

Zigzag Zoology: Rips Zigzags for Homology Inference

Steve Oudot, Donald Sheehy

▶ **To cite this version:**

Steve Oudot, Donald Sheehy. Zigzag Zoology: Rips Zigzags for Homology Inference. [Research Report] RR-8141, INRIA. 2012, pp.48. hal-00755280

HAL Id: hal-00755280
https://inria.hal.science/hal-00755280

Submitted on 23 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

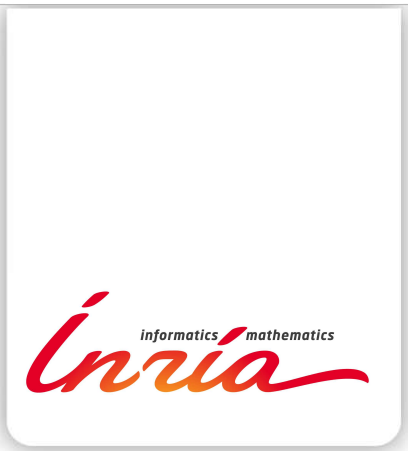


Zigzag Zoology: Rips Zigzags for Homology Inference

Steve Y Oudot, Donald R. Sheehy

RESEARCH REPORT
N° 8141
November 2012
Project-Team Geometrica

ISSN 0249-6399 ISRN INRIA/RR--8141--FR+ENG



Zigzag Zoology: Rips Zigzags for Homology Inference

Steve Y Oudot, Donald R. Sheehy

Project-Team Geometrica

Research Report n° 8141 — November 2012 — 48 pages

Abstract: For points sampled near a compact set X, the persistence barcode of the Rips filtration built from the sample contains information about the homology of X as long as X satisfies some geometric assumptions. The Rips filtration is prohibitively large, however zigzag persistence can be used to keep the size linear. We present several species of Rips-like zigzags and compare them with respect to the signal-to-noise ratio, a measure of how well the underlying homology is represented in the persistence barcode relative to the noise in the barcode at the relevant scales. Some of these Rips-like zigzags have been available as part of the Dionysus library for several years while others are new. Interestingly, we show that some species of Rips zigzags will exhibit less noise than the (non-zigzag) Rips filtration itself. Thus, the Rips zigzag can offer improvements in both size complexity and signal-to-noise ratio.
Along the way, we develop new techniques for manipulating and comparing persistence barcodes from zigzag modules. We give methods for reversing arrows and removing spaces from a zigzag. We also discuss factoring zigzags and a kind of interleaving of two zigzags that allows their barcodes to be compared. These techniques were developed to provide our theoretical analysis of the signal-to-noise ratio of Rips-like zigzags, but they are of independent interest as they apply to zigzag modules generally.

Key-words: Zigzag persistence, homology inference, Rips filtrations, applied topology

RESEARCH CENTRE
SACLAY – ÎLE-DE-FRANCE

Parc Orsay Université
4 rue Jacques Monod
91893 Orsay Cedex

Zoologie des zigzags à base de complexes de Rips pour l'inférence homologique

Résumé : Pour des points échantillonnés près d'un compact X, le code-barre de la filtration de Rips construite sur les points contient de l'information à propos de l'homologie de X sous certaines hypothèses géométriques. Toutefois, le coût de construction de la filtration de Rips est prohibitif, mais la persistence des zigzags peut permettre de le rendre linéaire en le nombre de points de données. Nous présentons plusieurs types de zigzags basés sur des complexes de Rips, et nous les comparons à l'aune de leur rapport signal sur bruit. Certains de ces zigzags sont disponibles dans la bibliothèque Dionysus depuis plusieurs années, tandis que d'autres sont nouveaux. Il est intéressant d'observer que certains ont des codes-barres avec significativement moins de bruit que celui de la filtration de Rips standard. Ainsi, les zigzags à base de complexes de Rips permettent-ils non seulement de réduire la complexité mais également d'améliorer le résultat de l'approche. Dans notre analyse nous développons de nouveaux outils pour manipuler les zigzags et comparer leurs codes-barres. En particulier, nous fournissons des méthodes pour inverser des flèches ou enlever des espaces dans un zigzag, tout en contrôlant l'impact sur son code-barre. Nous parlons également de factorisation et d'entrelacement de zigzags. Ces outils sont la clef de voûte de notre analyse, et ils présentent un intérêt indépendant puisqu'ils s'appliquent aux zigzags en général, et non pas seulement à ceux étudiés ici.

Mots-clés : Persistance des zigzags, inférence homologique, filtrations de Rips, topologie appliquée

1 Introduction

The goal of homology inference is to extract the homology of a space from a finite sample. The problem is ill-posed in general, but under the right geometric assumptions about the input and the underlying space, one can compute an object called a *persistence barcode* which provably contains information about the underlying homology. Indeed, homology inference was and continues to be one of the main motivations for the development of topological persistence theory.

The barcode is computed from a sequence of simplicial complexes, for which two main challenges arise. The first challenge is to guarantee that the simplicial complexes remain small. Commonly used methods produce complexes that quickly become too large to fit in memory. The topological signal is the information about the underlying space contained in the barcode. The second challenge is to decrease noise in the barcode that can obscure the topological signal while still guaranteeing that the signal remains. We confront both of these challenges, analyze several approaches that give linear size data structures, and provide theoretical guarantees on the signal-to-noise ratio in the barcodes.

The standard persistence theory applies to nested, parameterized families of simplicial complexes called *filtrations*. The persistence algorithm takes a filtration and produces a barcode describing all the changes in homology as one goes from one complex to the next in the filtration [14, 23].

Persistent homology has an important connection with geometric inference results that describe conditions when homology inference is possible using a union of balls centered at the sample points (see the survey by Chazal and Cohen-Steiner [4]). The *(Vietoris-)Rips filtration* $\{\mathcal{R}_\alpha\}_{\alpha\geq 0}$ is useful when these conditions are met. It is defined to have a simplex in $\mathcal{R}_\alpha$ for every subset of points with diameter at most α. So, the filtration parameter is the geometric scale and the existing theory guarantees the existence of some range of scales for which the barcode encodes the homology of the underlying space. The barcode of this filtration thus has an elegant multi-scale interpretation of the results as being "the homology of the input point cloud across scales."

The immediate drawback to using the Rips filtration is that it quickly becomes so large that it no longer fits in main memory. The scale of this breaking point varies with the input data as well as with the filtration and computer used, however, it is observed to happen early enough so that not all the interesting homological information hidden in the data can be discovered — see [19] for a compelling example. Recent research looks at how to reduce the size of the filtration elements and thereby to postpone the breaking point. For example Chazal and Oudot [9] use truncated filtrations on a nested sequence of subsets of the input points corresponding to samplings at different scales. They compute the barcodes of the Rips filtration of each subset restricted to a range of scales near the sampling scale of the subset. This can prevent the size blowup in the Rips filtrations because every subset looks like a uniform sample at the relevant scale. The lingering challenge from this work is to relate the bars in the resulting barcodes for different scales.

Taking advantage of the recent introduction of zigzag persistence by Carlsson and de Silva [3], Morozov suggested a simple way to connect the truncated Rips filtrations of consecutive subsamples together, to obtain a single long sequence of simplicial complexes connected by inclusions — called the *Morozov zigzag* hereafter. Zigzag persistence relaxes the condition that the family of complexes be a filtration and instead allows consecutive spaces to be included in either direction, forwards or backwards, so the sequence is a zigzag diagram rather than a filtration. The Morozov zigzag has been integrated into the Dionysus library [13] since early 2009, and as reported by its author from preliminary experiments [21], it has given surprisingly good results in practice. However, to date it comes with no theoretical guarantees, so a primary motivation of our paper

is to assess the theoretical quality of the results provided by this zigzag.

Existing methods for building sparse approximations to the Rips filtration have all focused on the size question, but have ignored the question of noise. For example, even the Rips filtration can have noise in the barcode at the scales where it represents the underlying topology. We show that a generalization of the Morozov zigzag not only recovers the topological signal but also provably eliminates noise in the relevant range.

Contributions. We provide the following theoretical guarantees for the Morozov zigzag:

- When the input point cloud P is sufficiently close (in the Hausdorff distance) to a compact set X with positive weak feature size in $\mathbb{R}^d$, there is a *sweet range* of geometric scales over which the persistence barcode of the Morozov zigzag exhibits the homology of X (technically, the offsets X^λ for an arbitrarily small $\lambda > 0$). That is, the barcode has long intervals spanning the entire sweet range, and their number is at least the dimension of the homology group $H_*(X^\lambda)$ — Theorem 5.3.
- There is a smaller (*sweeter*) range over which the number of spanning intervals is exactly $\dim H_*(X^\lambda)$. The other intervals in the sweeter range are ephemeral (length zero) and can therefore be ignored.
  - For the 0-th and 1-st homology, the sweeter range is as large as the sweet range. As a consequence, the 0-th and 1-st homology of X^λ can be read from the barcode of the Morozov zigzag when X has positive weak feature size — Theorem 5.6.
  - For the k-th homology with $k \geq 2$, our proof of existence of a sweeter range requires X to have positive μ-reach — Section 5.3.2. It remains an open question whether there exists a sweeter range for k-th homology when X has zero μ-reach and positive weak feature size. Although it is possible, there is no theoretical evidence to suggest that it always exists even for small, non-zero μ-reach.

This motivates the study of more elaborate variants of the Morozov zigzag that are less likely to carry topological noise in the sweet range, even when the underlying space X has zero μ-reach and positive weak feature size. We analyze three variants in the paper:

- The first one, called the *discretized Mozorov zigzag*, consists in considering only subsamples whose corresponding geometric scales are of the form c^i for a fixed constant c and an integer i. This discretization makes sure that the geometric scale drops significantly (by a factor of c) from one subsample to the next, so there is enough room in each connection between truncated filtrations to kill the noise.
- The second one, called the *oscillating Rips zigzag*, consists in somewhat relaxing the truncation parameter in the Rips filtrations before connecting them together. The effect is to leave enough room in every truncated filtration for the noise to be killed.
- Finally, the third one, called the *image Rips zigzag*, consists in taking a nested pair of Morozov zigzags with different filtration parameters, and in connecting them by canonical inclusions to obtain an image zigzag module at the homology level. Taking a pair of zigzags instead of single zigzag kills the noise in the same way as taking a pair of Rips complexes instead of a single Rips complex did in [9].

Each one of these variants comes with the desired guarantee that the sweet and sweeter ranges are equal, meaning that there is guaranteed to be only ephemeral noise in the sweet range even when the underlying space X merely has positive weak feature size. Consequently, the homology of X^λ can be inferred from its persistence barcode. Moreover, the noise within the sweet ranges is ephemeral, so the barcodes exhibit less noise than the standard (non-zigzag) Rips filtration itself. Thus, Rips zigzags offer improvements in both size complexity and signal-to-noise ratio compared to the Rips filtration. The price to pay compared to the basic Morozov zigzag is

a somewhat increased time and/or space complexity — Theorems 6.5 and 6.6. The overhead depends on the variant considered but it always remains bounded, so the variants are tractable alternatives to the Morozov zigzag in practice.

To prove the aforementioned results we develop new techniques for manipulating zigzag modules and comparing their persistence barcodes. More precisely:

- We show how arrows in a zigzag module can be reversed while preserving the persistence barcode — Theorem 3.1. Applied repeatedly, this result makes it possible to turn zigzag modules into standard persistence modules, eventually leading to a stability result for zigzag modules — Theorem 6.4.
- We give a method for removing spaces from a zigzag module while tracking the intervals in its barcode — Theorem 3.2. For instance, this result tells how the persistence barcode of a zigzag homology module evolves when an inclusion between simplicial complexes at the topological level is replaced by a sequence of elementary inclusions where one simplex is added or removed at a time.

These low-level manipulations on zigzag modules enable us to derive higher-level comparison theorems. In particular, we discuss factoring zigzag modules (Theorem 4.1), as well as a special kind of interleaving between zigzags that allows their barcodes to be compared (Theorem 4.2). These theorems are the cornerstones of our proofs of the aforementioned guarantees on the Morozov zigzag and its variants. However, they are also of independent interest as they apply to zigzag modules generally.

Related work. A different approach to the problem of building sparse filtrations for offsets of point clouds in Euclidean space was presented by Hudson et al. [19]. They used ideas from Delaunay refinement mesh generation to build linear size filtrations that provide provably good approximations to the persistence diagram of the offsets. However, that approach requires building a complex that covers the ambient space and includes simplices up to its dimension. Moreover, the construction requires the use of high degree predicates. In contrast, the new methods described here only depend on an intrinsic dimension of data and can be built using only distances comparisons.

Recently, Sheehy [22] proposed a method for building a sparse zigzag filtration whose barcode is provably close to that of the Rips filtration as well as a non-zigzagging variant achieving similar guarantees. Also, Dey et al. gave an alternative persistence algorithm for simplicial maps rather than inclusions, which is closely related to zigzag persistence [12]. Their approach similarly gives barcodes that are provably close to that of the Rips filtration. We obtain comparable space/time bounds to these results but get stronger guarantees regarding noise. Methods that approximate the Rips filtration directly can, in principle, have noise that is as large as the noise in the Rips filtration itself (or worse).

Paper layout. After introducing the necessary background in Section 2, we present our low-level zigzags manipulations in Section 3, then our high-level zigzags comparison theorems in Section 4. These theorems are used to analyze the properties of Morozov zigzags and their variants in Section 5. Finally, we discuss the impact and limitations of our results in Section 6, before presenting some experimental data in Section 7.

2 Background

Section 2.1 gives a brief overview of the concepts and results from zigzag persistence theory that will be used in the algebraic part of our analysis. Our terminology is the same as in [3] up to a few minor variants, and we refer the reader to that paper for a more in-depth treatment.

Section 2.2 introduces some concepts and results from the sampling theory for compact sets in Euclidean spaces, which will be used in the geometric part of our analysis. We refer the reader to [4] for a comprehensive survey on this topic.

2.1 Zigzag persistence

A *zigzag module* $\mathbb{V}$ is a finite diagram of finite-dimensional vector spaces over a same field $\mathbb{F}$:

$$\mathbb{V} = V_1 \overset{f_1}{\leftrightarrow} V_2 \overset{f_2}{\leftrightarrow} \cdots \overset{f_{n-1}}{\leftrightarrow} V_n,$$

where the notation $V_i \overset{f_i}{\leftrightarrow} V_{i+1}$ indicates that the linear map f_i can be either a forward map ($f_i : V_i \to V_{i+1}$) or a backward map ($f_i : V_i \leftarrow V_{i+1}$). An equivalent notation is $f_i : V_i \leftrightarrow V_{i+1}$. The sequence of map orientations (forwards or backwards) defines the *type* of the module $\mathbb{V}$. A *persistence module*, as defined in the standard (non-zigzag) persistence literature [23], is a zigzag module in which all the maps are oriented forwards. Thus, all persistence modules have the same type.

A *submodule* $\mathbb{W}$ of a zigzag module $\mathbb{V}$ is defined by subspaces $W_i \leq V_i$ such that for all i we have $f_i(W_i) \subseteq W_{i+1}$ if $f_i : V_i \leftrightarrow V_{i+1}$ is a forward map and $f_i(W_{i+1}) \subseteq W_i$ if f_i is a backward map. The maps in $\mathbb{W}$ are then the restrictions of the maps in $\mathbb{V}$ to the W_i's, which makes the types of $\mathbb{V}$ and $\mathbb{W}$ the same. $\mathbb{W}$ is called a *summand* of $\mathbb{V}$ if there exists another submodule $\mathbb{X}$ of $\mathbb{V}$ such that $V_i = W_i \oplus X_i$ for all i. In that case, we say that $\mathbb{V}$ is the *direct sum* of $\mathbb{W}$ and $\mathbb{X}$, written $\mathbb{V} = \mathbb{W} \oplus \mathbb{X}$. As pointed out in [3], all summands are submodules but not all submodules are summands.

A zigzag module $\mathbb{V}$ is called *indecomposable* if it admits no nonzero summands. It is known since Gabriel [15] that the indecomposable zigzag modules are the so-called *interval modules*. Given a module type τ and an integer interval $[b, d]$, the interval τ-module with birth time b and death time d is written $\mathbb{I}_\tau[b, d]$ and defined with spaces I_i such that $I_i = \mathbb{F}$ if $i \in [b, d]$ and $I_i = 0$ otherwise, and with identity maps between adjacent copies of the base field $\mathbb{F}$ and zero maps elsewhere (the maps are oriented according to τ). Another way of stating Gabriel's result is to say that every τ-module can be written as a direct sum of τ-intervals. Moreover, it follows from the Krull-Schmidt principle that this decomposition is unique up to a reordering of the terms — see Proposition 2.2 in [3]. We gather these facts into a single statement:

Theorem 2.1 (Interval Decomposition)**.** *Every τ-module can be written uniquely (up to reordering) as a direct sum of τ-intervals.*

Thus, as in standard (non-zigzag) persistence theory, the structure of a zigzag module $\mathbb{V}$ is fully and uniquely described by a multiset of integer intervals, called the *persistence barcode* of $\mathbb{V}$ and noted $\mathsf{Pers}(\mathbb{V})$. Given an interval $[b, d]$, we write $\mathrm{mult}([b, d],\ \mathbb{V})$ for the multiplicity of $[b, d]$ in the multiset $\mathsf{Pers}(\mathbb{V})$, which is also the number of copies of the interval module $\mathbb{I}_r[b, d]$ in the decomposition of $\mathbb{V}$.

Carlsson and de Silva gave a constructive proof of Theorem 2.1 — see [3, Thm. 4.1], which lead to an algorithm for computing the decompositions of zigzag modules. Among the concepts and results presented in their paper, the following ones play an important part here.

Given a zigzag module $\mathbb{V} = V_1 \overset{f_1}{\leftrightarrow} \cdots \overset{f_{n-1}}{\leftrightarrow} V_n$ and two integers $p \leq q \in [1, n]$, let $\mathbb{V}[p, q]$ denote the restriction of $\mathbb{V}$ to the index set $p \leq i \leq q$:

$$\mathbb{V}[p, q] = V_p \overset{f_p}{\leftrightarrow} V_{p+1} \overset{f_{p+1}}{\leftrightarrow} \cdots \overset{f_{q-1}}{\leftrightarrow} V_q,$$

and let $\mathsf{Pers}(\mathbb{V})|_{[p,q]}$ denote the restriction of $\mathsf{Pers}(\mathbb{V})$ to $[p, q]$:

$$\mathsf{Pers}(\mathbb{V})|_{[p,q]} = \{[b, d] \cap [p, q] \mid [b, d] \in \mathsf{Pers}(\mathbb{V})\}.$$

The restriction principle [3, Prop. 2.12] states that the restrictions of a module and of its barcode behave as expected, namely:

Theorem 2.2 (Restriction)**.** $\mathsf{Pers}(\mathbb{V}[p, q]) = \mathsf{Pers}(\mathbb{V})|_{[p,q]}$.

In other words,

$$\forall b \leq d \in [p, q],\ \mathrm{mult}([b, d],\ \mathbb{V}[p, q]) = \begin{cases} \mathrm{mult}([b, d], \mathbb{V}) \text{ if } p < b \text{ and } d < q, \\ \sum_{d' \geq q} \mathrm{mult}([b, d'], \mathbb{V}) \text{ if } p < b \text{ and } d = q, \\ \sum_{b' \leq p} \mathrm{mult}([b', d], \mathbb{V}) \text{ if } p = b \text{ and } d < q, \\ \sum_{[b',d'] \supseteq [p,q]} \mathrm{mult}([b', d'], \mathbb{V}) \text{ if } p = b \text{ and } d = q. \end{cases}$$

Given a zigzag module $\mathbb{V} = V_1 \overset{f_1}{\leftrightarrow} \cdots \overset{f_{n-1}}{\leftrightarrow} V_n$, the *right filtration* of $\mathbb{V}$, noted $\mathcal{R}_\mathbb{V}$, is a filtration (i.e. a nested sequence of subspaces) of the vector space V_n, defined recursively as follows:

- if $n = 1$, then $\mathcal{R}_\mathbb{V} = (0, V_1)$;
- else $(n > 1)$,

$$\mathcal{R}_\mathbb{V} = \begin{cases} (f_{n-1}(R_0), \cdots, f_{n-1}(R_{n-1}), V_n) \text{ if } f_{n-1} : V_{n-1} \to V_n \\ (0, f_{n-1}^{-1}(R_0), \cdots, f_{n-1}^{-1}(R_{n-1})) \text{ if } f_{n-1} : V_{n-1} \leftarrow V_n, \end{cases}$$

where $(R_0, \cdots, R_{n-1})$ is the right filtration of $\mathbb{V}[1, n-1]$.

We write $\mathcal{R}_\mathbb{V}[k]$ for the k-th element in the right filtration of $\mathbb{V}$. Note that the recursive definition maintains the filtration property, that is,

$$0 = \mathcal{R}_\mathbb{V}[0] \leq \mathcal{R}_\mathbb{V}[1] \leq \cdots \leq \mathcal{R}_\mathbb{V}[n-1] \leq \mathcal{R}_\mathbb{V}[n] = V_n.$$

The *left filtration* $\mathcal{L}_\mathbb{V}$ is defined symmetrically as the right filtration of the reversal of $\mathbb{V}$. It is thus a filtration of the space V_1.

The *birth-time index* $\mathrm{b}_\mathbb{V}$ is a vector of integers defined recursively as follows:

- if $n = 1$, then $\mathrm{b}_\mathbb{V} = (1)$;
- else $(n > 1)$,

$$\mathrm{b}_\mathbb{V} = \begin{cases} (b_1, \cdots, b_{n-1}, n) \text{ if } f_{n-1} : V_{n-1} \to V_n \\ (n, b_1, \cdots, b_{n-1}) \text{ if } f_{n-1} : V_{n-1} \leftarrow V_n, \end{cases}$$

where $(b_1, \cdots, b_{n-1})$ is the birth-time index of $\mathbb{V}[1, n-1]$.

In other words, $\mathrm{b}_\mathbb{V}$ stores a permutation of the index set $[1, n]$, and it derives from the type τ of the module $\mathbb{V}$ in a similar way as $\mathcal{R}_\mathbb{V}$ derives from $\mathbb{V}$ itself. The k-th element in $\mathrm{b}_\mathbb{V}$, denoted by $\mathrm{b}_\mathbb{V}[k]$, corresponds to the image of k through the permutation. The *death-time index* $\mathrm{d}_\mathbb{V}$ is defined (almost) symmetrically as $n + 1 - \mathrm{b}_{\bar{\mathbb{V}}}$, where $\bar{\mathbb{V}}$ is the reversal of $\mathbb{V}$. The Localization Theorem [3, Thm. 5.3] describes how these various quantities interplay with each other in the expression of the interval multiplicities in the barcode of $\mathbb{V}$:

Theorem 2.3 (Localization)**.** *Given an index $k \in [1,n]$, for all i,j in the range $1 \leq i \leq k$, $1 \leq j \leq n+1-k$,*

$$\begin{aligned}\mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}[1,k]}[i], \mathrm{d}_{\mathbb{V}[k,n]}[j]\right], \mathbb{V}) &= \dim(\mathcal{R}_{\mathbb{V}[1,k]}[i] \cap \mathcal{L}_{\mathbb{V}[k,n]}[j]) \\ &\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[i-1] \cap \mathcal{L}_{\mathbb{V}[k,n]}[j]) \\ &\quad + \dim(\mathcal{R}_{\mathbb{V}[1,k]}[i-1] \cap \mathcal{L}_{\mathbb{V}[k,n]}[j-1]) \\ &\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[i] \cap \mathcal{L}_{\mathbb{V}[k,n]}[j-1]).\end{aligned}$$

2.2 Critical Point Theory for Distance Functions

The geometric part of our analysis takes place in Euclidean space $\mathbb{R}^d$, where $\|\cdot\|$ denotes the Euclidean norm. The distance from a point y to a set $X \subset \mathbb{R}^d$ is $\mathrm{d}(y,X) = \inf_{x \in X} \|x-y\|$. When X is compact, the infimum becomes a minimum, and we let d_X denote the function *distance to X*.

$$\forall y \in \mathbb{R}^d,\ \mathrm{d}_X(y) \stackrel{\text{def}}{=} \mathrm{d}(y,X) = \min_{x \in X} \|x-y\|.$$

The α-offset of X is the locus of the points of $\mathbb{R}^d$ whose distance to X is at most α:

$$X^\alpha \stackrel{\text{def}}{=} \mathrm{d}_X^{-1}([0,\alpha]).$$

Although d_X may not be differentiable everywhere in $\mathbb{R}^d$, its gradient can be extended to be well-defined over all $\mathbb{R}^d$ [6]. The extended gradient is denoted ∇_X in the following. When working with offsets, it is useful to observe that for $y \notin X^\alpha$, $\mathrm{d}_{X^\alpha}(y) = \mathrm{d}_X(y) - \alpha$, so $\nabla_{X^\alpha}(y) = \nabla_X(y)$.

Definition 2.4. *A* critical point *of the distance function d_X to a compact set $X \subset \mathbb{R}^d$ is a point p of $\mathbb{R}^d \setminus X$ such that $\nabla_X(p) = 0$. Equivalently, a* critical point *is a point of $\mathbb{R}^d \setminus X$ that is in the convex hull of its nearest points in X. A number $r \in \mathbb{R}$ is a* critical value *if there exists a critical point p such that $\mathrm{d}_X(p) = r$.*

Definition 2.5. *The* weak feature size *of a compact set X, noted $\mathrm{wfs}(X)$, is the smallest critical value of its distance function d_X.*

Given $X \subseteq \mathbb{R}^d$ and $\beta \geq \alpha \geq 0$, we let $H_*(X^\beta_\alpha)$ denote the image of the homomorphism $H_*(X^\alpha) \to H_*(X^\beta)$ induced at the homology level[1] by the canonical inclusion $X^\alpha \hookrightarrow X^\beta$.

Lemma 2.6 ([9])**.** *Let X be a compact set and P a finite set in $\mathbb{R}^d$, such that $\mathrm{d}_H(X,P) < \varepsilon$ for some $\varepsilon < \frac{1}{4}\mathrm{wfs}(X)$. Then, $H_*(P^\beta_\alpha) \cong H_*(X^\lambda)$ for any $\alpha, \beta \in [\varepsilon, \mathrm{wfs}(X) - \varepsilon]$ such that $\beta - \alpha \geq 2\varepsilon$, and for any $\lambda \in (0, \mathrm{wfs}(X))$.*

For any finite sets $P \subseteq Q \subset \mathbb{R}^d$ and any non-negative parameters $\alpha, \alpha', \beta, \beta'$ such that $\alpha \leq \alpha'$, $\beta \leq \beta'$, $\alpha \leq \beta$ and $\alpha' \leq \beta'$, we have the following commutative diagram where all linear maps are induced by inclusions:

$$\begin{array}{ccc} H_*(P^\beta) & \to & H_*(Q^{\beta'}) \\ \uparrow & & \uparrow \\ H_*(P^\alpha) & \to & H_*(Q^{\alpha'}) \end{array} \tag{1}$$

This commutative diagram induces a homomorphism $H_*(P^\beta_\alpha) \to H_*(Q^{\beta'}_{\alpha'})$.

[1] Throughout the paper we use singular homology with coefficients in a field — omitted in our notations.

Lemma 2.7. *Let X be a compact set and $P \subseteq Q$ be finite sets in $\mathbb{R}^d$, such that $\mathrm{d}_H(X,P) < \varepsilon$ and $\mathrm{d}_H(Q,X) < \varepsilon$ for some $\varepsilon < \frac{1}{6}\mathrm{wfs}(X)$. Then, for any $\alpha, \alpha', \beta, \beta' \in [3\varepsilon, \mathrm{wfs}(X) - \varepsilon]$ such that $\beta - \alpha \geq 2\varepsilon$, $\beta' - \alpha' \geq 2\varepsilon$, $\alpha' \geq \alpha$ and $\beta' \geq \beta$, the linear map $H_*(P_\alpha^\beta) \to H_*(Q_{\alpha'}^{\beta'})$ induced by the diagram (1) is an isomorphism.*

Proof. According to Lemma 2.6, $H_*(P_\alpha^\beta)$ and $H_*(Q_{\alpha'}^{\beta'})$ are isomorphic vector spaces, therefore all we need to show is that $\operatorname{rank} H_*(P_\alpha^\beta) \to H_*(Q_{\alpha'}^{\beta'}) = \dim H_*(P_\alpha^\beta)$. We have the following commutative diagram where all the maps are induced by inclusions (note that $Q^{\alpha-2\varepsilon} \subseteq P^\alpha$ since $\mathrm{d}_H(P,Q) \leq \mathrm{d}_H(P,X) + \mathrm{d}_H(Q,X) \leq 2\varepsilon$):

$$\begin{array}{ccc} H_*(P^\beta) & \xrightarrow{b} & H_*(Q^{\beta'}) \\ {\scriptstyle a}\uparrow & & \uparrow{\scriptstyle d} \\ H_*(P^\alpha) & \xrightarrow{c} & H_*(Q^{\alpha'}) \\ & {\scriptstyle e}\nwarrow & \uparrow{\scriptstyle f} \\ & & H_*(Q^{\alpha-2\varepsilon}) \end{array}$$

The homomorphism $H_*(P_\alpha^\beta) \to H_*(Q_{\alpha'}^{\beta'})$ we are interested in is the restriction of b to $\operatorname{im} a$, whose rank is the same as the one of $b \circ a$. By composition and commutativity, we have

$$\operatorname{rank} d \circ f = \operatorname{rank} b \circ a \circ e \leq \operatorname{rank} b \circ a \leq \operatorname{rank} a,$$

and by Lemma 2.6 we have $\operatorname{rank} d \circ f = \operatorname{rank} a = \dim H_*(P_\alpha^\beta)$ since $\alpha - 2\varepsilon \geq \varepsilon$ and $\beta' \leq \mathrm{wfs}(X) - \varepsilon$. Hence, $\operatorname{rank} b \circ a = \dim H_*(P_\alpha^\beta)$. □

Combining the above analysis with the Persistent Nerve Lemma[2] [9, Lemma 3.4], we obtain the following result where the notation $H_*(\mathcal{C}_\alpha^\beta(P))$ stands for the image of the homomorphism $H_*(\mathcal{C}_\alpha(P)) \to H_*(\mathcal{C}_\beta(P))$ induced at the homology level by the inclusion $\mathcal{C}_\alpha(P) \hookrightarrow \mathcal{C}_\beta(P)$.

Theorem 2.8. *Let X be a compact set and P and Q be finite sets in $\mathbb{R}^d$ with $P \subseteq Q$, such that $\mathrm{d}_H(P,X) < \varepsilon$ and $\mathrm{d}_H(Q,X) < \varepsilon$.*

(i) *If $\varepsilon < \frac{1}{4}\mathrm{wfs}(X)$, then for any $\alpha, \beta \in [\varepsilon, \mathrm{wfs}(X) - \varepsilon]$ such that $\beta - \alpha \geq 2\varepsilon$, for any $\lambda \in (0, \mathrm{wfs}(X))$, the spaces $H_*(\mathcal{C}_\alpha^\beta(P))$ and $H_*(X^\lambda)$ are isomorphic.*

(ii) *If $\varepsilon < \frac{1}{6}\mathrm{wfs}(X)$, then for any $\alpha, \alpha', \beta, \beta' \in [3\varepsilon, \mathrm{wfs}(X) - \varepsilon]$ such that $\beta - \alpha \geq 2\varepsilon$, $\beta' - \alpha' \geq 2\varepsilon$, $\alpha' \geq \alpha$ and $\beta' \geq \beta$, the homomorphism $H_*(\mathcal{C}_\alpha^\beta(P)) \to H_*(\mathcal{C}_{\alpha'}^{\beta'}(Q))$ induced by the following commutative diagram (where the maps are induced by inclusions) is an isomorphism.*

$$\begin{array}{ccc} H_*(\mathcal{C}_\beta(P)) & \to & H_*(\mathcal{C}_{\beta'}(Q)) \\ \uparrow & & \uparrow \\ H_*(\mathcal{C}_\alpha(P)) & \to & H_*(\mathcal{C}_{\alpha'}(Q)) \end{array}$$

[2] We are in fact using an extended version of the Persistent Nerve Lemma, stated in [8], where the index sets of the open covers may differ.

3 Manipulating Zigzag Modules

Suppose we have a zigzag module $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_n$ and we want to reverse the map $V_k \leftrightarrow V_{k+1}$ for some arbitrary index k in the range $[1, n-1]$, while preserving the persistence barcode of $\mathbb{V}$. The following theorem shows that this is always possible, moreover with a reverse map that is closely tied to the original map.

Theorem 3.1 (Arrow Reversal)**.** *Let $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_k \overset{f}{\leftrightarrow} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ be a zigzag module. Then, there is a map $g : V_k \leftrightarrow V_{k+1}$ oriented opposite to f, such that $f \circ g|_{\operatorname{im} f} = \mathbb{1}_{\operatorname{im} f}$ and $g \circ f|_{\operatorname{im} g} = \mathbb{1}_{\operatorname{im} g}$, and the zigzag module $\mathbb{V}^*$ obtained from $\mathbb{V}$ by replacing the submodule $V_k \overset{f}{\leftrightarrow} V_{k+1}$ by $V_k \overset{g}{\leftrightarrow} V_{k+1}$ has the same persistence barcode as $\mathbb{V}$.*

Observe that when f is injective, g is surjective and $g \circ f$ is the identity over the domain of f. Conversely, when f is surjective, g is injective and $f \circ g$ is the identity over the codomain of f. The analysis of these special cases will be the first step in our proof of the theorem.

Suppose we have a zigzag module $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_n$ and we want to remove a space V_k from the sequence while preserving most of the persistence barcode of $\mathbb{V}$ — except intervals bounded at index k since these disappear from the sequence. The following theorem shows that this is always possible, and the map h connecting V_{k-1} and V_{k+1} can be chosen with some other useful properties.

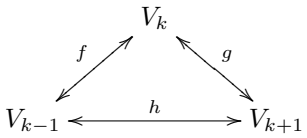
Theorem 3.2 (Space Removal)**.** *Let $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_{k-1} \overset{f}{\leftrightarrow} V_k \overset{g}{\leftrightarrow} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ be a zigzag module. Then, there is a map $h : V_{k-1} \leftrightarrow V_{k+1}$ such that the zigzag module $\mathbb{V}^*$ obtained from $\mathbb{V}$ by replacing the submodule $V_{k-1} \overset{f}{\leftrightarrow} V_k \overset{g}{\leftrightarrow} V_{k+1}$ by $V_{k-1} \overset{h}{\leftrightarrow} V_{k+1}$ has a persistence barcode that derives from the barcode of $\mathbb{V}$ as follows:*

(a) for $i \leq k-1$, $\operatorname{mult}([i, k-1], \mathbb{V}^*) = \operatorname{mult}([i, k-1], \mathbb{V}) + \operatorname{mult}([i, k], \mathbb{V})$,

(b) for $j \geq k+1$, $\operatorname{mult}([k+1, j], \mathbb{V}^*) = \operatorname{mult}([k+1, j], \mathbb{V}) + \operatorname{mult}([k, j], \mathbb{V})$,

(c) for any other interval $[i, j]$ with $i, j \neq k$, $\operatorname{mult}([i, j], \mathbb{V}^*) = \operatorname{mult}([i, j], \mathbb{V})$.

Furthermore, the map h makes the following triangle commute when applicable:



More precisely:

- $h = g \circ f$ *when* $V_{k-1} \overset{f}{\to} V_k \overset{g}{\to} V_{k+1}$,
- $h = f \circ g$ *when* $V_{k-1} \overset{f}{\leftarrow} V_k \overset{g}{\leftarrow} V_{k+1}$,
- $f = g \circ h$ *when* $V_{k-1} \overset{f}{\to} V_k \overset{g}{\leftarrow} V_{k+1}$ *and* $\operatorname{im} f \subseteq \operatorname{im} g$,
- $g = f \circ h$ *when* $V_{k-1} \overset{f}{\to} V_k \overset{g}{\leftarrow} V_{k+1}$ *and* $\operatorname{im} g \subseteq \operatorname{im} f$,
- $f = h \circ g$ *when* $V_{k-1} \overset{f}{\leftarrow} V_k \overset{g}{\to} V_{k+1}$ *and* $\ker g \subseteq \ker f$,
- $g = h \circ f$ *when* $V_{k-1} \overset{f}{\leftarrow} V_k \overset{g}{\to} V_{k+1}$ *and* $\ker f \subseteq \ker g$.

Note that the condition $\operatorname{im} f \subseteq \operatorname{im} g$ is not restrictive in the sense that it is a requirement for the existence of maps $h : V_{k-1} \to V_{k+1}$ such that $f = g \circ h$. Similarly, $\operatorname{im} g \subseteq \operatorname{im} f$ is required for $g = f \circ h$, $\ker g \subseteq \ker f$ is required for $f = h \circ g$, $\ker f \subseteq \ker g$ is required for $g = h \circ f$. When these conditions are not satisfied, the theorem still provides maps $h : V_{k-1} \leftrightarrow V_{k+1}$ (both orientations are possible) such that assertions (a)–(c) hold, however the triangle cannot commute.

Theorem 3.2 has the following high-level interpretation, which corresponds to the behavior observed in non-zigzag persistence theory:

- intervals $[k,k]$ in $\mathsf{Pers}(\mathbb{V})$ disappear in $\mathsf{Pers}(\mathbb{V}^*)$,
- intervals $[i,k]$ with $i<k$ in $\mathsf{Pers}(\mathbb{V})$ become $[i,k-1]$ in $\mathsf{Pers}(\mathbb{V}^*)$,
- intervals $[k,j]$ with $j>k$ in $\mathsf{Pers}(\mathbb{V})$ become $[k+1,j]$ in $\mathsf{Pers}(\mathbb{V}^*)$,
- all other intervals in $\mathsf{Pers}(\mathbb{V})$ remain unchanged in $\mathsf{Pers}(\mathbb{V}^*)$.

The rest of Section 3 is devoted to the proofs of Theorems 3.1 and 3.2, which are interleaved to some extent. Here is the outline:

- Section 3.1 focuses on the reversal of injective or surjective maps, and proves Theorem 3.1 in this special case.
- Section 3.2 focuses on the composition of successive maps with same orientation in a zigzag module, and proves Theorem 3.2 in this special case.
- Section 3.3 combines these restricted results together to obtain the more general versions stated above.

3.1 Reversing Injective and Surjective Maps

Lemma 3.3. *Let $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_k \xrightarrow{f} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ be a zigzag module such that the map f is surjective. Then, there exists an injective map $g : V_{k+1} \to V_k$ such that $f \circ g = \mathbb{1}_{V_{k+1}}$ and the zigzag module $\mathbb{V}^* = V_1 \leftrightarrow \cdots \leftrightarrow V_k \xleftarrow{g} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ has the same persistence barcode as $\mathbb{V}$. Conversely, if $\mathbb{V}^*$ is given, with g injective, then there exists a surjective map $f : V_k \to V_{k+1}$ such that $f \circ g = \mathbb{1}_{V_{k+1}}$ and the corresponding module $\mathbb{V}$ has the same persistence barcode as $\mathbb{V}^*$.*

Note that by reading the zigzag modules from right to left instead of from left to right in the statement of the lemma, we obtain the same guarantees for when the surjective map f is oriented backwards and the injective map g is oriented forwards.

The proof of the lemma proceeds in two steps corresponding to Lemmas 3.4 and 3.5 below. Given a surjective map $f : V_k \to V_{k+1}$, we first show that an injective map $g : V_{k+1} \to V_k$ can be built, such that $f \circ g = \mathbb{1}_{V_{k+1}}$ and g^{-1} *pushes* the right filtration of the submodule $V_1 \leftrightarrow \cdots \leftrightarrow V_k$ into V_{k+1} in the same way as f does. Then, we show that these properties imply that the persistence barcode of $\mathbb{V}$ is preserved when f is replaced by g. While step 1 relies only on standard arguments of linear algebra, step 2 requires the use of the Restriction Theorem 2.2 and Localization Theorem 2.3.

Lemma 3.4. *Let $V_1 \leftrightarrow \cdots \leftrightarrow V_k$ and $V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ be two zigzag modules. Given any surjective map $f : V_k \to V_{k+1}$, there exists an injective map $g : V_{k+1} \to V_k$ such that:*

(a) $f \circ g = \mathbb{1}_{V_{k+1}}$,

(b) $\forall i = 0, \cdots, k$, $(\ker f \cap R_i) \oplus (\operatorname{im} g \cap R_i) = R_i$, *where $(R_0, \cdots, R_k)$ is the right filtration of $V_1 \leftrightarrow \cdots \leftrightarrow V_k$.*

Conversely, given any injective map $g : V_{k+1} \to V_k$, there exists a surjective map $f : V_k \to V_{k+1}$ such that these conditions are satisfied.

Proof. Assuming $f : V_k \to V_{k+1}$ is given, choose any complement of $\ker f$ in V_k, and let $\operatorname{im} g$ be that complement. Then, $f|_{\operatorname{im} g}$ is an isomorphism onto V_{k+1}, so let $g = (f|_{\operatorname{im} g})^{-1}$. This map is an injection $V_{k+1} \to V_k$, with $f \circ g = \mathbb{1}_{V_{k+1}}$, thus it satisfies condition (a). In order to satisfy condition (b) as well, the choice of $\operatorname{im} g$ as a complement of $\ker f$ must be made in a way that is *compatible* with the right filtration $(R_0, \cdots, R_k)$. We define $\operatorname{im} g$ as a direct sum $\operatorname{im} g = \bigoplus_{i=0}^{k} J_i$, where the subspaces J_i are defined by induction:

- $J_0 = R_0 = 0$,
- $\forall i \geq 1$, let J_i be any complement of $R_{i-1} + (\ker f \cap R_i)$ in R_i.

An easy induction (left to the reader) shows that the J_i's are in direct sum with one another and with $\ker f$, and that $(\ker f \cap R_i) \oplus (\operatorname{im} g \cap R_i) = (\ker f \cap R_i) \oplus \bigoplus_{j=0}^{i} J_j = R_i$ for all $i = 0, \cdots, k$, thus satisfying condition (b). In particular, when $i = k$ we have $(\ker f \cap R_k) \oplus (\operatorname{im} g \cap R_k) = R_k$, so $\ker f \oplus \operatorname{im} g = V_k$.

Assuming now that $g : V_{k+1} \to V_k$ is given, choose any complement of $\operatorname{im} g$ in V_k, and let $\ker f$ be that complement. Define now $f|_{\operatorname{im} g}$ to be the inverse function of g over $\operatorname{im} g$ (recall that g is an isomorphism onto its image). Then, $f = 0 \oplus f|_{\operatorname{im} g} : \ker f \oplus \operatorname{im} g \to V_{k+1}$ is a surjection $V_k \to V_{k+1}$, with $f \circ g = \mathbb{1}_{V_{k+1}}$, thus it satisfies condition (a). In order to satisfy condition (b) as well, the choice of $\ker f$ as a complement of $\operatorname{im} g$ in V_k must be made in a way that is *compatible* with the right filtration $(R_0, \cdots, R_k)$. We define $\ker f$ as a direct sum $\ker f = \bigoplus_{i=0}^{k} J_i$, where the subspaces J_i are defined by induction:

- $J_0 = R_0 = 0$,
- $\forall i \geq 1$, let J_i be any complement of $R_{i-1} + (\operatorname{im} g \cap R_i)$ in R_i.

An easy induction (left to the reader) shows that the J_i's are in direct sum with one another and with $\operatorname{im} g$, and that $(\ker f \cap R_i) \oplus (\operatorname{im} g \cap R_i) = (\bigoplus_{j=0}^{i} J_j) \oplus (\operatorname{im} g \cap R_i) = R_i$ for all $i = 0, \cdots, k$, thus satisfying condition (b). In particular, when $i = k$ we have $(\ker f \cap R_k) \oplus (\operatorname{im} g \cap R_k) = R_k$, so $\ker f \oplus \operatorname{im} g = V_k$. □

Lemma 3.5. *Let $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_k \xrightarrow{f} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ be a zigzag module, and let $\mathbb{V}^*$ be obtained from $\mathbb{V}$ by replacing the submodule $V_k \xrightarrow{f} V_{k+1}$ by $V_k \xleftarrow{g} V_{k+1}$. Assume that the following conditions are met:*

(a) g is injective and f is surjective, with $f \circ g = \mathbb{1}_{V_{k+1}}$,

(b) for all $i = 0, \cdots, k$, $(\ker f \cap R_i) \oplus (\operatorname{im} g \cap R_i) = R_i$, where $(R_0, \cdots, R_k)$ is the right filtration of $V_1 \leftrightarrow \cdots \leftrightarrow V_k$.

Then, $\mathbb{V}$ and $\mathbb{V}^*$ have identical persistence barcodes.

Proof. An interval $[i, j]$ of $\mathsf{Pers}(\mathbb{V})$ or $\mathsf{Pers}(\mathbb{V}^*)$ can be of five different types:

(i) $i, j < k$,
(ii) $i, j > k + 1$,
(iii) $i \leq k < k + 1 \leq j$,
(iv) $i = k + 1$,
(v) $j = k$.

Intervals of type (i) are easily dealt with by restriction to $[1, k]$. By the Restriction Theorem 2.2, for all intervals $[i, j]$ of $\mathbb{V}[1, k] = V_1 \leftrightarrow \cdots \leftrightarrow V_k = \mathbb{V}^*[1, k]$ with $j \leq k - 1$, we have

$$\operatorname{mult}([i, j], \mathbb{V}) = \operatorname{mult}([i, j], \mathbb{V}[1, k]) = \operatorname{mult}([i, j], \mathbb{V}^*[1, k]) = \operatorname{mult}([i, j], \mathbb{V}^*).$$

The same goes for intervals of type (ii) by restriction to $[k + 1, n]$.

Intervals of type (iii) and (iv) are dealt with by localizing at index $k + 1$. Given the right filtration $\mathcal{R}_{\mathbb{V}[1,k]} = \mathcal{R}_{\mathbb{V}^*[1,k]} = (R_0, \cdots, R_k)$, we have

$$\mathcal{R}_{\mathbb{V}[1,k+1]} = (f(R_0), \cdots, f(R_k), V_{k+1}),$$
$$\mathcal{R}_{\mathbb{V}^*[1,k+1]} = (0, g^{-1}(R_0), \cdots, g^{-1}(R_k)).$$

For all $i = 0, \cdots, k$, the hypotheses (a) and (b) imply $g^{-1}(R_i) = f(R_i)$ because $f(R_i) = f(\ker f \cap R_i) + f(\operatorname{im} g \cap R_i) = f(\operatorname{im} g \cap R_i) = f \circ g(g^{-1}(R_i)) = g^{-1}(R_i)$. Now, given the birth-time index $\mathrm{b}_{\mathbb{V}[1,k]} = \mathrm{b}_{\mathbb{V}^*[1,k]} = (b_1, \cdots, b_k)$, we have

$$\mathrm{b}_{\mathbb{V}[1,k+1]} = (b_1, \cdots, b_k, k + 1),$$
$$\mathrm{b}_{\mathbb{V}^*[1,k+1]} = (k + 1, b_1, \cdots, b_k).$$

Thus, modulo some boundary effect at positions 0 and $k+1$, the right filtration and birth-time index of $\mathbb{V}^*[1,k+1]$ are obtained from the ones of $\mathbb{V}[1,k+1]$ by a right-shift of the elements. In addition, since the submodules $\mathbb{V}[k+1,n]$ and $\mathbb{V}^*[k+1,n]$ are identical, the left filtrations $\mathcal{L}_{\mathbb{V}[k+1,n]}$ and $\mathcal{L}_{\mathbb{V}^*[k+1,n]}$ are the same, and so are the death-time indices $\mathrm{d}_{\mathbb{V}[k+1,n]}$ and $\mathrm{d}_{\mathbb{V}^*[k+1,n]}$.

It follows then easily from the Localization Theorem 2.3 that $\mathrm{mult}([i,j],\mathbb{V}) = \mathrm{mult}([i,j],\mathbb{V}^*)$ for all $1 \leq i \leq k < k+1 \leq j \leq n$. Indeed, letting l, r be such that $\mathrm{b}_{\mathbb{V}[1,k+1]}[l] = i$ and $\mathrm{d}_{\mathbb{V}[k+1,n]}[r] = j$, we have $1 \leq l \leq k$ and $1 \leq r \leq n-k$, and Theorem 2.3 says that

$$
\begin{aligned}
\mathrm{mult}([i,j],\mathbb{V}) &= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}[1,k+1]}[l], \mathrm{d}_{\mathbb{V}[k+1,n]}[r]\right], \mathbb{V}) \\
&= \dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[l] \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[l-1] \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r]) \\
&\quad + \dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[l-1] \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[l] \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1]) \\
&= \dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[l+1] \cap \mathcal{L}_{\mathbb{V}^*[k+1,n]}[r]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[l] \cap \mathcal{L}_{\mathbb{V}^*[k+1,n]}[r]) \\
&\quad + \dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[l] \cap \mathcal{L}_{\mathbb{V}^*[k+1,n]}[r-1]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[l+1] \cap \mathcal{L}_{\mathbb{V}^*[k+1,n]}[r-1]) \\
&= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}^*[1,k+1]}[l+1], \mathrm{d}_{\mathbb{V}^*[k+1,n]}[r]\right], \mathbb{V}^*) \\
&= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}[1,k+1]}[l], \mathrm{d}_{\mathbb{V}[k+1,n]}[r]\right], \mathbb{V}^*) \\
&= \mathrm{mult}([i,j],\mathbb{V}^*).
\end{aligned}
$$

Thus, intervals of type (iii) have the same multiplicity in $\mathbb{V}$ as in $\mathbb{V}^*$.

Intervals of type (iv) correspond to the case where $i = k+1$ in the above analysis, and they must be handled separately due to the boundary effect. Note that $\mathrm{b}_{\mathbb{V}[1,k+1]}[k+1] = k+1 = \mathrm{b}_{\mathbb{V}^*[1,k+1]}[1]$, so for any $k+1 \leq j \leq n$ we let $1 \leq r \leq n-k$ be such that $\mathrm{d}_{\mathbb{V}[k+1,n]}[r] = j$, and by

the Localization Theorem 2.3 we have

$$\begin{aligned}
\mathrm{mult}([k+1,j],\mathbb{V}) &= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}[1,k+1]}[k+1],\mathrm{d}_{\mathbb{V}[k+1,n]}[r]\right],\mathbb{V}) \\
&= \dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[k+1]\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r]) \\
&\quad -\dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[k]\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r]) \\
&\quad +\dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[k]\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r-1]) \\
&\quad -\dim(\mathcal{R}_{\mathbb{V}[1,k+1]}[k+1]\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r-1]) \\
&= \dim(V_{k+1}\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r]) \\
&\quad -\dim(V_{k+1}\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r]) \\
&\quad +\dim(V_{k+1}\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r-1]) \\
&\quad -\dim(V_{k+1}\cap\mathcal{L}_{\mathbb{V}[k+1,n]}[r-1]) \\
&= 0 \\
&= \dim(0\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r]) \\
&\quad -\dim(0\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r]) \\
&\quad +\dim(0\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r-1]) \\
&\quad -\dim(0\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r-1]) \\
&= \dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[1]\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r]) \\
&\quad -\dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[0]\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r]) \\
&\quad +\dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[0]\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r-1]) \\
&\quad -\dim(\mathcal{R}_{\mathbb{V}^*[1,k+1]}[1]\cap\mathcal{L}_{\mathbb{V}^*[k+1,n]}[r-1]) \\
&= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}^*[1,k+1]}[1],\mathrm{d}_{\mathbb{V}^*[k+1,n]}[r]\right],\mathbb{V}^*) \\
&= \mathrm{mult}([k+1,j],\mathbb{V}^*).
\end{aligned}$$

Thus, intervals of type (iv) have the same multiplicity in $\mathbb{V}$ as in $\mathbb{V}^*$.

Finally, the case of intervals of type (v) is handled by localizing at index k. The submodules $\mathbb{V}[1,k]$ and $\mathbb{V}^*[1,k]$ are identical and so are the corresponding right filtrations and birth-time indices. It is easy to check that the left filtrations of $\mathbb{V}[k,n]$ and $\mathbb{V}^*[k,n]$ are such that

$$\begin{aligned}
\mathcal{L}_{\mathbb{V}[k,n]}[0] &= 0, \\
\mathcal{L}_{\mathbb{V}[k,n]}[1] &= \ker f, \\
\mathcal{L}_{\mathbb{V}^*[k,n]}[n+1-k] &= V_k, \text{ and} \\
\mathcal{L}_{\mathbb{V}^*[k,n]}[n-k] &= \mathrm{im}\, g.
\end{aligned}$$

Moreover,

$$\mathrm{d}_{\mathbb{V}[k,n]}[1] = \mathrm{d}_{\mathbb{V}^*[k,n]}[n+1-k] = k.$$

Let l be the index of i in $\mathrm{b}_{\mathbb{V}[1,k]}$, that is, $\mathrm{b}_{\mathbb{V}[1,k]}[l] = \mathrm{b}_{\mathbb{V}^*[1,k]}[l] = i$. Recall that hypothesis (b) says that for all $m = 0 \ldots k$

$$\mathcal{R}_{\mathbb{V}[1,k]}[m] = (\mathcal{R}_{\mathbb{V}[1,k]}[m]\cap\ker f)\oplus(\mathcal{R}_{\mathbb{V}[1,k]}[m]\cap\mathrm{im}\, g).$$

We can thus compute the multiplicity of intervals of type (v) as follows:

$$\begin{aligned}
\mathrm{mult}([i,k],\mathbb{V}) &= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}[1,k]}[l], \mathrm{d}_{\mathbb{V}[k,n]}[1]\right], \mathbb{V}) \\
&= \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l] \cap \mathcal{L}_{\mathbb{V}[k,n]}[1]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l] \cap \mathcal{L}_{\mathbb{V}[k,n]}[0]) \\
&\quad + \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1] \cap \mathcal{L}_{\mathbb{V}[k,n]}[0]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1] \cap \mathcal{L}_{\mathbb{V}[k,n]}[1]) \\
&= \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l] \cap \ker f) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l] \cap 0) \\
&\quad + \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1] \cap 0) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1] \cap \ker f) \\
&= \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l] \cap \ker f) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1] \cap \ker f) \\
&= \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l] \cap \operatorname{im} g) \\
&\quad + \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1] \cap \operatorname{im} g) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}[1,k]}[l-1]) \\
&= \dim(\mathcal{R}_{\mathbb{V}^*[1,k]}[l] \cap \mathcal{L}_{\mathbb{V}^*[k,n]}[n+1-k]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}^*[1,k]}[l] \cap \mathcal{L}_{\mathbb{V}^*[k,n]}[n-k]) \\
&\quad + \dim(\mathcal{R}_{\mathbb{V}^*[1,k]}[l-1] \cap \mathcal{L}_{\mathbb{V}^*[k,n]}[n-k]) \\
&\quad - \dim(\mathcal{R}_{\mathbb{V}^*[1,k]}[l-1] \cap \mathcal{L}_{\mathbb{V}^*[k,n]}[n+1-k]) \\
&= \mathrm{mult}(\left[\mathrm{b}_{\mathbb{V}^*[1,k]}[l], \mathrm{d}_{\mathbb{V}^*[k,n]}[n+1-k]\right], \mathbb{V}^*) \\
&= \mathrm{mult}([i,k],\mathbb{V}^*).
\end{aligned}$$

Thus, intervals of type (v) have the same multiplicity in $\mathbb{V}$ as in $\mathbb{V}^*$. □

3.2 Composing Maps

Lemma 3.6. *Let $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_{k-1} \xrightarrow{f} V_k \xrightarrow{g} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ be a zigzag module, and let $\mathbb{V}^*$ be obtained from $\mathbb{V}$ by replacing the submodule $V_{k-1} \xrightarrow{f} V_k \xrightarrow{g} V_{k+1}$ by $V_{k-1} \xrightarrow{g\circ f} V_{k+1}$. Then, the persistence barcode of $\mathbb{V}^*$ derives from the one of $\mathbb{V}$ through assertions (a)–(c) of Theorem 3.2.*

By reading the zigzag modules from right to left instead of from left to right in the statement of the lemma, we obtain the same guarantees for when the maps f and g are oriented backwards.

Proof. Localize at index $k+1$. Given the right filtration $\mathcal{R}_{\mathbb{V}[1,k-1]} = \mathcal{R}_{\mathbb{V}^*[1,k-1]} = (R_0, \cdots, R_{k-1})$, we have

$$\begin{aligned}
\mathcal{R}_{\mathbb{V}[1,k+1]} &= (g\circ f(R_0), \cdots, g\circ f(R_{k-1}), g(V_k), V_{k+1}), \\
\mathcal{R}_{\mathbb{V}^*[1,k+1]} &= (g\circ f(R_0), \cdots, g\circ f(R_{k-1}), V_{k+1}).
\end{aligned}$$

Now, given the birth-time index $\mathrm{b}_{\mathbb{V}[1,k-1]} = \mathrm{b}_{\mathbb{V}^*[1,k-1]} = (b_1, \cdots, b_{k-1})$, we have

$$\begin{aligned}\mathrm{b}_{\mathbb{V}[1,k+1]} &= (b_1, \cdots, b_{k-1}, k, k+1),\\ \mathrm{b}_{\mathbb{V}^*[1,k+1]} &= (b_1, \cdots, b_{k-1}, k+1).\end{aligned}$$

Since the submodules $\mathbb{V}[k+1,n]$ and $\mathbb{V}^*[k+1,n]$ are identical, the left filtrations $\mathcal{L}_{\mathbb{V}[k+1,n]}$ and $\mathcal{L}_{\mathbb{V}^*[k+1,n]}$ are the same, and so are the death-time indices $\mathrm{d}_{\mathbb{V}[k+1,n]}$ and $\mathrm{d}_{\mathbb{V}^*[k+1,n]}$. Therefore, for any $k+1 \leq j \leq n$ we let $1 \leq r \leq n-k$ be such that $\mathrm{d}_{\mathbb{V}[k+1,n]}[r] = j$, and by the Localization Theorem 2.3 we have

$$\begin{aligned}\mathrm{mult}([k+1,j],\mathbb{V}) + \mathrm{mult}([k,j],\mathbb{V}) &= \dim(V_{k+1} \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r])\\ &\quad - \dim(g(V_k) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r])\\ &\quad + \dim(g(V_k) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1])\\ &\quad - \dim(V_{k+1} \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1])\\ &\quad + \dim(g(V_k) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r])\\ &\quad - \dim(g \circ f(V_{k-1}) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r])\\ &\quad + \dim(g \circ f(V_{k-1}) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1])\\ &\quad - \dim(g(V_k) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1])\\ &= \dim(V_{k+1} \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r])\\ &\quad - \dim(g \circ f(V_{k-1}) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r])\\ &\quad + \dim(g \circ f(V_{k-1}) \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1])\\ &\quad - \dim(V_{k+1} \cap \mathcal{L}_{\mathbb{V}[k+1,n]}[r-1])\\ &= \mathrm{mult}([k+1,j],\mathbb{V}^*),\end{aligned}$$

which proves assertion (b). The proof of (a) is symmetric, based on localization at index $k-1$.

To prove (c), we need to consider all other intervals $[i,j]$ with $i,j \in \{1,\cdots,n\} \setminus \{k\}$, which can be of the following three types:
(c_1) $1 \leq i \leq k-1 < k+1 \leq j \leq n$,
(c_2) $1 \leq i \leq j \leq k-2$,
(c_3) $k+2 \leq i \leq j \leq n$.
Intervals of type (c_1) are handled by localizing once again at index $k+1$. Referring to the expressions of $\mathcal{R}_{\mathbb{V}[1,k+1]}$, $\mathcal{R}_{\mathbb{V}^*[1,k+1]}$, $\mathrm{b}_{\mathbb{V}[1,k+1]}$, $\mathrm{b}_{\mathbb{V}^*[1,k+1]}$, $\mathcal{L}_{\mathbb{V}[k+1,n]} = \mathcal{L}_{\mathbb{V}^*[k+1,n]}$, and $\mathrm{d}_{\mathbb{V}[k+1,n]} = \mathrm{d}_{\mathbb{V}^*[k+1,n]}$ given above, we have $i = b_l$ for some $l \leq k-1$ and therefore the expressions of $\mathrm{mult}([i,j],\mathbb{V})$ and $\mathrm{mult}([i,j],\mathbb{V}^*)$ given by the Localization Theorem 2.3 are identical.

Intervals of types (c_2) and (c_3) are handled using restriction. Consider the submodule $\mathbb{V}[1,k-1] = V_1 \leftrightarrow \cdots \leftrightarrow V_{k-1} = \mathbb{V}^*[1,k-1]$. By the Restriction Theorem 2.2, for all $1 \leq i \leq j \leq k-2$ we have

$$\mathrm{mult}([i,j],\mathbb{V}) = \mathrm{mult}([i,j],\mathbb{V}[1,k-1]) = \mathrm{mult}([i,j],\mathbb{V}^*[1,k-1]) = \mathrm{mult}([i,j],\mathbb{V}^*).$$

Similarly, restriction to the submodule $\mathbb{V}[k+1,n] = V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n = \mathbb{V}^*[k+1,n]$ gives

$$\mathrm{mult}([i,j],\mathbb{V}) = \mathrm{mult}([i,j],\mathbb{V}[k+1,n]) = \mathrm{mult}([i,j],\mathbb{V}^*[k+1,n]) = \mathrm{mult}([i,j],\mathbb{V}^*)$$

for all $k+2 \leq i \leq j \leq n$. This concludes the proofs of assertion (c) and of the lemma. □

3.3 Proofs of the Arrow Reversal and Space Removal Theorems

We will now combine Lemmas 3.3 and 3.6 together to prove the more general versions stated in Theorems 3.1 and 3.2.

Proof of Theorem 3.1. Assume without loss of generality that $f : V_k \to V_{k+1}$ (the case where f is oriented backwards is symmetric and reduces to this one by reading the zigzag module from right to left). Consider the module $\mathbb{W} = V_1 \leftrightarrow \cdots \leftrightarrow V_k \xrightarrow{f_1} V_{k+1/2} \xrightarrow{f_2} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$, where $V_{k+1/2} = \operatorname{im} f$, $f_1 = f$ and f_2 is the canonical inclusion $\operatorname{im} f \hookrightarrow V_{k+1}$. Since $f = f_2 \circ f_1$, the persistence barcodes of $\mathbb{V}$ and $\mathbb{W}$ are related to each other through the matching given by Lemma 3.6.

Then, we apply Lemma 3.3 twice, once for the surjective map f_1 and again for the injective map f_2. We thus obtain a module $\mathbb{W}^* = V_1 \leftrightarrow \cdots \leftrightarrow V_k \xleftarrow{g_1} V_{k+1/2} \xleftarrow{g_2} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$ that has the same persistence barcode as $\mathbb{W}$.

Finally, consider the module $\mathbb{V}^* = V_1 \leftrightarrow \cdots \leftrightarrow V_k \xleftarrow{g} V_{k+1} \leftrightarrow \cdots \leftrightarrow V_n$, where $g = g_1 \circ g_2$. Its persistence barcode is related to the one of $\mathbb{W}^*$ through the matching given by Lemma 3.6.

Together, these relations imply that for any $1 \leq i \leq j \leq k-1$ or $k+2 \leq i \leq j \leq n$ or $1 \leq i \leq k < k+1 \leq j \leq n$,

$$\operatorname{mult}([i,j],\mathbb{V}) = \operatorname{mult}([i,j],\mathbb{W}) = \operatorname{mult}([i,j],\mathbb{W}^*) = \operatorname{mult}([i,j],\mathbb{V}^*).$$

Moreover, for any $i \in [1,k]$,

$$\begin{aligned}\operatorname{mult}([i,k],\mathbb{V}) &= \operatorname{mult}([i,k],\mathbb{W}) + \operatorname{mult}([i,k+1/2],\mathbb{W}) \\ &= \operatorname{mult}([i,k],\mathbb{W}^*) + \operatorname{mult}([i,k+1/2],\mathbb{W}^*) \\ &= \operatorname{mult}([i,k],\mathbb{V}^*).\end{aligned}$$

And symmetrically, for any $j \in [k+1,n]$,

$$\begin{aligned}\operatorname{mult}([k+1,j],\mathbb{V}) &= \operatorname{mult}([k+1,j],\mathbb{W}) + \operatorname{mult}([k+1/2,j],\mathbb{W}) \\ &= \operatorname{mult}([k+1,j],\mathbb{W}^*) + \operatorname{mult}([k+1/2,j],\mathbb{W}^*) \\ &= \operatorname{mult}([k+1,j],\mathbb{V}^*).\end{aligned}$$

Thus, $\mathbb{V}$ and $\mathbb{V}^*$ have identical persistence barcodes.

Recall now from Lemma 3.3 that $f_1 \circ g_1 = g_2 \circ f_2 = \mathbb{1}_{\operatorname{im} f}$. Therefore,

$$f \circ g|_{\operatorname{im} f} = f_2 \circ f_1 \circ g_1 \circ g_2|_{\operatorname{im} f} = f_2 \circ \mathbb{1}_{\operatorname{im} f} \circ g_2|_{\operatorname{im} f} = f_2 \circ g_2|_{\operatorname{im} f} = f_2 \circ g_2|_{\operatorname{im} f_2} = \mathbb{1}_{\operatorname{im} f_2} = \mathbb{1}_{\operatorname{im} f},$$
$$g \circ f|_{\operatorname{im} g} = g_1 \circ g_2 \circ f_2 \circ f_1|_{\operatorname{im} g} = g_1 \circ \mathbb{1}_{\operatorname{im} f} \circ f_1|_{\operatorname{im} g} = g_1 \circ f_1|_{\operatorname{im} g} = g_1 \circ f_1|_{\operatorname{im} g_1} = \mathbb{1}_{\operatorname{im} g_1} = \mathbb{1}_{\operatorname{im} g}.$$

□

Proof of Theorem 3.2. Cases $V_{k-1} \xleftarrow{f} V_k \xleftarrow{g} V_{k+1}$ and $V_{k-1} \xrightarrow{f} V_k \xrightarrow{g} V_{k+1}$ are already handled by Lemma 3.6, therefore we only need to analyze the other two cases. Since their proofs are essentially the same, we will focus on case $V_{k-1} \xrightarrow{f} V_k \xleftarrow{g} V_{k+1}$ and leave case $V_{k-1} \xleftarrow{f} V_k \xrightarrow{g} V_{k+1}$ as an exercise to the reader.

First, we apply Theorem 3.1 to reverse the map g in $\mathbb{V}$, which gives a map $g' : V_k \to V_{k+1}$ such that $g \circ g'|_{\operatorname{im} g} = \mathbb{1}_{\operatorname{im} g}$ and the persistence barcode of $\mathbb{V}$ is preserved.

Next, we remove the space V_k by composing the maps f and g', which gives the zigzag module $\mathbb{V}^*$ with $h = g' \circ f : V_{k-1} \to V_{k+1}$. By Lemma 3.6, the persistence barcodes of $\mathbb{V}$ and $\mathbb{V}^*$ satisfy assertions (a)–(c) of Theorem 3.2.

In addition, we have

$$\forall x \in V_{k-1},\ g \circ h(x) = g \circ g' \circ f(x) = g \circ g'(f(x)),$$

which is equal to $f(x)$ if we further assume that $\operatorname{im} f \subseteq \operatorname{im} g$ (recall that $g \circ g'|_{\operatorname{im} g} = \mathbb{1}_{\operatorname{im} g}$).

Symmetrically, starting from $\mathbb{V}$ we can reverse f to obtain a map $f' : V_k \to V_{k-1}$, and then compose this map with g. This gives another zigzag module $\mathbb{V}^*$ with $h = f' \circ g : V_{k+1} \to V_{k-1}$, whose persistence barcode is related to the one of $\mathbb{V}$ through assertions (a)–(c). Furthermore, a short calculation as above gives $g = f \circ h$ if we further assume that $\operatorname{im} g \subseteq \operatorname{im} f$. □

4 Comparing Zigzag Modules

4.1 Factorization

Consider the following diagram of vector spaces and linear maps:

$$\begin{array}{ccccccccc} W_1 & \leftrightarrow & W_2 & \leftrightarrow & \cdots & \leftrightarrow & W_{n-1} & \leftrightarrow & W_n \\ \uparrow & & \uparrow & & & & \uparrow & & \uparrow \\ V_1 & \leftrightarrow & V_2 & \leftrightarrow & \cdots & \leftrightarrow & V_{n-1} & \leftrightarrow & V_n \\ \uparrow & & \uparrow & & & & \uparrow & & \uparrow \\ U_1 & \leftrightarrow & U_2 & \leftrightarrow & \cdots & \leftrightarrow & U_{n-1} & \leftrightarrow & U_n \end{array} \tag{2}$$

Let $\mathbb{U} = U_1 \leftrightarrow \cdots \leftrightarrow U_n$, $\mathbb{V} = V_1 \leftrightarrow \cdots \leftrightarrow V_n$, and $\mathbb{W} = W_1 \leftrightarrow \cdots \leftrightarrow W_n$ be the zigzag modules represented on the three lines of the diagram. The following result is an easy consequence of the Interval Decomposition Theorem 2.1, or rather the uniqueness part of it.

Theorem 4.1 (Factorization)**.** *Assume that the following conditions are met:*

- *the zigzag modules $\mathbb{U}$, $\mathbb{V}$ and $\mathbb{W}$ are of the same type,*
- *all the quadrangles in the diagram (2) commute,*
- *for all $k \in [1, n]$, the maps $U_k \to V_k \to W_k$ compose into an isomorphism $U_k \to W_k$.*

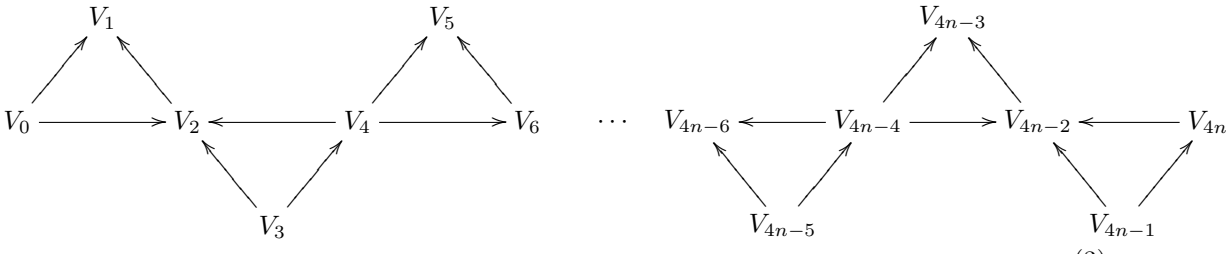
Then, $\mathsf{Pers}(\mathbb{U}) \subseteq \mathsf{Pers}(\mathbb{V})$.

The name *Factorization Theorem* is coined afer the fact that (2) factors the morphism between zigzag modules $\mathbb{U} \to \mathbb{W}$ through $\mathbb{V}$.

Proof. Let $\mathbb{J}$ be the submodule of $\mathbb{V}$ formed by the images J_k of the maps $U_k \to V_k$. This is a well-defined submodule thanks to the commutativity of the bottom quadrangles in the diagram (2). Similarly, let $\mathbb{K}$ be the submodule of $\mathbb{V}$ formed by the kernels K_k of the maps $V_k \to W_k$, which is well-defined thanks to the commutativity of the upper quadrangles in (2). Since the maps $U_k \to V_k \to W_k$ compose into an isomorphism, we have $V_k = J_k \oplus K_k$ for all $k \in [1, n]$. Thus, $\mathbb{V} = \mathbb{J} \oplus \mathbb{K}$, from which follows that $\mathsf{Pers}(\mathbb{J}) \subseteq \mathsf{Pers}(\mathbb{V})$ by the Interval Decomposition Theorem 2.1 (more precisely the uniqueness part of it). Now, each map $U_k \to V_k$ being an isomorphism onto J_k, we have $\mathsf{Pers}(\mathbb{J}) = \mathsf{Pers}(\mathbb{U})$. □

4.2 Interleaving

Consider now the following diagram of vector spaces and linear maps:



(3)

Let $\mathbb{V} = V_0 \to V_1 \leftarrow V_2 \leftarrow V_3 \to V_4 \to V_5 \leftarrow V_6 \leftarrow \cdots \leftarrow V_{4n-6} \leftarrow V_{4n-5} \to V_{4n-4} \to V_{4n-3} \leftarrow V_{4n-2} \leftarrow V_{4n-1} \to V_{4n}$ be the oscillating zigzag, and let $\mathbb{W} = V_0 \to V_2 \leftarrow V_4 \to V_6 \leftarrow \cdots \to V_{4n-6} \leftarrow V_{4n-4} \to V_{4n-2} \leftarrow V_{4n}$ be the axis zigzag in (3). The following result provides guarantees for this special kind of interleaving between zigzag modules.

Theorem 4.2 (Interleaving)**.** *Assume that the following conditions are met for all $k \in [0, n-1]$:*

- *all the triangles in the diagram (3) commute,*
- *all maps* $V_{4k} \to V_{4k+2}$ *and* $V_{4k+2} \leftarrow V_{4k+4}$ *are isomorphisms,*
- *all maps* $V_{4k+1} \leftarrow V_{4k+2}$ *are injective,*
- *all maps* $V_{4k+3} \to V_{4k+4}$ *are surjective.*

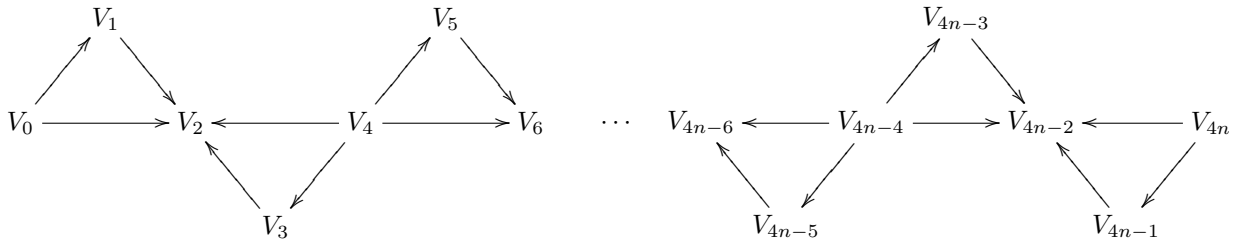
Then,

(a) every interval in $\mathsf{Pers}(\mathbb{W})$ *is of type* $[0, 4n]$*,*
(b) every interval in $\mathsf{Pers}(\mathbb{V})$ *is of type* $[0, 4n]$ *or* $[2k+1, 2k+1]$ *for some* $k \in [0, 2n-1]$*,*
(c) $\mathrm{mult}([0, 4n], \mathbb{V}) = \mathrm{mult}([0, 4n], \mathbb{W})$*.*

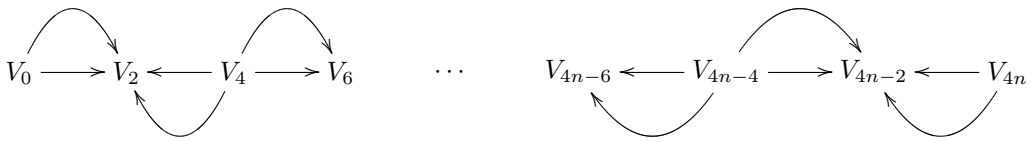
Rephrased in terms of information theory, the conclusion of the theorem says that the persistence barcodes of $\mathbb{V}$ and $\mathbb{W}$ contain the same signal, formed by a given number of copies of the full-length interval $[0, 4n]$, however, while $\mathsf{Pers}(\mathbb{W})$ contains only the signal, $\mathsf{Pers}(\mathbb{V})$ may contain additional ephemeral noise formed by intervals of length zero starting and ending at odd indices in the range $[0, 4n]$.

Proof. We will use arrow reversal and composition to turn $\mathbb{V}$ into $\mathbb{W}$, while tracking the changes in the persistence barcode.

We first apply Theorem 3.1 on every injective map $V_{4k+1} \leftarrow V_{4k+2}$ and on every surjective map $V_{4k+3} \to V_{4k+4}$, to get a new zigzag $\mathbb{V}^* = V_0 \to V_1 \to V_2 \leftarrow V_3 \leftarrow V_4 \to V_5 \to V_6 \leftarrow \cdots \to V_{4n-6} \leftarrow V_{4n-5} \leftarrow V_{4n-4} \to V_{4n-3} \to V_{4n-2} \leftarrow V_{4n-1} \leftarrow V_{4n}$ that has the same persistence barcode as $\mathbb{V}$. Moreover, the map $f : V_{4k+1} \to V_{4k+2}$ provided by Theorem 3.1 when reversing $g : V_{4k+1} \leftarrow V_{4k+2}$ satisfies $f \circ g = \mathbb{1}_{V_{4k+2}}$, while the map $g' : V_{4k+3} \leftarrow V_{4k+4}$ provided when reversing $f' : V_{4k+3} \to V_{4k+4}$ satisfies $f' \circ g' = \mathbb{1}_{V_{4k+4}}$. It follows that every triangle commutes in the resulting diagram involving $\mathbb{V}^*$ and $\mathbb{W}$:



Now, for k ranging from 0 to $n-1$, we compose $V_{4k} \to V_{4k+1} \to V_{4k+2}$ into a single map $V_{4k} \to V_{4k+2}$, and similarly we compose $V_{4k+2} \leftarrow V_{4k+3} \leftarrow V_{4k+4}$ into a single map $V_{4k+2} \leftarrow V_{4k+4}$. Since composition preserves commutativity of the subdiagrams, the following diagram involving $\mathbb{W}$ (straight path) and the newly obtained zigzag $\mathbb{W}^*$ (curved path) commutes:



Hence, the zigzags $\mathbb{W}$ and $\mathbb{W}^*$ are identical. It suffices then to prove assertions (a)–(c) with $\mathbb{V}$ replaced by $\mathbb{V}^*$ and $\mathbb{W}$ by $\mathbb{W}^*$, because $\mathsf{Pers}(\mathbb{V}) = \mathsf{Pers}(\mathbb{V}^*)$ and $\mathsf{Pers}(\mathbb{W}) = \mathsf{Pers}(\mathbb{W}^*)$. Recall that $\mathbb{W}^*$ is obtained from $\mathbb{V}^*$ by removing the spaces V_{2k+1} for all $k \in [0, 2n-1]$, so the Space Removal Theorem 3.2 relates their persistence barcodes.

• Assertion (a) follows from the fact that all maps in the zigzag module $\mathbb{W}^*$ are isomorphisms.

• Assertion (b) follows from (a) and Theorem 3.2. Indeed, for any k, l such that $0 < k \leq l < 2n$, we have $\mathrm{mult}([2k, 2l], \mathbb{W}^*) = \mathrm{mult}([2k, 2l], \mathbb{V}^*) + \mathrm{mult}([2k-1, 2l], \mathbb{V}^*) + \mathrm{mult}([2k-1, 2l+$

$1], \mathbb{V}^*) + \mathrm{mult}([2k, 2l+1], \mathbb{V}^*)$ by Theorem 3.2, and this sum is equal to zero by assertion (a). Therefore, $\mathrm{mult}([2k, 2l], \mathbb{V}^*) = \mathrm{mult}([2k-1, 2l], \mathbb{V}^*) = \mathrm{mult}([2k-1, 2l+1], \mathbb{V}^*) = \mathrm{mult}([2k, 2l+1], \mathbb{V}^*) = 0$. Similarly, $\mathrm{mult}([2k, 4n], \mathbb{V}^*) = \mathrm{mult}([2k-1, 4n], \mathbb{V}^*) = 0$ for any $0 < k \leq 2n$, and $\mathrm{mult}([0, 2l], \mathbb{V}^*) = \mathrm{mult}([0, 2l+1], \mathbb{V}^*) = 0$ for any $0 \leq l < 2n$. Hence, for any $0 \leq i \leq j \leq 4n$, we have $\mathrm{mult}([i, j], \mathbb{V}^*) = 0$ unless $i = 0$ and $j = 4n$, or $i = j = 2k+1$ for some $k \in [0, 2n-1]$.
• Assertion (c) follows directly from Theorem 3.2. □

5 Rips Zigzags

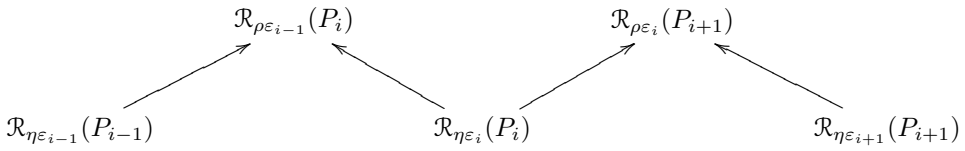
Let P be a finite point cloud in some metric space, and suppose that the matrix of pairwise distances between the points of P is known. Given an ordering $(p_1, \cdots, p_n)$ on the points of P, let $P_i := \{p_1, \ldots, p_i\}$ denote the ith prefix, and define the *ith geometric scale* for $i = 1 \ldots n$ as

$$\varepsilon_i \stackrel{\text{def}}{=} \mathrm{d}_H(P_i, P).$$

Since the prefix P_i grows as i increases, we have

$$\varepsilon_1 \geq \varepsilon_2 \geq \cdots \geq \varepsilon_n = 0.$$

Given a choice of multipliers $\rho \geq \eta$, Chazal and Oudot [9] proposed to do homological inference from P using the sequence of short filtrations $\mathcal{R}_{\eta\varepsilon_i}(P_i) \hookrightarrow \mathcal{R}_{\rho\varepsilon_i}(P_i)$. The invention of zigzag persistence makes it possible to replace this sequence of unrelated short filtrations by a single long zigzag filtration, a representative portion of which is depicted below.



(4)

The zigzag module induced at the homology level by this zigzag filtration is referred to as the *oscillating Rips zigzag* (*oR-ZZ* for short) hereafter. Note that from a computational point of view, the smaller ρ the smaller the maximum complex size in the zigzag. In addition, the closer η to ρ the fewer simplex additions and deletions during the zigzag calculation. Therefore, as a rule of thumb, one should try to make ρ as small as possible while η as close to ρ as possible.

Before proceeding with the analysis of the oscillating Rips zigzag in Section 5.2 and of its variants in the subsequent sections, we first make a short detour and study another zigzag over the sequence of vertex sets $P_1, \cdots, P_n$ that will play a central role in our analysis.

5.1 The image Čech zigzag

Canonical inclusions between Čech complexes give the following pair of horizontal zigzags connected by vertical arrows, where each zigzag alternately adds one point to the vertex set and reduces the geometric scale.

$$\begin{array}{ccccccccc}
\cdots & \leftarrow & H_*(\mathcal{C}_{\rho\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}_{\rho\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}_{\rho\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{C}_{\eta\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}_{\eta\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}_{\eta\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots
\end{array}$$

This commutative diagram induces the following zigzag of images, referred to as the *image Čech zigzag* hereafter.

$$\cdots \leftarrow H_*(\mathcal{C}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_i)) \rightarrow H_*(\mathcal{C}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_{i+1})) \leftarrow H_*(\mathcal{C}^{\rho\varepsilon_{i+1}}_{\eta\varepsilon_{i+1}}(P_{i+1})) \rightarrow \cdots \tag{5}$$

Theorem 5.1. *Given a choice of multipliers ρ, η such that $\rho > 5$ and $3 < \eta < \rho - 2$, suppose $P \subset \mathbb{R}^d$ and there is some compact set $X \subset \mathbb{R}^d$ such that $\mathrm{d}_H(P, X) < \varepsilon$ with $\varepsilon < \min\left\{\frac{\eta-3}{6\eta}, \frac{\eta-3}{3\rho+\eta}, \frac{\rho-\eta-2}{6(\rho-\eta)}, \frac{\rho-\eta-2}{3\rho-\eta}\right\}\mathrm{wfs}(X)$. Then, for any $k < l$ such that*

$$\max\left\{\frac{3\varepsilon}{\eta-3}, \frac{2\varepsilon}{\rho-\eta-2}\right\} \leq \varepsilon_k, \varepsilon_l < \min\left\{\frac{1}{6}\mathrm{wfs}(X) - \varepsilon, \frac{1}{\rho+1}\left(\mathrm{wfs}(X) - \varepsilon\right)\right\},$$

the image Čech zigzag restricted to $H_*(\mathcal{C}^{\rho\varepsilon_k}_{\eta\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{C}^{\rho\varepsilon_l}_{\eta\varepsilon_l}(P_l))$ *contains only isomorphisms, and its spaces are isomorphic to* $H_*(X^\lambda)$ *for any* $\lambda \in (0, \mathrm{wfs}(X))$*. Therefore, its persistence barcode is made only of full-length intervals, whose number equals the dimension of* $H_*(X^\lambda)$*.*

Proof. By the triangle inequality, for any $i \in [1,n]$ we have $\mathrm{d}_H(P_i, X) \leq \mathrm{d}_H(P_i, P) + \mathrm{d}_H(P, X) < \varepsilon_i + \varepsilon$. Since the geometric scale ε_i decreases with i, we have $\varepsilon_k \geq \varepsilon_i \geq \varepsilon_l$ for all $i \in [k,l]$, and therefore our hypotheses imply that $\varepsilon_i + \varepsilon < \frac{1}{6}\mathrm{wfs}(X)$, that $\eta\varepsilon_i$ and $\rho\varepsilon_i$ belong to the interval $[3(\varepsilon_i+\varepsilon), \mathrm{wfs}(X) - (\varepsilon_i+\varepsilon)]$, and that $\rho\varepsilon_i - \eta\varepsilon_i \geq 2(\varepsilon_i+\varepsilon)$. Thus, the hypotheses of Theorem 2.8 (ii) are satisfied within the range $[k,l]$, and so the result follows from that theorem. □

5.2 Analysis of the oscillating Rips zigzag

The following result gives conditions on η, ρ for the persistence barcode of the oR-ZZ to exhibit the homology of the shape underlying the input point cloud P when the latter lies in Euclidean space $\mathbb{R}^d$ ($d \geq 1$). The proof relies on Theorems 2.8 (i) and 5.1, as well as on the fact that Čech and Rips complexes are interleaved as follows in $\mathbb{R}^d$ — see [11] for a proof[3].

$$\forall \alpha \geq 0,\ \mathcal{C}_{\frac{\alpha}{2}}(P) \subseteq \mathcal{R}_\alpha(P) \subseteq \mathcal{C}_{\vartheta_d \alpha}(P),\ \text{where } \vartheta_d = \sqrt{\frac{d}{2(d+1)}}. \tag{6}$$

Theorem 5.2. *Let* ρ *and* η *be multipliers such that* $\rho > 10$ *and* $\frac{3}{\vartheta_d} < \eta < \frac{\rho-4}{2\vartheta_d}$*. Let* $X \subset \mathbb{R}^d$ *be a compact set and let* $P \subset \mathbb{R}^d$ *be such that* $\mathrm{d}_H(P,X) < \varepsilon$ *with*

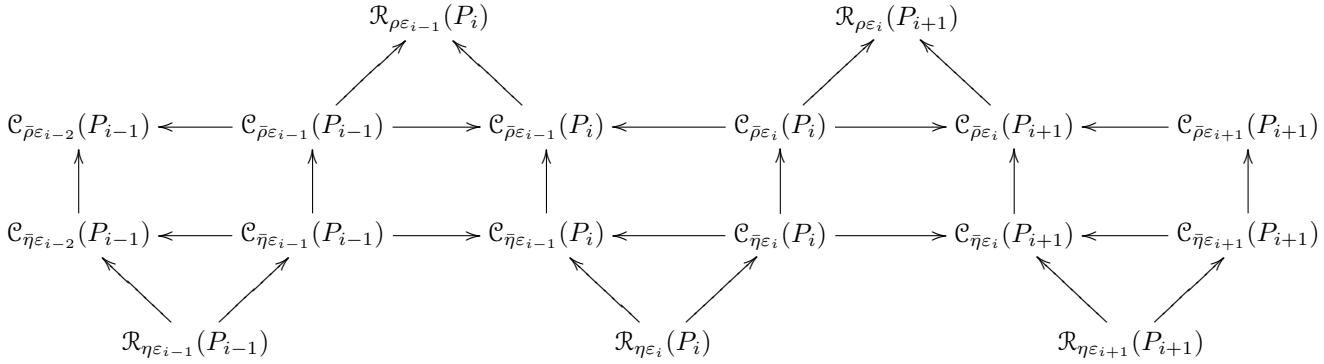
$$\varepsilon < \min\left\{ \frac{\vartheta_d\eta - 3}{6\vartheta_d\eta},\ \frac{\eta - 3/\vartheta_d}{3\rho+\eta},\ \frac{\rho - 2\vartheta_d\eta - 4}{6(\rho - 2\vartheta_d\eta)},\ \frac{\rho - 2\vartheta_d\eta - 4}{(4\vartheta_d+1)\rho - 2\vartheta_d\eta} \right\} \mathrm{wfs}(X).$$

Then, for any $k < l$ *such that*

$$\max\left\{ \frac{3\varepsilon}{\vartheta_d\eta - 3},\ \frac{4\varepsilon}{\rho - 2\vartheta_d\eta - 4} \right\} \leq \varepsilon_k, \varepsilon_l < \min\left\{ \frac{1}{6}\mathrm{wfs}(X) - \varepsilon,\ \frac{1}{\vartheta_d\rho + 1}\left(\mathrm{wfs}(X) - \varepsilon\right) \right\},$$

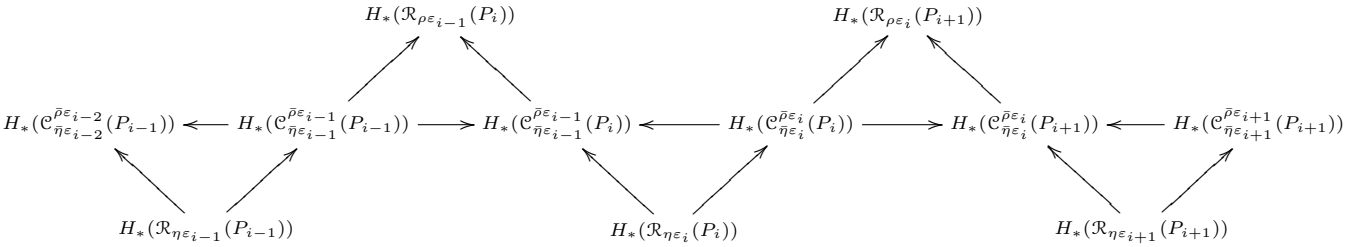
the oR-ZZ restricted to $H_*(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}_{\eta\varepsilon_l}(P_l))$ *has a persistence barcode made only of full-length intervals and ephemeral (length zero) intervals, the number of full-length intervals being equal to the dimension of* $H_*(X^\lambda)$ *for any* $\lambda \in (0, \mathrm{wfs}(X))$*.*

Proof. Let $\bar\rho = \frac{\rho}{2}$ and $\bar\eta = \vartheta_d\eta$. Our hypotheses imply $\frac{\rho}{2} \geq \vartheta_d\eta$, so we can use (6) to factor the inclusion maps in (4) through Čech complexes with multipliers $\bar\eta$ and $\bar\rho$ as follows.



[3]Our definition of the Rips complex differs from the one in [11] by a factor of 2 in the parameter value. This explains the slight discrepancy between our chain of inclusions and the one in [11]. Note that $\frac{1}{2} \leq \vartheta_d < \frac{1}{\sqrt{2}}$.

This commutative diagram induces the following interleaving between the oscillating Rips and image Čech zigzags at the homology level (note that the triangles still commute).



This diagram has the same shape as the one in (3). Let $\mathbb{V}$ be the oscillating zigzag $H_*(\mathcal{C}^{\bar{\rho}\varepsilon_k}_{\bar{\eta}\varepsilon_k}(P_k)) \to H_*(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_k}_{\bar{\eta}\varepsilon_k}(P_{k+1})) \leftarrow H_*(\mathcal{R}_{\eta\varepsilon_{k+1}}(P_{k+1})) \to H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{k+1}}_{\bar{\eta}\varepsilon_{k+1}}(P_{k+1})) \to \cdots \to H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{l-1}}_{\bar{\eta}\varepsilon_{l-1}}(P_{l-1})) \to H_*(\mathcal{R}_{\rho\varepsilon_{l-1}}(P_l)) \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{l-1}}_{\bar{\eta}\varepsilon_{l-1}}(P_l)) \leftarrow H_*(\mathcal{R}_{\eta\varepsilon_l}(P_l)) \to H_*(\mathcal{C}^{\bar{\rho}\varepsilon_l}_{\bar{\eta}\varepsilon_l}(P_l))$, and let $\mathbb{W}$ be the axis zigzag $H_*(\mathcal{C}^{\bar{\rho}\varepsilon_k}_{\bar{\eta}\varepsilon_k}(P_k)) \to H_*(\mathcal{C}^{\bar{\rho}\varepsilon_k}_{\bar{\eta}\varepsilon_k}(P_{k+1})) \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{k+1}}_{\bar{\eta}\varepsilon_{k+1}}(P_{k+1})) \to \cdots \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{l-1}}_{\bar{\eta}\varepsilon_{l-1}}(P_{l-1})) \to H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{l-1}}_{\bar{\eta}\varepsilon_{l-1}}(P_l)) \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_l}_{\bar{\eta}\varepsilon_l}(P_l))$.

Our hypotheses involving $\bar{\rho}$ and $\bar{\eta}$ are clearly stronger than the ones in Theorem 5.1, so we can apply that theorem and deduce that the spaces in $\mathbb{W}$ are isomorphic to $H_*(X^\lambda)$, with all the arrows in $\mathbb{W}$ being isomorphisms. One can also check that the hypotheses of Theorem 2.8 (i) are satisfied for all $i \in [k,l]$ so we can apply that theorem and deduce that the inclusions $\mathcal{C}_{\frac{\eta}{2}\varepsilon_i}(P_i) \hookrightarrow \mathcal{R}_{\eta\varepsilon_i}(P_i) \hookrightarrow \mathcal{C}_{\bar{\eta}\varepsilon_i}(P_i) \hookrightarrow \mathcal{C}_{\bar{\rho}\varepsilon_i}(P_i)$ compose into a map whose rank is equal to $\dim H_*(X^\lambda) = \dim H_*(\mathcal{C}^{\bar{\rho}\varepsilon_i}_{\bar{\eta}\varepsilon_i}(P_i))$ at the homology level. It follows that the map $H_*(\mathcal{R}_{\eta\varepsilon_i}(P_i)) \to H_*(\mathcal{C}^{\bar{\rho}\varepsilon_i}_{\bar{\eta}\varepsilon_i}(P_i))$ in $\mathbb{V}$ is surjective. Similarly, $\mathcal{C}_{\bar{\eta}\varepsilon_i}(P_{i+1}) \hookrightarrow \mathcal{C}_{\bar{\rho}\varepsilon_i}(P_{i+1}) \hookrightarrow \mathcal{R}_{\rho\varepsilon_i}(P_{i+1}) \hookrightarrow \mathcal{C}_{\vartheta_d\rho\varepsilon_i}(P_{i+1})$ compose into a map whose rank is equal to $\dim H_*(X^\lambda) = \dim H_*(\mathcal{C}^{\bar{\rho}\varepsilon_i}_{\bar{\eta}\varepsilon_i}(P_{i+1}))$ at the homology level, so $H_*(\mathcal{C}^{\bar{\rho}\varepsilon_i}_{\bar{\eta}\varepsilon_i}(P_{i+1})) \to H_*(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1}))$ in $\mathbb{V}$ is injective. Thus are satisfied all of the hypotheses of the Interleaving Theorem 4.2, which implies that the persistence barcode of $\mathbb{V}$ has only full-length intervals and some intervals of length zero. The full intervals are in same number as the ones in the barcode of $\mathbb{W}$, whose number is precisely $\dim H_*(X^\lambda)$.

To complete the proof of the theorem, we need to remove the Čech complexes from the oscillating zigzag module $\mathbb{V}$ in order to recover the oR-ZZ. We first restrict $\mathbb{V}$ to the subsequence $H_*(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_k}_{\bar{\eta}\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{C}^{\bar{\rho}\varepsilon_{l-1}}_{\bar{\eta}\varepsilon_{l-1}}(P_l)) \leftarrow H_*(\mathcal{R}_{\eta\varepsilon_l}(P_l))$, thus removing the Čech complexes standing at either ends of the zigzag. Since $k < l$, the Restriction Theorem 2.2 tells us that the full-length intervals in the barcode of the thus shortened zigzag $\mathbb{V}^*$ are in bijection with the ones in the barcode of $\mathbb{V}$, while the intervals of length zero can only be shortened. We then compose the incoming and outgoing maps at Čech complexes in the sequence, to obtain the restriction of the oR-ZZ to $H_*(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}_{\eta\varepsilon_l}(P_l))$. By the Space Removal Theorem 3.2, only the intervals starting or ending at a Čech complex can be affected by this operation, and these can only be shortened. Therefore, the full-length intervals remain in same number as in the barcode of $\mathbb{V}^*$, while the intervals of length zero can only be shortened[4]. The conclusion of the theorem follows. □

5.3 The Morozov zigzag

Following the intuition that η should be made as close to ρ as possible to optimize for speed, the following limit case of the oscillating Rips zigzag where the multipliers ρ, η are equal has been integrated into the Dionysus library [13] since early 2009.

$$\cdots \leftarrow \mathcal{R}_{\rho\varepsilon_{i-1}}(P_{i-1}) \to \mathcal{R}_{\rho\varepsilon_{i-1}}(P_i) \leftarrow \mathcal{R}_{\rho\varepsilon_i}(P_i) \to \mathcal{R}_{\rho\varepsilon_i}(P_{i+1}) \leftarrow \mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1}) \to \cdots \qquad (7)$$

[4] In fact, the persistence barcode is left unchanged by this operation, because as already observed, none of the intervals start or end at a Čech complex.

The zigzag module induced at the homology level by this zigzag filtration is referred to as the *Morozov zigzag* (*M-ZZ*) hereafter. As reported by its author [21], it has given good results in preliminary experiments, despite the fact that $\eta = \rho$ clearly violates the conditions of our theoretical guarantees (Theorem 5.2). Below we provide some weaker guarantees that may explain its good behavior in practice so far.

We begin with a guarantee that the *signal* is present in the barcode of the zigzag throughout a *sweet range* whose bounds are not as good as, yet of the same order of magnitude as, the ones worked out in Theorem 5.2.

Theorem 5.3. *Given a choice of multiplier $\rho > 10$, suppose $P \subset \mathbb{R}^d$ and there is some compact set $X \subset \mathbb{R}^d$ such that $\mathrm{d}_H(P, X) < \varepsilon$ with $\varepsilon < \frac{\rho-10}{(3+10\vartheta_d)\rho}\mathrm{wfs}(X)$. Then, for any $k < l$ such that*

$$\frac{10\varepsilon}{\rho - 10} \leq \varepsilon_k, \varepsilon_l < \min\left\{\frac{1}{6}\mathrm{wfs}(X) - \varepsilon,\ \frac{5}{(1+5\vartheta_d)\rho + 5}(\mathrm{wfs}(X) - \varepsilon)\right\},$$

the M-ZZ restricted to $H_*(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}_{\rho\varepsilon_l}(P_l))$ has a number of full-length intervals that is at least the dimension of $H_*(X^\lambda)$ for any $\lambda \in (0, \mathrm{wfs}(X))$.

Proof. Let $\mathbb{V}$ be the restriction of the M-ZZ to $H_*(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}_{\rho\varepsilon_l}(P_l))$. Let also $\mathbb{U}$ and $\mathbb{W}$ be the restrictions to the same indices of the image Čech zigzags of parameters (η_1, ρ_1) and (η_2, ρ_2) respectively, where

$$\begin{aligned}
\eta_1 &= \frac{\rho}{2} - 2(1 + \frac{\varepsilon}{\varepsilon_l}) \\
\rho_1 &= \frac{\rho}{2} \\
\eta_2 &= \vartheta_d\rho \\
\rho_2 &= \vartheta_d\rho + 2(1 + \frac{\varepsilon}{\varepsilon_l})
\end{aligned}$$

Note that $\eta_1 \leq \rho_1 = \frac{\rho}{2} \leq \vartheta_d\rho = \eta_2 \leq \rho_2$, so the canonical inclusions between these Čech complexes induce homomorphisms between the spaces of $\mathbb{U}$ and $\mathbb{W}$ of same index. Moreover, by (6) the inclusions $\mathcal{C}_{\rho_1\varepsilon_i}(Q) \hookrightarrow \mathcal{C}_{\eta_2\varepsilon_i}(Q)$ factor through the Rips complexes $\mathcal{R}_{\rho\varepsilon_i}(Q)$, so the homomorphisms from $\mathbb{U}$ to $\mathbb{W}$ factor through $\mathbb{V}$, thus giving a commutative diagram of the same form as (2).

Observe now that $3 < \eta_1 < \rho_1 - 2$ and $3 < \eta_2 < \rho_2 - 2$. Moreover, basic calculations show that the assumption on $\varepsilon_k, \varepsilon_l$ made in Theorem 5.1 is satisfied both with $(\eta, \rho) = (\eta_1, \rho_1)$ and with $(\eta, \rho) = (\eta_2, \rho_2)$. Hence, $\mathsf{Pers}(\mathbb{U})$ and $\mathsf{Pers}(\mathbb{W})$ contain only full-length intervals, and the number of intervals in each barcode is exactly the dimension of $H_*(X^\lambda)$.

In addition, the assumptions of Theorem 2.8 (ii) are satisfied for any $i \in [k, l]$, with $\alpha = \eta_1\varepsilon_i$, $\beta = \rho_1\varepsilon_i$, $\alpha' = \eta_2\varepsilon_i$, $\beta' = \rho_2\varepsilon_i$, $P = Q = P_i$ or $P = Q = P_{i+1}$. Hence, the vertical arrows in (2) compose into isomorphisms, and thus the Factorization Theorem 4.1 implies that $\mathsf{Pers}(\mathbb{U}) \subseteq \mathsf{Pers}(\mathbb{V})$. The conclusion follows. □

Remark 5.4. *As pointed out to us by Marc Glisse [16], the approach adopted in this proof can be extended to work with the oscillating Rips zigzag, thereby providing an alternate proof of Theorem 5.2. However, the bounds on the sweet range and (more importantly) on the Rips parameters η, ρ obtained this way are strictly worse than the ones derived in Theorem 5.2.*

According to Theorem 5.3, there is a sweet range throughout which the signal persists in the Morozov zigzag. The resilience of the noise within this range is not well understood though.

Typically, if an index i in the sweet range is such that $\varepsilon_i \geq (2\vartheta_d + \frac{4}{\rho})\varepsilon_{i+1} + \frac{4}{\rho}\varepsilon$, then the canonical inclusion $\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1}) \hookrightarrow \mathcal{R}_{\rho\varepsilon_i}(P_{i+1})$ factors through $\mathcal{C}_{\vartheta_d\rho\varepsilon_{i+1}}(P_{i+1}) \hookrightarrow \mathcal{C}_{\frac{\rho}{2}\varepsilon_i}(P_{i+1})$, with $\frac{\rho}{2}\varepsilon_i - \vartheta_d\rho\varepsilon_{i+1} \geq 2(\varepsilon + \varepsilon_{i+1})$, so Theorem 2.8 (i) implies

$$\operatorname{rank} H_*(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1})) \to H_*(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})) \leq \operatorname{rank} H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_{i+1}}(P_{i+1})) \to H_*(\mathcal{C}_{\frac{\rho}{2}\varepsilon_i}(P_{i+1})) = \dim(H_*(X^\lambda)).$$

In other words, only the signal can go through the link $H_*(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})) \leftarrow H_*(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1}))$, and the noise gets killed. Thus, such indices i with large drops in the geometric scale are desireable, however their existence within the sweet range remains questionable in full generality[5]. For now we will tackle the noise issue from a different perspective and add further restrictive conditions (considered independently): on the one hand, we will focus only on the 0-th and 1-st homology groups; on the other hand, we will assume the shape X underlying the data points to have positive μ-reach for some large enough value μ.

5.3.1 0-th or 1-st homology

As far as only 0-th or 1-st homology is concerned, we can take advantage of the following simple observation.

Lemma 5.5. *For any $Q \subset \mathbb{R}^d$ and any $\beta \geq \alpha \geq 0$,*

$$\operatorname{rank} H_0(\mathcal{R}_\alpha(Q)) \to H_0(\mathcal{R}_\beta(Q)) = \operatorname{rank} H_0(\mathcal{C}_{\frac{\alpha}{2}}(Q)) \to H_0(\mathcal{C}_{\frac{\beta}{2}}(Q)),$$
$$\operatorname{rank} H_1(\mathcal{R}_\alpha(Q)) \to H_1(\mathcal{R}_\beta(Q)) \leq \operatorname{rank} H_1(\mathcal{C}_{\frac{\alpha}{2}}(Q)) \to H_1(\mathcal{C}_{\frac{\beta}{2}}(Q)),$$

where the homomorphisms are induced at the homology level by canonical inclusions between the complexes.

Proof. Recall from [14] that for any finite simplicial complexes $X \subseteq Y$, the rank of the homomorphism induced at the r-th homology level by the canonical inclusion $X \subseteq Y$ is given by

$$\operatorname{rank} H_r(X) \to H_r(Y) = \dim \frac{Z_r(X)}{Z_r(X) \cap B_r(Y)}, \tag{8}$$

where $Z_r(X)$ denotes the space of r-cycles in X and $B_r(Y)$ denotes the space of r-boundaries in Y (both are subgroups of the space of r-chains in Y).

When $Q \subset \mathbb{R}^d$, it follows from the definitions of Čech and Rips complexes that $\mathcal{C}_{\frac{\gamma}{2}}(Q)$ and $\mathcal{R}_\gamma(Q)$ have the same 1-skeleton, given any $\gamma \geq 0$. Hence, $Z_0(\mathcal{C}_{\frac{\alpha}{2}}(Q)) = Z_0(\mathcal{R}_\alpha(Q))$ and $B_0(\mathcal{C}_{\frac{\beta}{2}}(Q)) = B_0(\mathcal{R}_\beta(Q))$, which implies by (8) that the maps $H_0(\mathcal{R}_\alpha(Q)) \to H_0(\mathcal{R}_\beta(Q))$ and $H_0(\mathcal{C}_{\frac{\alpha}{2}}(Q)) \to H_0(\mathcal{C}_{\frac{\beta}{2}}(Q))$ have same rank.

The definitions of Čech and Rips complexes also imply that the 2-skeleton of $\mathcal{R}_\gamma(Q)$ contains the one of $\mathcal{C}_{\frac{\gamma}{2}}(Q)$, while their 1-skeleton is the same as mentioned previously. Hence, $Z_1(\mathcal{C}_{\frac{\alpha}{2}}(Q)) = Z_1(\mathcal{R}_\alpha(Q))$ and $B_1(\mathcal{C}_{\frac{\beta}{2}}(Q)) \subseteq B_1(\mathcal{R}_\beta(Q))$, which implies by (8) that $\operatorname{rank} H_1(\mathcal{R}_\alpha(Q)) \to H_1(\mathcal{R}_\beta(Q)) \leq \operatorname{rank} H_1(\mathcal{C}_{\frac{\alpha}{2}}(Q)) \to H_1(\mathcal{C}_{\frac{\beta}{2}}(Q))$. □

Letting now $P, \rho, \varepsilon, \varepsilon_k, \varepsilon_l$ follow the hypotheses of Theorem 5.3, we have by Lemma 5.5 and for every index $i \in [k, l]$

$$\operatorname{rank} H_0(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1})) \to H_0(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})) = \operatorname{rank} H_0(\mathcal{C}_{\frac{\rho}{2}\varepsilon_{i+1}}(P_{i+1})) \to H_0(\mathcal{C}_{\frac{\rho}{2}\varepsilon_i}(P_{i+1})) = \dim(H_0(X^\lambda)),$$
$$\operatorname{rank} H_1(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1})) \to H_1(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})) \leq \operatorname{rank} H_1(\mathcal{C}_{\frac{\rho}{2}\varepsilon_{i+1}}(P_{i+1})) \to H_1(\mathcal{C}_{\frac{\rho}{2}\varepsilon_i}(P_{i+1})) = \dim(H_1(X^\lambda)).$$

[5] Nevertheless, they inspired the discretization scheme presented in Section 5.4.

Hence, in 0-th or 1-st homology the noise in the Morozov zigzag is killed when going through the link $\mathcal{R}_{\rho\varepsilon_i}(P_{i+1}) \leftarrow \mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1})$. More precisely, given $r \in \{0,1\}$, call $\mathbb{V}$ the restriction of the Morozov zigzag to $H_r(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_r(\mathcal{R}_{\rho\varepsilon_l}(P_l))$. On the one hand, the Restriction Theorem 2.2 implies that the total multiplicity of the intervals including $[H_r(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})),\ H_r(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1}))]$ in $\mathbb{V}$ is at most $\dim(H_r(X^\lambda))$. On the other hand, Theorem 5.3 implies that the multiplicity of the full-length interval in $\mathbb{V}$ is precisely $\dim(H_r(X^\lambda))$. It follows that among the intervals containing $[H_r(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})),\ H_r(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1}))]$, only the full-length one has non-zero multiplicity. Thus, $\mathsf{Pers}_r(\mathbb{V})$ contains only full-length intervals and intervals of type $[H_r(\mathcal{R}_{\rho\varepsilon_i}(P_i)),\ H_r(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1}))]$. These are not ephemeral in the index scale of $\mathbb{V}$, however they become so once represented on the scale of the geometric scales. Hence,

Theorem 5.6. *Given a choice of multiplier $\rho > 10$, suppose $P \subset \mathbb{R}^d$ and there is some compact set $X \subset \mathbb{R}^d$ such that $\mathrm{d}_H(P,X) < \varepsilon$ with $\varepsilon < \frac{\rho-10}{(3+10\vartheta_d)\rho}\mathrm{wfs}(X)$. Then, for any $k < l$ such that*

$$\frac{10\varepsilon}{\rho-10} \leq \varepsilon_k, \varepsilon_l < \min\left\{\frac{1}{6}\mathrm{wfs}(X) - \varepsilon,\ \frac{5}{(1+5\vartheta_d)\rho+5}(\mathrm{wfs}(X)-\varepsilon)\right\},$$

the zigzag module induced by (7) at the r-th homology level ($r \in \{0,1\}$), once restricted to $H_r(\mathcal{R}_{\rho\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_r(\mathcal{R}_{\rho\varepsilon_l}(P_l))$, has a persistence barcode made only of two types of intervals:

- *full-length intervals (the signal), whose number is equal to the dimension of $H_r(X^\lambda)$ for any $\lambda \in (0, \mathrm{wfs}(X))$,*
- *intervals of type $[H_r(\mathcal{R}_{\rho\varepsilon_i}(P_i)),\ H_r(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1}))]$ (the noise), which are ephemeral (length zero) on the scale of the geometric scales.*

5.3.2 Sampled compact sets of positive μ-reach

Thus far, we have only considered the case of compact sets X for which $\mathrm{wfs}(X)$ is positive. Now, we consider the stronger assumption that X has positive μ-reach. Recall that the μ-reach of a compact set X is the infimum of distances from X to points outside of X where the gradient of the distance to X is less than μ (see [6]). Attali et al. showed that for a compact set X with μ-reach R and a sample P of points with $\mathrm{d}_H(P,X) \leq \varepsilon$, if ε is sufficiently small and μ sufficiently large then for some values of α, the Rips complex $\mathcal{R}(P,\alpha)$ is homotopy equivalent to X^λ for $\lambda \in (0,R)$ (see [1, Theorem 14]). The immediate consequence of their result is that for a multiplier ρ and an index i, $\mathcal{R}_{\rho\varepsilon_i}(P_i)$ and $\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})$ are both homotopy equivalent to X^λ for $\eta \in (0,R)$ whenever

$$\frac{1+\mu(1-\mu)-\sqrt{1-\mu(2-\mu)\left(\frac{(2\vartheta_d\rho+1)\varepsilon_i+\varepsilon}{R}\right)^2}}{\mu(2-\mu)}R < 2\rho\varepsilon_i - 2\vartheta_d\varepsilon_i - 2(\varepsilon_i+\varepsilon).$$

This condition depends on μ, the μ-reach of X, the multiplier ρ, the Hausdorff distance of the sample ε, and the range of values ε_i for which the result holds. However, Attali et al. show that there do exist values for which the condition is satisfied.

We do not derive the space of valid assignment of constants here, but merely note that this result implies that for a sufficiently close sample of a set with bounded μ-reach, there is a multiplier ρ and a range of scales for which the M-ZZ exhibits no noise. This holds because Theorem 5.3 implies that the signal is present in the sweet range and the Attali et al. result shows that every space in the strictly smaller range has the same homology as X^λ. We call this the *sweeter range*.

Since the sweeter range depends on the μ-reach rather than the weak feature size, it may be arbitrarily smaller than the sweet range. However, when the input permits a sweeter range, the guarantees regarding the signal-to-noise ratio are correspondingly stronger. This will hold true for all of the Rips Zigzags discussed in this paper.

5.4 Discretized Morozov zigzag

We now describe a discretization scheme for the Morozov zigzag that ensures that the desired *large geometric scale drop* condition mentioned in Section 5.3 is satisfied.

Given a map $\zeta : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, referred to as the *scale drop function* hereafter, we select a subset of the indices $1, \cdots, n$ in the ordered point cloud $P = \{p_1, \cdots, p_n\}$ by the following iterative procedure:

$$\begin{aligned} &\text{let } n_1 = 1, \\ \forall i \geq 1,\ &\text{let } n_{i+1} = \min\left\{ j > n_i \ \mid\ \varepsilon_j \leq \zeta(\varepsilon_{n_i}) \cdot \varepsilon_{n_i} \right\}. \end{aligned}$$

Note that $n_r = n$ for some index r since $\varepsilon_n = 0 < \varepsilon_{n-1}$. We then build the following discretized version of the Morozov zigzag filtration (7):

$$\mathcal{R}_{\rho\varepsilon_{n_1}}(P_{n_1}) \to \mathcal{R}_{\rho\varepsilon_{n_1}}(P_{n_2}) \leftarrow \cdots \to \mathcal{R}_{\rho\varepsilon_{n_{k-1}}}(P_{n_k}) \leftarrow \mathcal{R}_{\rho\varepsilon_{n_k}}(P_{n_k}), \tag{9}$$

where $P_{n_1} = P_1 = \{p_1\}$ and $P_{n_r} = P_n = P$. The zigzag induced at the homology level is called the *discretized Morozov zigzag* (*dM-ZZ*) hereafter.

The proof of Theorem 5.3 applies verbatim to the dM-ZZ, with indices $1, 2, \cdots, n$ replaced by $n_1, n_2, \cdots, n_r$. Thus, for any $n_k < n_l$ such that $\varepsilon_{n_k}, \varepsilon_{n_l}$ lie within the sweet range defined in Theorem 5.3, the restriction $\mathbb{V}$ of the dM-ZZ to $H_*(\mathcal{R}_{\rho\varepsilon_{n_k}}(P_{n_{k+1}})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}_{\rho\varepsilon_{n_l}}(P_{n_l}))$ has a persistence barcode with at least $\dim(H_*(X^\lambda))$ full-length intervals.

Now, if we further assume that the scale drop function ζ satisfies the following condition:

$$\forall i \in [1, r-1],\ \zeta(\varepsilon_{n_i}) \leq \frac{\rho}{2\vartheta_d\rho + 4} - \frac{2}{\vartheta_d\rho + 2} \cdot \frac{\varepsilon}{\varepsilon_{n_i}}, \tag{10}$$

then, the *large geometric scale drop* condition mentioned in Section 5.3 is satisfied for all $i \in [k, l]$, that is,

$$\varepsilon_{n_i} \geq (2\vartheta_d + \frac{4}{\rho})\varepsilon_{n_{i+1}} + \frac{4}{\rho}\varepsilon,$$

so we have

$$\operatorname{rank} H_*(\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}})) \leftarrow H_*(\mathcal{R}_{\rho\varepsilon_{n_{i+1}}}(P_{n_{i+1}})) \leq \dim(H_*(X^\lambda)),$$

and therefore by the Restriction Theorem 2.2 the total multiplicity of the intervals including $[H_*(\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}})),\ H_*(\mathcal{R}_{\rho\varepsilon_{n_{i+1}}}(P_{n_{i+1}}))]$ in $\mathsf{Pers}(\mathbb{V})$ is at most $\dim(H_*(X^\lambda))$. It follows that among these intervals only the full-length one has non-zero multiplicity. Thus, $\mathsf{Pers}(\mathbb{V})$ contains only full-length intervals and intervals of type $[H_*(\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_i})),\ H_*(\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}}))]$. These are not ephemeral in the index scale of $\mathbb{V}$, however they become so once represented on the scale of the geometric scales. Hence,

Theorem 5.7. *Given a choice of multiplier $\rho > 10$, suppose $P \subset \mathbb{R}^d$ and there is some compact set $X \subset \mathbb{R}^d$ such that $\mathrm{d}_H(P, X) < \varepsilon$ with $\varepsilon < \frac{\rho-10}{(3+10\vartheta_d)\rho}\mathrm{wfs}(X)$. Then, for any choice of scale drop function $\zeta : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ that satisfies (10), for any $n_k < n_l$ such that*

$$\frac{10\varepsilon}{\rho - 10} \leq \varepsilon_{n_k}, \varepsilon_{n_l} < \min\left\{ \frac{1}{6}\mathrm{wfs}(X) - \varepsilon,\ \frac{5}{(1 + 5\vartheta_d)\rho + 5}(\mathrm{wfs}(X) - \varepsilon) \right\},$$

the dM-ZZ restricted to $H_*(\mathcal{R}_{\rho\varepsilon_{n_k}}(P_{n_{k+1}})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}_{\rho\varepsilon_{n_l}}(P_{n_l}))$*, has a persistence barcode made only of two types of intervals:*

- *full-length intervals (the signal), whose number is equal to the dimension of* $H_*(X^\lambda)$ *for any* $\lambda \in (0, \mathrm{wfs}(X))$*,*
- *intervals of type* $[H_*(\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_i})),\ H_*(\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}}))]$ *(the noise), which are ephemeral (length zero) on the scale of the geometric scales.*

Note that ε usually remains unknown in practice, so the user cannot merely set $\zeta(\varepsilon_{n_i})$ to be the quantity $\frac{\rho}{2\vartheta_d\rho+4} - \frac{2}{\vartheta_d\rho+2} \cdot \frac{\varepsilon}{\varepsilon_{n_i}}$. The bounds given in Theorem 5.7 suggest to let ζ be the constant map

$$\zeta = \frac{3\rho + 20}{10(\vartheta_d\rho + 2)}, \tag{11}$$

so that $\zeta(\varepsilon_{n_i})$ satisfies the condition of (10) as long as $\varepsilon_{n_i} \geq \frac{10\varepsilon}{\rho-10}$. Thus, the conclusion of the theorem continues to hold within the same sweet range. In fact, any smaller value could be chosen for $\zeta(\varepsilon_{n_i})$ without affecting the sweet range. Nevertheless, the larger $\zeta(\varepsilon_{n_i})$ the better in general, since the smaller the geometric scale gap the more chances there are that the set of discretization values ε_{n_i} intersects the sweet range, and furthermore the smaller the geometric scale gap the smaller the complexes involved in the discretized Morozov zigzag filtration (9).

5.5 Image Rips zigzag

We end this section with another variant of our Rips zigzag construction, called the *image Rips zigzag* (*iR-ZZ*). It is the same as (5) with Čech complexes replaced by Rips complexes. Given two multipliers $\rho \geq \eta \geq 0$, canonical inclusions between Rips complexes give the following commutative diagram where a pair of Morozov zigzags are connected together by vertical arrows.

$$\begin{array}{ccccccccc}
\cdots & \leftarrow & H_*(\mathcal{R}_{\rho\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{R}_{\eta\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{R}_{\eta\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{R}_{\eta\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots
\end{array}$$

This commutative diagram induces the following image Rips zigzag, where the notation $H_*(\mathcal{R}^\beta_\alpha(Q))$ stands for the image of the homomorphism $H_*(\mathcal{R}_\alpha(Q)) \to H_*(\mathcal{R}_\beta(Q))$ induced at the homology level by the inclusion $\mathcal{R}_\alpha(Q) \hookrightarrow \mathcal{R}_\beta(Q)$.

$$\cdots \leftarrow H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_i)) \rightarrow H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_{i+1})) \leftarrow H_*(\mathcal{R}^{\rho\varepsilon_{i+1}}_{\eta\varepsilon_{i+1}}(P_{i+1})) \rightarrow \cdots \tag{12}$$

Image Rips zigzags have been available in the Dionysus library [13] since 2009, with no theoretical guarantee on their behavior. Here we provide a guarantee on the output that is similar to (and even slightly better than) the one obtained for the oscillating Rips zigzag.

Theorem 5.8. *Given a choice of multipliers* ρ, η *such that* $\rho > 10$ *and* $\frac{3}{\vartheta_d} < \eta < \frac{\rho-4}{2\vartheta_d}$*, suppose* $P \subset \mathbb{R}^d$ *and there is some compact set* $X \subset \mathbb{R}^d$ *such that* $\mathrm{d}_H(P, X) < \varepsilon$ *with* $\varepsilon < \min\left\{\frac{\vartheta_d\eta-3}{6\vartheta_d\eta}, \frac{\eta-3/\vartheta_d}{3\rho+\eta}, \frac{\rho-2\vartheta_d\eta-4}{6(\rho-2\vartheta_d\eta)}, \frac{\rho-2\vartheta_d\eta-4}{(4\vartheta_d+1)\rho-2\vartheta_d\eta}\right\}\mathrm{wfs}(X)$*. Then, for any* $k < l$ *such that*

$$\max\left\{\frac{3\varepsilon}{\vartheta_d\eta - 3}, \frac{4\varepsilon}{\rho - 2\vartheta_d\eta - 4}\right\} \leq \varepsilon_k, \varepsilon_l < \min\left\{\frac{1}{6}\mathrm{wfs}(X) - \varepsilon, \frac{1}{\vartheta_d\rho + 1}\left(\mathrm{wfs}(X) - \varepsilon\right)\right\},$$

the iR-ZZ restricted to $H_*(\mathcal{R}^{\rho\varepsilon_k}_{\eta\varepsilon_k}(P_{k+1})) \leftarrow \cdots \leftarrow H_*(\mathcal{R}^{\rho\varepsilon_l}_{\eta\varepsilon_l}(P_l))$ *contains only isomorphisms, and its spaces are isomorphic to* $H_*(X^\lambda)$ *for any* $\lambda \in (0, \mathrm{wfs}(X))$*. Therefore, its persistence barcode is made only of full-length intervals, whose number equals the dimension of* $H_*(X^\lambda)$*.*

Proof. Our hypotheses imply $\frac{\rho}{2} \geq \vartheta_d\eta$, so we can use (6) to obtain the following diagram interleaving Čech and Rips zigzags (all arrows are inclusions).

$$\begin{array}{ccccccccccc}
\cdots & \leftarrow & H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{R}_{\rho\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{R}_{\rho\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{C}_{\frac{\rho}{2}\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}_{\frac{\rho}{2}\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}_{\frac{\rho}{2}\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{C}_{\vartheta_d\eta\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}_{\vartheta_d\eta\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}_{\vartheta_d\eta\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{R}_{\eta\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{R}_{\eta\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{R}_{\eta\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{C}_{\frac{\eta}{2}\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}_{\frac{\eta}{2}\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}_{\frac{\eta}{2}\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots
\end{array}$$

This commutative diagram induces the following interleaving between the image Rips and image Čech zigzags at the homology level (note that the quadrangles still commute).

$$\begin{array}{ccccccccc}
\cdots & \leftarrow & H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_i}_{\vartheta_d\eta\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_i}_{\vartheta_d\eta\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_{i+1}}_{\vartheta_d\eta\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \\
\cdots & \leftarrow & H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_i)) & \rightarrow & H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_{i+1})) & \leftarrow & H_*(\mathcal{R}^{\rho\varepsilon_{i+1}}_{\eta\varepsilon_{i+1}}(P_{i+1})) & \rightarrow & \cdots
\end{array} \tag{13}$$

The hypotheses of Theorem 5.1 are satisfied with (η, ρ) replaced by $(\vartheta_d\eta, \vartheta_d\rho)$, so we can apply that theorem and deduce that the spaces in the image Čech zigzag restricted to the index range $[k, l]$ are isomorphic to $H_*(X^\lambda)$, and that all the arrows in this restricted zigzag are isomorphisms.

We will now show that the restriction of the iR-ZZ to the same range is isomorphic to the restriction of the image Čech zigzag, which boils down to proving that the vertical arrows $H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_i)) \to H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_i}_{\vartheta_d\eta\varepsilon_i}(P_i))$ and $H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_{i+1})) \to H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_i}_{\vartheta_d\eta\varepsilon_i}(P_{i+1}))$ in (13) are isomorphisms for all $i \in [k, l]$.

First, one can check that the hypotheses of Theorem 2.8 (i) are satisfied for all $i \in [k, l]$ so we can apply that theorem together with rank inequalities induced by composition, to obtain

$$\dim(H_*(X^\lambda)) = \dim(H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_i}_{\frac{\eta}{2}\varepsilon_i}(P_i))) \leq \dim(H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_i))) \leq \dim(H_*(\mathcal{C}^{\frac{\rho}{2}\varepsilon_i}_{\vartheta_d\eta\varepsilon_i}(P_i))) = \dim(H_*(X^\lambda)).$$

The same holds when P_i is replaced by P_{i+1}. Thus, the spaces in the restriction of the iR-ZZ to the index range $[k, l]$ are isomorphic to $H_*(X^\lambda)$.

Second, notice that

$$\operatorname{rank} H_*(\mathcal{R}^{\rho\varepsilon_i}_{\eta\varepsilon_i}(P_i)) \to H_*(\mathcal{C}^{\vartheta_d\rho\varepsilon_i}_{\vartheta_d\eta\varepsilon_i}(P_i)) = \operatorname{rank} H_*(\mathcal{R}_{\eta\varepsilon_i}(P_i)) \to H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_i}(P_i)),$$

which by composition is sandwiched between

$$\operatorname{rank} H_*(\mathcal{C}_{\frac{\eta}{2}\varepsilon_i}(P_i)) \to H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_i}(P_i))$$

and

$$\operatorname{rank} H_*(\mathcal{C}_{\vartheta_d\eta\varepsilon_i}(P_i)) \to H_*(\mathcal{C}_{\vartheta_d\rho\varepsilon_i}(P_i)),$$

which by Theorem 2.8 (i) are both equal to $\dim(H_*(X^\lambda))$. The same holds when P_i is replaced by P_{i+1}. Hence, the vertical arrows in (13) are isomorphisms when $i \in [k, l]$, which concludes the proof of the theorem. □

6 Discussion

6.1 Caveats regarding zigzags manipulations

A few technical comments on the Arrow Reversal Theorem 3.1 and Space Removal Theorem 3.2 are in order at this point, to help prevent potential misuses of these results.

The statement of Theorem 3.1 claims the existence of a suitable reverse map g, not its uniqueness. It clearly appears from our proof that many such maps may exist — see the construction of the kernel of g in the proof of Lemma 3.4. Thus, Theorem 3.1 does not provide a canonical way of reversing arrows in zigzags. Nevertheless, this freedom in the choice of g remains limited. In particular, taking an arbitrary map g such that $f \circ g|_{\operatorname{im} f} = \mathbb{1}_{\operatorname{im} f}$ and $g \circ f|_{\operatorname{im} g} = \mathbb{1}_{\operatorname{im} g}$ is not sufficient in general to guarantee the preservation of the persistence barcode. Consider the following example

$$\mathbb{V} = V_1 \xrightarrow{f} V_2 \xleftarrow{h} V_3$$

where $V_1 = V_3 = \mathbb{F}$, $V_2 = \mathbb{F}^2$, f maps $\mathbb{F}$ isomorphically to the first coordinate space in $\mathbb{F}^2$, and h maps $\mathbb{F}$ isomorphically to the second coordinate space in $\mathbb{F}^2$. Suppose we take $g : V_2 \to V_1$ such that $\ker g$ is the diagonal in $\mathbb{F}^2$ while $g|_{\operatorname{im} f} = f^{-1}$. Then, both conditions $f \circ g|_{\operatorname{im} f} = \mathbb{1}_{\operatorname{im} f}$ and $g \circ f|_{\operatorname{im} g} = \mathbb{1}_{\operatorname{im} g}$ are satisfied, but it is not hard to see that the barcodes of $\mathbb{V}$ and of $\mathbb{V}^*$ are different, and in particular that $\operatorname{mult}([1,3], \mathbb{V}) = 0$ while $\operatorname{mult}([1,3], \mathbb{V}^*) = 1$.

The same caveats apply to the statement of Theorem 3.2. When the maps f and g are oriented the same way (either forwards or backwards), h is defined uniquely as their composition. However, when f and g are oriented differently, there is some freedom in the choice of h, which can be seen when the Arrow Reversal Theorem 3.1 is invoked in our proof. However, this freedom remains limited, and taking an arbitrary map making the triangle in the statement of Theorem 3.2 commute is not sufficient in general to make assertions (a) through (c) hold. This is true even in very special cases such as when $V_{k-1} = V_{k+1}$, $f = g$, and h is taken to be the identity of V_{k-1}. Consider the following example

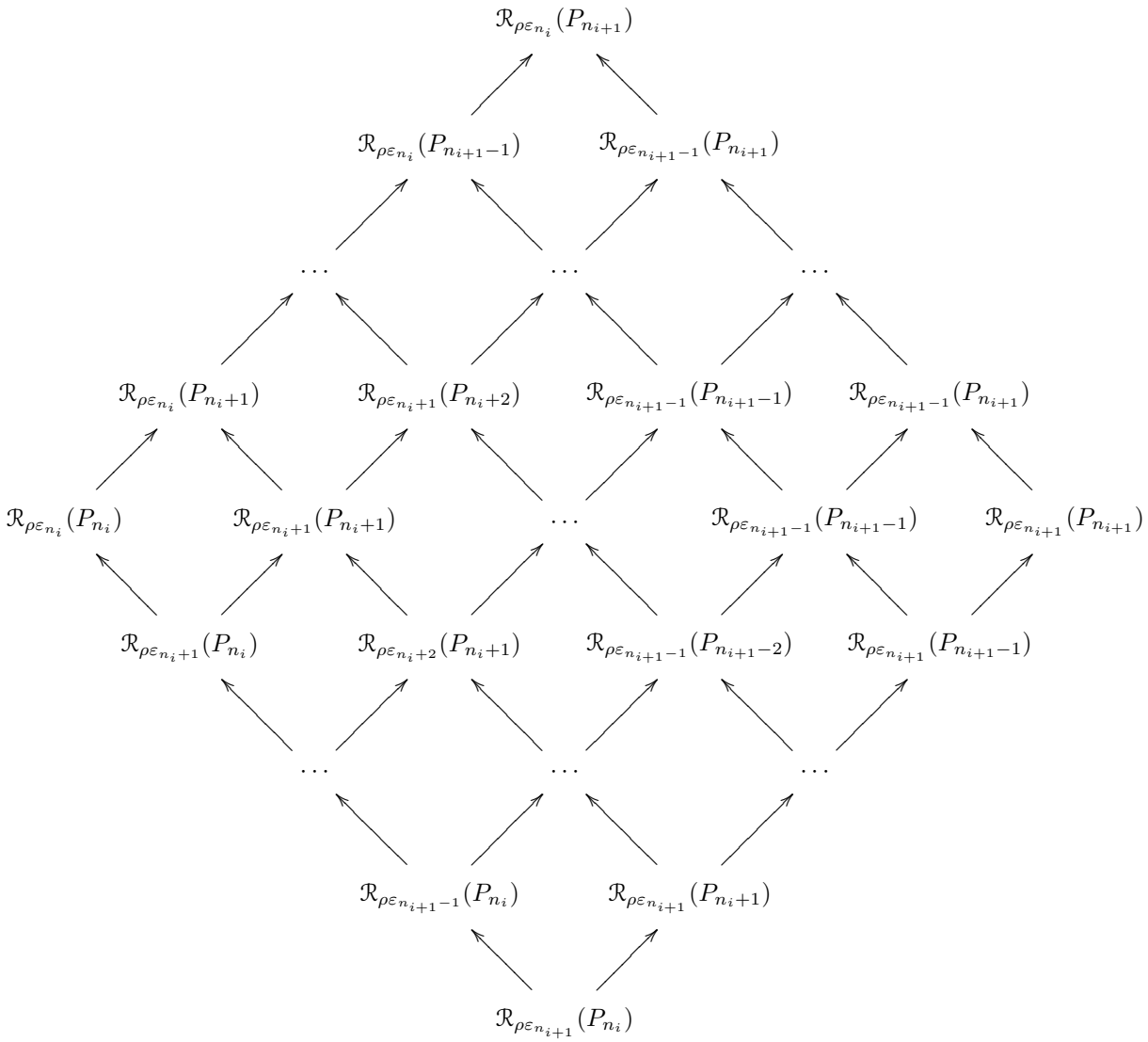
$$\mathbb{V} = V_1 \xrightarrow{r} V_2 \xrightarrow{f} V_3 \xleftarrow{f} V_4 \xleftarrow{s} V_5$$

where $k = 3$, $V_1 = V_3 = V_5 = \mathbb{F}$, $V_2 = V_4 = \mathbb{F}^2$, f projects $\mathbb{F}^2$ onto its first coordinate space, r maps $\mathbb{F}$ isomorphically to the first coordinate space in $\mathbb{F}^2$, and s maps $\mathbb{F}$ isomorphically to the diagonal in $\mathbb{F}^2$. Suppose $h : V_2 \to V_4$ is chosen to be the identity of $\mathbb{F}^2$. Then, the triangle formed by $V_2 \xrightarrow{f} V_3 \xleftarrow{f} V_4$ and $V_2 \xrightarrow{h} V_4$ clearly commutes, but a calculation shows that $\operatorname{mult}([1,5], \mathbb{V}) = 1$ while $\operatorname{mult}([1,5], \mathbb{V}^*) = 0$, thus contradicting assertion (c) of Theorem 3.2. Note that it also contradicts the intuition from (non-zigzag) persistence theory that adding an intermediate space V_k in-between V_{k-1} and V_{k+1} cannot increase the number of topological features that persist throughout the interval $[k-1, k+1]$.

6.2 Discretized Morozov zigzag, pyramid and weak diamonds

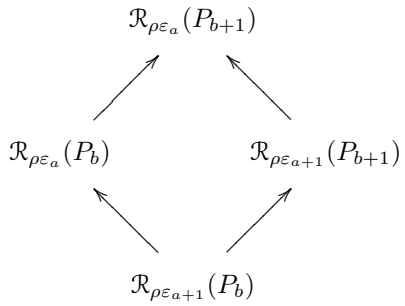
Since discretizing the Morozov zigzag results in killing most of the topological noise within the sweet range (see Theorem 5.7), it is tempting to try to relate the Morozov zigzag to its discretization. A natural way to do so is through the following double pyramid of Rips complexes,

which can be built for every consecutive discretization indices n_i, n_{i+1}.

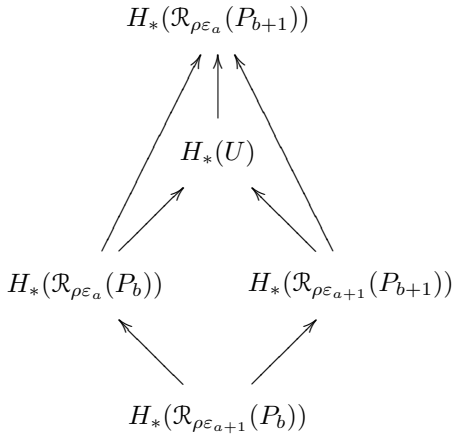


Following an up-right arrow in this pyramid means adding one point to the vertex set, while following an up-left arrow means increasing the Rips parameter. Given indices a, b within the

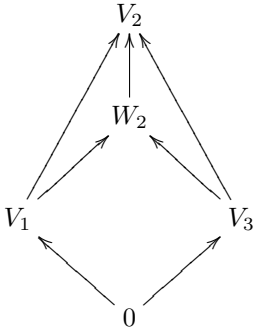
range $[n_i, n_{i+1})$, the diamond



is not Mayer-Vietoris in general since $\mathcal{R}_{\rho\varepsilon_a}(P_{b+1})$ may be a strict superset of the union U of $\mathcal{R}_{\rho\varepsilon_a}(P_b)$ and $\mathcal{R}_{\rho\varepsilon_{a+1}}(P_{b+1})$. Nevertheless, $\mathcal{R}_{\rho\varepsilon_{a+1}}(P_b)$ does coincide with the intersection, hence we get the following commutative diagram with an exact diamond at the homology level.



Intuition from classical persistence theory suggests that going through the superset $\mathcal{R}_{\rho\varepsilon_a}(P_{b+1})$ rather than through U can only kill more topological features. Such a statement, if true, would induce a weak version of the diamond principle [3, Thm 5.6], and it could be used to travel up and down the pyramid to show that the topological signal persists throughout the sweet range while the topological noise does not in the Morozov zigzag. Unfortunately, as appealing as it may be, this approach does not work because the weak diamond principle turns out to be false. Here is an example illustrating what may happen at the algebraic level. Consider the following diagram



(14)

where $V_1 = V_2 = V_3 = \mathbb{F}$, $W_2 = \mathbb{F}^2$, the map $W_2 \to V_2$ projects $\mathbb{F}^2$ onto its first coordinate space, the map $V_1 \to W_2$ is an isomorphism onto the first coordinate space in $\mathbb{F}^2$, the map $V_3 \to W_2$ is an isomorphism onto the diagonal in $\mathbb{F}^2$, and the maps $V_1 \to V_2$ and $V_3 \to V_2$ are obtained by composition. Then, it is easily seen that the diamond is exact, while a quick calculation shows that the interval $[1, 3]$ has multiplicity 1 in the upper zigzag $V_1 \to V_2 \leftarrow V_3$ and 0 in the intermediate zigzag $V_1 \to W_2 \leftarrow V_3$. As in Section 6.1, the change in multiplicity occurs when going from $V_1 \to W_2 \to V_2 \leftarrow W_2 \leftarrow V_3$ to $V_1 \to W_2 \xrightarrow{\mathbb{1}} W_2 \leftarrow V_3$.

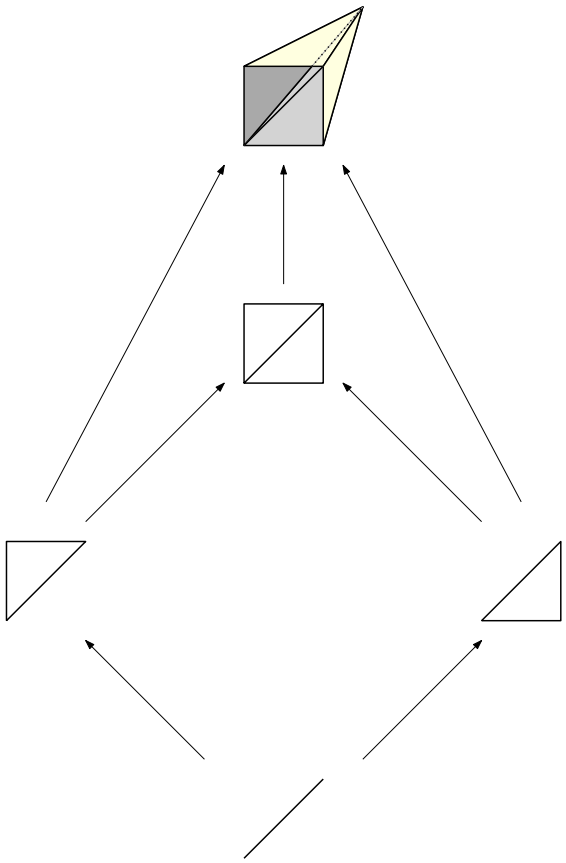


Figure 1: A diagram of simplicial complexes where all the maps are canonical inclusions.

Such scenarios at the algebraic level can derive from diagrams of simplicial complexes at the topological level. Take for instance the example of Figure 1, which induces the same diagram as in (14) at the 1-homology level. This example can easily be adapted so the complexes are Rips complexes, and we leave the details as an exercise to the reader.

6.3 Zigzags rectification, extrusion, and interleaving

The Arrow Reversal Theorem 3.1 provides a way of turning zigzag modules into regular persistence modules without affecting their persistence barcode. We call this operation a *rectification*. More precisely,

Definition 6.1. *Given a zigzag module* $\mathbb{V}$*, a* rectification *of* $\mathbb{V}$ *is a regular persistence module composed of the same sequence of vector spaces and the same forward maps as* $\mathbb{V}$*, and whose persistence barcode is the same as the one of* $\mathbb{V}$*.*

Theorem 3.1 guarantees the existence of such rectifications for any zigzag module $\mathbb{V}$.

Now, a rectification $\mathbb{V}^*$ of $\mathbb{V}$ is indexed over the same integer interval $[1, n]$, not over all of $\mathbb{R}$, which makes it unsuitable for interleavings with other persistence modules [5, 7]. Turning $\mathbb{V}^*$ into a module indexed over $\mathbb{R}$ can be done by *extrusion*, meaning that every space V_i in the sequence is replaced by the indexed family $\{V_i\}_{\alpha\in[i,i+1)}$ linked by identity maps, and that two constant families $\{0\}_{\alpha<1}$ and $\{0\}_{\alpha\geq n+1}$ are appended at either ends of the module, the rest of the maps being defined by composition. More formally,

Definition 6.2. *Given a persistence module*

$$\mathbb{V} = V_1 \overset{v_1^2}{\to} V_2 \overset{v_2^3}{\to} \cdots \overset{v_{n-1}^n}{\to} V_n,$$

the extrusion *of* $\mathbb{V}$ *is the persistence module* $\bar{\mathbb{V}}$ *indexed over* $\mathbb{R}$ *such that*

$$\begin{array}{rcll} \bar{V}_\alpha & = & V_{\lfloor\alpha\rfloor} & \text{if } 1 \leq \alpha < n+1 \\ & & 0 & \text{otherwise} \\ \bar{v}_\alpha^\beta & = & \mathbb{1}_{V_{\lfloor\alpha\rfloor}} & \text{if } 1 \leq \alpha \leq \beta < n+1 \text{ and } \lfloor\alpha\rfloor = \lfloor\beta\rfloor \\ & & v_{\lfloor\beta\rfloor-1}^{\lfloor\beta\rfloor} \circ \cdots \circ v_{\lfloor\alpha\rfloor}^{\lfloor\alpha\rfloor+1} & \text{if } 1 \leq \alpha \leq \beta < n+1 \text{ and } \lfloor\alpha\rfloor < \lfloor\beta\rfloor \\ & & 0 & \text{otherwise} \end{array}$$

Together, Definitions 6.1 and 6.2 give the following notion of interleaving between zigzags.

Definition 6.3. *Two zigzag modules* $\mathbb{V}$ *and* $\mathbb{W}$ *are* δ*-interleaved for some* $\delta \geq 0$ *if there are rectifications* $\mathbb{V}^*$ *of* $\mathbb{V}$ *and* $\mathbb{W}^*$ *of* $\mathbb{W}$ *whose extrusions are* δ*-interleaved in the sense of [5, 7].*

Once again, Theorem 3.1 guarantees the existence of rectifications for any zigzag module, so two zigzag modules $\mathbb{V}, \mathbb{W}$ indexed respectively over $[1, n]$ and over $[1, m]$ are always δ-interleaved for some $\delta \leq \frac{1}{2}\max\{n, m\}$.

Extruding the rectifications of $\mathbb{V}$ and $\mathbb{W}$ has a predictable effect on their persistence barcodes: every interval $[k, l]$ becomes $[k, l+1)$. This operation corresponds to a mere switch from the closed interval representation introduced specifically for zigzag modules in [3], and the standard half-open interval representation used in the literature on non-zigzag persistence [14, 23]. Defining the bottleneck distance between $\mathsf{Pers}(\mathbb{V})$ and $\mathsf{Pers}(\mathbb{W})$ to be the usual bottleneck distance between the standard representations of these barcodes, we obtain the following guarantee as a direct corollary of the stability theorem from non-zigzag persistence theory [5, 7, 10].

Theorem 6.4. *If two zigzag modules are* δ*-interleaved in the sense of Definition 6.3, then their persistence barcodes (in the half-open interval representation) are* δ*-close in the bottleneck distance.*

One may rightfully ask two questions at this point:

1. Why use extrusion specifically, and not another method, to turn the rectifications into persistence modules over $\mathbb{R}$? Extrusion has two properties that are appealing in our context: first, it is transparent in the sense that it does not affect the persistence barcodes (up to

a change in the representation); second, it is reversible in the sense that discretizing the extrusion of a persistence module $\mathbb{V}$ over its original integer index set gives $\mathbb{V}$ itself. Now, any operation with these two properties (or equivalent ones) would do for turning the rectifications into persistence modules over $\mathbb{R}$.

2. Why define the rectification of a zigzag $\mathbb{V}$ as in Definition 6.1, and not merely as an arbitraty regular persistence module with the same persistence barcode as $\mathbb{V}$? Or on the contrary, why not impose the additional constraint that the backward maps of $\mathbb{V}$ and their corresponding forward maps in the rectification satisfy all the properties stated in Theorem 3.1? We do not have a final answer to these questions. Basically, Definition 6.1 looks more canonical to us than the proposed variants. However, working out the "right" notion of interleaving between zigzag modules remains a wide open question at present, and our only claim here is that the Arrow Reversal Theorem may shed new light on it by suggesting a new approach via a reduction to standard persistence.

6.4 Riemannian manifolds and Alexandrov spaces

The analysis of the various Rips zigzags of Section 5 assumes the vertex set is in Euclidean space, Hausdorff-close to some compact set. This hypothesis is made for the sake of convenience as it allows us to refer to the sampling theory for compact sets developped by Chazal and co-authors — see [4] for a survey — in the geometric part of our analysis. However, it does not reflect the variety of scenarios in which Rips zigzags can be used.

Since their construction only requires a matrix of pairwise distances as input, Rips zigzags are applicable in any metric space. Although they come with no theoretical guarantees in such a generality, there are many contexts in which something can be said about their persistence barcodes. Finite sampes of a compact subset of $\mathbb{R}^d$ is but one example. Another important example is when P is sampled from a compact Riemannian manifold, or more generally from a compact Alexandrov space with curvature bounded from above, or even more generally from any compact length space X with positive convexity radius — see e.g. [2] for an introduction to these spaces. It is beyond the scope of this paper to redo our analysis in this context, however for completeness we provide high-level directions on how to adapt it:

- Compact spaces with positive convexity radius admit finite covers with (small enough) convex metric balls, so the Nerve Lemma and its persistent variant hold. One can then reproduce the results of Section 2.2, and in particular Theorem 2.8, with X being the whole space instead of some compact subset, and with the convexity radius of X playing the role formerly played by the weak feature size[6]. This settles most of the geometric aspects of the analysis of Section 5.
- The one part of the analysis that still remains to be adapted is when the results by Attali, Lieutier and Salinas [1] are invoked to bound the amplitude of the noise in the barcode of the Morozov zigzag. These results require P to be sitting close to some compact set with positive μ-reach in $\mathbb{R}^d$. However, it turns out that they follow previous work by Hausmann [18] and Latschev [20], which focuses precisely on cases where P lies on Riemannian manifolds or more general length spaces, and whose results apply directly to the present context.
- Once these geometric aspects have been adapted, the rest of the analysis can be reproduced and theorems similar to the ones of Section 5 can be derived for the oscillating Rips, image Rips, Morozov, and discretized Morozov zigzags. While the lower bounds of the sweet

[6]The analysis is even simpler in this case since Čech complexes of suitable parameter carry the same homological information as X, as opposed to the images of inclusions between Čech complexes.

ranges still depend on the sampling density parameter ε, the upper bounds now depend on the convexity radius rather than on the weak feature size.

6.5 Theoretical comparison between Rips zigzags

The results of Section 5 induce the following classification of the various Rips zigzags introduced in the paper, in terms of the theoretical quality of the their output.

1. The iR-ZZ has the widest sweet range of all, and within this range the output persistence barcode has no topological noise at all.

2. The oR-ZZ has the same sweet range as the iR-ZZ, and the topological noise remains ephemeral within this range.

3. The dM-ZZ has a somewhat smaller sweet range, although the difference with the other ones is only by a constant factor. Within this range, the topological noise is not ephemeral, but it appears so if the persistence barcode is represented on the scale of the geometric scales.

4. Finally, the M-ZZ has the same sweet range as its discretized version, except that it is only proved that the signal persists throughout this range. Whether or not the noise can be bounded in full generality remains an open question at present. For now, what can be said is that under stronger assumptions on the input (finite sampling of a compact set with positive μ-reach), and within a smaller (*sweeter*) range, any topological noise disappears from the barcode of the M-ZZ.

In light of the discussion about interval representations held in Section 6.3, the reader should keep in mind that the transience of the noise within the sweet ranges of the oR-ZZ and dM-ZZ depends on the use of the closed interval representation of persistence barcodes introduced in [3]. Should the classical half-open interval representation be used instead, intervals corresponding to the noise would not be ephemeral, and their length on the logarithmic scale of geometric scales could be up to constant.

The above classification must be compared against the requirements of each zigzag in terms of computing resources: runtime and memory usage.

Memory usage. The relevant parameter here is the multiplier ρ, which conditions the size of the biggest complex in the zigzag. Assuming that the ordering of the points of P is by furthest point sampling[7], we have that every prefix P_i is an ε_i-sparse ε_i-sample of P (see e.g. [17, Lemma 4.1] for a proof), from which ensues the following guarantee.

Theorem 6.5. *Suppose P is sitting in some metric space of doubling dimension d. Then, for any $k \geq 0$, the number of k-simplices in the current complex at any time of the construction of the M-ZZ of parameter ρ is at most $2^{O(kd \log \rho)}n$, where n is the cardinality of P. The same bound applies to the oR-ZZ and iR-ZZ of parameter ρ, regardless of the value of parameter $\eta \leq \rho$. Finally, given a scale drop function ζ bounded from below by some quantity $\zeta_0 > 0$, the number of k-simplices in the current complex at any time of the construction of the dM-ZZ is at most $2^{O(kd \log \frac{\rho}{\zeta_0})}n$.*

[7] Arbitrary orderings may lead to local oversampling and thus to an uncontrolled local growth of the complex, regardless of the zigzag considered.

Proof. We prove the result in the case of the dM-ZZ, the other cases following by letting $\zeta_0 = 1$. Let $1 = n_1 < n_2 < \cdots < n_{r-1} < n_r = n$ be the discretization steps. Since each prefix P_{n_i} is ε_{n_i}-sparse, a standard ball packing argument shows that every point $p \in P_{n_i}$ is connected to at most $2^{O(d \log \rho)}$ neighbors in the Rips complex $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_i})$. These neighbors can form at most $\begin{pmatrix} 2^{O(d\log\rho)} \\ k \end{pmatrix} = 2^{O(kd\log\rho)}$ k-simplices with p. Thus, the total number of k-simplies in $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_i})$ is at most $2^{O(kd\log\rho)}n_i \leq 2^{O(kd\log\rho)}n$.

Let us now bound the size of $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}})$. It follows from the definition of n_{i+1} that $P_{n_{i+1}-1}$ is $\zeta_0\varepsilon_{n_i}$-sparse. Then, the same ball packing argument as above shows that every point $p \in P$ has at most $2^{O(d\log\frac{\rho}{\zeta_0})}$ points of $P_{n_{i+1}-1}$ within distance $\rho\varepsilon_{n_i}$. Applying this result to every $p \in P_{n_{i+1}}$, and observing that the set difference $P_{n_{i+1}} \setminus P_{n_{i+1}-1}$ consists only of the point $p_{n_{i+1}}$, we deduce that every vertex of $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}})$ has at most $2^{O(d\log\frac{\rho}{\zeta_0})}$ neighbors, as previously. The rest of the analysis follows. □

This result allows us to make the following observations:

- Since the theoretical lower bounds on ρ worked out in Section 5 are all constant ($\rho > 10$), one is allowed to set ρ to some constant value in practice and benefit from the guarantees on the quality of the output. Meanwhile, Theorem 6.5 ensures that the number of k-simplices in the current complex remains at most $2^{O(kd)}n$ throughout the construction of the M-ZZ, oR-ZZ, or iR-ZZ, where d is the doubling dimension of the ambient space. This bound is as good as the ones achieved with other lightweight structures for homology inference, like for instance the sparse Rips filtration [22].
- The theoretical lower bounds on ρ being the same for the M-ZZ, oR-ZZ and iR-ZZ, the same value of parameter ρ can be used in practice, resulting in the maximum complex sizes being equal. Note however that the full data structure in the case of the iR-ZZ may be twice as big due to the fact that it must maintain two Morozov zigzags: one of parameter ρ, the other of parameter $\eta \leq \rho$.
- When using the dM-ZZ, one can also set ζ to be a constant map equal to some constant value as in (11), thus benefiting from the theoretical guarantees on the quality of the output while maintaining the number of k-simplices in the current complex below $2^{O(kd)}n$ throughout the construction of the zigzag. Note however that the exact complex size is bigger than the one achieved with the other types of Rips zigzags when $\zeta < 1$.
- In cases where P is sitting Hausdorff-close to some lower-dimensional subspace X, Theorem 6.5 can be applied to bound the size of the current complex inside the sweet range corresponding to X. The bound thus obtained is $2^{O(kd')}n$ for the number of k-simplices, where $d' < d$ is the dimension of X. Hence, if P approximates a low-dimensional object in high dimensions, then the size of the data structure scales up only with the dimension of the object within the corresponding sweet range. For this reason, Rips zigzags are computed backwards in practice, starting at $i = n$ and ending at $i = 0$. This way, smaller scales are considered first, which makes the dimensionality of the data (and thus the size of the current complex) scale up reasonably until the sweet ranges of all the lower-dimensional structures underlying the data have been entirely spanned by the algorithm.

Focusing back on our initial classification of the Rips zigzags, we can conclude that the M-ZZ and oR-ZZ have the best performances in terms of memory usage. Then comes the iR-ZZ, which may require twice as big a data structure even though the maximum complex size is the same as for the other two. Finally, the dM-ZZ incurs an overhead in terms of memory usage, which is kept constant (depending exponentially on the doubling dimension though) when ζ is lower-bounded by a positive constant.

Runtime. Assuming again that the ordering of the points of P is by furthest point sampling, we have the following guarantee.

Theorem 6.6. *Suppose P is sitting in some metric space of doubling dimension d. Then, for any $k \geq 0$, the total number of k-simplices inserted in the current complex throughout the construction of the M-ZZ of parameter ρ is at most $2^{O(kd \log \rho)}n$, where n is the cardinality of P. The same bound applies to the iR-ZZ of parameter ρ, regardless of the value of parameter $\eta \leq \rho$. For the oR-ZZ of parameters η, ρ, the bound becomes $2^{O(kd \log \rho)}n^2$. Finally, given a scale drop function ζ bounded from below by some quantity $\zeta_0 > 0$, the bound for the dM-ZZ is $2^{O(kd \log \frac{\rho}{\zeta_0})}n$.*

Proof. We begin with the M-ZZ, for which we will use a simple charging argument. Observe that simplex insertions occur only when a forward arrow is encountered in the zigzag. For any such arrow, the current complex is enlarged by adding a new vertex p_{i+1} and connecting it to the rest of the complex. By the same packing argument as in the proof of Theorem 6.5, p_{i+1} forms at most $2^{O(d \log \rho)}$ edges with the points of P_i, therefore the number of k-simplices in its star in $\mathcal{R}_{\rho\varepsilon_i}(P_{i+1})$ is at most $2^{O(kd \log \rho)}$. This is also the number of k-simplices created at this stage of the algorithm. Hence, the total number of k-simplices inserted throughout the process is at most $2^{O(kd \log \rho)}n$. This bound applies also to the iR-ZZ, which maintains two Morozov zigzags: one of parameter ρ, the other of parameter $\eta \leq \rho$.

The case of the dM-ZZ is similar, with the additional twist that more than one vertex is added to the current complex when going from $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_i})$ to $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}})$. Nevertheless, as observed in the proof of Theorem 6.5, the points of $P_{n_{i+1}-1}$ are $\zeta_0\varepsilon_{n_i}$-sparse, so the number of edges in the star of any point of $P_{n_{i+1}}$ in $\mathcal{R}_{\rho\varepsilon_{n_i}}(P_{n_{i+1}})$ is at most $2^{O(d \log \frac{\rho}{\zeta_0})}$, and the number of k-simplices is bounded by $2^{O(kd \log \frac{\rho}{\zeta_0})}$. Hence, the total number of k-simplices inserted at this stage of the algorithm is at most $2^{O(d \log \frac{\rho}{\zeta_0})}(n_{i+1} - n_i)$. The result follows.

Finally, the case of the oR-ZZ with parameters $\eta < \rho$ is trickier. Due to the fact that both the vertex set and the Rips parameter increase when a forward arrow is encountered in the zigzag, we cannot simply charge the new simplex insertions to the newly added vertices: former vertices also form new simplices together. In fact our bound is obtained by a cruder argument: for any forward arrow, the current complex contains at most $2^{O(kd \log \rho)}n$ k-simplices before and after following the arrow. Hence, in the worst case the complex is rebuilt entirely, which means inserting at most $2^{O(kd \log \rho)}n$ k-simplices in total. Since this is true for any one of the $n-1$ forward arrows, the claimed quadratic bound follows. □

This result confirms the intuition stated at the beginning of Section 5 that the relevant parameter for runtime is the multiplier η: the closer it is to ρ, the fewer simplex insertions and deletions occur during the zigzag calculation. In this respect, the M-ZZ and iR-ZZ offer the best performances, with a slight advantage to the M-ZZ in practice due to the fact that it maintains only one zigzag (and not a pair of zigzags) and does not need to compute image persistence. Then comes the dM-ZZ, which makes a trade-off between enlarging the complexes and reducing the number of steps in the zigzag. As observed for the maximum complex size, the overhead incured is kept constant (depending exponentially on the doubling dimension though) if the scale drop function ζ is lower-bounded by a positive constant. Finally comes the oR-ZZ, for which the effect of reducing η is that many simplices are inserted then immediately deleted from the complex when building the zigzag, thus inducing an overhead in terms of computation time. The upper bound computed in Theorem 6.6 suggests an overhead up to linear in n, however the tightness of this bound needs to be assessed.

Wrap-up. To conclude our discussion on the comparison of the various Rips zigzags, let us mention that our bounds on space and time complexities suggest to use the Morozov zigzag

first when dealing with a new data set. This is especially true if the data are supposed to be sampled from "simple" shapes (i.e. ones with positive reach or μ-reach), in which case the quality of the output should be comparable to the one achieved with the other zigzags (albeit with smaller sweet ranges). Then, in cases where the quality of the result is not sufficient and one needs to obtain cleaner (less noisy) barcodes, one of the three aforementioned variants of the Morozov zigzag may be considered: discretized Morozov, oscillating Rips, or image Rips. Our theorems guarantee that the corresponding barcodes should exhibit sweet ranges with ephemeral noise or even no noise at all, a clear improvement in quality compared to the M-ZZ but also to the standard Rips filtration. Note however that our theoretical analysis does not designate one variant of the M-ZZ as being clearly better than the others, so in practice thorough experiments will be required to decide which one is the best in a given scenario.

7 Experiments

The Rips package of the C++ library Dionysus [13] provides efficient implementations of the M-ZZ and iR-ZZ. We built our implementations of the dM-ZZ and oR-ZZ around this package. We also slightly modified the code for the M-ZZ and iR-ZZ so that it outputs barcodes with closed intervals on a log base 2 scale. Our code and input data are available at `http://geometrica.saclay.inria.fr/data/Steve.Oudot/Rips-ZZ/Rips-ZZ.tgz`.

We have only conducted preliminary experiments with artificial data; nevertheless we observe some noteworthy phenomena. These call for further experimental validation, which we intend to carry out in the near future. For now, we will comment on some experiments using a representative example taken from [17]. This data set and its variants have been used several times in subsequent work, so it serves as a benchmark for us. It is composed of $10,000$ points spread out uniformly along a helicoidal curve drawn on a torus in 3-space. The underlying space varies with the geometric scale at which the data is considered. Among other structures, one hopes to find both the curve and the torus from the data. For reference, the major and minor radii of the torus are 4 and 1 respectively, so its diameter is 10 and its reach is 1. The helicoid has 100 periods in total, so its reach is roughly $3\frac{\pi}{100} \approx 0.094$.

In our experiments it was not possible to choose parameters that fully satisfy the hypotheses stated in our theorems. In particular, $\rho > 10$ was too large for the simplicial complexes to fit within the computer memory, whatever the choice of Rips zigzag among the ones of Section 5. The M-ZZ was run with parameter $\rho = 3$, the dM-ZZ with parameters $\rho = 3$ and $\zeta = \frac{1}{\sqrt{2}}$, the or-ZZ and iR-ZZ with parameters $\eta = 3$ and $\rho = 4$.

The obtained barcodes are shown in Figure 2 — from top to bottom: M-ZZ, dM-ZZ, oR-ZZ, iR-ZZ. The barcodes are plotted against the log base 2 of the geometric scale. The left column shows the genuine barcodes, the right column shows the same barcodes with all their intervals of length at most 10^{-2} removed. Each barcode has three parts, corresponding to the 0-th (top), 1-st (center) and 2-nd (bottom) homology generators. Note that every input point is inserted as a vertex in the zigzag at geometric scale 0, so the corresponding interval in the barcode starts at $-\infty$ on $\log_2$ scale. This interval needs to be cut before being plotted, and we chose arbitrarily to cut it at the stage where the first edge appears in the zigzag. This introduces an artifact in the 0-th homology part of the barcodes, where all the intervals but one appear as ephemeral whereas they actually extend to $-\infty$.

Commentary on the results. Since the geometric structures underlying the point cloud all have positive reach, one would expect from the discussion of Section 5.3.2 that sweeter ranges appear in the persistence barcode of the Morozov zigzag, one for each of these structures. This is indeed what happens in practice, where fairly long intervals without noise reveal the homology of the curve and that of the torus. Thus, the M-ZZ already does a pretty good job at revealing the homology of the spaces underlying the data set.

The main flaw in the barcode is that the boundaries of the sweeter ranges are plagued with a significant amount of noise that degrades the signal-over-noise (SNR) ratio. The worth of each of the variants of the Morozov zigzag can be measured by its cleaning effect on these parts of the barcode, or equivalently by the amount by which it increases the signal to noise ratio. Note that we cannot ask for the noise to be completely removed since the hypotheses of our theorems are not quite satisfied by the choice of parameters. Nevertheless, it seems reasonable to expect a significant noise reduction. In this respect, the dM-ZZ performs best. Its barcode shown in Figure 2 looks clean, and the noise appears as ephemeral. This first impression is confirmed by a quantitative analysis, reported in Table 1, which shows that the noise is indeed ephemeral so the SNR of the torus is infinte, both in 1-st and in 2-nd homology. The SNR of the curve has

also been increased compared to the M-ZZ.

zigzag	curve SNR (β_1)	transition SNR (β_1)	torus SNR (β_1 / β_2)
M-ZZ	1.78	N/A	3.77 / 18.5
dM-ZZ	2.26	N/A	∞ / ∞
oR-ZZ	1.86	220	10.60 / 1868
iR-ZZ	1.59	220	10.60 / 22.7

Table 1: *Signal-over-noise ratios obtained with the Rips zigzags of Section 5. For the curve, the SNR is measured by the ratio between the lengths of the first and second longest intervals in the β_1 barcode. For the torus, the SNR is measured by the ratio between the lengths of the second and third longest intervals in the β_1 barcode, as well as the first and second longest intervals in the β_2 barcode. For the transition, the SNR is measured by the ratio between the lengths of the third (counted with multiplicity) and fourth intervals in the β_1 barcode.*

According to Table 1, the second best variant of the M-ZZ in terms of SNR is the oR-ZZ, while the third place goes to the iR-ZZ. This trend is confirmed by the qualitative study of the barcodes of Figure 2, which reveals that the one of the iR-ZZ has no sweet range corresponding to the curve alone[8], an anomaly that makes the barcode of the iR-ZZ qualitatively worse than the one of the oR-ZZ. Compared to the dM-ZZ, these two barcodes do exhibit the transition phase that occurs between the curve and the torus, whereas the dM-ZZ misses it completely due to the discretization. This can be viewed either as a feature or as a drawback, depending on the context and on the type of structural information sought by the user.

These observations somewhat contradict the theoretical classification discussed in Section 6.5, however they are moderated by two facts: first, as mentioned previously, the hypotheses of our theorems are not quite satisfied by the choice of parameters; second, the worst-case bounds on the sweet(er) ranges given in the theorems of Section 5 are not relevant here since the data set admits underlying structures with fairly simple geometry (positive reach) and thus clearly does not correspond to a worst-case scenario.

Comparison with the standard Rips filtration. For completeness we show in Figure 3 the persistence barcode obtained using the standard Rips filtration. The latter had to be truncated at a parameter value of 0.5, since beyond that limit the total memory usage reached 6.5 GB and forced the computer to begin disk-swapping. For comparison, the M-ZZ with $\rho = 3$ used only a couple hundred MB, the oR-ZZ and iR-ZZ with $\rho = 4$ used barely more than 1 GB, and the dM-ZZ with $\rho = 3$ and $\zeta = \frac{1}{\sqrt{2}}$ used less than 1.5 GB.

Within the range $[0, 0.5]$ ($[-\infty, -1]$ on a $\log_2$ scale), the barcode of the standard Rips filtration exhibits the homology of the curve and of the torus as expected. However, the sweet range corresponding to the torus remains fairly small and comparable in length to the range showing the transition between the curve and the torus. This makes the SNR pretty poor, and it is fair to say that under such circumstances the presence of a torus underlying the data may remain questionable in the user's mind. Using a larger truncation value would definitely improve the SNR. Nevertheless, the barcode would remain pretty noisy in the range corresponding to the transition between the curve and the torus, and therefore much less clean than the ones obtained with the variants of the Morozov zigzag. This provides experimental evidence for our claim that Rips zigzags may produce cleaner results than the standard Rips filtration itself.

[8]We currently do not have a final explanation to this surprising fact. It may be due to the use of image persistence in a zigzag context, or simply to a bug in the code of the Dionysus Rips package. This question will be investigated further.

Another point to make is that, due to the truncation, only the spaces underlying the input data at small scales are visible in the barcode of the standard Rips filtration. For instance, the doughnut corresponding to the inside of the torus is not visible, whereas it clearly appears in the barcodes of Figure 2 right after the end of the sweet range of the torus, when the second 1-st homology generator and the 2-nd homology die, leaving only the main 1-st homology generator and the 0-th homology generator. This doughnut is one of the candidate underlying spaces for the point cloud, whose presence is captured by the barcodes of the Rips zigzags but not by the one of the Rips filtration because of memory constraints.

zigzag	parameters	add time (s)	remove time (s)	total time (s)
M-ZZ	$\rho = 3$	17	109	185
	$\rho = 4$	155	759	1277
dM-ZZ	$\rho = 3$, $\zeta = 1/\sqrt{2}$	53	305	535
	$\rho = 3$, $\zeta = 1/2$	214	2282	3430
oR-ZZ	$\eta = 3$, $\rho = 3.3$	8229	4167	16383
	$\eta = 3$, $\rho = 4$	58756	37983	121988
iR-ZZ	$\eta = 3$, $\rho = 3.3$	95	277	516
	$\eta = 3$, $\rho = 4$	734	13289	14481

Table 2: *Timings (in seconds) achieved for the curve-torus data set on an eight-core Intel Xeon CPU @2.40GHz — only one core was used at a time.*

zigzag	parameters	max complex size	# insertions
M-ZZ	$\rho = 3$	132932	899188
	$\rho = 4$	759880	5282607
dM-ZZ	$\rho = 3$, $\zeta = 1/\sqrt{2}$	1001168	3324320
	$\rho = 3$, $\zeta = 1/2$	8177204	17206098
oR-ZZ	$\eta = 3$, $\rho = 3.3$	249016	275054345
	$\eta = 3$, $\rho = 4$	759880	1678134869
iR-ZZ	$\eta = 3$, $\rho = 3.3$	249016	2775503
	$\eta = 3$, $\rho = 4$	759880	8965889

Table 3: *Maximum complex size (in number of simplices) and total number of inserted simplices on the curve-torus data set.*

Computing resources consumption. Table 2 shows the timings achieved by the four types of Rips zigzags on the curve-torus data. For each zigzag type we reported the runtime with two different sets of parameters, including the one used to produce the results of Figure 2. As expected, the basic Morozov zigzag is the most efficient, but the striking fact here is that the runtime of the oR-ZZ is slower than the ones of the other zigzags by an order of magnitude or even more. Actually, this is not so surprising in light of Theorem 6.6, which claims a quadratic runtime in the number of data points for the oR-ZZ, versus only a linear runtime for the other zigzags. This trend is confirmed in Table 3, where the total number of inserted (and deleted) simplices in the oR-ZZ is orders of magnitude higher than the one achieved with the other zigzags. The interpretation is straightforward: to go from the geometric scale ε_i to ε_{i-1}, the oR-ZZ increases the Rips parameter from $\eta\varepsilon_i$ to $\rho\varepsilon_{i-1}$ and then decreases it immediately to $\eta\varepsilon_{i-1}$. Hence, most of the time is spent inserting simplices and removing them immediately from the current complex.

Concerning the memory usage, Table 3 corroborates the claim made in Theorem 6.5 that the M-ZZ, oR-ZZ and iR-ZZ are roughly equivalent, while dM-ZZ uses more memory. In practice the difference is an order of magnitude, which can make the use of thd dM-ZZ problematic in cases where most of the main memory is already in use with the basic Morozov zigzag.

8 Conclusion

In this paper, we explored several Rips-like zigzags that achieve both small size and bounded noise for homology inference. The proofs relied on general new techniques for manipulating and comparing zigzag modules. We hope that these techniques will find further use and stimulate new research and applications of zigzag persistence.

Acknowledgements

The authors wish to thank Marc Glisse for suggesting the approach used in the proof of Theorem 5.3. They also thank Frédéric Chazal for helpful discussions. This work was supported by the European project CG-Learning No. 255827.

References

[1] Dominique Attali, André Lieutier, and David Salinas. Vietoris-Rips complexes also provide topologically correct reconstructions of sampled shapes. *Computational Geometry: Theory and Applications*, 2012. Accepted (short version appeared in *Proc. SoCG'11*).

[2] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[3] Gunnar Carlsson and Vin de Silva. Zigzag persistence. *Foundations of Computational Mathematics*, 10(4):367–405, 2010.

[4] Frédéric Chazal and David Cohen-Steiner. Geometric inference. In *Tesselations in the Sciences*. Springer-Verlag, 2012. To appear.

[5] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J. Guibas, and Steve Y. Oudot. Proximity of persistence modules and their diagrams. In *Proc. 25th Annu. Symposium on Computational Geometry*, pages 237–246, 2009.

[6] Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in Euclidean space. *Discrete & Computational Geometry*, 41:461–479, 2009.

[7] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Y. Oudot. The structure and stability of persistence modules. Research Report arXiv:1207.3674 [math.AT], July 2012.

[8] Frédéric Chazal, Leonidas J. Guibas, Steve Y. Oudot, and Primoz Skraba. Analysis of scalar fields over point cloud data. *Discrete and Computational Geometry*, 46(4):743–775, December 2011.

[9] Frédéric Chazal and Steve Y. Oudot. Towards persistence-based reconstruction in Euclidean spaces. In *Proceedings of the 24th ACM Symposium on Computational Geometry*, pages 232–241, 2008.

[10] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, January 2007.

[11] Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic and Geometric Topology*, 7:339–358, 2007.

[12] Tamal K. Dey, Fengtao Fan, and Yusu Wang. Computing topological persistence for simplicial maps. Research Report arXiv:1208.5018 [cs.CG], August 2012.

[13] Dionysus. Dmitriy Morozov (`http://www.mrzv.org/software/dionysus/`).

[14] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002.

[15] Peter Gabriel. Unzerlegbare Darstellungen I. *Manuscripta Mathematica*, 6:71–103, 1972.

[16] Marc Glisse. Personal communication.

[17] L. J. Guibas and S. Y. Oudot. Reconstruction using witness complexes. *Discrete and Computational Geometry*, 40(3):325–356, 2008.

[18] Jean-Claude Hausmann. On the Vietoris-Rips complexes and a cohomology theory for metric spaces. *Prospects in topology, Ann. of Math. Stud.*, 138:175–188, 1995.

[19] Benoit Hudson, Gary L. Miller, Steve Y. Oudot, and Donald R. Sheehy. Topological inference via meshing. In *Proceedings of the 2010 annual symposium on Computational geometry*, SoCG '10, pages 277–286, New York, NY, USA, 2010. ACM.

[20] Janko Latschev. Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6):522–528, 2001.

[21] Dmitriy Morozov. Personal communication.

[22] Donald R. Sheehy. Linear-size approximations to the vietoris-rips filtration. In *Proceedings of the 2012 symposuim on Computational Geometry*, SoCG '12, pages 239–248, New York, NY, USA, 2012. ACM.

[23] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete Comput. Geom.*, 33(2):249–274, 2005.

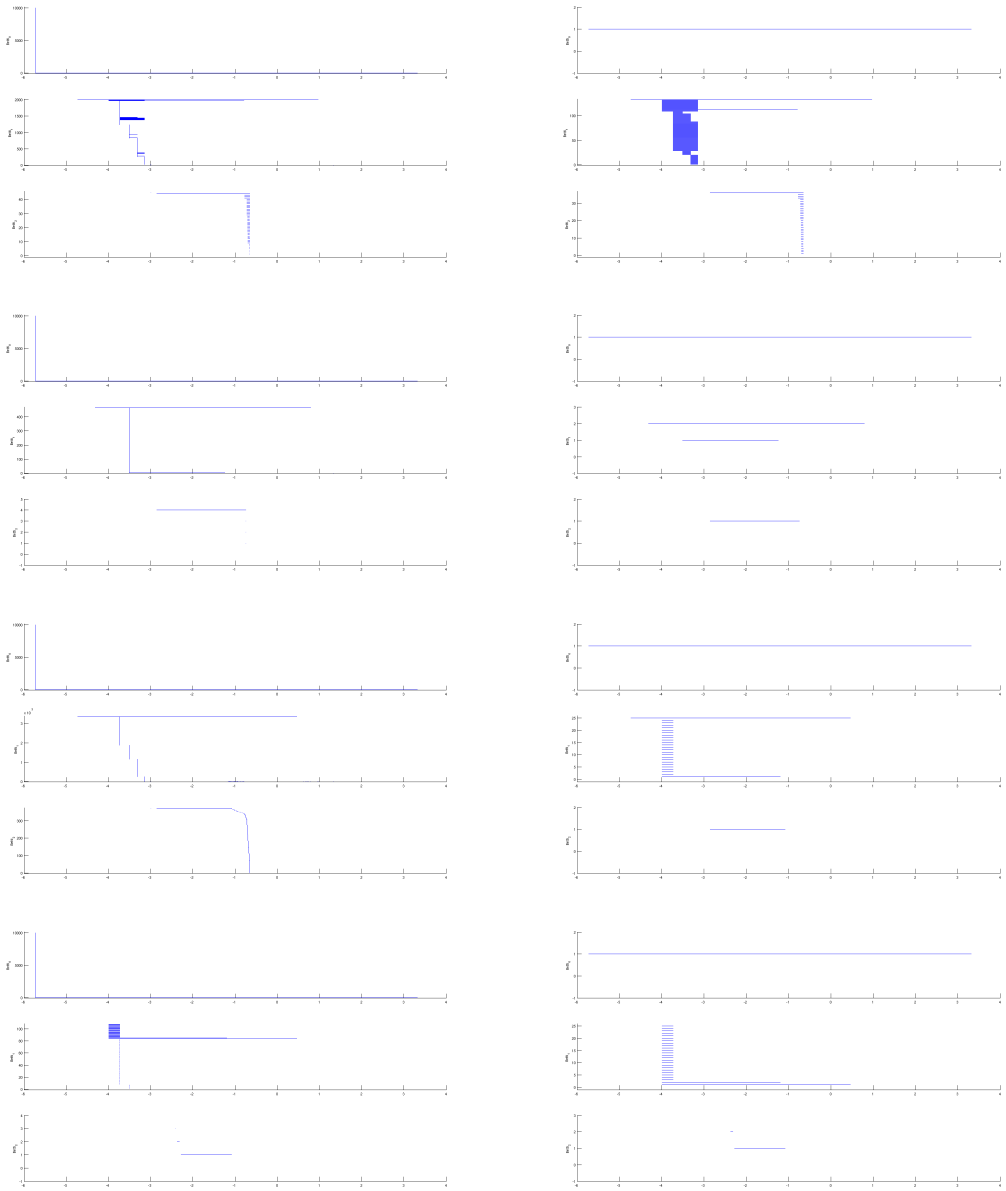


Figure 2: *Persistence barcodes computed by the Rips zigzags of Section 5 on the curve-torus data set, and plotted by the* px_homology *function of the PLEX 2.5 library. Intervals are plotted against the log base* 2 *of the geometric scale. The left column shows the genuine persistence barcodes, the right column shows the same barcodes after removing the intervals of length at most* 10^{-2}*. From top to bottom: M-ZZ, dM-ZZ, oR-ZZ, iR-ZZ.*

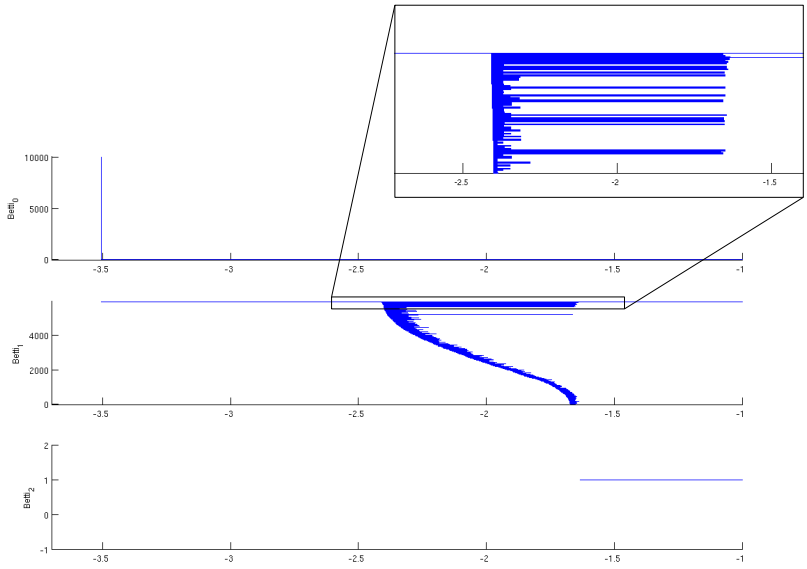
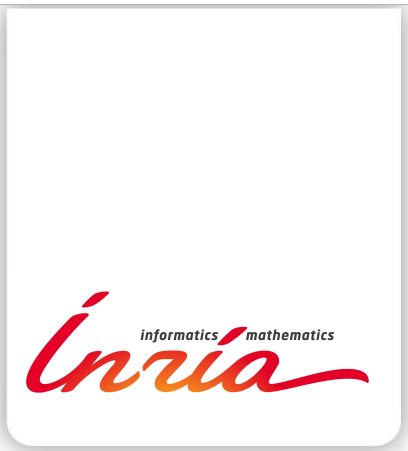


Figure 3: *Persistence barcode obtained from the standard Rips filtration truncated at parameter value* 0.5. *The barcode is plotted on a logarithmic scale. Intervals of length at most* 10^{-2} *have been removed, and intervals in* 0*-th homology that start at* $-\infty$ *have been cut to start at the stage where the first edge appears in the filtration.*



RESEARCH CENTRE
SACLAY – ÎLE-DE-FRANCE

Parc Orsay Université
4 rue Jacques Monod
91893 Orsay Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399