Rips filtrations for quasimetric spaces and asymmetric functions with stability results

KATHARINE TURNER

The Rips filtration over a finite metric space and its corresponding persistent homology are prominent methods in topological data analysis to summarise the "shape" of data. Crucial to their use is the stability result that says if X and Y are finite metric spaces then the (bottleneck) distance between the persistence diagrams constructed via the Rips filtration is bounded by $2d_{GH}(X, Y)$ (where d_{GH} is the Gromov–Hausdorff distance). A generalisation of the Rips filtration to any *symmetric* function $f : X \times X \to \mathbb{R}$ was defined by Chazal, de Silva and Oudot ([Geom. Dedicata 173 (2014) 193–214](#)), where they showed it was stable with respect to the correspondence distortion distance. Allowing asymmetry, we consider four different persistence modules, definable for pairs (X, f) where $f : X \times X \to \mathbb{R}$ is any real valued function. These generalise the persistent homology of the symmetric Rips filtration in different ways. The first method is through symmetrisation. For each $a \in [0, 1]$ we can construct a symmetric function $\mathrm{sym}_a(f)(x, y) = a \min\{d(x, y), d(y, x)\} + (1-a) \max\{d(x, y), d(y, x)\}$. We can then apply the standard theory for symmetric functions and get stability as a corollary. The second method is to construct a filtration $\{\mathcal{R}^{\mathrm{dir}}(X)_t\}$ of ordered tuple complexes where $(x_0, x_2, \ldots, x_p) \in \mathcal{R}^{\mathrm{dir}}(X)_t$ if $d(x_i, x_j) \le t$ for all $i \le j$. Both our first two methods have the same persistent homology as the standard Rips filtration when applied to a metric space, or more generally to a symmetric function. We then consider two constructions using an associated filtration of directed graphs or preorders. For each t we can define a directed graph $\{D(X)_t\}$ where directed edges $x \to y$ are included in $D(X)_t$ whenever $\max\{f(x, y), f(x, x), f(y, y)\} \le t$ (note this is when $d(x, y) \le t$ for $f = d$ a quasimetric). From this we construct a preorder where $x \le y$ if there is a path from x to y in $D(X)_t$. We build persistence modules using the strongly connected components of the graphs $D(X)_t$, which are also the equivalence classes of the associated preorders. We also consider persistence modules using a generalisation of poset topology to preorders.

The Gromov–Hausdorff distance, when expressed via correspondence distortions, can be naturally extended as a correspondence distortion distance to set–function pairs (X, f). We prove that all these new constructions enjoy the same stability as persistence modules built via the original persistent homology for symmetric functions.

1 Introduction

The Rips filtration over a finite metric space (X, d) is a filtration of simplicial complexes $\{\mathcal{R}(X, d)_t\}_{t \in [0, \infty)}$, where $\mathcal{R}(X, d)_t$ is the clique complex over the graph whose vertex set is X and edge set $\{[x, y] : d(x, y) \leq t\}$. It adds topological structure to an otherwise disconnected set of points. The persistent homology of the Rips filtration is widely used in topological data analysis because it encodes useful information about the geometry and topology of the underlying metric space; see Chazal, Cohen-Steiner. Guibas, Mémoli and Oudot [4], Ghrist [10], Lee, Chung, Kang, Kim and Lee [14] and Xia and Wei [20]. There are many potential applications for studying data whose structure is a quasimetric space. Examples include the web hyperlink quasimetric space, road networks, and quasimetrics induced from weighted directed graphs found throughout science (for example biological interaction graphs — see Klamt and von Kamp [12] — or the connections in neural systems; see Kaiser [11] and Reimann, Nolte, Scolamiero, Turner, Perin, Chindemi, Dłotko, Levi, Hess and Markram [18]). More generally we wish to define and show stability of Rips filtrations for sublevel sets of any (not necessarily symmetric) function $f : X \times X \to \mathbb{R}$.

Historically the Rips filtration was defined as a special increasing family of simplicial complexes built from a finite metric space. A metric space is a set X equipped with a distance function $d : X \times X \to \mathbb{R}$ that satisfies the following properties:

  (i)   **Nonnegativity** $d(x, y) \geq 0$ for all $x \in X$.

  (ii) **Symmetry** $d(x, y) = d(y, x)$ for all $x, y \in X$.

  (iii) **Triangle inequality** $d(x, z) \leq d(x, y) + d(y, z)$.

  (iv) **Identity of indiscernibles** $d(x, y) = 0 = d(y, x)$ if and only if $x = y$.

For any $r \geq 0$ we define the Rips complex of X at length scale r, denoted by $\mathcal{R}(X, d)_r$, as the abstract simplicial complex where $[x_0, x_1, \ldots, x_k] \in \mathcal{R}(X, d)_r$ whenever $d(x_i, x_j) \leq r$ for all i and j. We can think of $\mathcal{R}(X, d)_r$ as adding a topological structure of length scale r. It is easy to check that if $r \leq s$ then $\mathcal{R}(X, d)_r \subset \mathcal{R}(X, d)_s$. We thus can define the Rips filtration of X as the increasing family of simplicial complexes $\{\mathcal{R}(X, d)_r\}_{r \in [0, \infty)}$.

Two classic types of examples of Rips filtrations are examples that come from finite point clouds sitting inside some larger space (such as Euclidean space) and examples built from graphs. If $X \subset \mathbb{R}^d$ is a set of points then it inherits a finite metric space structure from that of $\mathbb{R}^d$; the distance function is just the restriction of the Euclidean

distance function to the set X. Given a graph G (with or without lengths on the edges) we can let the vertices of the graph be the finite set X and then construct a distance function on X by defining $d(x, y)$ as the shortest path length of all the paths from x to y in G.

From the Rips filtration we can produce a persistence module which describes its persistent homology. A persistence module is a family of vector spaces $\{V_t : t \in \mathbb{R}\}$ equipped with linear maps $\phi_s^t \colon V_s \to V_t$ for each pair $s \leq t$ such that $\phi_t^t = \mathrm{id}$ and $\phi_s^t = \phi_r^t \circ \phi_s^r$ whenever $s \leq r \leq t$. The persistence module we construct from the persistent homology of a Rips filtration over (X, d) has vector spaces $\{H_*(\mathcal{R}(X, d)_t)\}_{t \in [0, \infty)}$ along with maps on homology induced by inclusions, $\phi_s^t = \iota_* \colon H_*(\mathcal{R}(X, d)_s) \to H_*(\mathcal{R}(X, d)_t)$ when $s \leq t$.

Arguably the most important theoretical results in topological data analysis are the stability theorems. These stability results come in a variety of forms but generally say that if two sets of input data are close then various persistence modules computed from them are also close. To be specific we need to quantify what is meant by "close" for these different kinds of objects.

We can measure how close persistence modules are via whether there exist suitable families of interleaving maps. This distance is closely related to the bottleneck distance between the corresponding persistence diagrams or barcodes. Two persistence modules, $(\{V_t\}, \{\phi_s^t\})$ and $(\{U_t\}, \{\psi_s^t\})$, are called ϵ–interleaved when there exist families of linear maps $\{\alpha_t \colon V_t \to U_{t+\epsilon}\}$ and $\{\beta_t \colon U_t \to V_{t+\epsilon}\}$ satisfying natural commuting conditions. There is a pseudometric on the space of persistence modules called the interleaving distance, d_{int}, which is the infimum of the set of $\epsilon > 0$ such that there exists an ϵ–interleaving. More details about the interleaving distance are provided in Section 3. In this paper we will be considering a variety of different persistence modules, but we will always use the interleaving distance to quantify "closeness".

Gromov–Hausdorff distance is a classical distance between metric spaces. There are many equivalent formulations of Gromov–Hausdorff distance but for the purposes of this paper we will focus on that using correspondences. The set $\mathcal{M} \subset X \times Y$ is a *correspondence* between X and Y if for all $x \in X$ there exists some $y \in Y$ with $(x, y) \in \mathcal{M}$ and for all $y \in Y$ there is some $x \in X$ with $(x, y) \in \mathcal{M}$. Using correspondences we can define the Gromov–Hausdorff distance between X and Y as

$$(1\text{-}1) \quad d_{\mathrm{GH}}(X, Y) = \frac{1}{2} \inf_{\{\text{correspondences } \mathcal{M}\}} \sup_{(x_1, y_1), (x_2, y_2) \in \mathcal{M}} |d_X(x_1, x_2) - d_Y(y_1, y_2)|.$$

Here $\sup_{(x_1,y_1),(x_2,y_2)\in\mathcal{M}} |d_X(x_1,x_2) - d_Y(y_1,y_2)|$ is the distortion of the correspondence $\mathcal{M}$. We can define the *correspondence distortion distance* between set–function pairs $(X, f\colon X\times X\to\mathbb{R})$ and $(Y, g\colon Y\times Y\to\mathbb{R})$ by

$$d_{\mathrm{CD}}((X,f),(Y,g))$$
$$= \tfrac{1}{2} \inf_{\mathcal{M}\text{ correspondence between }X\text{ and }Y} \sup_{(x_1,y_1),(x_2,y_2)\in\mathcal{M}} |f(x_1,x_2) - g(y_1,y_2)|.$$

This agrees with the standard definition for the Gromov–Hausdorff distance when (X,d_X) and (Y,d_Y) are metric spaces. More background and details about the correspondence distortion distance are presented in Section 2.

Useful as the Rips filtration for finite metric spaces is, there are scenarios where the input is not a finite metric space. For example, it is common in data analysis to considers data sets X equipped with a dissimilarity measure. A dissimilarity measure is a map $d_X\colon X\times X\to\mathbb{R}$ that satisfies $d_X(x,x)=0$ and $d_X(x,y)=d_X(y,x)$ for all $x,y\in X$, but is not required to satisfy any of the other metric space axioms. In [5], Chazal, de Silva and Oudot generalised the notion of a Rips filtration to cover dissimilarity measures and more generally for any symmetric function $f\colon X\times X\to\mathbb{R}$. Just as in the finite metric space case, the Rips complex of X with parameter r, denoted by $\mathcal{R}(X,f)_r$, is defined as the abstract simplicial complex where $[x_0, x_1, \ldots, x_k]\in\mathcal{R}(X,f)_r$ whenever $f(x_i, x_j)\le r$ for all i and j (including $i=j$).

Persistent homology can be applied to any increasing family of topological spaces, so it is then natural to define persistence modules from the persistent homology of Rips filtrations built from any symmetric function. This was shown to be stable in [5].

Theorem *Let $f\colon X\times X\to\mathbb{R}$ and $g\colon Y\times Y\to\mathbb{R}$ be symmetric functions and $\mathcal{R}(X,f)$ and $\mathcal{R}(Y,g)$ their corresponding Rips filtrations. If $d_{\mathrm{CD}}((X,f),(Y,g))$ is finite then for all $\epsilon > d_{\mathrm{CD}}((X,f),(Y,g))$, the kth homology persistence modules of $\mathcal{R}(X,f)$ and $\mathcal{R}(Y,g)$ are ϵ–interleaved. In particular, when (X,d_X) and (Y,d_Y) are compact metric spaces, $\mathcal{R}(X,d_X)$ and $\mathcal{R}(Y,d_Y)$ are ϵ–interleaved for all $\epsilon > 2d_{\mathrm{GH}}(X,Y)$.*

The proofs of the interleaving results in [5] didn't have any requirement on the function $f\colon X\times X\to\mathbb{R}$ except that it had to be symmetric. The purpose of this paper is to complete this generalisation procedure to lose that symmetry requirement. However, there are multiple ways to use asymmetry information, and so we have explored a variety of different constructions.

One method is to study related symmetric functions. We can take our original function f and construct a parametric family of related symmetric functions $\mathrm{sym}_a(f)$ where

$a \in [0, 1]$ and

$$\text{sym}_a(f)(x, y) = a \min\{f(x, y), f(y, x)\} + (1 - a) \max\{f(x, y), f(y, x)\}.$$

We can then construct the Rips filtration as in [5] for the set–function pair $(X, \text{sym}_a(f))$. Notably, if f is a symmetric function to begin with then $\text{sym}_a(f) = f$ for all $a \in [0, 1]$ and hence this symmetrisation process does give a generalisation of Rips filtrations to any set–function pair. We can show that the correspondence distortion distance between $(X, \text{sym}_a(f))$ and $(Y, \text{sym}_a(g))$ is bounded by that between (X, f) and (Y, g). We gain stability for these persistence modules constructed through this symmetrisation process as a corollary.

A limitation with using a filtration of simplicial complexes is that a simplex is an inherently symmetric object. An alternative is to use ordered tuple complexes (shortened to OT complexes). An OT complex K is a sets of ordered tuples $(v_0, v_1, \ldots, v_p)$ such that if $(v_0, v_1, \ldots, v_p) \in K$ then $(v_0, v_1, \ldots, \widehat{v}_i, \ldots, v_p) \in K$ for all i. Note that repetitions of the v_j are allowed. Chain complexes, boundary maps, homology and persistent homology can analogously be defined for OT complexes. We will define the *directed Rips filtration* of OT complexes for $f \colon X \times X \to \mathbb{R}$, as the filtration $\{\mathcal{R}^{\text{dir}}(X, f)_t\}$ of ordered tuple complexes where $(x_0, x_2, \ldots, x_p) \in \mathcal{R}^{\text{dir}}(X, f)_t$ if $f(x_i, x_j) \leq t$ for all $i \leq j$. We call the persistence module produced using the OT homology of the directed Rips filtration the *directed Rips persistence module*.

For each simplicial complex there is a canonical OT complex with isomorphic homology group. Furthermore, since these homology isomorphisms commute with the maps induced by inclusion, the persistence modules of these corresponding complexes are also isomorphic. This implies that these directed Rips filtrations are truly generalisations of the Rips filtration built from a symmetric function. We will prove that the persistence modules constructed from these Rips filtrations are stable with respect to the correspondence distortion distance.

The third generalisation considers connected components. The standard dimension 0 homology can be viewed as the vector space whose elements are linear combinations of connected components in the 1–skeleton (ie the graph containing the 0– and 1–cells). When working with directed graphs there are two notions of connected components: weakly and strongly connected. Completely analogous to the traditional connected components story, we can consider vector space whose elements are formal linear combinations the equivalence classes of *strongly* connected components in the directed graph which is the 1–skeleton of the directed Rips filtration.

Given a function $f\colon X \times X \to \mathbb{R}$, for each real number t we can create a directed graph D_t related to the sublevel set $f^{-1}(-\infty, t]$. The graph D_t should have vertex set $\{x \in X \mid f(x, x) \leq t\}$ and directed edge set $\{x \to y \mid \max\{f(x, x), f(x, y), f(y, y)\} \leq t\}$. We can not include a directed edge $x \to y$ just when $f(x, y) \leq t$ because of the closure conditions a directed graph has to satisfy. For each $t \in \mathbb{R}$ we have a vector space V_t of the formal linear combinations the equivalence classes of strongly connected components (SCCs) of D_t. Whenever $s \leq t$ we have an inclusion map $D_s \subset D_t$ which induces a linear map from V_s to V_t. This process directly constructs a persistence module, which we call the *strongly connected components persistence module*. We prove that these persistence modules are stable with respect to the correspondence distortion distance. We also provide some pseudocode on how to compute the barcode decomposition of the strongly connected components persistence module using a modification of the union-find algorithm.

We also note that the persistence modules generated from formal linear combinations of the weakly connected components have already been covered as the dimension 0 persistent homology of the filtration by sublevel sets of $\mathrm{sym}_1(f)$.

Our fourth method uses the directed graphs described above to create a filtration of preorders. Given a directed graph D over vertices X we say $x \leq y$ if there is a path from x to y. From a filtration of directed graphs we obtain a filtration of preorders. We then can construct persistence modules using poset topology (which can be generalised for all preorders, not just posets, discussed in the appendix). We will call these *preorder persistence modules*. We prove that these preorder persistence modules are stable with respect to the correspondence distortion distance. If $f\colon X \times X \to \mathbb{R}$ is a symmetric function, then the dimension 0 preorder persistence module is the same as that of the persistent homology of the standard Rips filtration $\mathcal{R}(X, f)$ and its higher-dimensional preorder persistence modules are always trivial. This implies that preorder persistence modules are describing asymmetry information.

1.1 Related other works

Other related work involves approaches in topological data analysis for incorporating asymmetry information. Ordered set homology is used in [18] in order to study the topology of brain networks. There has been a series of papers by Chowdhury and Mémoli [6; 8; 7] about other constructions of persistence modules which incorporate asymmetry information.

2 Directed graphs, quasi- and pseudometric spaces and the correspondence distortion distance

The original stability result in topological data analysis for Rips filtrations was for filtrations of simplicial complexes built from metric spaces and the bound between persistence modules in terms of the Gromov–Hausdorff distance. This was generalised in [5] to consider symmetric functions and the bound between the functions was the correspondence distance. However, there are many applications where asymmetry naturally arises, of which important examples are quasimetric spaces, such as those constructed as the path metric of some directed graph (with or without weights on the directed edges).

Definition 1 A *directed graph* is an ordered pair $D = (V, A)$ where V is a set whose elements are called *vertices* and A is a set of ordered pairs of vertices called *directed edges* or arrows. A *weighted directed graph* is a directed graph where each arrow is given a nonnegative weight.

Note that a graph can be thought of as a directed graph such that whenever a directed edge $v \to w$ is in A its opposite direction $w \to v$ must also be in A.

Definition 2 Let X be a nonempty set and $d \colon X \times X \to \mathbb{R}$. Consider the following potential properties of d:

   (1) $d(x, x') \geq 0$ for all $x, x' \in X$.

   (2) $d(x, x') = d(x', x)$ for all $x, x' \in X$.

   (3)   For all $x, x' \in X$, $x = x'$ if and only if $d(x, x') = 0$ and $d(x', x) = 0$.

   (4) $d(x, x'') \leq d(x, x') + d(x', x'')$ for all $x, x', x'' \in X$.

If (X, d) satisfies (1), (2), (3) and (4), it is called a metric space. If (X, d) satisfies (1), (3) and (4), it is called a *quasimetric space* and we can call d a *quasimetric*. If (X, d) satisfies (1), (2) and (4), it is called a *pseudometric space* and we can call d a *pseudometric*. If (X, d) satisfies (1) and (4), it is called a *pseudoquasimetric space* and we can call d a *pseudoquasimetric*.

We can build examples of these different types of spaces using weighted directed graphs. Given a weighted directed graph $D = (V, A)$ and two vertices $x, y \in V$, we call $x = v_0, v_1, v_2, \dots, v_m = y$ a path from x to y if all of the arrows $v_i \to v_{i+1}$ are in A. The length of that path $(x = v_0, v_1, v_2, \dots, v_m = y)$ is the sum of the weights

$\sum_{i=0}^{m-1} w(v_i \to v_{i+1})$. Construct $d\colon V \times V \to \mathbb{R}$ by setting $d(x, y)$ to be the length of the shortest path from x to y (and ∞ if no path exists). Since each arrow has nonnegative weight, the function d automatically satisfies (1) in Definition 2. By considering the concatenation of paths, we can easily see that d also automatically satisfies (4) in Definition 2. Thus, (V, d) must always be a quasipseudometric space.

More generally, we can consider any function $f\colon X \times X \to \mathbb{R}$ not necessarily satisfying any of the properties (1)–(4). It is in this most general setting that we will prove stability theorems.

The *Gromov–Hausdorff distance* between metric spaces (X, d_X) and (Y, d_Y) is often defined by
$$d_{\mathrm{GH}}(X, Y) = \inf_{Z,\, f\colon X \to Z,\, g\colon Y \to Z} d_{H,Z}(f(X), g(Y)),$$
where the infimum is taken over all metric spaces Z and isometric embeddings f and g to Z from X and Y, respectively, and $d_{H,Z}$ is the Hausdorff distance between subsets of Z. It is a standard result that the Gromov–Hausdorff distance is a metric on the space of compact metric spaces.

A useful alternative, but equivalent, formula for the Gromov–Hausdorff distance can be given through correspondences. The set $\mathcal{M} \subset X \times Y$ is a *correspondence* between X and Y if for all $x \in X$ there exists some $y \in Y$ with $(x, y) \in \mathcal{M}$ and for all $y \in Y$ there is some $x \in X$ with $(x, y) \in \mathcal{M}$. Using correspondences we can write

$$d_{\mathrm{GH}}(X, Y)$$
$$= \tfrac{1}{2} \inf_{\mathcal{M} \text{ correspondence between } X \text{ and } Y} \; \sup_{(x_1,y_1),(x_2,y_2)\in\mathcal{M}} |d_X(x_1, x_2) - d_Y(y_1, y_2)|.$$

More generally, given functions $f\colon X \times X \to \mathbb{R}$ and $g\colon Y \times Y \to \mathbb{R}$ we can call $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M}) = \sup_{(x_1,y_1),(x_2,y_2)\in\mathcal{M}} |f(x_1, x_2) - g(y_1, y_2)|$ the *distortion* of the correspondence $\mathcal{M}$. We can then define the correspondence distortion distance by minimising this correspondence distortion.

Definition 3 For set–function pairs $(X, f\colon X \times X \to \mathbb{R})$ and $(Y, g\colon Y \times Y \to \mathbb{R})$ the *correspondence distance* between them can be defined as

$$d_{\mathrm{CD}}((X, f), (Y, g))$$
$$= \tfrac{1}{2} \inf_{\mathcal{M} \text{ correspondence between } X \text{ and } Y} \mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$$
$$= \tfrac{1}{2} \inf_{\mathcal{M} \text{ correspondence between } X \text{ and } Y} \; \sup_{(x_1,y_1),(x_2,y_2)\in\mathcal{M}} |f(x_1, x_2) - g(y_1, y_2)|.$$

This agrees with the standard definition for the Gromov–Hausdorff distance when (X, d_X) and (Y, d_Y) are metric spaces. It is straightforward to verify that d_{CD} is a pseudometric on the space of all set–function pairs and a metric on the space of finite quasimetric spaces. The proofs are analogous to that for metric spaces discussed in [1].

3 Background: persistence modules

In this section we will cover some background theory on persistence modules and the interleaving distance between persistence modules. This is important because the interleaving distance between persistence modules bounds the bottleneck distance between their corresponding persistence diagrams. To introduce and motivate the concepts we will provide a brief summary of the theory of persistent homology. We will omit most of the details as we will be phrasing all results in later sections in terms of persistence modules. For more details about the history and applications of persistent homology we refer the reader to [19; 10; 9; 2].

Persistent homology describes how the homology groups evolve over an increasing family of topological spaces. Throughout this section let $K = \{K_t\}$ denote a family of reasonable topological spaces such that $K_s \subset K_t$ whenever $s \leq t$. Given $s \leq t$ the kth–*dimensional persistent homology group* for K from s to t consists of the kth–dimensional homology classes in K_s that "persist" until K_t, that is, $Z_k(K_s)/(Z_k(K_t) \cup B_k(K_s))$. This is isomorphic to the image of the induced map on homology $\iota_*\colon H_k(K_s) \to H_k(K_t)$ from the inclusion $K_s \subset K_t$.

Barcodes and persistence diagrams were introduced as discrete summaries of persistent homology information. Each barcode consists of a multiset of real intervals called bars. The barcode corresponding to the kth–dimensional persistent homology of K is $\{I_1, I_2, \ldots, I_n\}$ if, for all $s \leq t$, the dimension of $\mathrm{im}(\iota_*\colon H_k(K_s) \to H_k(K_t))$ equals the number of bars in $\{I_1, I_2, \ldots, I_n\}$ that contain $[s, t)$. The corresponding persistence diagram is the multiset $\{(a_i, b_i)\}$ of points in $\mathbb{R}^2$, where a_i and b_i are the endpoints of the bar I_i, alongside infinitely many copies of every point along the diagonal (these diagonal points are acting the role of empty intervals).

Barcodes and persistence diagrams have played a prominent role in applied topology as topological summaries of data. In particular, they can provide insight into the "shape" of point cloud data through the persistent homology of the Rips filtration over that point cloud. Much of the power behind the use of barcodes and persistence diagrams comes

from stability theorems, such as the stability theorem for the persistent homology of the Rips filtration over a finite metric space.
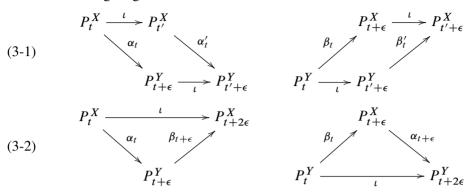
Persistence, such as persistent homology of a filtration of simplicial complexes, can be defined directly at an algebraic level. In [21], Zomorodian and Carlsson introduced the concept of a persistence module and proved that barcodes (and equivalently persistence diagrams) can be defined for persistence modules satisfying reasonable finiteness conditions. It was shown in [3] that we can define a distance between persistence modules (called the interleaving distance) and that the interleaving distance between persistence modules is a bound on the bottleneck distance of their corresponding persistence diagrams. Throughout this paper we will work directly with persistence modules.

Definition 4 Let R be a commutative ring with unity. A *persistence module* over $A \subset \mathbb{R}$ is a family $\{P_t\}_{t \in A}$ of R–modules indexed by real numbers, together with a family of homomorphism $\{\iota_t^s \colon P_t \to P_s\}$ such that $\iota_t^r = \iota_s^r \circ \iota_t^s$ for all $t \leq s \leq r$ and $\iota_t^t = \mathrm{id}\, P_t$.

If R is a field then the P_t are all vector fields and the ι_t^s are linear maps. As is standard in topological data analysis, we will assume throughout that R is the fixed field F (usually taken to be $\mathbb{F}_2$ for computational reasons). In the theory of persistence modules there are technical requirements about tameness. We say $\mathcal{P}$ is *tame* if rank ι_t^s is always finite for any $s < t$. A sufficient condition for tameness is that X is finite, which is almost always true in any application. It is less straightforward in the constructions involving asymmetry to provide other nice sufficient conditions which would ensure the resulting persistence modules are tame (see the future directions). When the persistence modules are tame, the interleaving results will immediately imply a stability theorem for the persistence diagrams/barcodes.

The space of persistence modules is a pseudometric space under the interleaving distance function. Here we will define the interleaving distance between two persistence modules as the infimum of $\epsilon > 0$ such that they are ϵ–interleaved. In this we slightly differ from [3], where they define both strongly and weakly ϵ–interleaved, both of which are weaker than our notion of interleaving. More details about the pseudometric space structure of persistence modules and how the interleaving distance between persistence modules relates to the distances between corresponding persistence diagrams can be found in [3; 10; 21].

Definition 5 Two persistence modules $\mathcal{P}^X$ and $\mathcal{P}^Y$ are ϵ–*interleaved* if there exist families of homomorphisms $\{\alpha_t\colon P_t^X \to P_{t+\epsilon}^Y\}_{t\in\mathbb{R}}$ and $\{\beta_t\colon P_t^Y \to P_{t+\epsilon}^X\}_{t\in\mathbb{R}}$ such that the following diagrams commute:

(3-1)

$$
\begin{array}{ccc}
P_t^X & \overset{\iota}{\longrightarrow} & P_{t'}^X \\
& \alpha_t \searrow \quad \alpha_t' \searrow & \\
& P_{t+\epsilon}^Y \underset{\iota}{\longrightarrow} P_{t'+\epsilon}^Y &
\end{array}
\qquad
\begin{array}{ccc}
P_{t+\epsilon}^X & \overset{\iota}{\longrightarrow} & P_{t'+\epsilon}^X \\
\beta_t \nearrow & & \beta_t' \nearrow \\
P_t^Y \underset{\iota}{\longrightarrow} P_{t'+\epsilon}^Y & &
\end{array}
$$

(3-2)

$$
\begin{array}{ccc}
P_t^X & \overset{\iota}{\longrightarrow} & P_{t+2\epsilon}^X \\
\alpha_t \searrow & \beta_{t+\epsilon} \nearrow & \\
& P_{t+\epsilon}^Y &
\end{array}
\qquad
\begin{array}{ccc}
& P_{t+\epsilon}^X & \\
\beta_t \nearrow & & \alpha_{t+\epsilon} \searrow \\
P_t^Y & \underset{\iota}{\longrightarrow} & P_{t+2\epsilon}^Y
\end{array}
$$

Definition 6 Two persistence modules $\mathcal{P}^X$ and $\mathcal{P}^Y$ are *isomorphic* if they are 0–interleaved.

The diagrams in (3-1) and (3-2) are slightly different from those given in [3] but the diagrams here commuting will imply that theirs also commute.

If $\mathcal{P}^X$ and $\mathcal{P}^Y$ are ϵ_1–interleaved and $\mathcal{P}^Y$ and $\mathcal{P}^Z$ are ϵ_2–interleaved then composing homomorphisms shows that $\mathcal{P}^X$ and $\mathcal{P}^Z$ are $(\epsilon_1+\epsilon_2)$–interleaved. We can define a pseudodistance on the space of persistence modules, called the *interleaving distance*, where the interleaving distance between $\mathcal{P}^X$ and $\mathcal{P}^Y$ is the infimum of the set of $\epsilon > 0$ such that $\mathcal{P}^X$ and $\mathcal{P}^Y$ are ϵ–interleaved. It is worth noting that two persistence modules might have interleaving distance 0 and yet not be 0–interleaved (and thus not isomorphic).

4 Existing stability results and Rips filtrations constructed from related symmetric functions

In this section we will recall the definition for the Rips filtration of a metric space and more generally for sublevel sets of a symmetric function $f\colon X \times X \to \mathbb{R}$. We will also recall the existing stability results for their persistent homology. Given a function $f\colon X \times X \to \mathbb{R}$ we construct a family of related symmetric functions $\mathrm{sym}_a(f)$ (for $a \in [0, 1]$). We show that the persistent homology constructed from the $\mathrm{sym}_a(f)$ is stable as a corollary of the stability results for symmetric functions under the correspondence distortion distance.

Definition 7 Given a set X and a symmetric function $f\colon X \times X \to \mathbb{R}$, the *Rips filtration* of (X, f) is a family of finite simplicial complexes $\mathcal{R}(X, f) = \{\mathcal{R}(X, f)_t\}_{t \geq 0}$ with $\mathcal{R}(X, f)_t$ the clique complex on the graph with vertices $X_t = \{x \in X : f(x, x) \leq t\}$ and edges $\{[x_1, x_2] \in X_t \times X_t : f(x_1, x_2) \leq t\}$.[1]

Theorem 8 Let $f\colon X \times X \to \mathbb{R}$ and $g\colon Y \times Y \to \mathbb{R}$ be symmetric functions and $\mathcal{R}(X, f)$ and $\mathcal{R}(Y, g)$ their corresponding Rips filtrations. If $d_{\mathrm{CD}}((X, f), (Y, g))$ is finite then for all $\epsilon > d_{\mathrm{CD}}((X, f), (Y, g))$, the kth homology persistence modules of $\mathcal{R}(X, f)$ and $\mathcal{R}(Y, g)$ are ϵ–interleaved. In particular, when (X, d_X) and (Y, d_Y) are compact metric spaces, $\mathcal{R}(X, d_X)$ and $\mathcal{R}(Y, d_Y)$ are ϵ–interleaved for all $\epsilon > 2d_{\mathrm{GH}}(X, Y)$.

Since the only condition required is symmetry of the filtration function, one approach for analysing general functions $f\colon X \times X \to \mathbb{R}$ is to construct related symmetric functions. We will consider a one-parameter family of possible symmetric filtrations. We then prove stability for the Rips filtrations of these symmetric constructions in terms of the correspondence distortion distance between the original set–function pairs.

Definition 9 Let (X, f) be a finite set $X = \{x_1, \ldots, x_N\}$ equipped with function $f\colon X \times X \to \mathbb{R}$. For any $a \in [0, 1]$ we can define a symmetric function

$$\mathrm{sym}_a(f)\colon X \times X \to \mathbb{R},$$
$$(x, y) \mapsto a \min\{f(x, y), f(y, x)\} + (1 - a) \max\{f(x, y), f(y, x)\}.$$

Since $\mathrm{sym}_a(f)$ is symmetric we can construct its Rips filtration $\{\mathcal{R}(X, \mathrm{sym}_a(f))_t\}$, where $\mathcal{R}(X, \mathrm{sym}_a(f))_t$ is the simplicial complex containing $[x_0, x_1, \ldots, x_p]$ whenever $\mathrm{sym}_a(f)(x_i, x_j) \leq t$ for all i and j. We call this the *Rips filtration under* sym_a. If f is a symmetric function then $\mathrm{sym}_a(f) = f$ for all a, which implies that the Rips filtration under sym_a generalises the symmetric Rips filtration.

As a corollary of the stability for symmetric functions we have stability for the symmetrised functions.

Corollary 10 Fix $a \in [0, 1]$ and a homology dimension k. Let (X, f) and (Y, g) be set–function pairs such that $d_{\mathrm{CD}}((X, f), (Y, g))$ is finite. Let P^X and P^Y be

---

[1]Readers need to be warned that sometimes the Rips filtration is defined by adding the edge $[x_1, x_2]$ when $d_X(x_1, x_2) \leq \frac{1}{2}t$ instead of $d_X(x_1, x_2) \leq t$, so sometimes results may differ from here by a corresponding factor of 2.

the corresponding k^{th}–dimensional homology persistence modules constructed from the corresponding Rips filtrations under sym_a $(\mathcal{R}(X, \text{sym}_a(f))$ and $\mathcal{R}(Y, \text{sym}_a(g))$, respectively). Then $d_{\text{int}}(P^X, P^Y) \leq 2d_{\text{CD}}((X, \text{sym}_a(f)), (Y, \text{sym}_a(g)))$.

Unfortunately this method of constructing Rips filtrations is somewhat crude. We can show that in the process of symmetrising we dampen dissimilarities. This is not surprising as the space of symmetric functions is much smaller than that of functions generally. In particular, we will show in Theorem 12 that $d_{\text{CD}}((X, \text{sym}_a(f)), (Y, \text{sym}_a(g))) \leq 2d_{\text{CD}}((X, f), (Y, g))$ for all $a \in [0, 1]$. There are many examples where this inequality is strict. For asymmetric functions, $d_{\text{CD}}((X, \text{sym}_a(f)), (Y, \text{sym}_a(g)))$ is often significantly smaller than $2d_{\text{CD}}((X, f), (Y, g))$. Suppose $X = Y$, $f \colon X \times X \to \mathbb{R}$ is an antisymmetric function and $g = -f$. Then, by construction, $\text{sym}_a(f) = \text{sym}_a(g)$ for all a but for nonzero f, we generally have $d_{\text{CD}}((X, f), (X, -f)) > 0$.

The dampening process through symmetrisation is encapsulated in the following lemma:

Lemma 11 Let $w, \widehat{w}, z, \widehat{z} \in \mathbb{R}$. Then

(i) $|\max\{w, \widehat{w}\} - \max\{z, \widehat{z}\}| \leq \max\{|w - z|, |\widehat{w} - \widehat{z}|\}$,

(ii) $|\min\{w, \widehat{w}\} - \min\{z, \widehat{z}\}| \leq \max\{|w - z|, |\widehat{w} - \widehat{z}|\}$.

Proof We can prove (i) through a series of cases. If $w \leq \widehat{w}$ and $z \leq \widehat{z}$ then $|\max\{w, \widehat{w}\} - \max\{z, \widehat{z}\}| = |\widehat{w} - \widehat{z}|$. If $w \geq \widehat{w}$ and $z \geq \widehat{z}$ then $|\max\{w, \widehat{w}\} - \max\{z, \widehat{z}\}| = |w - z|$.

If $w \leq \widehat{w}$ and $z \geq \widehat{z}$, then

$$|\max\{w, \widehat{w}\} - \max\{z, \widehat{z}\}| = |\widehat{w} - z| \leq \begin{cases} |\widehat{w} - \widehat{z}| & \text{if } \widehat{z} \leq w, \\ |w - z| & \text{if } \widehat{z} \geq w \end{cases}$$
$$\leq \max\{|w - z|, |\widehat{w} - \widehat{z}|\}.$$

Reversing the roles of the letters, we also see that

$$|\max\{w, \widehat{w}\} - \max\{z, \widehat{z}\}| \leq \max\{|w - z|, |\widehat{w} - \widehat{z}|\}$$

whenever $w \geq \widehat{w}$ and $z \leq \widehat{z}$

We can infer (ii) from (i) by replacing each of w, $\widehat{w}$, z and $\widehat{z}$ by their negatives. $\square$

Theorem 12 Fix $a \in [0, 1]$ and a homology dimension k. Let (X, f) and (Y, g) be set–function pairs. Then $d_{\text{CD}}((X, \text{sym}_a(f)), (Y, \text{sym}_a(g))) \leq 2d_{\text{CD}}((X, f), (Y, g))$.

Proof It is sufficient to show that $\text{dis}_{(X, \text{sym}_a(f)), (Y, \text{sym}_a(g))}(\mathcal{M}) \leq \text{dis}_{(X, f), (Y, g)}(\mathcal{M})$ for every correspondence $\mathcal{M}$.

Fix some correspondence $\mathcal{M} \subset X \times Y$ and let $(x_1, y_1), (x_2, y_2) \in \mathcal{M}$. From [Lemma 11](#) (using $w = f(x_1, x_2)$, $\hat{w} = f(x_2, x_1)$, $z = g(y_1, y_2)$ and $\hat{z} = g(y_2, y_1)$) we know that both

$$|\min\{f(x_1, x_2), f(x_2, x_1)\} - \min\{g(y_1, y_2), g(y_2, y_1)\}|$$
$$\leq \max\{|f(x_1, x_2) - g(y_1, y_2)|, |f(x_2, x_1) - g(y_2, y_1)|\}$$

and

$$|\max\{f(x_1, x_2), f(x_2, x_1)\} - \max\{g(y_1, y_2), g(y_2, y_1)\}|$$
$$\leq \max\{|f(x_1, x_2) - g(y_1, y_2)|, |f(x_2, x_1) - g(y_2, y_1)|\}.$$

Taking a convex combination of these equations tells us that

$$(4\text{-}1) \quad |\operatorname{sym}_a(f)(x, \hat{x}) - \operatorname{sym}_a(g)(y, \hat{y})|$$
$$\leq \max\{|f(x, \hat{x}) - g(y, \hat{y})|, |f(\hat{x}, x) - g(\hat{y}, y)|\}.$$

By taking the supremum on both sides over all pairs $\{(x, y), (\hat{x}, \hat{y})\} \in \mathcal{M}$ we see that

$$\operatorname{dis}_{(X, \operatorname{sym}_a(f)), (Y, \operatorname{sym}_a(g))}(\mathcal{M}) \leq \operatorname{dis}_{(X, f), (Y, g)}(\mathcal{M}). \qquad \square$$

5 Persistent homology of OT complexes

Ordered tuple complexes are an alternative to simplicial complexes. We will find them useful as they have more flexibility with regard to order; we can have asymmetric roles within the same tuple.

Definition 13 An *ordered tuple* is a sequence of $(v_0, v_1, v_2, \dots, v_n)$, potentially including repeats. A *ordered tuple complex* (shortened to *OT complex*) is a collection K of ordered tuples such that if $(v_0, v_1, v_2, \dots, v_n) \in K$ then $(v_0, \dots, \hat{v}_i, \dots, v_n) \in K$ for all i (where $(v_0, \dots, \hat{v}_i, \dots, v_n)$ is the ordered tuple with v_i removed).

It is worth emphasising that each ordered tuple is determined by the ordered sequence and not just the underlying vertices; (v_1, v_2, v_3) and (v_3, v_1, v_2) are distinct and not even linearly related.

The ideas of homology and persistent homology naturally extend to OT complexes. Throughout $\mathbb{F}$ will be a fixed field.

Definition 14 Given an OT complex K we can build a chain complex $C_*(K)$ where $C_p(K)$ is the set of all the $\mathbb{F}$–linear combinations of the ordered tuples in K with

length $p + 1$. This is an $\mathbb{F}$–vector space whose basis vectors are the ordered tuples in K of length $p + 1$. We define a boundary map $\partial_p \colon C_p(K) \to C_{p-1}(K)$ by

$$\partial_p((v_0, v_1, v_2, \ldots, v_p)) = \sum_{i=0}^{k} (-1)^i (v_0, \ldots, \widehat{v}_i, \ldots, v_p)$$

and extending linearly. We define the kth homology group of the OT complex K as $H_k(K) = \ker(\partial_{k-1})/\operatorname{im}(\partial_k)$.

When $K_1 \subset K_2$ are both OT complexes, the inclusion of chains induces a map on their homology groups, $\iota_* \colon H_*(K_1) \to H_*(K_2)$.

Definition 15   We say $\mathcal{K} = \{K_t\}$ is filtration of OT complexes if $K_t \subset K_r$ whenever $t \leq r$. We define the kth–*dimensional ordered tuple persistence module* corresponding to $\mathcal{K}$ as follows:

- For each t set the vector space $V_t = H_k(K)$ computed over $\mathbb{F}$.

- For each pair $s \leq t$ we have a linear map induced from inclusion,

$$\iota_{t \to s} \colon H_*(K_s) \to H_*(K_t).$$

It is easy to check that this does satisfy the requirements of a persistence module.

We can define the directed Rips filtration as a filtration of OT complexes where the condition for when an ordered tuple is included is dependent on the order in which the points in the tuple appear. From this filtration of OT complexes we can construct directed Rips persistence modules.

Definition 16   Let (X, f) be a set–function pair. Set $\{\mathcal{R}^{\text{dir}}(X, f)_t\}$ to be the filtration of OT complexes where $(v_0, v_1, \ldots, v_p) \in \mathcal{R}^{\text{dir}}(X, f)_t$ when $f(v_i, v_j) \leq t$ for all $i \leq j$. We call $\{\mathcal{R}^{\text{dir}}(X, f)_t\}$ the *directed Rips filtration* of (X, f). For each dimension k, we will define the kth–*dimensional directed Rips persistence module* as the kth–dimensional ordered tuple persistence module of $\{\mathcal{R}^{\text{dir}}(X, f)_t\}$.

We claim that these directed Rips persistence modules are a generalisation of the Rips persistence modules constructed from symmetric functions. To do this we need to recall some classical relationships between the homology of OT complexes and simplicial complexes. Indeed, a common first example of an OT complex is via a simplicial complex. For a simplicial complex K there is an OT complex K^{OT} where

$(v_0, v_1, \ldots, v_p) \in K^{\mathrm{OT}}$ whenever $[v_0, v_1 \ldots, v_p]$, after removing any repeats, is a simplex in K. In [17], Munkres calls the chain complex $C_*(K^{\mathrm{OT}})$ the *ordered chain complex of* K, and shows that the simplicial homology of K and the OT complex homology of K^{OT} are isomorphic. This isomorphism result holds also for persistence modules of filtrations of simplicial complexes as the isomorphisms on homology groups commute with the induced maps on homology by inclusions. This implies that if $f\colon X \times X \to \mathbb{R}$ is a symmetric function then the kth–dimensional ordered tuple persistence module of $\{\mathcal{R}^{\mathrm{dir}}(X, f)_t\}$ is isomorphic to the kth–dimensional persistence module of $\{\mathcal{R}(X, f)_t\}$.

5.1 Stability of the directed Rips persistence modules

We will want to prove that the directed Rips persistence modules enjoy stability with respect to the correspondence distortion distance. To do this we will compare set–function pairs over different sets via their induced set–function pairs over a common set constructed via a fixed correspondence.

Given functions $f\colon X \times X \to \mathbb{R}$ and $g\colon Y \times Y \to \mathbb{R}$ along with a correspondence $\mathcal{M} \subset X \times Y$, we can pull back the functions f and g to corresponding functions on $\mathcal{M} \times \mathcal{M}$ via the projection maps

$$f^{\mathcal{M}}\colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}, \quad (x_1, y_1) \times (x_2, y_2) \mapsto f(x_1, x_2),$$
$$g^{\mathcal{M}}\colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}, \quad (x_1, y_1) \times (x_2, y_2) \mapsto g(y_1, y_2).$$

The proof of the following lemma follows directly from the definitions of $f^{\mathcal{M}}$ and $g^{\mathcal{M}}$.

Lemma 17 Let (X, f) and (Y, g) be set–function pairs and $\mathcal{M} \subset X \times Y$ a correspondence. Then

$$\|f^{\mathcal{M}} - g^{\mathcal{M}}\|_{\infty} = 2 \operatorname{dis}_{(X,f),(Y,g)}(\mathcal{M}).$$

We will also need to prove that the directed Rips persistence modules over (X, f) and $(\mathcal{M}, f^{\mathcal{M}})$ are isomorphic. To do this we will introduce the notion of the expansion of an OT complex.

Definition 18 Let K be an OT complex. We say that K is *closed under adjacent repeats* if whenever $(v_0, v_1, \ldots, v_p) \in C_p(K)$ then $(v_0, \ldots, v_i, v_i, \ldots, v_p) \in C_{p+1}(K)$ for all $i = 0, 1, \ldots, p$.

It is worth observing that, by construction, $\{\mathcal{R}^{\mathrm{dir}}(X, f)_t\}$ is closed under adjacent repeats for any set–function pair (X, f).

Definition 19 Let K and $\tilde{K}$ be OT complexes, both closed under adjacent repeats, over vertex sets V and $\tilde{V}$, respectively. We say that $\tilde{K}$ is an *expansion* of K if there exists a surjective map $\pi\colon \tilde{V} \to V$ and an injective map $\iota\colon V \to \tilde{V}$ such that $\pi \circ \iota = \mathrm{id}_V$ and $(v_0, v_1, \ldots, v_p) \in \tilde{K}$ if and only if $(\pi(v_0), \pi(v_1), \ldots, \pi(v_p)) \in K$.

Let $\mathcal{K} = \{K_t\}$ and $\tilde{\mathcal{K}} = \{\tilde{K}_t\}$ be filtrations of OT complexes over vertex sets V and $\tilde{V}$, respectively. We say that $\tilde{\mathcal{K}}$ is an *expansion* of $\mathcal{K}$ if there exists a surjective map $\pi\colon \tilde{V} \to V$ and an injective map $\iota\colon V \to \tilde{V}$ such that $\pi \circ \iota = \mathrm{id}_V$ and, for all t, $(v_0, v_1, \ldots, v_p) \in \tilde{K}_t$ if and only if $(\pi(v_0), \pi(v_1), \ldots, \pi(v_p)) \in K_t$.

Proposition 20 *If $\mathcal{K} = \{K_t\}$ and $\tilde{\mathcal{K}} = \{\tilde{K}_t\}$ are filtrations of OT complexes such that $\tilde{\mathcal{K}}$ is an expansion of $\mathcal{K}$ then the OT persistence modules of $\mathcal{K}$ and $\tilde{\mathcal{K}}$ are isomorphic.*

Proof Without loss of generality we can relabel the points in V to consider it as a subset of $\tilde{V}$ (relabelling $v \in V$ as $\iota(v) \in \tilde{V}$). In this case ι is the inclusion map and π is a projection map.

Both $\pi\colon \tilde{K}_t \to K_t$ and $\iota\colon K_t \to \tilde{K}_t$ induce chain maps, $\pi_\#\colon C_*(\tilde{K}_t) \to C_*(K_t)$ and $\iota_\#\colon C_*(K_t) \to C_*(\tilde{K}_t)$. Observe that $\pi_\# \circ \iota_\# = \mathrm{id}\colon C_*(K_t) \to C_*(K_t)$, so $\pi_* \circ \iota_* = \mathrm{id}\colon H_*(K_t) \to H_*(K_t)$ for all t.

Suppose $(v_0, v_1, \ldots, v_i, \ldots, v_p) \in C_p(\tilde{K}_t)$. To construct a prism operator later we want to show that

$$(v_0, v_1, \ldots, v_i, \pi(v_i), \ldots, \pi(v_p)) \in C_{p+1}(\tilde{K}_t).$$

To do this we use that $\tilde{K}_t$ is closed under adjacent repeats, the definition of expansions (twice) and the property that π is a projection (so $\pi(\pi(v_j)) = \pi(v_j)$):

$(v_0, v_1, \ldots, v_i, \ldots, v_p) \in C_p(\tilde{K}_t)$
$$\implies (v_0, v_1, \ldots, v_i, v_i, \ldots, v_p) \in C_{p+1}(\tilde{K})$$
$$\implies (\pi(v_0), \pi(v_1), \ldots, \pi(v_i), \pi(v_i), \ldots, \pi(v_p)) \in C_{p+1}(K_t)$$
$$\implies (\pi(v_0), \pi(v_1), \ldots, \pi(v_i), \pi(\pi(v_i)), \ldots, \pi(\pi(v_p))) \in C_{p+1}(K_t)$$
$$\implies (v_0, v_1, \ldots, v_i, \pi(v_i), \ldots, \pi(v_p)) \in C_{p+1}(\tilde{K}_t).$$

Consider the prism operator

$$P((v_0, v_1, \ldots, v_p)) = \sum_{i=0}^{p} (-1)^i ((v_0, v_1, \ldots, v_i, \pi(v_i), \pi(v_{i+1}), \ldots, \pi(v_p))).$$

Routine algebra shows that $\partial P + P\partial = i_\# \circ \pi_\# - \mathrm{id}$ and thus $i_\# \circ \pi_\#$ is chain homotopic to the identity. This implies $i_* \circ \pi_*\colon H_*(\tilde{K}_t) \to H_*(\tilde{K}_t)$ is the identity function.

The chain maps $\pi_\#$ and $i_\#$ commute with the inclusion maps for the filtrations of OT complexes and hence the following diagrams commute:

$$
\begin{array}{ccc}
H_*(\tilde{K}_s) \xrightarrow{\iota_*} H_*(\tilde{K}_t) & \quad & H_*(K_s) \xrightarrow{\iota_*} H_*(K_t) \\
\Big\downarrow{\scriptstyle\pi_*} \qquad \Big\downarrow{\scriptstyle\pi_*} & & \Big\downarrow{\scriptstyle i_*} \qquad \Big\downarrow{\scriptstyle i_*} \\
H_*(K_s) \xrightarrow{\iota_*} H_*(K_t) & & H_*(\tilde{K}_s) \xrightarrow{\iota_*} H_*(\tilde{K}_t)
\end{array}
$$

Since $i_* \circ \pi_* = \mathrm{id}\colon H_*(\tilde{K}_t) \to H_*(\tilde{K}_t)$ and $\pi_* \circ i_* = \mathrm{id}\colon H_*(K_t) \to H_*(K_t)$ for all t we see that $\mathcal{K}$ and $\tilde{\mathcal{K}}$ are isomorphic. $\qquad\square$

Theorem 21 *Let (X, f) and (Y, g) be set–function pairs such that $d_{\mathrm{CD}}((X, f), (Y, g))$ is finite. Let P^X and P^Y be the corresponding k^{th}–dimensional homology persistence modules constructed from the corresponding directed Rips filtrations $\{\mathcal{R}^{\mathrm{dir}}(X, f)_t\}$ and $\{\mathcal{R}^{\mathrm{dir}}(Y, g)_t\}$. Then $d_{\mathrm{int}}(P^X, P^Y) \le 2 d_{\mathrm{CD}}((X, f), (Y, g))$.*

Proof Since $d_{\mathrm{CD}}((X, f), (Y, g))$ is finite, there exists some correspondence $\mathcal{M} \subset X \times Y$ with $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ finite. Fix a correspondence $\mathcal{M} \subset X \times Y$ with $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ finite. From this correspondence construct directed Rips filtrations $\{\mathcal{R}^{\mathrm{dir}}(\mathcal{M}, f^{\mathcal{M}})_t\}$ and $\{\mathcal{R}^{\mathrm{dir}}(\mathcal{M}, g^{\mathcal{M}})_t\}$ with corresponding k^{th}–dimensional persistence modules $P^{(X,\mathcal{M})}$ and $P^{(Y,\mathcal{M})}$.

By construction $\{\mathcal{R}^{\mathrm{dir}}(\mathcal{M}, f^{\mathcal{M}})_t\}$ is an expansion of $\{\mathcal{R}^{\mathrm{dir}}(X, f)_t\}$ and thus by [Proposition 20](#) we know that the persistence modules P^X and $P^{(X,\mathcal{M})}$ are isomorphic. Similarly, we can also show that P^Y and $P^{(Y,\mathcal{M})}$ are isomorphic.

By [Lemma 17](#) we know $\|f^{\mathcal{M}} - g^{\mathcal{M}}\|_\infty \le 2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$. There is an inclusion

$$
\mathcal{R}^{\mathrm{dir}}(\mathcal{M}, f^{\mathcal{M}})_t \subset \mathcal{R}^{\mathrm{dir}}(\mathcal{M}, g^{\mathcal{M}})_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}
$$

for all t as

$$
(v_0, v_1, \ldots, v_n) \in \mathcal{R}^{\mathrm{dir}}(\mathcal{M}, f^{\mathcal{M}})_t
$$
$$
\implies f^{\mathcal{M}}(v_i, v_j) \le t \qquad\qquad \text{for all } i \le j
$$
$$
\implies g^{\mathcal{M}}(v_i, v_j) \le t + \mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M}) \qquad \text{for all } i \le j
$$
$$
\implies (v_0, v_1, \ldots, v_n) \in \mathcal{R}^{\mathrm{dir}}(\mathcal{M}, g^{\mathcal{M}})_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}.
$$

Symmetrically, there are also inclusions

$$\mathcal{R}^{\mathrm{dir}}(\mathcal{M}, g^{\mathcal{M}})_t \subset \mathcal{R}^{\mathrm{dir}}(\mathcal{M}, f^{\mathcal{M}})_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$$

for all t. These inclusion maps induce a $2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ interleaving between $P^{(X,\mathcal{M})}$ and $P^{(Y,\mathcal{M})}$. This implies that P^X and P^Y are $2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$–interleaved.

By considering the infimum of the interleavings constructed by correspondences we see that $d_{\mathrm{int}}(P^X, P^Y)$ is at most $2d_{\mathrm{CD}}((X, f), (Y, g))$. □

5.2 Comparison to ordered-set persistent homology

It is possible to construct homology groups and persistence modules using ordered sets instead of ordered tuples. As a preemptive attempt to reduce confusion, this section will compare this ordered-tuple persistent homology to ordered-set persistent homology. In ordered-set homology we effectively restrict our chains to only contain ordered tuples where there are no repeats. We can still define homology, persistent homology and persistence modules. Furthermore, in some applications this may better reflect the connectivity structure (such as in the analysis of the blue brain project in [18]) but there are two important reasons why we are not considering ordered-set persistence modules as a generalisation of the Rips persistence modules. The first reason is that when we restrict to symmetric functions we do not get persistence modules isomorphic to the standard Rips persistence modules. The second reason is that these persistence modules are not stable with respect to the correspondence distortion distance.

For example, consider the set $X = \{x, y\}$ with the f the zero function. For $t < 0$, then, the corresponding ordered sets complexes are empty with trivial homology. The ordered tuple complexes and Rips simplicial complexes are also empty and have trivial homology. For $t \geq 0$, the corresponding ordered set complex consists of the ordered sets (x), (y), (x, y) and (y, x). It has nontrivial 1–dimensional homology. To see this first observe that $(x, y) + (y, x)$ is a cycle but the space of 2–chains is trivial, so there are no nontrivial 1–chain boundaries. In comparison, the Rips simplicial complex is $[x, y]$, which has no nontrivial 1–cycles. The ordered-tuple complex is more complicated but everything ends up cancelling each other. For example, this cycle of concern in the ordered-set homology, $(x, y) + (y, x)$, is a boundary in the setting of OT homology: $(x, y) + (y, x) = \partial((x, y, x) + (x, x, x))$.

To see that the ordered-set persistence modules are not stable with respect to the correspondence distortion distance, compare (X, f) in the example in the paragraph

above to the single-point space $Z = \{z\}$ with function $g(z) = 0$. The first-dimensional ordered set homology for Z is also trivial and so its first-dimensional persistence module is also trivial. The correspondence $\{(x, z), (y, z)\} \subset X \times Z$ has zero distortion but the ordered-set persistence modules are not ϵ–interleaved for any ϵ.

6 Persistence modules via strongly connected components and preorder homology

In this section we will consider constructions using an associated filtration of directed graphs or preorders. For each t we can define a directed graph $\{D(X)_t\}$ where $x \to y$ is included in $D(X)_t$ when $\max\{f(x, y), f(x, x), f(y, y)\} \le t$. From a directed graph we can induce a natural preorder via the existence of paths. That is a preorder where $x \le y$ if there is a path from x to y in $D(X)_t$. We will construct persistence modules using the strongly connected components of the graphs $D(X)_t$, which are also the equivalence classes of the associated preorders. We also consider persistence modules using ordered-tuple complexes constructed over preorders.

Let us first introduce the construction of directed graphs and preorders from set–function pairs.

Definition 22 Let X be a set with a binary relationship $\le$. Consider the following potential properties of $(X, \le)$:

  (i) **Reflexive** $x \le x$ for all $x \in X$.

 (ii) **Antisymmetric**   For all $x, y \in X$, if $x \le y$ and $y \le x$ then $x = y$.

(iii) **Transitive**   For all $x, y, z \in X$, if $x \le y$ and $y \le z$ then $x \le z$.

We say that $(X, \le)$ is a *poset* is it satisfies (i), (ii) and (iii). We say $(X, \le)$ is a *preorder* if it satisfies (i) and (iii).

There is a natural equivalence relation on X where $x \sim y$ when $x \le y$ and $y \le x$. If we quotient a preorder by this equivalence relation we are left with a poset.

One way to construct preorders is via directed graphs. Given a directed graph $G = (V, E)$ and vertices $x, y \in V$, we say there is a *path* from x to y when there is a finite sequence of vertices $x = x_0, x_1, \ldots, x_n = y$ such that (x_i, x_{i+1}) is a directed edge. To create a preorder on V we declare that $x \le y$ whenever there is a path from x to y. The strongly connected components of a directed graph are the equivalence classes of

points where $v \sim w$ when there exists both a path from v to w and a path from w to v. Thus, we see that the equivalence classes of this poset are precisely the strongly connected components of the directed graph it was built from. Suppose we start with a directed graph and we consider the preorder defined by the existence of paths. If we quotient by the equivalence relation to get a poset, then on the directed graph level we are collecting the vertices into the strongly connected components and then we have directed edges between these strongly connected components if there is a path between them. This will create an acyclic directed graph.

We will first need to construct directed graphs from the sublevel sets of a set–function pair. From this we can consider filtrations of directed graphs and of preorders.

Definition 23   Given a set–function pair (X, f) there is a natural filtration of directed graphs $\{\mathcal{D}(X)_t : t \in [0, \infty)\}$ associated to X by setting $\mathcal{D}(X, f)_t$ to the directed graph with vertices $\{x \in X : f(x, x) \leq t\}$ and including the directed edge $x \to y$ whenever $\max\{f(x, x), f(y, y), f(x, y)\} \leq t$. We will call this the *associated filtration of directed graphs* of (X, f).

It is necessary for the inclusion rule for the directed edges to occur at the maximum of $\{f(x, x), f(y, y), f(x, y)\}$ (rather than at $f(x, y)$, which may occur earlier) to ensure that $D(X, f)_t$ will satisfy the closure conditions for a directed graph. In the case where $f = d$ is a quasimetric, $d(x, x) = 0 = d(y, y)$ and $d(x, y) \geq 0$ and so the edge from x to y is included at $t = d(x, y)$.

We define a filtration of preorders to be a parametrised family of preorders

$$\{(X_t, \leq_t) : t \in \mathbb{R}\}$$

such that for all $s \leq t$ we have $X_s \subset X_t$ and if $x, y \in X_s$ with $x \leq_s y$ then $x \leq_t y$. From a filtration of associated graphs for a set–function pair we can construct a natural filtration of preordered spaces as follows:

Definition 24   Let (X, f) be a set–function pair and let $\{D(X, f)_t\}$ be its associated filtration of directed graphs. For each $t \geq 0$ construct a preordered space $(X_t, \leq_t)$ with X_t the set of points in $D(X, f)_t$ and $x \leq_t y$ when there exists a path in $D(X, f)_t$ from x to y. We call this the *associated filtration of preorders*.

The following is a useful lemma for proving the interleaving results for the persistence modules constructed with strongly connected components or with preorder homology:

Lemma 25 *Let X and Y be sets and (X, f) and (Y, g) be set–function pairs with $d_{\mathrm{CD}}((X, f), (Y, g))$ finite. Let $\mathcal{D}(X, f) = \{D(X, f)_t\}$ and $\mathcal{D}(Y, g) = \{D(Y, g)_t\}$ be the associated filtrations of directed graphs. Let $\mathcal{M} \subset X \times Y$ be a correspondence with $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ finite.*

(i) *If $(x, y) \in \mathcal{M}$ and $x \in D(X, f)_t$ then $y \in D(Y, g)_{t+\mathrm{dis}(\mathcal{M})}$.*

(ii) *If $(x_1, x_2), (y_1, y_2) \in \mathcal{M}$ and there exists a directed path from x_1 to x_2 in $D(X, f)_t$ then there exists a directed path from y_1 to y_2 in $D(Y, g)_{t+\mathrm{dis}(\mathcal{M})}$.*

Proof (i) If $x \in D(X, f)_t$ then $f(x, x) \leq t$. Since $(x, y) \in \mathcal{M}$, we know $g(y, y) \leq t + \mathrm{dis}(\mathcal{M})$ and hence $y \in D(Y, g)_{t+\mathrm{dis}(\mathcal{M})}$.

(ii) Suppose that there is a path from x_1 to x_2 in $D(X, f)_t$. This means that there exists a sequence of points $(x_1 = a_1, a_2, \ldots, a_k = x_2)$ in X such that $f(a_i, a_{i+1}) \leq t$. There exists a sequence of points in Y, $y_1 = b_1, b_2, \ldots, b_k = y_2$, where $(a_i, b_i) \in \mathcal{M}$. By (i) we know each of the b_i lie in $D(Y, g)_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$. Since each $(a_i, b_i) \in \mathcal{M}$, we have

$$|f(a_i, a_{i+1}) - g(b_i, b_{i+1})| \leq \mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$$

for each i and hence $(y_1 = b_1, b_2, \ldots, b_k = y_2)$ is a path in $D(Y, g)_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$. $\square$

The lemma can be rewritten in terms of preorders; for $(x_1, y_1), (x_2, y_2) \in \mathcal{M}$, if $x_1 \leq_t^f x_2$ then $y_1 \leq_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}^g y_2$.

6.1 Strongly connected components persistence

Dimension 0 persistent homology is all about tracking the evolution of connected components. For directed graphs, unlike graphs, there is choice in how to interpret what a connected component is, with each interpretation providing their own corresponding persistence module. Here we will consider the persistence of weakly and strongly connected components.

Weakly connected components are the components of the graph when the directions are forgotten. Given a filtration of a directed graph by edge weights, the weakly connected persistence would be the same as the dimension 0 persistent homology of the Rips filtration under sym_1 in Section 4, and to the dimension 0 directed Rips persistence module in Section 5.

Studying strongly connected components will provide new information. Recall the strongly connected components of a directed graph are the equivalence classes of points where $v \sim w$ when there exists both a path from v to w and a path from w to v. Given a filtration of directed graphs we can construct a persistence module based on linear combinations of strongly connected components (analogous to dimension 0 homology being interpreted as the space of formal linear combinations of connected components).

Definition 26 We call $\mathcal{D} = \{D_t : t \in \mathbb{R}\}$ a filtration of directed graphs if D_t is directed graph for all t such that if $s \leq t$ then D_s is a directed subgraph of D_t. Given a filtration of directed graphs $\mathcal{D} = \{D_t\}$, let $[v]_t$ denote the strongly connected component of D_t containing v. We define the *strongly connected persistence module* corresponding to $\mathcal{D}$ as follows:

- For each $t \in \mathbb{R}$ set the vector space V_t to be the vector space of finite linear combinations of strongly connected components $\big($that is, elements are of the form $\sum_{i=1}^{k} \lambda_i [v_i]_t$ with $\lambda_i \in \mathbb{F}\big)$.

- For each pair $t \leq s$ we have a linear map induced from inclusion,

$$\iota_{t \to s}\left(\sum_{i=1}^{k} \lambda_i [v_i]_t \right) = \sum_{i=1}^{k} \lambda_i [v_i]_s.$$

We will now check that the strongly connected component persistence module does satisfy the requirements of a persistence module. Whenever we have an inclusion of directed graphs $D_t \subset D_s$, whenever there is a path from v to w in D_t, there is also a path from v to w in D_s. This implies that the maps $\iota_{t \to s}$ are well defined. Furthermore, for $u \leq t \leq s$ we have $\iota_{t \to s}\big(\iota_{u \to t}\big(\sum_{i=1}^{k} \lambda_i [v_i]_u\big)\big) = \sum_{i=1}^{k} \lambda_i [v_i]_s = \iota_{u \to s}\big(\sum_{i=1}^{k} \lambda_i [v_i]_u\big)$. Whenever the directed graphs D_t are all finite (which is true in almost any application) we automatically know that the V_t are all finite-dimensional and hence the strongly connected persistence module is tame.

We can create strongly connected persistence modules from set–function pairs via its associated filtration of directed graphs.

Theorem 27 *Let X and Y be sets and (X, f) and (Y, g) be set–function pairs with $d_{\mathrm{CD}}((X, f), (Y, g))$ finite. Let $\mathcal{D}(X, f) = \{D(X, f)_t\}$ and $\mathcal{D}(Y, g) = \{D(Y, g)_t\}$ be the associated filtrations of directed graphs. Let $\mathcal{P}^X$ and $\mathcal{P}^Y$ be the strongly connected component persistence modules for $\mathcal{D}(X, f)$ and $\mathcal{D}(Y, g)$. Then $d_{\mathrm{int}}(\mathcal{P}^X, \mathcal{P}^Y) \leq d_{\mathrm{CD}}((X, f), (Y, g))$.*

Proof Fix a correspondence $\mathcal{M} \subset X \times Y$ with $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ finite.

Construct a map $\alpha \colon X \to Y$ where for each x we arbitrarily fix a representative from $\{y \in Y : (x, y) \in \mathcal{M}\}$, and construct a map $\beta \colon Y \to X$ where for each y we arbitrarily fix a representative from $\{x \in X : (x, y) \in \mathcal{M}\}$.

If $[x_1]_t = [x_2]_t$ then there exist paths in $D(X, f)_t$ from x_1 to x_2 and from x_2 to x_1. By Lemma 25 there exist paths in $D(Y, g)_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$ from $\alpha(x_1)$ to $\alpha(x_2)$ and from $\alpha(x_2)$ to $\alpha(x_1)$. This means that α induces a well-defined linear map

$$\alpha_* \colon P_t^X \to P_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}^Y, \quad [x]_t \mapsto [\alpha(x)]_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}.$$

Similarly, β induces a linear map $\beta_* \colon P_t^Y \to P_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}^X$ where $[y]_t \mapsto [\beta(y)]_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$.

It only remains to show that α_* and β_* satisfy an $2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ interleaving. That (3-1) commutes follows directly from the construction of α and β.

Let $f(x, x) = t$, whence $x \in D(X, f)_t$. From our construction of α and β we know that $(x, \alpha(x))$ and $(\beta(\alpha(x)), \alpha(x))$ are both in $\mathcal{M}$. By Lemma 25 this implies that there are directed paths in both directions between $\beta(\alpha(x))$ and x in the directed graph $D(X, f)_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$, and hence they lie in the same strongly connected component.

Similarly, for every $y \in Y$ with $g(y, y) = t$, we know that $\alpha(\beta(y))$ and y lie in the same strongly connected component in $D(Y, g)_{t+2\,\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$. This ensures that we satisfy (3-2).

By taking the infimum over all correspondences we see that the interleaving distance between P^X and P^Y is bounded above by $2d_{\mathrm{CD}}((X, f), (Y, g))$.   □

We provide some pseudocode (the algorithm in the appendix) for an algorithm that computes the interval decomposition of the strongly connected component persistence module from a filtration of directed graphs. It is a modification of the union-find algorithm used to compute the standard dimension 0 persistent homology. In the union-find algorithm each connected component is represented by a root vertex with an additional data of its birth time. The main difference for strongly connected components is that we have to also keep track of when directed paths exist between the various strongly connected components. These are stored as a list of the root vertices of "in" and "out" connected components. Here "in" means a connected component that has a path pointing into the current component and "out" means there is a path pointing

out of the current component. Note that for any root vertex these in and out sets are disjoint, as being in both would imply they are the same strongly connected component. The main challenge in this modification is to ensure that at each stage the list of in and out strongly connected components listed by the root vertices are referred to by their root vertex.

6.2 OT complexes constructed using the preorder structure

In the theory of partially ordered sets ("posets"), the order complex of a poset is the set of all finite chains. Its homology contains important information about the poset. Preorders are a generalisation of posets where we drop the antisymmetry condition. Poset homology naturally extends to preorders, where we will call it preorder homology. It is easier and more flexible to construct filtrations of preorders than of posets.

From the associated filtration of directed graphs of a set–function pair we can create a filtration of preorders which we will call the preorder Rips filtration. From the filtration of preorders we can construct persistence modules using preorder homology to generate preorder Rips persistence modules. These persistence modules enjoys stability with respect to the correspondence distortion distance. The homology dimension 0 preorder Rips persistence module is isomorphic to that of its weakly connected components, its directed Rips persistence module and the standard Rips persistence module under sym_1. If the input is a symmetric function then its higher-dimensional preorder Rips persistence modules are all trivial, showing that preorder Rips persistence module describes asymmetry information.

In this paper we will generalise to preorders some constructions normally defined for posets. The homology of a poset has been defined and studied via its corresponding Alexandrov topology. Preorders are in bijective correspondence with Alexandrov topologies, with the antisymmetry condition (which is the axiom that makes a preorder a poset) translating to those topologies that are T_0. For each preorder there is a canonical poset over its equivalence classes, and this poset corresponds to the Kolmogorov quotient of the Alexandrov topology of that original preorder. Because these quotient spaces are weakly homotopy equivalent, standard references for Alexandrov topology often state they will restrict their analysis to T_0 spaces/posets (eg [16; 15]). It is for this reason that definitions are usually only stated for posets and not more generally for preorders. In the appendix we will go into more detail into this background material and justify why the definitions given in this section are the natural generalisation of those traditionally given for posets.

Let us now construct an OT complex from a preorder.

Definition 28 Given a preorder $(X, \leq)$, let $\mathcal{O}(X, \leq)$ be the OT complex containing $(x_0, x_1, \ldots, x_p)$ when $x_0 \leq x_1 \leq \cdots \leq x_p$. We call $\mathcal{O}(X, \leq)$ the *preorder OT complex* of $(X, \leq)$.

Definition 29 Given a preorder $(X, \leq)$, its associated order complex $\Delta(X, \leq)$ is an abstract simplicial complex whose vertices are the elements of X and whose faces are the chains (subsets where each pair is comparable) of $(X, \leq)$.

From a filtration of preorders we can construct a filtration of OT complexes. From this, persistence modules can be constructed as standard with OT homology classes as the vector spaces and induced maps from inclusions as the transition maps.

Definition 30 Let $\mathcal{O}(X, f) = \{\mathcal{O}(X, f)_t\}$ be the filtration of OT complexes corresponding to the filtration of posets $\{(X_t, \leq_t)\}$. We call $\mathcal{O}(X, f)$ the *preorder filtration* of (X, f).

In the appendix we see that the simplicial homology of the order complex $\Delta(X, \leq)$ is naturally isomorphic to the homology of the preorder OT complex $\mathcal{O}(X, \leq)$. Moreover, isomorphisms between the simplicial homology of the order complexes and the homology of the preorder OT complexes will extend to persistent homology as they commute with the maps on homology induced by inclusions.

Definition 31 We define the kth–dimensional *preorder persistence module* corresponding to the filtration of preorders $\mathcal{X} = \{(X_t, \leq_t)\}$ as the dimension k OT homology persistence module for the filtration of OT complexes $\{\mathcal{O}(X, \leq_t)\}_{t \in \mathbb{R}}$.

Just as in the previous constructions in this paper we can prove that the corresponding persistence modules built from functions $f \colon X \times X \to \mathbb{R}$ and $g \colon Y \times Y \to \mathbb{R}$ are stable with respect to the correspondence distortion distance.

Theorem 32 *Let (X, f) and (Y, g) be set–function pairs with preorder Rips filtrations $\mathcal{O}(X, f)$ and $\mathcal{O}(Y, g)$. Let $\mathcal{P}^X$ and $\mathcal{P}^Y$ be the kth–dimensional persistence modules for $\mathcal{O}(X, f)$ and $\mathcal{O}(Y, g)$, respectively. Then $d_{\text{int}}(\mathcal{P}^X, \mathcal{P}^Y) \leq 2d_{\text{CD}}((X, f), (Y, g))$.*

Proof Since $d_{\mathrm{CD}}((X, f), (Y, g))$ is finite, there exists some correspondence $\mathcal{M} \subset X \times Y$ with $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ finite. Fix a correspondence $\mathcal{M} \subset X \times Y$ with $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ finite. From this correspondence construct preorder Rips filtrations $\{\mathcal{O}(\mathcal{M}, f^{\mathcal{M}})_t\}$ and $\{\mathcal{O}(\mathcal{M}, g^{\mathcal{M}})_t\}$ with corresponding k^{th}–dimensional persistence modules $P^{(X,\mathcal{M})}$ and $P^{(Y,\mathcal{M})}$.

By construction, $\{\mathcal{O}(\mathcal{M}, f^{\mathcal{M}})_t\}$ is an expansion of $\{\mathcal{O}(X, f)_t\}$ and thus by Proposition 20 we know that the persistence modules P^X and $P^{(X,\mathcal{M})}$ are isomorphic. Similarly we can also show that P^Y and $P^{(Y,\mathcal{M})}$ are isomorphic.

If $((x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)) \in \mathcal{O}(\mathcal{M}, f^{\mathcal{M}})_t$ then $x_0, \ldots, x_n \in D(X)_t$ and there exist directed paths from x_i to x_j in $D(X, f)_t$ for all $i \leq j$. By Lemma 25 there must exist a directed path from y_i to y_j in $D(Y, g)_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$ for all $i \leq j$. This implies that $\mathcal{O}(\mathcal{M}, f^{\mathcal{M}})_t \subset \mathcal{O}(\mathcal{M}, g^{\mathcal{M}})_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$ for all t. Similarly, $\mathcal{O}(\mathcal{M}, g^{\mathcal{M}})_t \subset \mathcal{O}(\mathcal{M}, f^{\mathcal{M}})_{t+\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})}$.

These inclusion maps induce a $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$ interleaving between $P^{(X,\mathcal{M})}$ and $P^{(Y,\mathcal{M})}$. This implies that P^X and P^Y are $\mathrm{dis}_{(X,f),(Y,g)}(\mathcal{M})$–interleaved.

By considering the infimum of the interleavings constructed by correspondences we see that the interleaving distance between P^X and P^Y is bounded above by $2d_{\mathrm{CD}}((X, f), (Y, g))$. $\qquad\square$

As shown in Theorem 36 (in the appendix), the simplicial homology of the order complex is naturally isomorphic to the OT homology of $\mathcal{O}(X, \leq)$. Furthermore, this isomorphism result holds also for persistence modules of filtrations of simplicial complexes as the isomorphisms on homology groups commute with the induced maps on homology by inclusions. This implies that interval decomposition of the kth preorder persistence modules can be computed via the simplicial persistent homology over the filtration of simplicial complexes $\{\Delta(X_t, \leq_t^f)\}$.

7 Future directions

There are many future directions related to the research in this paper. Examples include:

- Applying the constructions in this paper to quasimetric spaces to see what they reveal about their quasimetric structure, or to use as a method of getting a lower bound on the correspondence distortion distance between different quasimetric spaces.

- Adapting these methods to construct persistence modules for sublevel set filtrations of special functions on quasimetric spaces and proving related stability results. For example, we conjecture that all four constructions built from a suitably defined sublevel set of the extremity function of a quasimetric space (analogous to constructions in [4]) could have correspondence distortion distance stability with respect to the original quasimetric distance functions. This would provide another way of capturing the "shape" of a quasimetric space.

- Finding nice sufficient conditions on functions $f \colon X \times X \to \mathbb{R}$, with $|X|$ infinite, as to when these various Rips constructions create tame persistence modules. Even when restricting to the case of quasimetric spaces it is not even clear how we should define an ϵ–sampling or compactness. In the symmetric case, definitions have been used to describe sufficient conditions for metric spaces that result in tame persistence modules (such as in [5]).

- Algorithmic techniques for computing OT persistent homology efficiently. In particular, is there a related filtration of simplicial complexes that have isomorphic OT persistent homology, at least in low homology dimensions?

Appendix

A.1 Algorithm to compute interval decomposition of the strongly connected persistence module

INPUT: List L of vertices $V = \{v_1, v_2, \ldots, v_n\}$ and directed edges

$$\{(v_{i_1} \to v_{j_1}), \ldots, (v_{i_m} \to v_{j_m})\},$$

each with a real-valued height such that $h(v_i \to v_j) \geq \max\{h(v_i), h(v_j)\}$. These vertices and directed edges are ordered in a combined list L by increasing height values. All the vertices at a height value occur before the edges at that same height.

OUTPUT: Interval decomposition of the strongly connected component persistence module from filtration of sublevel sets of the height function

```
1:  function FIND(x)
2:      while root(x) != x do
3:          x = root(x)
4:      end while
5:      return x
6:  end function
```

7: **procedure** UNION(v_{tail}, v_{head}, height)

8: $\quad W = v_{\text{tail}}.\text{in} \cap v_{\text{head}}.\text{out}$

9: $\quad \widehat{w} :=$ earliest $w \in W$ to appear in list L

10: $\quad$ **for** $w \in W$, $w \neq \widehat{w}$ **do**

11: $\quad\quad$ root$(w) = \widehat{w}$

12: $\quad\quad$ **if** $h(w) <$ height **then**

13: $\quad\quad\quad$ append $[h(w), \text{height})$ to BARCODE

14: $\quad\quad$ **end if**

15: $\quad\quad \widehat{w}.\text{in} = \{\text{FIND}(x) \text{ for } x \in v_{\text{tail}}.\text{in}\}$ $\quad \triangleright$ An SCC has a path to $\widehat{w}$ if and only if it has a path to v_{tail}.

16: $\quad\quad \widehat{w}.\text{out} = \{\text{FIND}(x) \text{ for } x \in v_{\text{head}}.\text{out}\}$ $\quad \triangleright$ An SCC has a path from $\widehat{w}$ if and only if it has a path from v_{head}.

17: $\quad\quad$ **for** $x \in \widehat{w}.\text{in}$ **do**

18: $\quad\quad\quad x.\text{out} = \{\text{FIND(y) for } y \in x.\text{out} \cup \widehat{w}.\text{out}\}$

19: $\quad\quad$ **end for**

20: $\quad\quad$ **for** $x \in \widehat{w}.\text{out}$ **do**

21: $\quad\quad\quad x.\text{in} = \{\text{FIND(y) for } y \in x.\text{in} \cup \widehat{w}.\text{in}\}$

22: $\quad\quad$ **end for**

23: $\quad$ **end for**

24: **end procedure**

25:

26: **procedure** UPDATEINOUT(v_{tail}, v_{head}, height)

27: $\quad$ **for** $x \in v_{\text{tail}}.\text{in}$ **do**

28: $\quad\quad x.\text{out} = \{\text{FIND(y) for } y \in v_{\text{head}}.\text{out} \cup x.\text{out}\}$

29: $\quad$ **end for**

30: $\quad$ **for** $y \in v_{\text{head}}.\text{out}$ **do**

31: $\quad\quad y.\text{in} = \{\text{FIND(x) for } x \in v_{\text{tail}}.\text{in} \cup y.\text{in}\}$

32: $\quad$ **end for**

33: **end procedure**

34:

35: **for** $i = 1$ to length(L) **do**

36: $\quad$ **if** $L(i)$ is a vertex v_k **then**

37: $\quad\quad$ Add a vertex to A. Label it with (height $= h(L(i))$, root $= v_k$, in $= \{v_k\}$, out $= \{v_k\}$).

38:        **if** $L(i)$ is a directed edge $v_j \to v_k$ **then**

39:              $v_{\text{tail}} := \text{FIND}(v_j)$, $v_{\text{head}} := \text{FIND}(v_k)$

40:            **if** $v_{\text{head}} \notin v_{\text{tail}}.\text{out}$ **then**

41:                **if** $v_{\text{head}} \notin v_{\text{tail}}.\text{in}$ **then**   ▷ We need to update the paths between SCCs.

42:                    $\text{UPDATEINOUT}(v_{\text{tail}}, v_{\text{head}})$

43:                **end if**

44:                **if** $v_{\text{head}} \in v_{\text{tail}}.\text{in}$ **then**   ▷ This is when various SCCs merge.

45:                    $\text{UNION}(v_{\text{tail}}, v_{\text{head}}, h(L(i)))$

46:                **end if**

47:            **end if**

48:          **end if**

49:      **end if**

50: **end for**

51: RemainingComponents := $\{\text{FIND}(x) \text{ for } x \in V\}$ ▷ Final set of strongly connected components.

52: **for** $x \in$ RemainingComponents **do**

53:      Append $[h(x), \infty)$ to BARCODE

54: **end for**

A.2 Homology of a poset

There are multiple ways to compute the homology of a poset, including via Alexandrov topological spaces and order simplicial complexes. For each preorder there is a canonical poset we call its equivalence class poset. In this subsection we show that the definitions of homology of a poset can naturally be extended to preorders. Furthermore, the resulting homology of a preorder is naturally isomorphic to the homology of its equivalence class poset. This justifies the constructions in Section 6.2.

An *Alexandrov topology* is a topology in which the intersection of any family of open sets is open. It is an axiom of topology that the intersection of any finite family of open sets is open; in Alexandrov topologies the finiteness restriction is dropped. Given an Alexandrov topology we can construct a special preorder, called its specialisation preorder.

Definition 33 Let $X = (X, \tau)$ be an Alexandrov space. The *specialisation preorder* on X is the preorder where $x \leq y$ if and only if x is in the closure of $\{y\}$.

In the other direction, given a preorder $(X, \leq)$ there is a unique Alexandrov topology whose specialisation preorder is $(X, \leq)$. To construct this, let the open sets τ on X be the upper sets,

$$\tau = \{U \subset X : \forall x, y \in X \text{ if } x \leq y \text{ and } x \in U \text{ then } y \in U\}.$$

The corresponding closed sets for τ are the lower sets,

$$\{S \subseteq X : \forall x, y \in X \text{ if } x \in S \text{ and } y \leq x \text{ then } y \in S\}.$$

The topology τ is generated by the sets $U_x = \{y : x \leq y\}$.

A topological space X is a T_0 *space* if for any pair of points in X there exists an open set containing one and only one of them. It is an exercise to see how the antisymmetry condition of posets directly corresponds to the Alexandrov topologies that are T_0.

We can construct T_0 spaces by taking Kolmogorov quotients. The *Kolmogorov quotient* of a topological space is defined as its quotient by the equivalence relation of topological indistinguishability, equipped with the quotient topology.

There is a natural way of constructing a poset from a preorder by using quotients. For $(X, \leq)$ a preorder, define an equivalence relation $x \sim y$ when $x \leq y$ and $y \leq x$. Let $\widetilde{X} = X/\sim$ be the quotient space on these equivalence classes. It is easy to check that the binary relation $\leq$ is now well defined on $\widetilde{X}$ and that $(\widetilde{X}, \leq)$ is a poset. We will call $(\widetilde{X}, \leq)$ the *equivalence class poset* of $(X, \leq)$. The following lemma states the relationship between a preorder and its equivalence class poset is analogous to taking the Kolmogorov quotient of its Alexandrov topology. The proof for finite spaces is Lemmas 8 and 9 in [16], and the extension to infinite spaces can be proved similarly (see [13]).

Lemma 34 *Let $(X, \leq)$ be a finite preorder with equivalence class poset $(\widetilde{X}, \leq)$. The Alexandrov topology of $(\widetilde{X}, \leq)$ is the Kolmogorov quotient of the Alexandrov topology of $(X, \leq)$. Furthermore, the Alexandrov topologies of $(X, \leq)$ and $(\widetilde{X}, \leq)$ are homotopy equivalent.*

Since homology is defined up to weak homotopy equivalence, often in analysis researchers restrict their analysis from general topological spaces to T_0 spaces as they do not lose any homological information by taking the Kolmogorov quotient. Thus, many definitions are stated as for posets even though they could be defined for all preorders.

One definition of the homology of a poset is the singular homology of the Alexandrov topology which has that poset as its specialisation order. Since the specialisation orders of Alexandrov topologies provide a one-to-one correspondence between Alexandrov topologies and preorders, we can generalise this to define the homology of a preorder as the singular homology of the Alexandrov topology which has that preorder as its specialisation order.

A *chain* in a poset is defined as a subset of elements which are all pairwise comparable. Note that there is no order of the elements given as part of the information of the chain but that the transitivity of a preorder will ensure that there exists a total ordering of any chain. In a poset the antisymmetry condition ensures that this order is unique. In a general preorder multiple possible orders might be possible.

In a poset, the unique ordering of elements in a chain means we can define chain complexes and homology groups for a poset directly via chains. We thus say that an m−*chain* of a poset P is a totally ordered subset $c = (x_0 < x_1 < \cdots < x_m)$ of P written in order. We can construct a chain complex by setting $C_j(P, R)$ to be the R−module freely generated by j−chains, and defining boundary maps $\partial_j \colon C_j(P) \to C_{j-1}(P)$ by $\partial_j(x_0 < x_1 < \cdots < x_m) = \sum_{i=0}^{m}(-1)^i(x_0 < x_1 < \cdots \hat{x}_i \cdots < x_m)$ and extending linearly.

We can observe that this chain complex is exactly that for ordered sets (see Section 5.2). If we specify the order of each chain, we can extend this definition to preorders as the OS homology. Generally the OS homology and the OT homology are not isomorphic (see Section 5.2). However, in the special case of posets they do define isomorphic homology groups, as proved below in Theorem 35.

An alternative definition for the homology of a poset is via the construction of its associated order simplex. The associated order complex $\Delta(X, \leq)$ for the poset $(X, \leq)$ is the abstract simplicial complex whose vertices are the elements of X and whose faces are the chains (subsets where each pair is comparable) of $(X, \leq)$. The definition of the associated order complex of a preorder given in Section 6.2 restricts to the standard definition for posets.

The following theorem presents some relationships between these different homology constructions:

Theorem 35 *Let $(\widetilde{X}, \leq)$ be a poset. The following homology groups are isomorphic:*

   (i)   *OS homology of the finite chains of $(\widetilde{X}, \leq)$.*

(ii) *OT homology of the preorder OT complex $\mathcal{O}(\widetilde{X}, \leq)$.*

(iii) *Simplicial homology of the order complex $\Delta(\widetilde{X}, \leq)$.*

(iv) *Singular homology of the Alexandrov topology with specialisation order $(\widetilde{X}, \leq)$.*

Proof The proof that (ii) and (iv) are isomorphic is in [16, Theorem 2]. The isomorphism between (i) and (iii) is via the unique total orderings of each simplex in the order complex. It is the induced map on homology of $(x_0 < x_1 < \cdots < x_k) \mapsto [x_0, x_1, \ldots, x_k]$.

We will now prove that (i) is isomorphic to (ii). The set of ordered tuples forms a basis B for $\mathrm{OT}(\widetilde{X}, \leq)$. Define $\Phi\colon B \to \{\text{subcomplexes of } \mathrm{OT}(\widetilde{X}, \leq)\}$ by setting $\Phi(\tau)$ to be the subcomplex of $\mathrm{OT}(\widetilde{X}, \leq)$ containing only ordered tuples with elements within τ. Since τ is an ordered tuple, it has a smallest element x and for any $\alpha \in \Phi(\tau)$ the ordered tuple concatenating x in front of α (which we will denote by $(x\alpha)$) is also an element in $\Phi(\tau)$. Given a boundary α, we can see that $\partial(x\alpha) = \alpha - (x\partial(\alpha)) = \alpha$. This implies that Φ is an acyclic carrier.

Set $f\colon \mathrm{OT}(\widetilde{X}, \leq) \to \mathrm{OT}(\widetilde{X}, \leq)$ by $f(\tau)$ the identity when τ does not contain repeats (ie lives in $OS(\widetilde{X}, \leq)$ and $f(\tau) = 0$ if τ contains a repeat). Then f commutes with the boundary map because all repeats of a particular element in a tuple must be consecutive when working with posets. It is this claim that does not hold more generally between OT complexes and OS complexes. Since both f and the identity map are both carried by Φ, the acyclic carrier theorem (see [17]) ensures that f and the identity map are chain homotopic and hence the OS homology of the finite chains of $(\widetilde{X}, \leq)$ and the OT homology of preorder OT complex $\mathcal{O}(\widetilde{X}, \leq)$ are isomorphic. $\qquad\square$

Each of these four different constructions of homology groups for posets can be generalised to preorders. Three of these generalise in a way that the homology groups are invariant under taking equivalence class posets (or equivalently under taking Kolmogorov quotients). The OS homology of chains is the odd one out in this respect. A counterexample is the preorder $X = \{x, y\}$ with $x \leq y$ and $y \leq x$. It has nontrivial OS homology in dimension one but its equivalence class poset $\widetilde{X} = \{[x]\}$ has trivial OS homology in dimension one.

Theorem 36 *Let $(X, \leq)$ be a preorder with equivalence class poset $(\widetilde{X}, \leq)$. Then:*

(a) *The preorder OT complex $\mathcal{O}(X, \leq)$ is an expansion of $\mathcal{O}(\widetilde{X}, \leq)$ and hence has the same OT homology.*

(b)   *There is a natural projection map from $\Delta(X, \leq)$ to $\Delta(\widetilde{X}, \leq)$. This projection map induces an isomorphism on their simplicial homology groups.*

(c)   *The singular homology of the Alexandrov topology with specialisation order $(X, \leq)$ is isomorphic to the singular homology of the Alexandrov topology with specialisation order $(\widetilde{X}, \leq)$.*

Proof   (a)   The OT complexes $\mathcal{O}(X, \leq)$ and $\mathcal{O}(\widetilde{X}, \leq)$ are closed under adjacent repeats by construction. The quotient map sending X to its equivalence class poset $\widetilde{X}$ shows that $\mathcal{O}(X, \leq)$ is an expansion of $\mathcal{O}(\widetilde{X}, \leq)$. We conclude that they are isomorphic by applying Proposition 20.

(b)   Construct a map $f \colon \widetilde{X} \to X$ by fixing a representative $x \in X$ for each equivalence class $[x] \in \widetilde{X}$. We can embed $\Delta(\widetilde{X}, \leq)$ into $\Delta(X, \leq)$ via the induced map of f. A straight line homotopy provides a deformation from $\Delta(X, \leq)$ to $f(\Delta(\widetilde{X}, \leq))$. The result then follows because deformation retractions induce isomorphisms on homology classes.

(c)   The proof follows from Lemma 34 as homotopic topological spaces have isomorphic singular homology groups. $\square$

Combining these theorems we conclude that the OT homology of preorder OT complexes, simplicial homology of the associated order complex of a preorder and the singular homology of the Alexandrov topology of a preorder are all isomorphic. These isomorphisms extend to persistent homology as they commute with the maps on homology induced by inclusions.

References

[1]   **D Burago**, **Y Burago**, **S Ivanov**, *A course in metric geometry*, Graduate Studies in Mathematics 33, Amer. Math. Soc., Providence, RI (2001) MR

[2]   **G Carlsson**, *Topology and data*, Bull. Amer. Math. Soc. 46 (2009) 255–308 MR

[3]   **F Chazal**, **D Cohen-Steiner**, **M Glisse**, **L J Guibas**, **S Y Oudot**, *Proximity of persistence modules and their diagrams*, from "Proceedings of the twenty-fifth annual symposium on computational geometry", ACM, New York (2009) 237–246

[4]   **F Chazal**, **D Cohen-Steiner**, **L J Guibas**, **F Mémoli**, **S Y Oudot**, *Gromov–Hausdorff stable signatures for shapes using persistence*, Computer Graphics Forum 28 (2009) 1393–1403

[5] **F Chazal**, **V de Silva**, **S Oudot**, *Persistence stability for geometric complexes*, Geom. Dedicata 173 (2014) 193–214 MR

[6] **S Chowdhury**, **F Mémoli**, *A functorial Dowker theorem and persistent homology of asymmetric networks*, preprint (2016) arXiv

[7] **S Chowdhury**, **F Mémoli**, *Convergence of hierarchical clustering and persistent homology methods on directed networks*, preprint (2017) arXiv

[8] **S Chowdhury**, **F Mémoli**, *Persistent path homology of directed networks*, from "Proceedings of the twenty-ninth annual ACM–SIAM symposium on discrete algorithms" (A Czumaj, editor), SIAM, Philadelphia (2018) 1152–1169 MR

[9] **H Edelsbrunner**, **J Harer**, *Persistent homology: a survey*, from "Surveys on discrete and computational geometry" (J E Goodman, J Pach, R Pollack, editors), Contemp. Math. 453, Amer. Math. Soc., Providence, RI (2008) 257–282 MR

[10] **R Ghrist**, *Barcodes: the persistent topology of data*, Bull. Amer. Math. Soc. 45 (2008) 61–75 MR

[11] **M Kaiser**, *A tutorial in connectome analysis: topological and spatial features of brain networks*, Neuroimage 57 (2011) 892–907

[12] **S Klamt**, **A von Kamp**, *Computing paths and cycles in biological interaction graphs*, BMC Bioinformatics 10 (2009) art. id. 181

[13] **M J Kukieła**, *On homotopy types of Alexandroff spaces*, Order 27 (2010) 9–21 MR

[14] **H Lee**, **M K Chung**, **H Kang**, **B-N Kim**, **D S Lee**, *Discriminative persistent homology of brain networks*, from "2011 IEEE international symposium on biomedical imaging: from nano to macro" (S Wright, X Pan, M Liebling, editors), IEEE, Piscataway, NJ (2011) 841–844

[15] **J P May**, *Finite spaces and simplicial complexes*, lecture notes (2003) Available at http://www.math.uchicago.edu/~may/MISCMaster.html

[16] **M C McCord**, *Singular homology groups and homotopy groups of finite topological spaces*, Duke Math. J. 33 (1966) 465–474 MR

[17] **J R Munkres**, *Elements of algebraic topology*, volume 2, Addison-Wesley, Menlo Park, CA (1984) MR

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

[19] **S Weinberger**, *What is… persistent homology?*, Notices Amer. Math. Soc. 58 (2011) 36–39 MR

[20] **K Xia**, **G-W Wei**, *Persistent homology analysis of protein structure, flexibility, and folding*, Int. J. Numer. Methods Biomed. Eng. 30 (2014) 814–844 MR

[21]   **A Zomorodian**, **G Carlsson**, *Computing persistent homology*, Discrete Comput. Geom. 33 (2005) 249–274   MR

Mathematical Sciences Institute, Australian National University
Canberra, ACT, Australia

katharine.turner@anu.edu.au