, 1152, 416] $\beta_0 = 2$
		Polyhedra: [11, 31, 33, 12] $\beta_0 = 1$	Polyhedra: [11, 31, 33, 12] $\beta_0 = 1$
		Polyhedra: [105, 201, 96] $\beta_0 = 1, \beta_1 = 1$	Polyhedra: [12, 6] $\beta_0 = 6$
		Polyhedra: [37, 291, 895, 1300, 900, 240] $\beta_0 = 1$	Polyhedra: [37, 291, 895, 1300, 900, 240] $\beta_0 = 1$
		Polyhedra: [2385, 11505, 20390, 15840, 4560] $\beta_0 = 1, \beta_1 = 11$	Polyhedra: [160, 380, 360, 120] $\beta_0 = 20$
		Polyhedra: [9280, 29400, 30900, 10800] $\beta_0 = 5, \beta_1 = 25$	Empty fiber
		Polyhedra: [37, 319, 1157, 2152, 2170, 1132, 240] $\beta_0 = 1$	Polyhedra: [37, 319, 1157, 2152, 2170, 1132, 240] $\beta_0 = 1$

Figure 5.9: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a tree.

For lower star filters on arbitrary subdivisions of the circle it is proven in [MW21][22] that the fiber is made of circular components. Our computations (Figure 5.10) suggest that this property holds as well when allowing general filters and adding dangling edges to the circle.

[22] Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021

When K is homotopy equivalent to a bouquet of two circles (Figure 5.11), the fiber itself has trivial homology in degree higher than 1

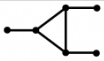
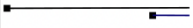
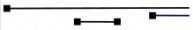
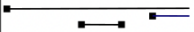
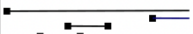
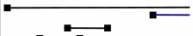
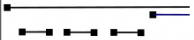
Complex K	Barcode D	Fiber
		Polyhedra: [9, 21, 12] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [25, 91, 106, 40] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [16, 64, 80, 32] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [25, 150, 305, 260, 80] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [68, 358, 664, 526, 152] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [183, 1338, 3591, 4572, 2808, 672] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [224, 224] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [192, 192] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [288, 288] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [2100, 4500, 2400] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [1560, 3400, 1840] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [3420, 7420, 4000] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [13440, 46272, 52416, 19584] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [8736, 31104, 36192, 13824] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [27360, 96672, 111936, 42624] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [10080, 21216, 11136] $\beta_0 = 8, \beta_1 = 8$

Figure 5.10: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is homotopy equivalent to a circle.

and we observe cases (indicated in red) where some facets are not top-dimensional polyhedra.

In light of all the previous calculations, we can conjecture that when K is a graph the fiber $\mathrm{PH}^{-1}(D)$ has trivial homology in degrees higher than 1. However, when K is the 2-skeleton of the 3-simplex (Figure 5.12), for some barcodes D the fiber has non-trivial degree 3 homology. Therefore in general the fiber $\mathrm{PH}^{-1}(D)$ may have higher non-trivial homologies than the base complex K.

Let K be an arbitrary connected simplicial complex, and let D_K be the barcode with one infinite bar $(0, +\infty)$ in degree 0, with no finite bars, followed by infinite bars $(1, +\infty)$ of multiplicity $\beta_p(K)$ in each degree $p \geqslant 1$. In all the examples computed with $\mathbb{Z}_2$ coefficients by our algorithm, the fiber $\mathrm{PH}^{-1}(D_K)$ and the base complex K have the same Betti numbers (with coefficients in Z_2). This motivates the following conjecture.

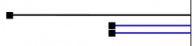
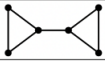
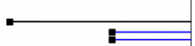
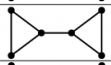
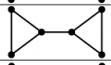
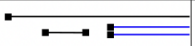
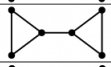
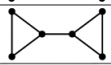
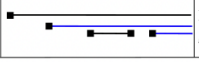
Complex K	Barcode D	Fiber	
		All filtrations	Lower Star filtrations
		Polyhedra: [42, 313, 718, 672, 224] $\beta_0 = 1, \beta_1 = 2$	Polyhedra: [15, 32, 16] $\beta_0 = 1, \beta_1 = 2$ Facets: [0, 4, 16]
		Polyhedra: [236, 1378, 2814, 2440, 768] $\beta_0 = 2, \beta_1 = 2$	Polyhedra: [52, 108, 56] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [2416, 9040, 10872, 4256] $\beta_0 = 1, \beta_1 = 9$	Polyhedra: [8] $\beta_0 = 8$
		Polyhedra: [6908, 24800, 28952, 11072] $\beta_0 = 2, \beta_1 = 14$	Polyhedra: [16] $\beta_0 = 16$
		Polyhedra: [380, 1364, 1592, 608] $\beta_0 = 4, \beta_1 = 4$	Empty fiber
		Polyhedra: [25, 124, 170, 72] $\beta_0 = 1, \beta_1 = 2$	Polyhedra: [11, 20, 8] $\beta_0 = 1, \beta_1 = 2$ Facets: [0, 2, 8]
		Polyhedra: [98, 386, 480, 192] $\beta_0 = 3, \beta_1 = 3$	Polyhedra: [12,12] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [572, 1298, 720] $\beta_0 = 1, \beta_1 = 7$	Polyhedra: [4] $\beta_0 = 4$
		Polyhedra: [1372, 3012, 1632] $\beta_0 = 4, \beta_1 = 12$	Empty fiber
		Polyhedra: [36, 84, 48] $\beta_0 = 4, \beta_1 = 4$	Empty fiber
		Polyhedra: [54, 531, 1726, 2562, 1792, 480] $\beta_0 = 1, \beta_1 = 2$	Polyhedra: [21, 70, 80, 32] $\beta_0 = 1, \beta_1 = 2$ Facets: [0, 4, 0, 32]
		Polyhedra: [356, 2730, 7790, 10576, 6920, 1760] $\beta_0 = 2, \beta_1 = 2$	Polyhedra: [92, 296, 324, 120] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [6832, 37982, 75144, 63712, 19712] $\beta_0 = 1, \beta_1 = 7$	Polyhedra: [64, 64] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [21378, 112012, 212402, 174352, 52576] $\beta_0 = 2, \beta_1 = 10$	Polyhedra: [144, 144] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [1342, 6626, 11948, 9384, 2720] $\beta_0 = 2, \beta_1 = 2$	Polyhedra: [24, 24] $\beta_0 = 4, \beta_1 = 4$

Figure 5.11: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is homotopy equivalent to a bouquet of two circles.

Conjecture. *Let K be a simplicial complex. Then the fiber $\mathrm{PH}^{-1}(D_K)$ and K have the same Betti numbers.*

5.6.2 CW Complexes

In this section K is a surface with a CW structure: the torus (Fig. 5.13), the Klein bottle (Fig. 5.14), the real projective plane (Fig.5.16), the Möbius strip (Fig. 5.17), the cylinder (Fig.5.18) and the Dunce Hat (Fig. 5.19). Indeed from section 5.5 our algorithm adapts to CW complexes and more generally to based chain complexes. This is a precious feature since simplicial triangulations of our surfaces have many simplices, hence our algorithm struggles to compute the associated fibers, while it handles cellular decompositions which are much smaller.

For cellular triangulations that are too small (e.g. the first two decompositions of the torus in Fig. 5.13), fibers are not interesting. This is why we consider cellular decompositions that are not minimal and have sufficiently many simplices for the fibers to be interesting. For

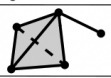
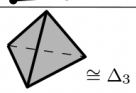
Complex K	Barcode D	Fiber
		Polyhedra: [64, 858, 4480, 10860, 13320, 8064, 1920] $\beta_0 = 1, \beta_2 = 1$
		Polyhedra: [3936, 29928, 84516, 113244, 72960, 18240] $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \beta_3 = 1$
		Polyhedra: [3936, 29928, 84516, 113244, 72960, 18240] $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \beta_3 = 1$
		Polyhedra: [87588, 422112, 743892, 571392, 162048] $\beta_0 = 1, \beta_1 = 2, \beta_2 = 27, \beta_3 = 2$
		Polyhedra: [324, 1512, 2628, 2016, 576] $\beta_0 = 4,\ \beta_1 = 8,\ \beta_2 = 4$
		Polyhedra: [6696, 33984, 62520, 49824, 14592] $\beta_0 = 4,\ \beta_1 = 14,\ \beta_2 = 13$
		Polyhedra: [6552, 33984, 63336, 50880, 14976] $\beta_0 = 4,\ \beta_1 = 14,\ \beta_2 = 13$
		Polyhedra: [11304, 59184, 111048, 89664, 26496] $\beta_0 = 2,\ \beta_1 = 28,\ \beta_2 = 26$
		Polyhedra: [183, 3262, 22283, 72710, 127173, 122937, 62088, 12816] $\beta_0 = 1, \beta_2 = 1$
 $\cong \Delta_3$		Polyhedra: [65, 922, 5338, 15340, 24180, 21384, 9984, 1920] $\beta_0 = 1$

Figure 5.12: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is homotopy equivalent to a 2-sphere.

such fibers, the remarks of section 5.6.1 about simplicial complexes apply as well. In particular, the fiber $\mathrm{PH}^{-1}(D_K)$ and the base complex K have the same Betti numbers.

We also find novel behaviours:

- For some CW complexes that are topological manifolds, such as the real projective plane and the Klein bottle, there are fibers whose facets do not consist only in top-dimensional polyhedra. In particular these fibers are not manifolds.

- For some CW complexes that are topological manifolds, such as the real projective plane, there are fibers whose connected components don't have the same homotopy type. This situation is detected whenever $\beta_0(\mathrm{PH}^{-1}(D)) \geqslant 2$ and $\beta_p(\mathrm{PH}^{-1}(D)) = 1$ for some $p \geqslant 1$.

- For some spaces like the Klein bottle and the real projective plane whose homology with coefficients in $\mathbb{Z}_2$ differ from that with coefficients in $\mathbb{Q}$, the fibers $\mathrm{PH}^{-1}(D)$ strongly depend on the choice field. Namely, the number of polyhedra in the fibers, the dimensions of the facets and the Betti numbers are not the same whether the persistence module associated to D is computed with coefficients in $\mathbb{Z}_2$ or $\mathbb{Q}$.

- The dunce hat is contractible but some fibers have non-trivial 2-dimensional homology.

5.7 *Connection with Simple Homotopy Theory*

In this section we show that collapsibility of a complex K, which is a combinatorial and stronger notion of contractibility, is equivalent to the fiber $\mathrm{PH}^{-1}(D)$ over a well-chosen barcode D being non-empty. In particular we can use our algorithm for computing $\mathrm{PH}^{-1}(D)$ to determine whether K is collapsible.

Given τ, σ two simplices, $\tau \subseteq \sigma$ and $\dim \sigma = \dim \tau + 1$, such that σ is a maximal face of K and no other simplex contains τ, we say that τ is a *free face*. The operation of removing τ, σ is called an *elementary collapse*, and if $L := K \backslash \{\tau \subseteq \sigma\}$ is the resulting complex we write $K \searrow L$. Finally K is said to be *collapsible* if there is a sequence of elementary collapses from K to one of its vertices:

$$K = L_n \searrow L_{n-1} \searrow L_{n-2} \searrow \cdots \searrow L_1 \searrow L_0 = \{v\}.$$

Collapsibility implies contractibility but the reverse is false: the dunce hat and the house with two rooms are instances of contractible 2-complexes that are not collapsible. However we have the following well-known Zeeman's conjecture, appropriately phrased in [AB17][23] for simplicial complexes:

[23] Karim A Adiprasito and Bruno Benedetti. Subdivisions, shellability, and collapsibility of products. *Combinatorica*, 37(1):1–30, 2017

Conjecture (Zeeman in [Zee63])**.** *Let K be a contractible 2-complex. Then after taking finitely many barycentric subdivisions the product $K \times I$ is collapsible.*

This conjecture remains open and implies the 3-dimensional conjecture[24].

[24] E Christopher Zeeman. On the dunce hat. *Topology*, 2(4):341–358, 1963

Next we bridge the question of the collapsibility of a complex K to the fiber of PH over barcodes D that are *elementary*: those have 1 infinite bar (b_0, ∞) in dimension 0 followed by $\frac{\sharp K-1}{2}$ non-overlapping intervals (b_i, d_i), that is:

$$b_0 < b_1 < d_1 < b_2 < d_2 < \cdots < b_i < d_i < \cdots < b_{\frac{\sharp K-1}{2}} < d_{\frac{\sharp K-1}{2}}$$

Proposition 5.7.1. *Let K be a contractible complex. Then K is collapsible if and only if there exists an elementary barcode D with nonempty fiber, i.e. $\mathrm{PH}^{-1}(D) \neq \varnothing$.*

Proof. If K is collapsible let $K = L_n \searrow L_{n-1} \searrow L_{n-2} \searrow \cdots \searrow L_1 \searrow L_0 = \{v\}$ be a sequence of elementary collapses, with notations $n = \frac{\sharp K-1}{2}$ and $L_{i+1} = L_i \cup \{\tau_i \subseteq \sigma_i\}$, and define a filter f by $f(v) := 0$, $f(\tau_i) := 2i + 1$ and $f(\sigma_i) := 2i + 2$. Then the barcode of f is clearly elementary by definition of an elementary collapse.

Conversely, let $D = \{(b_0, \infty)\} \cup \{(b_i, d_i)\}_{1 \leqslant i \leqslant \frac{\sharp K - 1}{2}}$ be an elementary barcode and f a filter in the fiber, i.e. $\mathrm{PH}(f) = D$. Since D has exactly $\sharp K$ distinct endpoint values, f establishes a bijection $v \mapsto b_0$, $(\tau_i, \sigma_i) \mapsto (b_i, d_i)$, from simplices of K to these endpoints. In particular τ_n and then σ_n are the last two simplices to enter the sub-level set filtration of f, so that σ_n is a maximal face and no other simplex can contain τ_n. However σ_n itself contains τ_n because the $\dim \tau_n = \dim \sigma_n - 1 = \dim(b_n, d_n)$-cycle created by τ_n becomes a boundary when adding σ_n in the filtration. Thus removing τ_n and σ_n defines an elementary collapse, and we conclude by induction. □

5.8 *Adaptation for persistent (relative) (co)homology*

In addition to the homology functor, the relative homology, cohomology, and relative cohomology functors engender distinct persistence modules of their own, each of which determines a barcode and thus a new persistence map. We claim that the procedure described to compute the fiber of the persistent homology map in this work also suffices to compute the fibers of these other maps.

Let (C, E) be a based, finite-dimensional, $\Bbbk$-linear chain complex equipped with a filter $f : E \to \mathrm{I}$ that surjects onto a finite subset of the unit interval I, denoted $\Gamma = \{\gamma_0 < \cdots < \gamma_m\}$, where $\gamma_0 = 0$ and $\gamma_m < 1$. Write L_t for the linear span of $\{e \in E : f(e) \leqslant t\}$, which forms a subcomplex of C by hypothesis.

From these data we can construct four distinct sequences of vector spaces and homomorphisms, induced by either inclusion or quotient:

$$
\begin{array}{rccccccc}
\mathrm{H}_*(L): & & H_*(L_{\gamma_0}) & \longrightarrow & \cdots & \longrightarrow H_*(L_{\gamma_m}) & \longrightarrow & H_*(L_1) \\
\mathrm{H}^*(L): & & H^*(L_{\gamma_0}) & \longleftarrow & \cdots & \longleftarrow H^*(L_{\gamma_m}) & \longleftarrow & H^*(L_1) \\
\mathrm{H}_*(L_1, L): & H_*(L_1) \longrightarrow H_*(L_1, L_{\gamma_0}) & \longrightarrow & \cdots & \longrightarrow H_*(L_1, L_{\gamma_m}) & & & \\
\mathrm{H}^*(L_1, L): & H^*(L_1) \longleftarrow H^*(L_1, L_{\gamma_0}) & \longleftarrow & \cdots & \longleftarrow H^*(L_1, L_{\gamma_m}) & & &
\end{array}
$$

We refer to these as the homology, cohomology, relative homology, and relative cohomology persistence modules, respectively.

A classic result of [DSMVJ11][25] states that the barcode for $\mathrm{H}_*(L)$ uniquely determines the barcodes for $\mathrm{H}^*(L), \mathrm{H}_*(L_1, L)$, and $\mathrm{H}^*(L_1, L)$.

To compute the fiber of one of these other maps, therefore, one must simply convert the barcode to the associated PH barcode and apply any algorithm that is specialized to compute fibers for $\mathrm{H}_*(L)$. Barcodes in persistent homology can be converted into barcodes for the other three standard persistence modules as follows[26]

1. $\mathrm{H}_*(L)$ from $\mathrm{H}^*(L)$: no change

[25] Vin De Silva, Dmitriy Morozov, and Mikael Vejdemo-Johansson. Dualities in persistent (co) homology. *Inverse Problems*, 27(12):124003, 2011

[26] For full details see [DSMVJ11]. The authors of that work stipulate that (C, E) be the chain complex of a filtered CW complex, equipped with the standard basis of cells; however no proofs make use of this added restriction on (C, E), and the results are easily seen to hold for arbitrary based filtered complexes.

2. $\mathrm{H}_*(L)$ from $\mathrm{H}_*(L_1, L)$ or $\mathrm{H}^*(L_1, L)$: subtract 1 from the homology degree of each finite bar; replace each infinite bar of form $(-\infty, a)$ with $[a, \infty)$, leaving degree unchanged

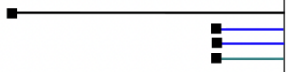
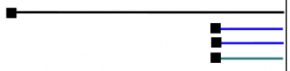
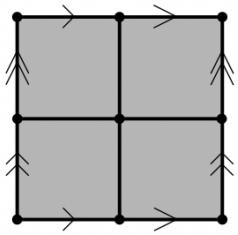
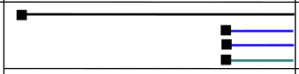
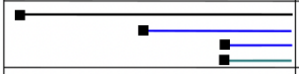
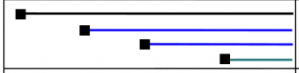
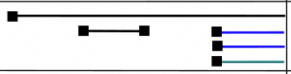
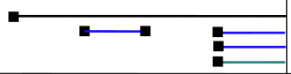
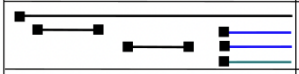
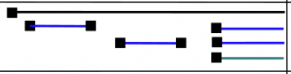
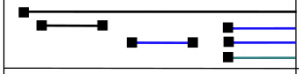
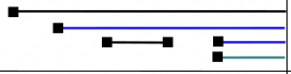
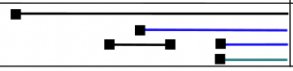
CW complex (Torus)	Barcode D	Fiber
		Polyhedra: [2] $\beta_0 = 2$
		Polyhedra: [12, 12] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [64, 448, 832, 576, 128] $\beta_0 = 1, \beta_1 = 2, \beta_2 = 1$ Facets: [0, 0, 0, 128, 128]
		Polyhedra: [1824, 19776, 77344, 146688, 146688, 74752, 15360] $\beta_0 = 1, \beta_1 = 4, \beta_2 = 5, \beta_3 = 2$
		Polyhedra: [1472, 8928, 20096, 21856, 11776, 2560] $\beta_0 = 6, \beta_1 = 8, \beta_2 = 2$ Facets: [0, 0, 0, 448, 0, 2560]
		Polyhedra: [18496, 153856, 505280, 852352, 786560, 378880, 74752] $\beta_0 = 6, \beta_1 = 14, \beta_2 = 10, \beta_3 = 2$
		Polyhedra: [2368, 6800, 5824, 1408] $\beta_0 = 1, \beta_1 = 17$ Facets: [0, 0, 1408, 1408]
		Polyhedra: [64, 256, 320, 128] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [2112, 3072, 896] $\beta_0 = 2, \beta_1 = 66$ Facets: [0, 896, 896]
		Empty Fiber
		Polyhedra: [1152, 2560, 1408] $\beta_0 = 8, \beta_1 = 8$
		Polyhedra: [2240, 7872, 10432, 6336, 1536] $\beta_0 = 8, \beta_1 = 16, \beta_2 = 8$ Facets: [0, 0, 320, 0, 1536]
		Polyhedra: [8320, 29824, 39936, 24320, 5888] $\beta_0 = 16, \beta_1 = 32, \beta_2 = 16$ Facets: [0, 0, 896, 0, 5888]

Figure 5.13: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the torus.

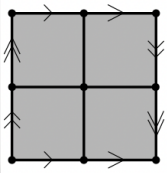
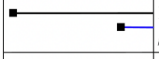
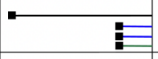
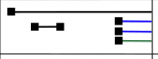
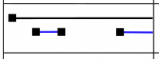
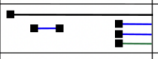
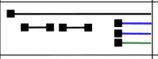
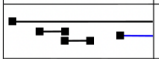
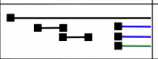
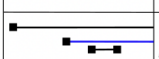
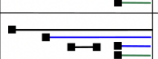
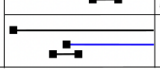
CW complex (Klein Bottle)	Barcode D Field: $\mathbb{Q}$	Fiber	Barcode D Field: $\mathbb{Z}_2$	Fiber
		Polyhedra: [768, 4768, 10848, 11744, 6176, 1280] $\beta_0 = 1, \beta_1 = 2, \beta_2 = 1$		Polyhedra: [64, 448, 832, 576, 128] $\beta_0 = 1, \beta_1 = 2, \beta_2 = 1$ Facets: [0, 0, 0, 128,128]
		Polyhedra: [13512, 52312, 74400, 46624, 11008] $\beta_0 = 1, \beta_1 = 17$		Polyhedra: [2368, 6800, 5824, 1408] $\beta_0 = 1, \beta_1 = 17$ Facets: [0, 0, 1408, 1408]
		Polyhedra: [192, 896, 1472, 1024, 256] $\beta_0 = 4, \beta_1 = 4$		Polyhedra: [64, 256, 320, 128] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [10656, 26784, 22592, 6528] $\beta_0 = 2, \beta_1 = 66$		Polyhedra: [2112, 3072, 896] $\beta_0 = 2, \beta_1 = 66$ Facets: [0, 896, 896]
		Polyhedra: [9120, 23008, 19456, 5632] $\beta_0 = 2, \beta_1 = 66$		Polyhedra: [1792, 2624, 768] $\beta_0 = 2, \beta_1 = 66$ Facets: [0, 768, 768]
		Empty Fiber		Empty Fiber
		Polyhedra: [3456, 9984, 9344, 2816] $\beta_0 = 8, \beta_1 = 8$		Polyhedra: [1152, 2560, 1408] $\beta_0 = 8, \beta_1 = 8$
		Polyhedra: [1088, 4000, 5472, 3328, 768] $\beta_0 = 2, \beta_1 = 4, \beta_2 = 2$		Polyhedra: [2080, 6688, 7776, 3936, 768] $\beta_0 = 6, \beta_1 = 12, \beta_2 = 6$ Facets: [0, 0, 160, 768, 768]
		Polyhedra: [3936, 14752, 20480, 12608, 2944] $\beta_0 = 4, \beta_1 = 8, \beta_2 = 4$		Polyhedra: [7616, 25088, 29632, 15104, 2944] $\beta_0 = 12, \beta_1 = 24, \beta_2 = 12$ Facets: [0, 0, 448, 2944, 2944]

Figure 5.14: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the Klein bottle, and when persistence modules are computed with coefficients in the field Q or $\mathbb{Z}_2$.

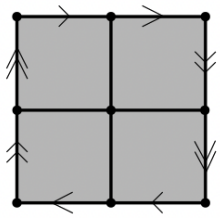
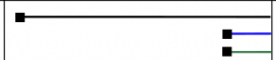
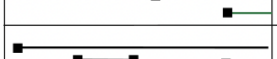
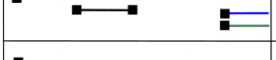
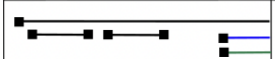
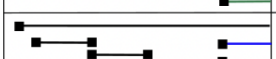
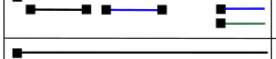
CW complex (Projective Plane)	Barcode D Field $\mathbb{Z}_2$	Fiber
		Polyhedra: [105, 1156, 3844, 5592, 3760, 960] $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1$
		Polyhedra: [3312, 46384, 238576, 620656, 908896, 762336, 342848, 64256] $\beta_0 = 1, \beta_1 = 2, \beta_2 = 2, \beta_3 = 1$
		Polyhedra: [10412, 52084, 91896, 69104, 18880] $\beta_0 = 1, \beta_1 = 3, \beta_2 = 2$
		Polyhedra: [408, 2336, 4664, 3952, 1216] $\beta_0 = 1, \beta_1 = 2, \beta_2 = 2, \beta_3 = 1$
		Polyhedra: [26400, 81288, 81648, 26816] $\beta_0 = 1, \beta_1 = 57$
		Polyhedra: [17232, 53816, 54384, 17856] $\beta_0 = 1, \beta_1 = 57$
		Empty Fiber
		Polyhedra: [12904, 45832, 53072, 20160] $\beta_0 = 4, \beta_1 = 20$
		Polyhedra: [25816, 205296, 656280, 1089408, 996288, 477888, 94208] $\beta_0 = 2, \beta_1 = 5, \beta_2 = 4, \beta_3 = 1$
		Polyhedra: [243592, 1443624, 3373600, 3895488, 2226752, 504832] $\beta_0 = 3, \beta_1 = 30, \beta_2 = 27$

Figure 5.15: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the real projective plane, and when persistence modules are computed with coefficients in the field $\mathbb{Z}_2$.

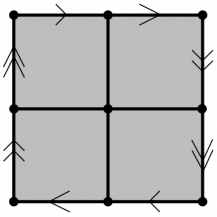
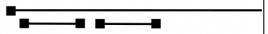
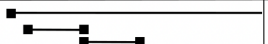
CW complex (Projective Plane)	Barcode D Field $\mathbb{Q}$	Fiber
		Polyhedra: [106, 1261, 5000, 9436, 9352, 4720, 960] $\beta_0 = 1$
		Polyhedra: [11348, 67224, 151436, 165432, 88752, 18880] $\beta_0 = 1, \beta_1 = 3, \beta_2 = 2$
		Polyhedra: [4000, 51104, 250784, 635840, 918080, 764512, 342848, 64256] $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 1$ Facets: [0, 0, 0, 0, 0, 2176, 0, 64256]
		Polyhedra: [31732, 122700, 176040, 111920, 26816] $\beta_0 = 1$, $\beta_1 = 33$
		Polyhedra: [20092, 78964, 114632, 73648, 17856] $\beta_0 = 1$, $\beta_1 = 33$
		Polyhedra: [4056, 33304, 109880, 188040, 177040, 87296, 17664] $\beta_0 = 4$, $\beta_1 = 8$, $\beta_2 = 4$
		Polyhedra: [2904, 24856, 84536, 148104, 142096, 71168, 14592] $\beta_0 = 4$, $\beta_1 = 8$, $\beta_2 = 4$

Figure 5.16: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the real projective plane, and when persistence modules are computed with coefficients in the field Q.

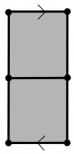
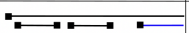
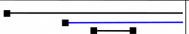
CW complex (Möbius Strip)	Barcode D	Fiber
		Polyhedra: [372, 2396, 5664, 6392, 3520, 768] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [7028, 27764, 40644, 26448, 6528] $\beta_0 = 1, \beta_1 = 13$
		Polyhedra: [96, 448, 736, 512, 128] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [5744, 14904, 13168, 4032] $\beta_0 = 1, \beta_1 = 25$
		Polyhedra: [7184, 18568, 16352, 4992] $\beta_0 = 1, \beta_1 = 25$
		Polyhedra: [936, 3492, 4908, 3120, 768] $\beta_0 = 8, \beta_1 = 8$
		Polyhedra: [216, 828, 1188, 768, 192] $\beta_0 = 4, \beta_1 = 4$
		Polyhedra: [1152, 2880, 2496, 768] $\beta_0 = 8, \beta_1 = 8$

Figure 5.17: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the Möbius strip.

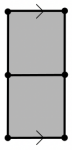
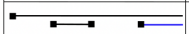
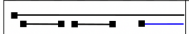
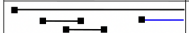
CW complex (Cylinder)	Barcode D	Fiber
		Polyhedra: [328, 2040, 4760, 5352, 2944, 640] $\beta_0 = 1, \beta_1 = 1$
		Polyhedra: [5612, 22528, 33520, 22112, 5504] $\beta_0 = 1, \beta_1 = 5$
		Polyhedra: [96, 448, 736, 512, 128] $\beta_0 = 2, \beta_1 = 2$
		Polyhedra: [4432, 11776, 10592, 3264] $\beta_0 = 2, \beta_1 = 18$
		Polyhedra: [5712, 15104, 13536, 4160] $\beta_0 = 2, \beta_1 = 18$
		Polyhedra: [1904, 7056, 9760, 6080, 1472] $\beta_0 = 4, \beta_1 = 8, \beta_2 = 4$
		Polyhedra: [496, 1824, 2528, 1584, 384] $\beta_0 = 2, \beta_1 = 4, \beta_2 = 2$
		Polyhedra: [1120, 2840, 2480, 768] $\beta_0 = 4, \beta_1 = 12$

Figure 5.18: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the cylinder.

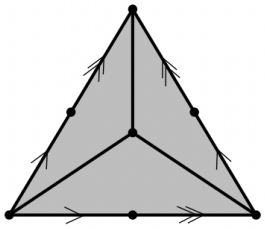
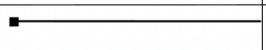
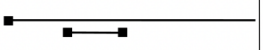
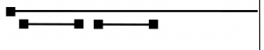
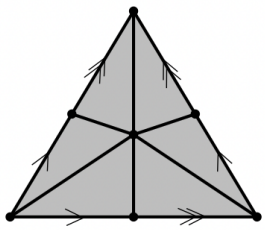
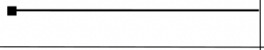
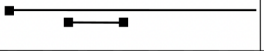
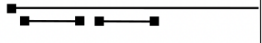
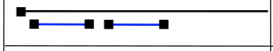
CW complex (Dunce Hat)	Barcode D	Fiber
		Polyhedra: [15, 54, 64, 24] $\beta_0 = 1$
		Polyhedra: [120, 220, 96] $\beta_0 = 2, \beta_1 = 6$
		Polyhedra: [163, 790, 1356, 1020, 288] $\beta_0 = 1, \beta_1 = 4$
		Polyhedra: [34, 24] $\beta_0 = 10$
		Polyhedra: [34, 24] $\beta_0 = 10$
		Empty Fiber
		Empty Fiber
		Polyhedra: [796, 2370, 2436, 864] $\beta_0 = 2, \beta_1 = 4$
		Polyhedra: [288, 828, 828, 288] $\beta_0 = 8, \beta_1 = 10, \beta_2 = 2$
		Polyhedra: [41, 243, 515, 508, 244, 48] $\beta_0 = 1$ Facets: [0, 0, 0, 20, 24, 48]
		Polyhedra: [500, 1518, 1720, 904, 192] $\beta_0 = 1, \beta_1 = 11$ Facets: [0, 0, 80, 96, 192]
		Polyhedra: [1956, 19600, 81350, 183396, 242534, 188044, 79216, 14016] $\beta_0 = 1, \beta_1 = 1$ Facets: [0, 0, 0, 144, 480, 0, 0, 14016]
		Polyhedra: [152, 272, 184, 48] $\beta_0 = 16$ Facets: [0, 20, 24, 48]
		Polyhedra: [152, 272, 184, 48] $\beta_0 = 16$ Facets: [0, 20, 24, 48]
		Polyhedra: [3088, 21522, 63178, 99484, 88132, 41456, 8064] $\beta_0 = 4, \beta_1 = 4$ Facets: [0, 0, 0, 96, 0, 0, 8064]
		Polyhedra: [504, 2796, 6324, 7248, 4176, 960] $\beta_0 = 4, \beta_1 = 4$

Figure 5.19: Some statistics about fibers $\mathrm{PH}^{-1}(D)$ when K is a CW decomposition of the Dunce Hat.

6
Fiber of Persistent Homology on Morse functions

Abstract

Let f be a Morse function on a smooth compact manifold M with boundary. The path component $\mathrm{PH}_f^{-1}(D)$ containing f of the space of Morse functions giving rise to the same Persistent Homology $D = \mathrm{PH}(f)$ is shown to be the same as the orbit of f under pre-composition $\phi \mapsto f \circ \phi$ by diffeomorphisms of M which are isotopic to the identity. Consequently we derive topological properties of the fiber $\mathrm{PH}_f^{-1}(D)$: In particular we compute its homotopy type for many compact surfaces M. In the 1-dimensional settings where M is the unit interval or the circle we extend the analysis to continuous functions and show that the fibers are made of contractible and circular components respectively.

6.1 Introduction

Persistent Homology is a central and computable descriptor in Topological Data Analysis (TDA) which has been applied to a large variety of data science problems. Namely the persistence map PH associates to a real-valued function f on a topological space X a so-called barcode that captures the topological variations of its sub level-sets [1]. It is natural to ask how much information can be recovered from persistent homology: Given a barcode D what can we say about the fiber $\mathrm{PH}^{-1}(D)$?

In this note $X = M$ is an arbitrary smooth (finite-dimensional) compact manifold with boundary ∂M and $f : M \to \mathbb{R}$ is a Morse function, i.e. f has a prescribed constant value a_j on each boundary component of ∂M_j and has isolated critical points none of which belong to ∂M. Denote by $D := \mathrm{PH}(f)$ the associated barcode. Pre-composing f with isotopies $\phi \mapsto f \circ \phi$, we get an orbit $\mathcal{O}_{\mathrm{Id}}(f)$ inside the space of Morse functions. Our core contribution (Theorem 6.3.1) is the equality between this orbit and the the path connected component $\mathrm{PH}_f^{-1}(D)$ containing f in the fiber of PH over D:

$$\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f).$$

This result crucially relies on Mather's fibration theorem for smooth mappings [2], which we slightly adapt to the case of Morse functions with equal critical values using results of Cerf [3].

We can then put at work the abundant literature about the homotopy type of the orbit $\mathcal{O}_{\mathrm{Id}}(f)$, especially the work of Maksymenko [4]: the mapping $\phi \mapsto f \circ \phi$ in fact defines a locally trivial fibration from isotopies $\mathcal{D}_{\mathrm{Id}}(M)$ to the orbit $\mathcal{O}_{\mathrm{Id}}(f)$, with fiber $\mathcal{S}_{\mathrm{Id}}(f)$ the isotopies stabilising f, i.e. $f \circ \phi = f$. Hence a long exact sequence links the fiber $\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f)$ to the well-studied diffeomorphism groups of the manifold M. In particular for any compact surface M, we compute $\pi_n(\mathrm{PH}_f^{-1}(D))$ for $n \geqslant 2$ (Proposition 6.4.5), and if furthermore D has distinct interval endpoints we derive the complete homotopy type of the fiber $\mathrm{PH}_f^{-1}(D)$ (Propositions 6.4.7 and 6.4.8).

Variations of this setting have already been addressed. In the discrete setting where $X = K$ is a finite simplicial complex and f is compatible with face inclusions, the fiber $\mathrm{PH}^{-1}(D)$ is a complex of polyhedra [5]. In the restricted case where K is a line complex, each path connected component of $\mathrm{PH}^{-1}(D)$ is contractible [6], and it is homeomorphic to a circle in the case where K is a subdivision of the unit circle [7].

In the analogous, continuous 1-dimensional setting where X is the interval or the circle and f is continuous, each component in the fiber is contractible and circular respectively as we show in the appendix. For the unit interval it is possible to count the number of path connected components in $\mathrm{PH}^{-1}(D)$ by means of the combinatorics of the barcode [8].

[1] Herbert Edelsbrunner and John Harer. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008; and Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005

[2] John N Mather. Stability of C^∞ mappings: II. infinitesimal stability implies stability. *Annals of Mathematics*, pages 254–291, 1969

[3] Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970

[4] Sergiy Maksymenko. Homotopy types of stabilizers and orbits of Morse functions on surfaces. *Annals of Global Analysis and Geometry*, 29(3):241–285, 2006

[5] Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021

[6] Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020

[7] Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021

[8] Justin Curry. The fiber of the persistence map for functions on the interval. *Journal of Applied and Computational Topology*, 2(3-4):301–321, 2018

For higher dimensional X analyzing the fiber is a challenging problem: already for Morse functions on the 2-sphere $X = \mathbb{S}^2$ new tools have been designed to describe the fiber $\mathrm{PH}^{-1}(D)$, and allowed for conjectures on the number of path connected components [9]. However the higher dimensional homotopy groups of $\mathrm{PH}^{-1}(D)$ remain unknown, from which stems the motivation of this work.

[9] Michael J Catanzaro, Justin M Curry, Brittany Terese Fasy, Jānis Lazovskis, Greg Malen, Hans Riess, Bei Wang, and Matthew Zabka. Moduli spaces of Morse functions for persistence. *Journal of Applied and Computational Topology*, pages 1–33, 2020

6.2 *Stability of Morse functions*

We fix a d-dimensional compact smooth manifold M with boundary ∂M, whose path connected components are denoted by ∂M_j. Given another smooth manifold $\mathcal{N}$, we denote by $C^\infty(M, \mathcal{N})$ the space of smooth maps from M to $\mathcal{N}$ equipped with the C^∞ Whitney topology. We denote by $\mathcal{D}(M) \subseteq C^\infty(M, M)$ the diffeomorphisms of M, and by $\mathcal{D}_{\mathrm{Id}}(M)$ its subspace of isotopies, i.e. the path connected component of the identity map on M.

Given real values a_j, a smooth map f belongs to the space $\mathrm{Morse}(M) \subseteq C^\infty(M, \mathbb{R})$ of *Morse functions* on M if:

- The Hessian of f is non-degenerate at critical points, all of which belong to $M \backslash \partial M$; and
- The restrictions $f_{|\partial M_j}$ to each boundary component ∂M_j are constant with prescribed value a_j.

Then $\mathcal{D}(M)$ acts on $C^\infty(M, \mathbb{R})$ by pre-composition and we denote by $\mathcal{O}(f) \subseteq \mathrm{Morse}(M)$ the orbit of f, and by $\mathcal{O}_{\mathrm{Id}}(f) \subseteq \mathcal{O}(f)$ the orbit of f under the restricted action of isotopies. Next we adapt Mather's stability of smooth mappings [Mat69][10] to the case of Morse functions with equal critical points by combining results of Cerf:

[10] John N Mather. Stability of C^∞ mappings: II. infinitesimal stability implies stability. *Annals of Mathematics*, pages 254–291, 1969

Proposition 6.2.1. *Let $f \in \mathrm{Morse}(M)$ and $\mathrm{Morse}_f(M)$ be the subspace of Morse functions with the same critical values as f. Then there exists a neighborhood $\mathcal{U} \subseteq \mathrm{Morse}_f(M)$ of f contained in the orbit $\mathcal{O}_{\mathrm{Id}}(f)$, that is:*

$$\forall g \in \mathcal{U},\ \exists \phi \in \mathcal{D}_{\mathrm{Id}}(M),\ g = f \circ \phi. \tag{6.1}$$

Proof. From [Cer70, Appendix, § 1, Proposition 1] the statement holds for a neighborhood $\mathcal{U}_f$ of f in the subspace $\mathrm{Morse}_f(M; \mathrm{Crit}(f), \partial M) \subseteq \mathrm{Morse}_f(M)$ of functions with the same critical points $p_1, \cdots, p_n$ as f and derivatives $\partial f^k_{|\partial M}$ of any order k equal those of f on the boundary ∂M:

Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970

$$\forall g \in \mathcal{U}_f,\ \exists \phi_g \in \mathcal{D}_{\mathrm{Id}}(M),\ g = f \circ \phi_g. \tag{6.2}$$

For the general case where critical points and derivatives on the boundary are allowed to vary we simply find a diffeomorphism sending them back to $p_1, \cdots, p_n$ and $\partial f^k_{|\partial M}$ and apply the above result.

Namely, from [Cer61, Theorem 5] the action of isotopies on embeddings defines a locally trivial fibration over the space of orbits: in particular we can find disjoint neighborhoods $U_1, \cdots, U_n$ of $p_1, \cdots, p_n$ and continuously associate to $(p'_1, \cdots, p'_n) \in U_1 \times \cdots \times U_n$ a diffeomorphism $\psi_{(p'_1,\cdots,p'_n)} \in \mathcal{D}_{\mathrm{Id}}(M)$ such that $\psi_{(p_1,\cdots,p_n)} = \mathrm{Id}$ and:

$$\forall (p'_1, \cdots, p'_n) \in U_1 \times \cdots \times U_n, \ \ \forall 1 \leqslant i \leqslant n, \ \ \psi_{(p'_1,\cdots,p'_n)}(p_i) = p'_i.$$

Let $\mathcal{U} \subseteq \mathrm{Morse}_f(M)$ be a neighborhood of f for which any $g \in \mathcal{U}$ has critical points $\mathrm{Crit}(g)$ in $U_1, \cdots, U_n$. In particular in this case $g \circ \psi_{\mathrm{Crit}(g)}$ has the same critical points $p_1, \cdots, p_n$ as f, so it remains to deal with the boundary ∂M.

Let ∂M_j be a boundary component. By flowing along the normalized gradient of f (or its inverse) from the boundary ∂M_j we get a collar $V_j \cong \partial M_j \times [0, \alpha)$ that is adapted to f in the sense that $f(x,t) = a_j \pm t$, w.l.o.g. $f(x,t) = a_j + t$. For g in a small neighborhood $\mathcal{V} \subseteq \mathrm{Morse}_f(M)$ of f, the derivative of g is non-zero in the normal direction over V_j, and we have $g(\partial M_j \times [0, \frac{\alpha}{2})) \subseteq [a_j, a_j + \alpha)$. Therefore we can continuously associate to $g \in \mathcal{V}$ the embedding $\iota_g : (x,t) \in \partial M_j \times [0, \frac{\alpha}{2}) \mapsto (x, g(x,t) - a_j) \in V_j$ that preserves ∂M_j, is transverse to ∂M_j and satisfies $g = f \circ \iota_g$ on $\partial M_j \times [0, \frac{\alpha}{2})$. Therefore by using again the local triviality of embeddings that preserve and are transverse to the boundary (see [Cer61, Theorem 5]), up to shrinking $\mathcal{V}$, we can extend ι_g to a diffeomorphism χ_g. Moreover the extensions χ_g can be assumed to induce the identity outside V_j. By repeating this process for each boundary component ∂M_j, we can continuously associate to $g \in \mathcal{V}$ a diffeomorphism χ_g such that g and $f \circ \chi_g$ agrees on a collar neighborhood of the boundary.

Jean Cerf. Topologie de certains espaces de plongements. *Bulletin de la Société Mathématique de France*, 89:227–380, 1961

By reducing the neighorhoods U_i and V_j to avoid overlaps, we have that for any g in $\mathcal{U} \cap \mathcal{V}$ the Morse function $g \circ \psi_{\mathrm{Crit}(g)} \circ \chi^{-1}_{g \circ \psi_{\mathrm{Crit}(g)}}$ has the same critical points as f and agrees with f on a neighborhood of the boundary ∂M, in particular it belongs to $\mathrm{Morse}_f(M; \mathrm{Crit}(f), \partial M)$. Hence by Eq. (6.2), up to shrinking $\mathcal{U} \cap \mathcal{V}$, we conclude that:

$$\begin{aligned} \forall g \in \mathcal{U} \cap \mathcal{V}, \ g &= (g \circ \psi_{\mathrm{Crit}(g)} \circ \chi^{-1}_{g \circ \psi_{\mathrm{Crit}(g)}}) \circ \chi_{g \circ \psi_{\mathrm{Crit}(g)}} \circ \psi^{-1}_{\mathrm{Crit}(g)} \\ &= f \circ \phi_{g \circ \psi_{\mathrm{Crit}(g)} \circ \chi^{-1}_{g \circ \psi_{\mathrm{Crit}(g)}}} \circ \chi_{g \circ \psi_{\mathrm{Crit}(g)}} \circ \psi^{-1}_{\mathrm{Crit}(g)}. \end{aligned}$$ □

6.3 *Covering the fiber with isotopies*

Given $f \in \mathrm{Morse}(M)$, we get a nested sequence of sub level-sets $f^{-1}((-\infty, x])$. In turn, by applying homology in degree $0 \leqslant k \leqslant d$ with coefficients in an arbitrary field, we get the *persistent homology module of* f: the sequence of vector spaces $\mathrm{H}_k(f^{-1}(-\infty, x])$ with linear maps between them induced by inclusions, in other words a functor from the poset $(\mathbb{R}, \leqslant)$

to finite dimensional vector spaces. The *barcode* of f in degree k is the isomorphism class of this functor up to natural isomorphism. From [CB15] any such functor uniquely decomposes as a direct sum of functors $\bigoplus_{(b,d)\in D} \mathbb{I}_{[b,d)}$, with $[b,d) \subseteq \mathbb{R}$ an interval closed on the left and open on the right (hence possibly $d = +\infty$): each $\mathbb{I}_{[b,d)}$ consists of 1-dimensional vector spaces linked with identity maps on $[b,d)$, and it is the zero vector space everywhere outside of $[b,d)$. Therefore the barcode of f, denoted by $\mathrm{PH}_k(f)$, can be equivalently described as the multi-set D of pairs (b,d) indexing this decomposition, and will be described in this way in the rest of this document. By abuse of terminology we refer to pairs (b,d) as *intervals* or *bars* of the barcode $D = \mathrm{PH}_k(f)$. Intuitively (b,d) corresponds to the appeareance of a k-cycle in $f^{-1}((-\infty,b])$ that is further cancelled in $f^{-1}((-\infty,d])$ (or persists forever if $d = \infty$). We refer to [EH08, ZC05] for extensive treatments of the theory of Persistence.

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

Herbert Edelsbrunner and John Harer. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008; and Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005

In this work the persistence map is defined on Morse functions and returns the $d+1$ barcodes of interest:

$$\mathrm{PH} : f \in \mathrm{Morse}(M) \longmapsto [\mathrm{PH}_0(f), \cdots, \mathrm{PH}_d(f)] \in \mathbf{Bar}^{d+1}.$$

We assume that **Bar** is equipped with its natural *bottleneck metric* which turns PH into a continuous map by the Stability Theorem [11]. Given a barcode D and a Morse function $f \in \mathrm{Morse}(M)$ such that $\mathrm{PH}(f) = D$, we denote by $\mathrm{PH}_f^{-1}(D)$ the path connected component of the fiber $\mathrm{PH}^{-1}(D) \subseteq \mathrm{Morse}(M)$ containing f.

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

Theorem 6.3.1. *Let D be a barcode and $f \in \mathrm{PH}^{-1}(D)$. Then $\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f)$.*

Proof. Let $(\phi_t)_{0\leqslant t\leqslant 1}$ be a path in $\mathcal{D}_{\mathrm{Id}}(M)$. Each ϕ_t restricts to a homeomorphism between the sub level-sets of $f\circ\phi_t$ and f, hence it induces an isomorphism between the associated persistent homology modules. In turn $\mathrm{PH}(f\circ\phi_t) = \mathrm{PH}(f)$, so that $(f\circ\phi_t)_{0\leqslant t\leqslant 1}$ is a path in the fiber $\mathrm{PH}_f^{-1}(D)$, which implies $\mathcal{O}_{\mathrm{Id}}(f) \subseteq \mathrm{PH}_f^{-1}(D)$.

Conversely let $g \in \mathrm{PH}_f^{-1}(D)$ and let $(f_t)_{0\leqslant t\leqslant 1}$ be a path in the fiber $\mathrm{PH}_f^{-1}(D)$ joining f to g, thus $\mathrm{PH}(f_t) = D$ for each t. As is well-known, when M has no boundary there is a one-to-one correspondence between the set $\mathcal{D}$ of (bounded) interval endpoints in the barcode and the set $\mathcal{C}$ of critical values (counted with multiplicity) for Morse functions, e.g. because the associated persistent homology module and Morse-Smale complex are isomorphic [12].

[12] Serguei Barannikov. The framed Morse complex and its invariants. *American Mathematical Society Translations, Series 2*, 1994

When M has a boundary the correspondence adapts by adding in $\mathcal{C}$ the value a_j with multiplicity $\sum_i \beta_i(\partial M_j)$ for each boundary component ∂M_j that is a local minimum. Note that a Morse function is constant on ∂M_j and has no critical points there, so either it has ∂M_j as

a local minimum or as a local maximum, and this choice is fixed inside a path connected component of $\mathrm{Morse}(M)$.

Therefore each f_t has the same critical values as f, because the barcode $\mathrm{PH}(f_t) = D$ is constant. By Proposition 6.2.1, each $t \in [0;1]$ has a neighborhood $\mathrm{I}_t \subseteq [0,1]$ such that f_h can be written $f_h = f_t \circ \phi_h$ whenever $h \in \mathrm{I}_t$. By compactness $[0;1]$ is covered by finitely many such intervals. Therefore $g = f_1$ equals $f \circ \phi$, here ϕ is a finite composition of isotopies hence is itself an isotopy. Consequently $\mathrm{PH}_f^{-1}(D) \subseteq \mathcal{O}_{\mathrm{Id}}(f)$. □

6.4 *Topological properties of the fiber*

We derive direct consequences of Theorem 6.3.1 combined with the extensive study of $\mathcal{O}_{\mathrm{Id}}(f)$ by Maksymenko[13]. Denote by $\mathcal{S}_{\mathrm{Id}}(f)$ the subspace of isotopies ϕ preserving a Morse function f, i.e. $f \circ \phi = f$.[14]

[13] Sergiy Maksymenko. Homotopy types of stabilizers and orbits of Morse functions on surfaces. *Annals of Global Analysis and Geometry*, 29(3):241–285, 2006

[14] In [Mako6] the notation $\mathcal{S}_{\mathrm{Id}}(f)$ rather stands for the space of diffeomorphisms ϕ preserving f that are isotopic to Id_M *though* maps preserving f, thus it is the path connected component of Id_M in our $\mathcal{S}_{\mathrm{Id}}(f)$.

Proposition 6.4.1. *Assume that M is connected. Let D be a barcode and $f \in \mathrm{PH}^{-1}(D)$. Then the action of $\mathcal{D}_{\mathrm{Id}}(M)$ on $\mathrm{PH}_f^{-1}(D)$ defines a principal $\mathcal{S}_{\mathrm{Id}}(f)$-bundle, in particular a locally trivial fibration. Hence a homotopy long exact sequence:*

$$\cdots \to \pi_n(\mathcal{S}_{\mathrm{Id}}(f)) \to \pi_n(\mathcal{D}_{\mathrm{Id}}(M)) \to \pi_n(\mathrm{PH}_f^{-1}(D)) \to \pi_{n-1}(\mathcal{S}_{\mathrm{Id}}(f)) \to \cdots \to \pi_0(\mathcal{D}_{\mathrm{Id}}(M)).$$

Proof. From [Mako6, Theorem 2.1] the action of $\mathcal{D}_{\mathrm{Id}}(M)$ on $\mathcal{O}_{\mathrm{Id}}(f)$ defines a principal $\mathcal{S}_{\mathrm{Id}}(f)$-bundle, and the fiber $\mathrm{PH}_f^{-1}(D)$ equals the orbit $\mathcal{O}_{\mathrm{Id}}(f)$ by Theorem 6.3.1. □

Consequently for $M = \mathrm{S}^1$ the path components of the fiber $\mathrm{PH}^{-1}(D)$ are homotopy equivalent to a circle:

Proposition 6.4.2. *Assume $M = \mathrm{S}^1$. Let D be a barcode and $f \in \mathrm{PH}^{-1}(D)$. Then $\mathrm{PH}_f^{-1}(D)$ is homotopy equivalent to S^1.*

Proof. From Proposition 6.4.1 $\mathrm{PH}_f^{-1}(D)$ is homeomorphic to $\mathcal{D}_{\mathrm{Id}}(\mathrm{S}^1)/\mathcal{S}_{\mathrm{Id}}(f)$. Let n be the number of minima of f, which is then also the number of maxima of f because $\chi(\mathrm{S}^1) = 0$. Without loss of generality we assume that the associated $2n$ critical points of f are evenly spaced on S^1. The space $\mathcal{D}_{\mathrm{Id}}(\mathrm{S}^1)$ of isotopies of the circle deformation retracts to S^1, i.e. the rotations of the circle. The subgroup $\mathcal{S}_{\mathrm{Id}}(f)$ of isotopies ϕ preserving f, that is $f \circ \phi = f$, is then (isomorphic to) the subgroup of rotations consisting of the $2n$-th roots of unity that preserve the sequence of extremal values of f. The result follows since the quotient of S^1 by a finite subgroup is again S^1. □

When $M = [0,1]$ recall that Morse functions have prescribed values a_0 and a_1 on the boundary points 0 and 1.

Proposition 6.4.3. *Assume $M = [0,1]$. Let D be a barcode and $f \in \mathrm{PH}^{-1}(D)$. Then $\mathrm{PH}_f^{-1}(D)$ is contractible.*

Proof. From Proposition 6.4.1 $\mathrm{PH}_f^{-1}(D)$ is homeomorphic to $\mathcal{D}_{\mathrm{Id}}([0,1])/\mathcal{S}_{\mathrm{Id}}(f)$. However $\mathcal{D}_{\mathrm{Id}}([0,1])$ deformation retracts on the identity diffeomorphism $\mathrm{Id}_{[0,1]}$ by straight-line interpolations, and $\mathcal{S}_{\mathrm{Id}}(f) = \{\mathrm{Id}_{[0,1]}\}$. □

Note that we could easily derive a similar statement for Morse functions on $[0,1]$ without boundary conditions. In the appendix we prove the analogues of Propositions 6.4.2 and 6.4.3 for continuous functions. The analogues for lower-star filtrations on the subdivided interval and circle have been proved in [CMW20] and [MW21] respectively.

Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020; and Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021

Remark 6.4.4. When $M = M_1 \sqcup M_2$ has more than one connected component, the path component $\mathrm{PH}_f^{-1}(D)$ in the fiber over $D = \mathrm{PH}(f)$ can be retrieved as the product of the path components of the fibers over $D_1 := \mathrm{PH}(f_{|M_1})$ and $D_2 := \mathrm{PH}(f_{|M_2})$ containing the restrictions $f_{|M_1}$ and $f_{|M_2}$ respectively:

$$\mathrm{PH}_f^{-1}(D) = \mathrm{PH}_{f_{|M_1}}^{-1}(D_1) \times \mathrm{PH}_{f_{|M_2}}^{-1}(D_2).$$

For this reason we focus our analysis to the interesting case where M is connected.

For the rest of the section we fix a compact connected surface M and a function f in the fiber over D, whose number of critical points of index 1 is denoted by c_1. We make use of the analysis of the orbit $\mathcal{O}_{\mathrm{Id}}(f)$ from [Mako6].

Proposition 6.4.5. *Assume that $c_1 > 0$. Then $\pi_n(\mathrm{PH}_f^{-1}(D)) = \pi_n(M)$ for $n \geqslant 3$ and $\pi_2(\mathrm{PH}_f^{-1}(D)) = 0$.*

Proof. $\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f)$ by Theorem 6.3.1, and by [Mako6, (2), Theorem 1.5] we have $\pi_n(\mathcal{O}_{\mathrm{Id}}(f)) = \pi_n(M)$ for $n \geqslant 3$ and $\pi_2(\mathcal{O}_{\mathrm{Id}}(f)) = 0$. □

Remark 6.4.6. From [Mako6, (2), Theorem 1.5] we can also derive a short exact sequence $0 \to \pi_1(\mathcal{D}_{\mathrm{Id}}(M)) \oplus \mathbb{Z}^{k_f} \to \pi_1(\mathrm{PH}_f^{-1}(D)) \to G \to 0$ where G is a finite group and the integer $k_f \geqslant 0$ depends on the component $\mathrm{PH}_f^{-1}(D)$ in the fiber, on the number c_1 of saddles and the surface M.

Proposition 6.4.7. *Assume that $c_1 = 0$. Then the homotopy type of the fiber $\mathrm{PH}_f^{-1}(D)$ is classified as follows:*

Surface M	$\mathbb{S}^2$	$\mathbb{S}^1 \times I$	$\mathbb{D}^2$
Fiber $\mathrm{PH}_f^{-1}(D)$	$\mathbb{S}^2$	$\{*\}$	$\{*\}$

Proof. $\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f)$ by Theorem 6.3.1, and the homotopy type of $\mathcal{O}_{\mathrm{Id}}(f)$ is computed in [Mako6, Theorem 1.9]. □

For instance the case where $f : \mathbb{S}^2 \to \mathbb{R}$ has no saddle ($c_1 = 0$) can be interpreted as follows: The fiber sequence $\mathcal{S}_{\mathrm{Id}}(f) \to \mathcal{D}_{\mathrm{Id}}(\mathbb{S}^2) \to \mathrm{PH}_f^{-1}(D)$ of Proposition 6.4.1 can be identified up to homotopy with the standard fiber sequence $\mathbb{S}^1 \to \mathrm{SO}(3) \to \mathbb{S}^2$. This is because $\mathcal{D}_{\mathrm{Id}}(\mathbb{S}^2)$ deformation retracts to $\mathrm{SO}(3)$ by [Sma59], and if without loss of generality we assume that f is the standard height function then $\mathcal{S}_{\mathrm{Id}}(f)$ consists of those rotations fixing the poles, so that it is fixed through the deformation retraction and $\mathcal{S}_{\mathrm{Id}}(f) \sim \mathbb{S}^1$.

Stephen Smale. Diffeomorphisms of the 2-sphere. *Proceedings of the American Mathematical Society*, 10(4):621–626, 1959

Proposition 6.4.8. *Assume that D has pairwise distinct bounded interval endpoints, and that $c_1 > 0$. Then we have the following homotopy types for the fiber $\mathrm{PH}_f^{-1}(D)$:*

Surface M	$\mathbb{S}^2$	*Projective Plane*	*Torus*	$\mathbb{S}^1 \times I$	$\mathbb{D}^2$
Fiber $\mathrm{PH}_f^{-1}(D)$	$\mathrm{SO}(3) \times (\mathbb{S}^1)^{c_1-1}$	$\mathrm{SO}(3) \times (\mathbb{S}^1)^{c_1-1}$	$(\mathbb{S}^1)^{c_1+1}$	$(\mathbb{S}^1)^{c_1}$	$(\mathbb{S}^1)^{c_1}$

When M is obtained from the surfaces in the above tables by removing finitely many 2-disks, then $\mathrm{PH}_f^{-1}(D) \sim (\mathbb{S}^1)^{c_1-1}$. If M is the Möbius strip, then $\mathrm{PH}_f^{-1}(D) \sim (\mathbb{S}^1)^{c_1}$. For other orientable surfaces M, we have $\mathrm{PH}_f^{-1}(D) \sim (\mathbb{S}^1)^{c_1+\chi(M)}$. For the remaining non-orientable surfaces, we have $\mathrm{PH}_f^{-1}(D) \sim (\mathbb{S}^1)^{k_f}$ for some integer $k_f \leqslant c_1 + \chi(M)$, unless M is the Klein bottle in which case $k_f \leqslant c_1 + 1$.

Proof. $\mathrm{PH}_f^{-1}(D) = \mathcal{O}_{\mathrm{Id}}(f)$ by Theorem 6.3.1. Since D has distinct bounded interval endpoints, f has distinct critical points, and then the homotopy type of $\mathcal{O}_{\mathrm{Id}}(f)$ is computed in [Mako6, (2)&(3), Theorem 1.5]. □

Remark 6.4.9. When M has no boundary, $\partial M = \varnothing$, the number c_1 of saddles of Morse functions f in the fiber $\mathrm{PH}^{-1}(D)$ can be directly inferred from the barcode D. Namely, if we denote by k_D the number of intervals in D, then the quantity $\mathrm{k}_D - \beta_0 - \beta_2$ counts (i) all the intervals (b, d) of D in degree 1, which correspond by their birth value b to saddle points of f whose attaching handle increases the 1-dimensional homology of the sub level-set $f^{-1}((-\infty, b])$, and (ii) all the bounded intervals (b, d) of D in degree 0, which correspond by their death value $d < \infty$ to saddle points of f whose attaching handle decreases the 0-dimensional homology of the sub level-set $f^{-1}((-\infty, d])$. Hence $c_1 = \mathrm{k}_D - \beta_0 - \beta_2$. When $\partial M = \bigsqcup_j \partial M_j \neq \varnothing$, we can partition the boundary components ∂M_j into the sets $\partial M^{\min}$ (resp. $\partial M^{\max}$) of components ∂M_j that are local minimum (resp. maximum) of one (hence any) function f in the component of $\mathrm{PH}^{-1}(D)$ at stake. Since M

is a surface each ∂M_j is a circle, therefore if $\partial M_j \subseteq \partial M^{\min}$, then it corresponds in the barcode D to the births of one interval in degree 0 and one interval in degree 1. Otherwise $\partial M_j \subseteq \partial M^{\max}$ induces no topological change when entering the sub level-sets of f. Consequently the correspondence between critical points and interval endpoints adapts and yields $c_1 = \mathrm{k}_D - \beta_0 - \beta_2 - \#\partial M^{\min}$.

6.5 *Fiber of Persistent Homology for continuous maps on the circle and on the interval*

In this section the domain of the persistence map consists of continuous maps on the circle:

$$\mathrm{PH} : \mathcal{C}^0(\mathbb{S}^1, \mathbb{R}) \longrightarrow \mathbf{Bar}^2.$$

Note that in the codomain we record the two barcodes with non-trivial homology, those in degree 0 and 1. In fact the second barcode contains a unique unbounded interval starting at the maximum of the function on the circle.

We fix a barcode D with finitely many intervals. When $f = \underline{\mathrm{cst}}$ is constant it forms the fiber by itself over the trivial barcode $D = \mathrm{PH}(f)$ with only two infinite bars $(\underline{\mathrm{cst}}, +\infty)$, one in each degree 0 and 1. Other barcodes such that $\mathrm{PH}^{-1}(D) \neq \varnothing$ have one infinite interval $(b_0, +\infty)$ in degree 0, one infinite interval $(b_1, +\infty)$ with $b_0 < b_1$ in degree 1, finitely many bounded intervals in degree 0 with endpoints in $[b_0, b_1]$, and not other intervals. In the rest of this section we assume that D is non-trivial and denote by $(n-1)$, for some $n \geqslant 1$, its number of bounded intervals in degree 0.

Let $\mathrm{Aut}_{\leqslant}(\mathbb{S}^1)$ be the space of orientation-preserving homeomorphisms of the circle, and $\mathrm{End}_{\leqslant}(\mathbb{S}^1) = \overline{\mathrm{Aut}_{\leqslant}(\mathbb{S}^1)}$ be its closure in $\mathcal{C}^0(\mathbb{S}^1, \mathbb{R})$ in the compact-open topology. Given $f \in \mathcal{C}^0(\mathbb{S}^1, \mathbb{R})$ we have the pre-composition map $\phi \in \mathrm{End}_{\leqslant}(\mathbb{S}^1) \mapsto f \circ \phi \in \mathcal{C}^0(\mathbb{S}^1, \mathbb{R})$; we denote by $\mathcal{S}_{\mathrm{Id}}(f)$ the stabiliser of f and by $\mathcal{O}_{\mathrm{Id}}(f)$ its orbit.

Proposition 6.5.1. *The fiber* $\mathrm{PH}^{-1}(D)$ *has finitely many path connected components. In each such component* $\Omega(D)$ *there exists some* $f_{\Omega(D)} : \mathbb{S}^1 \to \mathbb{R}$ *such that:*

$$\Omega(D) = \mathcal{O}_{\mathrm{Id}}(f_{\Omega(D)}),$$

and then $\Omega(D)$ *is homeomorphic to the quotient* $\mathrm{End}_{\leqslant}(\mathbb{S}^1)/\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$*, and in particular is homotopy equivalent to* $\mathbb{S}^1$*.*

Unlike the smooth case the component $\Omega(D) \subseteq \mathrm{PH}^{-1}(D)$ in the fiber equals the orbit of a function only for a careful choice of function $f_{\Omega(D)}$: the requirement will be that $f_{\Omega(D)}$ is injective between its consecutive extrema. Nevertheless the fact that the pre-composition map induces

a homeomorphism from $\mathrm{End}_{\leqslant}(\mathbb{S}^1)/\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$ to the orbit $\mathcal{O}_{\mathrm{Id}}(f_{\Omega(D)})$ is reminiscent of the smooth case, and in fact with slightly more work it can be shown that it defines a $\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$-principal bundle. We state without proof the analogous and simpler result for the unit interval $[0,1]$, which works with or without fixed values on the boundary points 0 and 1.

Proposition 6.5.2. *For any barcode D the fiber $\mathrm{PH}^{-1}(D) \subseteq \mathcal{C}^0([0,1],\mathbb{R})$ has finitely many path connected components, each of which is contractible.*

Using a fixed orientation on $\mathbb{S}^1$ and going around starting from the north pole we can order the n minima and n maxima of a non-constant $f \in \mathcal{C}^0(\mathbb{S}^1,\mathbb{R})$ into a sequence $\mathrm{Val}(f)$ which we view as an element in $\mathbb{R}^{2n}$:

$$\mathrm{Val}(f) := m_1(f) < M_1(f) > \cdots > m_n(f) < M_n(f).$$

Associated to this sequence we have the sequence of critical sets of f:

$$\mathrm{Seq}(f) : c_1(f),\, d_1(f), \cdots,\, c_n(f),\, d_n(f).$$

Explicitly, each $c_i(f)$ (resp. $d_i(f)$) is a connected component of $f^{-1}(m_i(f))$ (resp. of $f^{-1}(M_i(f))$).

Proposition 6.5.3. *Let $f \in \mathrm{PH}^{-1}(D)$. Then f has $2n$ extrema, i.e. $\mathrm{Val}(f) \in \mathbb{R}^{2n}$. In addition, let Γ_n be the group of cyclic permutations on n elements, which acts on $\mathbb{R}^{2n}$ by cyclically permuting the n pairs of entries. Then the connected component $\Omega(D)$ in the fiber containing f is made of functions g whose sequence of extrema is the same as that of f up to a different ordering, that is:*

$$\Omega(D) = \left\{ g \in \mathcal{C}^0(\mathbb{S}^1,\mathbb{R}) \,|\, \mathrm{Val}(g) \in \Gamma_n.\mathrm{Val}(f) \right\} \tag{6.3}$$

We omit the proof of this elementary statement. So if $\Omega(D)$ is a component in the fiber, we can pick the following simple function $f_{\Omega(D)}$ in $\Omega(D)$, whose critical sets and extrema are denoted by c_i, d_i, m_i, M_i for simplicity: the critical sets c_i and d_i are singletons arranged on the regular $2n$-gon in $\mathbb{S}^1$ and on each circular arc $[c_i, d_i]$, $f_{\Omega(D)}$ restricts to the linear homeomorphism to $[m_i, M_i]$.

Proposition 6.5.4. *Let $\Omega(D) \subseteq \mathrm{PH}^{-1}(D)$ be a path component in the fiber. Then the pre-composition map $\phi \mapsto f_{\Omega(D)} \circ \phi$ induces a homeomorphism from $\mathrm{End}_{\leqslant}(\mathbb{S}^1)/\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$ to $\Omega(D)$.*

Proof. The map $\phi \in \mathrm{End}_{\leqslant}(\mathbb{S}^1) \mapsto f_{\Omega(D)} \circ \phi \in \Omega(D)$ is well-defined, i.e. $\mathrm{PH}(f \circ \phi) = \mathrm{PH}(f) = D$. This is because a homeomorphism $\phi \in \mathrm{Aut}_{\leqslant}(\mathbb{S}^1)$ restricts to a homeomorphism between the sub level-sets of $f_{\Omega(D)} \circ \phi$ and those of $f_{\Omega(D)}$, hence it induces an isomorphism of persistent homology modules and the equality of barcodes $\mathrm{PH}(f_{\Omega(D)} \circ$

$\phi) = \mathrm{PH}(f_{\Omega(D)})$, which holds as well for any $\phi \in \mathrm{End}_{\leqslant}(\mathbb{S}^1) = \overline{\mathrm{Aut}_{\leqslant}(\mathbb{S}^1)}$ by continuity of PH.

Let $f \in \Omega(D)$. From Proposition 6.5.3 there are cyclic permutations $\pi \in \Gamma_n$ such that $\mathrm{Val}(f) = \pi.\mathrm{Val}(f_{\Omega(D)})$. For each such permutation π there is a unique map $\phi^{f,\pi}$ satisfying both $f_{\Omega(D)} \circ \phi^{f,\pi} = f$ and $\phi^{f,\pi}(c_i(f)) = c_{\pi(i)}$ (and $\phi^{f,\pi}(d_i(f)) = d_{\pi(i)}$): It is defined on each circular arc $[c_i(f), d_i(f)]$ by

$$\phi^{f,\pi}_{|[c_i(f),d_i(f)]} := [(f_{\Omega(D)})_{|[c_{\pi(i)},d_{\pi(i)}]}]^{-1} \circ f_{|[c_i(f),d_i(f)]}, \tag{6.4}$$

and similarly on circular arcs $[d_{i-1}(f), c_i(f)]$. In particular for $f = f_{\Omega(D)}$ the set of such $\phi^{f,\pi}$ equals the group $\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$ of stabilisers. Therefore $\phi \mapsto f_{\Omega(D)} \circ \phi$ descends to a continuous bijection from $\mathrm{End}_{\leqslant}(\mathbb{S}^1)/\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$ to $\Omega(D)$.

Finally we show that the inverse is continuous. Let $f \in \Omega(D)$ and $\phi^{f,\pi}$ as in (6.4). Up to pre-composing f by a suitable homeomorphism the north pole does not belong to any extremal set $c_i(f), d_i(f)$. Consequently, for g in a small neighborhood $\mathcal{U} \subseteq \Omega(D)$ around f, we also have $\mathrm{Val}(g) = \pi.\mathrm{Val}(f_{\Omega(D)})$, hence we can define $\phi^{g,\pi} \in \mathrm{End}_{\leqslant}(\mathbb{S}^1)$ like in Eq. (6.4) and then $f_{\Omega(D)} \circ \phi^{g,\pi} = g$. Hence the map $g \in \mathcal{U} \longmapsto \phi^{g,\pi} \in \mathrm{End}_{\leqslant}(\mathbb{S}^1)$ is a local section, whose continuity is a consequence of the fact that on each circular arc $[c_i, d_i]$ the linear restriction $(f_{\Omega(D)})_{|[c_i,d_i]}$ and its inverse are Lipschitz, and of the fact that the maximal distance from points in the critical sets $c_i(g), d_i(g)$ to the critical sets $c_i(f), d_i(f)$ of f can be continuously tracked in a sufficiently small neighborhood $\mathcal{U}$ of f, see Fig. 6.1. The technical details are omitted. □

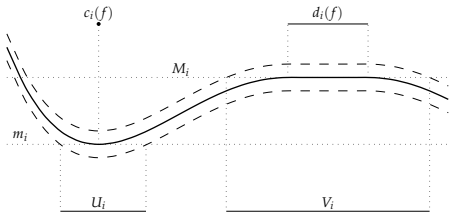


Figure 6.1: A piece of a continuous function $f : \mathbb{S}^1 \to \mathbb{R}$ and a small neighborhood $\mathcal{U}$ indicated by dashed curves. Any function in $\mathrm{PH}^{-1}(D)$ between the dashed curves must have a critical value in each U_i and V_i, provided the band between the dashed curves is thin enough to separate critical values. If g is such a function then these must be the only critical values. Then the critical value of g in U_i must be m_i, in V_i must be M_i, and so on.

Proof of Proposition 6.5.1. From Proposition 6.5.4 the pre-composition map $\phi \mapsto f_{\Omega(D)} \circ \phi$ induces a homeomorphism from $\mathrm{End}_{\leqslant}(\mathbb{S}^1)/\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$ to the path connected component $\Omega(D)$. Besides it is well-known that $\mathrm{End}_{\leqslant}(\mathbb{S}^1)$ deformation retracts to the group $\mathrm{SO}(2) \cong \mathbb{S}^1$ of orientation preserving rotations.[15] Recall that $f_{\Omega(D)}$ is a piece-wise linear interpolation between extremal values arranged on a regular $2n$-gon, therefore its stabiliser $\mathcal{S}_{\mathrm{Id}}(f_{\Omega(D)})$ is a finite subgroup of $\mathrm{SO}(2)$ which is preserved under the deformation retraction. Hence $\Omega(D)$ is homotopy equivalent to the quotient of $\mathrm{SO}(2) \cong \mathbb{S}^1$ by a finite subgroup, so it is in fact homotopy equivalent to $\mathbb{S}^1$. □

[15] For instance the deformation retract of $\mathrm{Aut}_{\leqslant}(\mathbb{S}^1)$ in [H+74, Theorem 1.1.2] extends to $\mathrm{End}_{\leqslant}(\mathbb{S}^1)$.

Bibliography

[AB17] Karim A Adiprasito and Bruno Benedetti. Subdivisions, shellability, and collapsibility of products. *Combinatorica*, 37(1):1–30, 2017.

[ABC+16] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pages 265–283, 2016.

[ABS13] Hedy Attouch, Jérôme Bolte, and Benar Fux Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Math. Program.*, 137(1-2, Ser. A):91–129, 2013.

[AEK+17] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research*, 18(1):218–252, 2017.

[AGH+09a] Dominique Attali, Marc Glisse, Samuel Hornus, Francis Lazarus, and Dmitriy Morozov. Persistence-sensitive simplication of functions on surfaces in linear time. In *Topological Methods in Data Analysis and Visualization (TopoInVis 2009)*. Springer, 2009.

[AGH+09b] Dominique Attali, Marc Glisse, Samuel Hornus, Francis Lazarus, and Dmitriy Morozov. Persistence-sensitive simplication of functions on surfaces in linear time. In *Topological Methods in Data Analysis and Visualization (TopoInVis 2009)*. Springer, 2009.

[AHHH01] Ali N Akansu, Paul A Haddad, Richard A Haddad, and Paul R Haddad. *Multiresolution signal decomposition: transforms, subbands, and wavelets*. Academic press, 2001.

[ASC11] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *2011 IEEE international conference on computer vision workshops (ICCV workshops)*, pages 1626–1633. IEEE, 2011.

[Bar94a] S. A. Barannikov. The framed Morse complex and its invariants. In *Singularities and bifurcations*, volume 21 of *Adv. Soviet Math.*, pages 93–115. Amer. Math. Soc., Providence, RI, 1994.

[Bar94b] Serguei Barannikov. The framed Morse complex and its invariants. *American Mathematical Society Translations, Series 2*, 1994.

[BBBN20] Francisco Belchí, Jacek Brodzki, Matthew Burfitt, and Mahesan Niranjan. A numerical measure of the instability of Mapper-type algorithms. *Journal of Machine Learning Research*, 21(202):1–45, 2020.

[BBMW21] Adam Brown, Omer Bobrowski, Elizabeth Munch, and Bei Wang. Probabilistic convergence and stability of random Mapper graphs. *Journal of Applied and Computational Topology*, 5:99–140, 2021.

[BC90] Michael J Best and Nilotpal Chakravarti. Active set algorithms for isotonic regression; a unifying framework. *Mathematical Programming*, 47(1):425–439, 1990.

[BCLO19] James V Burke, Frank E Curtis, Adrian S Lewis, and Michael L Overton. The gradient sampling methodology. *invited survey for* INFORMS Computing Society Newsletter, *Research Highlights*, 2019.

[BDLS07] Jérôme Bolte, Aris Daniilidis, Adrian Lewis, and Masahiro Shiota. Clarke subgradients of stratifiable functions. *SIAM Journal on Optimization*, 18(2):556–572, 2007.

[BDP+15] Paolo Bajardi, Matteo Delfino, André Panisson, Giovanni Petri, and Michele Tizzoni. Unveiling patterns of international communities in a global city using mobile phone data. *EPJ Data Science*, 4:1–17, 2015.

[Ber75] Dimitri P Bertsekas. *Nondifferentiable optimization via approximation*. Springer, 1975.

[Bet18] Leo M Betthauser. *Topological Reconstruction of Grayscale Images*. PhD thesis, University of Florida, 2018.

[BFM+20] Robin Lynne Belton, Brittany Terese Fasy, Rostik Mertz, Samuel Micka, David L Millman, Daniel Salinas, Anna Schenfisch, Jordan Schupbach, and Lucia Williams. Reconstructing embedded graphs from persistence diagrams. *Computational Geometry*, page 101658, 2020.

[BGGSSG20a] Rickard Brüel-Gabrielsson, Vignesh Ganapathi-Subramanian, Primoz Skraba, and Leonidas J Guibas. Topology-aware surface reconstruction for point clouds. In *Computer Graphics Forum*, volume 39, pages 197–207. Wiley Online Library, 2020.

[BGGSSG20b] Rickard Brüel-Gabrielsson, Vignesh Ganapathi-Subramanian, Primoz Skraba, and Leonidas J Guibas. Topology-aware surface reconstruction for point clouds. In *Computer Graphics Forum*, volume 39 No. 5, pages 197–207. Wiley Online Library, 2020.

[BGK15] Subhrajit Bhattacharya, Robert Ghrist, and Vijay Kumar. Persistent homology for path planning in uncertain environments. *IEEE Transactions on Robotics*, 31(3):578–590, 2015.

[BGKP13] Holger Boche, Mijail Guillemard, Gitta Kutyniok, and Friedrich Philipp. Signal analysis with frame theory and persistent homology. *Proceedings of Sampling Theory and Applications (SampTA'13)*, 2013.

[Bih84] A Bihain. Optimization of upper semidifferentiable functions. *Journal of Optimization Theory and Applications*, 44(4):545–568, 1984.

[BK10] Michael M Bronstein and Iasonas Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1704–1711. IEEE, 2010.

[BKM14] Adil Bagirov, Napsu Karmitsa, and Marko M Mäkelä. *Introduction to Nonsmooth Optimization: theory, practice and software*. Springer, 2014.

[BKSW18] Paul Breiding, Sara Kališnik, Bernd Sturmfels, and Madeleine Weinstein. Learning algebraic varieties from samples. *Revista Matemática Complutense*, 31(3):545–593, 2018.

[BL15] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *Journal of Computational Geometry*, 6(1):162–191, 2015.

[BLO05] James V Burke, Adrian S Lewis, and Michael L Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM Journal on Optimization*, 15(3):751–779, 2005.

[BLS07] Türker Biyikoglu, Josef Leydold, and Peter F Stadler. *Laplacian eigenvectors of graphs: Perron-Frobenius and Faber-Krahn type theorems*. Springer, 2007.

[BMM+16] Paul Bendich, James S Marron, Ezra Miller, Alex Pieloch, and Sean Skwerer. Persistent homology analysis of brain artery trees. *The annals of applied statistics*, 10(1):198, 2016.

[BOS+05] Karsten M Borgwardt, Cheng Soon Ong, Stefan Schönauer, SVN Vishwanathan, Alex J Smola, and Hans-Peter Kriegel. Protein function prediction via graph kernels. *Bioinformatics*, 21(suppl_1):i47–i56, 2005.

[Bub15] Peter Bubenik. Statistical topological data analysis using persistence landscapes. *The Journal of Machine Learning Research*, 16(1):77–102, 2015.

[CB15] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications*, 14(05):1550066, 2015.

[CBK09] Moo K Chung, Peter Bubenik, and Peter T Kim. Persistence diagrams of cortical surface data. In *International Conference on Information Processing in Medical Imaging*, pages 386–397. Springer, 2009.

[CCBdS16] Frédéric Chazal, William Crawley-Boevey, and Vin de Silva. The observable structure of persistence modules. *Homology Homotopy Appl.*, 18(2):247–265, 2016.

[CCF+20] Michael J Catanzaro, Justin M Curry, Brittany Terese Fasy, Jānis Lazovskis, Greg Malen, Hans Riess, Bei Wang, and Matthew Zabka. Moduli spaces of Morse functions for persistence. *Journal of Applied and Computational Topology*, pages 1–33, 2020.

[CCG91] Sheng Chen, Colin FN Cowan, and Peter M Grant. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on neural networks*, 2(2):302–309, 1991.

[CCG+21] Mathieu Carriere, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hariprasad Kannan, and Yuhei Umeda. Optimizing persistent homology based functions. In *International Conference on Machine Learning*, pages 1294–1303. PMLR, 2021.

[CCI+20] Mathieu Carrière, Frédéric Chazal, Yuichi Ike, Théo Lacombe, Martin Royer, and Yuhei Umeda. Perslay: A neural network layer for persistence diagrams and new graph topological signatures. In *International Conference on Artificial Intelligence and Statistics*, pages 2786–2796. PMLR, 2020.

[CCO17] Mathieu Carriere, Marco Cuturi, and Steve Oudot. Sliced wasserstein kernel for persistence diagrams. In *International Conference on Machine Learning*, volume 70, pages 664–673. PMLR, 2017.

[CCR13] Joseph Minhow Chan, Gunnar Carlsson, and Raul Rabadan. Topology of viral evolution. *Proceedings of the National Academy of Sciences*, 110(46):18566–18571, 2013.

[CCSG+09] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the twenty-fifth annual Symposium on Computational Geometry*, pages 237–246. ACM, 2009.

[CD20] Padraig Corcoran and Bailin Deng. Regularization of persistent homology gradient computation. In *NeurIPS 2020 Workshop on Topological Data Analysis and Beyond*, 2020.

[CdSGO16] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The structure and stability of persistence modules*. SpringerBriefs in Mathematics. Springer, 2016.

[Cer61] Jean Cerf. Topologie de certains espaces de plongements. *Bulletin de la Société Mathématique de France*, 89:227–380, 1961.

[Cer70] Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la

pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970.

[CFL+14] Frédéric Chazal, Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, and Larry Wasserman. Stochastic convergence of persistence landscapes and silhouettes. In *Proceedings of the thirtieth annual symposium on Computational geometry*, pages 474–483, 2014.

[CG97] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997.

[CH13] Corrie J Carstens and Kathy J Horadam. Persistent homology of collaboration networks. *Mathematical problems in engineering*, 2013, 2013.

[Chu16] Charles K Chui. *An introduction to wavelets*. Elsevier, 2016.

[CIDSZ08] Gunnar Carlsson, Tigran Ishkhanov, Vin De Silva, and Afra Zomorodian. On the local behavior of spaces of natural images. *International journal of computer vision*, 76(1):1–12, 2008.

[CK16] Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788*, 2016.

[CL19] Yu-Min Chung and Austin Lawson. Persistence curves: A canonical framework for summarizing persistence diagrams. *arXiv preprint arXiv:1904.07768*, 2019.

[Cla90] Frank H Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990.

[CLR16] Pablo G Cámara, Arnold J Levine, and Raul Rabadan. Inference of ancestral recombination graphs through topological data analysis. *PLoS computational biology*, 12(8):e1005071, 2016.

[CM19] Mathieu Carrière and Bertrand Michel. Statistical analysis of Mapper for stochastic and multivariate filters. In *CoRR*. arXiv:1912.10742, 2019.

[CMO18] Mathieu Carrière, Bertrand Michel, and Steve Oudot. Statistical analysis and parameter selection for Mapper. *Journal of Machine Learning Research*, 19(12):1–39, 2018.

[CMT18] Justin Curry, Sayan Mukherjee, and Katharine Turner. How many directions determine a shape and other sufficiency results for two topological transforms. *arXiv preprint arXiv:1805.09782*, 2018.

[CMW20] Jacek Cyranka, Konstantin Mischaikow, and Charles Weibel. Contractibility of a persistence map preimage. *Journal of Applied and Computational Topology*, 4(4):509–523, 2020.

[CNBW19] Chao Chen, Xiuyan Ni, Qinxun Bai, and Yusu Wang. A topological regularizer for classifiers via persistent homology. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 2573–2582, 2019.

[CO17] Mathieu Carrière and Steve Oudot. Structure and stability of the one-dimensional Mapper. *Foundations of Computational Mathematics*, 18(6):1333–1396, 2017.

[COB+19] James R Clough, Ilkay Oksuz, Nicholas Byrne, Julia A Schnabel, and Andrew P King. Explicit topological priors for deep-learning based image segmentation using persistent homology. In *International Conference on Information Processing in Medical Imaging*, pages 16–28. Springer, 2019.

[Coh12] Marshall M Cohen. *A course in simple-homotopy theory*, volume 10. Springer Science & Business Media, 2012.

[COO15] Mathieu Carrière, Steve Oudot, and Maks Ovsjanikov. Stable topological signatures for points on 3d shapes. In *Computer Graphics Forum*, volume 34, pages 1–12. Wiley Online Library, 2015.

[CP16] Justin Curry and Amit Patel. Classification of constructible cosheaves. *arXiv preprint arXiv:1603.01587*, 2016.

[CR20] Mathieu Carrière and Raúl Rabadán. Topological data analysis of single-cell Hi-C contact maps. In *Abel Symposia*, volume 15, pages 147–162. Springer-Verlag, 2020.

[CSEH07] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007.

[CSEH09] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Extending persistence using poincaré and lefschetz duality. *Foundations of Computational Mathematics*, 9(1):79–103, 2009.

[CSEHM10] David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko. Lipschitz functions have l p-stable persistence. *Foundations of Computational Mathematics*, 10(2):127–139, 2010.

[CSEM06] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the twenty-second annual Symposium on Computational Geometry*, pages 119–126. ACM, 2006.

[Cur17] Carina Curto. What can topology tell us about the neural code? *Bulletin of the American Mathematical Society*, 54(1):63–78, 2017.

[Cur18] Justin Curry. The fiber of the persistence map for functions on the interval. *Journal of Applied and Computational Topology*, 2(3-4):301–321, 2018.

[CWRW15] Yen-Chi Chen, Daren Wang, Alessandro Rinaldo, and Larry Wasserman. Statistical analysis of persistence intensity functions. *arXiv preprint arXiv:1510.02502*, 2015.

[CZ09] Gunnar Carlsson and Afra Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, 2009.

[DB16] Steven Diamond and Stephen Boyd. Cvxpy: A python-embedded modeling language for convex optimization. *The Journal of Machine Learning Research*, 17(1):2909–2913, 2016.

[DBF14] Yuri Dabaghian, Vicky L Brandt, and Loren M Frank. Reconceiving the hippocampal map as a topological template. *Elife*, 3:e03476, 2014.

[DC19] Vincent Divol and Frédéric Chazal. The density of expected persistence diagrams and its kernel based estimation. *Journal of Computational Geometry*, 10(2):127–153, 2019.

[DCCW+10] Daniel DeWoskin, Joan Climent, I Cruz-White, Mariel Vazquez, Catherine Park, and Javier Arsuaga. Applications of computational homology to the analysis of treatment response in breast cancer patients. *Topology and its Applications*, 157(1):157–164, 2010.

[DD03] Paul D Dobson and Andrew J Doig. Distinguishing enzyme structures from non-enzymes without alignments. *Journal of Molecular Biology*, 330(4):771–783, 2003.

[DDKL20] Damek Davis, Dmitriy Drusvyatskiy, Sham Kakade, and Jason D Lee. Stochastic subgradient method converges on tame functions. *Foundations of Computational Mathematics*, 20(1):119–154, 2020.

[DFF15] Barbara Di Fabio and Massimo Ferri. Comparing persistence diagrams through complex vectors. In *International Conference on Image Analysis and Processing*, pages 294–305. Springer, 2015.

[DFP+11] Stanley Durrleman, Pierre Fillard, Xavier Pennec, Alain Trouvé, and Nicholas Ayache. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage*, 55(3):1073–1090, 2011.

[DGS01] Lijin Ding, Ardeshir Goshtasby, and Martin Satter. Volume image registration by template matching. *Image and Vision Computing*, 19(12):821–832, 2001.

[DIL15] Dmitriy Drusvyatskiy, Alexander D Ioffe, and Adrian S Lewis. Clarke subgradients for directionally lipschitzian stratifiable functions. *Mathematics of Operations Research*, 40(2):328–349, 2015.

[DL20] Vincent Divol and Théo Lacombe. Understanding the topology and the geometry of the space of persistence diagrams via optimal partial transport. *Journal of Applied and Computational Topology*, pages 1–53, 2020.

[DLdCD+91] Asim Kumar Debnath, Rosa L Lopez de Compadre, Gargi Debnath, Alan J Shusterman, and Corwin Hansch. Structure-activity relationship of mutagenic aromatic and heteroaromatic nitro compounds. correlation with molecular orbital energies and hydrophobicity. *Journal of Medicinal Chemistry*, 34(2):786–797, 1991.

[DSG07] Vin De Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7(1):339–358, 2007.

[DSMVJ11] Vin De Silva, Dmitriy Morozov, and Mikael Vejdemo-Johansson. Dualities in persistent (co) homology. *Inverse Problems*, 27(12):124003, 2011.

[DSW15] Tamal K. Dey, Dayu Shi, and Yusu Wang. Comparing graphs via persistence distortion. In *31st International Symposium on Computational Geometry*, volume 34 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 491–506. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2015.

[EH08] Herbert Edelsbrunner and John Harer. Persistent homology-a survey. *Contemporary mathematics*, 453:257–282, 2008.

[EH10a] H. Edelsbrunner and J. Harer. *Computational Topology: an Introduction*. American Mathematical Society, Providence, RI, 2010.

[EH10b] Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. American Mathematical Soc., 2010.

[ELZ02] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002.

[ESR16] Kevin Emmett, Benjamin Schweinhart, and Raul Rabadan. Multiscale topology of chromatin folding. In *Proceedings of the 9th EAI International Conference on Bio-Inspired Information and Communications Technologies (Formerly BIONETICS)*, page 177–180. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2016.

[Fed69] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[FGG04a] Antonio Fuduli, Manlio Gaudioso, and Giovanni Giallombardo. A DC piecewise affine model and a bundling technique in nonconvex nonsmooth minimization. *Optimization Methods and Software*, 19(1):89–102, 2004.

[FGG04b] Antonio Fuduli, Manlio Gaudioso, and Giovanni Giallombardo. Minimizing nonconvex nonsmooth functions via cutting planes and proximity control. *Siam journal on optimization*, 14(3):743–756, 2004.

[FK88] Alfred Frölicher and Andreas Kriegl. *Linear spaces and differentiation theory*. Pure and Applied Mathematics (New York). John Wiley & Sons, Ltd., Chichester, 1988. A Wiley-Interscience Publication.

[FMM+19] Brittany Terese Fasy, Samuel Micka, David L Millman, Anna Schenfisch, and Lucia Williams. Persistence diagrams for efficient simplicial complex reconstruction. *arXiv preprint arXiv:1912.12759*, 2019.

[GB09] Joseph Gubeladze and Winfried Bruns. *Polytopes, rings, and K-theory*. Springer, 2009.

[GG73] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities*. Springer-Verlag, New York-Heidelberg, 1973. Graduate Texts in Mathematics, Vol. 14.

[GGB16] Chad Giusti, Robert Ghrist, and Danielle S Bassett. Two's company, three (or more) is a simplex. *Journal of computational neuroscience*, 41(1):1–14, 2016.

[GHI+15] Marcio Gameiro, Yasuaki Hiraoka, Shunsuke Izumi, Miroslav Kramar, Konstantin Mischaikow, and Vidit Nanda. A topological measurement of protein compressibility. *Japan Journal of Industrial and Applied Mathematics*, 32(1):1–17, 2015.

[GHO16] Marcio Gameiro, Yasuaki Hiraoka, and Ippei Obayashi. Continuation of point clouds via persistence diagrams. *Physica D: Nonlinear Phenomena*, 334:118–132, 2016.

[Gid17] Marian Gidea. Topological data analysis of critical transitions in financial networks. In *3rd International Winter School and Conference on Network Science*, pages 47–59. Springer International Publishing, 2017.

[GLM18] Robert Ghrist, Rachel Levanger, and Huy Mai. Persistent homology and euler integral transforms. *Journal of Applied and Computational Topology*, 2(1-2):55–60, 2018.

[GM88] Mark Goresky and Robert MacPherson. *Stratified Morse theory*. Ergebnisse der Mathematik, volume 14. Springer-Verlag, Berlin, 1988.

[GNDS20a] Rickard Brüel Gabrielsson, Bradley J Nelson, Anjan Dwaraknath, and Primoz Skraba. A topology layer for machine learning. In *International Conference on Artificial Intelligence and Statistics*, pages 1553–1563. PMLR, 2020.

[GNDS20b] Rickard Brüel Gabrielsson, Bradley J Nelson, Anjan Dwaraknath, and Primoz Skraba. A topology layer for machine learning. In *International Conference on Artificial Intelligence and Statistics*, pages 1553–1563. PMLR, 2020.

[Gol77] AA Goldstein. Optimization of lipschitz continuous functions. *Mathematical Programming*, 13(1):14–22, 1977.

[Gra95] Amara Graps. An introduction to wavelets. *IEEE computational science and engineering*, 2(2):50–61, 1995.

[H+74] Mary-Elizabeth Hamstrom et al. Homotopy in homeomorphism spaces, *TOP* and *PL*. *Bulletin of the American Mathematical Society*, 80(2):207–230, 1974.

[HGR+20a] Christoph Hofer, Florian Graf, Bastian Rieck, Marc Niethammer, and Roland Kwitt. Graph filtration learning. In *International Conference on Machine Learning*, pages 4314–4323. PMLR, 2020.

[HGR+20b] Christoph Hofer, Florian Graf, Bastian Rieck, Marc Niethammer, and Roland Kwitt. Graph filtration learning. In *International Conference on Machine Learning*, pages 4314–4323. PMLR, 2020.

[HHP21] Haibin Hang and Gregory Henselman-Petrusek. Exact homological algebra for computational topology (ExHACT). https://github.com/ExHACT, 2021.

[HKN19] Christoph D Hofer, Roland Kwitt, and Marc Niethammer. Learning representations of persistence barcodes. *Journal of Machine Learning Research*, 20(126):1–45, 2019.

[HKND19] Christoph Hofer, Roland Kwitt, Marc Niethammer, and Mandar Dixit. Connectivity-optimized representation learning via persistent homology. In *International Conference on Machine Learning*, pages 2751–2760. PMLR, 2019.

[HLSC19] Xiaoling Hu, Fuxin Li, Dimitris Samaras, and Chao Chen. Topology-preserving deep image segmentation. In *Advances in Neural Information Processing Systems*, pages 5658–5669, 2019.

[HMM07] Napsu Haarala, Kaisa Miettinen, and Marko M Mäkelä. Globally convergent limited memory bundle method for large-scale nonsmooth optimization. *Mathematical Programming*, 109(1):181–205, 2007.

[HNH+16] Yasuaki Hiraoka, Takenobu Nakamura, Akihiko Hirata, Emerson G Escolar, Kaname Matsue, and Yasumasa Nishiura. Hierarchical structures of amorphous solids characterized by persistent homology. *Proceedings of the National Academy of Sciences*, 113(26):7035–7040, 2016.

[HRG14] Nan Hu, Raif M Rustamov, and Leonidas Guibas. Stable and informative spectral signatures for graph matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2305–2312, 2014.

[HSS17] Elias Salomão Helou, Sandra A Santos, and Lucas EA Simões. On the local convergence analysis of the gradient sampling method for finite max-functions. *Journal of Optimization Theory and Applications*, 175(1):137–157, 2017.

[Hur55] Witold Hurewicz. On the concept of fiber space. *Proceedings of the National Academy of Sciences of the United States of America*, 41(11):956, 1955.

[HVG11] David K Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.

[Iof09] AD Ioffe. An invitation to tame optimization. *SIAM Journal on Optimization*, 19(4):1894–1917, 2009.

[IZ13] Patrick Iglesias-Zemmour. *Diffeology*. Mathematical Surveys and Monographs, volume 185. American Mathematical Society, Providence, RI, 2013.

[JHP22] Leygonie Jacob and Gregory Henselman-Petrusek. Software Companion to Algorithmic Reconstruction of the Fiber of Persistent Homology on Cell Complexes. https://github.com/Eetion/phfibre, 2022.

[JM00] Sarang C Joshi and Michael I Miller. Landmark matching via large deformation diffeomorphisms. *IEEE transactions on image processing*, 9(8):1357–1370, 2000.

[Kac20] Oleg Kachan. Persistent homology-based projection pursuit. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 856–857, 2020.

[Kal19] Sara Kališnik. Tropical coordinates on the space of persistence barcodes. *Found. Comput. Math.*, 19(1):101–129, 2019.

[KB14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[KGKM13] M Kramar, Arnaud Goullet, Lou Kondic, and Konstantin Mischaikow. Persistence of force networks in compressed granular media. *Physical Review E*, 87(4):042207, 2013.

[KGKM14] Miroslav Kramár, Arnaud Goullet, Lou Kondic, and Konstantin Mischaikow. Quantifying force networks in particulate systems. *Physica D: Nonlinear Phenomena*, 283:37–55, 2014.

[KHF16] Genki Kusano, Yasuaki Hiraoka, and Kenji Fukumizu. Persistence weighted gaussian kernel for topological data analysis. In *International Conference on Machine Learning*, pages 2004–2013, 2016.

[Kiw06] Krzysztof C Kiwiel. *Methods of descent for nondifferentiable optimization*, volume 1133. Springer, 2006.

[Kiw07] Krzysztof C Kiwiel. Convergence of the gradient sampling algorithm for nonsmooth nonconvex optimization. *SIAM Journal on Optimization*, 18(2):379–388, 2007.

[KM97] Andreas Kriegl and Peter W. Michor. *The convenient setting of global analysis*. Mathematical Surveys and Monographs, volume 53. American Mathematical Society, Providence, RI, 1997.

[KM12] Nils Kriege and Petra Mutzel. Subgraph matching kernels for attributed graphs. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ICML'12, page 291–298, Madison, WI, USA, 2012. Omnipress.

[KNBNH16] Violeta Kovacev-Nikolic, Peter Bubenik, Dragan Nikolić, and Giseon Heo. Using persistent homology and dynamical distances to analyze protein binding. *Statistical applications in genetics and molecular biology*, 15(1):19–38, 2016.

[LEF+16] Louis-David Lord, Paul Expert, Henrique M Fernandes, Giovanni Petri, Tim J Van Hartevelt, Francesco Vaccarino, Gustavo Deco, Federico Turkheimer, and Morten L Kringelbach. Insights into brain architectures from the homological scaffolds of functional connectivity networks. *Frontiers in systems neuroscience*, 10:85, 2016.

[Les15] Michael Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics*, 15(3):613–650, 2015.

[LH13] Chunyuan Li and A Ben Hamza. A multiresolution descriptor for deformable 3d shape retrieval. *The Visual Computer*, 29(6-8):513–524, 2013.

[LOT21] Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021.

[LP18] Derek Lo and Briton Park. Modeling the spread of the zika virus using topological data analysis. *Plos one*, 13(2):e0192120, 2018.

[LPRS08] Gregory Leibon, Scott Pauls, Daniel Rockmore, and Robert Savell. Topological structures in the equities market network. *Proceedings of the National Academy of Sciences*, 105(52):20589–20594, 2008.

[LT21] Jacob Leygonie and Ulrike Tillmann. The fiber of persistent homology for simplicial complexes. *arXiv preprint arXiv:2104.01372*, 2021.

[LV98] Ladislav Lukšan and Jan Vlček. A bundle-Newton method for nonsmooth unconstrained minimization. *Mathematical Programming*, 83(1-3):373–391, 1998.

[Mak06] Sergiy Maksymenko. Homotopy types of stabilizers and orbits of Morse functions on surfaces. *Annals of Global Analysis and Geometry*, 29(3):241–285, 2006.

[Mat69] John N Mather. Stability of C^∞ mappings: II. infinitesimal stability implies stability. *Annals of Mathematics*, pages 254–291, 1969.

[Mat12] John Mather. Notes on topological stability. *Bull. Amer. Math. Soc. (N.S.)*, 49(4):475–506, 2012.

[MBGY14] Clément Maria, Jean-Daniel Boissonnat, Marc Glisse, and Mariette Yvinec. The gudhi library: Simplicial complexes and persistent homology. In *International congress on mathematical software*, pages 167–174. Springer, 2014.

[MHRB20a] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. Topological autoencoders. In *International Conference on Machine Learning*, pages 7045–7054. PMLR, 2020.

[MHRB20b] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. Topological autoencoders. In *International conference on machine learning*, pages 7045–7054. PMLR, 2020.

[Mic80] Peter W. Michor. *Manifolds of differentiable mappings*. Shiva Mathematics Series, volume 3. Shiva Publishing Ltd., Nantwich, 1980.

[Mic20] Samuel Adam Micka. *Searching and Reconstruction: Algorithms with Topological Descriptors*. PhD thesis, Montana State University, 2020.

[Mif77] Robert Mifflin. An algorithm for constrained optimization with semismooth functions. *Mathematics of Operations Research*, 2(2):191–207, 1977.

[Mil63] John Milnor. Morse theory. *Princeton University Press, Princeton, N.J.*, pages vi+153, 1963.

[Mil06] David A Miller. Popaths and holinks. *Journal of Homotopy and Related Structures*, 1(1):1–9, 2006.

[MM20] Yuri I. Manin and Matilde Marcolli. Nori diagrams and persistent homology. *Math. Comput. Sci.*, 14(1):77–102, 2020.

[MW16] Elizabeth Munch and Bei Wang. Convergence between categorical representations of Reeb space and Mapper. In *32nd International Symposium on Computational Geometry (SoCG 2016)*, volume 51, pages 53:1–53:16. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016.

[MW21] Konstantin Mischaikow and Charles Weibel. Persistent homology with non-contractible preimages. *arXiv preprint arXiv:2105.08130*, 2021.

[MZR16] Slobodan Maletić, Yi Zhao, and Milan Rajković. Persistent topological features of dynamical systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(5):053105, 2016.

[NAK16] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2014–2023, New York, New York, USA, 20–22 Jun 2016. PMLR.

[New18] Mark Newman. *Networks*. Oxford university press, 2018.

[Nic11] Liviu Nicolaescu. *An invitation to Morse theory*. Universitext. Springer, New York, second edition, 2011.

[NLC11] Monica Nicolau, Arnold J Levine, and Gunnar Carlsson. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proceedings of the National Academy of Sciences*, 108(17):7265–7270, 2011.

[NLV+17] Takashi Nagano, Yaniv Lubling, Csilla Várnai, Carmel Dudley, Wing Leung, Yael Baran, Netta Mendelson-Cohen, Steven Wingett, Peter Fraser, and Amos Tanay. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*, 547:61–67, 2017.

[Nol14] Dominikus Noll. Convergence of non-smooth descent methods using the Kurdyka-Lojasiewicz inequality. *J. Optim. Theory Appl.*, 160(2):553–572, 2014.

[OPT+17] Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6(1):17, 2017.

[OS17] Steve Oudot and Elchanan Solomon. Barcode embeddings for metric graphs. *arXiv:1712.03630*, 2017.

[OS20a] Steve Oudot and Elchanan Solomon. Inverse problems in topological persistence. In *Topological Data Analysis*, pages 405–433. Springer, 2020.

[OS20b] Steve Oudot and Elchanan Solomon. Inverse problems in topological persistence. In *Topological Data Analysis*, pages 405–433. Springer, 2020.

[Oud15] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, RI, 2015.

[Pan89] Pierre Pansu. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Annals of Mathematics*, pages 1–60, 1989.

[PDHH15] Jose A Perea, Anastasia Deckard, Steve B Haase, and John Harer. Sw1pers: Sliding windows and 1-persistence scoring; discovering periodicity in gene expression time series data. *BMC bioinformatics*, 16(1):1–12, 2015.

[PHR16] Florian T Pokorny, Majd Hawasly, and Subramanian Ramamoorthy. Topological trajectory classification with filtrations of simplicial complexes and persistent homology. *The International Journal of Robotics Research*, 35(1-3):204–223, 2016.

[PMRS17] Siddharth Pal, Terrence J Moore, Ram Ramanathan, and Ananthram Swami. Comparative topological signatures of growing collaboration networks. In *International Workshop on Complex Networks*, pages 201–209. Springer, 2017.

[PS70] Jacob Palis and Stephen Smale. Structural stability theorems. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 223–231. Amer. Math. Soc., Providence, R.I., 1970.

[PS91] Jooyoung Park and Irwin W Sandberg. Universal approximation using radial-basis-function networks. *Neural computation*, 3(2):246–257, 1991.

[PS16] Leonid Polterovich and Egor Shelukhin. Autonomous Hamiltonian flows, Hofer's geometry and persistence modules. *Selecta Math. (N.S.)*, 22(1):227–296, 2016.

[PSO18] Adrien Poulenard, Primoz Skraba, and Maks Ovsjanikov. Topological function optimization for continuous shape matching. In *Computer Graphics Forum*, volume 37, pages 13–25. Wiley Online Library, 2018.

[Qui88] Frank Quinn. Homotopically stratified sets. *Journal of the American Mathematical Society*, 1(2):441–499, 1988.

[RBB19] Bastian Rieck, Christian Bock, and Karsten Borgwardt. A persistent weisfeiler-lehman procedure for graph classification. In *International Conference on Machine Learning*, pages 5448–5458. PMLR, 2019.

[RCK+17] Abbas Rizvi, Pablo Cámara, Elena Kandror, Thomas Roberts, Ira Schieren, Tom Maniatis, and Raúl Rabadán. Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nature Biotechnology*, 35:551–560, 2017.

[RG19] Raif M Rustamov and Leonidas J Guibas. Wavelets on graphs via deep learning. In *Vertex-Frequency Analysis of Graph Signals*, pages 207–222. Springer, 2019.

[RHBK15] Jan Reininghaus, Stefan Huber, Ulrich Bauer, and Roland Kwitt. A stable multi-scale kernel for topological

machine learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4741–4748, 2015.

[RPDB11] Julien Rabin, Gabriel Peyré, Julie Delon, and Marc Bernot. Wasserstein barycenter and its application to texture mixing. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 435–446. Springer, 2011.

[SCM+14] Nikhil Singh, Heather D Couture, JS Marron, Charles Perou, and Marc Niethammer. Topological descriptors of histology images. In *International Workshop on Machine Learning in Medical Imaging*, pages 231–239. Springer, 2014.

[Sho12] Naum Zuselevich Shor. *Minimization methods for non-differentiable functions*, volume 3. Springer Science & Business Media, 2012.

[Sma59] Stephen Smale. Diffeomorphisms of the 2-sphere. *Proceedings of the American Mathematical Society*, 10(4):621–626, 1959.

[SMC07] Gurjeet Singh, Facundo Mémoli, and Gunnar Carlsson. Topological methods for the analysis of high dimensional data sets and 3D object recognition. In *4th Eurographics Symposium on Point-Based Graphics (SPBG 2007)*, pages 91–100. The Eurographics Association, 2007.

[SOG09] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer graphics forum*, volume 28 No. 5, pages 1383–1392. Wiley Online Library, 2009.

[SOW03] Jeffrey J Sutherland, Lee A O'brien, and Donald F Weaver. Spline-fitting with a genetic algorithm: A method for developing classification structure- activity relationships. *Journal of Chemical Information and Computer Sciences*, 43(6):1906–1915, 2003.

[SSVL+11] Nino Shervashidze, Pascal Schweitzer, Erik Jan Van Leeuwen, Kurt Mehlhorn, and Karsten M Borgwardt. Weisfeiler-lehman graph kernels. *Journal of Machine Learning Research*, 12(9), 2011.

[SVP+09] Nino Shervashidze, SVN Vishwanathan, Tobias Petri, Kurt Mehlhorn, and Karsten Borgwardt. Efficient graphlet kernels for large graph comparison. In *Artificial Intelligence and Statistics*, pages 488–495, 2009.

[SWB21] Yitzchak Solomon, Alexander Wagner, and Paul Bendich. A fast and robust method for global topological functional optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 109–117. PMLR, 2021.

[TBI97] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*, volume 50. Siam, 1997.

[Tho69] René Thom. Ensembles et morphismes stratifiés. *Bulletin of the American Mathematical Society*, 75(2):240–284, 1969.

[TKH+15] Dane Taylor, Florian Klimm, Heather A Harrington, Miroslav Kramár, Konstantin Mischaikow, Mason A Porter, and Peter J Mucha. Topological data analysis of contagion maps for examining spreading processes on networks. *Nature communications*, 6(1):1–11, 2015.

[TMB14a] Katharine Turner, Sayan Mukherjee, and Doug M. Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[TMB14b] Katharine Turner, Sayan Mukherjee, and Doug M Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[TMMH14] Katharine Turner, Yuriy Mileyko, Sayan Mukherjee, and John Harer. Fréchet means for distributions of persistence diagrams. *Discrete & Computational Geometry*, 52(1):44–70, 2014.

[Tre09] David Treumann. Exit paths and constructible stacks. *Compositio Mathematica*, 145(6):1504–1532, 2009.

[Ume17] Yuhei Umeda. Time series classification via topological data analysis. *Information and Media Technologies*, 12:228–239, 2017.

[UZ16] Michael Usher and Jun Zhang. Persistent homology and Floer-Novikov theory. *Geom. Topol.*, 20(6):3333–3430, 2016.

[VAB13] Ramanarayan Vasudevan, Aaron Ames, and Ruzena Bajcsy. Persistent homology for automatic determination of human-data based cost of bipedal walking. *Nonlinear Analysis: Hybrid Systems*, 7(1):101–115, 2013.

[VL01] Jan Vlček and Ladislav Lukšan. Globally convergent variable metric method for nonconvex nondifferentiable unconstrained minimization. *Journal of Optimization Theory and Applications*, 111(2):407–430, 2001.

[VSKB10] S Vichy N Vishwanathan, Nicol N Schraudolph, Risi Kondor, and Karsten M Borgwardt. Graph kernels. *The Journal of Machine Learning Research*, 11:1201–1242, 2010.

[VZ17] Saurabh Verma and Zhi-Li Zhang. Hunt for the unique, stable, sparse and fast feature learning on graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 87–97, Red Hook, NY, USA, 2017. Curran Associates Inc.

[Wei52] André Weil. Sur les théoremes de de rham. *Comment. Math. Helv*, 26(1):119–145, 1952.

[Wei94] Shmuel Weinberger. *The topological classification of stratified spaces*. University of Chicago Press, 1994.

[Woo08] Jonathan Woolf. The fundamental category of a stratified space. *arXiv preprint arXiv:0811.2580*, 2008.

[WWK08] Nikil Wale, Ian A Watson, and George Karypis. Comparison of descriptor spaces for chemical compound retrieval and classification. *Knowledge and Information Systems*, 14(3):347–375, 2008.

[XLM18] Kelin Xia, Zhiming Li, and Lin Mu. Multiscale persistent functions for biomolecular structure characterization. *Bulletin of mathematical biology*, 80(1):1–31, 2018.

[XSC+19] Bingbing Xu, Huawei Shen, Qi Cao, Yunqi Qiu, and Xueqi Cheng. Graph wavelet neural network. *arXiv preprint arXiv:1904.07785*, 2019.

[XW14] Kelin Xia and Guo-Wei Wei. Persistent homology analysis of protein structure, flexibility, and folding. *International journal for numerical methods in biomedical engineering*, 30(8):814–844, 2014.

[YKAY16] Jaejun Yoo, Eun Young Kim, Yong Min Ahn, and Jong Chul Ye. Topological persistence vineyard for dynamic functional brain connectivity during resting and gaming stages. *Journal of neuroscience methods*, 267:1–13, 2016.

[YL21] Ka Man Yim and Jacob Leygonie. Optimization of spectral wavelets for persistence-based graph classification. *Frontiers in Applied Mathematics and Statistics*, 7:16, 2021.

[YV15] Pinar Yanardag and SVN Vishwanathan. Deep graph kernels. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1365–1374, 2015.

[YZY+17] Tao Yang, Feipeng Zhang, Galip Yardımcı, Fan Song, Ross Hardison, William Noble, Feng Yue, and Qunhua Li. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Research*, 27(11):1939–1949, 2017.

[ZC05] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.

[Zee63] E Christopher Zeeman. On the dunce hat. *Topology*, 2(4):341–358, 1963.

[Zhu13] Xiaojin Zhu. Persistent homology: An introduction and a new text representation for natural language processing. In *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

[ZW19] Qi Zhao and Yusu Wang. Learning metrics for persistence-based summaries and applications for graph classification. In *Advances in Neural Information Processing Systems*, pages 9859–9870, 2019.

[ZWX+18] Zhen Zhang, Mianzhi Wang, Yijian Xiang, Yan Huang, and Arye Nehorai. Retgk: Graph kernels based on return probabilities of random walks. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 3968–3978, Red Hook, NY, USA, 2018. Curran Associates Inc.

Acknowledgements

First and foremost I would like to thank my supervisors Ulrike Tillmann and Heather Harrington for their invitation to spend these three wonderful years at the Mathematical Institute of Oxford by their side. Thank you both for so many things, in particular for always encouraging imagination and liberty of thought. Ulrike, I was struck by your kindness: with our endless scientific conversations and successive collaborations I felt with you as with a member of my family to whom I could ask without counting. You taught me how to better think about mathematics.

In fact, due to the international crisis, I worked remotely during the last year and a half. During the first part of the thesis I immersed myself into Oxford's magic atmosphere. I loved that I could enter any class or seminar room and learn from the experts about topics ranging from Mathematics to Computer Science and Physics. Along the way, in the corridors, your office, Vidit, would always be open to lost souls. Your late night Perverse Sheaf Theory seminar interspersed with breaks at the Royal Oak's pub exemplified both your humoristic view on mathematics and your infinite curiosity. Surely Naya, Oliver, Barbara, Finn, Tadas, Nicholas, Ambrose, you would agree that our brains sweated a lot during these sessions. Barbara, Naya, Ambrose and Ollie sometimes I laugh just remembering our coffee breaks, travels and puzzlements; I value a lot your distinct forms of folly. Similarly Alexis I can't tell if you are right in your mind considering all the strange mathematical convictions that you would defend uninterruptedly despite delicious brunches in your plate. I warmly keep souvenirs of the diners at Wolfson college and lunches at the Mathematical Institute, of the visits to my colleagues at DATASHAPE in Paris, EPFL in Lausanne and IST in Vienna, and of the fascinating people I met at these various occasions.

One delight of the thesis has been to collaborate with others. Steve, I already knew from my undergraduate studies that you were a dedicated teacher, and during the thesis I learnt that you were a marvelous person to work with as well. You have diligently shared an impressive amount of knowledge, and our intense sessions during Summer 2019 involving frequent travels between France and England form a memorable experience. Ollie, let's face it, you and I bearly understand each other, yet during our conversations we would generate ideas and solve problems. This witnesses a distinct form of intellectual connection. Ambrose, we learnt together the hard way, and the result was very much worth the effort. David, we saw how semi-formal reasonings allow to take a leap into uncharted and deep territory. Greg, I think we found our way towards harmony, and our journey continues. Théo and Mathieu I was first perturbed by your relaxed approaches, and then I realised we would make substantial progress with seemingly no efforts. Sebastiano you are an intriguing aesthete, and I enjoyed debating math with you in Swansea, Kyoto, Vienna, Malaga, Zoom or Paris.

Irradié, Claire, par ta vitalité naturelle et surnaturelle, il a été facile de me consacrer à cette thèse, et je concevrai les aventures les plus folles en te sachant à mes côtés. Tu es ma chance inouïe. Mes grands-parents aussi. Nana, le récit de ta carrière indépendante, plusieurs fois par semaine durant nos appels ou lors de mes visites à Paris, me confort(er)a toujours dans une persévérance imperturbable de 'taupe monomanique'.

Si Rebecca tu deviens Présidente de la République, Maman se vantera certes mais ne manquera pas de signaler que son fils est aussi docteur ; c'est toi Dede qui me l'a appris, tu peux dire 'le contraire' mais pas le contraire! De même je ne conteste plus que ce que je cherche perpétuellement, malgré moi, c'est le bonheur. Ton exemple et le tien également Maman me donnent la v(o)ie et le tournis. Ce manuscript est pour vous, Maman Nana Dede Rebecca et Claire, en ersatz du témoignage indicible de mon amour.

Je n'ai pas la place pour retranscrire fidèlement l'influence de mes familles adorées qui m'a conduit aux portes de cette aventure mathématique. Cela surtout pour des raisons démographiques: Papa en particulier, tu ne me facilites pas la tâche en donnant la vie selon une loi de Poisson de paramètre $\frac{1}{7}$. Tu es largement blanchi par les enseignements que je tire de tes qualités, en particulier de ta capacité singulière de re-création, et aussi par la mise en relation déterminante avec Bernard, à qui je dois une ouverture d'esprit mathématique et une vision humoresthétique de la vie. À vous mes frères et soeurs Rebecca, Adrien, Tamina, Thomien, Martin, Automna, Danil et Paulina, mon soutien vous sera bien entendu indéfectible, quoique proportionnel à ce que vous voudrez bien lire de la présente thèse. D'ailleurs je suis davantage soutenu que je ne soutiens: Eugène, Matthieu, Albert, J.-O., Juliette, Eliot, Anouk, Dominique, Amatxi, Aitatxi, Chantal, Victoire, Laurent, Siméon, Vincent, Daniel, Mioara ... je pense fortement à vous.

Ma pensée échoue finalement aux amitiés et à leur rendre justice. Mouss le tribut -avions ratés, voile de kite cassée, altercations- de nos plans *in extremis* n'aura jamais raison de notre entêtement à s'amuser et de notre amitié libre. Je jubile de savoir où te mènera ta bougeotte, cette hargne intérieure indéfinie et inébranlable, peut-être à trouver et aménager la terre ou l'île qui nous unira tous. On y trouvera Beber, le plus poétique des esprits de contradiction, à qui je dois beaucoup et confierai n'importe quoi, et Zouzoul, créature dangereusement attachante qui tient du mystique, qui me doit beaucoup et confie sans vergogne ses aventures compromettantes. On y accueillera depuis Paris les déesses Bibosh, Dim-Dim et Didou et contrairement à Pâris je me défile: vous êtes toutes trois les plus belles et à jamais dans mon coeur. Lucas tu es un ami incroyable -toutes nos discussions, tes conseils voire même tes influences sur cette thèse, je les chéris- et un co-fondateur idéal: ta souplesse d'esprit, tes convictions, ton écoute attentive et tes compétences concrètes nous ont permis de développer des projets de folie. Enfin merci à toi le Groloulouz', tu m'as également beaucoup appris, comme le fait de 'rentrer dans le métro, avec un pass Navigo, tranquille', ou 'd'envoyer un mail, via gmail', et j'en passe. Travailler avec toi toutes ces années au café, en Corse, n'importe où, fut et sera un amusement sans limite donnant un sens indescriptible à ma vie.

Aux êtres si chers qui m'ont déterminé jadis et à ceux me déterminent encore aujourd'hui, je vous témoigne ma gratitude et vous souhaite tous les bonheurs, et si mon coeur a ses raisons que la raison ignore, il se remémore globalement tout, et localement, pêle-mêle, Big D, Lena, Sash', Dr. Liloï, Jean de Dieu, Iled, Jules, Samuel, Luc, Bibaum, Tobaï, The Tchômen, Léo, Gose-Bird, Kimai, Zaz, Commissaire, ... [*Error List index out of range: LaTeX can't compile the complete list of people you are indebted to.*]