Journal of Applied and Computational Topology (2023) 7:771–813
https://doi.org/10.1007/s41468-023-00124-x

A directed persistent homology theory for dissimilarity functions

David Méndez[1,2] · **Rubén J. Sánchez-García**[1,3]

Received: 3 August 2020 / Revised: 28 April 2023 / Accepted: 4 May 2023 /
Published online: 2 June 2023
© The Author(s) 2023

Abstract
We develop a theory of persistent homology for directed simplicial complexes which detects persistent directed cycles in odd dimensions. We relate directed persistent homology to classical persistent homology, prove some stability results, and discuss the computational challenges of our approach. Our directed persistent homology theory is motivated by homology with semiring coefficients: by explicitly removing additive inverses, we are able to detect directed cycles algebraically.

Keywords Persistent homology · Dissimilarity functions · Directed simplicial complexes · Directed cycles

Mathematics Subject Classification 55N31 · 55N35 · 55U10 · 16Y60

1 Introduction

Persistent homology is one of the most successful tools in Topological Data Analysis (Carlsson 2009), with recent applications in numerous scientific domains such as biology, medicine, neuroscience, robotics, and many others (Otter et al. 2017). In its most common implementation, persistent homology is used to infer topological properties of the metric space underlying a finite point cloud using two steps (Edelsbrunner and Harer 2010):

✉ David Méndez
david.mendez@research.fchampalimaud.org

Rubén J. Sánchez-García
R.Sanchez-Garcia@soton.ac.uk

1 School of Mathematical Sciences, University of Southampton, SO17 1BJ Southampton, UK

2 Champalimaud Centre for the Unknown, Av. Brasília s/n, 1400-038 Lisbon, Portugal

3 The Alan Turing Institute, London NW1 2DB, UK

(1) Build a filtration of simplicial complexes from distances, or similarities, between data points.
(2) Compute the singular homology of each of the simplicial complexes in the filtration, along with the linear maps induced in homology by their inclusions. The resulting *persistence module* can normally be represented using a *persistence diagram* or a *persistence barcode*.

A fundamental limitation of persistent homology, and in fact homology, is its inability to incorporate directionality, which can be important in some real-world applications (see examples below). For instance, homology cannot in principle distinguish between directed and undirected cycles (Fig. 1). Although the definition of homology, namely the differential or boundary operator, requires a choice of orientation for the simplices, the resulting homology is independent of this choice. Previous attempts have had some partial success at this issue (see 'Related work' below), however, as far as we know, there is no homology theory able to exactly detect directed cycles up to boundary equivalence.

The main difficulty occurs at the algebraic level: opposite orientations on a simplex correspond to additive inverse elements in the coefficient ring or field. Our key insight is to retain the submodule of homology generated by those classes that only have non-negative coefficients on elementary chains. In this way, we can define a directed homology module (Definition 3.6) of directed simplicial complexes (Definition 3.1) that detects homology classes of directed 1-cycles, as desired (Fig. 1, Proposition 3.24).

This insight arises from the use of homology with semiring coefficients, where additive inverses may not exist. Indeed, the main results in this paper can be understood and reformulated in full generality using homology theory with semiring coefficients, see Appendix 1.

Our directed homology theory can be extended to the persistent setting. More concretely, we define two persistence modules associated to filtrations of directed simplicial complexes: an undirected one using the whole homology module (Definition 4.2) and a directed one using the submodule of directed homology (Definition 4.3). We show that the bars in the directed barcodes can be matched one-to-one with bars in the undirected barcodes, with the directed bars possibly having a later birth (Proposition 4.7, Figs. 2, 5, 6).

Our motivation to develop directed persistent homology is the applications to real-world data sets where asymmetry or directionality, in particular the presence of directed cycles, plays a fundamental role. Examples include biological neural networks (directed synaptic connections), time series data (directed temporal connections) or biological molecular networks such as protein-protein interaction networks (inhibitory and excitatory connections). Mathematically, we can model these situations using a finite set V of data points, and an arbitrary function $d_V : V \times V \to \mathbb{R}$, which we call *dissimilarity function* (Sect. 2.2) and which, crucially, may not be symmetric. Our main example of a directed simplicial complex is the directed Rips complex (Definition 4.12) of (V, d_V), which becomes the input of our suggested directed persistent homology pipeline.

In the context of dissimilarity functions, we prove a stability result with respect to the bottleneck distance between (directed) persistent diagrams and the correspondence

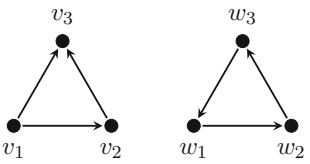


Fig. 1 Two examples of directed simplicial complexes X (left) and Y (right). Our directed submodules of homology over the rings $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ detect directed 1-cycles: $H_1^{\mathrm{Dir}}(X, R) = \{0\}$, while $H_1^{\mathrm{Dir}}(Y, R) = R$ generated by the directed 1-cycle $[w_1, w_2] + [w_2, w_3] + [w_3, w_1]$

distortion distance (a generalization of the Gromov-Hausdorff distance) between dissimilarity functions over finite sets (Theorem 4.19). We finish this article by discussing the computational challenges of our approach.

Related work

A first approach to incorporating directionality would be to encode the asymmetry of a data set in the simplicial complex built from it. In Turner (2019), the author uses *ordered tuple complexes* or *OT-complexes*, which are generalizations of simplicial complexes where simplices are ordered tuples of vertices. Similar ideas have been successfully used in the field of neuroscience to show the importance of *directed cliques* of neurons (Reimann et al. 2017), showing that directionality in neuron connectivity plays a crucial role in the structure and function of the brain. In Chowdhury and Mémoli (2018a), the authors use so-called *Dowker filtrations* to develop persistent homology for asymmetric networks, and note (Chowdhury and Mémoli 2018a, Remark 35) that a non-trivial 1-dimensional persistence diagram associated to Dowker filtrations suggests the presence of directed cycles. However, they also note that such persistence diagram may be non-trivial even if directed cycles are not present. Further progress can be achieved by combining the ideas above with homology theories that are themselves sensitive to asymmetry. In Chowdhury and Mémoli (2018b), the authors develop persistent homology for directed networks using the theory of path homology of graphs (Grigor'yan et al. 2020). This persistent homology theory shows stability with respect to the distance between asymmetric networks (Carlsson et al. 2014), and is indeed able to tell apart digraphs with isomorphic underlying graphs but different orientations on the edges. Nonetheless, cycles in path homology do not correspond to directed cycles, see Chowdhury and Mémoli (2018b, Example 10). Another significant contribution can be found in Turner (2019), where four new approaches to persistent homology are developed for asymmetric data, all of which are shown to be stable. Each approach is sensitive to asymmetry in a different way. For example, one of them uses a generalization of poset homology to preorders to detect strongly connected components in digraphs, a feature our implementation does not have (see Proposition 3.18). More interestingly for our purposes, in Turner (2019, Section 5) the author introduces a persistent homology approach that builds a *directed Rips filtration* of ordered tuple complexes associated to *dissimilarity functions* $d_V : V \times V \to \mathbb{R}$. This is indeed sensitive to asymmetry and yields a persistent homology pipeline that

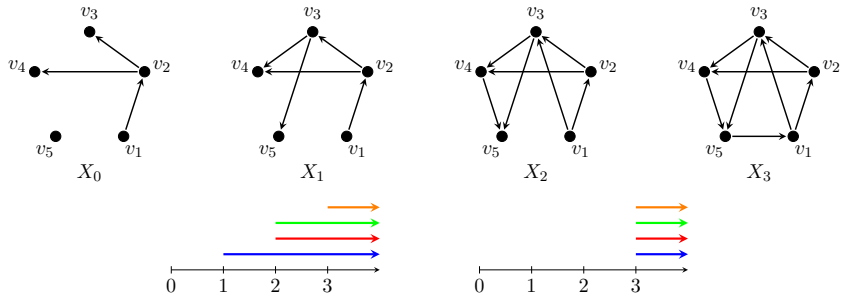


Fig. 2 A filtration of a directed simplicial complex (top, left to right) and corresponding undirected (bottom, left) and directed (bottom, right) persistence barcodes. Every bar in the directed barcode corresponds to a unique bar in the undirected barcode (shown here by matching colours) (colour figure online)

is stable with respect to the correspondence distortion distance, a generalization of the Gromov-Hausdorff distance to dissimilarity functions (see Turner 2019, Section 2 and Sect. 2.2). However, undirected cycles are also detected by this homology theory. In contrast, in this article we produce a theory of persistent homology associated to ordered tuple complexes that only detects directed cycles and which is also stable with respect to the correspondence distortion distance.

Overview of results

We introduce modules of undirected and directed homology associated to directed simplicial complexes (Definition 3.4) and show that they satisfy some desirable properties. Furthermore, we prove that for certain coefficient rings, closed paths will only have non-trivial directed homology in dimension 1 when their edges form a directed cycle, as we set up to do (Proposition 3.24). Our modules of directed homology can be defined in dimensions greater than one, although they are zero in all even dimensions (Proposition 3.26) except in dimension 0 where they detect weakly connected components (Proposition 3.18).

Next, we extend our homology theory to the persistence setting. In particular, we introduce, for each dimension, two persistence modules associated to the same filtration: an *undirected* persistence module (Definition 4.2), which is analogous to the one introduced in Turner (2019, Section 5), and the submodule generated by directed classes, which we call the *directed* persistence module (Definition 4.3). The directed persistence module gives rise to bars associated to homology classes that can be represented by directed cycles, as illustrated in Fig. 2 and Examples 4.8 to 4.11. We also establish a relation between the undirected and directed persistence barcodes of the same filtration. Indeed, in Proposition 4.7 we show that every bar in the directed barcode corresponds to a unique bar in the undirected barcode that dies at the same time. The bar in the directed barcode may, however, be born later, and some bars in the undirected barcode may be left unmatched (see Figs. 2, 5 and 6).

Having all the necessary ingredients, we can provide a complete workflow for directed persistent homology: given a dissimilarity function d_V on a finite set V, we construct its directed Rips filtration (Definition 4.12), and take the correspond-

ing undirected and directed n-dimensional persistence modules, respectively denoted $\mathcal{H}_n(V, d_V)$ and $\mathcal{H}_n^{\mathrm{Dir}}(V, d_V)$ (Definition 4.13). Both of these persistence modules have persistence diagrams and barcodes (Definition 4.15) and the associated persistence modules are stable with respect to the correspondence distortion distance (Sect. 2.2 and Theorem 4.19).

We finish the article with a discussion of the computational challenges of calculating the directed persistent homology of a dissimilarity function. The Standard Algorithm can easily be adapted to compute the barcodes of the undirected persistent homology of the dissimilarity function, but the computation of the directed persistence barcodes is more challenging. A direct approach results in an extreme ray enumeration problem (Sect. 5), a well-known problem in computational geometry. Developing an algorithm that computes the directed homology classes, rather than all directed cycles, remains an open problem.

We include an Appendix on homology with semiring coefficients and on how it gives rise to the ideas introduced in this article. This provides additional motivation and context for our results, and presents them in full generality.

Outline of the article

In Sect. 2, we introduce the necessary background in persistence modules (2.1) and dissimilarity functions (2.2). In Sect. 3 we introduce undirected and directed homology modules of directed simplicial complexes. We then use these homology modules to introduce directed and undirected persistent homology in Sect. 4, where we also show our stability results with respect to dissimilarity functions. In Sect. 5, we discuss the computational implementation of directed persistent homology and lay out some future research directions. Finally, Appendix 1 introduces homology with semiring coefficients and shows how it relates to directed homology.

2 Persistence modules and dissimilarity functions

Our goal is to extend persistent homology to directed simplicial complexes such as those constructed from a dissimilarity function. In this section, we introduce the necessary background regarding persistent homology (Sect. 2.1) and dissimilarity functions (Sect. 2.2).

2.1 Persistence modules, diagrams and barcodes

In this section, we introduce persistence modules and their associated persistence diagrams and barcodes. We follow the exposition in Chowdhury and Mémoli (2018b), as it is simple and focused on persistence modules arising in the context we are interested in: that of persistent homology of finite simplicial complexes. Many of the results below hold in more generality (Chazal et al. 2016) but we chose to keep the exposition simple.

Let R be a ring with unity and let $T \subseteq \mathbb{R}$ be a subset.

Definition 2.1 A *persistence R-module* over T, written $\mathcal{V} = \left(\{V^\delta\}, \{\nu_\delta^{\delta'}\}\right)_{\delta \leq \delta' \in T}$, is a family of R-modules $\{V^\delta\}_{\delta \in T}$ and homomorphisms $\nu_\delta^{\delta'} \colon V^\delta \to V^{\delta'}$, whenever $\delta \leq \delta' \in T$, such that

(1) for every $\delta \in T$, ν_δ^δ is the identity map, and
(2) for every $\delta \leq \delta' \leq \delta'' \in T$, $\nu_{\delta'}^{\delta''} \circ \nu_\delta^{\delta'} = \nu_\delta^{\delta''}$.

Let $\mathcal{V} = \left(\{V^\delta\}, \{\nu_\delta^{\delta'}\}\right)_{\delta \leq \delta' \in T}$ and $\mathcal{W} = \left(\{W^\delta\}, \{\mu_\delta^{\delta'}\}\right)_{\delta \leq \delta' \in T}$ be two persistence R-modules over T. A *morphism of persistence R-modules* $f \colon \mathcal{V} \to \mathcal{W}$ is a family of morphisms of R-modules $\{f^\delta \colon V^\delta \to W^\delta\}_{\delta \in T}$ such that for every $\delta \leq \delta' \in T$ we have a commutative diagram

$$\begin{array}{ccc} V^\delta & \xrightarrow{\nu_\delta^{\delta'}} & V^{\delta'} \\ \downarrow{\scriptstyle f^\delta} & & \downarrow{\scriptstyle f^{\delta'}} \\ W^\delta & \xrightarrow{\mu_\delta^{\delta'}} & W^{\delta'}. \end{array}$$

Let us assume that R is a field, thus $\mathcal{V} = \left(\{V^\delta\}, \{\nu_\delta^{\delta'}\}\right)_{\delta \leq \delta' \in T}$ is a persistence vector space, and that V^δ is finite-dimensional, for every $\delta \in \mathbb{R}$. Furthermore, we suppose that there exists a finite subset $\{\delta_0, \delta_1, \ldots, \delta_n\} \subseteq T$ for which

(1) if $\delta \in T$, $\delta \leq \delta_0$, then $V^\delta = 0$,
(2) if $\delta \in T \cap [\delta_{i-1}, \delta_i)$ for some $1 \leq i \leq n$, the map $\nu_{\delta_{i-1}}^\delta$ is an isomorphism, while $\nu_\delta^{\delta_i}$ is not, and
(3) if $\delta, \delta' \in T$, $\delta_n \leq \delta < \delta'$, then $\nu_\delta^{\delta'} \colon V^\delta \to V^{\delta'}$ is an isomorphism.

Remark 2.2 When R is a field, either of these restrictions alone (namely, each V^δ being finite-dimensional or the existence of a finite subset $\{\delta_0, \delta_1, \ldots, \delta_n\}$ as above) is enough to associate a *persistence barcode* and a *persistence diagram* to $\mathcal{V}$ (Chazal et al. 2016, Theorem 2.8). We nevertheless assume both restrictions since they hold for persistence diagrams associated to finite simplicial complexes, which are precisely the sort of objects that arise from data, and such assumption makes the exposition simpler.

Let us now introduce persistence barcodes and diagrams. First, for simplicity, we consider $\mathcal{V}$ indexed by the natural numbers:

$$\mathcal{V} = \left(\{V^{\delta_i}\}, \left\{\nu_{\delta_i}^{\delta_{i+1}}\right\}\right)_{i \in \mathbb{N}},$$

where $V^{\delta_k} = V^{\delta_n}$ for all $k \geq n$ and $\nu_{\delta_k}^{\delta_l}$ is the identity whenever $k, l \geq n$. (This clearly contains all the information in $\mathcal{V}$.)

By Edelsbrunner et al. (2015, Basis Lemma), we can find bases B_i of the vector spaces V^{δ_i}, $i \in \mathbb{N}$, such that

(1) $\nu_{\delta_i}^{\delta_{i+1}}(B_i) \subseteq B_{i+1} \cup \{0\}$,

(2) $\operatorname{rank}(v_{\delta_i}^{\delta_{i+1}}) = |v_{\delta_i}^{\delta_{i+1}}(B_i) \cap B_{i+1}|$, and
(3) each $w \in \operatorname{Im}\left(v_{\delta_i}^{\delta_{i+1}}\right) \cap B_{i+1}$ is the image of exactly one element $v \in B_i$.

Such bases are called *compatible bases*. Elements in B_i that are mapped to an element in B_{i+1} correspond to linearly independent elements of V^{δ_i} that 'survive' until the next step in the persistence vector space. Similarly, elements in a basis B_i which are not in B_{i-1} are considered to be 'born' at index i. Formally, we define

$$L := \left\{(b,i) \mid b \in B_i, b \notin \operatorname{Im}\left(v_{\delta_{i-1}}^{\delta_i}\right), i > 0\right\} \cup \{(b,0) \mid b \in B_0\}.$$

Given $(b,i) \in L$, we call i the *birth index* of the basis element b. This element 'survives' on subsequent bases until it is eventually mapped to zero. When this happens, the number of steps taken until the class dies is its *death index* of b. Formally, the death index of (b,i) is

$$\ell(b,i) := \max\left\{j \in \mathbb{N} \mid \left(v_{\delta_{j-1}}^{\delta_j} \circ \cdots \circ v_{\delta_{i+1}}^{\delta_{i+2}} \circ v_{\delta_i}^{\delta_{i+1}}\right)(b) \in B_j\right\}.$$

We allow $\ell(b,i) = \infty$, if b is in every basis B_j for all $j \geq i$, so that the death index takes values in $\overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$. Using this information, we can introduce the *persistence barcode* of $\mathcal{V}$. We call a pair (X, m) a *multiset* if X is a set and $m\colon X \to \overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$ is a function. We call $m(x)$ the *multiplicity* of $x \in X$.

Definition 2.3 Consider a persistence vector space $\mathcal{V} = \left(\{V^{\delta_i}\}, \left\{v_{\delta_i}^{\delta_{i+1}}\right\}\right)_{i\in\mathbb{N}}$ as above. The *persistence barcode of* $\mathcal{V}$ is defined as the multiset of intervals

$$\begin{aligned}\operatorname{Pers}(\mathcal{V}) := &\left\{[\delta_i, \delta_{j+1}) \mid \exists (b,i) \in L,\, j \in \mathbb{N} \text{s.t.} \ell(b,i) = j\right\} \\ &\cup \left\{[\delta_i, +\infty) \mid \exists (b,i) \in L \text{s.t.} \ell(b,i) = +\infty\right\},\end{aligned}$$

with the multiplicity of $[\delta_i, \delta_{j+1})$ (respectively $[\delta_i, +\infty)$) being the number of elements $(b,i) \in L$ such that $\ell(b,i) = j$ (respectively $\ell(b,i) = +\infty$).

Thus, persistence barcodes encode the birth and death of elements in a family of compatible bases. Crucially, and even though compatible bases are not unique, the number of birth and death events at each step is determined by the rank of the linear maps $v_{\delta_i}^{\delta_{i+1}}$ and their compositions, and it is thus independent of the choice of compatible bases. Furthermore, the persistence barcode of a persistence vector space completely determines both the dimension of the vector spaces V^{δ_i} and the ranks of the linear maps between them, hence it determines the persistence vector space up to isomorphism (Chazal et al. 2016, Theorem 2.8).

Each interval in $\operatorname{Pers}(\mathcal{V})$ is called a *persistence interval*. Persistence barcodes can be represented by stacking horizontal lines, each of which represents a persistence interval or bar. The endpoints of the line in the horizontal axis correspond to the endpoints of the interval it represents, whereas the vertical axis has no significance other than being able to represent every persistence interval at once. Persistence bars can be stacked in any order, although they are usually ordered by their birth. This representation is what gives persistence barcodes their name.

An alternative characterization of a persistence vector space is its *persistence diagram*. Let us write $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ for the extended real line.

Definition 2.4 The *persistence diagram* of the persistence vector space $\mathcal{V}$ is the multiset

$$\mathrm{Dgm}(\mathcal{V}) := \left\{(\delta_i, \delta_{j+1}) \in \overline{\mathbb{R}}^2 \mid [\delta_i, \delta_{j+1}) \in \mathrm{Pers}(\mathcal{V})\right\} \cup \left\{(\delta_i, +\infty) \in \overline{\mathbb{R}}^2 \mid [\delta_i, +\infty) \in \mathrm{Pers}(\mathcal{V})\right\}.$$

The multiplicity of a point in $\mathrm{Dgm}(\mathcal{V})$ is the multiplicity of the corresponding interval in $\mathrm{Pers}(\mathcal{V})$.

A crucial property for applications of persistent homology to real data is that a small perturbation of the input (a data set, encoded as filtration of simplicial complexes) results in a small perturbation of the output (its persistence module). Algebraically, this amounts to proving that a small perturbation of the persistence diagram results in a small perturbation of the associated persistence module. In order to state this stability result, we therefore need to introduce distances between persistence diagrams, and persistence modules, respectively.

To measure how far apart two persistence diagrams are, we can use the bottleneck distance between multisets of $\overline{\mathbb{R}}^2$. Let Δ^∞ denote the multiset of $\overline{\mathbb{R}}^2$ consisting of every point in the diagonal counted with infinite multiplicity. A bijection of multisets $\varphi\colon (X, m_X) \to (Y, m_Y)$ is a bijection of sets $\varphi\colon \cup_{x\in X} \sqcup_{i=1}^{m_X(x)} x \to \cup_{y\in Y} \sqcup_{i=1}^{m_Y(y)} y$, that is, a bijection between the sets obtained when counting each element in both X and Y with its multiplicity.

Definition 2.5 Let A and B be two multisets in $\overline{\mathbb{R}}^2$. The *bottleneck distance* between A and B is defined as

$$d_B(A, B) := \inf_{\varphi} \left\{ \sup_{a\in A} \|a - \varphi(a)\|_\infty \right\},$$

where the infimum is taken over all bijections of multisets $\varphi\colon A \cup \Delta^\infty \to B \cup \Delta^\infty$ and where $\|\cdot\|_\infty$ denotes the ℓ^∞-norm in $\overline{\mathbb{R}}^2$.

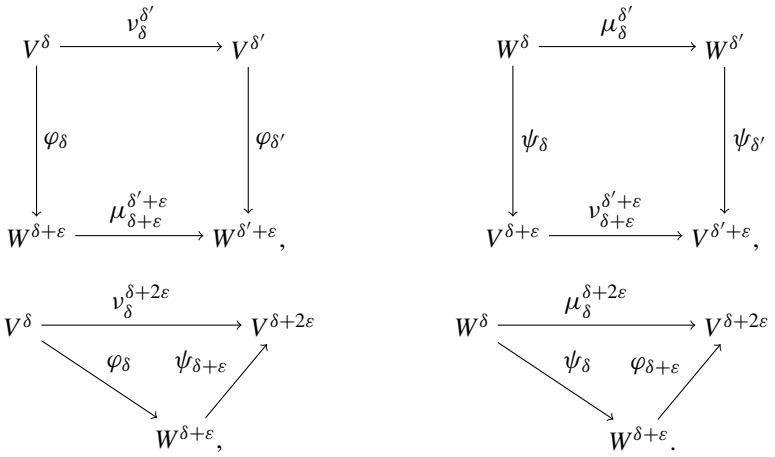
Note that the bottleneck distance between two persistence diagrams defined from persistence barcodes can only be finite if the diagrams have the same amount of bars of form $[a, \infty)$ for $a \in \mathbb{R}$.

We also need a distance between persistence vector spaces. To that purpose, we use the *interleaving distance*, introduced in Chazal et al. (2009).

Definition 2.6 Let $\mathcal{V} = \left(\{V^\delta\}, \{\nu_\delta^{\delta'}\}\right)_{\delta\le\delta'\in\mathbb{R}}$ and $\mathcal{W} = \left(\{W^\delta\}, \{\mu_\delta^{\delta'}\}\right)_{\delta\le\delta'\in\mathbb{R}}$ be two persistence vector spaces, and $\varepsilon \ge 0$. We say that $\mathcal{V}$ and $\mathcal{W}$ are *ε-interleaved* if there exist two families of linear maps

$$\{\varphi_\delta\colon V^\delta \to W^{\delta+\varepsilon}\}_{\delta\in\mathbb{R}}, \quad \{\psi_\delta\colon W^\delta \to V^{\delta+\varepsilon}\}_{\delta\in\mathbb{R}}$$

such that the following diagrams are commutative for all $\delta' \ge \delta \in \mathbb{R}$:



The *interleaving distance* between $\mathcal{V}$ and $\mathcal{W}$ is then defined as

$$d_I(\mathcal{V}, \mathcal{W}) = \inf\{\varepsilon \geq 0 \mid \mathcal{V} \text{ and } \mathcal{W} \text{ are } \varepsilon - \text{interleaved}\}.$$

The authors in Chazal et al. (2009) show that the interleaving distance is a pseudo-metric (a zero distance between distinct points may occur) in the class of persistence vector spaces. Moreover, they show the following Algebraic Stability Theorem.

Theorem 2.7 (Chazal et al. 2009) *Let* $\mathcal{V}=\big(\{V^{\delta}\}, \{\nu_{\delta}^{\delta'}\}\big)_{\delta\leq\delta'\in\mathbb{R}}$ *and* $\mathcal{W}=\big(\{W^{\delta}\}, \{\mu_{\delta}^{\delta'}\}\big)_{\delta\leq\delta'\in\mathbb{R}}$ *be two persistence vector spaces. Then,*

$$d_B\big(\mathrm{Dgm}(\mathcal{V}), \mathrm{Dgm}(\mathcal{W})\big) \leq d_I(\mathcal{V}, \mathcal{W}).$$

2.2 Dissimilarity functions and the correspondence distortion distance

Our objective is to define a theory of persistent homology able to detect directed cycles modulo boundaries. Therefore, instead of building filtrations (of simplicial complexes) from finite metric spaces, we are interested in filtrations (of directed simplicial complexes) built from arbitrary dissimilarity functions. We will follow (Turner 2019), where they receive the name of set-function pairs, and Carlsson et al. (2014), Chowdhury and Mémoli (2018a, b), where they are referred to as dissimilarity networks or asymmetric networks.

In this section, we introduce the necessary background on dissimilarity functions, and the correspondence distortion distance between them, which is a generalization of the Gromov-Hausdorff distance to the asymmetric setting. We finish with a reformulation of this distance, established in Carlsson et al. (2014), which we will need to prove our persistent homology stability result.

Definition 2.8 Let V be a finite set. A *dissimilarity function* (V, d_V) on V is a function $d_V : V \times V \to \mathbb{R}$. The value of d_V on a pair (v_1, v_2) is called the *distance*, or *dissimilarity, from* v_1 *to* v_2.

Note that no restrictions are imposed on d_V, thus it may not be symmetric, the triangle inequality may not hold, and the distance from a point to itself may not be zero. These functions are also referred to as *asymmetric networks* (Carlsson et al. 2014; Chowdhury and Mémoli 2018a, b) since they may be represented as a network with vertex set V and an edge from v_1 to v_2 with weight $d_V(v_1, v_2)$ for every $(v_1, v_2) \in V \times V$. Thus, dissimilarity functions are very flexible and allow for the modelling of widely different problems, including any problem that can be modelled using a network.

We would like to build (directed) simplicial complexes and, ultimately, persistence diagrams from dissimilarity functions. In order to check the stability of our constructions, we need a way to measure how close two such objects are. When comparing networks with the same vertex sets, a natural choice is the ℓ^∞ norm. However, we are interested in comparing dissimilarity functions on different vertex sets. To that end, we consider the ℓ^∞ norm over all possible pairings (quantified by a binary relation) between vertex sets, following ideas similar to those behind the definition of the Gromov-Hausdorff distance.

Definition 2.9 Let (V, d_V) and (W, d_W) be two dissimilarity functions and let θ be a non-empty binary relation between V and W, that is, an arbitrary subset $\theta \subseteq V \times W$. The *distortion* of the relation θ is defined as

$$\operatorname{dis}(\theta) := \max_{(v_1,w_1),(v_2,w_2)\in\theta} |d_V(v_1, v_2) - d_W(w_1, w_2)|.$$

A *correspondence* between V and W is a relation θ between these sets such that $\pi_V(\theta) = V$ and $\pi_W(\theta) = W$, where $\pi_V : V \times W \to V$ is the projection onto V, and similarly for π_W. That is, θ is a correspondence if every element of V is related to at least an element of W, and vice-versa. The set of all correspondences between V and W is denoted $\mathcal{R}(V, W)$.

The *correspondence distortion distance* (Turner 2019) between (V, d_V) and (W, d_W) is defined as

$$d_{\mathrm{CD}}\big((V, d_V), (W, d_W)\big) = \frac{1}{2} \min_{\theta\in\mathcal{R}(V,W)} \operatorname{dis}(\theta).$$

This notion agrees with that of the Gromov–Hausdorff distance when (V, d_V) and (W, d_W) are metric spaces (Burago et al. 2001,Section 7.3.3).

We will use a reformulation of the correspondence distortion that can be found in Chowdhury and Mémoli (2018a). In order to introduce it, we need to define the distortion and co-distortion of maps between sets endowed with dissimilarity functions.

Definition 2.10 Let (V, d_V) and (W, d_W) be any two dissimilarity functions and let $\varphi : V \to W$ and $\psi : W \to V$ be maps of sets. The *distortion* of φ (with respect to d_V and d_W) is defined as

$$\operatorname{dis}(\varphi) := \max_{v_1,v_2\in V} \big|d_V(v_1, v_2) - d_W\big(\varphi(v_1), \varphi(v_2)\big)\big|.$$

The *co-distortion* of φ and ψ (with respect to d_V and d_W) is defined as

$$\operatorname{codis}(\varphi, \psi) := \max_{(v,w)\in V\times W} \left| d_V\big(v, \psi(w)\big) - d_W\big(\varphi(v), w\big) \right|.$$

Note that codistortion is not necessarily symmetrical, namely, $\operatorname{codis}(\varphi, \psi)$ and $\operatorname{codis}(\psi, \varphi)$ may be different if either of the dissimilarity functions are asymmetric.

Finally, we have the following reformulation of the correspondence distortion distance.

Proposition 2.11 (Chowdhury and Mémoli 2018a, Proposition 9) *Let* (V, d_V) *and* (W, d_W) *be any two dissimilarity functions. Then,*

$$d_{\mathrm{CD}}\big((V, d_V), (W, d_W)\big) = \frac{1}{2} \min_{\substack{\varphi:\, V\to W,\\ \psi:\, W\to V}} \big\{ \max\{\operatorname{dis}(\varphi), \operatorname{dis}(\psi), \operatorname{codis}(\varphi, \psi), \operatorname{codis}(\psi, \varphi)\} \big\}.$$

3 Directed homology of directed simplicial complexes

In this section, we introduce and study a family of subrings of the homology rings of simplicial complexes which encode directionality information and in particular detect directed cycles (Fig. 1). We cannot introduce such a family using (undirected) simplicial complexes, as they cannot encode directionality information of the simplices. One approach is to use the so-called *ordered-set complexes*, where simplices are sets with a total order on the vertices. They generalize simplicial complexes, which can be encoded as fully symmetric ordered-set complexes, that is, ordered-set complexes where if a set of vertices forms a simplex, it must do so with every possible order. However, persistent homology of ordered-set complexes is not stable (Remark 4.20).

To achieve stability, we use one further generalization, called *ordered tuple complexes* or *OT-complexes* in Turner (2019), and called *directed simplicial complexes* in this article (Definition 3.1). The only difference is that arbitrary repetitions of vertices are allowed in the ordered tuples representing simplices. Clearly, any ordered-set complex is a directed simplicial complex. Furthermore, a (undirected) simplicial complex can be encoded as a fully symmetric (as above) directed simplicial complex where every possible repetition of vertices is also included. When doing so, the (undirected) simplicial complex and its associated directed simplicial complex have isomorphic homologies over rings (Remark 3.5), and morphisms of simplicial complexes can be lifted to morphisms between their associated directed simplicial complexes (Remark 3.9). Finally, and crucially for us, we will be able to use the above-mentioned subrings encoding information on directed cycles to introduce a theory of persistent homology for dissimilarity functions which is stable with respect to the correspondence-distortion distance (Sect. 4). Note that, as ordered-set complexes are directed simplicial complexes, the results stated here apply to homology computations on ordered-set complexes as well.

3.1 Chain complexes of directed simplicial complexes

In this section, we introduce directed simplicial complexes and their associated chain complexes and homology modules, including modules of directed homology.

Definition 3.1 A *directed simplicial complex*, or *ordered tuple complex*, is a pair (V, X) where V is a finite set of *vertices*, and X is a family of tuples $(x_0, x_1, \ldots, x_n)$ of elements of V such that if $(x_0, x_1, \ldots, x_n) \in X$, then $(x_0, x_1, \ldots, \widehat{x_i}, \ldots, x_n) \in X$ for every $i = 0, 1, \ldots, n$. (Here and throughout the rest of the paper, $\widehat{x_i}$ indicates that x_i is omitted from the list.)

We will denote the directed simplicial complex (V, X) just by X, and assume that every vertex belongs to at least one directed simplex (so that V is uniquely determined from X). Elements of X of length $n + 1$ are called *n-simplices*, and the subset of the n-simplices of X is denoted by X_n. Note that arbitrary repetitions of vertices in a tuple are allowed.

Note that directed simplicial complexes are more general objects than simplicial complexes, in a way that allows directed simplicial complexes to encode asymmetric information. Note further that it is possible to define a geometric realization for these objects in which simplices with vertex repetitions are collapsed to a simplex with no repetitions, thus simplices with repeated vertices play a role similar to that of the degenerate simplices in simplicial sets.

We make use of the following terminology when dealing with directed simplicial complexes. Elements of X of length $n + 1$ are called *n-simplices*, and the subset of the n-simplices of X is denoted by X_n. An n-simplex obtained by removing vertices from an m-simplex, $n \leq m$, is said to be a *face* of the m-simplex. Such face is called *proper* if $n < m$, or equivalently, if at least vertex of the m-simplex is removed. A directed simplicial complex is said to be n-dimensional, denoted $\dim(X) = n$, if X_{n+1} is trivial (empty) but X_n is not. A collection Y of simplices of X that is itself a directed simplicial complex is said to be a *directed simplicial subcomplex* (or, simply, *subcomplex*) of X, denoted $Y \subseteq X$. Note that the vertex set of Y may be strictly smaller than that of X.

We now introduce chain complexes associated to a directed simplicial complex. These definitions are straight generalizations of those for (undirected) simplicial complexes.

Definition 3.2 Let R be a commutative ring. The *n-dimensional chains* of X are defined as the elements of the free R-module generated by the n-simplices,

$$C_n(X, R) = R\big(\{[x_0, x_1, \ldots, x_n] \mid (x_0, x_1, \ldots, x_n) \in X\}\big).$$

We call the elements $[x_0, x_1, \ldots, x_n] \in C_n(X, R)$, $(x_0, x_1, \ldots, x_n) \in X$, *elementary n-chains*.

Definition 3.3 Let X be a directed simplicial complex. For each $n > 0$, we define a morphism of R-modules $\partial_n \colon C_n(X, R) \to C_{n-1}(X, R)$ by

$$\partial_n([x_0, x_1, \ldots, x_n]) = \sum_{i=0}^{n} (-1)^i [x_0, x_1, \ldots, \widehat{x_i}, \ldots, x_n]$$

For $n = 0$, let $\partial_0 \colon C_0(X, R) \to \{0\}$ be the trivial map.

A straightforward computation analogous to the one for chain complexes in simplicial homology shows that $\{C_n(X, R), \partial_n\}$ is a chain complex of R-modules. In particular, we can define homology groups.

Definition 3.4 Let X be a directed simplicial complex, R a commutative ring, and $n \geq 0$. The n-dimensional *cycles*, *boundaries*, and *homology* of X are $Z_n(X, R) = \{x \in C_n(X, R) \mid \partial_n(x) = 0\}$, $B_n(X, R) = \operatorname{Im} \partial_{n+1}$ and the quotient $H_n(X, R) = Z_n(X, R)/B_n(X, R)$, respectively.

Remark 3.5 If X is an (undirected) simplicial complex, we can define an ordered-set simplicial complex X^{OT} where $(x_0, x_1, \ldots, x_n) \in X^{\mathrm{OT}}$ whenever $\{x_0, x_1, \ldots, x_n\}$, after removing any repetition of vertices, is a simplex in X. The chain complex $C_*(X^{\mathrm{OT}}, R)$ is called the *ordered chain complex* of X in Munkres (1984), and its homology is isomorphic to the singular homology of X over R.

We now discuss how to only retain the information associated to directed cycles. Note that when introducing the differential for the chain complexes of simplicial complexes, an orientation is chosen for each of the simplices. However, undirected cycles can still be produced because the chosen orientation for a simplex can be reverted by changing the sign of the corresponding elementary chain. As such, homology of (undirected) simplicial complexes is independent of these choices of orientations.

In the case of directed simplicial complexes, orientations are not chosen, and we can ensure that we are only retaining information regarding directed cycles by making sure that all the chains have positive coefficients, when the coefficient ring is chosen appropriately. However, we also require the introduced object to be a submodule of the entire module of homology. Thus, we arrive at the following definition.

Definition 3.6 Let X be a directed simplicial complex and let $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. The set of *$n$-dimensional directed cycles* of X over R is

$$Z_n^{\mathrm{Dir}}(X, R) = \left\{ \sum_i \lambda_i c_i \in Z_n(X, R) \mid \lambda_i \geq 0, c_i \text{ elementary } n\text{-chain} \right\}.$$

The nth dimensional module of directed homology, $H_n^{\mathrm{Dir}}(X, R)$, is the submodule of $H_n(X, R)$ generated by the directed cycles,

$$H_n^{\mathrm{Dir}}(X, R) = \left\langle \left\{ [x] \mid x \in Z_n^{\mathrm{Dir}}(X, R) \right\} \right\rangle.$$

Remark 3.7 The rings $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are used to simplify the exposition: the definition of $Z_n^{\mathrm{Dir}}(X, R)$, and thus of $H_n^{\mathrm{Dir}}(X, R)$, can be made in an analogous way for any ring R that is the Grothendieck completion of a cancellative, zerosumfree semiring Λ (see Appendix 1).

In view of Definition 3.6 it is clear that, in dimension 1, directed circuits in the directed graph underlying a directed simplicial complex will give rise to directed cycles. In Sect. 3.3 we will show how these directed submodules of homology behave in a desirable way.

3.2 Functoriality of directed homology

In this section, we prove that homology and directed homology are functors from the category of directed simplicial complexes to the category of graded R-modules. We also show that two morphisms allowing for the construction of the prism operator must induce the same map on homology, a result we will need to prove our persistent homology stability result. We begin by introducing morphisms of directed simplicial complexes.

Definition 3.8 A *morphism of directed simplicial complexes* $f\colon (V, X) \to (W, Y)$ or just $f\colon X \to Y$, is a map of sets $f\colon V \to W$ such that if $(x_0, x_1, \ldots, x_n) \in X$ then $\big(f(x_0), f(x_1), \ldots, f(x_n)\big) \in Y$.

Remark 3.9 This definition is stricter than the classical notion of morphism of simplicial complexes, where morphisms are allowed to take simplices to simplices of lower dimension, as directed simplicial complexes can intrinsically account for vertex repetitions. However, note that if X and Y are (undirected) simplicial complexes, a map $f\colon X \to Y$ is a morphism of simplicial complexes if and only if $f\colon X^{\mathrm{OT}} \to Y^{\mathrm{OT}}$ (see Remark 3.5) is a morphism of directed simplicial complexes.

Definition 3.10 Let $f\colon X \to Y$ be a morphism of directed simplicial complexes. Then f induces morphisms of R-modules $C(f) = \{C_n(f)\}$, defined as

$$\begin{aligned} C_n(f)\colon C_n(X, R) &\longrightarrow C_n(Y, R) \\ [x_0, x_1, \ldots, x_n] &\longmapsto [f(x_0), f(x_1), \ldots, f(x_n)]. \end{aligned}$$

We will often abbreviate $C_n(f) = f_n$.

Proposition 3.11 *If $f\colon X \to Y$ is a morphism of directed simplicial complexes, the family of maps $\{C_n(f)\}$ is a R-homomorphism of chain complexes. Therefore, it induces a map $H_n(f)\colon H_n(X, R) \to H_n(Y, R)$. Furthermore, $H_n(f)$ restricts to a map $H_n^{\mathrm{Dir}}(f)\colon H_n^{\mathrm{Dir}}(X, R) \to H_n^{\mathrm{Dir}}(Y, R)$.*

Proof Let $[x_0, x_1, \ldots, x_n] \in C_n(X, R)$ be a simplex. Simple computations show that $C_{n-1}(f)\partial_n = \partial_n C_n(f)$, thus inducing a map $H_n(f)\colon H_n(X, R) \to H_n(Y, R)$.

Now let $x \in Z_n^{\mathrm{Dir}}(X, R)$. We can write $x = \sum_i \lambda_i x_i$, where $\lambda_i \geq 0$ and x_i is an elementary n-chain. Then, $C_n(f)(x) = \sum_i \lambda_i C_n(f)(x_i)$. Clearly, $C_n(f)$

takes elementary n-chains to elementary n-chains, thus $C_n(f)(x) \in Z_n^{\mathrm{Dir}}(Y, R)$. It is then clear from the definition of directed homology that $H_n(f)$ restricts to $H_n(f)\colon H_n(X, R) \to H_n(Y, R)$. □

Remark 3.12 Let X and Y be (undirected) simplicial complexes and let R be a ring. If $f\colon X \to Y$ is a morphism of (undirected) simplicial complexes, the map induced on homology by $f\colon X \to Y$ is the same as the map induced on homology by $f\colon X^{\mathrm{OT}} \to Y^{\mathrm{OT}}$ (see Remarks 3.5 and 3.9).

Corollary 3.13 (Directed) *Homology is a functor from the category of directed simplicial complexes to the category of graded R-modules. In particular, isomorphic directed simplicial complexes have isomorphic (directed) homologies.*

We finish this section by showing a sufficient condition for two morphisms to induce the same map on homology. We will need this result to prove that our definition of persistent homology is stable.

Lemma 3.14 *Let R be a commutative ring. Let X and Y be two directed simplicial complexes and let $f, g\colon X \to Y$ be morphisms of directed simplicial complexes such that if $(x_0, x_1, \ldots, x_n) \in X$, then $\big(f(x_0), f(x_1), \ldots, f(x_i), g(x_i), \ldots, g(x_n)\big) \in Y$ for every $i = 0, 1, \ldots, n$. Then, $H_n(f) = H_n(g)$ for every $n \geq 0$. Consequently, $H_n^{\mathrm{Dir}}(f) = H_n^{\mathrm{Dir}}(g)$, for every $n \geq 0$.*

Proof For $x = [x_0, x_1, \ldots, x_n] \in C_n(X, R)$ an elementary n-chain, define

$$s_n[x_0, x_1, \ldots, x_n] = (-1)^i \sum_{i=0}^{n} [f(x_0), \ldots, f(x_i), g(x_i), \ldots, g(x_n)].$$

We show that this family of maps provides a chain homotopy between f and g. Namely, we prove that $C_n(g) - C_n(f) = \partial_{n+1} s_n + s_{n-1}\partial_n$, for all n. On the one hand,

$$s_{n-1}\partial_n(x) = \sum_{i=0}^{n} (-1)^i \left[\sum_{j=0}^{i-1} (-1)^j \big[f(x_0), f(x_1), \ldots, f(x_j), g(x_j), \ldots, \widehat{g(x_i)}, \ldots, g(x_n)\big] + \sum_{j=i+1}^{n} (-1)^{j+1} \big[f(x_0), f(x_1), \ldots, \widehat{f(x_i)}, \ldots, f(x_j), g(x_j), \ldots, g(x_n)\big] \right]. \tag{3.1}$$

On the other hand,

$$\partial_{n+1} s_n(x) = \sum_{j=0}^{n} (-1)^j \left[\sum_{i=0}^{j} (-1)^i \big[f(x_0), f(x_1), \ldots, \widehat{f(x_i)}, \ldots, f(x_j), g(x_j), \ldots, g(x_n)\big] + \sum_{i=j}^{n} (-1)^{i+1} \big[f(x_0), f(x_1), \ldots, f(x_j), g(x_j), \ldots, \widehat{g(x_i)}, \ldots, g(x_n)\big] \right].$$

By exchanging the roles of the indices in this last equation, we obtain that

$$\partial_{n+1}s_n(x) = \sum_{i=0}^{n}\left[\sum_{j=0}^{i}(-1)^{i+1}(-1)^{j}\big[f(x_0), f(x_1), \ldots, f(x_j), g(x_j), \ldots, \widehat{g(x_i)}, \ldots, g(x_n)\big] + \sum_{j=i}^{n}(-1)^{i}(-1)^{j}\big[f(x_0), f(x_1), \ldots, \widehat{f(x_i)}, \ldots, f(x_j), g(x_j), \ldots, g(x_n)\big]\right]. \tag{3.2}$$

Now, adding Eqs. (3.1) and (3.2),

$$\partial_{n+1}s_n(x) + s_{n-1}\partial_n(x) = \sum_{i=0}^{n}[f(x_0), f(x_1), \ldots, f(x_{i-1}), g(x_i), \ldots, g(x_n)] - \sum_{i=0}^{n}\big[f(x_0), f(x_1), \ldots, f(x_i), g(x_{i+1}), \ldots, g(x_n)\big].$$

It is now straightforward to check that this sum amounts to

$$[g(x_0), g(x_1), \ldots, g(x_n)] - [f(x_0), f(x_1), \ldots, f(x_n)] = C_n(g)(x) - C_n(f)(x),$$

and we are done. □

3.3 Basic computations and examples

In this section, we explore some basic properties of this homology theory. We begin by studying the relation between homology and connectivity.

Definition 3.15 Let (V, X) be a directed simplicial complex and $v, w \in V$ be two vertices. A *path* from v to w in X a sequence of vertices $v = x_0, x_1, \ldots, x_n = w$ such that either (x_i, x_{i+1}) or (x_{i+1}, x_i) is a simplex, for all $1 \leq i \leq n-1$. The *(weakly) connected components* of X are the equivalence classes of the equivalence relation $v \sim w$ if there is a path in X from v to w. We call X *(weakly) connected* if it has only one connected component, that is, every pair of vertices can be connected by a path.

Note that this notion of connectedness ignores the direction of the edges (1-simplices). The next result shows that we can compute the homology of each connected component independently. The proof follows by observing that the chain complex $C_n(X, R)$ is clearly the direct sum of the chain complexes associated to each of the (weakly) connected components of X. In particular, it is immediate that any directed cycle can be decomposed as a sum of directed cycles in each of the directed components.

Proposition 3.16 *The (directed) homology of a directed simplicial complex X is isomorphic to the direct sum of the (directed) homology modules of its (weakly) connected components.*

We next show that the 0th (directed) homology counts the number of (weakly) connected components of a directed simplicial complex X. We start with a lemma.

Lemma 3.17 *Let R be a commutative ring and X be a directed simplicial complex. For any vertex v and any $\lambda \in R$, $\lambda[v]$ is homologous to the trivial class if and only if $\lambda = 0$.*

Proof It is clear that $\lambda[v]$ is a cycle as ∂_0 is trivial. On the other hand, it is clear that it must be the trivial class if $\lambda = 0$. Now assume that $\lambda[v]$ is the trivial class. We shall prove that $\lambda = 0$.

Assume that X has $n+1$ vertices, namely $V = \{x_0 = v, x_1, \ldots, x_n\}$. Then if $\lambda[x_0]$ is trivial, there exists a 1-chain

$$x = \sum_{i,j=0}^{n} \lambda_{ij}[x_i, x_j]$$

such that $\lambda[x_0] = \partial(x)$, where we are assuming that $\lambda_{ij} = 0$ whenever $[x_i, x_j]$ is not a 1-chain. By computing the differentials,

$$\lambda[x_0] = \sum_{i,j=0}^{n} \lambda_{ij}[x_j] - \sum_{i,j=0}^{n} \lambda_{ij}[x_i].$$

Now, since $\{[x_0], [x_1], \ldots, [x_n]\}$ is a basis of $C_0(X, R)$, the coefficients of each element in the basis must be equal in both sides of the equality. Namely, for each $j \neq 0$, $0 = \sum_{i=0}^{n}(\lambda_{ij} - \lambda_{ji})$, and $\lambda = \sum_{i=0}^{n}(\lambda_{i0} - \lambda_{0i})$. Adding these equations, we deduce that $\lambda = 0$. □

Proposition 3.18 *Let X be a directed simplicial complex and let R be a commutative ring. Then $H_0(X, R) \cong R^k$, where k is the number of (weakly) connected components of X. Furthermore, if $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, then $H_0(X, R) = H_0^{\mathrm{Dir}}(X, R) \cong R^k$.*

Proof By Proposition 3.16 we only need to show that if X is connected, then $H_0(X, R) \cong R$. Let v, w be any two vertices and let us show that $[v]$ and $[w]$ are homologous 0-cycles. Indeed, since X is connected, we can find vertices $v = v_0, v_1, \ldots, v_m = w$ such that either $[v_i, v_{i+1}]$ or $[v_{i+1}, v_i]$ are 1-chains, for all $i = 0, 1, \ldots, n-1$. If either $[v_i, v_{i+1}]$ or $[v_{i+1}, v_i]$ are chains, then $[v_i]$ and $[v_{i+1}]$ are homologous. As this holds for every $i = 0, 1, \ldots, m-1$, $[v]$ is homologous to $[w]$. Now, if $\sum_{i=1}^{n} \lambda_i[x_i]$ is any 0-cycle, it is homologous to $\sum_{i=1}^{n} \lambda_i[v]$. Then, by Lemma 3.17, $\lambda[v]$ is not homologous to $\mu[v]$ whenever $\lambda \neq \mu$, thus $H_0(X, R) \cong R$.

For the second statement, if $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, we can compute $H_n^{\mathrm{Dir}}(X, R)$. In such case, $[v]$ is a directed cycle, thus $\langle [v] \rangle = H_0(X, R) \subseteq H_0^{\mathrm{Dir}}(X, R)$. It then follows that $H_0(X, R) = H_0^{\mathrm{Dir}}(X, R)$. □

Remark 3.19 The fact that 0th directed homology ignores edge directions seems counter-intuitive, as we set up this framework to detect directed features, namely directed cycles. However, this is a natural consequence of homology being an equivalence relation. In particular, as mentioned in the proof, if either $[u, v]$ or $[v, u]$ are

1-chains, $[u]$ and $[v]$ are homologous, thus directionality of the 1-chains does not affect 0th directed homology. However, this only occurs in dimension 0, as a consequence of 0-chains always being cycles. Thus, when working with 1-simplices, the symmetry in the homology relation does not affect the detection of asymmetry in the data, as such asymmetry is encoded in the cycles themselves. It is also worth mentioning that a persistent homology able to detect strongly connected components in directed graphs has been introduced in Turner (2019).

We now show that the homology of a point is trivial for positive indices.

Example 3.20 Let R be a commutative ring and let X be the directed simplicial complex with vertex set $V = \{x_0\}$ and simplices $X = \{(x_0)\}$. Then,

$$H_n(X, R) = \begin{cases} R, & \text{if } n = 0, \\ 0, & \text{if } n > 0. \end{cases}$$

Indeed, X is connected, thus $H_0(X, R) = R$ by Proposition 3.18. For $n \geq 1$, there are no n-chains, hence $H_n(X, R) = 0$. In this case, it is also immediate to check that if $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, $H_n^{\text{Dir}}(X, R) = H_n(X, R)$, for every n.

Directed simplicial complexes whose homology is isomorphic to that of the point are called *acyclic*.

Definition 3.21 A directed simplicial complex X is *acyclic* if

$$H_n(X, R) = \begin{cases} R, & \text{if } n = 0, \\ 0, & \text{if } n > 0. \end{cases}$$

If $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, we say that X is *directionally acyclic* if

$$H_n^{\text{Dir}}(X, R) = \begin{cases} R, & \text{if } n = 0, \\ 0, & \text{if } n > 0. \end{cases}$$

Note that acyclic implies directionally acyclic, as directed homology is a submodule of the module of homology, for every n.

Next, we show that a directed simplicial complex consisting of one m-dimensional simplex along with all of its faces (note that this is an ordered-set complex) is acyclic.

Proposition 3.22 *Let X be the directed simplicial complex consisting of the simplex $(x_0, x_1, \ldots, x_m)$ and all of its faces. Then X is acyclic.*

Proof As a consequence of how X is defined, if $x \in C_n(X, R)$ is a chain so that x_0 does not participate in any of its elementary n-chains (as there are no repetitions in the simplices of X), then x_0 can be added as the first element of every elementary

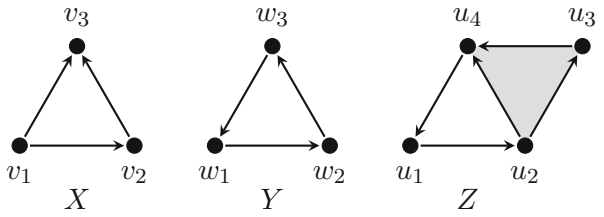


Fig. 3 Directed homology detects directed 1-cycles modulo boundaries. In these examples, $H_1^{\mathrm{Dir}}(X, R) = 0$ while $H_1^{\mathrm{Dir}}(Y, R) = H_1^{\mathrm{Dir}}(Z, R) = R$ respectively generated by $[w_1, w_2] + [w_2, w_3] + [w_3, w_1]$ and $[u_1, u_2] + [u_2, u_4] + [u_4, u_1]$ (or $[u_1, u_2] + [u_2, u_3] + [u_3, u_4] + [u_4, u_1]$). To be a directed 1-cycle, the directions of the involved 1-simplices must form circuit. This does not hold for (undirected) homology

n-chain in x, giving us an $(n+1)$-chain which we denote $x_0 x \in C_{n+1}(X, R)$. Simple computations show that

$$\partial(x_0 x) = x - x_0 \partial(x).$$

Now take $x \in Z_n(X, R)$ any cycle and decompose it as $x = x_0 y + z$, where x_0 does not participate in any of the chains in either y or z. Using the formula above,

$$\partial(x) = \partial(x_0 y + z) = y - x_0 \partial(y) + \partial(z) = 0.$$

In particular, the chains in which x_0 does not participate must add up to zero, namely $y + \partial(z) = 0$. Finally, consider the chain $x_0 z$. We have that

$$\partial(x_0 z) = z - x_0 \partial(z) = z + x_0 y = x.$$

Thus, $x \in Z_n(X, R)$ is homologous to zero, and the result follows. □

We now illustrate the ability of this homology theory to detect directed cycles. We do so through some simple examples.

Example 3.23 Let $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. Let X, Y and Z be the directed simplicial complexes represented in Fig. 3. These three complexes are connected, so their 0th homologies are R, both directed and undirected. Furthermore, neither X nor Y have k-simplices for $k \geq 2$, so $H_k(X, R)$ and $H_k(Y, R)$ are trivial (zero) for every $k \geq 2$. This in particular implies that $H_k^{\mathrm{Dir}}(X, R)$ and $H_k^{\mathrm{Dir}}(Y, R)$ are trivial as well. Although Z has one 2-simplex, it is not a 2-cycle, thus $H_2(Z, R)$ is also trivial.

To compute the first homology, we need to consider the 1-simplices and their differentials. We list them below.

X	∂
$[v_1, v_2]$	$v_2 - v_1$
$[v_2, v_3]$	$v_3 - v_2$
$[v_1, v_3]$	$v_3 - v_1$

Y	∂
$[w_1, w_2]$	$w_2 - w_1$
$[w_2, w_3]$	$w_3 - w_2$
$[w_3, w_1]$	$w_1 - w_3$

Z	∂
$[u_1, u_2]$	$u_2 - u_1$
$[u_2, u_3]$	$u_3 - u_2$
$[u_3, u_4]$	$u_4 - u_3$
$[u_4, u_1]$	$u_1 - u_4$
$[u_2, u_4]$	$u_4 - u_2$

We start with X. Note that $\lambda_1[v_1, v_2] + \lambda_2[v_2, v_3] + \lambda_3[v_1, v_3]$, $\lambda_1, \lambda_2, \lambda_3 \in R$, is a cycle for X if and only if $\lambda_1 = \lambda_2 = -\lambda_3$. Since there are no 2-cycles in X, different 1-cycles cannot be homologous. We deduce that $H_1(X, R) \cong R$. However, no 1-cycle can have non-negative coefficients for the three elementary 1-cycles, thus $H_1^{\mathrm{Dir}}(X, R) = \{0\}$.

Regarding Y, note that $\mu_1[w_1, w_2] + \mu_2[w_2, w_3] + \mu_3[w_3, w_1]$, $\mu_1, \mu_2, \mu_3 \in R$, is a cycle if and only if $\mu_1 = \mu_2 = \mu_3$. Since there are no 2-simplices in Y, we deduce that $H_1(Y, R) = R$. Furthermore, the coefficients can be simultaneously positive. Namely, $[w_1, w_2] + [w_2, w_3] + [w_3, w_1]$ is a directed cycle, which generates the entirety of $H_1(X, R)$. We deduce that $H_1^{\mathrm{Dir}}(X, R) \cong R$. Note that X and Y have isomorphic homology rings, but their directed homology rings are different.

Finally, for Z, similar computations show that $Z_1(Z, R)$ is the free R-module generated by $x_1 = [u_1, u_2] + [u_2, u_3] + [u_3, u_4] + [u_4, u_1]$ and $x_2 = [u_1, u_2] + [u_2, u_4] + [u_4, u_1]$, both of which are directed cycles. In particular, $Z_1^{\mathrm{Dir}}(Z, R)$ consists of linear combinations of those two cycles, with positive coefficients. Note, however, that generating sets of $Z_1(X, R)$ containing fewer than two directed cycles exist. This shows that computing a generating set for the cycles may not be enough to compute the full set of directed cycles.

Continuing with the homology computations, note that in this case we have a 2-simplex, $y = [u_2, u_3, u_4]$, for which $\partial(y) = [u_3, u_4] - [u_2, u_4] + [u_2, u_3] = x_1 - x_2$. Namely, x_1 and x_2 are homologous. Therefore, $H_1(Z, R) = H_1^{\mathrm{Dir}}(Z, R) = R$, generated by either x_1 or x_2. In particular, both the directed and undirected homologies of Y and Z are isomorphic, and we can see how 2-simplices can make directed cycles equivalent, as expected.

Fig. 4 A directed simplicial complex with just two vertices can have a non-trivial 1-cycle

More generally, if $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, a *polygon* will give rise to a non-trivial directed homology class in dimension 1 if, and only if, the cycle can be traversed following the direction of the edges.

Proposition 3.24 *Let X be a 1-dimensional directed simplicial complex with vertices $v_0, v_1, \ldots, v_n$, $n \geq 1$, and with 1-simplices $e_0, e_1, \ldots, e_n$ where either $e_i = (v_i, v_{i+1})$ or $e_i = (v_{i+1}, v_i)$, $i = 0, 1, \ldots, n$, and we write $v_{n+1} = v_0$. Then, $H_1(X, R) = R$. If $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, then $H_1^{\mathrm{Dir}}(X, R)$ is non-trivial if and only if either $e_i = (v_i, v_{i+1})$ for all i or $e_i = (v_{i+1}, v_i)$ for all i, in which case $H_1^{\mathrm{Dir}}(X, R) \cong R$.*

Proof We first prove that $H_1(X, R) = R$. Let $x = \sum_{i=0}^{n} \lambda_i [e_i]$ be a cycle. If x is non-trivial, we can assume without loss of generality that $\lambda_0 \neq 0$. We can further assume that $e_0 = (v_0, v_1)$. Thus, $\partial(\lambda_0[e_0]) = \lambda_0[v_1] - \lambda_0[v_0]$. The only other edge in which e_1 participates is e_1. We deduce that $\lambda_1 = \lambda_0$, if $e_1 = (v_1, v_2)$, or $\lambda_1 = -\lambda_0$, if $e_1 = (v_2, v_1)$. By proceeding iteratively, we deduce that $\lambda_i = \lambda_0$, if $e_i = (v_i, v_{i+1})$, or that $\lambda_i = -\lambda_0$, if $e_i = (v_{i+1}, v_i)$. It is now straightforward to check that, with λ_i defined in terms of λ_0 in such way, $i = 1, \ldots, n$, $x = \sum_{i=0}^{n} \lambda_i [e_i]$ is a non-trivial cycle. Since there are no 2-chains, $H_1(X, R) \cong R$.

Regarding the directed homology, note that if $e_0 = (v_0, v_1)$, the coefficients of the cycle x can only be all positive if and only if $e_i = (v_i, v_{i+1})$. In such case, $H_1^{\mathrm{Dir}}(X, R) = H_1(X, R) \cong R$, and otherwise $H_1^{\mathrm{Dir}}(X, R) = 0$. The case in which $e_0 = (v_1, v_0)$ is analogous. □

Remark 3.25 Note that Proposition 3.24 applies to polygons with only two vertices v_1 and v_2 (see Fig. 4), which are allowed in a directed simplicial complex. Indeed, an immediate computation shows that $[v_1, v_2] + [v_2, v_1]$ is a directed 1-cycle. This result shows that directed homology is not a homotopy invariant for the geometric realization of directed simplicial complexes, as all of the polygons considered in Proposition 3.24 give rise to homotopic geometric realizations.

We end this section by noting that, although our initial aim was a directed homology group able to detect directed 1-cycles, we have a directed homology theory defined in all dimensions. However, we can show that the directed homology is trivial in even dimensions other than 0.

Proposition 3.26 *Let X be a directed simplicial complex $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. Then $H_{2n}^{\mathrm{Dir}}(X, R) = \{0\}$, for every $n \geq 1$.*

Proof We will show that no non-trivial directed cycles may exist in such dimensions. In order to do so, consider the morphism of R-modules

$$\begin{aligned} \varphi_n \colon C_n(X, R) &\longrightarrow R \\ [v_0, v_1, \ldots, v_n] &\longmapsto 1. \end{aligned}$$

Thus, if $x = \sum_{i=0}^{k} \lambda_i x_i$ where x_i is an elementary n-chain, $\varphi_n(x) = \sum_{i=0}^{k} \lambda_i$.

Now assume that x_i is an elementary $2n$-chain, for some $n \geq 1$. In this case, $\partial(\lambda_i x_i) = \lambda_i \partial(x_i)$, where $\partial(x_i)$ consists of the sum of $\lfloor \frac{2n}{2} \rfloor = n$ elementary n-chains with positive sign and $\lfloor \frac{2n-1}{2} \rfloor = n-1$ elementary n-chains with negative sign. Thus, $\varphi_{2n-1}(\lambda_i x_i) = \lambda_i(n-(n-1)) = \lambda_i$.

To prove our claim, let $x = \sum_{i=0}^{k} \lambda_i x_i \in Z_{2n}(X, R)$ be a cycle where each x_i is an elementary $2n$-cycle. Then, given that $\partial(x) = 0$, $\varphi_{2n-1}\big(\partial(x)\big) = \sum_{i=0}^{k} \lambda_i = 0$. If x is directed, then $\lambda_i \geq 0$, for all i, thus we deduce that $\lambda_i = 0$, for all i. Namely, the trivial cycle is the only directed cycle, and $H_{2n}^{\mathrm{Dir}}(X, R) = \{0\}$. □

Remark 3.27 The proof of this result is based on the fact that the number of positive coefficients in the boundary of an elementary n-chain is larger than the number of negative coefficients, if n is even. This may also be related to the fact that there is no obvious way to define directed n-cycles for $n > 1$. For the purposes of this article, it suffices to consider 1-cycles and $H_1(X, R)$. However, note that non-trivial (homological) cycles do exist in all odd dimensions. For instance, the elementary $(2n-1)$-chain obtained by repeating a vertex $2n$ times is always a cycle.

4 Persistent directed homology

In this section, we introduce a theory of persistent homology for directed simplicial complexes which comes in two flavours: one takes into account the directionality of the complex, whereas the other one is analogous to persistent homology in the classical setting. We thus have, associated to the same filtration of directed simplicial complexes, two persistence modules which produce two different barcodes. Both persistent homology theories show stability (see Theorem 4.19) and they are closely related. Indeed, directed cycles are undirected cycles as well, thus every bar in a directed persistence barcode can be uniquely matched with a bar in the corresponding undirected one, although the undirected bar may be born sooner, see Proposition 4.7.

4.1 Persistence modules associated to a directed simplicial complex

Let us begin by introducing filtrations of directed simplicial complexes.

Definition 4.1 Let X be a directed simplicial complex (Definition 3.1). A *filtration* of X is a family of subcomplexes $(X_\delta)_{\delta \in T}$, $T \subseteq \mathbb{R}$, such that if $\delta \leq \delta' \in T$, then X_δ is a subcomplex of $X_{\delta'}$, and such that $X = \cup_{\delta \in T} X_\delta$. Note that for $\delta \leq \delta'$, the inclusion $i_\delta^{\delta'} \colon X_\delta \to X'_\delta$ is a morphism of directed simplicial complexes.

We now introduce undirected persistence modules.

Definition 4.2 Let $(X_\delta)_{\delta \in T}$ be a filtration of a directed simplicial complex X and let R be a commutative ring. The n-dimensional *undirected persistence R-module* of X is the persistence R-module

$$\big(\{H_n(X_\delta, R)\}, \{H_n(i_\delta^{\delta'})\}\big)_{\delta \leq \delta' \in T}.$$

The functoriality of H_n makes this a persistence module.

Now, in order to retain the information on directionality, we take the submodule of directed classes.

Definition 4.3 Let $(X_\delta)_{\delta\in T}$ be a filtration of a directed simplicial complex X and let $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. The n-dimensional *directed persistence R-module* of X is the persistence R-module

$$\left(\{H_n^{\mathrm{Dir}}(X_\delta, R)\}, \{H_n^{\mathrm{Dir}}(i_\delta^{\delta'})\}\right)_{\delta\leq\delta'\in T},$$

where $H_n^{\mathrm{Dir}}(X_\delta, R)$ are defined as in Definition 3.6, $i_\delta^{\delta'} : X_\delta \to X_{\delta'}$ are the inclusion maps, and $H_n^{\mathrm{Dir}}(i_\delta^{\delta'})$ are the restrictions of the maps $H_n(i_\delta^{\delta'})$ to directed homology, as introduced in Proposition 3.11.

Remark 4.4 Let $(X_\delta)_{\delta\in T}$ be a filtration of a directed simplicial complex X and let $\delta \leq \delta' \in T$. If $R \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, by applying Proposition 3.11 to the map $i_\delta^{\delta'} : X_\delta \to X_{\delta'}$, we have a commutative diagram

$$\begin{array}{ccc} H_n(X_\delta, R) & \xrightarrow{H_n(i_\delta^{\delta'})} & H_n(X_{\delta'}, R) \\ \uparrow & & \uparrow \\ H_n^{\mathrm{Dir}}(X_\delta, R) & \xrightarrow[H_n^{\mathrm{Dir}}(i_\delta^{\delta'})]{} & H_n^{\mathrm{Dir}}(X_{\delta'}, R). \end{array}$$

In particular, the injections $\left\{H_n^{\mathrm{Dir}}(X_\delta, R) \hookrightarrow H_n(X_\delta, R)\right\}_{\delta\in T}$ give rise to a monomorphism of persistence modules.

We can now introduce persistence diagrams and barcodes associated to filtrations of directed simplicial complexes. As these were only introduced for fields in Sect. 2.1, from this point on, we will assume that R is a field. In particular, when taking directed homology, we will assume that $R \in \{\mathbb{Q}, \mathbb{R}\}$.

Definition 4.5 Let $(X_\delta)_{\delta\in T}$ be a filtration of a directed simplicial complex X where $T \subseteq \mathbb{R}$ is finite. The *persistence diagrams* associated to the n-dimensional undirected and directed persistence R-modules of X are respectively denoted $\mathrm{Dgm}_n(X, R)$ and $\mathrm{Dgm}_n^{\mathrm{Dir}}(X, R)$. Similarly, the respective *barcodes* are denoted $\mathrm{Pers}_n(X, R)$ and $\mathrm{Pers}_n^{\mathrm{Dir}}(X, R)$.

Remark 4.6 Note that $H_0^{\mathrm{Dir}}(X_\delta, R) \cong H_0(X_\delta, R)$ for every $\delta \in T$, as shown in Proposition 3.18. As a consequence, the 0-dimensional directed and undirected persistence R-modules of a directed simplicial complex X are isomorphic. In particular, they have the same persistence barcodes and diagrams, and they measure the connectivity of the simplicial complex at each stage of the filtration.

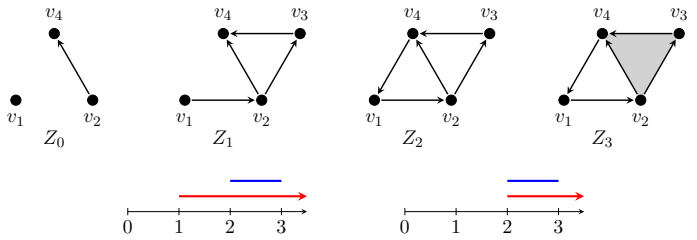


Fig. 5 A filtration of the directed simplicial complex Z in Example 3.23 (Fig. 3) and its associated undirected (bottom, left) and directed (bottom, right) 1-dimensional persistence barcodes. The shorter undirected bar (blue) is also directed, while the longer undirected bar (red) becomes directed at $\delta = 2$ (colour figure online)

The next result establishes the relation between the undirected and directed persistence barcodes and diagrams of a persistence module, and it is thus key to understanding the directed persistence barcodes.

Proposition 4.7 *Let $(X_\delta)_{\delta\in T}$ be a filtration of directed simplicial complexes. Then, there is an injective matching sending each bar in $\mathrm{Pers}_n^{\mathrm{Dir}}(X, R)$ to a bar in $\mathrm{Pers}_n(X, R)$ which contains it. Furthermore, matched bars must die at the same time.*

Proof This result is an immediate consequence of Remark 4.4 and Bauer and Lesnick (2014, Proposition 6.1). □

Note that a bar in the directed persistence barcode of a filtration may be born after the one it is matched with in the undirected one, and some bars in the undirected barcode may remain unmatched (see Examples below and Figs. 5 and 6).

We now introduce some examples to illustrate Proposition 4.7 and the general behaviour of the directed persistence barcodes.

Example 4.8 Let us illustrate how undirected and directed persistence modules and their barcodes can be different. First, consider the directed simplicial complexes X and Y in Fig. 3 (see Example 3.23). Regardless of the filtration chosen for X, the lack of directed cycles means that the 1-dimensional directed persistence module is trivial. However, at the end of the filtration, there is a cycle in homology, thus there is a bar in the undirected persistence barcode. In the case of Y, the only 1-cycle is directed, so the undirected and directed persistence modules associated to any filtration of Y are isomorphic, and different from those of X.

Example 4.9 Consider now the directed simplicial complex Z in Fig. 3 (see Example 3.23). Let $T = \{0, 1, 2, \dots\}$ and consider the filtration of Z given by $(Z_\delta)_{\delta\in T}$, as illustrated in Fig. 5, where $Z_\delta = Z$ for every $\delta \geq 3$.

The undirected and directed 1-dimensional persistence modules of this filtration are not isomorphic. Indeed, in Z_1 there is clearly an undirected cycle, whereas $Z_1^{\mathrm{Dir}}(Z_1, R)$ is trivial, thus $H_1^{\mathrm{Dir}}(Z_1, R) = \{0\}$. However, $H_1(Z_2, R) \cong H_1^{\mathrm{Dir}}(Z_2, R)$, as both vector spaces are generated by the classes of $[v_1, v_2] + [v_2, v_3] + [v_3, v_4] + [v_4, v_1]$ and $[v_2, v_4] + [v_4, v_1] + [v_1, v_2]$. These two classes become equivalent in Z_3. The undirected and directed 1-dimensional persistence barcodes of this filtration, shown in Fig. 5, illustrate how undirected classes may become directed.

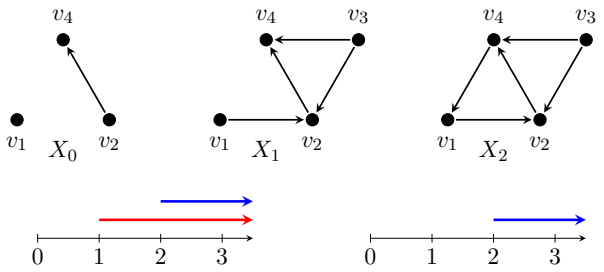


Fig. 6 A filtration of directed simplicial complexes and its associated undirected (bottom, left) and directed (bottom, right) 1-dimensional persistence barcodes. The shorter undirected bar (blue) is also directed, while the longer undirected bar (red) never becomes directed (colour figure online)

The last example also shows an important difference between classical and directed persistence modules and barcodes. Namely, in the undirected setting, the addition of one simplex to the filtration can either cause the birth of a class in the dimension of the added simplex, or kill a class in the preceding dimension. This simple idea is in fact the basis of the Standard Algorithm for computing persistent homology. However, the addition of only one simplex to a directed simplicial complex can cause the birth of several classes in directed homology, as shown in the previous example at $\delta = 2$, and also in the following example.

Example 4.10 Let $T = \{0, 1, 2, \dots\}$ and consider the filtration of simplicial complexes $(X_\delta)_{\delta \in T}$ illustrated in Fig. 2 in the Introduction, where $X_\delta = X_3$ for every $\delta \geq 3$.

It is clear that $Z_1^{\mathrm{Dir}}(X_j, R)$ is trivial for $j = 0, 1, 2$, whereas undirected cycles appear as early as X_1. However, by adding the edge from v_5 to v_1, several directed cycles are born at once. Namely, $Z_1^{\mathrm{Dir}}(X_3, R)$ contains the cycles

$$
\begin{aligned}
&[v_1, v_2] + [v_2, v_3] + [v_3, v_4] + [v_4, v_5] + [v_5, v_1],\\
&[v_1, v_2] + [v_2, v_4] + [v_4, v_5] + [v_5, v_1],\\
&[v_1, v_2] + [v_2, v_3] + [v_3, v_5] + [v_5, v_1],\\
&[v_1, v_3] + [v_3, v_4] + [v_4, v_5] + [v_5, v_1],\\
&[v_1, v_3] + [v_3, v_5] + [v_5, v_1].
\end{aligned}
$$

Straightforward computations show that $H_1(X_3, R) = R^4$, and four of the cycles above are linearly independent, thus $H_1^{\mathrm{Dir}}(X_3, R) = H_1(X_3, R) = R^4$. Consequently, at $\delta = 3$ every class (including the birthing one) becomes directed.

Our last example shows an undirected homology class that never becomes directed.

Example 4.11 Let $T = \{0, 1, 2, \dots\}$ and consider the filtration of simplicial complexes $(X_\delta)_{\delta \in T}$ illustrated in Fig. 6, where $X_\delta = X_2$ for $\delta \geq 2$.

Clearly, $Z_1^{\mathrm{Dir}}(X_j, R) = \{0\}$ for $j = 0, 1$, whereas $Z_1^{\mathrm{Dir}}(X_2, R)$ contains the cycle $[v_1, v_2] + [v_2, v_4] + [v_4, v_1]$. The undirected class represented by this element is thus directed, but there is a linearly independent class in undirected homology, $[v_3, v_2] + [v_2, v_4] - [v_3, v_4]$, which never becomes directed. Its bar in the barcode is thus unmatched.

4.2 Directed persistent homology of dissimilarity functions

In this section, we introduce the undirected and directed persistence diagrams and barcodes associated to dissimilarity functions (Definition 4.13) and prove their stability with respect to the bottleneck distance (Theorem 4.19).

Let us begin by introducing the *directed Rips filtration* of directed simplicial complexes associated to a dissimilarity function (Turner 2019, Definition 16).

Definition 4.12 Let (V, d_V) be a dissimilarity function. The *directed Rips filtration* of (V, d_V) is the filtration of directed simplicial complexes $\left(\mathcal{R}^{\mathrm{Dir}}(V, d_V)\right)_{\delta\in\mathbb{R}}$ where $(v_0, v_1, \ldots, v_n) \in \mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta$ if and only if $d_V(v_i, v_j) \leq \delta$, for all $0 \leq i \leq j \leq n$. It is clearly a filtration with the inclusion maps $i_\delta^{\delta'} : \mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta \to \mathcal{R}^{\mathrm{Dir}}(V, d_V)_{\delta'}$ for all $\delta \leq \delta'$.

Let us now introduce the persistent homology modules associated to such a filtration. Assume that $R \in \{\mathbb{Q}, \mathbb{R}\}$.

Definition 4.13 Let (V, d_V) be a dissimilarity function and consider its associated directed Rips filtration $\left(\mathcal{R}^{\mathrm{Dir}}(V, d_V)\right)_{\delta\in\mathbb{R}}$. For each $n \geq 0$, the *n-dimensional undirected persistence* R-module of (V, d_V) is

$$\mathcal{H}_n(V, d_V) := \left(\{H_n(\mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta, R)\}, \{H_n(i_\delta^{\delta'})\}\right)_{\delta\leq\delta'\in\mathbb{R}}.$$

Similarly, the *n-dimensional directed persistence* R-module of (V, d_V) is defined as

$$\mathcal{H}_n^{\mathrm{Dir}}(V, d_V) := \left(\{H_n^{\mathrm{Dir}}(\mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta, R)\}, \{H_n^{\mathrm{Dir}}(i_\delta^{\delta'})\}\right)_{\delta\leq\delta'\in\mathbb{R}}.$$

Remark 4.14 The persistence module $\mathcal{H}_n(V, d_V)$ associated to the directed Rips filtration of (V, d_V) is precisely the persistence module studied in Turner (2019, Section 5), hence the remarks made there hold for the undirected persistence module. In particular, if (V, d_V) is a (finite) metric space, $\mathcal{R}^{\mathrm{Dir}}(V, d_V)$ is the (classical) Vietoris-Rips filtration of (V, d_V). Furthermore, in this case, it can easily be seen that $\mathcal{H}_n(V, d_V) = \mathcal{H}_n^{\mathrm{Dir}}(V, d_V)$. Thus, these persistence modules generalize the persistence modules associated to the Vietoris-Rips filtration of a metric space. Also note that $\mathcal{R}^{\mathrm{Dir}}(V, d_V)$ is *closed under adjacent repeats* (Turner 2019, Definition 18). Namely, given $(v_0, v_1, \ldots, v_n) \in \mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta$, $(v_0, v_1, \ldots, v_i, v_i, \ldots, v_n) \in \mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta$, for any $i = 0, 1, \ldots, n$. Using this, it can be proven that the homology of this object must be trivial in dimensions larger than $|V| + 1$.

As R is a field and since V is finite, both of these persistence modules fulfil the assumptions in Sect. 2.1. Namely, their indexing sets can be chosen to be finite, corresponding to the threshold values where new simplices are added to the simplicial complex. Furthermore, no simplex is added to the filtration until the threshold value reaches the minimum of the images of the dissimilarity function. Finally, and even though the directed simplicial complex $\mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta$ may have infinite simplices due to arbitrary repetitions of vertices being allowed, it always has a finite number of simplices in a given dimension n, thus its n-dimensional homology is always finite-dimensional. As a consequence, we can introduce the following.

Definition 4.15 Let (V, d_V) be a dissimilarity function. For each $n \geq 0$, the n-dimensional *persistence diagrams* associated to the persistence R-modules $\mathcal{H}_n(V, d_V)$ and $\mathcal{H}_n^{\mathrm{Dir}}(V, d_V)$ are respectively denoted by $\mathrm{Dgm}_n(V, d_V)$ and $\mathrm{Dgm}_n^{\mathrm{Dir}}(V, d_V)$. Similarly, their associated *persistence barcodes* are denoted by $\mathrm{Pers}_n(V, d_V)$ and $\mathrm{Pers}_n^{\mathrm{Dir}}(V, d_V)$.

Of course, Proposition 4.7 holds for these barcodes, namely, every bar in $\mathrm{Pers}_n^{\mathrm{Dir}}(V, d_V)$ can be uniquely matched with one in $\mathrm{Pers}_n(V, d_V)$ which dies at the same time, although the directed bar may be born later.

We now use results from Sect. 2.2 to show that both these persistent homology constructions are stable. The proof is split in several lemmas. Let (V, d_V) and (W, d_W) be two dissimilarity functions on respective sets V and W and define $\eta = 2d_{\mathrm{CD}}\big((V, d_V), (W, d_W)\big)$, where d_{CD} is the correspondence distortion distance (Definition 2.9). By Proposition 2.11, we can find maps $\varphi\colon V \to W$ and $\psi\colon W \to V$ such that $\mathrm{dis}(\varphi), \mathrm{dis}(\psi), \mathrm{codis}(\varphi, \psi), \mathrm{codis}(\psi, \varphi) \leq \eta$. To simplify notation in the proofs below, denote $\mathcal{R}^{\mathrm{Dir}}(V, d_V)_\delta = X_V^\delta$ and $\mathcal{R}^{\mathrm{Dir}}(W, d_W)_\delta = X_W^\delta$, for all $\delta \in \mathbb{R}$.

Lemma 4.16 *For each $\delta \in \mathbb{R}$, the maps φ and ψ induce morphisms of directed simplicial complexes*

$$\begin{aligned} \varphi_\delta\colon X_V^\delta &\longrightarrow X_W^{\delta+\eta} & \qquad \psi_\delta\colon X_W^\delta &\longrightarrow X_V^{\delta+\eta} \\ x &\longmapsto \varphi(x), & x &\longmapsto \psi(x). \end{aligned}$$

Proof Let us prove the statement for φ_δ (the proof is analogous for ψ_δ). Let $(x_0, x_1, \ldots, x_n)$ be an n-simplex in X_V^δ. Then $d_V(x_i, x_j) \leq \delta$ for all $1 \leq i \leq j \leq n$. Since $\mathrm{dis}(\varphi) \leq \eta$, we have that, for all $v_1, v_2 \in V$,

$$\big|d_V(v_1, v_2) - d_W\big(\varphi(v_1), \varphi(v_2)\big)\big| \leq \eta.$$

Choosing $v_1 = x_i$ and $v_2 = x_j$, we have

$$d_W\big(\varphi(x_i), \varphi(x_j)\big) \leq \eta + d_V(x_i, x_j) \leq \delta + \eta \text{ for all } 1 \leq i \leq j \leq n.$$

Consequently, $\big(\varphi(x_0), \varphi(x_1), \ldots, \varphi(x_n)\big) \in X_W^{\delta+\eta}$ and the result follows. □

Lemma 4.17 *For $\delta \leq \delta' \in \mathbb{R}$ consider the inclusion maps $i_\delta^{\delta'}\colon X_V^\delta \hookrightarrow X_V^{\delta'}$ and $j_\delta^{\delta'}\colon X_W^\delta \hookrightarrow X_W^{\delta'}$. The following are commutative diagrams of morphisms of directed simplicial complexes:*

$$\begin{array}{ccc} X_V^\delta & \xrightarrow{i_\delta^{\delta'}} & X_V^{\delta'} \\ \downarrow{\scriptstyle \varphi_\delta} & & \downarrow{\scriptstyle \varphi_{\delta'}} \\ X_W^{\delta+\eta} & \xrightarrow{j_{\delta+\eta}^{\delta'+\eta}} & X_W^{\delta'+\eta}, \end{array} \qquad \begin{array}{ccc} X_W^\delta & \xrightarrow{j_\delta^{\delta'}} & X_W^{\delta'} \\ \downarrow{\scriptstyle \psi_\delta} & & \downarrow{\scriptstyle \psi_{\delta'}} \\ X_V^{\delta+\eta} & \xrightarrow{i_{\delta+\eta}^{\delta'+\eta}} & X_V^{\delta'+\eta}. \end{array}$$

Proof We prove that the first diagram is commutative (the proof for the second diagram is analogous). Let $x \in V$. Since $i_\delta^{\delta'}$ is an inclusion, $(\varphi_{\delta'} \circ i_\delta^{\delta'})(x) = \varphi_{\delta'}(x) = \varphi(x)$. Similarly, since $j_{\delta+\eta}^{\delta'+\eta}$ is an inclusion, $(j_{\delta+\eta}^{\delta'+\eta} \circ \varphi_\delta)(x) = j_{\delta+\eta}^{\delta'+\eta}\big(\varphi(x)\big) = \varphi(x)$. □

Lemma 4.18 *With the same notation as in Lemmas 4.16 and 4.17, for every $\delta \in \mathbb{R}$, the following diagrams of morphisms of directed simplicial complexes induce commutative diagrams on homology.*

$$\begin{array}{ccc} X_V^\delta & \xrightarrow{i_\delta^{\delta+2\eta}} & X_V^{\delta+2\eta} \\ & {\scriptstyle\varphi_\delta} \searrow \quad \nearrow {\scriptstyle\psi_{\delta+\eta}} & \\ & X_W^{\delta+\eta}, & \end{array} \qquad \begin{array}{ccc} X_W^\delta & \xrightarrow{j_\delta^{\delta+2\eta}} & X_V^{\delta+2\eta} \\ & {\scriptstyle\psi_\delta} \searrow \quad \nearrow {\scriptstyle\varphi_{\delta+\eta}} & \\ & X_W^{\delta+\eta}. & \end{array}$$

Proof Again, we only prove the result for the first diagram, as the proof for the second diagram is analogous. We show that it is commutative up to homotopy by showing that the maps $i_\delta^{\delta+2\eta}$ and $\psi_{\delta+\eta} \circ \varphi_\delta$ satisfy the hypothesis of Lemma 3.14.

Take a simplex $\sigma = (x_0, x_1, \ldots, x_n) \in X_V^\delta$, thus $d_V(x_i, x_j) \leq \delta$, for all $1 \leq i \leq j \leq n$. On the one hand, $i_\delta^{\delta+2\eta}$ is an inclusion, so $i_\delta^{\delta+2\eta}(\sigma) = \sigma$. On the other hand, since $\psi_{\delta+\eta} \circ \varphi_\delta$ is a morphism of directed simplicial complexes (Definition 3.8), $\big(\psi(\varphi(x_0)), \psi(\varphi(x_1)), \ldots, \psi(\varphi(x_n))\big) \in X_V^{\delta+2\eta}$. This implies that $d_V\big(\psi(\varphi(x_i)), \psi(\varphi(x_j))\big) \leq \delta + 2\eta$, for all $1 \leq i \leq j \leq n$.

Now recall that $\operatorname{codis}(\varphi, \psi) \leq \eta$, thus for all $v \in V$ and $w \in W$,

$$\big|d_V\big(v, \psi(w)\big) - d_W\big(\varphi(v), w\big)\big| \leq \eta.$$

Then, for $1 \leq i \leq j \leq n$, by taking $v = x_i$ and $w = \varphi(x_j)$,

$$d_V\big(x_i, \psi(\varphi(x_j))\big) \leq \eta + d_W\big(\varphi(x_i), \varphi(x_j)\big) \leq \delta + 2\eta.$$

As a consequence of the inequalities above, we have shown that for every $0 \leq i \leq n$,

$$\big(x_0, x_1, \ldots, x_i, \psi(\varphi(x_i)), \psi(\varphi(x_{i+1})), \ldots, \psi(\varphi(x_n))\big) \in X_V^{\delta+2\eta}.$$

Therefore, the maps $i_\delta^{\delta+2\eta}$ and $\psi_{\delta+\eta} \circ \varphi_\delta$ satisfy the hypothesis of Lemma 3.14, thus they induce the same map on homology. The result follows. □

We now have everything we need to prove the stability results.

Theorem 4.19 *Let $R \in \{\mathbb{Q}, \mathbb{R}\}$. Let (V, d_V) and (W, d_W) be two dissimilarity functions on finite sets V and W. Then, for all $n \geq 0$,*

$$d_B\big(\operatorname{Dgm}_n(V, d_V), \operatorname{Dgm}_n(W, d_V)\big) \leq 2d_{\mathrm{CD}}\big((V, d_V), (W, d_W)\big)$$

and

$$d_B\big(\mathrm{Dgm}_n^{\mathrm{Dir}}(V,d_V),\mathrm{Dgm}_n^{\mathrm{Dir}}(W,d_V)\big)\leq 2d_{\mathrm{CD}}\big((V,d_V),(W,d_W)\big).$$

Proof Define $\eta = 2d_{\mathrm{CD}}\big((V,d_V),(W,d_W)\big)$. By Theorem 2.7, it suffices to show that the persistence modules $\mathcal{H}_n(V,d_V)$ and $\mathcal{H}_n(W,d_W)$ (respectively $\mathcal{H}_n^{\mathrm{Dir}}(V,d_V)$ and $\mathcal{H}_n^{\mathrm{Dir}}(W,d_W)$) are η-interleaved. Comparing Definition 2.5 (for $\varepsilon=\eta$) and Lemmas 4.16, 4.17 and 4.18, the result follows by using the functoriality of homology and, in the directed case, Proposition 3.11. □

Remark 4.20 Recall from Remark 4.14 that the persistence module $\mathcal{H}_n(V,d_V)$ associated to the directed Rips filtration of (V,d_V) is the persistence module studied in Turner (2019, Section 5). Thus, the remarks made in Turner (2019, Section 5.2) hold for these persistence modules, meaning that a result analogous to Theorem 4.19 would not hold if we were using ordered-set complexes instead of directed simplicial complexes, as mentioned at the beginning of Sect. 3. This justifies our definition of directed simplicial complex (Definition 3.1).

5 Computational challenges

Let (V,d_V) be a dissimilarity function in a finite set V. In this section, we study the algorithmic aspects of computing both $\mathcal{H}_n(V,d_V)$ and $\mathcal{H}_n^{\mathrm{Dir}}(V,d_V)$. For simplicity, we assume that $R=\mathbb{R}$.

The computation of the persistence barcodes for the undirected persistent homology $\mathcal{H}_n(V,d_V)$ can be made using the Standard Algorithm (Edelsbrunner et al. 2001; Zomorodian and Carlsson 2005), which also applies to ordered simplicial complexes. Write $\mathcal{R}^{\mathrm{Dir}}(V,d_V)_\delta = X_V^\delta$, $\delta\in\mathbb{R}$ for the directed Rips filtration (Definition 4.12). As V is finite, the number of simplices in X_V^δ up to dimension $n+1$ is finite, say N. We can list them $\{\sigma_1,\sigma_2,\ldots,\sigma_N\}$ in such a way that $i<j$ if σ_i is a (proper) face of σ_j, or if σ_j appears 'later' in the filtration, thus having a *compatible ordering*. Then, we can represent the differential using a sparse $N\times N$ matrix M over $\mathbb{R}$, where the (i,j)-entry M_{ij} is the coefficient of σ_j in the differential of σ_i. Such matrix M is an upper-triangular matrix. At this stage, the Standard Algorithm can be used to compute $\mathrm{Dgm}_k(V,d_V)$ by reducing M at once using column operations.

We cannot directly adapt the Standard Algorithm to compute $\mathrm{Dgm}_k^{\mathrm{Dir}}(V,d_V)$: we can try to find reduction operations that generate directed cycles using operations with positive coefficients only, but this is not enough to guarantee that we find all the possible directed cycles. For instance, once one column has been eliminated, we may be preventing the participation of the simplex it represents in directed cycles not involving the columns we are using in the elimination. To avoid such a problem, we would need to keep track of every possible way in which column eliminations can be produced using only positive coefficients. This is very similar to the *double description algorithm* for the computation of extreme rays of a polyhedral cone (Fukuda et al. 1995; Motzkin et al. 1953). Indeed, the directed cycles $Z_n^{\mathrm{Dir}}(X,\mathbb{R})$ are the non-negative solutions to

the linear equation $Mx = 0$, that is,

$$\{x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N \mid Mx = 0, x \geq 0\}, \tag{5.1}$$

where a vector $x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$ represents the cycle $\sum_{i=1}^{N} x_i \sigma_i$, M is the matrix corresponding to the differential, and $x \geq 0$ means that every entry in x is non-negative, that is, $x_i \geq 0$ for all i. Equivalently, directed cycles are the extreme rays of the unbounded polyhedral cone consisting of the intersection of the solution space to the linear system $Mx = 0$ with the first quadrant of $\mathbb{R}^N$. The enumeration of the extreme rays (generators of the solution set) of a polyhedral cone such as Eq. (5.1) is a well-known albeit hard problem in computational geometry. Several algorithms, including the double description algorithm, and modifications thereof, exist (Terzer 2009).

Note, however, that we do not require the computation of all the directed cycles. Indeed, it would be enough to calculate them up to homology, which suggests that efficient computations may be possible.

In summary, although important challenges remain on the computation of the persistence diagrams associated to directed persistence modules, our objective was to provide the necessary groundwork for the study of directed persistent homology of asymmetric data sets, and we are hopeful that we are opening up an exciting new approach for future research into the topological properties of asymmetric data sets.

Acknowledgements This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1 The first author was partially supported by Ministerio de Economía y Competitividad (Spain) Grants Nos. MTM2016-79661-P and MTM2016-78647-P and by Ministerio de Ciencia, Innovación y Universidades (Spain) Grant No. PGC2018-095448-B-I00.

Funding Open access funding provided by FCT|FCCN (b-on).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: Directed persistent homology via semiring homology

As mentioned in the Introduction, our ideas of directed homology are based on homology with semiring coefficients, which can be introduced for directed simplicial complexes using chain complexes of semimodules and their associated homologies, building on work by Patchkoria (1977, 2014). In this appendix, we introduce this homology theory and explain how it relates to the modules of directed homology introduced in Definition 3.6.

We remark that other perspectives of directionality in topological spaces could be used instead. One such important perspective is that of directed algebraic topology,

which is extensively used in the study of concurrency in computer science (Fajstrup et al. 2016). In this approach, directionality information is encoded by means of distinguished paths, forming a category dTop where objects are topological spaces together with their directed paths (Fajstrup et al. 2016, Definition 4.1). These spaces can be modelled using cubical complexes, and their homology including directional information can be computed using chain complexes over preordered abelian groups (Grandis 2009, Chapter 2).

In this theory, the homology groups of a space are those of ordinary homology, but directionality information is recorded in the preorder given to those groups. Therefore, it would be interesting for future work to study potential definitions of persistent homology over preordered abelian groups. Nonetheless, we focus here on homology over semimodules, since it provides us with the information we require while keeping the exposition simpler.

A.1 Semirings, semimodules and their completions

Let us begin by introducing the algebraic structures that we need, following (Golan 1999).

Definition A.1 A *semiring* $\Lambda = (\Lambda, +, \cdot)$ is a set Λ together with two operations such that

- $(\Lambda, +)$ is an abelian monoid whose identity element we denote 0_Λ,
- $(\Lambda, \cdot)$ is a monoid whose identity element we denote 1_Λ,
- $\cdot$ is distributive with respect to $+$ from either side,
- $0_\Lambda \cdot \lambda = \lambda \cdot 0_\Lambda = 0_\Lambda$, for all $\lambda \in \Lambda$.

A semiring Λ is *commutative* if $(\Lambda, \cdot)$ is a commutative monoid, and *cancellative* if $(\Lambda, +)$ is a cancellative monoid, that is, $\lambda + \lambda' = \lambda + \lambda''$ implies $\lambda' = \lambda''$ for all $\lambda, \lambda', \lambda'' \in \Lambda$. A semiring Λ is a *semifield* if every $0_\Lambda \neq \lambda \in \Lambda$ has a multiplicative inverse. A semiring is *zerosumfree* if no element other than 0_Λ has an additive inverse.

Example A.2 Every ring is a semiring. The non-negative integers, rationals and reals with their usual operations, respectively denoted $\mathbb{N}$, $\mathbb{Q}^+$ and $\mathbb{R}^+$, are cancellative zerosumfree commutative semirings which are not rings.

Definition A.3 Let Λ be a semiring. A *(left)* Λ*-semimodule* is an abelian monoid $(A, +)$ with identity element 0_A together with a map $\Lambda \times A \to A$ which we denote $(\lambda, a) \mapsto \lambda a$ and such that for all $\lambda, \lambda' \in \Lambda$ and $a, a' \in A$,

- $(\lambda\lambda')a = \lambda(\lambda' a)$,
- $\lambda(a + a') = \lambda a + \lambda a'$,
- $(\lambda + \lambda')a = \lambda a + \lambda' a$,
- $1_\Lambda a = a$,
- $\lambda 0_A = 0_A = 0_\Lambda \lambda$.

A non-empty subset B of a left Λ-semimodule A is a *subsemimodule* of A if B is closed under addition and scalar multiplication, which implies that B is a left Λ-semimodule with identity element $0_A \in B$. If A and B are Λ-semimodules, a Λ-homomorphism is a map $f : A \to B$ such that for all $a, a' \in A$ and for all $\lambda \in \Lambda$,

- $f(a + a') = f(a) + f(a')$,
- $f(\lambda a) = \lambda f(a)$.

Clearly, $f(0_A) = 0_B$.

A Λ-semimodule A is *cancellative* if $a + a' = a + a''$ implies $a' = a''$, and *zerosumfree* if $a + a' = 0$ implies $a = a' = 0$, for all $a, a', a'' \in A$.

The Cartesian product of Λ-semimodules is a Λ-semimodule. The definition of direct sum of modules can be generalized to the framework of Λ-semimodules, and a finite direct sum of Λ-semimodules is isomorphic to the Cartesian product of the same Λ-semimodules. Quotient Λ-semimodules can be defined using congruence relations.

Definition A.4 Let A be a left Λ-semimodule. An equivalence relation ρ on A is a Λ-*congruence* if, for all $a, a' \in \Lambda$, and all $\lambda \in \Lambda$,

- if $a \sim_\rho a'$ and $b \sim_\rho b'$, then $(a + b) \sim_\rho (a' + b')$, and
- if $a \sim_\rho a'$, then $\lambda a \sim_\rho \lambda a'$.

If ρ is a Λ-congruence relation on A and we write a/ρ for the class of an element $a \in A$, then $A/\rho = \{a/\rho \mid a \in A\}$ inherits a Λ-semimodule structure by setting $(a/\rho) + (a'/\rho) = (a + a')/\rho$ and $\lambda(a/\rho) = (\lambda a)/\rho$, for all $a, a' \in A$ and $\lambda \in \Lambda$. The left Λ-semimodule A/ρ is called the *factor semimodule* of A by ρ. Note that the quotient map $A \to A/\rho$ is a surjective Λ-homomorphism.

If B is a subsemimodule of a Λ-semimodule A, then it determines a Λ-congruence $\sim_B$ by setting $a \sim_B a'$ if there exist $b, b' \in B$ such that $a + b = a' + b'$. Classes in this quotient are denoted a/B, and the factor semimodule is denoted A/B. If Λ is a ring, this is just the usual quotient module of A by B.

Example A.5 The following are examples of factor semimodules.

- If Λ is a ring and $B \leq A$ are Λ-modules, the factor semimodule A/B is the usual module quotient of A by B.
- In the $\mathbb{N}$-module $\mathbb{N} \times \mathbb{N}$, consider the relation ρ such that $(u, v) \sim_\rho (x, y)$ if $u + y = v + x$. It is immediate to show that ρ is a $\mathbb{N}$-congruence, and the factor semimodule $(\mathbb{N} \times \mathbb{N})/\rho$ is isomorphic to $\mathbb{Z}$. Indeed, consider the map $(\mathbb{N} \times \mathbb{N})/\rho \to \mathbb{Z}$ taking $(u, v) \mapsto u - v$. It is clearly a map of $\mathbb{N}$-semimodules, $(u, v) \sim_\rho (x, y) \Leftrightarrow u + y = v + x \Leftrightarrow u - v = x - y$ shows that it is well-defined and injective, and it is clearly surjective.
- Also in $\mathbb{N} \times \mathbb{N}$, consider the submodule $B = \{0\} \times \mathbb{N}$. Then $(u, v) \sim_B (x, y)$ if there exist $(0, n), (0, m) \in \{0\} \times \mathbb{N}$ such that $(u, v) + (0, n) = (x, y) + (0, m)$. From this, we deduce that $(u, v) \sim_B (x, y)$ if and only if $u = x$. Namely, $(\mathbb{N} \times \mathbb{N})/(\{0\} \times \mathbb{N})$ is isomorphic to $\mathbb{N}$.

Clearly, if Λ is a semiring then it is a Λ-semimodule over itself. The idea of free Λ-semimodules comes about naturally just as in the case of rings, and we will make extensive use of these objects.

Definition A.6 Let Λ be a semiring, A a left Λ-semimodule and $V = \{v_1, v_2, \ldots, v_n\}$ a finite subset of A. The set V is a *generating set* for A if every element in A is a linear combination of elements of V. The *rank* of a Λ-semimodule A, denoted $\mathrm{rank}(A)$, is

the least n for which there is a set of generators of A with cardinality n, or infinity, if not such n exists.

The set V is *linearly independent* if for $\lambda_1, \lambda_2, \ldots, \lambda_n \in \Lambda$ and $\mu_1, \mu_2, \ldots, \mu_n \in \Lambda$ such that $\sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \mu_i v_i$, then $\lambda_i = \mu_i$, for $i = 1, 2, \ldots, n$, and it is called *linearly dependent* otherwise. We call V is a *basis* of A if it is a linearly independent generating set of A.

The Λ-semimodule A is a *free* Λ*-semimodule* if it admits a basis V. We denote a free Λ-semimodule with basis $\{v_1, v_2, \ldots, v_n\}$ by $\Lambda(v_1, v_2, \ldots, v_n)$.

Remark A.7 It is easy to check that free Λ-semimodules over a cancellative semiring Λ are themselves cancellative, as are subsemimodules of cancellative Λ-semimodules. The quotient of a cancellative Λ-semimodule over any of its subsemimodules is also cancellative, see Golan (1999, Proposition 15.24).

Note that if Λ is a ring, free Λ-semimodules are just free Λ-modules. Thus, as not every module over a ring is free, clearly not every Λ-semimodule is free. On the other hand, Λ^n (the direct sum, or product, of n copies of Λ) is clearly a free Λ-semimodule, for every $n \geq 1$. Another key property is that if A is a free Λ-semimodule with basis V and B is another Λ-semimodule, each map $V \to B$ uniquely extends to a Λ-homomorphism $A \to B$, see Golan (1999, Proposition 17.12).

Remark A.8 If Λ is a commutative semiring and A is a Λ-semimodule admitting a finite basis, every basis has the same cardinality, which coincides with the rank of A (Tan 2014, Theorem 3.4). In fact, for every integer $n > 0$, every finitely generated subsemimodule of $\mathbb{N}^n$ has a unique basis, whereas basis for semimodules of $(\mathbb{R}^+)^n$ are unique up to non-zero multiples, see e.g. Kim and Roush (1980, Theorem 2.1).

We finish this section by extending the Grothendieck construction from abelian monoids to semirings and semimodules.

Definition A.9 Let M be an abelian monoid. Consider the congruence relation ρ in $M \times M$ defined as

$$(u, v) \sim_\rho (x, y) \Leftrightarrow \text{there exists } z \in M \text{ such that } u + y + z = v + x + z.$$

Let $[u, v]$ denote the equivalence class of (u, v). Then $M \times M/\rho$ becomes a group with the componentwise addition. This group is called the *Grothendieck group* or *group completion* of M. There is a canonical homomorphism of monoids $k_M \colon M \to K(M)$ defined as $k_M(x) = [x, 0]$. If M is a cancellative monoid, then k_M is injective.

Given an element $[u, v] \in K(M)$ we can interpret u as its *positive part* and v as its *negative part*, and the relation $\sim$ becomes the obvious one. The identity element is then $0 = [x, x]$ for any element $x \in M$, and the inverse of $[x, y]$ is $[y, x]$ for any $x, y \in M$.

If $f \colon M \to N$ is a morphism of monoids, there is a morphism of groups $K(f) \colon K(M) \to K(N)$ that takes $[u, v]$ to $[f(u), f(v)]$. In fact, K is a functor from the category of abelian monoids to the category of abelian groups. Furthermore, $K(f) \circ k_M = k_N \circ f$.

As shown in Example A.5, $K(\mathbb{N}) \cong \mathbb{Z}$. Similarly, one can show that $K(\mathbb{R}^+) \cong \mathbb{R}$. Crucially for us, this construction can be extended to semirings and semimodules.

Definition A.10 Let Λ be a semiring. The group completion $K(\Lambda)$ becomes a ring with the operation $[x_1, x_2] \cdot [y_1, y_2] = [x_1y_1 + x_2y_2, x_2y_1 + x_1y_2]$, and $k_\Lambda \colon \Lambda \to K(\Lambda)$ is in fact a morphism of semirings, which is injective if Λ is cancellative. We call $K(\Lambda)$ the *ring completion* of Λ. If $f \colon \Lambda \to \Lambda'$ is a morphism of semirings, the map $K(f) \colon K(\Lambda) \to K(\Lambda')$ as introduced in Definition A.9 is a morphism of rings, and K is a functor from the category of semirings to the category of rings.

If A is a Λ-semimodule, then the abelian group $K(A)$ together with the operation $K(\Lambda) \times K(A) \to K(A)$ given by $[\lambda_1, \lambda_2][a_1, a_2] = [\lambda_1 a_1 + \lambda_2 a_2, \lambda_1 a_2 + \lambda_2 a_1]$ is a $K(\Lambda)$-module, the $K(\Lambda)$-module completion of A. Again, if $f \colon A \to B$ is a Λ-morphism, the map $K(f) \colon K(A) \to K(B)$ becomes a morphism of $K(\Lambda)$-modules. Furthermore, if A is a free Λ-semimodule with basis $\{v_i \mid i \in I\}$, it becomes immediate that $K(A)$ is a free $K(\Lambda)$-module with basis $\{[v_i, 0] \mid i \in I\}$.

A.2 Chain complexes of semimodules

Now that we have introduced the necessary algebraic structures, we move on to chain complexes of semimodules over semirings, following Patchkoria (2014). This theory is a natural generalization of the classical theory of chain complexes of modules and, in fact, they give rise to the same cycles, boundaries and homologies when the semimodules are modules over a ring.

In order to introduce chain complexes in the context of semimodules an immediate problem arises: alternating sums cannot be defined as elements in a semimodule may not have inverses. The solution is to use two maps, a *positive* and *negative* part, for the differentials.

Definition A.11 Let Λ be a semiring and consider a sequence of Λ-semimodules and homomorphisms indexed by $n \in \mathbb{Z}$

$$X \colon \cdots \rightrightarrows X_{n+1} \underset{\partial^-_{n+1}}{\overset{\partial^+_{n+1}}{\rightrightarrows}} X_n \underset{\partial^-_n}{\overset{\partial^+_n}{\rightrightarrows}} X_{n-1} \rightrightarrows \cdots .$$

We say that $X = \{X_n, \partial^+_n, \partial^-_n\}$ is a *chain complex* of Λ-semimodules if

$$\partial^+_n \partial^+_{n+1} + \partial^-_n \partial^-_{n+1} = \partial^+_n \partial^-_{n+1} + \partial^-_n \partial^+_{n+1}.$$

As in the classical case, chain complexes of Λ-semimodules give rise to a Λ-semimodule of homology.

Definition A.12 Let $X = \{X_n, \partial^+_n, \partial^-_n\}$ be a chain complex of Λ-semimodules. The Λ-semimodule of *cycles* of X is

$$Z_n(X, \Lambda) = \{x \in X_n \mid \partial^+_n(x) = \partial^-_n(x)\}.$$

The nth *homology* of X is then the quotient Λ-semimodule

$$H_n(X, \Lambda) = Z_n(X, \Lambda)/\rho_n(X, \Lambda)$$

where $\rho_n(X, \Lambda)$ is the following Λ-congruence relation on $Z_n(X, \Lambda)$:

$$\begin{aligned} x \sim_{\rho_n(X,\Lambda)} y \Leftrightarrow \exists u, v \in X_{n+1} \text{ s.t. } x + \partial^+_{n+1}(u) + \partial^-_{n+1}(v) \\ = y + \partial^+_{n+1}(v) + \partial^-_{n+1}(u). \end{aligned}$$

We will omit from now on the coefficient semiring Λ from the notation when it is clear from the context.

Remark A.13 The definition of cycle is a direct generalization of the classical definition. For the boundary relation, note that we may need two different chains u and v in order to establish two classes as homologous. Intuitively, these two classes are the 'positive' and 'negative' part of w, where $x = y + \partial(w)$ in the classical setting.

Remark A.14 If $X = \{X_n, \partial^+_n, \partial^-_n\}$ is a chain complex of Λ-semimodules, then

$$\cdots \longrightarrow K(X_{n+1}) \xrightarrow{K(\partial^+_{n+1}) - K(\partial^-_{n+1})} K(X_n) \xrightarrow{K(\partial^+_n) - K(\partial^-_n)} K(X_{n-1}) \longrightarrow \cdots$$

is a chain complex of $K(\Lambda)$-modules. If furthermore the Λ-semimodules X_n are cancellative for all n, the converse is also true.

If Λ is a ring, then $K(\Lambda) = \Lambda$ and the functor K acts as the identity on Λ-modules. Therefore, $X = \{X_n, \partial^+_n, \partial^-_n\}$ is a chain complex of Λ-semimodules if and only if

$$\cdots \longrightarrow X_{n+1} \xrightarrow{\partial^+_{n+1} - \partial^-_{n+1}} X_n \xrightarrow{\partial^+_n - \partial^-_n} X_{n-1} \longrightarrow \cdots$$

is a chain complex of Λ-modules. In this case, it is clear that the homology semimodules introduced in Definition A.12 are precisely the usual homology modules of X.

We remark that a similar idea of decomposition of differentials in a positive and negative components has been used by Steiner to relate ω-categories with the so-called *augmented directed complexes* (Steiner 2004). These structures share, in fact, a deep similarity with Patchkoria's chain complexes with semiring coefficients.

We now turn our attention to maps between complexes.

Definition A.15 Let $X = \{X_n, \partial^+_n, \partial^-_n\}$ and $X' = \{X'_n, \partial^+_n, \partial^-_n\}$ be two chain complexes of Λ-semimodules. A sequence $f = \{f_n\}$ of Λ-homomorphisms $f_n \colon X_n \to X'_n$ is said to be a *morphism* from X to X' if

$$\partial^+_n f_n + f_{n-1}\partial^-_n = \partial^-_n f_n + f_{n-1}\partial^+_n.$$

It is clear that for such a map $f_n\big(Z_n(X)\big)$ is a Λ-subsemimodule of $Z_n(X')$. Furthermore, if X_n and X'_n are cancellative Λ-semimodules, f is also compatible with

the congruence relations $\rho_n(X)$ and $\rho_n(X')$, so it induces a map

$$H_n(f)\colon H_n(X) \to H_n(X').$$

It is then easy to check that homology H_* is a functor from the category of chain complexes of cancellative Λ-semimodules and morphisms to the category of graded Λ-semimodules.

Remark A.16 Let Λ be a semiring and let $\{X_n, \partial_n^+, \partial_n^-\}$ be a chain complex of Λ-semimodules. The family of canonical maps $k_{X_n}\colon X_n \to K(X_n)$ gives rise to a morphism from $\{X_n, \partial_n^+, \partial_n^-\}$ to $\{K(X_n), K(\partial_n^+), K(\partial_n^-)\}$ and, therefore, to a morphism of Λ-semimodules

$$H(k_X)\colon H_n(X) \to H_n\big(K(X)\big),$$

which takes the class of x to the class of $[x, 0]$. Furthermore, if the Λ-semimodules X_n are cancellative, $H_n(k_X)$ is injective, which in particular implies that $H_n(X)$ is a cancellative Λ-semimodule. Also note that, by Remark A.14, the homology of $\{K(X_n), K(\partial_n^+), K(\partial_n^-)\}$ and the usual homology of $\{K(X_n), K(\partial_n^+) - K(\partial_n^-)\}$ are isomorphic as $K(\Lambda)$-modules.

Finally, we discuss chain homotopies.

Definition A.17 Let $f = \{f_n\}$ and $g = \{g_n\}$ be morphisms from $X = \{X_n, \partial_n^+, \partial_n^-\}$ to $X' = \{X'_n, \partial_n^+, \partial_n^-\}$. We say that f is *homotopic* to g if there exist Λ-homomorphisms $s_n^+, s_n^-\colon X_n \to X'_{n+1}$ such that

$$\begin{aligned}&\partial_{n+1}^+ s_n^- + \partial_{n+1}^- s_n^+ + s_{n-1}^- \partial_n^+ + s_{n-1}^+ \partial_n^- + g_n\\ &\quad = \partial_{n+1}^+ s_n^+ + \partial_{n+1}^- s_n^- + s_{n-1}^+ \partial_n^+ + s_{n-1}^- \partial_n^- + f_n,\end{aligned}$$

for all n. The family $\{s_n^+, s_n^-\}$ is called a *chain homotopy* from f to g, and we write $(s^+, s^-)\colon f \simeq g$.

We then have the following result.

Proposition A.18 (Patchkoria 2014, Proposition 3.3) *Let $f, g\colon X \to X'$ be morphisms between chain complexes of cancellative Λ-semimodules. If f is homotopic to g, then $H_n(f) = H_n(g)$.*

Homotopy equivalences are defined in the usual way and they induce isomorphisms on homology. We finish with a remark that the homotopy of maps behaves well with respect to semimodule completion.

Remark A.19 If a morphism $f\colon X \to X'$ is homotopic to $g\colon X \to X'$, then $K(f)\colon K(X) \to K(X')$ is homotopic to $K(g)\colon K(X) \to K(X')$. Furthermore, if both X_n and X'_n are cancellative Λ-semimodules, for all n, and X_n is a free Λ-semimodule for all n, the converse is also true.

A.3 Chain complexes of semimodules of directed simplicial complexes

In this section, we introduce chain complexes of semimodules and homology semimodules of directed simplicial complexes, Definition 3.1. The results introduced here directly generalize those in Sect. 3.

Definition A.20 The *n-dimensional chains* of a directed simplicial complex X are defined as the elements of the free Λ-semimodule generated by (i.e. with basis) the n-simplices,

$$C_n(X, \Lambda) = \Lambda\big(\{[x_0, x_1, \ldots, x_n] \mid (x_0, x_1, \ldots, x_n) \in X\}\big).$$

We call the elements $[x_0, x_1, \ldots, x_n] \in C_n(X, \Lambda)$, $(x_0, x_1, \ldots, x_n) \in X$, *elementary n-chains*.

Remark A.21 Note that if Λ is a cancellative semiring, the Λ-semimodule $C_n(X, \Lambda)$ is cancellative for every n, as it is a free semimodule over a cancellative semiring.

We now define the positive and negative differentials on $C_n(X, \Lambda)$.

Definition A.22 Let X be a directed simplicial complex. For each $n > 0$, we define morphisms of Λ-semimodules $\partial_n^+, \partial_n^- : C_n(X, \Lambda) \to C_{n-1}(X, \Lambda)$ by

$$\partial_n^+([x_0, x_1, \ldots, x_n]) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} [x_0, x_1, \ldots, \widehat{x_{2i}}, \ldots, x_n], \text{ and}$$

$$\partial_n^-([x_0, x_1, \ldots, x_n]) = \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} [x_0, x_1, \ldots, \widehat{x_{2i+1}}, \ldots, x_n].$$

For $n = 0$, let $\partial_0^+, \partial_0^- : C_0(X, \Lambda) \to \{0\}$ be the trivial maps, by definition.

Proposition A.23 *Let X be a directed simplicial complex and Λ be a cancellative semiring. Then $\{C_n(X, \Lambda), \partial_n^+, \partial_n^-\}$ is a chain complex of modules of Λ-semimodules.*

Proof By Remark A.21, the Λ-semimodule $C_n(X, \Lambda)$ is cancellative for every n. Thus, by Remark A.14, it is enough to prove that $\big\{K\big(C_n(X, \Lambda)\big), K(\partial_n^+) - K(\partial_n^-)\big\}$ is a chain complex of $K(\Lambda)$-modules. Note that $K\big(C_n(X, \Lambda)\big)$ is a free $K(\Lambda)$-module whose basis is given by elements $\big[[x_0, x_1, \ldots, x_n], 0\big]$ such that $[x_0, x_1, \ldots, x_n]$ is an elementary n-chain. Thus, it suffices to show that the composition $\big(K(\partial_n^+) - K(\partial_n^-)\big) \circ \big(K(\partial_{n-1}^+) - K(\partial_{n-1}^-)\big)$ is trivial (zero) on these elements. Since

$$\big(K(\partial_n^+) - K(\partial_n^-)\big)\big[[x_0, x_1, \ldots, \widehat{x_i}, \ldots, x_n], 0\big] = \sum_{i=0}^{n} (-1)^i \big[[x_0, x_1, \ldots, x_n], 0\big],$$

the proof is now a straightforward computation analogous to the standard one for chain complexes in simplicial homology. □

Proposition A.23 allows us to define the homology of a directed simplicial complex with coefficients on a semiring.

Definition A.24 Let X be a directed simplicial complex and $n \geq 0$ and Λ a cancellative semiring. The n-dimensional *homology* of X, written $H_n(X, \Lambda)$, is the nth homology Λ-semimodule of the chain complex $\{C_n(X, \Lambda), \partial_n^+, \partial_n^-\}$.

Remark A.25 If Λ is a ring, by defining $\partial = \partial^+ - \partial^+$, $\{C_n(X, \Lambda), \partial\}$ is a chain complex in the usual sense. In fact, it is precisely the chain complex introduced in Definition 3.3. Thus, as a consequence of Remark A.14, the homology theory introduced in Definition A.24 coincides with the one introduced in Definition 3.4. Therefore, homology with semiring coefficients of directed simplicial complexes is a non-trivial generalization of homology with ring coefficients.

The following result will become useful later on, so we recorded it here.

Proposition A.26 *Let Λ be a cancellative semiring and X a directed simplicial complex. The chain complexes of $K(\Lambda)$-semimodules $\left\{K\left(C_n(X, \Lambda)\right), K(\partial_n^+), K(\partial_n^-)\right\}$ and $\left\{C_n\left(X, K(\Lambda)\right), \partial_n^+, \partial_n^-\right\}$ are isomorphic.*

Proof Take $n \geq 0$ an integer. By definition, $C_n\left(X, K(\Lambda)\right)$ is the free $K(\Lambda)$-module over the elementary n-chains $[x_0, x_1, \ldots, x_n]$. Similarly, $K\left(C_n(X, \Lambda)\right)$ is a free $K(\Lambda)$-module with a basis given by the elements $\left[[x_0, x_1, \ldots, x_n], 0\right]$ where $[x_0, x_1, \ldots, x_n]$ is an elementary n-chain. Consider the $K(\Lambda)$-morphisms

$$\begin{aligned} \alpha_n \colon C_n\left(X, K(\Lambda)\right) &\longrightarrow K\left(C_n(X, \Lambda)\right) & \qquad \beta_n \colon K\left(C_n(X, \Lambda)\right) &\longrightarrow C_n\left(X, K(\Lambda)\right) \\ [x_0, x_1, \ldots, x_n] &\longmapsto \left[[x_0, x_1, \ldots, x_n], 0\right], & \left[[x_0, x_1, \ldots, x_n], 0\right] &\longmapsto [x_0, x_1, \ldots, x_n]. \end{aligned}$$

Immediate computations show that the families of $K(\Lambda)$-morphisms $\{\alpha_n\}$ and $\{\beta_n\}$ are morphisms of $K(\Lambda)$-semimodules which are inverses of each other, and the claim follows. □

A.4 Functoriality of semiring homology

In this section, we prove that homology is a functor from the category of directed simplicial complexes to the category of graded Λ-semimodules. We also show that two morphisms allowing for the construction of the prism operator must induce the same map on homology, a result we will need to prove our persistent homology stability result. Recall the definition of morphism of directed simplicial complexes, Definition 3.8.

Definition A.27 Let $f \colon X \to Y$ be a morphism of directed simplicial complexes. Then f induces morphisms of Λ-semimodules $C(f) = \{C_n(f)\}$,

$$\begin{aligned} C_n(f) \colon C_n(X, \Lambda) &\longrightarrow C_n(Y, \Lambda) \\ [x_0, x_1, \ldots, x_n] &\longmapsto [f(x_0), f(x_1), \ldots, f(x_n)]. \end{aligned}$$

We will often abbreviate $C_n(f) = f_n$.

Proposition A.28 *If $f\colon X \to Y$ is a morphism of directed simplicial complexes, the family of maps $\{C_n(f)\}$ is a Λ-homomorphism of chain complexes. Therefore, it induces a map $H_n(f)\colon H_n(X,\Lambda) \to H_n(Y,\Lambda)$.*

Proof Simple computations show that $f_{n-1}\partial_n^+ = \partial_n^+ f_n$ and that $f_{n-1}\partial_n^- = \partial_n^- f_n$. □

Remark A.29 As a consequence of Proposition A.26, the map $C_n(f)\colon C_n\big(X, K(\Lambda)\big) \to C_n\big(Y, K(\Lambda)\big)$ is precisely $K\big(C_n(f)\big)\colon K\big(C_n(X,\Lambda)\big) \to K\big(C_n(Y,\Lambda)\big)$.

Corollary A.30 *Homology is a functor from the category of directed simplicial complexes to the category of graded Λ-semimodules. In particular, isomorphic directed simplicial complexes have isomorphic homologies. Note further that when Λ is a ring, the homology functor is precisely the functor introduced in Sect. 3.2.*

We finish this section by showing a sufficient condition for two morphisms to induce the same map on homology. We will need this result to prove that our definition of persistent homology is stable.

Lemma A.31 *Let Λ be a cancellative semiring. Let X and Y be two directed simplicial complexes and let $f, g\colon X \to Y$ be morphisms of directed simplicial complexes such that if $(x_0, x_1, \ldots, x_n) \in X$, then $\big(f(x_0), f(x_1), \ldots, f(x_i), g(x_i), \ldots, g(x_n)\big) \in Y$ for every $i = 0, 1, \ldots, n$. Then, $H_n(f) = H_n(g)$ for every $n \geq 0$.*

Proof For $x = [x_0, x_1, \ldots, x_n] \in C_n(X,\Lambda)$ an elementary n-chain, define

$$s_n^+[x_0, x_1, \ldots, x_n] = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} [f(x_0), \ldots, f(x_{2i}), g(x_{2i}), \ldots, g(x_n)], \text{ and}$$

$$s_n^-[x_0, x_1, \ldots, x_n] = \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} [f(x_0), \ldots, f(x_{2i+1}), g(x_{2i+1}). \ldots, g(x_n)].$$

Now recall that since Λ is cancellative, the canonical map $k_n\colon C_n(X,\Lambda) \to K\big(C_n(X,\Lambda)\big)$ is injective for all n. We will show that $(s^+, s^-)\colon C(f) \simeq C(g)$ by proving that both sides of the equality in Definition A.17 have the same images through k_n. We will also make use of the isomorphism $K\big(C_n(X,\Lambda)\big) \cong C_n\big(X, K(\Lambda)\big)$ established in Proposition A.26, and that for any Λ-morphism $h\colon C_n(X,\Lambda) \to C_n(X,\Lambda)$, $k_n \circ h = K(h) \circ k_n$.

By abuse of notation, we write $k_n(x) = x = [x_0, x_1, \ldots, x_n] \in C_n\big(X, K(\Lambda)\big)$. Denote $\partial_n = K(\partial_n^+) - K(\partial_n^-)$ and $s_n = K(s_n^+) - K(s_n^-)$. Then,

$$\begin{aligned}\partial_n(x) &= \sum_{i=0}^{n} (-1)^i [x_0, x_1, \ldots, \widehat{x_i}, \ldots, x_n], \quad s_n(x) \\ &= \sum_{j=0}^{n} (-1)^j \big[f(x_0), f(x_1), \ldots, f(x_j), g(x_j), \ldots, g(x_n)\big].\end{aligned}$$

Note that, by appropriately grouping the terms in the equality in Definition A.17, it suffices to show that $K(g_n) - K(f_n)$ and $\partial_{n+1}s_n + s_{n-1}\partial_n$ have the same image on x. These computations are analogous to the ones in Lemma 3.14. □

A.5 Comparison to directed (persistent) homology

In Sect. 3, directed homology modules were introduced by taking submodules of homology generated by cycles with only non-negative coefficients on the elementary chains. This was done for the rings $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. These three rings have in common that they are the Grothendieck's completion of a cancellative, zerosumfree semiring. Namely, $\mathbb{Z} = K(\mathbb{N})$, $\mathbb{Q} = K(\mathbb{Q}^+)$, $\mathbb{R} = K(\mathbb{R}^+)$. It turns out that homology with coefficients in this class of semirings has the ability to detect directed cycles, and what we have introduced in the main body of this article is a particular case of this fact.

Proposition A.32 *Let $\Lambda \in \{\mathbb{N}, \mathbb{Q}^+, \mathbb{R}^+\}$ and let X be a directed simplicial complex. Consider $Z_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$ the set of directed cycles introduced in Definition 3.6. Then, $Z_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big) \cong Z_n(X, \Lambda)$ as Λ-semimodules.*

Proof First note that for the considered semirings, Λ can be seen as the non-negative elements of $K(\Lambda)$. Using this identification, it is straightforward to show that $Z_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$ is a Λ-semimodule. In particular, linear combinations of directed cycles with non-negative coefficients are directed cycles.

By abuse of notation, we will denote elementary n-chains in the same way in $C_n(X, \Lambda)$ and in $C_n\big(X, K(\Lambda)\big)$. Assume that $\sum_{i=1}^n \lambda_i c_i \in Z_n(X, \Lambda)$. Then clearly $\sum_{i=1}^n \lambda_i c_i \in Z_n\big(X, K(\Lambda)\big)$ (see Proposition A.26). Furthermore, since the involved coefficients are non-negative, $\sum_{i=1}^n \lambda_i c_i \in Z_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$. Consequently, we can introduce a map

$$\begin{aligned} \varphi\colon Z_n(X, \Lambda) &\longrightarrow Z_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big) \\ \sum_{i=1}^n \lambda_i c_i &\longmapsto \sum_{i=1}^n \lambda_i c_i. \end{aligned}$$

It is immediate to check that φ is a morphism of Λ-semimodules. It is also injective. We only need to prove that it is surjective. For that, it would be enough to prove that if $\sum_{i=1}^n \lambda_i c_i \in Z_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$, then $\sum_{i=1}^n \lambda_i c_i \in Z_n(X, \Lambda)$. To that purpose, note that $x = \sum_{i=1}^n \lambda_i c_i$ can be regarded as an element of $C_n(X, \Lambda)$ as the involved coefficients are non-negative, and thus, in Λ. We only need to prove that $\partial^+(x) = \partial^-(x)$. Since Λ is cancellative, it is enough to prove that $K\big(\partial^+(x)\big) = K\big(\partial^-(x)\big)$, or equivalently, that $K\big(\partial^+(x)\big) - K\big(\partial^-(x)\big) = 0$. But this follows from Proposition A.26 and from the fact that $x \in Z_n\big(X, K(\Lambda)\big)$. □

In general, we could use homology with coefficients in a cancellative, zerosumfree semiring to detect directed cycles. Namely, we can prove an analogous to Proposition 3.24 for homology with semiring coefficients.

Proposition A.33 *Let Λ be a cancellative, zerosumfree semiring. Let X be a 1-dimensional directed simplicial complex with vertex set $\{v_0, v_1, \ldots, v_n\}$, $n \geq 1$, and with 1-simplices $e_0, e_1, \ldots, e_n$ where either $e_i = (v_i, v_{i+1})$ or $e_i = (v_{i+1}, v_i)$, $i = 0, 1, \ldots, n$, and where $v_{n+1} = v_0$. Then, $H_1(X, \Lambda) = \Lambda$ is non-trivial if and only if either $e_i = (v_i, v_{i+1})$ for all i or $e_i = (v_{i+1}, v_i)$ for all i, and $H_1(X, \Lambda) = \{0\}$ in any other case.*

Proof Let $[e_i]$ denote the elementary 1-chain associated to e_i. Let $x = \sum_{i=0}^{n} \lambda_i [e_i]$ be any non-trivial 1-cycle. We can assume without loss of generality that $\lambda_0 \neq 0$. If $e_0 = (v_0, v_1)$, then $\lambda_0[v_1]$ is a non-trivial summand in $\partial^+(e_0)$. Since Λ is zerosumfree, such summand does not have an inverse, thus $\lambda_0[v_1]$ must be a summand in $\partial^-(x)$. Now note that e_1 is the only other simplex involving the vertex v_1. Furthermore, $[v_1]$ appears in $\partial^-[e_1]$ if and only if $e_1 = (v_1, v_2)$, in which case $\partial^-[e_1] = \lambda_1 v_1$. We further deduce that $\lambda_1 = \lambda_0$.

By proceeding iteratively, we deduce that if x is non-trivial, necessarily $e_i = (v_i, v_{i+1})$ and $\lambda_0 = \lambda_1 = \cdots = \lambda_n = \lambda$, in which case $x = \sum_{i=0}^{n} \lambda [v_i, v_{i+1}]$. Simple computations show that such x is indeed a cycle, and since there are no 2-simplices, it cannot be homologous to any other cycle. Thus, $H_1(X, \Lambda) = \Lambda$. The case where $e_0 = (v_1, v_0)$ is analogous. □

The question would then be how to use this information in the setting of persistent homology. Recall from Remark A.16 that the family of canonical maps $k_{C(X,\Lambda)} \colon C_n(X, \Lambda) \to K\big(C_n(X, \Lambda)\big)$ gives rise to a morphism of Λ-semimodules $k_{C(X,\Lambda)} \colon H_n(X, \Lambda) \to H_n\big(K(C(X, \Lambda))\big)$, which takes the class of x to the class of $[x, 0]$. As a consequence of Proposition A.26, we can also rewrite $H_n\big(K(C(X, \Lambda))\big) = H_n\big(C(X, K(\Lambda))\big)$. Therefore, the map $k_{C(X,\Lambda)}$ above can be regarded as a morphism $k_{C(X,\Lambda)} \colon H_n(X, \Lambda) \to H_n\big(X, K(\Lambda)\big)$. We denote this map by k_X. We then have the following result.

Proposition A.34 *Let X be a directed simplicial complex and let $\Lambda \in \{\mathbb{N}, \mathbb{Q}^+, \mathbb{R}^+\}$. The directed homology module introduced in Definition 3.6, $H_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$, is the submodule of $H_n\big(X, K(\Lambda)\big)$ generated by* $\operatorname{Im} k_X$.

Proof Note that all representatives of classes in $H_n(X, \Lambda)$ are in $Z_n(X, \Lambda)$. Thus, the image through k_X corresponds to a subset of the classes generated by positive cycles, as a consequence of Proposition A.32. Namely, $\operatorname{Im} k_X \subseteq H_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$. To prove the result, it would be enough to prove that every class of a directed cycle is in $\operatorname{Im} k_X$. But this follows from the definition of $H_n(X, \Lambda)$. □

More generally, given Λ a cancellative, zerosumfree semiring, we can define $H_n^{\mathrm{Dir}}\big(X, K(\Lambda)\big)$ as the submodule of $H_n\big(X, K(\Lambda)\big)$ generated by $\operatorname{Im} k_X$, and such module would detect directed cycles as exhibited by Proposition A.34. If $K(\Lambda)$ is a field, we can also use these homology modules to define persistence vector spaces for which we can prove stability results in a way analogous to the results in Sect. 4.

The results of this paper arose from the ideas in the constructions above, which are introduced in their full generality using homology with semiring coefficients. This shows that the idea of introducing a submodule of directed homology by taking cycles with non-negative coefficients has a deeper algebraic meaning, providing more context

for the results in this paper and presenting them in their full generality. Furthermore, similar ideas could be considered for semirings which are not zerosumfree, perhaps allowing for the detection of additional information that persistent homology with ring coefficients may miss. The matter of whether a theory of persistence semimodules could provide more information could also be looked into.

References

Bauer, U., Lesnick, M.: Induced matchings of barcodes and the algebraic stability of persistence. In: Proceedings of the Thirtieth Annual Symposium on Computational Geometry (New York, NY, USA), SOCG'14, Association for Computing Machinery. pp. 355–364 (2014)

Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry. American Mathematical Society, Providence, RI (2001)

Carlsson, G.: Topology and data. Bull. Am. Math. Soc. (N.S.) **46**(2), 255–308 (2009)

Carlsson, G., Mémoli, F., Ribeiro, A., Segarra, S.: Hierarchical quasi-clustering methods for asymmetric networks. In: Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32, ICML'14, JMLR.org. pp. II-352–II-360 (2014)

Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L.J., Oudot, S.: Proximity of persistence modules and their diagrams. In: Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry (New York, NY, USA), SCG '09, Association for Computing Machinery. pp. 237–246 (2009)

Chazal, F., de Silva, V., Glisse, M., Oudot, S.: The Structure and Stability of Persistence Modules. SpringerBriefs in Mathematics. Springer, Cham (2016)

Chowdhury, S., Mémoli, F.: A functorial Dowker theorem and persistent homology of asymmetric networks. J. Appl. Comput. Topol. **2**(1–2), 115–175 (2018)

Chowdhury, S., Mémoli, F.: Persistent path homology of directed networks. In: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (USA), SODA'18, Society for Industrial and Applied Mathematics, pp. 1152–1169 (2018)

Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological Persistence and Simplification, Discrete and Computational Geometry and Graph Drawing (Columbia, SC, 2001), vol. 28, pp. 511–533 (2002)

Edelsbrunner, H., Harer, J.L.: Computational Topology: An Introduction. American Mathematical Society, Providence, RI (2010)

Edelsbrunner, H., Jabłoński, G., Mrozek, M.: The persistent homology of a self-map. Found. Comput. Math. **15**(5), 1213–1244 (2015)

Fajstrup, L., Goubault, E., Haucourt, E., Mimram, S., Raussen, M.: Directed Algebraic Topology and Concurrency. Springer, Berlin (2016)

Fukuda, K., Prodon, A.: Double description method revisited, Combinatorics and computer science (Brest: Lecture Notes in Computer Science vol. 1120. Springer, Berlin 1996), pp. 91–111 (1995)

Golan, J.S.: Semirings and Their Applications. Kluwer Academic Publishers, Dordrecht (1999)

Grandis, M.: Directed Algebraic Topology: Models of Non-reversible Worlds. Cambridge University Press, Cambridge (2009)

Grigor'yan, A., Lin, Y., Muranov, Y., Yau, S.-T.: Path Complexes and their Homologies. J. Math. Sci. **248**, 564–599 (2020)

Kim, K.H., Roush, F.W.: Generalized fuzzy matrices. Fuzzy Sets Syst. **4**(3), 293–315 (1980)

Motzkin, T.S., Raiffa, H., Thompson, G.L., Thrall, R.M.: The Double Description Method, Contributions to the Theory of Games, Annals of Mathematics Studies, no. 28, Princeton University Press, Princeton, NJ, , vol. 2, pp. 51–73 (1953)

Munkres, J.R.: Elements of Algebraic Topology. Addison-Wesley Publishing Company, Menlo Park, CA (1984)

Otter, N., Porter, M.A., Tillmann, U., Grindrod, P., Harrington, H.A.: A roadmap for the computation of persistent homology. EPJ Data Sci. **6**(1), 17 (2017)

Patchkoria, A.: Cohomology of monoids with coefficients in semimodules. Bull. Georgian Acad. Sci. **86**(3), 545–548 (1977)

Patchkoria, A.: Cohomology monoids of monoids with coefficients in semimodules I. J. Homotopy Relat. Struct. **9**(1), 239–255 (2014)

Reimann, M.W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., Markram, H.: Cliques of neurons bound into cavities provide a missing link between structure and function. Front. Comput. Neurosci. **11**, 48 (2017)

Steiner, R.: Omega-categories and chain complexes. Homol. Homotopy Appl. **6**(1), 175–200 (2004)

Tan, Y.-J.: Bases in semimodules over commutative semirings. Linear Algebra Appl. **443**, 139–152 (2014)

Terzer, M.: Large scale methods to enumerate extreme rays and elementary modes, Ph.D. thesis (2009), ETH Zürich, Switzerland

Turner, K.: Rips filtrations for quasimetric spaces and asymmetric functions with stability results. Algebr. Geom. Topol. **19**(3), 1135–1170 (2019)

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

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.