Generalized persistence diagrams for persistence modules over posets

Woojin Kim[1] · Facundo Mémoli[2]

© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

When a category $\mathscr{C}$ satisfies certain conditions, we define the notion of *rank invariant* for arbitrary poset-indexed functors $F : \mathbf{P} \to \mathscr{C}$ from a category theory perspective. This generalizes the standard notion of rank invariant as well as Patel's recent extension. Specifically, the barcode of any interval decomposable persistence modules $F : \mathbf{P} \to \mathbf{vec}$ of finite dimensional vector spaces can be extracted from the rank invariant by the principle of inclusion-exclusion. Generalizing this idea allows freedom of choosing the indexing poset $\mathbf{P}$ of $F : \mathbf{P} \to \mathscr{C}$ in defining Patel's generalized persistence diagram of F. Of particular importance is the fact that the generalized persistence diagram of F is defined regardless of whether F is interval decomposable or not. By specializing our idea to zigzag persistence modules, we also show that the zeroth level set barcode of a Reeb graph can be obtained in a purely set-theoretic setting without passing to the category of vector spaces. This leads to a promotion of Patel's semicontinuity theorem about type $\mathscr{A}$ persistence diagram to Lipschitz continuity theorem for the category of sets.

Keywords Generalized persistence diagrams · Generalized persistence modules · Zigzag persistence · Reeb graphs · Rank invariant · Limits and colimits

Mathematics Subject Classification 55N32 · 16G20 · 05E10

✉ Facundo Mémoli
memoli@math.osu.edu

Woojin Kim
woojin@math.duke.edu

1 Department of Mathematics, Duke University, Durham, USA

2 Department of Mathematics and Department of Computer Science and Engineering, The Ohio State University, Columbus, USA

Contents

1 Introduction

The notion of persistence diagram of Cohen-Steiner et al. (2007) was recently extended by Patel to the setting of constructible persistence modules $F : \mathbf{R} \to \mathscr{C}$ where $\mathscr{C}$ is an essentially small, symmetric monoidal category with images (Patel 2018). In a nutshell, the persistence diagram of $F : \mathbf{R} \to \mathscr{C}$ is obtained via Möbius inversion of the map sending each pair $s \leq t$ in $\mathbf{R}$ to the *image* of the morphism $F(s \leq t)$ in $\mathscr{C}$. This map can be regarded as a generalization of both the *rank invariant* (Carlsson and Zomorodian 2009) and the *persistent homology group* (Cohen-Steiner et al. 2007).

   In this paper we further generalize the notions of rank invariant and persistence diagram to the setting of locally finite poset-indexed functors $F : \mathbf{P} \to \mathscr{C}$, where, besides being symmetric, monoidal, essentially small and having images, the target category $\mathscr{C}$ is assumed to be bicomplete (see Convention 2.3 for another admissible set of assumptions about $\mathscr{C}$). A motivation for considering such level of generality is, as suggested by Patel, the possibility of studying torsion in data, e.g. persistent homology groups with integer coefficients. Specifically, we identify an explicit formula for the generalized persistence diagram of F in terms of the generalized rank invariant of F that we will define. We in particular prove that, given an *interval decomposable* diagram of vector spaces, its rank invariant and its persistence diagram can be translated into each other, and either of them contains enough information to reconstruct the diagram up to isomorphism.

   One consequence of our framework is a novel method for computing the zeroth level set barcode of a Reeb graph (Sect. 4). This method is obtained by re-interpreting the barcode of a zigzag module (Carlsson et al. 2009) as the generalized persistence diagram. Furthermore, since the generalized persistence diagram is defined even for zigzag modules *valued in categories other than the category of vector spaces*, our framework has significant potential to be exploited in the theoretical and algorithmic study on zigzag persistence (Botnan 2017; Botnan and Lesnick 2018; Carlsson and De Silva 2010; Carlsson et al. 2019; Curry and Patel 2020; Elchesen and Mémoli 2019; Milosavljević et al. 2011; Oudot and Sheehy 2015), and also in its applications to mobile sensor networks, image processing, analysis of time-varying metric

spaces/graphs (Adams and Carlsson 2015; Corcoran and Jones 2016; Kim and Mémoli 2017, 2018a; Mata et al. 2015), etc.

1.1 Our contributions

By **P** denote any connected, locally finite poset (Definitions 2.5 and 2.14), and by $\mathscr{C}$ denote any symmetric monoidal, essentially small, bicomplete category with images (see Convention 2.3 for another admissible set of assumptions about $\mathscr{C}$).

(i)   We generalize the notion of rank invariant to functors $\mathbf{P} \to \mathscr{C}$ (i.e. generalized persistence modules (Bubenik and Scott 2014)). In particular, our construction generalizes Patel's rank function (Patel 2018): see Definition 3.5 and Example 3.6 (i). For any zigzag diagram of vector spaces, we show that our construction of the rank invariant is not only more faithful than the one introduced by Puuska (2020), but also becomes a *complete* invariant: see Example 3.6 (ii), Theorem 3.14, Remark 3.15, and Appendix C.

In the sequel, we assume that the poset **P** is connected and *essentially finite*—a condition which is slightly stronger condition than being locally-finiteness; cf. Definition 3.10. Also, let **vec** be the category of finite-dimensional vector spaces over a fixed field $\mathbb{F}$.

(ii)   We extend the notion of generalized persistence diagram by Patel (2018) to arbitrary functors $\mathbf{P} \to \mathscr{C}$. In particular, for any interval decomposable persistence module $F : \mathbf{P} \to \mathbf{vec}$, the persistence diagram/barcode of F can be extracted from the rank invariant by the *principle of inclusion-exclusion* in combinatorics (Theorem 3.14). This implies that our construction of rank invariant is *complete* invariant for interval decomposable persistence modules. Also, as a by-product, we obtain a novel necessary condition for the interval decomposability of persistence modules $\mathbf{P} \to \mathbf{vec}$: See Remark 3.16 and Example 3.20.

(iii) It is well known that a Reeb graph can be seen as a zigzag persistence in the category of sets (Curry and Patel 2020; De Silva et al. 2016). In this respect, we show that the 0-th level set barcode of a Reeb graph can be computed in a purely set-theoretic setting *without passing to the category of vector spaces* (Sect. 4). As a corollary, we partially promote the semicontinuity theorem by Patel to a Lipschitz continuity theorem (Corollary 4.16) and this continuity theorem further extends to a certain class of Reeb graphs (Theorem 4.21). These results indicate that the semicontinuity theorem by Patel is open to further improvement.

1.2 Key ideas: Möbius inversion and a categorical view of the rank invariant

This section aims at highlighting the main ideas of this paper, without discussing technical details. Specifically, we briefly overview how to generalize the notions of rank invariant and persistence diagram to the setting of finite poset-indexed diagrams; strictly speaking, we do not actually require the indexing poset to be finite, but *locally finite* (for defining the rank invariant) or *essentially finite* (for defining the persistence diagram; Definition 3.10).

Let **P** be a finite and connected poset (Definition 2.5). Consider a diagram $F : \mathbf{P} \to \mathscr{C}$ where $\mathscr{C}$ is a symmetric monoidal, bicomplete category with images (all these terms are defined in Sect. 2.1). By combining the limit cone of F and the colimit cocone of F, one obtains the canonical limit-to-colimit map $\phi_F : \varprojlim F \to \varinjlim F$.

Let **Con(P)** be the collection of all subposets I of $\hat{\mathbf{P}}$ such that the restriction of the Hasse diagram of **P** to I is a connected graph (Definition 2.16). For $I \in \mathbf{Con(P)}$, we consider the restricted diagram $F|_I : I \to \mathscr{C}$ and define $\mathrm{im}(F|_I)$ to be the *image* of the canonical map $\phi_{F|_I} : \varprojlim F|_I \to \varinjlim F|_I$. Let $\mathscr{I}(\mathscr{C})$ denote the collection of all isomorphism classes of $\mathscr{C}$. Now, as a generalization of the *rank invariant* (a.k.a. *rank function*) in Carlsson and Zomorodian (2009), McCleary and Patel (2020), Patel (2018), we define the rank invariant of F as the function

$$\mathrm{rk}(F) : \mathbf{Con(P)} \to \mathscr{I}(\mathscr{C}),$$

which sends $I \in \mathbf{Con(P)}$ to the isomorphism class of $\mathrm{im}(F|_I)$. By the assumptions on $\mathscr{C}$, the codomain $\mathscr{I}(\mathscr{C})$ is a symmetric monoid.[1] Let $\mathscr{A}(\mathscr{C})$ be the group completion of $\mathscr{I}(\mathscr{C})$, known as the *Grothendieck group* of $\mathscr{C}$. We define the persistence diagram of F as the *Möbius inversion* of $\mathrm{rk}(F)$, which amounts to a function $\mathbf{Con(P)} \to \mathscr{A}(\mathscr{C})$ (Definition 3.13). These notions of rank invariant and persistence diagram generalize those appearing in Botnan and Lesnick (2018), Carlsson et al. (2009); Carlsson and Zomorodian (2009), Cohen-Steiner et al. (2007), Patel (2018), Zomorodian and Carlsson (2005).

When restricted to the case when $\mathscr{C} = \mathbf{vec}$, related ideas have been or are currently being explored by several researchers: Botnan, Oppermann and Steen have been studying the rank of ϕ_F for counting "thin" summands in the indecomposable decomposition of $F : \mathbf{P} \to \mathbf{Vec}$. We remark that their work is addressing a version of Proposition 3.17, see announcement in Botnan (2017). Also, Chambers and Letscher (2018) define persistent homology for a filtration over a directed graph as the image of the canonical map between the limit and colimit. Their work also provides a variant of Proposition 3.17 (Chambers and Letscher 2018, Lemma 3.1).

Botnan and Lesnick exploited the notion of *left Kan extensions* for addressing stability of barcode of a zigzag module (Botnan and Lesnick 2018). We remark that $\mathrm{rk}(F)$ above can be interpreted as an invariant of F which is obtained by interconnecting the *left and right Kan extensions* of F along a certain restriction functor. We refer the interested readers to the extended version of this paper; (Kim and Mémoli 2018b, Sections E.2 and E.3).

1.2.1 Organization

In Sect. 2 we briefly review basic concepts from category theory and persistence theory, and we also recall basic terminology from order theory. In Sect. 3 we define the rank invariant and persistence diagram of a functor $\mathbf{P} \to \mathscr{C}$. In Sect. 4 we propose a novel viewpoint on Reeb graphs and their 0-th level set barcodes. This new perspective

---

[1] A symmetric monoid is a set S equipped with a binary operation which is associative and symmetric. Also, S must contain an identity element.

provides a new method to compute the 0-th level set barcodes of Reeb graphs without passing to the category of vector spaces. Also, we establish stability results for the persistence diagrams of merge trees and "untwisted" Reeb graphs. In Sect. 5 we discuss possible extensions.

2 Preliminaries

In Sect. 2.1 we set up terminology and notation relevant to category theory. In Sect. 2.2 we review the notion of poset-indexed persistence modules and their interval decomposability. Also, we introduce notation relevant to zigzag persistence modules. In Sect. 2.3 we introduce the notion of path-connected subposets of a given poset which will be instrumental in the definitions of generalized rank invariant and persistence diagram.

2.1 Category theory elements

2.1.1 Categories

Consult (Awodey 2010; Mac Lane 2013) for general definitions related to category theory. For any category $\mathscr{C}$, let $\mathrm{ob}(\mathscr{C})$ and $\mathrm{hom}(\mathscr{C})$ denote the class/set of all objects and that of all morphisms in $\mathscr{C}$, respectively. Let I be a small category, i.e. $\mathrm{ob}(I)$ and $\mathrm{hom}(I)$ are sets. For any two functors $F, G : I \rightarrow \mathscr{C}$, we write $F \cong G$ if F and G are naturally isomorphic. A functor $F : I \rightarrow \mathscr{C}$ will sometimes be referred to as a *diagram*. Since the domain I is small, we also refer to F as a *small* diagram. In particular, if $\mathrm{ob}(I)$ and $\mathrm{hom}(I)$ are finite sets, then $F : I \rightarrow \mathscr{C}$ will be called a finite diagram. A sub-diagram of F means the restriction $F|_J$ to a full subcategory J of I. The following categories will be of main interest in this paper.

(i) By **Vec** and **vec**, we mean the category of vector spaces and *finite* dimensional vector spaces, respectively with linear maps over a fixed field $\mathbb{F}$.
(ii) By **Set** and **set**, we mean the category of sets and *finite* sets, respectively with set maps.

Most of the following concepts can be found in Mitchell (1965), Patel (2018), Weibel (2013).
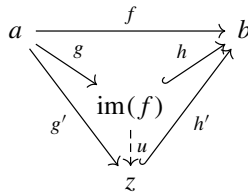
2.1.2 Symmetric monoidal category with images

A *symmetric monoidal category* $(\mathscr{C}, \square)$ is, in brief, a category $\mathscr{C}$ with a binary operation $\square$ on its object and an identity object $e \in \mathrm{ob}(C)$ satisfying the following properties:

– (Symmetry) $a \square b \cong b \square a$, for all $a, b \in \mathrm{ob}(C)$.
– (Associativity) $a \square (b \square c) \cong (a \square b) \square c$, for all $a, b, c \in \mathrm{ob}(C)$.
– (Identity) $a \square e \cong a$, for all $a \in \mathrm{ob}(C)$.

We refer the reader to Weibel (2013, p. 114) for the precise definition of a symmetric monoidal category.

A morphism $f : a \to b$ is said to be a *monomorphism* if f is left-cancellative: for any morphisms $k_1, k_2 : c \to a$, if $f \circ k_1 = f \circ k_2$, then $k_1 = k_2$. Such f is often written as $f : a \hookrightarrow b$. On the other hand, a right-cancellative morphism $g : a \to b$ is said to be an *epimorphism*, and often written as $g : a \twoheadrightarrow b$.

Definition 2.1 (*Images and epimorphic images*) A morphism $f : a \to b$ has an *image* if there exist a monomorphism $h : \mathrm{im}(f) \hookrightarrow b$ and a morphism $g : a \to \mathrm{im}(f)$ such that $f = h \circ g$, satisfying the following property: For any monomorphism $h' : z \hookrightarrow b$ and a morphism $g' : a \to z$ with $f = h' \circ g'$, there is a unique morphism $u : \mathrm{im}(f) \to z$ such that the following diagram commutes:

$$
\begin{array}{ccc}
a & \xrightarrow{\quad f \quad} & b \\
& g \searrow \quad \nearrow h & \\
& \mathrm{im}(f) & \\
g' \searrow & \downarrow u & \nearrow h' \\
& z &
\end{array}
$$

If, moreover, the morphism g is an epimorphism, then f is said to have an *epimorphic image*.

Remark 2.2 The morphism u in the diagram above is a monomorphism: Assume that there exist two morphisms $k_1, k_2 : c \to \mathrm{im}(f)$ with $u \circ k_1 = u \circ k_2$. Then,

$$
\begin{aligned}
u \circ k_1 = u \circ k_2 &\implies h' \circ (u \circ k_1) = h' \circ (u \circ k_2) \\
&\implies (h' \circ u) \circ k_1 = (h' \circ u) \circ k_2 \\
&\implies h \circ k_1 = h \circ k_2 \\
&\implies k_1 = k_2 \quad \text{since } h \text{ is a monomorphism.}
\end{aligned}
$$

A category in which every morphism has an image (resp. epimorphic image) is said to be a *category with images (resp. category with epimorphic images)*. See Mitchell (1965, p.12) for more about images.

2.1.3 Complete, cocomplete, essentially small categories (Mac Lane 2013)

A category $\mathscr{C}$ is called *complete* if every small diagram $F : I \to \mathscr{C}$ has a limit in $\mathscr{C}$. Likewise, a category $\mathscr{C}$ is called *cocomplete* if every small diagram has a colimit in $\mathscr{C}$. If a category $\mathscr{C}$ is both complete and cocomplete, then $\mathscr{C}$ is called *bicomplete*.[2] A category $\mathscr{C}$ is called *essentially small* if the collection of isomorphism classes of objects in $\mathscr{C}$ is a set.

Convention 2.3 Throughout this paper, unless otherwise stated, we always assume that $(\mathscr{C}, \square)$ (or simply $\mathscr{C}$) is essentially small and symmetric monoidal. Also, we always assume that at least one of conditions (a) and (b) below hold:

---

[2] Besides **Set** and **Vec**, examples include the category of groups, the category of abelian groups, the category of topological spaces.

(a) $\mathscr{C}$ is bicomplete and has images.
(b) $\mathscr{C}$ is a full subcategory[3] of a bicomplete category $\mathscr{D}$, where

- $\mathscr{D}$ has *epimorphic* images
- for any monomorphism $f : a \hookrightarrow b$ in $\mathscr{D}$, if $b \in \mathrm{ob}(\mathscr{C})$, then $a \in \mathrm{ob}(\mathscr{C})$.
- for any epimorphism $g : c \twoheadrightarrow d$ in $\mathscr{D}$, if $c \in \mathrm{ob}(\mathscr{C})$, then $d \in \mathrm{ob}(\mathscr{C})$.

For example, the categories (**set**, $\sqcup$) (where $\sqcup$ stands for disjoint union), (**vec**, $\oplus$) (where $\oplus$ stands for direct sum), and (**ab**, $\oplus$) (the category of finitely generated abelian groups) all conform with Convention 2.3. In particular, they are full subcategories of the bicomplete categories **Vec**, **Set**, and **Ab** (the category of abelian groups) respectively whose images are epimorphic.

2.1.4 Grothendieck groups

We review the notion of Grothendieck groups from Weibel (2013, Chapter II) and Patel (2018, Sect. 6). Fix a category $(\mathscr{C}, \square)$. Let $\mathscr{I}(\mathscr{C})$ be the set of all isomorphism classes of $\mathscr{C}$. For any $C \in \mathrm{ob}(\mathscr{C})$, let $[C]$ denote the isomorphism class of C in $\mathscr{I}(\mathscr{C})$. The binary operation in $\mathscr{I}(\mathscr{C})$ is defined as $[a]+[b] := [a \square b]$. The group completion $\mathscr{A}(\mathscr{C})$ of $\mathscr{I}(\mathscr{C})$ with respect to $+$ is said to be the *Grothendieck group* of $\mathscr{C}$. Hence, the notions of addition $+$ and subtraction $-$ are well-defined in $\mathscr{A}(\mathscr{C})$. When $\mathscr{C}$ is an abelian category, there is another type of Grothendieck group $\mathscr{B}(\mathscr{C})$.[4] However, it is beyond the scope of this paper to give a complete treatment of $\mathscr{B}(\mathscr{C})$.

Remark 2.4 For the monoid $(\mathscr{I}(\textbf{vec}), \bigoplus)$ (resp. $(\mathscr{I}(\textbf{set}), \bigsqcup))$ of isomorphism classes, we have the isomorphism

$$\left(\mathscr{I}(\textbf{vec}), \bigoplus \right) \cong (\mathbf{Z}_+, +) \ (\text{resp.} \left(\mathscr{I}(\textbf{set}), \bigsqcup \right) \cong (\mathbf{Z}_+, +)),$$

which sends each isomorphism class $[V] \in \mathscr{I}(\textbf{vec})$ to $\dim(V)$ (resp. $[A] \in \mathscr{I}(\textbf{set})$ to $|A|$). Extending this isomorphism, we obtain the group isomorphism $(\mathscr{A}(\textbf{vec}), \bigoplus) \cong (\mathbf{Z}, +)$ (resp. $(\mathscr{A}(\textbf{set}), \bigsqcup) \cong (\mathbf{Z}, +).$).

2.2 Interval decomposability and barcodes

Given any poset $\mathbf{P}$, we regard $\mathbf{P}$ as the category: Objects are elements of $\mathbf{P}$. For any $p, q \in \mathbf{P}$, there exists a unique morphism $p \to q$ if and only if $p \leq q$.

For $\mathbf{P}$ a poset and $\mathscr{C}$ an arbitrary category, $F : \mathbf{P} \to \mathscr{C}$ a functor, and $s \in \mathbf{P}$, let $F_s := F(s)$. Also, for any pair $s \leq t$ in $\mathbf{P}$, let $\varphi_F(s, t) : F_s \to F_t$ denote the morphism $F(s \leq t)$. Any functor $F : \mathbf{P} \to \textbf{vec}$ is said to be a $\mathbf{P}$-*indexed module*.

---

[3] A subcategory $\mathscr{C}'$ of a category $\mathscr{C}$ is *full* if for all $a, b \in \mathrm{ob}(\mathscr{C}')$, $\hom_{\mathscr{C}'}(a, b) = \hom_{\mathscr{C}}(a, b)$.

[4] Define the relation $\sim$ on $\mathscr{A}(\mathscr{C})$ as $[b] \sim [a] + [c]$ if there exists a short exact sequence $0 \to a \to b \to c \to 0$. Then, $\mathscr{B}(C)$ is defined as the quotient group $\mathscr{A}(\mathscr{C})/\sim$. The category **vec** is an instance of abelian categories, and it holds that $\mathscr{A}(\textbf{vec}) = \mathscr{B}(\textbf{vec})$ (Patel 2018, Example 6.2.1).

2.2.1 Interval modules and direct sums

We mostly follow the notation and definitions from Botnan and Lesnick (2018).

Definition 2.5 (*Connected posets*) A poset $\mathbf{P}$ is said to be *connected*, if for all $p, q \in \mathbf{P}$, there exists a sequence $p = p_0, \ldots, p_n = q$ in $\mathbf{P}$ such that p_i and p_{i+1} are comparable, i.e. $p_i \leq p_{i+1}$ or $p_{i+1} \leq p_i$, for $i = 0, \ldots, n-1$.

For a poset $\mathbf{P}$, non-empty *convex* connected subposets of $\mathbf{P}$ are said to be *intervals* of $\mathbf{P}$:

Definition 2.6 (*Intervals*) Given a poset $\mathbf{P}$, an *interval* $\mathscr{J}$ of $\mathbf{P}$ is any non-empty subset $\mathscr{J} \subset \mathbf{P}$ such that (1) $\mathscr{J}$ is connected, and (2) if $r, t \in \mathscr{J}$ and $r \leq s \leq t$, then $s \in \mathscr{J}$.
By $\mathbf{Int}(\mathbf{P})$, we denote the collection of all intervals of $\mathbf{P}$.

For $\mathscr{J}$ an interval of $\mathbf{P}$, the *interval module* $I^{\mathscr{J}} : \mathbf{P} \to \mathbf{vec}$ is the $\mathbf{P}$-indexed module where

$$I_t^{\mathscr{J}} = \begin{cases} \mathbb{F} & \text{if } t \in \mathscr{J}, \\ 0 & \text{otherwise.} \end{cases} \qquad \varphi_{I^{\mathscr{J}}}(s, t) = \begin{cases} \mathrm{id}_{\mathbb{F}} & \text{if } s, t \in \mathscr{J}, \ s \leq t, \\ 0 & \text{otherwise.} \end{cases}$$

Let F, G be $\mathbf{P}$-indexed modules. The *direct sum* $F \bigoplus G$ *of* F *and* G is the $\mathbf{P}$-indexed module defined as follows: for all $s \in \mathbf{P}$, $(F \bigoplus G)_s := F_s \bigoplus G_s$ and for all $s \leq t$ in $\mathbf{P}$, the linear map $\varphi_{F \bigoplus G}(s, t) : (F \bigoplus G)_s \to (F \bigoplus G)_t$ is defined by

$$\varphi_{F \bigoplus G}(s, t)(v, w) := \big(\varphi_F(s, t)(v), \varphi_G(s, t)(w)\big)$$

for all $(v, w) \in (F \bigoplus G)_s$. We say that a $\mathbf{P}$-indexed module F is *decomposable* if F is (naturally) isomorphic to $G_1 \bigoplus G_2$ for some non-trivial $\mathbf{P}$-indexed modules G_1 and G_2, and we denote it by $F \cong G_1 \bigoplus G_2$. Otherwise, we say that F is *indecomposable*.

Proposition 2.7 (Botnan and Lesnick 2018, Proposition 2.2) $I^{\mathscr{J}}$ *is indecomposable.*

2.2.2 Barcodes

Recall that *a multiset* is a collection of objects (called elements) in which elements may occur more than once. We call the number of instances of an element in a specific multiset *the multiplicity* of the element. For example, $A = \{x, x, y\}$ is a multiset and the multiplicity of x is two. Also, this multiset A is distinct from the multiset $\{x, y\}$.

Definition 2.8 (*Interval decomposability*) A $\mathbf{P}$-indexed module F is *interval decomposable* if there exists a multiset $\mathrm{barc}^{\mathbf{P}}(F)$ of intervals (Definition 2.6) of $\mathbf{P}$ such that

$$F \cong \bigoplus_{\mathscr{J} \in \mathrm{barc}^{\mathbf{P}}(F)} I^{\mathscr{J}}$$

Fig. 1 The posets **ZZ** in Definition 2.10
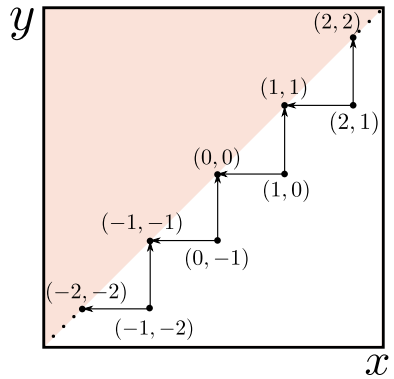
It is well-known that, by the theorem of Azumaya-Krull-Remak-Schmidt (Azumaya 1950), such a decomposition is unique up to a permutation of the terms in the direct sum. Therefore, the multiset barc$^\mathbf{P}(F)$ is unique if F is interval decomposable since a multiset is careless of the order of its elements.

Definition 2.9 (*Barcodes*) We call barc$^\mathbf{P}(F)$ in Definition 2.8 the *barcode* of F.

Given any two posets $\mathbf{P}$, $\mathbf{P}'$, we assume that by default the product $\mathbf{P} \times \mathbf{P}'$ is equipped with the partial order where $(p, p') \leq (q, q')$ if and only if $p \leq p'$ and $q \leq q'$. Also, by $\mathbf{P}^{\mathrm{op}}$ we mean the opposite poset of $\mathbf{P}$, i.e. for any $p, q \in \mathbf{P}$, $p \leq q$ in $\mathbf{P}^{\mathrm{op}}$ if and only if $q \leq p$ in $\mathbf{P}$.

Definition 2.10 (*Zigzag poset*) The poset $\mathbf{ZZ} = \{(i, j) \in \mathbf{Z}^2 : j = i \text{ or } j = i - 1\}$ has the partial order inherited from $\mathbf{R}^{\mathrm{op}} \times \mathbf{R}$. See Fig. 1.

Definition 2.11 (*Zigzag module*) Any functor $F : \mathbf{ZZ} \to \mathbf{vec}$ is called a *zigzag module*.

Theorem 2.12 (Interval decomposability of 1-D persistence modules and zigzag modules Botnan 2017; Carlsson and De Silva 2010; Crawley-Boevey 2015) *For any* $\mathbf{P} \in \{\mathbf{R}, \mathbf{Z}, \mathbf{ZZ}\}$, $\mathbf{P}$-*indexed modules are interval decomposable. Also, for each interval* I *of* $\mathbf{ZZ}$ (*Definition 2.6*), I-*indexed modules are interval decomposable.*

2.2.3 The barcode associated to a zigzag module

By Theorem 2.12, any zigzag module $M : \mathbf{ZZ} \to \mathbf{vec}$ admits a *barcode* (Definition 2.9), which will be denoted by barc$^{\mathbf{ZZ}}(M)$.

Notation 2.13 (*Intervals of* $\mathbf{ZZ}$) The notation introduced in Botnan and Lesnick (2018) is useful for describing barcodes of zigzag modules: Letting $\preceq$ (resp. $\prec$) denote the partial order (resp. the strict partial order) on $\mathbf{Z}^2$ (not on $\mathbf{Z}^{\mathrm{op}} \times \mathbf{Z}$), any interval of $\mathbf{ZZ}$ falls into one of the following four types:

$$(b, d)_{\mathbf{ZZ}} := \{(i, j) \in \mathbf{ZZ} : (b, b) \prec (i, j) \prec (d, d)\} \quad \text{for } b < d \text{ in } \mathbf{Z} \cup \{-\infty, \infty\},$$

$$[b, d)_{\mathbf{ZZ}} := \{(i, j) \in \mathbf{ZZ} : (b, b) \preceq (i, j) \prec (d, d)\} \quad \text{for } b < d \text{ in } \mathbf{Z} \cup \{\infty\},$$
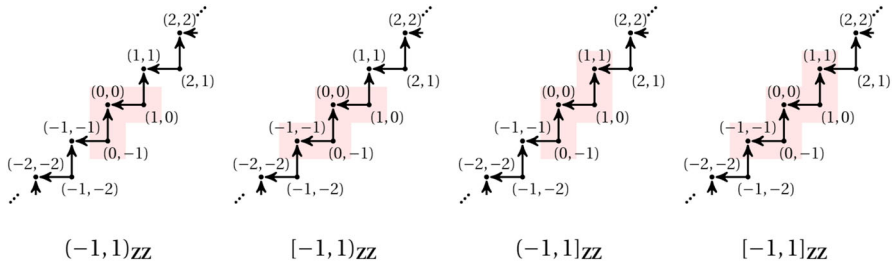
Fig. 2 The points falling into the shaded regions comprise the intervals $(-1, 1)_{\mathbf{ZZ}}, [-1, 1)_{\mathbf{ZZ}}, (-1, 1]_{\mathbf{ZZ}}$ and $[-1, 1]_{\mathbf{ZZ}}$ of the poset $\mathbf{ZZ}$, respectively in order

$$(b, d]_{\mathbf{ZZ}} := \{(i, j) \in \mathbf{ZZ} : (b, b) \prec (i, j) \preceq (d, d)\} \quad \text{for } b < d \text{ in } \mathbf{Z} \cup \{-\infty\},$$
$$[b, d]_{\mathbf{ZZ}} := \{(i, j) \in \mathbf{ZZ} : (b, b) \preceq (i, j) \preceq (d, d)\} \quad \text{for } b \leq d \text{ in } Z.$$

See Fig. 2 for examples. Specifically, we let $\langle b, d \rangle_{\mathbf{ZZ}}$ denote any of the above types of intervals. By utilizing this notation, the barcode of a zigzag module $M : \mathbf{ZZ} \to \mathbf{vec}$ can be expressed, for some index set J, as

$$\mathrm{barc}^{\mathbf{ZZ}}(M) = \left\{\!\!\left\{ \langle b_j, d_j \rangle_{\mathbf{ZZ}} : M \cong \bigoplus_{j \in J} I^{\langle b_j, d_j \rangle_{\mathbf{ZZ}}} \right\}\!\!\right\}.$$

Here, $\{\!\{\cdot\}\!\}$ is used instead of $\{\cdot\}$ to indicate $\mathrm{barc}^{\mathbf{ZZ}}(M)$ is a *multiset.*

2.3 Hasse diagram and path-connected subposets of a poset

We introduce the notion of *path-connected subposets* of a poset (this concept is *different* from that of *connected posets* in Definition 2.5). To that end, we begin by reviewing relevant terminology from lattice theory (Birkhoff 1940).

Definition 2.14 (*Locally finite posets*) A poset $\mathbf{P}$ is said to be *locally finite* if for all $p, q \in \mathbf{P}$ with $p \leq q$, the set $[p, q] := \{r \in \mathbf{P} : p \leq r \leq q\}$ is finite.

Let $\mathbf{P}$ be a poset and let $p, q \in \mathbf{P}$. We say that p *covers* q if $q < p$ and there is no $r \in \mathbf{P}$ such that $q < r < p$. If either [p covers q] or [q covers p], then we write $p \asymp q$.

Definition 2.15 (*Hasse diagram of a poset*) The *Hasse diagram* of a poset $\mathbf{P}$ is a simple graph on the vertex set $\mathbf{P}$, denoted by $\mathrm{Hasse}(\mathbf{P})$, where any two different vertices $p, q \in \mathbf{P}$ are adjacent if and only if $p \asymp q$.

If $\mathbf{P}$ is connected (Definition 2.5) and locally finite, then it is clear that $\mathrm{Hasse}(\mathbf{P})$ is a connected graph. Locally finiteness is an important assumption for studying the connectedness of the Hasse diagram of a poset. For example, even though $\mathbf{R}$ is a connected totally ordered set, the Hasse diagram of $\mathbf{R}$ contains no edge.

Definition 2.16 (*Path-connected subposets*) Let $\mathbf{P}$ be a poset. A subposet $I \subset \mathbf{P}$ is said to be *path-connected* in $\mathbf{P}$ if the induced subgraph of Hasse($\mathbf{P}$) on I is a connected graph.

By $\mathbf{Con}(\mathbf{P})$ we denote the collection of all path-connected subposets of $\mathbf{P}$. We equip $\mathbf{Con}(\mathbf{P})$ with the partial order given by inclusion, i.e. a pair $I, J \in \mathbf{Con}(\mathbf{P})$ satisfies $I \leq J$ if and only if $I \subset J$.

We emphasize the difference between connectivity and path-connectivity (Definitions 2.5 and 2.16) as follows: If a subposet $\mathbf{Q}$ is path-connected in $\mathbf{P}$, then $\mathbf{Q}$ itself is connected. However, in general, the converse is not true. For example, given the poset $\{a \leq b \leq c\}$, the subposet $\{a \leq c\}$ is connected itself, but *not* path-connected in $\{a \leq b \leq c\}$.

Remark 2.17 Let $\mathbf{P}$ be a locally finite poset. Note that the collection $\mathbf{Int}(\mathbf{P})$ of all intervals of $\mathbf{P}$ (Definition 2.6) is contained in $\mathbf{Con}(\mathbf{P})$. Also, $\mathbf{Int}(\mathbf{P})$ can sometimes be equal to $\mathbf{Con}(\mathbf{P})$, e.g. $\mathbf{P} = \mathbf{Z}$ or $\mathbf{P} = \mathbf{ZZ}$.

3 Generalized rank invariant and generalized persistence diagrams

In this section we introduce the notions of *generalized rank invariant* and *generalized persistence diagram.*

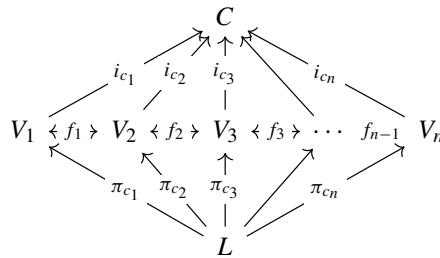### 3.1 Rank invariant for persistence modules over posets

In this section, we introduce the notion of *generalized rank invariant* of a diagram $\mathbf{P} \to \mathscr{C}$ when $\mathbf{P}$ is a locally finite, connected poset (as for assumptions on the target category $\mathscr{C}$, recall Convention 2.3).

3.1.1 Rank and rank invariant of a persistence module over a poset

Recall the notions of cone and cocone from category theory (Definitions A.1 and A.4). The following observation is the first step toward defining the rank invariant of a persistence module over a poset.

Proposition 3.1 (Canonical map from a cone to a cocone) *Let $\mathbf{P}$ be a connected poset and let $F : \mathbf{P} \to \mathscr{C}$. Let $(L, (\pi_a)_{a \in \mathbf{P}})$ and $(C, (i_a)_{a \in \mathbf{P}})$ be any cone and cocone of F, respectively. Then for any $a, b \in \mathbf{P}$, we have that $i_a \circ \pi_a = i_b \circ \pi_b$.*

Proof Fix $a, b \in \mathbf{P}$ and let $a = c_1, c_2, \ldots, c_n = b$ be any sequence in $\mathbf{P}$ such that c_j and c_{j+1} are comparable for $j = 1, \ldots, n - 1$. Let $V_j := M_{c_j}$ for each j. Then, we have the combined commutative diagram of $\left(L, (\pi_{c_j})_{j=1,\ldots,n}\right)$ and $\left(C, (i_{c_j})_{j=1,\ldots,n}\right)$ as follows:

$$C$$

$$i_{c_1} \quad i_{c_2} \quad i_{c_3} \qquad i_{c_n}$$

$$V_1 \leftarrow f_1 \rightarrow V_2 \leftarrow f_2 \rightarrow V_3 \leftarrow f_3 \rightarrow \cdots \quad f_{n-1} \quad V_n$$

$$\pi_{c_1} \quad \pi_{c_2} \quad \pi_{c_3} \qquad \pi_{c_n}$$

$$L$$

Without loss of generality, assuming f_1 is a map from V_1 to V_2, we prove that $i_{c_1} \circ \pi_{c_1} = i_{c_2} \circ \pi_{c_2}$. This is clear by tracking the commutativity of the diagram above:

$$i_{c_1} \circ \pi_{c_1} = (i_{c_2} \circ f_1) \circ \pi_{c_1} = i_{c_2} \circ (f_1 \circ \pi_{c_1}) = i_{c_2} \circ \pi_{c_2}.$$

Similarly, for each $k = 2, \ldots, n-1$, we have $i_{c_k} \circ \pi_{c_k} = i_{c_{k+1}} \circ \pi_{c_{k+1}}$, completing the proof.   □

Recall that a limit (resp. colimit) of a functor F is the terminal (resp. initial) object in the category of cones (resp. cocones) over F (Definitions A.2 and A.5). By virtue of Proposition 3.1, we can define:

Definition 3.2 (*Rank of a poset-indexed diagram*) Let $\mathbf{P}$ be a connected poset. The *rank* of an $F : \mathbf{P} \to \mathscr{C}$ is defined as the isomorphism class of the image of the canonical limit-to-colimit map $\psi_F : \varprojlim F \to \varinjlim F$.

In what follows, for any $F : \mathbf{P} \to \mathbf{vec}$, we will regard the codomain of $\mathrm{rk}(F)$ as $\mathbf{Z}_+$ by identifying $m \in \mathbf{Z}_+$ with $[\mathbb{F}^m] \in \mathscr{I}(\mathbf{vec})$ (Remark 2.4).

Remark 3.3 In Definition 3.2, assuming $\mathscr{C} = \mathbf{vec}$, the rank of F cannot exceed $\min_{\mathbf{a} \in \mathbf{P}} \dim(F_\mathbf{a})$ and is thus finite. More generally, $\mathrm{im}(\psi_F)$ belongs to $\mathscr{I}(\mathscr{C})$ for any category $\mathscr{C}$ satisfying Convention 2.3.

The example below justifies the use of the term *rank*.

Example 3.4 (i) For the singleton set $\{\bullet\}$, a functor $F : \{\bullet\} \to \mathbf{vec}$ amounts to a vector space $F_\bullet$ with the identity map $\mathrm{id}_{F_\bullet} : F_\bullet \to F_\bullet$. In this case, the rank of F is the dimension of $F_\bullet$.

(ii) A functor $G : \{a \leq b\} \to \mathbf{vec}$ amounts to the linear map $G_a \xrightarrow{g} G_b$, where $g = \varphi_G(a \leq b)$. Note that $\varprojlim G \cong G_a$ and $\varinjlim G \cong G_b$ and the rank of G is identical to that of g.

Definition 3.5 (*Generalized rank invariant*) Let $\mathbf{P}$ be a locally finite, connected poset. We define the *(generalized) rank invariant* of a diagram $F : \mathbf{P} \to \mathscr{C}$ as the map

$$\mathrm{rk}(F) : \mathbf{Con}(\mathbf{P}) \to \mathscr{I}(\mathscr{C})$$

which sends each $I \in \mathbf{Con}(\mathbf{P})$ to the rank of the sub-diagram $F|_I$.

In the following example, we will see that the generalized rank invariant is either *equivalent to* or *a refinement of* the rank invariant that have been considered in the literature (Carlsson and Zomorodian 2009; Cerri et al. 2013; Patel 2018; Puuska 2020):

Example 3.6 (i) (1-parameter persistence) Consider the integers $\mathbf{Z}$. We have $\mathbf{Con}(\mathbf{Z}) = \mathbf{Int}(\mathbf{Z})$ and the generalized rank invariant of a diagram $F : \mathbf{Z} \to \mathbf{vec}$ coincides with the standard rank invariant (Carlsson and Zomorodian 2009; Cohen-Steiner et al. 2007). This directly follows from the following observation: for any $a, b \in \mathbf{Z}$ with $a \leq b$, the rank of $\varphi_F(a, b)$ coincides with the rank of $F|_{[a,b]}$ given in Definition 3.2. This observation is still valid even if the target category of F is other than $\mathbf{vec}$, implying that the rank invariant of $\mathbf{Z} \to \mathscr{C}$ in Definition 3.5 is essentially equal to that of Patel (2018).

(ii) (Zigzag persistence) Consider the zigzag poset $\mathbf{ZZ}$. Note that $\mathbf{Con}(\mathbf{ZZ}) = \mathbf{Int}(\mathbf{ZZ})$. The generalized rank invariant of a diagram $F : \mathbf{ZZ} \to \mathbf{vec}$ contains more information than the rank invariant defined in Puuska (2020). See Appendix C. More interestingly, even though rk(F) is defined in way which makes no reference to whether F is interval decomposable or not , it will turn out that one can extract the barcode of F from rk(F) (see Fig. 4 (A)). This implies that the generalized rank invariant is a *complete* invariant for zigzag modules, i.e. any two zigzag modules that have the same rank invariant are isomorphic. This is a generalization of the completeness of the rank invariant for $\mathbf{Z}$-indexed persistence (Theorem B.5).

(iii) (Multiparameter persistence) Consider the 2-dimensional grid poset $\mathbf{Z}^2$ and a functor $F : \mathbf{Z}^2 \to \mathbf{vec}$. For any $(a, b), (c, d) \in \mathbf{Z}^2$ with $(a, b) \leq (c, d)$, the rank of $\varphi_F((a, b), (c, d))$ coincides with the rank of the restricted diagram $F|_{[(a,b),(c,d)]}$, in the sense of Definition 3.2. However, the standard rank invariant (Carlsson and Zomorodian 2009) does not record the rank of $F|_I$ when I is not in the form of $[(a, b), (c, d)]$. See Fig. 3 for an illustrative example. This implies that the generalized rank invariant is a refinement of the standard rank invariant for 2-parameter persistence modules $\mathbf{Z}^2 \to \mathbf{vec}$. Similar argument applies to diagrams $\mathbf{Z}^d \to \mathbf{vec}$ for $d > 2$.

3.1.2 Order-reversing property of the rank invariant

It is well-known that there exists a translation-invariant order $\leq$ on the Grothendieck group $\mathscr{A}(\mathscr{C})$ of $\mathscr{C}$ (Patel 2018, Sect. 6.1), Weibel (2013). For example, both $(\mathscr{A}(\mathbf{vec}), \leq)$ and $(\mathscr{A}(\mathbf{set}), \leq)$ are isomorphic to $(\mathbf{Z}, \leq)$. In analogy with the case of standard persistent modules,[5] if we assume that the target $\mathscr{C}$ satisfies certain properties, then the generalized rank invariant is also order-reversing. To establish this result, we first introduce a proposition which is a slight extension of Puuska (2020, Lemma 3.4):

Proposition 3.7 *Let $\mathscr{C}$ be an essentially small, symmetric monoidal category with epimorphic images (Definition 2.1). Also, assume that, for all $a, b \in \mathrm{ob}(\mathscr{C})$,*

---

[5] Given a diagram $F : \mathbf{Z} \to \mathbf{vec}$, it holds that rank $\varphi_F(i', j') \leq$ rank $\varphi_F(i, j)$ for $i' \leq i \leq j \leq j'$ in $\mathbf{Z}$.

Fig. 3 Consider $I \in \mathbf{Con}(\mathbf{Z}^2)$ consisting of 6 points in $\mathbf{Z}^2$, depicted as above. The standard rank invariant of $F : \mathbf{Z}^2 \to \mathbf{vec}$ does not record the rank of $F|_I$, whereas the generalized rank invariant of F does
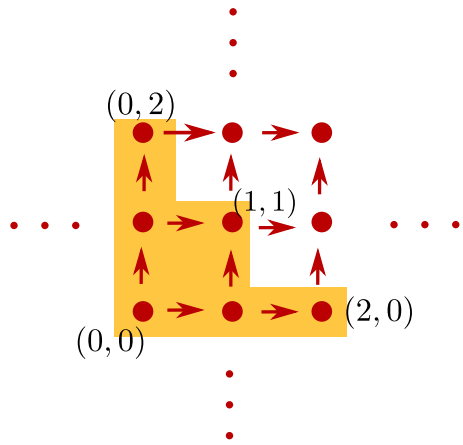
(i) $a \hookrightarrow b \Rightarrow [a] \leq [b]$ in $\mathscr{A}(\mathscr{C})$,
(ii) $a \twoheadrightarrow b \Rightarrow [b] \leq [a]$ in $\mathscr{A}(\mathscr{C})$.

Let $f : a \to b, g : a' \to a, h : b \to b'$ be morphisms in $\mathscr{C}$ and denote $f' = h \circ f \circ g$. Then, $\left[\mathrm{im}(f')\right] \leq [\mathrm{im}(f)]$ in $\mathscr{A}(\mathscr{C})$.

Proof By the universal properties of images (which are epimorphic), we have the following commutative diagram:

$$\tag{1}$$
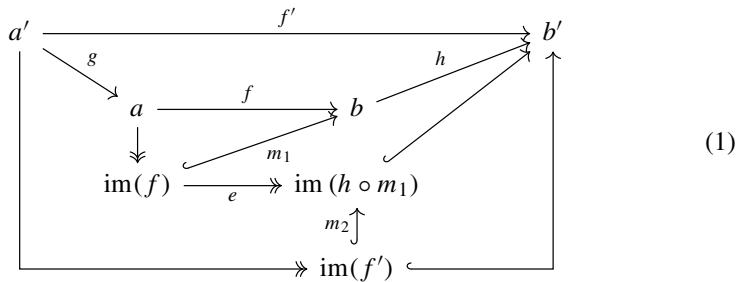
In particular, by the epimorphism e and the monomorphism m_2 (Remark 2.2) in the diagram, we have $\left[\mathrm{im}(f')\right] \leq [\mathrm{im}(h \circ m_1)] \leq [\mathrm{im}(f)]$, as desired. $\qquad\square$

Proposition 3.8 (Order-reversing property of the generalized rank invariant) *Let $\mathscr{C}$ be an essentially small, symmetric monoidal category satisfying the conditions in Convention 2.3 (b). Let $\mathbf{P}$ be any locally finite, connected poset and let $F : \mathbf{P} \to \mathscr{C}$ be a functor. Then, whenever I and I' in $\mathbf{Con}(\mathbf{P})$ are such that $I \subset I'$, it holds that $\mathrm{rk}(F)(I') \leq \mathrm{rk}(F)(I)$ in $\mathscr{A}(\mathscr{C})$.*

The order-reversing property of the rank invariant can sometimes be useful for the efficient computation of persistence diagrams we will define (Definition 3.13 and Example 3.20).

Proof of Proposition 3.8 Fix $I, I' \in \mathbf{Con}(\mathbf{P})$ such that $I \subset I'$. Notice that the limit $\left(\varprojlim F|_{I'}, (\pi'_a)_{a \in I'} \right)$ and colimit $\left(\varinjlim F|_{I'}, (i'_a)_{a \in I'} \right)$ can be restricted to a cone and a cocone over $F|_I$, respectively (Remark A.7). Then by the terminal property of $\left(\varprojlim F|_I, (\pi_a)_{a \in I} \right)$ and the initial property of $\left(\varinjlim F|_I, (i_a)_{a \in I} \right)$, there exist unique morphisms $p : \varprojlim F|_{I'} \to \varprojlim F|_I$ and $\iota : \varinjlim F_I \to \varinjlim F|_{I'}$ such that for every a in I, $\pi'_a = \pi_a \circ p$ and $i'_a = \iota \circ i_a$. Let $\psi_F(I) : \varprojlim F|_I \to \varinjlim F|_I$ and $\psi_M(I') : \varprojlim F|_{I'} \to \varinjlim F|_{I'}$ be the canonical LC maps. Fix $a \in I$. Then

$$\psi_M(I') = i'_a \circ \pi'_a = (\iota \circ i_a) \circ (\pi_a \circ p) = \iota \circ (i_a \circ \pi_a) \circ p = \iota \circ \psi_M(I) \circ p.$$

We first show that $\mathrm{im}(\psi_F(I))$ belongs to $\mathrm{ob}(\mathscr{C})$, not $\mathrm{ob}(\mathscr{D}) \setminus \mathrm{ob}(\mathscr{C})$. By the universal property of $\mathrm{im}(i_a \circ \pi_a)(= \mathrm{im}(\psi_F(I)))$ and $\mathrm{im}(i_a)$, there exist monomorphisms $\mathrm{im}(i_a \circ \pi_a) \hookrightarrow \mathrm{im}(i_a) \hookrightarrow \varinjlim F|_I$ (Remark 2.2). Also, since $\mathscr{C}$ has epimorphic images, we have an epimorphism from $F_a \in \mathrm{ob}(\mathscr{C})$ to $\mathrm{im}(i_a)$. Then, by assumption (b), $\mathrm{im}(i_a) \in \mathrm{ob}(\mathscr{C})$. Now, by assumption (b), $\mathrm{im}(i_a \circ p_a) \in \mathrm{ob}(\mathscr{C})$. By the same argument, one can check that $\mathrm{im}(\psi_F(I')) \in \mathrm{ob}(\mathscr{C})$.

Note that ι and p are morphisms of $\mathscr{D}$, but not necessarily $\mathscr{C}$. However, in diagram (1), by replacing f, f', g and h by $\psi_M(I)$, $\psi_M(I')$, p and ι respectively, and invoking assumptions (b) and (b) in Proposition 3.8, we have $\mathrm{rk}(F)(I') \leq \mathrm{rk}(F)(I)$ in $\mathscr{A}(\mathscr{C})$ as desired.   □

3.2 Persistence diagrams for persistence modules over posets

In this section we define the generalized persistence diagram of a functor $\mathbf{P} \to \mathscr{C}$ where $\mathbf{P}$ is a connected and *essentially finite* poset (defined below), and $\mathscr{C}$ satisfies the conditions in Convention 2.3.

3.2.1 Essentially finite posets

A locally finite poset $\mathbf{P}$ is said to be *essentially finite* if every $I \in \mathbf{Con}(\mathbf{P})$ has a finite *perimeter*:

Definition 3.9 (*Neighborhood and perimeter*) Let $\mathbf{P}$ be a locally finite poset. For $I \in \mathbf{Con}(\mathbf{P})$, we define the *neighborhood* of I as

$$\mathrm{nbd}(I) := \{p \in \mathbf{P} \setminus I : \text{there exists } q \in I \text{ such that } q \asymp p\}.$$

The *perimeter* o_I of I is defined as the cardinality of $\mathrm{nbd}(I)$ (note that if I is a singleton $\{p\}$, then $\mathrm{nbd}(I)$ is the neighborhood of the vertex p in the graph $\mathrm{Hasse}(\mathbf{P})$).

We remark that, for each interval I of $\mathbf{Z}$, the perimeter of I is either 0, 1 or 2: If $I = \mathbf{Z}$, then the perimeter is 0. For intervals of the form $I = [b, \infty)$, $b \in \mathbf{Z}$ or $I = (-\infty, d]$, $d \in \mathbf{Z}$, the perimeter of I is 1. If $I = [b, d]$ for some $b, d \in \mathbf{Z}$, then perimeter of I is 2. Similarly, for the poset $\mathbf{ZZ}$, the perimeter of any interval of $\mathbf{ZZ}$ is either 0, 1 or 2.

Definition 3.10 (*Essentially finite posets*) A poset $\mathbf{P} := (\mathbf{P}, \leq)$ is said to be *essentially finite* if

(i) $(\mathbf{P}, \leq)$ is locally finite, and
(ii) For each $I \in \mathbf{Con}(\mathbf{P})$, the perimeter o_I of I is finite.

Examples of essentially finite posets include all finite posets, the integers $\mathbf{Z}$, and $\mathbf{ZZ}$. For $d > 1$, the infinite d-dimensional integer grid $\mathbf{Z}^d$ is locally finite, but *not* essentially finite. For example, the interval $\mathbf{Z} \times \{0\}$ of $\mathbf{Z}^2$ has the infinite neighborhood $\mathbf{Z} \times \{1, -1\}$.

3.2.2 Persistence diagrams for generalized persistence modules

The following notation is useful for defining generalized persistence diagrams.

Notation 3.11 (*n-th entourage*) Let $\mathbf{P}$ be a connected, essentially finite poset and let $I \in \mathbf{Con}(\mathbf{P})$. Fix $n \in \mathbf{N}$. By I^n, we denote the set of all $J \subset \mathbf{P}$ such that $I \subset J \subset I \cup \mathrm{nbd}(I)$ and $|J \cap \mathrm{nbd}(I)| = n$. In other words, each $J \in I^n$ is obtained by adding n points of $\mathrm{nbd}(I)$ to I. We refer to I^n as the *n-th entourage* of I.

Remark 3.12 Let o_I be the perimeter of I (Definition 3.9).

(i) For example, $I^1 = \{I \cup \{p\} : p \in \mathrm{nbd}(I)\}$, and $|I^1| = o_I$. In general, for $n \leq o_I$, one has $|I^n| = \binom{o_I}{n}$, whereas for $n > o_I$, $I^n = \emptyset$.
(ii) Any $J \in I^n$ belongs to $\mathbf{Con}(\mathbf{P})$.

By Remark 2.4, we can identify the Grothendieck group $\mathscr{A}(\mathbf{vec})$ with the integer group $(\mathbf{Z}, +)$. Given an $F : \mathbf{Z} \to \mathbf{vec}$ and $[a, b] \in \mathbf{Int}(\mathbf{Z}) = \mathbf{Con}(\mathbf{Z})$ with $a < b$, let us recall that in the standard persistence diagram of F the multiplicity of $(a, b) \in \mathbf{Z}^2$ is defined as Cohen-Steiner et al. (2007):

$$\mathrm{rk}(F)([a, b]) - \mathrm{rk}(F)([a - 1, b]) - \mathrm{rk}(F)([a, b + 1]) + \mathrm{rk}(F)([a - 1, b + 1]).$$

This formula is a motivation for:

Definition 3.13 (*Generalized persistence diagram of* $\mathbf{P} \to \mathscr{C}$) Let $\mathbf{P}$ be a connected, essentially finite poset. Given an $F : \mathbf{P} \to \mathscr{C}$, let us define the persistence diagram $\mathrm{dgm}^{\mathbf{P}}(F) : \mathbf{Con}(\mathbf{P}) \to \mathscr{A}(\mathscr{C})$ of F by sending each $I \in \mathbf{Con}(\mathbf{P})$ to

$$\mathrm{dgm}^{\mathbf{P}}(F)(I) := \mathrm{rk}(F)(I) - \sum_{J \in I^1} \mathrm{rk}(F)(J) + \sum_{K \in I^2} \mathrm{rk}(F)(K) - \ldots$$
$$+ (-1)^{o_I} \sum_{L \in I^{o_I}} \mathrm{rk}(F)(L). \tag{2}$$

At the end of this section we will show that $\mathrm{dgm}^{\mathbf{P}}(F)$ is the Möbius inversion of $\mathrm{rk}(F)$ over the poset $\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$, generalizing the framework of Patel (2018). Now, by invoking the principle of inclusion-exclusion, we prove that the barcode of interval decomposable persistence modules (Definition 2.9) can be obtained via formula (2):
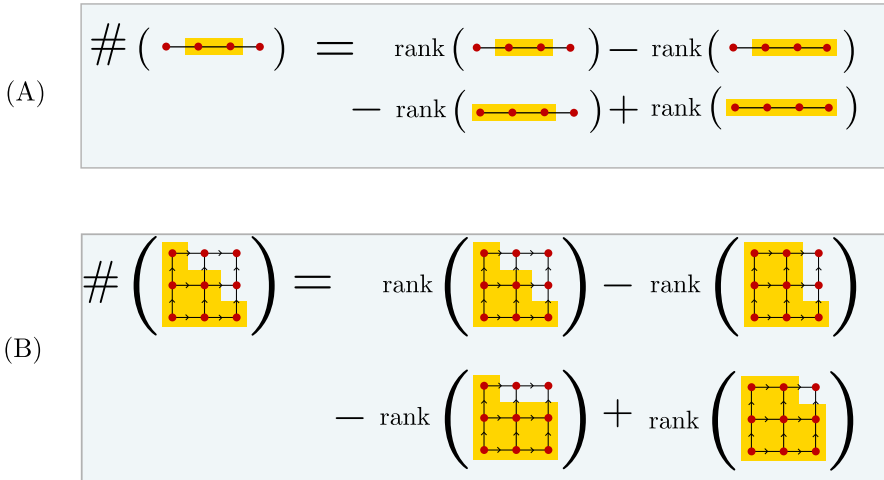
Fig. 4 **a** Let $\leftrightarrow$ denote $\leq$ or $\geq$. Given any **vec**-diagram M indexed by $\{1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4\}$, M is interval decomposable. The multiplicity of the interval $[2, 3]$ in the barcode of M is equal to $\mathrm{rank}(M|_{[2,3]}) - \mathrm{rank}(M|_{[2,4]}) - \mathrm{rank}(M|_{[1,3]}) + \mathrm{rank}(M|_{[1,4]})$. **b** Given any interval decomposable persistence module N over the grid $\{1, 2, 3\} \times \{1, 2, 3\}$, the multiplicity of the interval indicated by the LHS in the barcode of N can be computed in a similar way

Theorem 3.14 (Generalized persistence diagram recovers the barcode) *Let* **P** *be any connected, essentially finite poset and let* $F : \mathbf{P} \to \mathbf{vec}$ *be interval decomposable. For* $I \in \mathbf{Con}(\mathbf{P})$, *the value given by equation* (2) *is equal to the multiplicity of* I *in* $\mathrm{barc}^{\mathbf{P}}(F)$ *(this implies that if* $I \in \mathbf{Con}(\mathbf{P}) \setminus \mathbf{Int}(\mathbf{P})$, *then the value is zero).*

See Fig. 4 for illustrative examples of applications of Theorem 3.14. We remark that formula (2) can be simplified in special cases, e.g. see Botnan et al. (2020) for the case of rectangle decomposable $\mathbf{Z}^2$-indexed persistence modules. We will obtain a proof of Theorem 3.14 via an elementary argument resting upon the inclusion-exclusion principle. This proposition can also be obtained as a corollary to Propositions 3.18 and 3.19, based on standard results from incidence algebra.

Remark 3.15 Since the barcode of an interval decomposable $F : \mathbf{P} \to \mathbf{vec}$ is a complete invariant of F, Theorem 3.14 implies that the generalized rank invariant of F (Definition 3.5) is also a complete invariant for F.

Remark 3.16 Theorem 3.14 also implies the following: Given any $G : \mathbf{P} \to \mathbf{vec}$, (1) if there exists $I \in \mathbf{Con}(\mathbf{P}) \setminus \mathbf{Int}(\mathbf{P})$ such that $\mathrm{dgm}^{\mathbf{P}}(G)(I) \neq 0$, then G is *not* interval decomposable. (2) If there exists $I \in \mathbf{Int}(\mathbf{P})$ such that $\mathrm{dgm}^{\mathbf{P}}(G)(I) < 0$, then G is *not* interval decomposable: See Example 3.20. (3) The converse of statement in item (2) is not true: See Example 3.21.

Let us recall the following folklore fact for $\mathbf{Z}$-indexed modules: Let $F : \mathbf{Z} \to \mathbf{vec}$ and let $a, b \in \mathbf{Z}$ with $a \leq b$. Then, the rank of $\varphi_F(a, b)$ is equal to the total number of intervals $J \in \mathrm{barc}^{\mathbf{Z}}(F)$ such that $[a, b] \subset J$. In our setting, this fact generalizes to:

Proposition 3.17 (Barcode recovers the generalized rank invariant) *Let* **P** *be a connected, locally finite poset. For any interval decomposable* $F : \mathbf{P} \to \mathbf{vec}$ *and* $J \in \mathbf{Con}(\mathbf{P})$, $\mathrm{rk}(F)(J)$ *is equal to the total multiplicity of intervals* $K \in \mathrm{barc}^{\mathbf{P}}(F)$ *such that* $K \supset J$.

Here we provide a succinct proof of Proposition 3.17 *under the extra assumption* that $\varprojlim F|_J$ is of finite dimension. A general proof is deferred to the appendix due to its length (Sect. E).

Proof in the case where $\varprojlim F|_J$ *is finite dimensional.* Since F is interval decomposable, there exists an indexing set C such that $F \cong \bigoplus_{c \in C} I^{K_c}$, where each K_c is an interval of **P**. Let us recall that (1) colimits preserve direct sums, and (2) limits preserve direct products (Mac Lane 2013, Theorem V.5.1). By virtue of the extra assumption that $\varprojlim F|_J$ is finite dimensional (as well as for each $p \in \mathbf{P}$, F_p is finite dimensional), the notions of direct product and direct sum coincide in the category of cones over F. Therefore,

$$\mathrm{rk}(F)(J) = \mathrm{rank}(F|_J) = \mathrm{rank}\left(\bigoplus_{c \in C} I^{K_c}|_J\right) = \sum_{c \in C} \mathrm{rank}\left(I^{K_c}|_J\right).$$

Note that, for every $c \in C$, $\mathrm{rank}\left(I^{K_c}|_J\right) = \begin{cases} 1, & K_c \subset J \\ 0, & \text{otherwise.} \end{cases}$ (cf. Remark 3.3). Therefore, the right-most summation is the total multiplicity of intervals $K_c \in \mathrm{barc}^{\mathbf{P}}(F)$ such that $K_c \supset J$.   □

Proof of Theorem 3.14 Let $\mathrm{barc}^{\mathbf{P}}(F) = \{\!\!\{ J_c : F \cong \bigoplus_{c \in C} I^{J_c} \}\!\!\}$ for some indexing set C. Let $a_0(I)$ be the cardinality of the set $\{c \in C : I \subset J_c\}$. Also, let $a_1(I)$ be the cardinality of the set $\{c \in C : I \subsetneq J_c\}$. Since the difference $a_0(I) - a_1(I)$ is the multiplicity of I in $\mathrm{barc}^{\mathbf{P}}(F)$, it suffices to show that $a_0(I) - a_1(I)$ is identical to the value given by formula (2). First, by Proposition 3.17, $a_0(I) = \mathrm{rk}(F)(I)$.

Next we compute $a_1(I)$. For any $K \in \mathbf{Con}(\mathbf{P})$, let $B_F(K) := \{c \in C : K \subset J_c\}$.

Let $K \in \mathbf{Con}(\mathbf{P})$ be such that $I \subsetneq K$. Let us show that K must include at least one element of $\mathrm{nbd}(I)$: Pick any $k \in K \setminus I$ and any $p \in I$. Since the induced subgraph of $\mathrm{Hasse}(\mathbf{P})$ on K is connected, there exist $p = q_0, q_1, \ldots, q_m = k$ in K such that q_i and q_{i+1} are adjacent (i.e. $q_i \asymp q_{i+1}$) in the subgraph for all i. Since $p = q_0 \in I$ and $q_m = k \in K \setminus I$, there must be $i \in \{1, \ldots, m\}$ such that q_i belongs to $\mathrm{nbd}(I)$.

Now, we have the equality

$$\{c \in C : I \subsetneq J_c\} = \bigcup_{K \in I^1} B_F(K) \text{ (see Notation 3.11)},$$

and hence $a_1(I) = \left|\bigcup_{K \in I^1} B_F(K)\right|$. In particular, since F is a diagram of finite dimensional vector spaces, the set $\{c \in C : I \subsetneq J_c\}$ and all the sets $B_F(K)$ must be finite sets. Therefore, by the principle of inclusion-exclusion and Proposition 3.17, we have:

$$a_1(I) = \sum_{J \in I^1} |B_F(J)| - \sum_{K \in I^2} |B_F(K)| + \ldots + (-1)^{o_I - 1} \sum_{L \in I^{o_I}} |B_F(L)|$$

$$= \sum_{J \in I^1} \mathrm{rk}(F)(J) - \sum_{K \in I^2} \mathrm{rk}(F)(K) + \ldots + (-1)^{o_I - 1} \sum_{L \in I^{o_I}} \mathrm{rk}(F)(L),$$

completing the proof. $\qquad\square$

3.2.3 Möbius function and another interpretation of Definition 3.13

Let $\mathbf{Q}$ be a locally finite poset. The Möbius function $\mu_{\mathbf{Q}} : \mathbf{Q} \times \mathbf{Q} \to \mathbf{Z}$ of $\mathbf{Q}$ is defined[6] recursively (Rota 1964) as

$$\mu_{\mathbf{Q}}(p, q) = \begin{cases} 1, & p = q, \\ -\sum_{p \le r < q} \mu_{\mathbf{Q}}(p, r), & p < q, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Proposition 3.18 (Möbius inversion formula Rota 1964, Proposition 2 (p.344)) *Let $\mathbf{Q}$ be a locally finite poset and let k be a field. Suppose that an element $0 \in \mathbf{Q}$ exists with the property that $0 \le q$ for all $q \in \mathbf{Q}$. Consider a pair of functions $f, g : \mathbf{Q} \to k$ with the property that*

$$g(q) = \sum_{r \le q} f(r).$$

Then, $f(q) = \sum_{r \le q} g(r) \cdot \mu_{\mathbf{Q}}(r, q)$ for $q \in \mathbf{Q}$.

Under the assumption that $\mathbf{Con}(\mathbf{P})$ is locally finite, formula (2) is in fact the *Möbius inversion* of the rank invariant $\mathrm{rk}(F)$ over the poset $\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$:

Proposition 3.19 (Extension of Patel's generalized persistence diagrams) *Let $\mathbf{P}$ be a connected, essentially finite poset such that $(\mathbf{Con}(\mathbf{P}), \subset)$ is locally finite. Let $F : \mathbf{P} \to \mathscr{C}$. Let $\mu : \mathbf{Con}^{\mathrm{op}}(\mathbf{P}) \times \mathbf{Con}^{\mathrm{op}}(\mathbf{P}) \to \mathbf{Z}$ be the Möbius function of the poset $\mathbf{Con}(\mathbf{P})^{\mathrm{op}}$. Then, for $I \in \mathbf{Con}^{\mathrm{op}}(\mathbf{P})$*

$$\mathrm{dgm}^{\mathbf{P}}(F)(I) = \sum_{J \supset I} \mathrm{rk}(F)(J) \cdot \mu(J, I).$$

We prove Proposition 3.19 at the end of this section.

Example 3.20 Let $\mathbf{P} := \{a, b, c, d\}$ equipped with the partial order $\le := \{(a, b), (c, b), (d, b)\} \subset \mathbf{P} \times \mathbf{P}$. Then, Hasse($\mathbf{P}$) is shown as in (A) below. Consider $F : \mathbf{P} \to \mathbf{vec}$ given as in (B) below,

---

[6] More precisely, the codomain of $\mu_{\mathbf{Q}}$ is the multiple of 1 in a specified base ring rather than $\mathbf{Z}$.
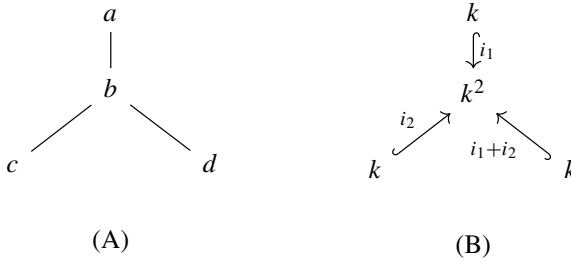
where $i_1, i_2 : k \to k^2$ are the canonical inclusions into the first factor and the second factor of k^2, respectively. We show that F is *not* interval decomposable.

By Remark 3.16, it suffices to show that, for $I := \{b\}$, one has $\mathrm{dgm}^{\mathbf{P}}(F)(I) < 0$. Observe that $I^1 = \{\{a, b\}, \{b, c\}, \{b, d\}\}$, $I^2 = \{\{a, b, c\}, \{a, b, d\}, \{b, c, d\}\}$, and $I^3 = \{\{a, b, c, d\}\} = \{\mathbf{P}\}$, $I^n = \{\emptyset\}$ for $n > 3$. Note that:

– $\mathrm{rk}(F)(I) = 2$, which is the dimension of $F(b) = k^2$ (cf. Example 3.4 (i)).
– $\mathrm{rk}(F)(J) = 1$ for every $J \in I^1$ (cf. Example 3.4 (ii)).
– $\mathrm{rk}(F)(J) = 0$ for every $J \in I^2$: the sub-diagram $F|_J$ amounts to a zigzag module of length 3 (Carlsson et al. 2009). It is not difficult to check that the barcode of $F|_J$ does not include the full interval J. This implies, by Proposition 3.17, $\mathrm{rk}(F)(J) = \mathrm{rk}(F|_J)(J) = 0$.
– $\mathrm{rk}(F)(\mathbf{P}) = 0$: Fix $J \in I^2$. By Proposition 3.8, $0 \leq \mathrm{rk}(F)(\mathbf{P}) \leq \mathrm{rk}(F)(J) = 0$.

Therefore, formula (2) gives that $\mathrm{dgm}^{\mathbf{P}}(F)(I) = 2 - (1+1+1) + (0+0+0) - 0 = -1$, completing the proof.

The non-negativity of the persistence diagram of a $G : \mathbf{P} \to \mathbf{vec}$ does not imply the interval-decomposability of G:

Example 3.21 Let $\mathbf{P}$ and $F : \mathbf{P} \to \mathbf{vec}$ respectively be the same as Example 3.20. Let G be the direct sum of F and the interval module $\mathscr{I}^{\{b\}} : \mathbf{P} \to \mathbf{vec}$ supported by $\{b\}$. Note that $\mathrm{rk}(G) = \mathrm{rk}(F) + \mathrm{rk}(\mathscr{I}^{\{b\}})$ and

$$\mathrm{dgm}^{\mathbf{P}}(G)(I) = \begin{cases} 0, & I = \{b\} \\ \mathrm{dgm}^{\mathbf{P}}(F)(I), & \text{otherwise.} \end{cases}$$

In particular, observe that $\mathrm{dgm}^{\mathbf{P}}(G)$ is non-negative even though G is not interval decomposable. Also, note that, for $H = I^{\{a,b\}} \bigoplus I^{\{b,c\}} \bigoplus I^{\{b,d\}}$, we have $\mathrm{rk}(G) = \mathrm{rk}(H)$ and in turn $\mathrm{dgm}^{\mathbf{P}}(G) = \mathrm{dgm}^{\mathbf{P}}(H)$. This shows that the generalized rank invariant and persistence diagram are *not* complete invariants beyond the class of interval decomposable persistence modules. The following is also noteworthy: for the intervals $I = \{a, b\}$, $\{b, c\}$, $\{b, d\}$, G does not admit $\mathscr{I}^I$ as a summand even though $\mathrm{dgm}^{\mathbf{P}}(G)(I) = 1$.

Remark 3.22 Given $F : \mathbf{P} \to \mathbf{vec}$, for each $K \in \mathbf{Int}(\mathbf{P})$, let $n_K := \mathrm{dgm}^{\mathbf{P}}(F)(K)$. When $\mathrm{dgm}^{\mathbf{P}}(F)$ is non-negative on $\mathbf{Int}(\mathbf{P})$, the direct sum $G = \bigoplus_{K \in \mathbf{Int}(\mathbf{P})} n_K \cdot I^K$ of interval modules could be seen as an "approximation" of F. In particular,

by Theorem 3.14 we know that $F \cong G$ whenever F is interval decomposable. This observation is connected to Asashiba et al. (2019), where the authors propose a method to approximate a diagram $F : \mathbf{R}^2 \to \mathbf{vec}$ by an interval decomposable $\delta^*(F) : \mathbf{R}^2 \to \mathbf{vec}$ whose (standard) rank invariant is the same as F. In particular, $F = \delta^*(F)$ whenever F is interval decomposable.

3.2.4 Towards a proof of Proposition 3.19

In order to prove Proposition 3.19, we will exploit the poset structure of $\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$. To this end, we briefly review the notion of *lattice*. See Birkhoff (1940) for a comprehensive treatment on this subject.

Let $\mathbf{L}$ be a poset, and let $S \subset \mathbf{L}$. An element $u \in \mathbf{L}$ is said to be an upper bound of S if $s \leq u$ for all $s \in S$. An upper bound u of S is said to be its least upper bound, or *join* if $u \leq x$ for each upper bound x of S. If a join of S exists, then it is unique. The concepts of lower bound and greatest lower bound are defined in a dual way. In particular, a greatest lower bound is said to be a *meet*.

Definition 3.23 (*Lattices*) Let $\mathbf{L}$ be a poset. $\mathbf{L}$ is said to be a *join-semi lattice* if for every pair $a, b \in \mathbf{L}$, the set $\{a, b\}$ has a join in $\mathbf{L}$. Dually, $\mathbf{L}$ is said to be a *meet-semi lattice* if for every pair $a, b \in \mathbf{L}$, the set $\{a, b\}$ has a meet in $\mathbf{L}$. If $\mathbf{L}$ is both join-semi lattice and meet-semi lattice, then $\mathbf{L}$ is said to be a *lattice*.

Lemma 3.24 ($\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$ is locally a lattice) *Let $\mathbf{P}$ be a connected, essentially finite poset such that $\mathbf{Con}(\mathbf{P})$ is locally finite. Let $I, J \in \mathbf{Con}^{\mathrm{op}}(\mathbf{P})$ with $J \supset I$. Then $[J, I] := \{K \in \mathbf{Con}^{\mathrm{op}}(\mathbf{P}) : J \supset K \supset I\}$ is a finite lattice.*

Proof By assumption, $\mathbf{Con}(\mathbf{P})$ is locally finite and thus $[J, I]$ is finite. Fix any $K_1, K_2 \in [J, I]$. Observe that $K_1 \cup K_2$ is the greatest lower bound for $\{K_1, K_2\}$ in the subposet $[J, I]$ of $\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$. Also, observe that the least upper bound for $\{K_1, K_2\}$ is the union of all subposets L in the subcollection $\{L \in \mathbf{Con}^{\mathrm{op}}(\mathbf{P}) : K_1 \cap K_2 \supset L \supset I\}$ of $[J, I]$ (note that this subcollection contains I and hence not empty). This shows that $[J, I]$ is a lattice. □

An *atom* in a poset is an element that covers a minimal element. A *dual atom* is an element that is covered by a maximal element (Rota 1964).

Proposition 3.25 (Rota 1964, Corollary (P. Hall), p. 349) *Let $\mathbf{L}$ be a finite lattice with $0, 1 \in \mathbf{L}$ such that $0 \leq l \leq 1$ for all $l \in \mathbf{L}$. If 0 is not the meet of dual atoms of $\mathbf{L}$, or if 1 is not the join of atoms, then $\mu_{\mathbf{L}}(0, 1) = 0$.*

Let $\mathbf{P}$ be a connected, essentially finite poset such that $\mathbf{Con}(\mathbf{P})$ is locally finite. Let $I, J \in \mathbf{Con}(\mathbf{P})$ with $J \supset I$. Observe that, in the finite lattice $\mathbf{L} = [J, I]$ which is a subposet of $\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$, all dual atoms belong to the first entourage I^1 of I (Notation 3.11).

Proof of Proposition 3.19 Invoking formula (3), by elementary induction, it follows that for any $n \in \{1, \ldots, o_I\}$ and for any $J \in \mathbf{Con}^{\mathrm{op}}(\mathbf{P})$ with $J \in I^n$, $\mu(J, I) = (-1)^n$.

Now pick any $J \in \mathbf{Con}^{\mathrm{op}}(\mathbf{P})$ such that $J \supset I$ and $J \notin \bigcup_{n=1}^{o_I} I^n$. This implies that J is not the union of elements in I^1 and in turn implies that J is not the meet of dual atoms of $[J, I]$. By (Rota 1964, Proposition 4 (p.345)), the restriction of the Möbius function of $\mathbf{Con}^{\mathrm{op}}(\mathbf{P})$ to $[J, I] \times [J, I]$ is equal to the Möbius function of the subposet $[J, I]$. Lemma 3.24 and Proposition 3.25 now directly imply that $\mu(J, I) = 0$. □

4 Computing the 0-th level set barcode of a Reeb graph within set-category

In this section we propose a novel method to compute the 0-th level set barcode of a Reeb graph within the category of sets (Sect. 4.2). This will be possible via a geometric interpretation of **ZZ**-indexed **set**-diagrams (Sect. 4.1). An improvement of a semicontinuity theorem by Patel will also follow thereby (Sect. 4.3).

4.1 Local analysis of Reeb graphs

4.1.1 Reeb graphs

Let $f : M \to \mathbf{R}$ be a Morse(-type) function (Definition G.1) on a manifold. The Reeb graph of f is a descriptor for the evolution of connected components of the level sets $f^{-1}(r)$ as $r \in \mathbf{R}$ varies (Reeb 1946). There have been not only numerous applications of Reeb graphs in shape analysis and visualization (Biasotti et al. 2008; Shinagawa et al. 1991) or in dynamic/high-dimensional data analysis (Buchin et al. 2015; Ge et al. 2011; Kim and Mémoli 2017, 2018a), but also a vast body of theoretical study on Reeb graphs (De Silva et al. 2016; Stefanou 2020), approximation/computation of Reeb graphs (Dey and Wang 2013; Harvey et al. 2010; Parsa 2012), metrics on Reeb graphs (Bauer et al. 2014, 2015, 2020; Carrière and Oudot 2017), and generalizations (Dey et al. 2016; Singh et al. 2007).

By a Reeb graph, we will refer to a *constructible* topological graph X equipped with a notion of "height" represented by a continuous function $f : X \to \mathbf{R}$. While we defer the rigorous definition of Reeb graphs to Appendix (Definitions G.1 and G.2), it is well-known that any diagram $M : \mathbf{ZZ} \to \mathbf{set}$ amounts to a Reeb graph and vice versa as we will see below in more detail (Curry and Patel 2020; De Silva et al. 2016).

Definition 4.1 (*Reeb-graph-realization of a zigzag diagram in* **set**) Let $M : \mathbf{ZZ} \to \mathbf{set}$. We define the *Reeb graph corresponding to M* as the pair $(\mathbf{Reeb}(M), \pi)$ consisting of a topological graph $\mathbf{Reeb}(M)$ and a map $\pi : \mathbf{Reeb}(M) \to \mathbf{R}$, described subsequently (referring to Example 4.2 may help the understanding of the description below).

1. For $(i, i) \in \mathbf{ZZ}$, let each element in $M_{(i,i)}$ become a vertex which lies over $i \in \mathbf{Z}(\subset \mathbf{R})$.
2. For $(i, i - 1) \in \mathbf{ZZ}$, let each element in $M_{(i,i-1)}$ become an edge which lies over the interval $[i - 1, i] \subset \mathbf{R}$.
3. For each comparable pair $(i, i - 1) \leq (i, i)$ (resp. $(i, i - 1) \leq (i - 1, i - 1)$) in $\mathbf{ZZ}$, the attaching map between the vertex set and the edge set is specified by $\varphi_M((i, i - 1), (i, i))$ (resp. $\varphi_M((i, i - 1), (i - 1, i - 1)))$.

Fig. 5 The Reeb graph
($\mathbf{Reeb}(M), \pi$) corresponding to
M in Example 4.2, where
$\pi : \mathbf{Reeb}(M) \to \mathbf{R}$ is the
projection map to the horizontal
real axis
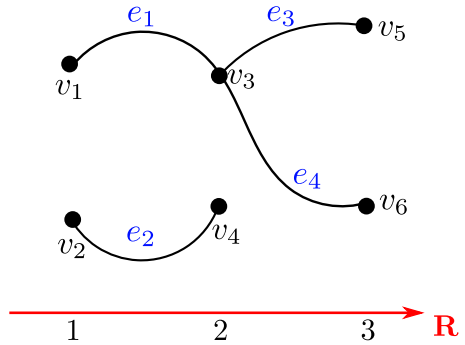
The space $\mathbf{Reeb}(M)$ is the quotient of the disjoint union of the spaces $M_{(i,i)} \times \{i\}$ and
$M_{(i,i-1)} \times [i-1, i]$ for all $i \in \mathbf{Z}$ with respect to the identifications

$$(\varphi_M((i, i-1), (i, i))(e), i) \sim (e, i) \text{ and}$$
$$(\varphi_M((i, i-1), (i-1, i-1))(e), i-1) \sim (e, i-1).$$

The map $\pi : \mathbf{Reeb}(M) \to \mathbf{R}$ is defined as the projection onto the second factor.

For any interval I of $\mathbf{ZZ}$, one can define the Reeb graph corresponding to a diagram
$I \to \mathbf{set}$ in the same way.

Example 4.2 Consider $M : \mathbf{ZZ} \to \mathbf{set}$ specified as follows:
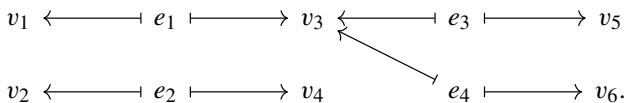
$$M_{(1,1)} = \{v_1, v_2\}, \quad M_{(2,1)} = \{e_1, e_2\}, \quad M_{(2,2)} = \{v_3, v_4\},$$
$$M_{(3,2)} = \{e_3, e_4\}, \quad M_{(3,3)} = \{v_5, v_6\},$$

and other $M_{(i,j)}$ are the empty set. The four maps

$$M_{(1,1)} \xleftarrow{\varphi_M((2,1),(1,1))} M_{(2,1)} \xrightarrow{\varphi_M((2,1),(2,2))} M_{(2,2)} \xleftarrow{\varphi_M((3,2),(2,2))} M_{(3,2)} \xrightarrow{\varphi_M((3,2),(3,3))} M_{(3,3)}$$

are defined as follows:

$$v_1 \longleftarrow\!\!\!| \; e_1 \; |\!\!\!\longrightarrow v_3 \; \xleftarrow{} \; e_3 \; |\!\!\!\longrightarrow v_5$$
$$v_2 \longleftarrow\!\!\!| \; e_2 \; |\!\!\!\longrightarrow v_4 \qquad e_4 \; |\!\!\!\longrightarrow v_6.$$

The Reeb graph corresponding to M is depicted in Fig. 5.

Definition 4.1 describes how to turn a functor $M : \mathbf{ZZ} \to \mathbf{set}$ into a Reeb graph. In
this respect, we will refer to any $M : \mathbf{ZZ} \to \mathbf{set}$ as a *Reeb graph*. On the other hand,
any Reeb graph can be expressed as a **set**-valued zigzag diagram over $\mathbf{R}$ (Curry and
Patel 2020; De Silva et al. 2016). This zigzag diagram over $\mathbf{R}$ is in turn equivalent to
a $\mathbf{ZZ}$-index **set**-diagram up to rescaling— this idea has already implicitly appeared in
Botnan and Lesnick (2018), Curry (2013). This means that the entire combinatorial

information of a Reeb graph X can be encoded into a certain $M : \mathbf{ZZ} \to \mathbf{set}$. In particular, the 0-th level set barcode of X (Carlsson et al. 2009) can be extracted from M up to rescaling of intervals.

4.1.2 Local analysis of Reeb graphs

We illustrate how to extract topological/combinatorial information from limits and colimits of interval restrictions of a Reeb graph $M : \mathbf{ZZ} \to \mathbf{set}$. For $k, l \in \mathbf{ZZ}$, assume that $x \in M_k$ and $y \in M_l$. We write $x \sim y$ if k and l are comparable and one of x and y is mapped to the other via the internal map between M_k and M_l.

We can explicitly represent the limit and colimit of M as follows.

(i) The limit of M is the pair $(L, (\pi_k)_{k \in \mathbf{ZZ}})$, where

$$L := \left\{ (x_k)_{k \in \mathbf{ZZ}} \in \prod_{k \in \mathbf{ZZ}} M_k : \text{ for comparable } k, l \in \mathbf{ZZ}, x_k \sim x_l \right\},$$

and each $\pi_k : L \to M_k$ is the canonical projection.

(ii) The colimit of M is the pair $(C, (i_k)_{k \in \mathbf{ZZ}})$ described as follows:

$$C := \left(\coprod_{k \in \mathbf{ZZ}} M_k \right) / \approx, \tag{4}$$

where $\approx$ is the equivalence relation generated by the relations $x_k \sim x_l$ for $x_k \in M_k$ and $x_l \in M_l$ with k, l being comparable. For the quotient map $q : \coprod_{k \in \mathbf{ZZ}} M_k \to C$, each i_k is the composition $M_k \hookrightarrow \coprod_{k \in \mathbf{ZZ}} M_k \xrightarrow{q} C$.

Let $I \in \mathbf{Int}(\mathbf{ZZ})$. For any diagram $N : I \to \mathbf{set}$, we can construct the limit and colimit of N in the same way; namely, in items (i) and (ii) above, replace M and $\mathbf{ZZ}$ by N and I, respectively. **In what follows, we use those explicit constructions whenever considering limits and colimits of (interval restrictions of) ZZ-indexed set-diagrams.**

Definition 4.3 (*Supports and full components*) Let $I \in \mathbf{Int}(\mathbf{ZZ})$ and let $N : I \to \mathbf{set}$. Let $c \in \varinjlim N$. We define the *support* of c as

$$\mathrm{supp}(c) := \{k \in I : \exists x_k \in N_k, \ i_k(x_k) = c\}.$$

In particular, if $\mathrm{supp}(c) = I$, we call c a *full component* of N.

Now, extending ideas in De Silva et al. (2016), we geometrically interpret elements of limits, colimits and images of the canonical LC maps of (interval restrictions of) **ZZ**-indexed **set**-diagrams.

Recall that, for $b, d \in \mathbf{Z}$ with $b < d$, $\langle b, d \rangle_{\mathbf{ZZ}}$ stands for an interval of $\mathbf{ZZ}$ (Notation 2.13). Similarly, for $b, d \in \mathbf{R}$ with $b < d$, $\langle b, d \rangle$ will stand for one of the real intervals (b, d), $[b, d]$, $(b, d]$, $[b, d)$. The terminology introduced in Definition 4.4 below will be illustrated in Example 4.5.

Definition 4.4 (*Local analysis of a Reeb graph*) Consider any $M : \mathbf{ZZ} \to \mathbf{set}$ and fix $I := \langle b, d \rangle_{\mathbf{ZZ}} \in \mathbf{Int}(\mathbf{ZZ})$.

(i) Each element in $\varinjlim M|_I$ is said to be a $\langle b, d \rangle$-component.
(ii) Each element in $\varprojlim M|_I$ is said to be a $\langle b, d \rangle$-section.[7]
(iii) If $c \in \varinjlim M|_I$ and $\mathrm{supp}(c) = I$, then c is said to be a $\langle b, d \rangle$-*full*-component.
(iv) Each element in the image of the canonical LC map $\psi_{M|_I} : \varprojlim M|_I \to \varinjlim M|_I$ is said to be a $\langle b, d \rangle$-full-component of *M with a section*.

Example 4.5 (*Geometric interpretation of Definition* 4.4) For $M : \mathbf{ZZ} \to \mathbf{set}$ in Example 4.2 (Fig. 5), note the following:

(i) There are two [1, 3]-components: The connected component containing the vertex v_1 and the one containing v_2. Also, there are two [1, 2]-components: these components correspond to the restrictions of the previous two components to $\pi^{-1}([1, 2])$.
(ii) The 5-tuples $(v_1, e_1, v_3, e_3, v_5)$ and $(v_1, e_1, v_3, e_4, v_6)$ are [1, 3]-sections. These tuples deserve to be called by such a name since there indeed exist continuous maps $s_1 : [1, 3] \to \mathbf{Reeb}(M)$ and $s_2 : [1, 3] \to \mathbf{Reeb}(M)$ satisfying (1) $\pi \circ s_1 = \pi \circ s_2 = \mathrm{id}_{[1,3]}$, and (2) $\mathrm{im}(s_1), \mathrm{im}(s_2)$ lie on $v_1 e_1 v_3 e_3 v_5$ and $v_1 e_1 v_3 e_4 v_6$ respectively in $\mathbf{Reeb}(M)$.
(iii) There exists only one [1, 3]-full-component; among the two [1, 3]-components considered in item (i), the one containing v_1 is the unique [1, 3]-full-component.
(iv) Via the canonical LC map, each [1, 3]-section is mapped to the unique [1, 3]-full-component, which contains the image of the [1, 3]-section. This demonstrates why the connected component containing v_1 is called a [1, 3]-full-component of *M with a section (cf. Example* 4.12).

4.2 Computing the 0-th level set barcode of a Reeb graph without homology

In this section we propose a novel method to compute the 0-th level set barcode of a Reeb graph within the category of sets.

4.2.1 The 0-th level set barcode of a Reeb graph

The 0-th level set barcode of a Reeb graph (Definition 4.7) captures the lifetime of all homological features in the Reeb graph (Bauer et al. 2014). Whereas measuring the interleaving distance (or equivalently functional distortion distance) between Reeb graphs is not easy De Silva et al. (2016, Sect. 5), (Bauer et al. 2015), the computation of the bottleneck distance between the 0-th level set barcodes of Reeb graphs can be efficiently carried out (Kerber et al. 2017). Moreover, the bottleneck distance between the 0-th level set barcodes of Reeb graphs is a tight lower bound for the interleaving distance between Reeb graphs up to multiplicative constant 2 (Bauer et al. 2014; Botnan and Lesnick 2018).

---

[7] A right-inverse of a morphism is called a section (Mac Lane 2013). See Example 4.5 (ii).

Definition 4.6 (*Linearization functor*) Let $L_{\mathbb{F}} : \textbf{Set} \rightarrow \textbf{Vec}$ be the *linearization functor* (a.k.a. free functor): For any set S, $L_{\mathbb{F}}(S)$ consists of formal linear combination $\sum_i a_i s_i, (a_i \in \mathbb{F}, s_i \in S)$ of finite terms of elements in S over the field $\mathbb{F}$. Also, given a set map $f : S \rightarrow T$, $L_{\mathbb{F}}(f)$ is the linear map from $L_{\mathbb{F}}(S)$ to $L_{\mathbb{F}}(T)$ obtained by linearly extending f.

Throughout this section, we identify both of the Grothendieck groups $\mathscr{A}(\textbf{vec})$ and $\mathscr{A}(\textbf{set})$ with the integer group $(\textbf{Z}, +)$ (Remark 2.4). Let us recall the 0-th level set barcode of a Reeb graph (Carlsson et al. 2009; Botnan and Lesnick 2018):

Definition 4.7 (*The 0-th level set barcode of a Reeb graph*) Given $M : \textbf{ZZ} \rightarrow \textbf{set}$, consider $L_{\mathbb{F}} \circ M : \textbf{ZZ} \rightarrow \textbf{vec}$. The 0-th level set barcode of M refers to $\text{barc}^{\textbf{ZZ}}(L_{\mathbb{F}} \circ M)$.

4.2.2 New method for computing the 0-th level set barcode of a Reeb graph

In order to establish a method to compute the 0-th level set barcode of a Reeb graph *without matrix operations* (Carlsson et al. 2009; Milosavljević et al. 2011), we introduce a new invariant for Reeb graphs.

Given $M : \textbf{ZZ} \rightarrow \textbf{set}$ and $I \in \textbf{Int}(\textbf{ZZ})$, we denote the number of *full components* of $M|_I$ by $\text{full}(M|_I)$ (Definition 4.3).

Definition 4.8 (*Full function of a Reeb graph*) Given $M : \textbf{ZZ} \rightarrow \textbf{set}$, the *full function* $\text{Full}(M) : \textbf{Int}(\textbf{ZZ}) \rightarrow \textbf{Z}_{\geq 0}$ is defined as $I \mapsto \text{full}(M|_I)$.

One can observe that $\text{Full}(M)(\langle b, d \rangle_{\textbf{ZZ}})$ is equal to the number of connected components of $\pi^{-1}\langle b, d \rangle (\subset \textbf{Reeb}(M))$ *whose images via π are* $\langle b, d \rangle$. For instance, consider $M : \textbf{ZZ} \rightarrow \textbf{set}$ whose corresponding Reeb graph is depicted as in Fig. 6a. $\text{Full}(M)([2, 3)_{\textbf{ZZ}})$ is 2, since both of the two connected components of $\pi^{-1}[2, 3)$ cover $[2, 3)$ via π. On the other hand, $\text{Full}(M)((1, 3)_{\textbf{ZZ}})$ is 1 since only one of the connected components of $\pi^{-1}(1, 3)$ covers the entire $(1, 3)$ via π.

The following proposition is the core observation for establishing a new method to compute the 0-th level set barcodes of Reeb graphs.

Proposition 4.9 *For any $M : \textbf{ZZ} \rightarrow \textbf{set}$,*

$$\text{Full}(M) = \text{rk}_{\textbf{vec}}(L_{\mathbb{F}} \circ M).$$

We prove Proposition 4.9 at the end of this section.

Let $I \in \textbf{Int}(\textbf{ZZ})$ be finite. Then $\text{nbd}(I)$ consists of two elements, say $\text{nbd}(I) := \{a, b\}$. Hence, we can write the first and second entourage of I as $I^1 = \{I \cup \{a\}, I \cup \{b\}\}$ and $I^2 = \{I \cup \{a, b\}\}$, respectively. By Theorems 2.12, 3.14, and Proposition 4.9, we directly have:

Corollary 4.10 *For any $M : \textbf{ZZ} \rightarrow \textbf{set}$ and any finite $I \in \textbf{Int}(\textbf{ZZ})$, the multiplicity of I in $\text{barc}^{\textbf{ZZ}}(L_{\mathbb{F}} \circ M)$ is*

$$\text{Full}(M)(I) - \text{Full}(M)(I \cup \{a\}) - \text{Full}(M)(I \cup \{b\}) + \text{Full}(M)(I \cup \{a, b\})$$

which is also equal to $\text{dgm}^{\textbf{ZZ}}_{\textbf{vec}}(L_{\mathbb{F}} \circ M)(I)$.
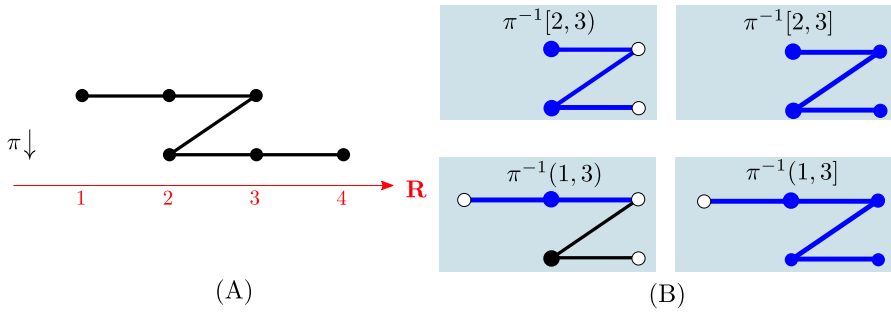
Fig. 6 **a** A Reeb graph (**Reeb**$(M), \pi$) which corresponds to M in Example 4.11. **b** Pre-images $\pi^{-1}(I)$ for 4 different choices of intervals $I \subset \mathbf{R}$. Blue components are the full components over the corresponding intervals

Note that, when $I \in \mathbf{Int}(\mathbf{ZZ})$ is infinite, the perimeter of I is either 0 or 1 and thus the formula for the multiplicity of I in barc$(L_\mathbb{F} \circ M)$ is even simpler than the one given above.

Example 4.11 Consider $M : \mathbf{ZZ} \to \mathbf{set}$ whose corresponding Reeb graph is depicted as in Fig. 6a. By Corollary 4.10, we have that

$$\mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{vec}}(L_\mathbb{F} \circ M)([2,3)_{\mathbf{ZZ}})$$
$$= \mathrm{Full}(M)([2,3)_{\mathbf{ZZ}}) - \mathrm{Full}(M)((1,3)_{\mathbf{ZZ}})$$
$$- \mathrm{Full}(M)([2,3]_{\mathbf{ZZ}}) + \mathrm{Full}(M)((1,3]_{\mathbf{ZZ}})$$
$$= 2 - 1 - 1 + 1$$
$$= 1,$$

which is the multiplicity of $[2,3)_{\mathbf{ZZ}}$ in barc$^{\mathbf{ZZ}}(L_\mathbb{F} \circ M)$. By applying the same strategy to the other intervals of $\mathbf{ZZ}$, we have:

$$\mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{vec}}(L_\mathbb{F} \circ M)(I) = \begin{cases} 1, & I \in \{[2,3)_{\mathbf{ZZ}}, (2,3]_{\mathbf{ZZ}}, [1,4]_{\mathbf{ZZ}}\}, \\ 0, & \text{otherwise,} \end{cases} \qquad (5)$$

and the barcode of $L_\mathbb{F} \circ M$ consists of the three intervals $[2,3)_{\mathbf{ZZ}}, (2,3]_{\mathbf{ZZ}}$ and $[1,4]_{\mathbf{ZZ}}$.

We remark that the Reeb graph from Example 4.11 was introduced in Adams and Carlsson (2015, Fig. 12) along the way to address the problem of searching for evasion paths in mobile sensor networks. The authors pointed out that, without careful homological considerations, one could make a wrong guess about the 0-th level set barcode of this Reeb graph (see also Curry 2013, Sect. 10). Example 4.11 demonstrates the usefulness of Corollary 4.10 in this context.

Given $M : \mathbf{ZZ} \to \mathbf{set}$, we also have another persistence diagram $\mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{set}}(M)$, which has nothing to do with the linearization functor $L_\mathbb{F}$. Next we show that $\mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{set}}(M)$ is not necessarily equal to $\mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{vec}}(L_\mathbb{F} \circ M)$:

Example 4.12 Consider $M : \mathbf{ZZ} \to \mathbf{set}$ whose corresponding Reeb graph is depicted as in Fig. 6a. By Definition 4.4 (iv), $\mathrm{rk}(M)(\langle b, d \rangle_{\mathbf{ZZ}})$ counts the number of $\langle b, d \rangle$-full components *with a section*. We claim that

$$\text{for } I \in \mathbf{Int}(\mathbf{ZZ}), \qquad \mathrm{rk}(M)(I) = \begin{cases} 0, & I \supsetneq [2, 3]_{\mathbf{ZZ}} \\ \mathrm{full}(M|_I), & \text{otherwise.} \end{cases} \tag{6}$$

Assume that $I := \langle b, d \rangle_{\mathbf{ZZ}} \in \mathbf{Int}(\mathbf{ZZ})$ strictly contains $[2, 3]_{\mathbf{ZZ}}$. Then, $\pi^{-1}\langle b, d \rangle$ has one connected component and it does not has a section, i.e. there is no map $s : \langle b, d \rangle \to \mathbf{Reeb}(M)$ such that $\pi \circ s = \mathrm{id}_{\langle b, d \rangle}$. On the other hand, if I does not strictly contain $[2, 3]_{\mathbf{ZZ}}$, then every $\langle b, d \rangle$-full-connected component has a section. These observations prove the claim in (6). Now, by Definition 3.13, it is not hard to obtain:

$$\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)(I) = \begin{cases} 1, & I = [2, 3)_{\mathbf{ZZ}} \text{ or } I = (2, 3]_{\mathbf{ZZ}} \\ -1, & I = [2, 3]_{\mathbf{ZZ}} \\ 0, & \text{otherwise.} \end{cases}$$

Now, observe that $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)$ is different from $\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ M)$ in (5).

Remark 4.13 (i) The persistence diagram of a $\mathbf{ZZ}$-indexed $\mathbf{set}$-diagram is not a complete invariant. For example, assume that $N : \mathbf{ZZ} \to \mathbf{set}$ is defined as

$$N|_{[2,3]_{\mathbf{ZZ}}} = M|_{[2,3]_{\mathbf{ZZ}}}, \qquad N_s = \emptyset \text{ for } s \notin [2, 3]_{\mathbf{ZZ}}.$$

Even though N is not isomorphic to M in Example 4.12, it is not difficult to check that $\mathrm{rk}(M) = \mathrm{rk}(N)$ and thus $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M) = \mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)$.

(ii) For every $M : \mathbf{ZZ} \to \mathbf{set}$, $\mathrm{rk}(M)(I) \leq \mathrm{rk}(L_{\mathbb{F}} \circ M)(I)$ for all $I = \langle b, d \rangle_{\mathbf{ZZ}}$. This directly follows from the fact that $\mathrm{rk}(M)(I)$ counts $\langle b, d \rangle$-full-components *with a section*, whereas $\mathrm{rk}(L_{\mathbb{F}} \circ M)(I)$ counts $\langle b, d \rangle$-full-components (Proposition 4.9).

In Sect. 4.3 we will discuss the stability of $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)$ under perturbations of $M : \mathbf{ZZ} \to \mathbf{set}$.

4.2.3 Extended persistence and a proof of Proposition 4.9

For proving Proposition 4.9, we briefly review the notion of *extended persistence*. See Cohen-Steiner et al. (2009) for details.

Let X be a topological space and let $f : X \to \mathbf{R}$ be a Morse-type function (Definition G.1) where $S := \{s_1, s_2, \ldots, s_n\}$ is the collection of critical points of f. Let us select a set of indices t_i which satisfy

$$-\infty < t_0 < s_1 < t_1 < s_2 < \cdots < t_{n-1} < s_n < t_n < \infty.$$

For $1 \le i \le j \le n$, let $X_i^j := f^{-1}([t_i, t_j])$. For $k \in \mathbf{Z}_+$, consider the following diagram of absolute and relative homology groups:

$$\mathrm{H}_k(X_0^0) \longrightarrow \mathrm{H}_k(X_0^1) \longrightarrow \cdots \longrightarrow \mathrm{H}_k(X_0^{n-1}) \longrightarrow \mathrm{H}_k(X_0^n)$$
$$\downarrow$$
$$\mathrm{H}_k(X_0^n, X_0^n) \longleftarrow \mathrm{H}_k(X_0^n, X_1^n) \longleftarrow \cdots \longleftarrow \mathrm{H}_k(X_0^n, X_{n-1}^n) \longleftarrow \mathrm{H}_k(X_0^n, X_n^n),$$

where each arrow stands for the map induced by the inclusion. The collection of pairs arising from this diagram are recorded in the so-called *k-th extended persistence diagram*. There are three different types of pairs: ordinary pairs arise between the elements in the top row of the sequence, relative pairs between the elements in the bottom row, and extended pairs span both rows. In particular, in the *k*-th extended persistence diagram,

(i) an ordinary pair of X_0^i and X_0^j ($i < j < n$) is recorded as $[s_i, s_{j+1})$,
(ii) a relative pair of (X_0^n, X_i^n) and (X_0^n, X_j^n) ($i < j < n$) is recorded as $[\bar{s}_{j+1}, \bar{s}_i)$,
(iii) an extended pair of X_0^i and (X_0^n, X_j^n) is recorded as $[s_i, \bar{s}_j)$.

Proof of Proposition 4.9 Fix $I := \langle b, d \rangle_{\mathbf{ZZ}} \in \mathbf{Int}(\mathbf{ZZ})$. We will prove that

$$\mathrm{full}(M)(I) = \mathrm{rk}_{\mathbf{vec}}(L_{\mathbb{F}} \circ M)(I). \tag{7}$$

By Proposition 3.17, $\mathrm{rk}_{\mathbf{vec}}(L_{\mathbb{F}} \circ M)(I)$ is equal to the multiplicity of I in $\mathrm{barc}^I(L_{\mathbb{F}} \circ M|_I)$. Let $(\mathbf{Reeb}(M|_I), \pi)$ be the Reeb graph corresponding to $M|_I$. By the bijection given in Carlsson et al. (2009, EP Equivalence Theorem (Table 1)), the copies of $I = \langle b, d \rangle_{\mathbf{ZZ}}$ in $\mathrm{barc}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ M|_I)$ one-to-one correspond to the copies of $[b, \bar{d})$ in the 0-th extended persistence diagram of $(\mathbf{Reeb}(M|_I), \pi)$. Also, the copies of $[b, \bar{d})$ in the 0-th extended persistence diagram of $(\mathbf{Reeb}(M|_I), \pi)$ one-to-one correspond to the connected components of $\mathbf{Reeb}(M|_I)$ whose image via π is $[b, d] \subset \mathbf{R}$ (Cohen-Steiner et al. 2009, p.83). These connected components are exactly the full components of $M|_I$ in the sense of Definition 4.3. Therefore, we have the equality in (7). □

4.3 The stability of set-persistence diagrams for merge trees and untwisted Reeb graphs

Patel's semicontinuity theorem Patel (2018) states that the so-called *type $\mathscr{A}$ persistence diagram* of $F : \mathbf{R} \to \mathscr{C}$ is stable to all *sufficiently small* perturbations of F. In this section we promote this theorem to a complete continuity theorem when $\mathscr{C} = \mathbf{set}$. Also, in turn we establish a stability result for **set**-persistence diagrams of a certain class of Reeb graphs.

Again, throughout this section, we identify both of the Grothendieck groups $\mathscr{A}(\mathbf{vec})$ and $\mathscr{A}(\mathbf{set})$ with the integer group $(\mathbf{Z}, +)$ (Remark 2.4).

4.3.1 Bottleneck stability for merge trees

One can encode all the combinatorial information of a *merge tree* (Morozov et al. 2013) into a **Z**-indexed **set**-diagram:

Definition 4.14 A functor $F : \mathbf{Z} \to \mathbf{set}$ is said to be a *(discrete) merge tree*.[8]

Recall from Sect. 4.2 that, given an $M : \mathbf{ZZ} \to \mathbf{set}$, the two persistence diagrams $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)$ and $\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ M)$ do not coincide in general. That is not the case for merge trees:

Proposition 4.15 *For any $F : \mathbf{Z} \to \mathbf{set}$, we have* $\mathrm{rk}(F) = \mathrm{rk}(L_{\mathbb{F}} \circ F)$ *and therefore,*

$$\mathrm{dgm}_{\mathbf{set}}^{\mathbf{Z}}(F) = \mathrm{dgm}_{\mathbf{vec}}^{\mathbf{Z}}(L_{\mathbb{F}} \circ F).$$

Before proving Proposition 4.15, note the following: given a set map $f : A \to B$, it directly follows from the definition of the linearization functor $L_{\mathbb{F}}$ that the cardinality of $\mathrm{im}(f)$ is equal to the rank of $L_{\mathbb{F}}(f) : L_{\mathbb{F}}(A) \to L_{\mathbb{F}}(B)$.

Proof of Proposition 4.15 Let $I := [a, b] \in \mathbf{Int}(\mathbf{Z})$. Then,

$$\mathrm{rk}(F)(I) = |\mathrm{im}(\varphi_F(a \leq b))| = \mathrm{rank}\,(L_{\mathbb{F}}(\varphi_F(a \leq b))) = \mathrm{rk}(L_{\mathbb{F}} \circ F)(I).$$

Since, $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{Z}}(F)$ (resp. $\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{Z}}(L_{\mathbb{F}} \circ F)$) is the Möbius inversion of $\mathrm{rk}(F)$ (resp. $\mathrm{rk}(L_{\mathbb{F}} \circ F)$) over the poset $(\mathbf{Int}(\mathbf{Z}), \supset) = (\mathbf{Con}(\mathbf{Z}), \supset)$ (Definition 3.13), we also have $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{Z}}(F) = \mathrm{dgm}_{\mathbf{vec}}^{\mathbf{Z}}(L_{\mathbb{F}} \circ F)$.   □

We remark that the stability of $\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{Z}}(L_{\mathbb{F}} \circ F)$ under perturbations of F was established in Morozov et al. (2013), whereas the semicontinuity of $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{Z}}(F)$ was established in Patel (2018, Theorem 8.1)[9] (we remark that both of those works use **R** as the indexing poset). Now, by virtue of Proposition 4.15, Patel (2018, Theorem 8.1) can be improved to a complete stability theorem: Let d_{B} be the bottleneck distance, and let $d_{\mathrm{I}}^{\mathscr{C}}$ be the interleaving distance between functors $\mathbf{Z} \to \mathscr{C}$ (Definitions D.2 and D.6 in Appendix). We have:

Corollary 4.16 *For $F, G : \mathbf{Z} \to \mathbf{set}$,*

$$d_{\mathrm{B}}\left(\mathrm{dgm}_{\mathbf{set}}^{\mathbf{Z}}(F), \mathrm{dgm}_{\mathbf{set}}^{\mathbf{Z}}(G)\right) \leq d_{\mathrm{I}}^{\mathbf{set}}(F, G).$$

Though this corollary directly follows from Proposition 4.15 and arguments similar to those is Morozov et al. (2013), we provide another concise proof:

---

[8] A *constructible* persistence module $F : \mathbf{R} \to \mathbf{set}$ (Definition F.1 in Appendix) is often said to be a *merge tree* (Morozov et al. 2013). In order to compute the Patel's persistence diagram of F, it suffices to consider a certain re-indexed diagram $D(F) : \mathbf{Z} \to \mathbf{set}$ of F. Such a re-indexing method is described in Appendix F (the paragraph *Re-indexing a constructible persistence module by* **Z**).

[9] This theorem states that, when $\mathscr{C}$ is essentially small symmetric monoidal category with images, the persistence diagram of $F : \mathbf{R} \to \mathscr{C}$ is stable to all *sufficiently small* perturbations of F.

Proof We have:

$$d_{\mathrm{B}}\left(\mathrm{dgm}^{\mathbf{Z}}_{\mathbf{set}}(F), \mathrm{dgm}^{\mathbf{Z}}_{\mathbf{set}}(G)\right) = d_{\mathrm{B}}\left(\mathrm{dgm}^{\mathbf{Z}}_{\mathbf{vec}}(L_{\mathbb{F}} \circ F), \mathrm{dgm}^{\mathbf{Z}}_{\mathbf{vec}}(L_{\mathbb{F}} \circ G)\right) \quad \text{by Proposition 4.15}$$

$$= d_{\mathrm{I}}^{\mathbf{vec}}(L_{\mathbb{F}} \circ F, L_{\mathbb{F}} \circ G) \quad \text{see below}$$

$$\leq d_{\mathrm{I}}^{\mathbf{set}}(F, G) \quad \text{by functoriality of } L_{\mathbb{F}}.$$

The second equality follows from the celebrated isometry theorem (Chazal et al. 2016; Lesnick 2015). $\qquad\qquad\square$

We remark that Corollary 4.16 is true even for constructible functors $F, G : \mathbf{R} \to \mathbf{set}$ (Definition F.1 in Appendix), which can be proved via trivial re-indexing of F, G by $\mathbf{Z}$.

4.3.2 Decomposition of a Reeb graph and untwisted Reeb graphs

We define *untwisted Reeb graphs*, a generalization of merge trees, and extend Corollary 4.16 to untwisted Reeb graphs.

A Reeb graph $M : \mathbf{ZZ} \to \mathbf{set}$ is said to be *empty* if $M_s = \emptyset$ for all $s \in \mathbf{ZZ}$. Let $N, N' : \mathbf{ZZ} \to \mathbf{set}$. The *disjoint union* $N \bigsqcup N' : \mathbf{ZZ} \to \mathbf{set}$ is defined as follows: for all $s \in \mathbf{ZZ}$, $(N \bigsqcup N')_s := N_s \sqcup N'_s$ and for all $s \leq t$ in $\mathbf{ZZ}$, $\varphi_{N \bigsqcup N'}(s, t) : (N \bigsqcup N')_s \to (N \bigsqcup N')_t$ is defined as $\varphi_N(s, t) \sqcup \varphi_{N'}(s, t)$. We say that a nonempty $M : \mathbf{ZZ} \to \mathbf{set}$ is *decomposable* if M is isomorphic to $N \bigsqcup N'$ for some nonempty $N, N' : \mathbf{ZZ} \to \mathbf{set}$. Otherwise, we say that M is *indecomposable*. We have similar definitions even when the indexing poset $\mathbf{ZZ}$ is replaced by any $I \in \mathbf{Int}(\mathbf{ZZ})$. In what follows, we will see that the decomposition of a nonempty $M : \mathbf{ZZ} \to \mathbf{set}$ into indecomposables parallels the topological decomposition of $\mathbf{Reeb}(M)$ (Definition 4.1) into path-connected components.

We can explicitly decompose M as follows: Let $J := \varinjlim M$, which is a nonempty set (not necessarily finite). According to the canonical construction in equation (4), each $j \in J$ is a subset of the disjoint union $\bigsqcup_{k \in \mathbf{ZZ}} M_k$. Note that $M \cong \bigsqcup_{j \in J} N^j$, where each $N^j : \mathbf{ZZ} \to \mathbf{set}$ is defined as $N^j_k = \{m \in M_k : m \in j\}$ for $k \in \mathbf{ZZ}$ and the internal morphisms of N are the canonical restrictions of the internal morphisms of M. Now we directly have:

Proposition 4.17 *Any nonempty $M : \mathbf{ZZ} \to \mathbf{set}$ decomposes into a disjoint union of indecomposables $\mathbf{ZZ} \to \mathbf{set}$, and the decomposition is unique up to a permutation of summands.*

Remark 4.18 (i) Recall that a nonempty $M : \mathbf{ZZ} \to \mathbf{set}$ gives rise to the Reeb graph $(\mathbf{Reeb}(M), \pi)$ (Definition 4.1). Given an indecomposable decomposition $M \cong \bigsqcup_{j \in J} N^j$, the path-connected components $(\mathbf{Reeb}(M), \pi)$ are precisely $\mathbf{Reeb}(N^j)$, $j \in J$ equipped with the restrictions of π.

(ii) Given any nonempty $M : \mathbf{ZZ} \to \mathbf{set}$, the following are equivalent: (a) a nonempty M is indecomposable, (b) $\mathbf{Reeb}(M)$ (Definition 4.1) contains only one path-connected component, (c) $\varinjlim M$ is a singleton.

Let $M : \mathbf{ZZ} \to \mathbf{set}$ be nonempty indecomposable. Then, $\varinjlim M$ is a singleton by Remark 4.18 (ii). By supp(M), we denote the support of the unique element of $\varinjlim M$ (Definition 4.3).

Definition 4.19 (*Untwisted Reeb graphs*) A nonempty $M : \mathbf{ZZ} \to \mathbf{set}$ is said to be *untwisted* if the following holds: For $I \in \mathbf{Int}(\mathbf{ZZ})$ such that $M|_I$ is nonempty, if $M|_I \cong \coprod_{j \in J} N^j$ for some indexing set J and nonempty indecomposables $N^j : I \to \mathbf{set}$, then $\varprojlim N^j|_{\mathrm{supp}(N^j)} \neq \emptyset$ for each $j \in J$.

Examples of untwisted Reeb graphs include M in Example 4.2 and merge trees,[10] while non-examples include the Reeb graph of Fig. 6a. Let us characterize untwisted Reeb graphs:

Proposition 4.20 *Let $M : \mathbf{ZZ} \to \mathbf{set}$ be nonempty and let $(\mathbf{Reeb}(M), \pi)$ be the corresponding Reeb graph (Definition 4.1). The following are equivalent:*

(i) M is untwisted.
(ii) $\mathrm{rk}(M) = \mathrm{rk}(L_{\mathbb{F}} \circ M)$ (and thus, $\mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{set}}(M) = \mathrm{dgm}^{\mathbf{ZZ}}_{\mathbf{vec}}(L_{\mathbb{F}} \circ M)$).
(iii) For each real interval $\langle b, d \rangle$ such that $\pi^{-1}(\langle b, d \rangle)$ is nonempty, the restriction of π to each connected component of $\pi^{-1}(\langle b, d \rangle)$ has a section.

Proof We will utilize geometric insight of Example 4.5 and Remark 4.18 throughout the proof.

(i)⇒(iii): It suffices to consider an interval $\langle b, d \rangle \subset \mathbf{R}$ such that $b, d \in \mathbf{Z} \cup \{-\infty, \infty\}$ and $\pi^{-1}(\langle b, d \rangle)$ is nonempty. Let $I := \langle b, d \rangle_{\mathbf{ZZ}}$ and consider the restricted diagram $M|_I$. Note that the assumption $\pi^{-1}(\langle b, d \rangle) \neq \emptyset$ implies that $M|_I$ is nonempty. Now assume that $M|_I$ has an indecomposable decomposition $M|_I \cong \coprod_{j \in J} N^j$. Then $\mathbf{Reeb}(N^j)$ are exactly the connected components of $\mathbf{Reeb}(M|_I)$ (Remark 4.18 (i)). Also, there exists a canonical bijection from $\varprojlim N^j|_{\mathrm{supp}(N^j)}$ to the sections of the restriction $\pi|_{\mathbf{Reeb}(N^j)}$ (cf. Example 4.5 (ii)). For each $j \in J$, since $\varprojlim N^j|_{\mathrm{supp}(N^j)} \neq \emptyset$, $\pi|_{\mathbf{Reeb}(N^j)}$ has at least one section.

(iii)⇒(ii): Fix $I = \langle b, d \rangle_{\mathbf{ZZ}}$. By assumption, every $\langle b, d \rangle$-full-component has a section. Invoking that $\mathrm{rk}(L_{\mathbb{F}} \circ M)(I)$ is the number of $\langle b, d \rangle$-full-components (Proposition 4.9), whereas $\mathrm{rk}(M)(I)$ stands for the number of $\langle b, d \rangle$-full-components *with a section* (Definition 4.4 (iv) and Example 4.5 (iv)), we have $\mathrm{rk}(M)(I) = \mathrm{rk}(L_{\mathbb{F}} \circ M)(I)$.

(ii)⇒(i): Fix any $I \in \mathbf{Int}(\mathbf{ZZ})$ such that $M|_I$ is nonempty and assume that $M|_I \cong \coprod_{j \in J} N^j$ for some indexing set J and nonempty indecomposables $N^j : I \to \mathbf{set}$.

Fix $j \in J$ and let $K := \mathrm{supp}(N^j) = \langle b, d \rangle_{\mathbf{ZZ}}$. By assumption, the number $\mathrm{rk}(M)(K)$ of $\langle b, d \rangle$-full-components of $\mathbf{Reeb}(M)$ is equal to the number $\mathrm{rk}(L_{\mathbb{F}} \circ M)(K)$ of $\langle b, d \rangle$-full-components of $\mathbf{Reeb}(M)$ *which have a section*. Therefore, every $\langle b, d \rangle$-full-component has a section. By the choice of K, $\mathbf{Reeb}(N^j)$ is a $\langle b, d \rangle$-full-component and thus $\mathbf{Reeb}(N^j)$ has a section. This means that $\varprojlim N^j|_K \neq \emptyset$, as desired. □

---

[10] Every $\mathbf{Z}$-indexed $\mathbf{set}$-diagram F can be converted into a $\mathbf{ZZ}$-indexed diagram $D(F)$ which contains the same combinatorial information as F (refer to the paragraph *Re-indexing $\mathbf{ZR}$-indexed diagram by $\mathbf{ZZ}$* in Appendix F). In this respect, every merge tree can be viewed as a Reeb graph.

4.3.3 Bottleneck stability for untwisted Reeb graphs

We extend the stability result in Corollary 4.16 to a class of *untwisted Reeb graphs*.

An *interleaving* distance $d_I^{\mathscr{C}}$ between zigzag modules $\mathbf{ZZ} \to \mathscr{C}$ is defined when $\mathscr{C}$ is cocomplete (Definition D.4 in Appendix). The following theorem can be proved in the same way as Corollary 4.16 except utilizing the algebraic stability of zigzag modules (Theorem D.8 in Appendix) in lieu of the isometry theorem.

Theorem 4.21 *For any untwisted Reeb graphs* $M, N : \mathbf{ZZ} \to \mathbf{set}$,

$$d_B\left(\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M), \mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)\right) \leq 2 \cdot d_I^{\mathbf{Set}}(M, N). \tag{8}$$

Proof We have:

$$\begin{aligned}
d_B\left(\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M), \mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)\right) &= d_B\left(\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ M), \mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ N)\right) \quad \text{by Proposition 4.20} \\
&\leq 2 \cdot d_I^{\mathbf{Vec}}(L_{\mathbb{F}} \circ M, L_{\mathbb{F}} \circ N) \quad \text{by Theorem D.8} \\
&\leq 2 \cdot d_I^{\mathbf{Set}}(M, N) \quad \text{by functoriality of } L_{\mathbb{F}}.
\end{aligned}$$

$\square$

We remark that the inequality in (8) is tight: $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)$ and $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)$ are equal to $\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ M)$ and $\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ N)$ respectively and the tightness of the following inequality is known (Botnan and Lesnick 2018):

$$d_B\left(\mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ M), \mathrm{dgm}_{\mathbf{vec}}^{\mathbf{ZZ}}(L_{\mathbb{F}} \circ N)\right) \leq 2 \cdot d_I^{\mathbf{set}}(M, N).$$

We provide a concrete example of which the both sides in (8) coincide:

Example 4.22 (*Tightness*) Let $M : \mathbf{ZZ} \to \mathbf{set}$ be defined as

$$M_{(0,0)} = \{v_1\}, \ M_{(1,1)} = \{v_2\}, \ M_{(1,0)} = \{e_1, e_2\}$$

and any other $M_{(i,j)}$ is the empty set. The maps $\varphi_M((1,0),(0,0)), \varphi_M((1,0),(1,1))$ are the unique surjections See Fig. 7a. One can check that $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)(I) = 1$ if $I = [1,2]_{\mathbf{ZZ}}$ or $I = (1,2)_{\mathbf{ZZ}}$, and $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)(I) = 0$ otherwise.

On the other hand, define $N : \mathbf{ZZ} \to \mathbf{set}$ as

$$N_{(0,0)} = \{v_1\}, \ N_{(1,1)} = \{v_2\}, \ N_{(1,0)} = \{e_1\}$$

and any other $N_{(i,j)}$ is the empty set. The maps $\varphi_N((1,0),(0,0)), \varphi_N((1,0),(1,1))$ are the unique bijections. See Fig. 7b. One can check that $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)(I) = 1$ if $I = [1,2]_{\mathbf{ZZ}}$ and $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)(I) = 0$ otherwise. Checking that

$$d_B\left(\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M), \mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(N)\right) = 1/2, \ \text{and} \ d_I^{\mathbf{set}}(M, N) = 1/4,$$
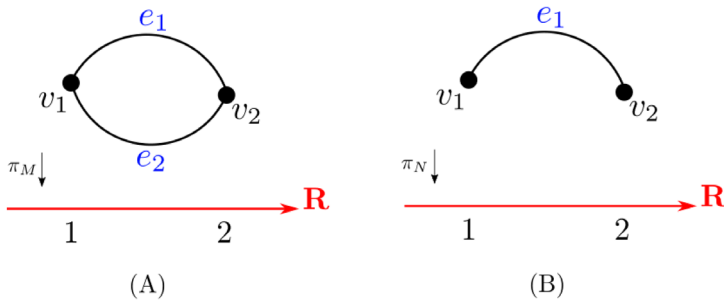
demonstrates the tightness of the inequality in (8).

Fig. 7 Illustration for Example 4.22. **a** The Reeb graph corresponding to M. **b** The Reeb graph corresponding to N. Observe that for every interval $I \in \mathbf{R}$, the preimages $\pi_M^{-1}(I)$ and $\pi_N^{-1}(I)$ consist solely of connected components which allow a section. This demonstrate that M and N are untwisted

5 Discussion

We have extended the notions of rank invariant and generalized persistence diagram due to Patel (2018) to the setting of generalized persistence modules $\mathbf{P} \to \mathscr{C}$, when $\mathbf{P}$ and $\mathscr{C}$ satisfy mild assumptions. The rank invariant and persistence diagram defined in this paper generalize those in Botnan and Lesnick (2018), Carlsson et al. (2009), Carlsson and Zomorodian (2009), Cohen-Steiner et al. (2007), Patel (2018), Zomorodian and Carlsson (2005). In particular, our construction of the rank invariant yields a complete invariant for interval decomposable persistence modules $\mathbf{P} \to \mathbf{vec}$ (which include the case of zigzag modules). Along the way, the 0-th level set barcode of a Reeb graph has been interpreted with a novel viewpoint. This leads to the promotion of a Patel's theorem and opens up the possibility of the existence of another efficient algorithm for computing the 0-th level set barcode of a Reeb graph or a Morse-type function.

Many questions remain unanswered:

(1) To what extent can we generalize stability results about persistence diagrams/barcodes beyond those results in Sect. 4.3 and those in Botnan and Lesnick (2018), Chazal et al. (2009), Cohen-Steiner et al. (2007), McCleary and Patel (2020), Patel (2018)? In particular, for arbitrary $M : \mathbf{ZZ} \to \mathbf{set}$: how to address the stability of $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)$ given that $\mathrm{dgm}_{\mathbf{set}}^{\mathbf{ZZ}}(M)$ could take negative values?[11]

(2) For persistence modules $\mathbf{P} \to \mathbf{vec}$ which are *not* interval decomposable, how faithful is the rank invariant?

(3) We wonder whether interval modules can be characterized in terms of persistence diagrams as follows. Given an *indecomposable* $F : \mathbf{P} \to \mathbf{vec}$, if $\mathrm{dgm}^{\mathbf{P}}(F)$ vanishes on $\mathbf{Con}(\mathbf{P}) \setminus \mathbf{Int}(\mathbf{P})$ and has non-negative values on $\mathbf{Int}(\mathbf{P})$: is F an interval module? This question is motivated from Example 3.21.

(4) In order to compute level set barcodes of Morse-type functions, algorithms in Carlsson et al. (2009), Milosavljević et al. (2011) make use of a sequence of

---

[11] We saw such a case in Example 4.12. Non-positive persistence diagrams also appear in Betthauser et al. (2021).

operations on matrices. Utilizing the results in Sect. 4.2, can we establish more efficient algorithms for specifically computing 0-th level set barcodes?

(5) An $\mathbf{R}^d$-indexed persistence module M can often be encoded as a $\mathbf{P}$-indexed module F_M for some connected finite poset $\mathbf{P}$; see Miller (2020, Sect. 1.3). Small perturbations to M in the interleaving distance (Definition D.2) are expected to result in small variations of $\mathrm{dgm}^{\mathbf{P}}(F_M)$. How can we quantify the *stability* of the assignment $M \mapsto \mathrm{dgm}^{\mathbf{P}}(F_M)$? Answering this question will enable us to utilize $\mathrm{dgm}^{\mathbf{P}}(F_M)$ in statistical studies of multiparameter persistence modules (e.g. efficient clustering methods for multiparameter persistence modules).

Note added in proof: See Section H of the arXiv version of this paper for a result establishing the stability of the generalized rank invariant (Definition 3.5) with respect to the interleaving distance.

Declaration

A Limits and Colimits

We recall the notions of limit and colimit (Mac Lane 2013). Throughout this section I will stand for a small category.

Definition A.1 (*Cone*) Let $F : I \rightarrow \mathscr{C}$ be a functor. A *cone* over F is a pair $\left(L, (\pi_x)_{x \in \mathrm{ob}(I)}\right)$ consisting of an object L in $\mathscr{C}$ and a collection $(\pi_x)_{x \in \mathrm{ob}(I)}$ of morphisms $\pi_x : L \rightarrow F(x)$ that commute with the arrows in the diagram of F, i.e. if $g : x \rightarrow y$ is a morphism in I, then $\pi_y = F(g) \circ \pi_x$ in $\mathscr{C}$, i.e. the diagram below commutes.

$$
\begin{array}{ccc}
F(x) & \xrightarrow{\ F(g)\ } & F(y) \\
& \underset{\pi_x}{\nwarrow} \quad \underset{\pi_y}{\nearrow} & \\
& L &
\end{array}
$$

In Definition A.1, the cone $\left(L, (\pi_x)_{x \in \mathrm{ob}(I)}\right)$ over F will sometimes be denoted simply by L, suppressing the collection $(\pi_x)_{x \in \mathrm{ob}(I)}$ of morphisms if no confusion can arise. A limit of a diagram $F : I \rightarrow \mathscr{C}$ is a terminal object in the collection of all cones:

Definition A.2 (*Limit*) Let $F : I \to \mathscr{C}$ be a functor. A *limit* of F is a cone over F, denoted by $\left(\varprojlim F, \ (\pi_x)_{x \in \mathrm{ob}(I)} \right)$ or simply $\varprojlim F$, with the following *terminal* property: If there is another cone $\left(L', (\pi'_x)_{x \in \mathrm{ob}(I)} \right)$ of F, then there is a *unique* morphism $u : L' \to \varprojlim F$ such that $\pi'_x = \pi_x \circ u$ for all $x \in \mathrm{ob}(I)$.

Remark A.3 It is possible that a diagram does not have a limit at all. However, if a diagram does have a limit then the terminal property of the limit guarantees its uniqueness up to isomorphism. For this reason, we will sometimes refer to a limit as *the* limit of a diagram.

Cocones and colimits are defined in a dual manner:

Definition A.4 (*Cocone*) Let $F : I \to \mathscr{C}$ be a functor. A *cocone* over F is a pair $\left(C, (i_x)_{x \in \mathrm{ob}(I)} \right)$ consisting of an object C in $\mathscr{C}$ and a collection $(i_x)_{x \in \mathrm{ob}(I)}$ of morphisms $i_x : F(x) \to C$ that commute with the arrows in the diagram of F, i.e. if $g : x \to y$ is a morphism in I, then $i_x = i_y \circ F(g)$ in $\mathscr{C}$, i.e. the diagram below commutes.

$$
\begin{array}{ccc}
 & C & \\
 i_x \nearrow & & \nwarrow i_y \\
 F(x) & \xrightarrow[\ F(g)\ ]{} & F(y)
\end{array}
$$

In Definition A.4, a cocone $\left(C, (i_x)_{x \in \mathrm{ob}(I)} \right)$ over F will sometimes be denoted simply by C, suppressing the collection $(i_x)_{x \in \mathrm{ob}(I)}$ of morphisms. A colimit of a diagram $F : I \to \mathscr{C}$ is an initial object in the collection of cocones over F:

Definition A.5 (*Colimit*) Let $F : I \to \mathscr{C}$ be a functor. A *colimit* of F is a cocone, denoted by $\left(\varinjlim F, \ (i_x)_{x \in \mathrm{ob}(I)} \right)$ or simply $\varinjlim F$, with the following *initial* property: If there is another cocone $\left(C', (i'_x)_{x \in \mathrm{ob}(I)} \right)$ of F, then there is a *unique* morphism $u : \varinjlim F \to C'$ such that $i'_x = u \circ i_x$ for all $x \in \mathrm{ob}(I)$.

Remark A.6 It is possible that a diagram does not have a colimit at all. However, if a diagram does have a colimit then the initial property of the colimit guarantees its uniqueness up to isomorphism. For this reason, we will sometimes refer to a colimit as *the* colimit of a diagram.

Remark A.7 (*Restriction of an indexing poset*) Let $\mathbf{P}$ be any poset and let $\mathbf{Q}$ be a subposet of $\mathbf{P}$. In categorical language, $\mathbf{Q}$ is a full subcategory of $\mathbf{P}$. Let $F : \mathbf{P} \to \mathscr{C}$ be a functor.

(i) Assume that the limit of the restriction $F|_{\mathbf{Q}}$ exists. For any cone $\left(L', (\pi'_p)_{p \in \mathbf{P}} \right)$ over F, its restriction $\left(L', (\pi'_p)_{p \in \mathbf{Q}} \right)$ is a cone over the restriction $F|_{\mathbf{Q}} : \mathbf{Q} \to \mathscr{C}$. Therefore, by the terminal property of the limit $\left(\varprojlim F|_{\mathbf{Q}}, (\pi_q)_{q \in \mathbf{Q}} \right)$, there exists the unique morphism $u : L' \to \varprojlim F|_{\mathbf{Q}}$ such that $\pi'_q = \pi_q \circ u$ for all $q \in \mathbf{Q}$.

(ii) Assume that the colimit of the restriction $F|_\mathbf{Q}$ exists. For any cocone $\left(C', (i'_p)_{p \in \mathbf{P}}\right)$ over F, its restriction $\left(C', (i'_p)_{p \in \mathbf{Q}}\right)$ is a cocone over the restriction $F|_\mathbf{Q} : \mathbf{Q} \to \mathscr{C}'$. Therefore, by the initial property of $\varinjlim F|_\mathbf{Q}$, there exists the unique morphism $u : \varinjlim F|_\mathbf{Q} \to C'$ such that such that $i'_q = u \circ i_q$ for all $q \in \mathbf{Q}$.

B The rank invariant of a standard persistence module

In this section we review some important (standard) results about the rank invariant of one-dimensional or multidimensional persistence modules. Recall the poset $\mathbf{U}$ from Definition 2.10.

Definition B.1 (*Rank invariant of a persistence module*) Let $F : \mathbf{R} \to \mathbf{vec}$ be any persistence module. The *rank invariant* of F is defined as the map $\mathrm{rk}(F) : \mathbf{U} \to \mathbf{Z}_+$ which sends each $\mathbf{u} = (u_1, u_2) \in \mathbf{U}$ to $\mathrm{rank}\,(F(u_1 \leq u_2))$.

Remark B.2 (*Category theoretical interpretation of the rank invariant*) Let $F : \mathbf{R} \to \mathbf{vec}$ be a persistence module. For any $\mathbf{u} = (u_1, u_2) \in \mathbf{U}$, it is not difficult to check that

$$\left(F_{u_1}, (\varphi_F(u_1, t))_{t \in [u_1, u_2]}\right) \text{ and } \left(F_{u_2}, (\varphi_F(t, u_2))_{t \in [u_1, u_2]}\right)$$

are a limit and a colimit of $F|_{[u_1, u_2]}$, respectively. The canonical LC map from F_{u_1} to F_{u_2} is definitely $\varphi_F(u_1, u_2)$, which is identical to $\varphi_F(t, u_2) \circ \varphi_F(u_1, t)$, for any $t \in [u_1, u_2]$. Therefore, $\mathrm{rk}(F)(\mathbf{u})$ can be regarded as the rank of the canonical LC map of $F|_{[u_1, u_2]}$.

Remark B.3 In Definition B.1, $\mathrm{rk}(F)(\mathbf{u})$ for $\mathbf{u} = (u_1, u_2)$ counts all the persistence features of the persistence module F which are born before or at u_1 and die after u_2. Also, when $u_1 = u_2$, $\mathrm{rk}(M)(\mathbf{u})$ is the dimension of F_{u_1}.

Remark B.4 (*Rank invariant is order-reversing*) In Definition B.1, for any pair $\mathbf{u} = (u_1, u_2) \leq \mathbf{u}' = \left(u'_1, u'_2\right)$ in $\mathbf{U}$, since

$$\varphi_F\left(u'_1, u'_2\right) = \varphi_F\left(u_2, u'_2\right) \circ \varphi_F(u_1, u_2) \circ \varphi_F\left(u'_1, u_1\right),$$

it holds that $\mathrm{rk}(F)\left(\mathbf{u}'\right) \leq \mathrm{rk}(F)(\mathbf{u})$. Therefore, the map $\mathrm{rk}(F) : \mathbf{U} \to \mathbf{Z}_+$ is an order-reversing map. This result generalizes to McCleary and Patel (2020, Proposition 4.4). Also, see Puuska (2020, Proposition 3.7).

Theorem B.5 (Completeness of the rank invariant for one-dimensional modules Carlsson and Zomorodian 2009) *The rank invariant defined in Definition B.1 is a complete invariant for one-dimensional persistence modules, i.e. if there are two* constructible *persistence modules $F, G : \mathbf{R} \to \mathbf{vec}$ such that $\mathrm{rk}(F) = \mathrm{rk}(G)$, then F and G are isomorphic (see Definition F.1 for the meaning of* constructible*).*

The rank invariant can also be defined for multidimensional modules $F : \mathbf{R}^n \to \mathbf{vec}$, $n > 1$: For any pair $\mathbf{a} \leq \mathbf{b}$ in $\mathbf{R}^n$, let $\mathrm{rk}(F)(\mathbf{a}, \mathbf{b}) := \mathrm{rank}\,(\varphi_F(\mathbf{a}, \mathbf{b}))$. This

defines a function from the set $\{(\mathbf{a}, \mathbf{b}) \in \mathbf{R}^n \times \mathbf{R}^n : \mathbf{a} \le \mathbf{b}\}$ to $\mathbf{Z}_+$. However, the map $\mathrm{rk}(F)$ is not a complete invariant for multidimensional modules, i.e. for any $n > 1$, there exists a pair of persistence modules $F, G : \mathbf{R}^n \to \mathbf{vec}$ that are not isomorphic but $\mathrm{rk}(M) = \mathrm{rk}(N)$ (Carlsson and Zomorodian 2009).

C Comparison with Ville Puuska's rank invariant

In (Puuska 2020), Ville Puuska considers the set

$$\mathrm{Dgm}_\mathbf{P} := \{(a, b) \in \mathbf{P} \times \mathbf{P} : a < b\}$$

and defines the rank invariant of a functor $F : \mathbf{P} \to \mathscr{C}$ as the map $dF : \mathrm{Dgm}_\mathbf{P} \to \mathscr{C}$ sending (a, b) to $\mathrm{im}\,(\varphi_F(a, b)) \in \mathrm{ob}(\mathscr{C})$. Even though this definition is a straightforward generalization of the rank invariant of Carlsson and Zomorodian (2009), when $\mathbf{P} = \mathbf{ZZ}$ and $\mathscr{C} = \mathbf{vec}$, this definition is not anywhere near a complete invariant of $F : \mathbf{P} \to \mathscr{C}$. Namely, there exists a pair of zigzag modules M, N such that $d_\mathrm{I}(M, N) = +\infty$ (Definition D.4) whereas $dM \cong dN$:

Example C.1 Consider the two zigzag modules $M, N : \mathbf{ZZ} \to \mathbf{vec}$ defined as follows: $M := I^{(-\infty, \infty)\mathbf{zz}}$ and N is defined as

$$\begin{aligned}
N_{(i,i)} &= \mathbb{F}, \qquad N_{(i+1,i)} = \mathbb{F}^2, \\
\varphi_N((i, i-1), (i, i)) &= \pi_1, \\
\varphi_N((i+1, i), (i, i)) &= \pi_2,
\end{aligned}$$

where $\pi_1, \pi_2 : \mathbb{F}^2 \to \mathbb{F}$ are the canonical projections to the first and the second coordinate, respectively. Note that $dM(a, b) \cong dN(a, b) \cong \mathbb{F}$ for all $(a, b) \in \mathrm{Dgm}_\mathbf{P}$. However, it is not difficult to check that $\mathrm{barc}^{\mathbf{ZZ}}(M) = \{(-\infty, \infty)_{\mathbf{ZZ}}\}$ and $\mathrm{barc}^{\mathbf{ZZ}}(N) = \{(i, i+2)_{\mathbf{ZZ}} : i \in \mathbf{Z}\}$. This implies that $d_\mathrm{B}\left(\mathrm{barc}^{\mathbf{ZZ}}(M), \mathrm{barc}^{\mathbf{ZZ}}(N)\right) = +\infty$, and in turn $d_\mathrm{I}(M, N) = +\infty$ by Theorem D.8.

D Interleaving distance and existing stability theorems

D.1 Interleaving distance

We review the *interleaving distance* between $\mathbf{R}^d$ (or $\mathbf{Z}^d$)-indexed functors and between $\mathbf{ZZ}$-indexed functors (Botnan and Lesnick 2018; Chazal et al. 2009; Lesnick 2015).

D.1.1 Natural transformations

We recall the notion of *natural transformations* from category theory (Mac Lane 2013): Let $\mathscr{C}$ and $\mathscr{D}$ be any categories and let $F, G : \mathscr{C} \to \mathscr{D}$ be any two functors. A natural transformation $\psi : F \Rightarrow G$ is a collection of morphisms $\psi_c : F_c \to G_c$ in

$\mathscr{D}$ for all objects $c \in \mathscr{C}$ such that for any morphism $f : c \to c'$ in $\mathscr{C}$, the following diagram commutes:

$$
\begin{array}{ccc}
F_c & \xrightarrow{F(f)} & F_{c'} \\
\downarrow{\psi_c} & & \downarrow{\psi_{c'}} \\
G_c & \xrightarrow{G(f)} & G_{c'}.
\end{array}
$$

Natural transformations $\psi : F \to G$ are considered as morphisms in the category $\mathscr{D}^{\mathscr{C}}$ of all functors from $\mathscr{C}$ to $\mathscr{D}$.

D.1.2 The interleaving distance between $\mathbf{R}^d$ (or $\mathbf{Z}^d$)-indexed functors

In what follows, for any $\varepsilon \in [0, \infty)$, we will denote the vector $\varepsilon(1, \ldots, 1) \in \mathbf{R}^d$ by $\boldsymbol{\varepsilon}$. The dimension d will be clearly specified in context.

Definition D.1 ($\mathbf{v}$-*shift functor*) Let $\mathscr{C}$ be any category. For each $\mathbf{v} \in [0, \infty)^n$, the $\mathbf{v}$-shift functor $(-)(\mathbf{v}) : \mathscr{C}^{\mathbf{R}^d} \to \mathscr{C}^{\mathbf{R}^d}$ is defined as follows:

(i) (On objects) Let $F : \mathbf{R}^d \to \mathscr{C}$ be any functor. Then the functor $F(\mathbf{v}) : \mathbf{R}^d \to \mathscr{C}$ is defined as follows: For any $\mathbf{a} \in \mathbf{R}^d$,

$$
F(\mathbf{v})_{\mathbf{a}} := F_{\mathbf{a}+\mathbf{v}}.
$$

Also, for another $\mathbf{a}' \in \mathbf{R}^d$ such that $\mathbf{a} \leq \mathbf{a}'$ we define

$$
\varphi_{F(\mathbf{v})}(\mathbf{a}, \mathbf{a}') := \varphi_F \left(\mathbf{a} + \mathbf{v}, \mathbf{a}' + \mathbf{v} \right).
$$

In particular, if $\mathbf{v} = \boldsymbol{\varepsilon} \in [0, \infty)^d$, then we simply write $F(\varepsilon)$ in lieu of $F(\boldsymbol{\varepsilon})$.

(ii) (On morphisms) Given any natural transformation $\psi : F \Rightarrow G$, the natural transformation $\psi(\mathbf{v}) : F(\mathbf{v}) \Rightarrow G(\mathbf{v})$ is defined as $\psi(\mathbf{v})_{\mathbf{a}} = \psi_{\mathbf{a}+\mathbf{v}} : F(\mathbf{v})_{\mathbf{a}} \to G(\mathbf{v})_{\mathbf{a}}$ for each $\mathbf{a} \in \mathbf{R}^d$.

For any $\mathbf{v} \in [0, \infty)^d$, let $\psi_F^{\mathbf{v}} : F \Rightarrow F(\mathbf{v})$ be the natural transformation whose restriction to each $F_{\mathbf{a}}$ is the morphism $\varphi_F(\mathbf{a}, \mathbf{a} + \mathbf{v})$ in $\mathscr{C}$. When $\mathbf{v} = \boldsymbol{\varepsilon}$, we denote $\psi_F^{\mathbf{v}}$ simply by ψ_F^{ε}.

Definition D.2 ($\mathbf{v}$-*interleaving between* $\mathbf{R}^d$-*indexed functors*) Let $\mathscr{C}$ be any category. Given any two functors $F, G : \mathbf{R}^d \to \mathscr{C}$, we say that they are $\mathbf{v}$-*interleaved* if there are natural transformations $f : F \Rightarrow G(\mathbf{v})$ and $g : G \Rightarrow F(\mathbf{v})$ such that

(i) $g(\mathbf{v}) \circ f = \psi_F^{2\mathbf{v}}$,
(ii) $f(\mathbf{v}) \circ g = \psi_G^{2\mathbf{v}}$.

In this case, we call (f, g) a $\mathbf{v}$-*interleaving pair*. When $\mathbf{v} = \varepsilon(1, \ldots, 1)$, we simply call (f, g) ε-*interleaving pair*. The interleaving distance between $d_{\mathrm{I}}^{\mathscr{C}}$ is defined as

$$
d_{\mathrm{I}}^{\mathscr{C}}(F, G) := \inf\{\varepsilon \in [0, \infty) : F, G \text{ are } \boldsymbol{\varepsilon}\text{-interleaved}\}, \tag{9}
$$

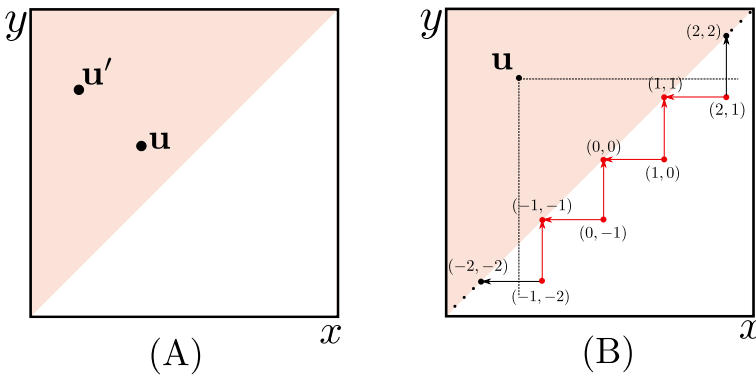**Fig. 8** **a** The shaded region stands for the poset **U**. For **u**, **u**′ ∈ **U** as marked in the figure, we have **u** ≤ **u**′. **b** Fixing **u** ∈ **U** as shown, the subposet **ZZ**[$\iota \leq$ **u**] is indicated by the red points and the red arrows

where we set $d_{\mathrm{I}}^{\mathscr{C}}(F, G) = \infty$ if there is no ε-interleaving pair between F and G for any $\varepsilon \in [0, \infty)$. Then $d_{\mathrm{I}}^{\mathscr{C}}$ is an extended pseudo-metric for $\mathscr{C}$-valued $\mathbf{R}^d$-indexed functors. **By replacing $\mathbf{R}^d$ by $\mathbf{Z}^d$ in Definitions** D.1 **and** D.2**, we similarly obtain the interleaving distance between $\mathbf{Z}^d$-indexed functors.**

Definition D.3 (*Poset* **U**) The poset $\mathbf{U} := \big\{(u_1, u_2) \in \mathbf{R}^2 : u_1 \leq u_2\big\}$ is equipped with the partial order inherited from $\mathbf{R}^{\mathrm{op}} \times \mathbf{R}$ (see Fig. 8 (A)), i.e. $(u_1, u_2) \leq (u_1', u_2')$ in **U** if and only if the interval $(u_1, u_2) \subset \mathbf{R}$ is contained in $(u_1', u_2') \subset \mathbf{R}$.

The reflection map $\mathbf{r} : \mathbf{R}^{\mathrm{op}} \to \mathbf{R}$ defined by $t \mapsto -t$ induces a poset isomorphism $\mathbf{R}^{\mathrm{op}} \times \mathbf{R} \to \mathbf{R}^2$ and this in turn induces an isomorphism $\mathscr{C}^{\mathbf{R}^2} \to \mathscr{C}^{\mathbf{R}^{\mathrm{op}} \times \mathbf{R}}$. Therefore, the notions of ε-interleaving pair and the interleaving distance $d_{\mathrm{I}}^{\mathscr{C}}$ on ob($\mathscr{C}^{\mathbf{R}^2}$) carry over to $\mathbf{R}^{\mathrm{op}} \times \mathbf{R}$-indexed or **U**-indexed functors.

D.1.3 Extension functor and the interleaving for ZZ-indexed functors

From Definition 2.10, recall the poset $\mathbf{ZZ} = \{(i, j) \in \mathbf{Z}^2 : j = i \text{ or } j = i - 1\} \subset \mathbf{R}^{\mathrm{op}} \times \mathbf{R}$. Let $\iota : \mathbf{ZZ} \hookrightarrow \mathbf{R}^{\mathrm{op}} \times \mathbf{R}$ be the canonical inclusion map. For any $\mathbf{u} = (u_1, u_2) \in \mathbf{U}$, let

$$\mathbf{ZZ}[\iota \leq \mathbf{u}] := \{a \in \mathbf{ZZ} : \iota(a) \leq \mathbf{u}\}$$

which is a subposet of **ZZ**. Observe that $\mathbf{ZZ}[\iota \leq \mathbf{u}]$ cannot be empty for any choice of $\mathbf{u} \in \mathbf{U}$. See Fig. 8b.

We review the definition of extension functor $E : \mathscr{C}^{\mathbf{ZZ}} \to \mathscr{C}^{\mathbf{U}}$ of Botnan and Lesnick (2018) for a cocomplete category $\mathscr{C}$. For any $M : \mathbf{ZZ} \to \mathscr{C}$, define the functor $\tilde{E}(M) : \mathbf{R}^{\mathrm{op}} \times \mathbf{R} \to \mathscr{C}$ as follows: For $\mathbf{a} \in \mathbf{R}^{\mathrm{op}} \times \mathbf{R}$,

$$\tilde{E}(M)(\mathbf{a}) := \varinjlim M|_{\mathbf{ZZ}[(\iota \leq \mathbf{a})]}.$$

Also, for any pair $\mathbf{a} \leq \mathbf{b}$ in $\mathbf{R}^{\mathrm{op}} \times \mathbf{R}$, since $\mathbf{ZZ}[\iota \leq \mathbf{a}]$ is a subposet of $\mathbf{ZZ}[\iota \leq \mathbf{b}]$, the linear map $\varinjlim M|_{\mathbf{ZZ}[(\iota \leq \mathbf{a})]} \to \varinjlim M|_{\mathbf{ZZ}[(\iota \leq \mathbf{b})]}$, is uniquely specified by the initial property of the colimit $\varinjlim M|_{\mathbf{ZZ}[(\iota \leq \mathbf{a})]}$ (Definition A.5 and Remark A.7). This $\tilde{E}(M)$ is called the *left Kan extension* of M along $\iota : \mathbf{ZZ} \hookrightarrow \mathbf{R}^{\mathrm{op}} \times \mathbf{R}$. Given $M, N : \mathbf{ZZ} \to \mathscr{C}$ and a natural transformation $\Gamma : M \to N$, universality of colimits also yields an induced morphism $\tilde{\Gamma} : \tilde{E}(M) \to \tilde{E}(N)$. Given any $M : \mathbf{ZZ} \to \mathscr{C}$, the functor $E(M) : \mathbf{U} \to \mathscr{C}$ is defined as the restriction $\tilde{E}(M)|_{\mathbf{U}} : \mathbf{U} \to \mathscr{C}$.

Definition D.4 (*Interleaving distance between zigzag modules* Botnan and Lesnick 2018) Let $\mathscr{C}$ be a cocomplete category. For any $M, N : \mathbf{ZZ} \to \mathscr{C}$,

$$d_{\mathrm{I}}(M, N) := d_{\mathrm{I}}^{\mathscr{C}}(E(M), E(N)).$$

D.2 Existing stability theorems

In this section we review the existing stability results from Botnan and Lesnick (2018), McCleary and Patel (2020).

D.2.1 Stability theorem of Patel

We recall McCleary and Patel (2020, Theorem 5.6):

Theorem D.5 (Stability for constructible persistence modules) *Let $\mathscr{C}$ be an essentially small, abelian[12] category. Let $F, G : \mathbf{R} \to \mathscr{C}$ be constructible (Definition F.1). Then,*

$$d_{\mathrm{B}}\left(\mathrm{dgm}_{\mathscr{C}}(F), \mathrm{dgm}_{\mathscr{C}}(G)\right) \leq d_{\mathrm{I}}(F, G).$$

We remark that the persistence diagram $\mathrm{dgm}_{\mathscr{C}}(F)$ can be computed via the following process: (1) Re-indexing of $F : \mathbf{R} \to \mathscr{C}$ to obtain $D(F) : \mathbf{Z} \to \mathscr{C}$ (Sect. F), (2) computing the persistence diagram of $D(F)$ by Definition 3.13, and (3) the rescaling of the persistence diagram (Sect. F).

D.2.2 Bottleneck distance

We regard the extended real line $\overline{\mathbf{R}} := \mathbf{R} \cup \{-\infty, \infty\}$ as a poset with the canonical order $\leq$. Let $\overline{\mathbf{U}} := \left\{(u_1, u_2) \in \overline{\mathbf{R}}^2 : u_1 \leq u_2\right\}$ be equipped with the partial order inherited from $\overline{\mathbf{R}}^{\mathrm{op}} \times \overline{\mathbf{R}}$. For $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2) \in \overline{\mathbf{U}}$, let

$$\|\mathbf{u} - \mathbf{v}\|_{\infty} := \max\left(|u_1 - v_1|, |u_2 - v_2|\right).$$

We can quantify the difference between two persistence diagrams or two barcodes of real intervals, using the bottleneck distance (Cohen-Steiner et al. 2007):

---

[12] A category $\mathscr{C}$ is *abelian* if morphisms and objects in $\mathscr{C}$ can be "added" and *kernels* and *cokernels* exist in $\mathscr{C}$ with some desirable properties. See Mac Lane (2013, p.198) for the precise definition.

Definition D.6 (*The bottleneck distance*) Let X_1, X_2 be multisets of points in $\overline{\mathbf{U}}$. Let $\alpha : X_1 \nrightarrow X_2$ be a matching, i.e. a partial injection. We call α an ε-*matching* if

(i) for all $\mathbf{u} \in \mathrm{dom}(\alpha)$, $\|\mathbf{u} - \alpha(\mathbf{u})\|_\infty \leq \varepsilon$,
(ii) for all $\mathbf{u} = (u_1, u_2) \in X_1 \setminus \mathrm{dom}(\alpha)$, $u_2 - u_1 \leq 2\varepsilon$,
(iii) for all $\mathbf{v} = (v_1, v_2) \in X_2 \setminus \mathrm{im}(\alpha)$, $v_2 - v_1 \leq 2\varepsilon$.

Their bottleneck distance $d_\mathrm{B}(X_1, X_2)$ is defined as the infimum of $\varepsilon \in [0, \infty)$ for which there exists an ε-matching $\alpha : X_1 \nrightarrow X_2$.

Recall that $\langle b, d \rangle_{\mathbf{ZZ}}$ for $b, d \in \mathbf{Z}$ denotes intervals of $\mathbf{ZZ}$ and that $\langle b, d \rangle$ for $b, d \in \mathbf{R}$ denotes intervals of $\mathbf{R}$.

Remark D.7 (1) Given a pair of multisets of intervals $\langle b, d \rangle_{\mathbf{ZZ}}$ of $\mathbf{ZZ}$, their bottleneck distance is defined by converting those multisets into the multisets of $\overline{\mathbf{U}}$ via the identification $\langle b, d \rangle_{\mathbf{ZZ}} \leftrightarrow (b, d) \in \overline{\mathbf{U}}$. (2) A map $\mathbf{Int}(\mathbf{ZZ}) \to \mathbf{Z}_+$ can be considered as a multiset of $\mathbf{Int}(\mathbf{ZZ})$ in an obvious way, and thus such maps can also be compared in d_B. (3) The bottleneck distance between multisets of real intervals $\langle b, d \rangle$ is also defined via the identification $\langle b, d \rangle \leftrightarrow (b, d) \in \overline{\mathbf{U}}$.

D.2.3 Algebraic stability for zigzag modules

Theorem D.8 (Bottleneck stability for zigzag modules Bjerkevik 2021; Botnan and Lesnick 2018) *For any* $M, N : \mathbf{ZZ} \to \mathbf{vec}$,

$$d_\mathrm{B}\left(\mathrm{barc}^{\mathbf{ZZ}}(M), \mathrm{barc}^{\mathbf{ZZ}}(N)\right) \leq 2 \cdot d_\mathrm{I}(M, N).$$

Let M be any zigzag module. By $\mathrm{barc}_{\mathbf{o}}^{\mathbf{ZZ}}(M)$, $\mathrm{barc}_{\mathbf{co}}^{\mathbf{ZZ}}(M)$, $\mathrm{barc}_{\mathbf{oc}}^{\mathbf{ZZ}}(M)$ and $\mathrm{barc}_{\mathbf{c}}^{\mathbf{ZZ}}(M)$, we mean the subcollection of $\mathrm{barc}^{\mathbf{ZZ}}(M)$, consisting solely of the points of the form $(b, d)_{\mathbf{ZZ}}$, $[b, d)_{\mathbf{ZZ}}$, $(b, d]_{\mathbf{ZZ}}$, and $[b, d]_{\mathbf{ZZ}}$, respectively.

Remark D.9 Suppose that the two zigzag modules M, N in Theorem D.8 are the *levelset zigzag persistent homology* of any two *Morse type* functions $f, g : X \to \mathbf{R}$ (Definition G.1). Then, the inequality in Theorem D.8 can be extended and strengthened as follows (Botnan and Lesnick 2018; Carlsson et al. 2019, Theorem 4.11): For each $\star \in \{\mathbf{o}, \mathbf{co}, \mathbf{oc}, \mathbf{c}\}$,

$$d_\mathrm{B}\left(\mathrm{barc}_{\star}^{\mathbf{ZZ}}(M), \mathrm{barc}_{\star}^{\mathbf{ZZ}}(N)\right) \leq d_\mathrm{I}(M, N)$$
$$\leq \|f - g\|_\infty.$$

Therefore, we also have

$$d_\mathrm{B}\left(\mathrm{barc}^{\mathbf{ZZ}}(M), \mathrm{barc}^{\mathbf{ZZ}}(N)\right) \leq \max_{\star} d_\mathrm{B}\left(\mathrm{barc}_{\star}^{\mathbf{ZZ}}(M), \mathrm{barc}_{\star}^{\mathbf{ZZ}}(N)\right)$$
$$\leq d_\mathrm{I}(M, N) \leq \|f - g\|_\infty,$$

where the maximum is taken over all $\star \in \{\mathbf{o}, \mathbf{co}, \mathbf{oc}, \mathbf{c}\}$.

E Proof of Proposition 3.17

In this section we prove Proposition 3.17. For notational simplicity, we set $\mathbf{P} = \mathbf{ZZ}$. We remark that the proof extends to arbitrary connected locally finite posets.

E.1 Canonical bases and *l*-type intervals

Suppose that $M : \mathbf{ZZ} \to \mathbf{vec}$ is given as $M = \bigoplus_{c \in C} I^{J_c}$ for an index set C. Then for each $(i, j) \in \mathbf{ZZ}$, the dimension of $M_{(i,j)}$ is the total multiplicity of intervals containing (i, j) in $\mathrm{barc}^{\mathbf{ZZ}}(M) = \{J_c : c \in C\}$. Given any $c \in C$, and $(i, j) \in J_c$, let $e_c^{(i,j)}$ denote the 1 in the field $\mathbb{F}$ which corresponds to the (i, j)-component of I^{J_c}. Then for each $(i, j) \in \mathbf{ZZ}$, the vector space $M_{(i,j)}$ admits the *canonical basis* $B_{(i,j)} = \left\{ e_c^{(i,j)} : J_c \text{ contains } (i, j) \right\}$ and hence every $v \in M_{(i,j)}$ can be uniquely expressed as a linear combination of the elements in $B_{(i,j)}$, i.e.

$$v = \sum_{J_c \ni (i,j)} a_c e_c^{(i,j)}, \qquad a_c \in \mathbb{F}. \tag{10}$$

This expression is called the *canonical expression* of v. The collection $\mathscr{B} = (B_{(i,j)})_{(i,j) \in \mathbf{ZZ}}$ is called the *canonical basis* of $M = \bigoplus_{c \in C} I^{J_c}$.

In order to prove Proposition 3.17, we identify a certain type of intervals of the poset $\mathbf{ZZ}$:

Definition E.1 (*l-type intervals*) Any interval I of $\mathbf{ZZ}$ is called *l-type*[13] if I is of the form $(b, d)_{\mathbf{ZZ}}$ for some $b \in \mathbf{Z} \cup \{-\infty\}$ and $d \in \mathbf{Z} \cup \{\infty\}$ (see Fig. 2).

Notation E.2 Let $M : \mathbf{ZZ} \to \mathbf{vec}$ be any zigzag module and let $(i, j), (i', j') \in \mathbf{ZZ}$. For $v_{(i,j)} \in M_{(i,j)}$ and $v_{(i',j')} \in M_{(i',j')}$, we write $v_{(i,j)} \sim v_{(i',j')}$ if $(i, j), (i', j')$ are comparable and either

$$v_{(i,j)} = \varphi_M((i', j'), (i, j))(v_{(i',j')}) \text{ (when } (i', j') \leq (i, j)) \text{ or}$$
$$v_{(i',j')} = \varphi_M((i, j), (i', j'))(v_{(i,j)}) \text{ (when } (i, j) \leq (i', j')).$$

Proof of Proposition 3.17 For simplicity of notation, we prove the proposition when $\mathbf{P} = \mathbf{ZZ}$ and $J = (-\infty, \infty) = \mathbf{ZZ}$. The proof straightforwardly extends to any connected, locally finite poset $\mathbf{P}$ and $J \in \mathrm{Con}(\mathbf{P})$. By Theorem 2.12, we may assume that $M = \bigoplus_{c \in C} I^{J_c}$, where $\mathrm{barc}^{\mathbf{ZZ}}(M) = \{J_c : c \in C\}$. In what follows, we identify $\mathbf{ZZ}$ with the integers $\mathbf{Z}$ via the bijection $(i, j) \mapsto i + j$. Therefore, by $\bigoplus_{k \in \mathbf{Z}} M_k$ and $\prod_{k \in \mathbf{Z}} M_k$, we will denote $\bigoplus_{(i,j) \in \mathbf{ZZ}} M_{(i,j)}$ and $\prod_{(i,j) \in \mathbf{ZZ}} M_{(i,j)}$ respectively.

(i) The limit of M is (isomorphic to) the pair $(V, (\pi_k)_{k \in \mathbf{Z}})$ described as follows:

$$V := \left\{ (v_k)_{k \in \mathbf{Z}} \in \prod_{k \in \mathbf{Z}} M_k : \forall k \in \mathbf{Z}, \ v_k \sim v_{k+1} \right\}. \tag{11}$$

---

[13] Given any $M : \mathbf{ZZ} \to \mathbf{vec}$, *l*-type intervals in $\mathrm{barc}^{\mathbf{ZZ}}(M)$ contribute to the dimension of the *limit* of M and that is the reason for the name '*l*'-type.

For each $k \in \mathbf{Z}$, the map $\pi_k : V \to M_k$ is the canonical projection.

(ii) The colimit of M is (isomorphic to) the pair $(U, (i_k)_{k \in \mathbf{Z}})$ described as follows: U is the quotient vector space $\left(\bigoplus_{k \in \mathbf{Z}} M_k \right) / W$, where W is the subspace of the direct sum $\bigoplus_{k \in \mathbf{Z}} M_k$ which is generated by the vectors of the form $(\cdots, 0, \ldots, 0, v_k, -v_{k+1}, 0, \ldots, 0, \cdots)$ with $v_k \sim v_{k+1}$ (Notation E.2). Let q be the quotient map from $\bigoplus_{k \in \mathbf{Z}} M_k$ to $U = \left(\bigoplus_{k \in \mathbf{Z}} M_k \right) / W$. For $k \in \mathbf{Z}$, let the map $\bar{i}_k : M_k \to \bigoplus_{k \in \mathbf{Z}} M_k$ be the canonical injection. Then $i_k : M_k \to U$ is the composition $q \circ \bar{i}_k$.

Let $0 \in \mathbf{Z}$. Since the canonical map $\psi_M : \varprojlim M \to \varinjlim M$ is equal to $i_0 \circ \pi_0$, it suffices to show that the dimension of the vector space $\mathrm{im}(i_0 \circ \pi_0)$ is equal to the cardinality of $(-\infty, \infty)_{\mathbf{ZZ}}$ in $\mathrm{barc}^{\mathbf{ZZ}}(M)$.

If M_0 is the zero space, then it is clear that there is no $(-\infty, \infty)_{\mathbf{ZZ}}$ in $\mathrm{barc}^{\mathbf{ZZ}}(M)$ and that $\mathrm{im}(\pi_0) = 0$. Therefore, $\mathrm{im}(i_0 \circ \pi_0) = 0$, and the statement directly follows.

Assume that M_0 is not trivial. Define

$$C_0 := \{ c \in C : 0 \in J_c \}.$$

Since M_0 is a finite dimensional vector space, $|C_0| = \dim(M_0)$ is finite and hence we can write $C_0 = \{c_1, \ldots, c_m\}$, where $\dim(M_0) = m$. Also, we can write $M_0 \cong \bigoplus_{j=1}^m \mathbb{F}_j$, where each $\mathbb{F}_j = \mathbb{F}$ is the component of the interval module $I^{J_{c_j}}$ at $0 \in \mathbf{ZZ}$. We identify M_0 with $\bigoplus_{j=1}^m \mathbb{F}_j$. Then, we claim that

$$\mathrm{im}(\pi_0) = \left\{ (a_1, \ldots, a_m) \in \bigoplus_{j=1}^m \mathbb{F}_j : a_j = 0 \text{ if } J_{c_j} \text{ is not } l\text{-type} \right\}. \tag{12}$$

We prove this equality at the end of the proof.

For $j = 1, \ldots, m$, let $e_j := (0, \ldots, 0, \underset{j-\mathrm{th}}{1}, 0, \ldots, 0) \in \bigoplus_{j=1}^m \mathbb{F}_j$. By equation (12), the set $B_0 = \{ e_j : J_{c_j} \text{ is } l\text{-type} \}$ is a basis of $\mathrm{im}(\pi_0)$. Therefore, the dimension of $\mathrm{im}(i_0 \circ \pi_0)$ is equal to the dimension of the space that is spanned by the image of B_0 under the map $i_0 : M_0 \to U$. By invoking item (ii) above, if $J_{c_j} \neq (-\infty, \infty)_{\mathbf{ZZ}}$, it follows that $i_0(e_j) = 0 \in U$.

Let $C_0^{\mathrm{full}} := \{ c \in C : J_c = (-\infty, \infty)_{\mathbf{ZZ}} \}$, which is a subset of $C_0 = \{c_1, \ldots, c_m\}$. Assuming that $C_0^{\mathrm{full}} \neq \emptyset$, suppose that $C_0^{\mathrm{full}} = \{c_1, \ldots, c_n\}$ for some $n \leq m$ without loss of generality. Invoking item (ii) above, the set $\{ i_0(e_1), \ldots, i_0(e_n) \}$ is linearly independent in U. Therefore, we have that

$$\mathrm{rank}(i_0 \circ \pi_0) = n = (\text{the multiplicity of } (-\infty, \infty)_{\mathbf{ZZ}} \text{ in } \mathrm{barc}^{\mathbf{ZZ}}(M)),$$

as desired.

Finally we prove equation (12). First we prove "$\subset$". Recall item (i) above and pick any $v = (v_k)_{k \in \mathbf{Z}} \in V$. Then $\pi_0(v) = v_0 = (a_1, \ldots, a_m)$. Suppose that $c_j \in C_0$ is such that J_{c_j} is not l-type. This implies that the interval J_{c_j} has an endpoint $r = (r_1, r_2) \in \mathbf{ZZ}$ where $r_1 = r_2 \in \mathbf{Z}$ and then either $(r_1 + 1, r_1)$ or $(r_1, r_1 - 1)$ is not in

J_{c_1}. Without loss of generality, assume that $s = (r_1 + 1, r_1) \in \mathbf{ZZ}$ does not belong to J_{c_1}. By the choice of $v = (v_k)_{k \in \mathbf{Z}}$, we have $v_0 \sim v_1 \sim \ldots \sim v_{2r_1} \sim v_{2r_1+1}$, and this leads to that $a_j = 0$.

Next we show "$\supset$". Pick any $(a_1, \ldots, a_m)$ in the RHS of equation (12). For each $k \in \mathbf{Z}$, define $v_k \in M_k$ using the canonical expression:

$$v_k := \sum_{J_c \ni k} b_c e_c^k, \qquad b_c \in \mathbb{F},$$

where $b_c = 0$ if $c \notin C_0$ and $b_c = a_j$ if $c = c_j \in C_0 = \{c_1, \ldots, c_m\}$. Let $v := (v_k)_{k \in \mathbf{Z}}$. Then, one can check that $\pi_0(v) = (a_1, \ldots, a_m)$, completing the proof.   □

F Constructible persistence modules and re-indexing

Patel generalizes the persistence diagram of Cohen-Steiner, Edelsbrunner, and Harer to the setting of *constructible* **R**-indexed diagrams valued in a symmetric monoidal category (Patel 2018).

Definition F.1 (*Constructible* **R**-*indexed diagrams*) Let $S = \{s_1 < s_2 < \ldots < s_n\}$ be a finite set of **R**. A diagram $F : \mathbf{R} \to \mathscr{C}$ is *S-constructible* if

  (i) for $p \leq q < s_1$, $\varphi_F(p, q)$ is the identity on e,
 (ii) for $s_i \leq p \leq q < s_{i+1}$, $\varphi_F(p, q)$ is an isomorphism,
(iii) for $s_n \leq p \leq q$, $\varphi_F(p, q)$ is an isomorphism.

If $G : \mathbf{R} \to \mathscr{C}$ is T-constructible for some finite set $T \subset \mathbf{R}$, then we call G *constructible*.

F.1 Re-indexing an R-indexed diagram by Z

Let $F : \mathbf{R} \to \mathscr{C}$ be S-constructible with $S = \{s_1 < s_2 < \ldots < s_n\}$. A functor $D(F) : \mathbf{Z} \to \mathscr{C}$ that contains all the algebraic information of F would be defined as follows:

$$D(F)_i = \begin{cases} e, & \text{for } i \leq 0, \\ F_{s_i}, & \text{for } i = 1, \ldots, n, \\ F_{s_n}, & \text{for } i \geq n + 1. \end{cases}$$

$$\varphi_{D(F)}(i, i+1) = \begin{cases} \mathrm{id}_e, & \text{for } i \leq 0, \\ \varphi_F(s_i, s_{i+1}), & \text{for } i = 0, 1, \ldots, n-1, \text{ where} \\ & \quad s_0 \text{ is arbitrarily chosen in } (-\infty, s_1), \\ \mathrm{id}_{F_{s_n}}, & \text{for } i \geq n. \end{cases}$$

When $\mathscr{C} = \mathbf{vec}$, there exists a bijection from the barcode of F to that of $D(F)$ via $[s_i, s_j) \mapsto [i, j-1]$ for $1 \leq i < j \leq n$, and $[s_i, \infty) \mapsto [i, \infty)$ for $1 \leq i \leq n$.

F.2 Re-indexing a Z-indexed diagram by ZZ

Let $F : \mathbf{Z} \to \mathscr{C}$. Let us define $L(F) : \mathbf{ZZ} \to \mathscr{C}$ as follows (Botnan and Lesnick 2018, Remark 4.5):

$$
\begin{aligned}
L(F)_{(i,i)} &= L(F)_{(i+1,i)} = F_i \\
\varphi_{L(F)}\left((i+1,i),(i,i)\right) &= \mathrm{id}_{F_i} \\
\varphi_{L(F)}\left((i-1,i),(i,i)\right) &= \varphi_F(i-1,i).
\end{aligned}
$$

When $\mathscr{C} = \mathbf{vec}$ and $F_i = 0$ for $i \leq 0$, there exists a bijection from the barcode of F to that of $L(F)$ via $[a,b] \mapsto [a,b+1)_{\mathbf{ZZ}}$ for $a,b \in \mathbf{Z}$ with $a \leq b$, and $[a,\infty) \mapsto [a,\infty)_{\mathbf{ZZ}}$ for $a \in \mathbf{Z}$.

G Rigorous definition of Reeb graphs

In order to introduce the definition of Reeb graphs, we begin by introducing the notion of *Morse type* functions from Botnan and Lesnick (2018), Carlsson et al. (2009).

Definition G.1 (*Morse type functions*) Let X be a topological space. We say that a continuous function $p : X \to \mathbf{R}$ is of *Morse type* if

(i) There exists a strictly increasing function $\mathscr{G} : \mathbf{Z} \to \mathbf{R}$ such that $\lim_{i \to +\infty} \mathscr{G}(i) = \infty$, $\lim_{i \to -\infty} \mathscr{G}(i) = -\infty$ and such that for each open interval $I_i = (\mathscr{G}(i), \mathscr{G}(i+1))$ there exist a topological space Y_i and a homeomorphism $h_i : I_i \times Y_i \to p^{-1}(I_i)$ with $f \circ h_i$ being the projection $I_i \times Y_i \to I_i$.

(ii) Each homeomorphism $h_i : I_i \times Y_i \to p^{-1}(I_i)$ extends to a continuous function

$$
\bar{h}_i : \bar{I}_i \times Y_i \to p^{-1}(\bar{I}_i),
$$

where $\bar{I}_i$ denotes the closure of I_i.

(iii) For all $t \in \mathbf{R}$ and $k \in \mathbf{Z}_+$, $\dim \mathrm{H}_k\left(p^{-1}(t)\right) < \infty$.

We introduce the definition of Reeb graphs (De Silva et al. 2016).

Definition G.2 (*Reeb graphs*) Let X be a topological space and let $p : X \to \mathbf{R}$ be of Morse type. If the topological spaces Y_i as in Definition G.1 (i) are finite sets of points with the discrete topology, then the pair (X, p) is said to be a *Reeb graph*.[14] See Fig. 5 (A) for an illustrative example.

References

Adams, H., Carlsson, G.: Evasion paths in mobile sensor networks. Int. J. Robot. Res. **34**(1), 90–104 (2015)

Asashiba, H., Escolar, E.G., Nakashima, K., Yoshiwaki, M.: On approximation of 2 d persistence modules by interval-decomposables (2019). arXiv preprint arXiv:1911.01637

---

[14] In De Silva et al. (2016), the authors use the term **R**-*graph* instead of Reeb graph.

Awodey, S.: Category Theory. Oxford University Press, Oxford (2010)

Azumaya, G.: Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt's theorem. Nagoya Math. J. **1**, 117–124 (1950)

Bauer, U., Ge, X., Wang, Y.: Measuring distance between Reeb graphs. In: Proceedings of the thirtieth annual symposium on Computational geometry, p. 464. ACM (2014)

Bauer, U., Landi, C., Mémoli, F.: The Reeb graph edit distance is universal. Found. Comput. Math. (2020). https://doi.org/10.1007/s10208-020-09488-3

Bauer, U., Munch, E., Wang, Y.: Strong Equivalence of the Interleaving and Functional Distortion Metrics for Reeb Graphs. In: Arge, L., Pach, J. (eds.) 31st International Symposium on Computational Geometry (SoCG 2015), vol. 34, pp. 461–475. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2015). https://doi.org/10.4230/LIPIcs.SOCG.2015.461. Retrieved from http://drops.dagstuhl.de/opus/volltexte/2015/5146

Betthauser, L., Bubenik, P., Edwards, P.B.: Graded persistence diagrams and persistence landscapes. Discrete Comput. Geom. (2021). https://doi.org/10.1007/s00454-021-00316-1

Biasotti, S., Giorgi, D., Spagnuolo, M., Falcidieno, B.: Reeb graphs for shape analysis and applications. Theor. Comput. Sci. **392**(1–3), 5–22 (2008)

Birkhoff, G.: Lattice theory, vol. 25. Am. Math. Soc. (1940)

Bjerkevik, H.B.: On the stability of interval decomposable persistence modules. Discrete Comput. Geom. **66**(1), 92–121 (2021)

Botnan, M.B.: Interval decomposition of infinite zigzag persistence modules. Proc. Am. Math. Soc. **145**(8), 3571–3577 (2017)

Botnan, M.B.: Multidimensional persistence, applied algebraic topology network (2017). https://www.youtube.com/watch?v=B3NPBSqH2lM&t=3257

Botnan, M.B., Lebovici, V., Oudot, S.: On Rectangle-Decomposable 2-Parameter Persistence Modules. In: 36th International Symposium on Computational Geometry (SoCG 2020), *Leibniz International Proceedings in Informatics (LIPIcs)*, vol. 164, pp. 22:1–22:16 (2020). Retrieved from https://drops.dagstuhl.de/opus/volltexte/2020/12180. https://doi.org/10.4230/LIPIcs.SoCG.2020.22

Botnan, M.B., Lesnick, M.: Algebraic stability of zigzag persistence modules. Algebraic Geomet. Topol. **18**(6), 3133–3204 (2018)

Bubenik, P., Scott, J.A.: Categorification of persistent homology. Dis. Comput. Geomet. **51**(3), 600–627 (2014)

Buchin, K., Buchin, M., van Kreveld, M.J., Speckmann, B., Staals, F.: Trajectory grouping structure. JoCG **6**(1), 75–98 (2015)

Carlsson, G., De Silva, V.: Zigzag persistence. Foundations Comput. Math. **10**(4), 367–405 (2010)

Carlsson, G., De Silva, V., Morozov, D.: Zigzag persistent homology and real-valued functions. In: Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry, pp. 247–256. ACM (2009)

Carlsson, G., de Silva, V., Kališnik, S., Morozov, D.: Parametrized homology via zigzag persistence. Algebraic Geomet. Topol. **19**(2), 657–700 (2019)

Carlsson, G., Zomorodian, A.: The theory of multidimensional persistence. Dis. Comput. Geomet. **42**(1), 71–93 (2009)

Carrière, M., Oudot, S.: Local Equivalence and Intrinsic Metrics between Reeb Graphs. In: Aronov, B., Katz, M.J. (eds.) 33rd International Symposium on Computational Geometry (SoCG 2017), vol. 77, pp. 25:1-25:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2017). https://doi.org/10.4230/LIPIcs.SoCG.2017.25. Retrieved from http://drops.dagstuhl.de/opus/volltexte/2017/7179

Cerri, A., Fabio, B.D., Ferri, M., Frosini, P., Landi, C.: Betti numbers in multidimensional persistent homology are stable functions. Math. Methods Appl. Sci. **36**(12), 1543–1557 (2013)

Chambers, E.W., Letscher, D.: Persistent homology over directed acyclic graphs. In: Research in Computational Topology, pp. 11–32. Springer (2018)

Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L.J., Oudot, S.: Proximity of persistence modules and their diagrams. In: Proceedings of the 25th Annual Symposium on Computational Geometry, pp. 237–246 (2009)

Chazal, F., De Silva, V., Glisse, M., Oudot, S.: The Structure and Stability of Persistence Modules. Springer, Berlin (2016)

Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. Dis. Comput. Geomet. **37**(1), 103–120 (2007)

Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Extending persistence using Poincaré and Lefschetz duality. Foundations Comput. Math. **9**(1), 79–103 (2009)

Corcoran, P., Jones, C.B.: Spatio-temporal modeling of the topology of swarm behavior with persistence landscapes. In: Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 1–4 (2016)

Crawley-Boevey, W.: Decomposition of pointwise finite-dimensional persistence modules. J. Algebra Appl. **14**(05), 1550066 (2015)

Curry, J.: Sheaves, cosheaves and applications (2013). arXiv preprint arXiv:1303.3255

Curry, J., Patel, A.: Classification of constructible cosheaves Theory Appl. Categories (2020)

De Silva, V., Munch, E., Patel, A.: Categorified Reeb graphs. Dis. Comput. Geomet. **55**(4), 854–906 (2016)

Dey, T.K., Mémoli, F., Wang, Y.: Multiscale mapper: topological summarization via codomain covers. In: Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 997–1013. SIAM (2016)

Dey, T.K., Wang, Y.: Reeb graphs: Approximation and persistence. Dis. Comput. Geomet. **49**(1), 46–73 (2013)

Elchesen, A., Mémoli, F.: The reflection distance between zigzag persistence modules. J. Appl. Comput. Topol. **3**(3), 185–219 (2019)

Ge, X., Safa, I.I., Belkin, M., Wang, Y.: Data skeletonization via reeb graphs. In: Advances in Neural Information Processing Systems, pp. 837–845 (2011)

Harvey, W., Wang, Y., Wenger, R.: A randomized o (m log m) time algorithm for computing reeb graphs of arbitrary simplicial complexes. In: Proceedings of the Twenty-Sixth Annual Symposium on Computational Geometry, pp. 267–276 (2010)

Kerber, M., Morozov, D., Nigmetov, A.: Geometry helps to compare persistence diagrams. J. Exp. Algorithmics **22**, 1–20 (2017)

Kim, W., Mémoli, F.: Stable signatures for dynamic graphs and dynamic metric spaces via zigzag persistence (2017). arXiv preprint arXiv:1712.04064

Kim, W., Mémoli, F.: Formigrams: Clustering summaries of dynamic data. In: Proceedings of 30th Canadian Conference on Computational Geometry (CCCG 2018) (2018)

Kim, W., Mémoli, F.: Generalized persistence diagrams for persistence modules over posets (2018). arXiv preprint arXiv:1810.11517v4

Lesnick, M.: The theory of the interleaving distance on multidimensional persistence modules. Foundations Comput. Math. **15**(3), 613–650 (2015). https://doi.org/10.1007/s10208-015-9255-y

Mac Lane, S.: Categories for the Working Mathematician, vol. 5. Springer, Berlin (2013)

Mata, G., Morales, M., Romero, A., Rubio, J.: Zigzag persistent homology for processing neuronal images. Pattern Recognit. Lett. **62**, 55–60 (2015)

McCleary, A., Patel, A.: Bottleneck stability for generalized persistence diagrams. Proc. Am. Math. Soc. **148**(7), 3149–3161 (2020)

Miller, E.: Homological algebra of modules over posets (2020). arXiv preprint arXiv:2008.00063

Milosavljević, N., Morozov, D., Skraba, P.: Zigzag persistent homology in matrix multiplication time. In: Proceedings of the Twenty-Seventh Annual Symposium on Computational Geometry, pp. 216–225. ACM (2011)

Mitchell, B.: Theory of Categories, vol. 17. Academic Press, Cambridge (1965)

Morozov, D., Beketayev, K., Weber, G.: Interleaving distance between merge trees. In: Proceedings of TopoInVis. (2013)

Oudot, S.Y., Sheehy, D.R.: Zigzag zoology: Rips zigzags for homology inference. Foundations Comput. Math. **15**(5), 1151–1186 (2015)

Parsa, S.: A deterministic $O(m \log m)$ time algorithm for the Reeb graph. In: ACM Sympos. Comput. Geom. (SoCG), pp. 269–276 (2012)

Patel, A.: Generalized persistence diagrams. J. Appl. Comput. Topol. **1**, 1–23 (2018)

Puuska, V.: Erosion distance for generalized persistence modules. Homol. Homotopy Appl. **22**(1), 233–254 (2020)

Reeb, G.: Sur les points singuliers d'une forme de pfaff completement integrable ou d'une fonction numerique [on the singular points of a completely integrable pfaff form or of a numerical function]. Comptes Rendus Acad. Sci. Paris **222**, 847–849 (1946)

Rota, G.C.: On the foundations of combinatorial theory i theory of Möbius functions. Probab. Theory Related Fields. **2**(4), 340–368 (1964)

Shinagawa, Y., Kunii, T., Kergosien, Y.L.: Surface coding based on morse theory. IEEE Comput. Graph. Appl. **11**(5), 66–78 (1991). https://doi.org/10.1109/38.90568

Singh, G., Mémoli, F., Carlsson, G.: Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. In: Botsch, M., Pajarola, R. (eds) pp. 91–100. Eurographics Association, Prague, Czech Republic (2007). https://doi.org/10.2312/SPBG/SPBG07/091-100

Stefanou, A.: Tree decomposition of reeb graphs, parametrized complexity, and applications to phylogenetics. J. Appl. Comput. Topol. (2020). https://doi.org/10.1007/s41468-020-00051-1

Weibel, C.A.: The K-Book: An Introduction to Algebraic K-Theory, vol. 145. American Mathematical Society, Providence (2013)

Zomorodian, A., Carlsson, G.: Computing persistent homology. Dis. Comput. Geometry **33**(2), 249–274 (2005)