On Computing Levelset Zigzag Homology

Nicolas Berkouk and Luca Nyckees

Abstract

Topological data analysis is a branch in the field of applied mathematics that has rapidly grown in the past few years. It offers toolboxes to study the shape of data and finds applications in various domains of machine learning [6]. Key concepts in topological data analysis are the notion of *persistent homology* [3] and its extended versions, namely *extended persistent homology* and *zigzag homology* [4], which offer a way of analysing the behavior of topological features along a diagram of spaces built from data. A central topic in topological data analysis is the computation of the corresponding persistence diagrams. We provide an efficient and visual algorithm aiming at computing zigzag persistence diagrams. The method we use is based on a diagram bijection between levelset zigzag persistence and extended persistence, introduced in [5]. A categorical setting for describing the bijection is described in [1].

Laboratory for Topology and Neuroscience
Spring semester of 2021

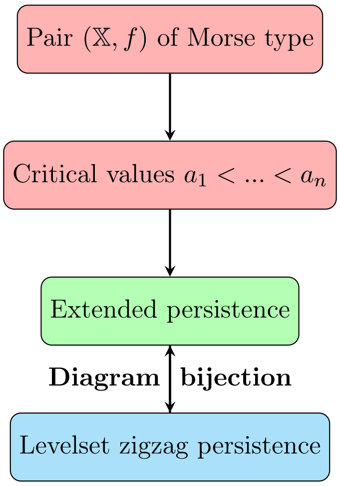

Figure 1: General pipeline illustration.

1 Introduction

A central task in applied algebraic topology, that is the key to various applications in topological data analysis, is the computation of persistence diagrams of a sequence of homology groups. We implement an efficient algorithm designed to compute levelset zigzag persistence diagrams of a sequence of homology groups. The method is based on the bijection between levelset zigzag persistence and extended persistence. The idea is to deduce results from the computation of extended persistence, which is already efficiently implemented in C++. The bijection mentioned above is based on the so-called *pyramid principle*. The mathematical objects used to introduce the pyramid principle and the diagram bijection were developed by two schools of thought, and a part of this work aims at establishing a clear connection between them. In Section 2, we introduce the concept of zigzag persistent homology, and refer to [4] for a complete introduction to the subject. In Section 3, we describe the basic setting used to understand the bijection and introduce the functioning of the pyramid principle by following the ideas of [5]. In Section 4, we introduce a categorical setting used to understand the same bijection with a more abstract viewpoint (we take inspiration from [1]). In Section 5, we make a connection between the two settings. We provide the algorithm implementation together with visual results in Section 6.

Notation. If unspecified, a vector space is a K-vector space where K is a fixed field.

2 Zigzag Persistent Homology

Zigzag persistent homology aims at generalizing the setting of persistent homology, by considering a class of diagrams that encapsulate the case of persistence modules.

Definition 2.1 (Zigzag module)**.** *Consider a diagram of vector spaces of the form*

$$V_0 \longleftrightarrow V_1 \longleftrightarrow V_2 \longleftrightarrow \cdots \longleftrightarrow V_n,$$

where each double-sided arrow $\longleftrightarrow$ is either a left arrow $\longleftarrow$ or a right arrow $\longrightarrow$, that is a linear map of vector spaces. We call such a diagram of vector spaces a zigzag module. The type τ *of a zigzag module represents the orientation of its arrows and is denoted by a sequence of f and g, respectively denoting forward and backward maps.*

Example 2.2 (Type of a module)**.** *The zigzag module below has type $\tau = gfg$.*

$$V_1 \longleftarrow V_2 \longrightarrow V_3 \longleftarrow V_4$$

Here, we study a specific type of zigzag modules formed by applying a homology functor $H_p(\cdot)$ to a diagram of embeddings between simplicial complexes of the form

$$\mathbb{X}_0 \longrightarrow \mathbb{X}_1 \longleftarrow \mathbb{X}_2 \longrightarrow \cdots \longleftarrow \mathbb{X}_n.$$

This yields, for a fixed $p \in \mathbb{N}$, the following zigzag module.

$$H_p(\mathbb{X}_0) \longrightarrow H_p(\mathbb{X}_1) \longleftarrow H_p(\mathbb{X}_2) \longrightarrow \cdots \longleftarrow H_p(\mathbb{X}_n)$$

Given a collection of simplicial complexes, there is a natural way to obtain such a zigzag module, as desribed in the next definition.

Definition 2.3 (Natural p-zigzag module)**.** *Let $\mathcal{X} = \{X_i\}_{i=0}^n$ be a collection of simplicial complexes. For a given $p \in \mathbb{N}$, we define the* natural p-zigzag module associated to $\mathcal{X}$ *to be the module formed by applying the functor $H_p(\cdot)$ to the following diagram.*

$$\begin{tikzcd}
& X_0 \cup X_1 & & X_1 \cup X_2 & & \dots & \\
X_0 \arrow[ur] & & X_1 \arrow[ul] \arrow[ur] & & X_2 \arrow[ul] \arrow[ur] & & X_n \arrow[ul]
\end{tikzcd}$$

Definition 2.4 (Remak decomposition)**.** *A* Remak decomposition *of a zigzag module $\mathbb{V}$ is a decomposition of the form $\mathbb{V} = \mathbb{W}_1 \oplus \cdots \oplus \mathbb{W}_n$, where all $\mathbb{W}_i$ are indecomposable.*

Remark 2.5. *Any zigzag module admits a Remak decomposition.* ♠

Definition 2.6 (Interval module)**.** *One can define the* type-τ interval module $\mathbb{I}_\tau(b,d)$ *as the only τ-module formed by the spaces*

$$I_i^\tau(b,d) = \begin{cases} \mathbb{F} & \text{if } b \leq i \leq d \\ 0 & \text{otherwise,} \end{cases}$$

and equipped with identity maps between two adjacent copies of and zero maps elsewhere.

Theorem 2.7. *Any zigzag module $\mathbb{V}$ admits a Remak decomposition of the form*

$$\mathbb{V} = \mathbb{I}_\tau(b_1,d_1) \oplus \cdots \oplus \mathbb{I}_\tau(b_N,d_N).$$

Remark 2.8. *By the Krull-Schmidt principle (see Proposition 2.2 in [4]), the Remak decomposition into interval modules of a given zigzag module $\mathbb{V}$ is an isomorphism invariant of $\mathbb{V}$. This allows us to formulate the next definition.* ♠

Definition 2.9 (Zigzag persistence)**.** *The persistence of a zigzag module $\mathbb{V}$ is defined as the multiset* $\mathrm{Pers}(\mathbb{V}) = \{[b_i,d_i] | i = 1,...,N\}$ *induced from a Remak decomposition*

$$\mathbb{V} = \mathbb{I}_\tau(b_1,d_1) \oplus \cdots \oplus \mathbb{I}_\tau(b_N,d_N).$$

3 The Pyramid Principle

In this section, we introduce the *pyramid principle* and express the diagram bijection between the levelset zigzag persistence and the extended persistence of a pair $(\mathbb{X}, f)$ of Morse type having finitely generated levelset homology. We mainly follow [5].

3.1 Background and notations

For the rest of this section, we consider a generic space $\mathbb{X}$ with real-valued continuous function $f : \mathbb{X} \to \mathbb{R}$, such that the pair $(\mathbb{X}, f)$ is of Morse type (*i.e.* has finitely many critical values) $a_1 < ... < a_n$. We also assume that homology groups of the form $H_p(f^{-1}((-\infty,t)))$ have finite rank for any $p \in \mathbb{N}$ and any $t \in \mathbb{R}$. Now, we choose values $s_0, ..., s_n \in \mathbb{R}$ that satisfy

$$-\infty < s_0 < a_1 < s_1 < \cdots < s_{n-1} < a_n < s_n < \infty,$$

and denote by $\mathbb{X}_i^j$ the space $f^{-1}([s_i,s_j])$ for any pair of indices $i \leq j$. We define the *levelset zigzag persistence* of the pair $(\mathbb{X}, f)$ to be the zigzag persistence of the diagram

$$\mathbb{X}_0^0 \to \mathbb{X}_0^1 \leftarrow \mathbb{X}_1^1 \to \cdots \leftarrow \mathbb{X}_{n-1}^{n-1} \to \mathbb{X}_{n-1}^n \leftarrow \mathbb{X}_n^n.$$

Definition 3.1 (Pyramid)**.** *We define the* pyramid *associated to* $(\mathbb{X}, f)$ *as the diagram of spaces below (drawn in Figure 2 for the case $n = 3$). The southern sequence corresponds to levelset zigzag persistence, while the left-to-right upward diagonal sequence corresponds to extended persistence.*

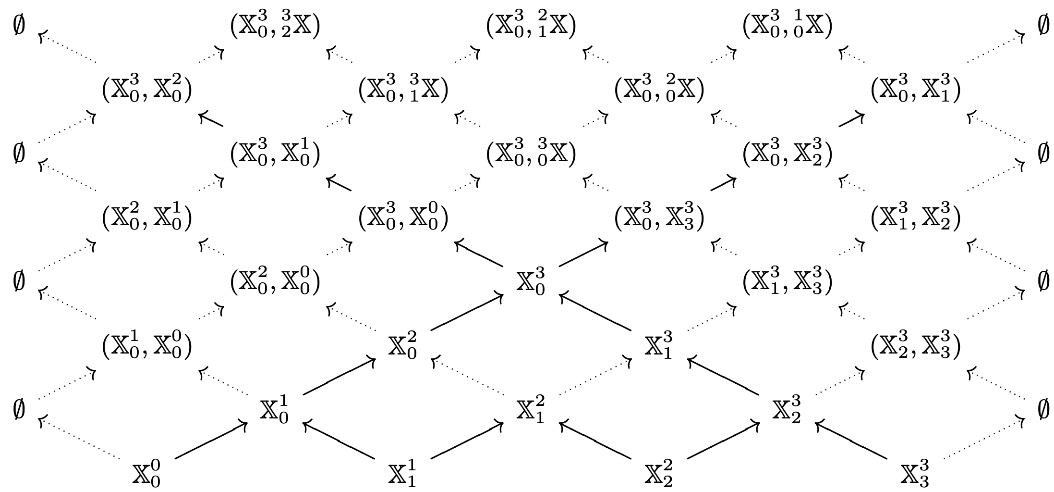

Figure 2: Pyramid for the case $n = 3$, taken from [5]. We set ${}_i^j\mathbb{X} := \mathbb{X}_0^i \cup \mathbb{X}_j^n$.

Remark 3.2. *The diagram of spaces defined above can be seen as a* pyramid from above *by thinking of it as a squared-basis pyramid with the space $\mathbb{X}_0^n$ as its top.* ♠

Remark 3.3. *Note that each diamond appearing in the pyramid is a relative Mayer-Vietoris diamond in the sense that it can be expressed as a square of the form*

$$\begin{tikzcd} & (\mathbb{A} \cup \mathbb{C}, \mathbb{B} \cup \mathbb{D}) & \\ (\mathbb{A}, \mathbb{B}) \arrow[ur] & & (\mathbb{C}, \mathbb{D}) \arrow[ul] \\ & (\mathbb{A} \cap \mathbb{C}, \mathbb{B} \cap \mathbb{D}) \arrow[ul] \arrow[ur] & \end{tikzcd}$$

♠

This property enables us to effectuate so-called *diamond moves*, that express a bijection between the persistence intervals of two general zigzag diagrams of spaces that differ by exactly one Mayer-Vietoris diamond. Like so, one can go, for example, from extended persistence to levelset zigzag persistence and vice versa via a sequence of bijections (diamond moves) between intermediary zigzag diagrams.

Theorem 3.4 (Pyramid Principle)**.** *There is an explicit bijection between the extended persistence and the levelset zigzag persistence of the pair $(\mathbb{X}, f)$, that respects homological dimension except for possible shifts of degree $d \in \{-1, 1\}$.*

Proof. A proof of a more general result can be found in [5] (see *Pyramid Theorem*). ■

Example 3.5 (Pyramidal Transformation)**.** *Let $(\mathbb{X}, f)$ be a pair of Morse type such that the interval $[\mathbb{X}_1^1, \mathbb{X}_2^2]$ appears in its levelset zigzag barcode. We denote the levelset zigzag sequence of $(\mathbb{X}, f)$ by $\mathrm{LZZ}(f)$, and its up-down sequence by $\mathrm{UD}(f)$. Going from one sequence to the other via pyramidal transformation consists in making three diamond moves.*

$$\begin{tikzcd}
\mathbb{X}_0^0 \arrow[r] & \mathbb{X}_0^1 & \mathbb{X}_1^1 \arrow[l] \arrow[r] & \mathbb{X}_1^2 \arrow[d, Rightarrow, "\textit{Step (1)}"] & \mathbb{X}_2^2 \arrow[l] \arrow[r] & \mathbb{X}_2^3 & \mathbb{X}_3^3 \arrow[l] & \mathrm{LZZ}(f) \\
\mathbb{X}_0^0 \arrow[r] & \mathbb{X}_0^1 \arrow[r] & \mathbb{X}_0^2 & \mathbb{X}_1^2 \arrow[l] \arrow[d, Rightarrow, "\textit{Step (2)}"] & \mathbb{X}_2^2 \arrow[l] \arrow[r] & \mathbb{X}_2^3 & \mathbb{X}_3^3 \arrow[l] & \\
\mathbb{X}_0^0 \arrow[r] & \mathbb{X}_0^1 \arrow[r] & \mathbb{X}_0^2 & \mathbb{X}_1^2 \arrow[l] \arrow[r] \arrow[d, Rightarrow, "\textit{Step (3)}"] & \mathbb{X}_1^3 & \mathbb{X}_2^3 \arrow[l] & \mathbb{X}_3^3 \arrow[l] & \\
\mathbb{X}_0^0 \arrow[r] & \mathbb{X}_0^1 \arrow[r] & \mathbb{X}_0^2 \arrow[r] & \mathbb{X}_0^3 & \mathbb{X}_1^3 \arrow[l] & \mathbb{X}_2^3 \arrow[l] & \mathbb{X}_3^3 \arrow[l] & \mathrm{UD}(f)
\end{tikzcd}$$

The steps indicated in the diagram above involve the following relative Mayer-Vietoris diamonds.

$$\begin{tikzcd}
& \mathbb{X}_0^2 & \\
\mathbb{X}_0^1 \arrow[ur] & (1) & \mathbb{X}_1^2 \arrow[ul] \\
& \mathbb{X}_1^1 \arrow[ul] \arrow[ur] &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \mathbb{X}_1^3 & \\
\mathbb{X}_1^2 \arrow[ur] & (2) & \mathbb{X}_2^3 \arrow[ul] \\
& \mathbb{X}_2^2 \arrow[ul] \arrow[ur] &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \mathbb{X}_0^3 & \\
\mathbb{X}_0^2 \arrow[ur] & (3) & \mathbb{X}_1^3 \arrow[ul] \\
& \mathbb{X}_1^2 \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

The interval $[\mathbb{X}_1^1, \mathbb{X}_2^2]$ is transformed according to the bijection expressed in the Pyramid theorem. We obtain the transformation below, whose step-wise birth and death travels are given in Figure 3. For example, during the first step, the birth coordinate travels from space $\mathbb{X}_1^1$ to space $\mathbb{X}_0^2$, while the death coordinate stays at space $\mathbb{X}_2^2$.

$$[\mathbb{X}_1^1, \mathbb{X}_2^2] \xmapsto{(1)} [\mathbb{X}_0^2, \mathbb{X}_2^2] \xmapsto{(2)} [\mathbb{X}_0^2, \mathbb{X}_1^3] \xmapsto{(3)} [\mathbb{X}_0^2, \mathbb{X}_1^3]$$

One can summarize the transformation as a direct sum of two applications, as follows. We stick to the case $n = 3$. Let $\alpha = \alpha_3 \circ \alpha_2 \circ \alpha_1$ and $\beta = \beta_3 \circ \beta_2 \circ \beta_1$ denote the travel of the birth and death coordinates respectively (so that α_i and β_i denote the travel of birth and death at step (i), respectively). Then the interval transformation

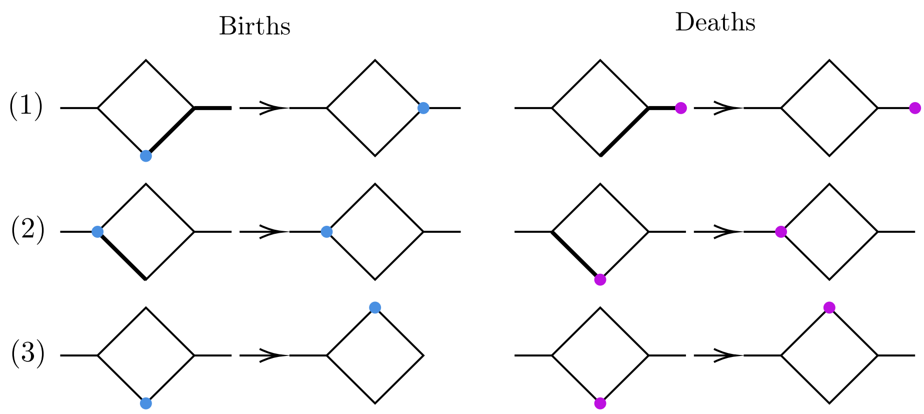

Figure 3: Birth and death travels along three consecutive diamond moves.

$$\begin{tikzcd}
\mathbb{X}_1^1 \arrow[r] \arrow[d] \arrow[rrr, bend left, "\alpha"] & \mathbb{X}_0^2 \arrow[r] \arrow[d] & \mathbb{X}_0^2 \arrow[r] \arrow[d] & \mathbb{X}_0^2 \arrow[d] \\
\mathbb{X}_1^1 \oplus \mathbb{X}_2^2 \arrow[r] & \mathbb{X}_0^2 \oplus \mathbb{X}_2^2 \arrow[r] & \mathbb{X}_0^2 \oplus \mathbb{X}_1^3 \arrow[r] & \mathbb{X}_0^2 \oplus \mathbb{X}_1^3 \\
\mathbb{X}_2^2 \arrow[r] \arrow[u] \arrow[rrr, bend right, "\beta"'] & \mathbb{X}_2^2 \arrow[r] \arrow[u] & \mathbb{X}_1^3 \arrow[r] \arrow[u] & \mathbb{X}_1^3 \arrow[u]
\end{tikzcd}$$

Figure 4: Decomposition of pyramidal transformation as a direct sum of birth and death travel, in the case of Example 3.5 ($n = 3$, LZZ(f) to UP(f), with $[\mathbb{X}_1^1, \mathbb{X}_2^2]$).

can be written as $\varphi := \alpha \oplus \beta$, if we consider intervals as direct sums, *i.e.* by looking at direct sums $\mathbb{A} \oplus \mathbb{B} \cong [\mathbb{A}, \mathbb{B}]$. The particular case presented in Example 3.5 can then be represented as in Figure 3.1, where φ is the only application making the diagram commute.

From the pyramid principle, one can deduce the following result, which described an explicit bijection between levelset zigzag persistence and extended persistence. The proof relies on case-by-case investigation, as in Example 3.5. Here, we naturally classify intervals appearing in the context of the extended persistence into four types as follows. We recall that extended persistence is a multiset of intervals deduced from the life and death of representatives in a sequence of absolute and relative homology groups of the form

$$H_p(\mathbb{X}_0^0) \to \cdots \to H_p(\mathbb{X}_0^n) \to H_p(\mathbb{X}_0^n, \mathbb{X}_n^n) \to \cdots \to H_p(\mathbb{X}_0^n, \mathbb{X}_0^n),$$

for some dimension $p \in \mathbb{N}$. Now, *Type I* (resp. *Type II*) stands for intervals whose lifespan is contained within the first (resp. second) part of the sequence. *Type III* (resp. *Type IV*) stands for intervals whose birth appears in the first part of the sequence (resp. death) and whose death (resp. birth) appears in the second one.

Theorem 3.6 (Diagram Bijection)**.** *One has the following correspondence between the intervals of extended persistence (left) and levelset zigzag persistence (right).*

Type	*Extended*	*Levelset zigzag*
I ($i < j$)	$[\mathbb{X}_0^i, \mathbb{X}_0^{j-1}]$	$[\mathbb{X}_{i-1}^i, \mathbb{X}_{j-1}^{j-1}]$
II ($i < j$)	$[(\mathbb{X}_0^n, \mathbb{X}_{j-1}^n), (\mathbb{X}_0^n, \mathbb{X}_i^n)]^+$	$[\mathbb{X}_i^i, \mathbb{X}_{j-1}^j]$
III ($i \leq j$)	$[\mathbb{X}_0^i, (\mathbb{X}_0^n, \mathbb{X}_j^n)]$	$[\mathbb{X}_{i-1}^i, \mathbb{X}_{j-1}^j]$
IV ($i < j$)	$[\mathbb{X}_0^j, (\mathbb{X}_0^n, \mathbb{X}_i^n)]^+$	$[\mathbb{X}_i^i, \mathbb{X}_{j-1}^{j-1}]$

Proof. This immediately follows from the proof of Theorem 3.4. ∎

4 A Categorical Setting

The authors of [1] define the pyramid concept, as well as way of expressing levelset zigzag persistence depending on extended persistence, in a more abstract setting - with a categorical background - that we introduce now. We aim at expressing the bijection between the two kinds of persistence in the presented setting.

Notation. We consider the inversed plane $P := \mathbb{R} \times \mathbb{R}^\circ$, where we have the posets $\mathbb{R} = (\mathbb{R}, \leq)$ and $\mathbb{R}^\circ = (\mathbb{R}, \geq)$. We obtain a poset relation on P by $(a, b) \preceq (c, d)$ if and only if $a \leq c$ and $b \geq d$. Moreover, given a point $m \in \mathbb{M}$, we define the sets

$$\uparrow(m) := \{u \in \mathbb{M} | m \preceq u\} \text{ and } \downarrow(m) := \{u \in \mathbb{M} | u \preceq m\}.$$

For a subset $S \subset \mathbb{M}$, we define in a similar way the sets

$$\uparrow(S) := \{u \in \mathbb{M} | m \preceq u \forall m \in S\} \text{ and } \uparrow(S) := \{u \in \mathbb{M} | u \preceq m \forall m \in S\}.$$

Definition 4.1 (Big Strip)**.** *We define a subspace* $\mathbb{M} \subset P$ *as the convex hull formed by the lines* $l_1 := \{(x, y) \in P | y = 1 - x\}$ *and* $l_2 := \{(x, y) \in P | y = -1 - x\}$*. We call* $\mathbb{M}$ *the* big strip.

We write the extended real line as $\bar{R} := [-\infty, \infty]$. Let $\blacktriangle : \bar{R} \to \mathbb{M}$ be an embedding such that the injected copy of $\bar{R}$ is orthogonal to l_1 and goes through the origin of

P. Furthermore, we write $\bigstar = \mathrm{Im}_{\blacktriangle}(\bar{R})$ for the injected copy of $\bar{R}$.

We let $T \in \mathrm{End}(\mathbb{M})$ be the map defined as follows. Let h_1 (resp. v_1) be the horizontal (resp. vertical) line passing through m. Let a be the intersection of l_2 and v_1. Let h_2 be the horizontal line passing through a, and let b be the intersection of l_1 and h_1. Finally, let v_2 be the vertical line passing through b. For $m \in \mathbb{M}$, we define $T(m)$ as the intersection point of v_2 and h_2.

Definition 4.2 (Fundamental Domain)**.** *We define the* fundamental domain *of* $\mathbb{M}$ *as*

$$D := \downarrow(\bigstar) \setminus T^{-1}(\downarrow(\bigstar)).$$

Remark 4.3. *We have an action* $\cdot : \mathbb{Z} \cong \langle T \rangle \circlearrowleft \mathbb{M}$ *defined by* $T \cdot m = T(m)$ *for all* $m \in \mathbb{M}$*. This action induces a tessellation of the big strip* $\mathbb{M}$ *by seing it as the orbit of* D*.* ♠

Definition 4.4. *We define the map*

$$\rho : \begin{cases} \mathbb{M} \to \mathcal{P} := \mathcal{T}_{\mathbb{R}} \times \mathcal{T}_{\mathbb{R}} \\ m \mapsto \left(\blacktriangle^{-1}(\mathrm{int}(\downarrow T(m))), \blacktriangle^{-1}(\mathbb{M} \setminus \uparrow m)\right), \end{cases}$$

where $\mathcal{T}_{\mathbb{R}}$ *denotes the standard topology on* $\mathbb{R}$*.*

Let $(\mathbb{X}, f)$ be a pair of Morse type. We can obtain a mapping F_f that assigns to each point of $\downarrow(\bigstar)$ a homology group by setting

$$F_f(m) = H_p(f^{-1}(\rho(m)))$$

for all $m \in \downarrow(\bigstar)$, where $p \in \mathbb{N}$ is the only integer such that $T^p(m) \in D$.

Notation. For a category $\mathcal{C}$, we let $\mathcal{C}^\circ$ denote the opposite category, with the same objects and arrows reversed. Let Vect_K denote the category of finite dimensional K-vector spaces with linear maps as morphisms, and let $\mathrm{Vect}_K^{\mathbb{Z}}$ denote the category of $\mathbb{Z}$-graded vector spaces over K.

Definition 4.5 (Operator $(\cdot)^{\mathbb{Z}}$)**.** *For a category* $\mathcal{C}$*, we define* $\mathcal{C}^{\mathbb{Z}}$ *as the category having for objects maps* $M^\bullet : \mathbb{Z} \to \mathcal{C} : n \mapsto M^n$*. Morphisms of* $\mathcal{C}^{\mathbb{Z}}$ *are naturally defined pointwise.*

Definition 4.6 (Evaluation)**.** *Let* $\mathcal{C}$ *be a category. We define the evaluation functor*

$$\mathrm{ev}^0 : \begin{cases} \mathcal{C}^{\mathbb{Z}} \to \mathcal{C} \\ M^\bullet \mapsto M^0 \end{cases}.$$

Next, we define the notion of *extended persistence diagram* associated to a pair $(\mathbb{X}, f)$ of Morse type, in the new setting. But first, we need to define the RISC functor associated to this type of pair, as well as the persistence diagram of a contravariant functor $F : \mathbb{M}^\circ \to \mathrm{Vect}_K$. We may consider the strip $\mathbb{M}$ as a category whose objects are the points on $\mathbb{M}$ and whose arrows $\to$ are given by the relations $\preceq$. Let $\mathrm{int}(\mathbb{M})$ denote the interior of the strip $\mathbb{M}$ and $\partial\mathbb{M}$ its boundary. Finally, we consider a general cohomology theory $\mathcal{H}^\bullet$ that takes values in the category $\mathrm{Vect}_K^{\mathbb{Z}}$, sending weak equivalences to isomorphisms.

Let $\Sigma : (\mathrm{Vect}_K^{\mathbb{Z}})^\circ \to (\mathrm{Vect}_K^{\mathbb{Z}})^\circ$ be the degree-shift endofunctor acting as $\Sigma(M^\bullet) = M^{\bullet-1}$. One can observe that $F' : D \to \left(\mathrm{Vect}_K^{\mathbb{Z}}\right)^\circ : m \mapsto \mathcal{H}^\bullet(f^{-1}(\rho(m)))$ is a contravariant functor that may be extended to a functor $F : \mathbb{M} :\to \left(\mathrm{Vect}_K^{\mathbb{Z}}\right)^\circ$ such that the square

$$\begin{tikzcd} \mathbb{M} \arrow[r, "F"] \arrow[d, "T"] & (\mathrm{Vect}_K^{\mathbb{Z}})^\circ \arrow[d, "\Sigma"] \\ \mathbb{M} \arrow[r, "F"] & (\mathrm{Vect}_K^{\mathbb{Z}})^\circ \end{tikzcd}$$

commutes, *i.e.* $\Sigma \circ F = F \circ T$.

Ultimately, we are looking for a functor whose target space is $\mathrm{Vect}_K^{\mathbb{Z}}$. Moreover, we note that the transformation T corresponds to degree-shifts and thus there is unnecessary information within the extended functor F. First, we consider the opposite functor $F^\circ : \mathbb{M}^\circ \to \mathrm{Vect}_K^{\mathbb{Z}}$. Second, we compose it with an evaluation map so as to obtain the desired functor

$$h(f) := \mathrm{ev}^0 \circ F^\circ : \mathbb{M}^\circ \to \mathrm{Vect}_K,$$

where there is no redundancy in the information it contains.

Definition 4.7 (RISC). *Let $f : \mathbb{X} \to \mathbb{R}$ be a continuous function. We define the* relative interlevel set cohomology *of $(\mathbb{X}, f)$ as the functor $h(f) := \mathrm{ev}^0 \circ F^\circ : \mathbb{M}^\circ \to \mathrm{Vect}_K$.*

Definition 4.8 (Persistence diagram). *Let $G : \mathbb{M}^\circ \to \mathrm{Vect}_K$ be a contravariant functor that vanishes on $\partial\mathbb{M}$. We define the* extended persistence diagram *of G as the multiset map*

$$\mathrm{Dgm}(G) : \begin{cases} \mathrm{int}(\mathbb{M}) \to \mathbb{N}_0 \\ m \mapsto \dim_K(G(m)) - \dim_K \left(\sum_{u \succ m} \mathrm{Im} G(m \preceq u)\right) \end{cases}$$

Definition 4.9 (Extended persistence diagram). *Let $f : \mathbb{X} \to \mathbb{R}$ be a continuous function with RISC $h(f)$ that is pointwise finite dimensional. We define the* extended persistence diagram *of $(\mathbb{X}, f)$ to be the multiset map $\mathrm{Dgm}(f) := \mathrm{Dgm}(h(f))$.*

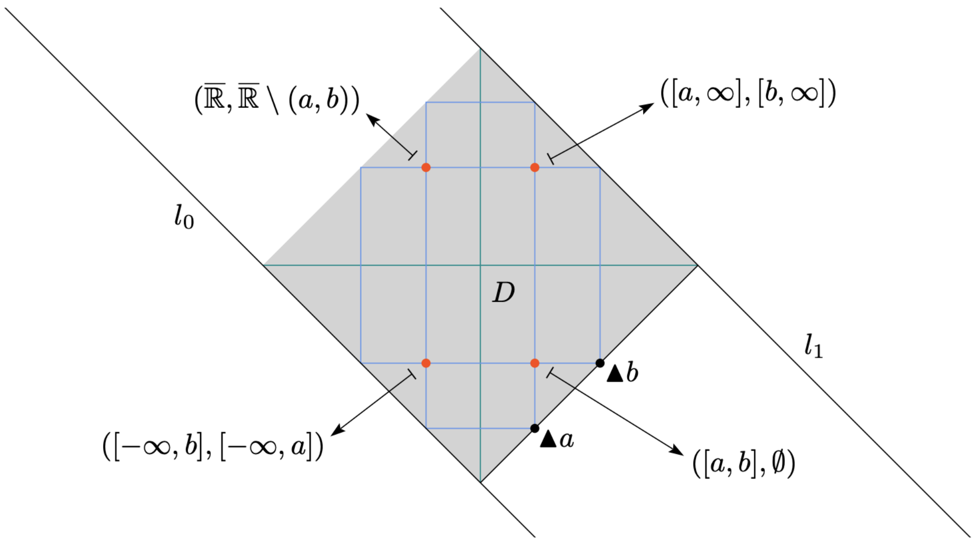

Figure 5: Schematic functioning of ρ (taken from [1].)

5 Making a Connection

5.1 About pyramids

First, we investigate the relation between the pyramid built on a pair $(\mathbb{X}, f)$ introduced in Section 3, which we denote by $\square_f$ and the strip $\mathbb{M}$.

Proposition 5.1. *Let $(\mathbb{X}, f)$ be a pair of Morse type. The pyramid $\square_f$ is equivalent to a sublattice embedded in the fundamental domain D of $\mathbb{M}$.*

Proof. Let $a_1 < ... < a_n$ be critical values of $f : \mathbb{X} \to \mathbb{R}$, with in-between regular values $-\infty < s_0 < a_1 < s_1 < \cdots < s_{n-1} < a_n < s_n < \infty$. We note that the pyramid $\square_f$ can be divided into four triangular regions, which are the faces of the pyramid. Now, each region contains a specific kind of relative or absolute space. Indeed, the south face contains spaces of the form $\mathbb{X}_i^j$ for $i \leq j$, which correspond to intervals of the form $[s_i, s_j]$. The north face contains pairs of the form $(\mathbb{X}_0^n, \mathbb{X}_0^i \cup \mathbb{X}_j^n)$, which correspond to pairs of the form $(\mathbb{R}, \mathbb{R} \setminus [s_i, s_j])$. The east face is made of pairs $(\mathbb{X}_0^i, \mathbb{X}_0^j)$ corresponding to pairs $((-\infty, s_i), (-\infty, s_j))$, and the west face is made of pairs $(\mathbb{X}_i^n, \mathbb{X}_j^n)$ corresponding to pairs $((s_i, \infty), (s_j, \infty))$. We conclude the proof by looking at the schematic functioning of the map $\rho : \mathbb{M} \to \mathcal{P}$ shown in Figure 5. ∎

5.2 Extracting a barcode

In this section, we show how one may extract an *extended persistence barcode* from the information contained in an extended persistence diagram $\mu : \mathrm{int}(\mathbb{M}) \to \mathbb{N}_0$. The first thing to note is that all the information we need is lies in how the injected real line ★ and its copies along $\mathbb{M}$ intersect with the fundamental blocks of μ. Now, to complete the reasoning, it remains to naturally extract a barcode from those intersections, which can form four types of intervals.

6 An Efficient Algorithm

We implement our algorithm via the bijection expressed in Theorem 3.6. Instead of dealing with intervals of spaces, we look at intervals of critical values and reformulate the diagram bijection theorem (*cf* Theorem 6.1).

Levelset zigzag persistence. Regarding the levelset zigzag persistence, we define the following correspondence. Every interval is either closed, open or half-open.

1. $[\mathbb{X}_{i-1}^{i}, \mathbb{X}_{j-1}^{j}] \leftrightarrow [a_i, a_j]$ for $i, j \in \{1, ..., n\}$
2. $[\mathbb{X}_{i-1}^{i}, \mathbb{X}_{j-1}^{j-1}] \leftrightarrow [a_i, a_j)$ for $i < j \in \{1, ..., n+1\}$
3. $[\mathbb{X}_{i}^{i}, \mathbb{X}_{j-1}^{j}] \leftrightarrow (a_i, a_j]$ for $i < j \in \{0, ..., n\}$
4. $[\mathbb{X}_{i}^{i}, \mathbb{X}_{j-1}^{j-1}] \leftrightarrow (a_i, a_j)$ for $i < j \in \{0, ..., n+1\}$

Extended persistence. As for the extended persistence, we define the notation change as follows. Intervals have four different forms (for *Type I* to *Type IV*).

1. $[\mathbb{X}_0^i, \mathbb{X}_0^{j-1}] \leftrightarrow [a_i, a_j)$ for $i, j \in \{1, ..., n\}$
2. $[(\mathbb{X}_0^n, \mathbb{X}_{j-1}^n), (\mathbb{X}_0^n, \mathbb{X}_i^n)]^+ \leftrightarrow [\bar{a}_j, \bar{a}_i)^+$ for $i < j \in \{1, ..., n+1\}$
3. $[\mathbb{X}_0^i, (\mathbb{X}_0^n, \mathbb{X}_j^n)] \leftrightarrow [a_i, \bar{a}_j)$ for $i < j \in \{0, ..., n\}$
4. $[\mathbb{X}_0^j, (\mathbb{X}_0^n, \mathbb{X}_i^n)]^+ \leftrightarrow [a_i, \bar{a}_j)^+$ for $i < j \in \{0, ..., n+1\}$

Theorem 6.1 (Diagram Bijection, new version). *One has the following correspondence between the intervals of extended persistence (left) and levelset zigzag persistence (right).*

Type	*Extended*	*Levelset zigzag*
I ($i < j$)	$[a_i, a_j)$	$[a_i, a_j)$
II ($i \leq j$)	$[\bar{a}_j, \bar{a}_i)^+$	$(a_i, a_j]$
III ($i < j$)	$[a_i, \bar{a}_j)$	$[a_i, a_j]$
IV ($i < j$)	$[a_j, \bar{a}_i)^+$	(a_i, a_j)

Proof. This is another formulation of Theorem 3.6. ∎

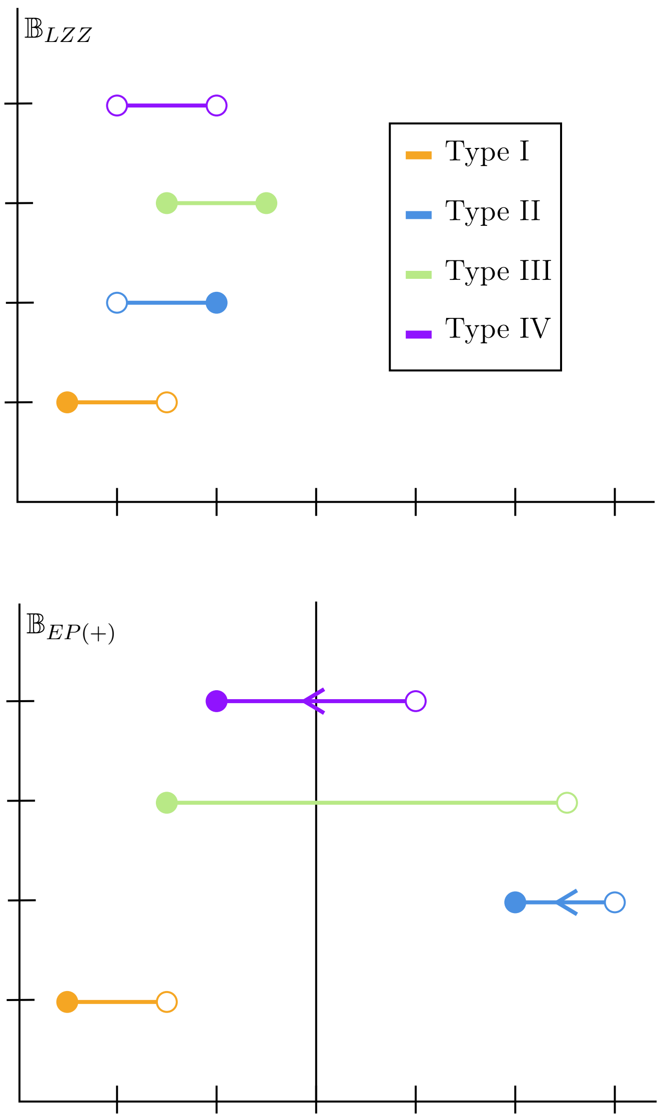

Figure 6: Illustrative example of the diagram bijection theorem. Barcodes of levelset zigzag persistence and extended persistence are denoted by $\mathbb{B}_{LZZ}$ and $\mathbb{B}_{EP(+)}$ respectively. In the extended persistence barcode, the central vertical segment separates points living in $\mathbb{R}$ from points living in $\mathbb{R}^{\circ}$. Intervals marked with an arrow correspond to intervals marked with a "+" exponent in Theorem 6.1.

Finally, the algorithm simply consists in computing the levelset zigzag persistence of a sequence of homology groups via Theorem 6.1, as shown in Figure 1's pipeline.

We consider the simple but very illustrative case of the circle S^1, paired with a height function.

Example 6.2 (The circle S^1)**.** *Let $\mathbb{X} = S^1$, and consider the height function $f : \mathbb{X} \to \mathbb{R}$ as defined in Figure 6.2, where a_1, a_2 are the critical values of the function f and s_0, s_1, s_2 are regular values of f such that $-\infty < s_0 < a_1 < s_1 < a_2 < s_2 < \infty$.*

Levelset zigzag persistence. *For $p \in \mathbb{N}$, we have levelset zigzag modules*

$$H_p(f^{-1}(s_0)) \to H_p(f^{-1}([s_0, s_1])) \leftarrow H_p(f^{-1}(s_1)) \to H_p(f^{-1}([s_1, s_2])) \leftarrow H_p(f^{-1}(s_2)).$$

The case $p = 0$ leads to the zigzag module

$$0 \to K \leftarrow K \oplus K \to K \leftarrow 0,$$

that decomposes into indecomposable summands as

$$(0 \to K \leftarrow K \to K \leftarrow 0) \oplus (0 \to 0 \leftarrow K \to 0 \leftarrow 0),$$

leading to the levelset zigzag barcode given by the intervals $[\mathbb{X}^1_0, \mathbb{X}^2_1]_0$ and $[\mathbb{X}^1_1, \mathbb{X}^1_1]_0$, where we use the 0 index to keep track of the dimensionality of the features.

Extended persistence. *For $p \in \mathbb{N}$, we have levelset zigzag modules*

$$\begin{array}{l} H_p(f^{-1}(s_0)) \longrightarrow H_p(f^{-1}([s_0, s_1])) \longrightarrow H_p(f^{-1}([s_0, s_2])) \longrightarrow \\ H_p(f^{-1}(S^1, \{s_2\})) \longrightarrow H_p(f^{-1}(S^1, [s_1, s_2])) \longrightarrow H_p(f^{-1}(S^1, S^1)). \end{array}$$

The case $p = 0$ gives the indecomposable module

$$0 \to K \to K \to K \to 0 \to 0,$$

and the case $p = 1$ gives the indecomposable module

$$0 \to 0 \to K \to K \to K \to 0,$$

leading to the extended barcode given by $[\mathbb{X}^1_0, (\mathbb{X}^2_0, \mathbb{X}^2_2)]_0$ and $[\mathbb{X}^2_0, (\mathbb{X}^2_0, \mathbb{X}^2_1)]_1$.

Correspondence. *One may refer to the table in Theorem 6.1 to obtain the matching*

$$\begin{cases} [\mathbb{X}^1_1, \mathbb{X}^1_1]_0 \leftrightarrow [\mathbb{X}^2_0, (\mathbb{X}^2_0, \mathbb{X}^2_1)]_1, \\ [\mathbb{X}^1_0, \mathbb{X}^2_1]_0 \leftrightarrow [\mathbb{X}^1_0, (\mathbb{X}^2_0, \mathbb{X}^2_2)]_0. \end{cases}$$

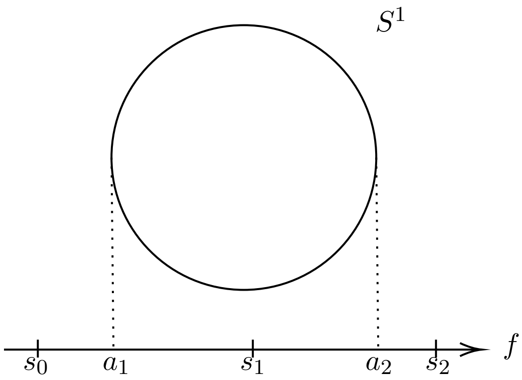

Figure 7: Height function on the circle S^1.

Interpretation. *The levelset zigzag persistence has a way of detecting cycles that is a bit less natural than standard persistence and extended persistence. In fact, the cycle formed by $\mathbb{X} = S^1$ is not represented with dimension-1 intervals. Instead, it creates a particular signature entirely encoded in the dimension-0 intervals.*

7 Defining a Bottleneck distance

The authors of [1] propose a way of defining a Bottleneck distance between extended persistence diagrams as introduced in their categorical setting (see Section 4). This work is contained in [2]. Here, we aim at introducing the Bottleneck distance and have a look at how it behaves on a concrete algorithmic scale.

Consider the strip $\mathbb{M}$ built as if l_1 intersects the x-axis at $(-\pi, 0)$ and l_2 intersects it at $(\pi, 0)$. Let $d_0 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+ \cup \{\infty\}$ be the unique extended metric satisfying

$$d_0(s,t) = \begin{cases} \tan t - \tan s \text{ if } [s,t] \cap (\frac{\pi}{2} + \pi\mathbb{Z}) = \emptyset, \\ \infty \text{ otherwise.} \end{cases}$$

Now, we define a metric on $\mathbb{R} \times \mathbb{R}$ based on d_0 as follows.

$$d_0^{\mathbb{R}\times\mathbb{R}} : \begin{cases} (\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \to \mathbb{R}^+ \cup \{\infty\} \\ ((s_1, s_2), (t_1, t_2)) \mapsto \max_{i=1,2} d_0(s_i, t_i) \end{cases}$$

Definition 7.1 (Metric d on the strip $\mathbb{M}$)**.** *We define an extended metric on $\mathbb{M}$ as*

$$d : \begin{cases} \mathbb{M} \times \mathbb{M} \to \mathbb{R}^+ \cup \{\infty\} \\ (s,t) \mapsto d_0^{\mathbb{R}\times\mathbb{R}}|_{\mathbb{M}\times\mathbb{M}}(s,t) \end{cases}$$

We give three lemmas that help getting a better understanding of how the metric $d : \mathbb{M} \times \mathbb{M} \to \mathbb{R}^+ \cup \{\infty\}$ acts. Before formulating those results, we introduce a notation that associates to a point of the interior of the strip a specific region of $\mathbb{M}$.

Notation. Let $m \in \mathbb{M}$. We denote by R_m the triangular region in which m is situated. For example, the fundamental domain D consists of four closed triangular regions (the north, east, south and west faces of the pyramid), whose interiors don't intersect. Moreover, let $\nu(m)$ denote the only integer $p \in \mathbb{N}$ such that $T^p(m) \in D$.

Lemma 7.2. *For all $s, t \in \mathrm{int}(\mathbb{M})$ such that $\nu(s) = \nu(t)$, we have*

$$d(s,t) = \infty \iff \mathrm{int}(R_s) \cap \mathrm{int}(R_t) = \emptyset.$$

Proof. This follows from the fact that for $s, t \in \mathrm{int}(\mathbb{M})$, one has $d(s,t) = \infty$ if and only if either $[s_1, t_1] \cap (\frac{\pi}{2} + \pi\mathbb{Z}) \neq \emptyset$ or $[s_2, t_2] \cap (\frac{\pi}{2} + \pi\mathbb{Z}) \neq \emptyset$. Indeed, the first condition above means that amongst s and t, one is in the upper part of its domain and the other one is in the bottom part. Similarly, the second condition means that one of them is in the left part of its domain and the other one is in the right part. ∎

Lemma 7.3. *For all $s, t \in \mathrm{int}(\mathbb{M})$ such that $|\nu(s) - \nu(t)| \geq 2$, one has $d(s,t) = \infty$.*

Lemma 7.4. *For all $s, t \in \mathrm{int}(\mathbb{M})$ such that $|\nu(s) - nu(t)| = 1$, one has $d(s,t) \neq \infty$ if and only if $|R_s \cap R_t| > 1$, i.e. if and only if the closed triangular regions R_s and R_t share a copy of the injected real line $\bigstar = \mathrm{Im}_{\blacktriangle}(\bar{R})$.*

Based on the introduced metric, we may now express the notion of *Bottleneck distance* between extended persistence diagrams. A δ-matching between two extended persistence diagrams is a partial matching of their vertices such that any two matched vertices $s, t \in \mathbb{M}$ satisfy $d(s,t) \leq \delta$ and unmatched vertices are at distance at most δ of the boundary $\partial\mathbb{M}$. For $\delta > 0$, we write $\mathcal{M}(\delta)$ for the set of pairs (μ_1, μ_2) of extended persistence diagrams for which there exists a δ-matching.

Definition 7.5 (Bottleneck distance)**.** *We define the* Bottleneck distance *between two extended peristence diagrams $\mu_1 : \mathrm{int}(\mathbb{M}) \to \mathbb{N}_0$ and $\mu_2 : \mathrm{int}(\mathbb{M}) \to \mathbb{N}_0$ as*

$$d_B(\mu_1, \mu_2) = \inf\{\delta > 0 \mid (\mu_1, \mu_2) \in \mathcal{M}(\delta)\}.$$

7.1 Stability

Suppose we have two barcodes of the form $\mathbb{B} = \{\mathbb{B}_{ord}, \mathbb{B}_{ext+}, \mathbb{B}_{ext-}, \mathbb{B}_{rel}\}$ and $\mathbb{B}'$. One may define a Bottleneck distance between $\mathbb{B}$ and $\mathbb{B}'$ by taking the maximum of the type-wise usual Bottleneck distances.

Question. If we denote by ϕ the transformation of an extended persistence barcode into a levelset zigzag barcode, does ϕ preserve the Bottleneck distance?

8 Conclusion

1. The metric question
2. Expression the diagram bijection in abstract setting
3. Diagram bijection with intuition ?
4. Conclusion
5. Tamal Dey paper : how do I make it fit in ?
6. Complete the first sections : diamond moves explained
7. Clean code
8. Making a connection section : illustration of pyramids DONE
9. Making a connection section : link between the two bijections
10. Example with height function on the circle. DONE
11. $\mathbb{M}$ induces a barcodes ON IT : think about how the injected copy of the real line into the strip intersects with the fundamental blocks :)
12. distance on $\mathbb{M} \times \mathbb{M}$ – alogrithmic version DONE
13. Nicolas : algebraic proof of recovering $\mathbb{M}$ from LZZ
14. Clarify $d : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ DONE

References

[1] Ulrich Bauer, Magnus Bakke Botnan, and Benedikt Fluhr. *Structure and Interleavings of Relative Interlevel Set Cohomology.* 2021. arXiv: 2108.09298 `[math.AT]`.

[2] Ulrich Bauer, Magnus Bakke Botnan, and Benedikt Fluhr. *Universality of the Bottleneck Distance for Extended Persistence Diagrams.* 2020. arXiv: 2007.01834 `[math.AT]`.

[3] Gunnar Carlsson. *Persistent Homology and Applied Homotopy Theory.* 2020. arXiv: 2004.00738 `[math.AT]`.

[4] Gunnar Carlsson and Vin de Silva. *Zigzag Persistence.* 2008. arXiv: 0812.0197 `[cs.CG]`.

[5] Gunnar Carlsson, Vin de Silva, and Dmitriy Morozov. "Zigzag Persistent Homology and Real-Valued Functions". In: *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry.* SCG '09. Aarhus, Denmark: Association for Computing Machinery, 2009, pp. 247–256. ISBN: 9781605585017. DOI: 10.1145/1542362.1542408. URL: https://doi.org/10.1145/1542362.1542408.

[6] Frédéric Chazal and Bertrand Michel. *An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists.* 2021. arXiv: 1710.04019 `[math.ST]`.