



# A functorial Dowker theorem and persistent homology of asymmetric networks

Samir Chowdhury<sup>1</sup> · Facundo Mémoli<sup>1,2</sup>

Received: 4 January 2018 / Accepted: 4 September 2018  
 © Springer Nature Switzerland AG 2018

## Abstract

We study two methods for computing network features with topological underpinnings: the Rips and Dowker persistent homology diagrams. Our formulations work for general networks, which may be asymmetric and may have any real number as an edge weight. We study the sensitivity of Dowker persistence diagrams to asymmetry via numerous theoretical examples, including a family of highly asymmetric cycle networks that have interesting connections to the existing literature. In particular, we characterize the Dowker persistence diagrams arising from asymmetric cycle networks. We investigate the stability properties of both the Dowker and Rips persistence diagrams, and use these observations to run a classification task on a dataset comprising simulated hippocampal networks. Our theoretical and experimental results suggest that Dowker persistence diagrams are particularly suitable for studying asymmetric networks. As a stepping stone for our constructions, we prove a functorial generalization of a theorem of Dowker, after whom our constructions are named.

**Keywords** Directed networks · Functorial Dowker theorem · Cycle networks · Functorial Nerve Theorem

**Mathematics Subject Classification** 55U99 · 68U05 · 55N35

## Contents

1	<a href="#">Introduction</a>	.....
1.1	<a href="#">Contributions and an overview of our approach</a>	.....
1.2	<a href="#">Implementations</a>	.....

---

✉ Facundo Mémoli  
 memoli@math.osu.edu

Samir Chowdhury  
 chowdhury.57@osu.edu

<sup>1</sup> Department of Mathematics, The Ohio State University, Columbus, OH, USA

<sup>2</sup> Department of Computer Science and Engineering, The Ohio State University, Columbus, OH, USA

1.3	Organization of the paper	.....
1.4	Notation	.....
2	Background on persistent homology	.....
2.1	Interleaving distance and stability of persistence vector spaces	.....
3	Background on networks and our network distance	.....
4	The Rips filtration of a network	.....
5	The Dowker filtration of a network	.....
5.1	The functorial Dowker theorem and equivalence of diagrams	.....
5.2	The equivalence between the finite FDT and the simplicial FNTs	.....
	Theorem 24 implies Theorem 25	.....
	Theorem 26 implies Theorem 24	.....
	Theorem 25 implies Theorem 26	.....
6	Dowker persistence diagrams and asymmetry	.....
6.1	Cycle networks	.....
6.2	Sensitivity to network transformations	.....
7	Implementation and an experiment on network classification	.....
7.1	Simulated hippocampal networks	.....
8	Discussion	.....
	Appendix A: Proofs	.....
	Appendix B: Higher dimensional Dowker persistence diagrams of cycle networks	.....
	References	.....

## 1 Introduction

Networks are used throughout the sciences for representing the complex relations that exist between the objects of a dataset (Newman 2003; Easley and Kleinberg 2010). Network data arises from applications in social science (Kumar et al. 2010; Easley and Kleinberg 2010), commerce and economy (Elliott et al. 2014; Easley and Kleinberg 2010; Acemoglu et al. 2015), neuroscience (Sporns 2011, 2012; Sporns and Kötter 2004; Rubinov and Sporns 2010; Pessoa 2014), biology (Barabási and Oltvai 2004; Huson et al. 2010), and defence (Masys 2014), to name a few sources. Networks are often directed, in the sense that weights attached to edges do not satisfy any symmetry property, and this asymmetry often precludes the applicability of many standard methods for data analysis.

Network analysis problems come in a wide range of flavors. One problem is in *exploratory data analysis*: given a network representing a dataset of societal, economic, or scientific value, the goal is to obtain insights that are meaningful to the interested party and that can help uncover interesting phenomena. Another problem is *network classification*: given a “bag” of networks representing multiple instances of different phenomena, one wants to obtain a clustering which groups the networks together according to the different phenomena they represent.

Because networks are often too complex to deal with directly, one typically extracts certain invariants of networks, and infers structural properties of the networks from properties of these invariants. While there are numerous such network invariants in the existing literature, there is growing interest in adopting a particular invariant arising from *persistent homology* (Frosini 1992; Robins 1999; Edelsbrunner et al. 2002; Zomorodian and Carlsson 2005), known as a *persistence diagram*, to the setting of networks. Persistence diagrams are used in the context of finite metric space or point cloud data to pick out *features of significance* while rejecting random noise (Edels-

brunner et al. 2002; Carlsson 2009). Since a network on  $n$  nodes is regarded, in the most general setting, as an  $n \times n$  matrix of real numbers, i.e. as a generalized metric space, it is conceivable that one should be able to describe persistence diagrams for networks as well.

The motivation for computing persistence diagrams of networks is at least two-fold: (1) comparing persistence diagrams has been shown to be a viable method for *shape matching* applications (Frosini 1992; Frosini and Landi 1999; Collins et al. 2004; Cohen-Steiner et al. 2007; Carlsson et al. 2005; Chazal et al. 2009), analogous to the network classification problem described above, and (2) persistence diagrams have been successfully applied to feature detection, e.g. in detecting the structure of protein molecules (see Krishnamoorthy et al. 2007; Xia and Wei 2014 and Edelsbrunner et al. 2002, §6) and solid materials (see Hiraoka et al. 2016) and might thus be a useful tool for exploratory analysis of network datasets.

We point the reader to Ghrist (2008), Edelsbrunner and Harer (2008), Carlsson (2009), Carlsson and Silva (2010), Weinberger (2011), Burghelca and Dey (2013) and Dey et al. (2014) for surveys of persistent homology and its applications, and some recent extensions.

Some extant approaches that obtain persistence diagrams from networks assume that the underlying network data actually satisfies metric properties (Lee et al. 2011; Khalid et al. 2014). A more general approach for obtaining persistence diagrams from networks is followed in Horak et al. (2009), Carstens and Horadam (2013), Giusti et al. (2015) and Petri et al. (2013), albeit with the restriction that the input data sets are required to be symmetric matrices.

Our chief goal is to devise notions of persistent homology that are directly applicable to asymmetric networks in the most general sense, and are furthermore capable of absorbing structural information contained in the asymmetry.

## 1.1 Contributions and an overview of our approach

In this paper, we study two types of persistence diagrams: the *Rips* and *Dowker* diagrams. We define both invariants in the setting of asymmetric networks with real-valued weights, without assuming any metric properties at all (not symmetry and not even that the matrix representing the networks weights vanishes on the diagonal). As a key step in defining the Dowker persistence diagram, we first define two dual constructions, each of which can be referred to as a Dowker persistence diagram, and then prove a *functorial Dowker theorem* which implies that these two possible diagrams are equivalent. Following the line of work in Chazal et al. (2009), where stability of Rips persistence diagrams arising from finite metric spaces was first established, we formulate similar stability results for the Rips and Dowker persistence diagrams of a network. Through various examples, in particular a family of *cycle networks*, we espouse the idea that Dowker persistence diagrams are more appropriate than Rips persistence diagrams for studying asymmetric networks. We test our methods by solving a network classification problem on a database of simulated hippocampal networks.

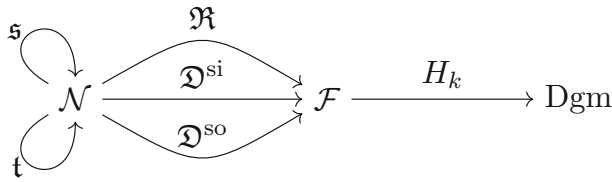
The first step in constructing a persistence diagram from a network is to construct a nested sequence of simplicial complexes, i.e. a simplicial filtration, which, in our

work, will be the *Rips* or *Dowker* filtrations associated to a network. Rips and Dowker simplicial complexes and their associated filtrations are classically defined for metric spaces (De Silva and Carlsson 2004; Ghrist 2014), and the generalization to networks that we use is a natural extension of the metric versions. After producing the simplicial filtrations, the standard framework of *persistent homology* takes over, and we obtain the Rips or Dowker persistence diagrams.

Practitioners of persistent homology might recall that there are *two* Dowker complexes (Ghrist 2014, p. 73), which we describe as the *source* and *sink* Dowker complexes. A subtle point to note here is that each of these Dowker complexes can be used to construct a persistence diagram. A folklore result in the literature about persistent homology of metric spaces, known as *Dowker duality*, is that the two persistence diagrams arising this way are equal (Chazal et al. 2014, Remark 4.8). In this paper we prove a stronger result—a functorial Dowker theorem—from which the duality follows easily. Furthermore, the context of this result is strictly more general than that of metric spaces (see below for a more thorough description of the functorial version of Dowker's theorem).

Providing a construction of Rips and Dowker persistence diagrams is not enough: in order for these invariants to be useful in practice, one must verify that the diagrams are *stable*. In this context, stability means the following: the dissimilarity between two Rips (resp. Dowker) persistence diagrams obtained from two networks should be bounded above by a function of the dissimilarity between the two networks. To our knowledge, stability is not addressed in the existing literature on producing persistence diagrams from networks. In our work, we provide stability results for both the Rips and Dowker persistence diagrams (Propositions 12 and 15). One key ingredient in our proof of this result is a notion of *network distance* that follows previous work in Carlsson et al. (2014) and Chowdhury and Mémoli (2015, 2016). This network distance is analogous to the Gromov–Hausdorff distance between metric spaces, which has previously been used to prove stability results for hierarchical clustering (Carlsson and Mémoli 2008, 2010) and Rips persistence diagrams obtained from finite metric spaces (Chazal et al. 2009, Theorem 3.1). The Gromov–Hausdorff distance was later used in conjunction with the Algebraic Stability Theorem of Chazal et al. (2009) to provide alternative proofs of stability results for Rips and Dowker persistence diagrams arising from metric spaces (Chazal et al. 2014). Our proofs also involve this Algebraic Stability Theorem, but the novelty of our approach lies in a reformulation of the network distance (Proposition 9) that yields direct maps between two networks, thus passing naturally into the machinery of the Algebraic Stability Theorem (without having to define auxiliary constructions such as multivalued maps, as in Chazal et al. (2014)).

A crucial issue that we point out in this paper is that even though we can construct both Rips and Dowker persistence diagrams out of asymmetric networks, Rips persistence diagrams appear to be *blind* to asymmetry, whereas Dowker persistence diagrams do exhibit sensitivity to asymmetry. In the case of Rips complexes, this purported insensitivity to asymmetry can be immediately seen from its definition. In the case of Dowker complexes, we argue about its sensitivity to asymmetry in two different ways. Firstly, we do so by explicitly computing Dowker persistence diagrams of multiple examples of asymmetric networks. In particular, we consider a family of highly asymmetric networks, the *cycle networks*, and by building upon results from Adamaszek



**Fig. 1** A schematic of some of the objects studied in this paper.  $\mathcal{N}$  is the collection of all weighted, directed networks (i.e. digraphs with possibly asymmetric real weights).  $\mathcal{F}$  is the collection of filtered simplicial complexes. Dgm is the collection of persistence diagrams. We study the Rips ( $\mathfrak{R}$ ) and Dowker ( $\mathfrak{D}^{si}$ ,  $\mathfrak{D}^{so}$ ) filtrations, each of which takes a network as input and produces a filtered simplicial complex.  $s$  and  $t$  denote the network transformations of symmetrization (replacing a pair of weights between two nodes by the maximum weight) and transposition (swapping the weights between pairs of nodes).  $\mathfrak{R}$  is insensitive to both  $s$  and  $t$ . But  $\mathfrak{D}^{si} \circ t = \mathfrak{D}^{so}$ ,  $\mathfrak{D}^{so} \circ s = \mathfrak{D}^{si}$ , and in general,  $\mathfrak{D}^{si}$  and  $\mathfrak{D}^{so}$  are *not* invariant under  $t$  (Theorem 43)

and Adams (2017) and Adamaszek et al. (2016) we prove a complete characterization result for the Dowker persistence diagrams—across all dimensions—of any network belonging to this family. These networks constitute directed analogues of circles and may be *motifs* of interest in different applications related to network data analysis. More specifically, appearance of nontrivial 1-dimensional persistence in the Dowker persistence diagram of asymmetric network data may suggest the presence of directed cycles in the data.

Some of our experimental results suggest that the Rips persistence diagrams of this family of networks are pathological, in the sense that they do not represent the signatures one would expect from the underlying dataset, which is a directed circle. Dowker persistence diagrams, on the other hand, are well-behaved in this respect in that they succeed at capturing relevant features. Secondly, we study the degree to which Dowker persistence diagrams are insensitive to changes (such as edge flips, or transposition) in the network structure. An overview of this thread of work is provided in Fig. 1.

**Dowker’s theorem and a functorial generalization** Let  $X, Y$  be two totally ordered sets, and let  $R \subseteq X \times Y$  be a nonempty relation. Then one defines two simplicial complexes  $E_R$  and  $F_R$  as follows. A finite subset  $\sigma \subseteq X$  belongs to  $E_R$  whenever there exists  $y \in Y$  such that  $(x, y) \in R$  for each  $x \in \sigma$ . Similarly a finite subset  $\tau \subseteq Y$  belongs to  $F_R$  whenever there exists  $x \in X$  such that  $(x, y) \in R$  for each  $y \in \tau$ . These constructions can be traced back to Dowker (1952), who proved the following result that we refer to as *Dowker’s theorem*:

**Theorem 1** (Dowker’s theorem; Theorem 1a, Dowker 1952) *Let  $X, Y$  be two totally ordered sets, let  $R \subseteq X \times Y$  be a nonempty relation, and let  $E_R, F_R$  be as above. Then for each  $k \in \mathbb{Z}_+$ ,*

$$H_k(E_R) \cong H_k(F_R).$$

There is also a strong form of Dowker’s theorem that Björner proves via the classical *nerve theorem* (Björner 1995, Theorems 10.6, 10.9). Below we write  $|X|$  to denote the geometric realization of a simplicial complex  $X$ .

**Theorem 2** (The strong form of Dowker’s theorem; Theorem 10.9, Björner 1995) *Under the assumptions of Theorem 1, we in fact have  $|E_R| \simeq |F_R|$ .*

The Functorial Dowker Theorem is the following generalization of the strong form of Dowker’s theorem: instead of a single nonempty relation  $R \subseteq X \times Y$ , consider any pair of nested, nonempty relations  $R \subseteq R' \subseteq X \times Y$ . Then there exist homotopy equivalences between the geometric realizations of the corresponding complexes that commute with the canonical inclusions, up to homotopy. We formalize this statement below.

**Theorem 3** (The functorial Dowker theorem (FDT)) *Let  $X, Y$  be two totally ordered sets, let  $R \subseteq R' \subseteq X \times Y$  be two nonempty relations, and let  $E_R, F_R, E_{R'}, F_{R'}$  be their associated simplicial complexes. Then there exist homotopy equivalences  $\Gamma_{|E_R|} : |F_R| \rightarrow |E_R|$  and  $\Gamma_{|E_{R'}|} : |F_{R'}| \rightarrow |E_{R'}|$  such that the following diagram commutes up to homotopy:*

$$\begin{array}{ccc}
 |F_R| & \xrightarrow{|\iota_E|} & |F_{R'}| \\
 \Gamma_{|E_R|} \downarrow \simeq & & \simeq \downarrow \Gamma_{|E_{R'}|} \\
 |E_R| & \xrightarrow{|\iota_F|} & |E_{R'}|
 \end{array}$$

*In other words, we have  $|\iota_F| \circ \Gamma_{|E_R|} \simeq \Gamma_{|E_{R'}|} \circ |\iota_E|$ , where  $\iota_E, \iota_F$  are the canonical inclusions.*

From Theorem 3 we automatically obtain Theorem 2 (the strong form of Dowker’s theorem) as an immediate corollary. The strong form does not appear in Dowker’s original paper (Dowker 1952), but Björner has given a proof using the nerve theorem (Björner 1995, Theorems 10.6, 10.9). Moreover, Björner writes in a remark following (Björner 1995, Theorem 10.9) that the nerve theorem and the strong form of Dowker’s theorem are equivalent, in the sense that one implies the other. We were not able to find an elementary proof of the strong form of Dowker’s theorem in the existing literature. However, such an elementary proof is provided by our proof of Theorem 3 (given in Sect. 5.1), which we obtained by extending ideas in Dowker’s original proof of Theorem 1.<sup>1</sup>

Whereas the Functorial Dowker Theorem and our elementary proof are of independent interest, it has been suggested in Chazal et al. (2014, Remark 4.8) that such a functorial version of Dowker’s theorem could also be proved using a functorial nerve theorem (Chazal and Oudot 2008, Lemma 3.4). Despite being an interesting possibility, we were not able to find a detailed proof of this claim in the literature. In addition, Björner’s remark regarding the equivalence between the nerve theorem and the strong form of Dowker’s theorem suggests the following question:

<sup>1</sup> A thread with ideas towards the proof of Theorem 2 was discussed in <https://ncatlab.org/nlab/show/Dowker%27s+theorem> (Accessed 24 Apr 2017), but the proposed strategy was incomplete. We have inserted an addendum in <https://ncatlab.org/nlab/show/Dowker%27s+theorem> proposing a complete proof with a slightly different construction.

**Question 1** Are the Functorial Nerve Theorem (FNT) of Chazal and Oudot (2008) and the Functorial Dowker Theorem (FDT, Theorem 3) equivalent?

This question is of fundamental importance because the Nerve Theorem is a crucial tool in the applied topology literature and its functorial generalizations are equally important in persistent homology. In general, the answer is *no*, and moreover, one (of the FNT and FDT) is not stronger than the other. The FNT of Chazal and Oudot (2008) is stated for paracompact spaces, which are more general than the simplicial complexes of the FDT. However, the FNT of Chazal and Oudot (2008) is stated for spaces with *finitely-indexed* covers, so the associated nerve complexes are necessarily finite. All the complexes involved in the statement of the FDT are allowed to be infinite, so the FDT is more general than the FNT in this sense.

To clarify these connections, we formulate a simplicial Functorial Nerve Theorem (Theorem 26) and prove it via a finite formulation of the FDT (Theorem 24). In turn, we show that the simplicial FNT implies the finite FDT, thus proving the equivalence of these formulations (Theorem 27).

**Remark 4** Dowker complexes are also known to researchers who use Q-analysis to study social networks (Johnson 2013; Atkin 1972, 1975). We perceive that viewing Dowker complexes through the modern lens of persistence will enrich the classical framework of Q-analysis by incorporating additional information about the *meaningfulness* of features, thus potentially opening new avenues in the social sciences.

An announcement of part of our work has appeared in Chowdhury and Mémoli (2016).

## 1.2 Implementations

Following work in Curto and Itskov (2008) and Dabaghian et al. (2012), we implement our methods in the setting of classifying simulated hippocampal networks. We simulate the activity pattern of hippocampal cells in an animal as it moves around arenas with a number of obstacles, and compile this data into a network which can be interpreted as the transition matrix for the time-reversal of a Markov process. The motivating idea is to ascertain whether, by just observing hippocampal activity and not using any higher reasoning ability, one might be able to determine the number of obstacles in the arena that the animal has just finished traversing. The results of computing Dowker persistence diagrams suggest that the hippocampal activity is indeed sufficient to accurately count the number of obstacles in each arena.

Our datasets and software are available on <https://research.math.osu.edu/networks/Datasets.html> as part of the `PersNet` software package.

## 1.3 Organization of the paper

Notation used globally is defined directly below. Section 2 contains the necessary background on persistent homology. Section 3 contains our formulations for networks, as well as some key ingredients of our stability results. Section 4 contains details about

the Rips persistence diagram. The first part of Sect. 5 contains details about the Dowker persistence diagram. Section 5.1 contains the Functorial Dowker Theorem. The connection between the simplicial Functorial Nerve Theorem and the finite Functorial Dowker Theorem is detailed in Sect. 5.2. In Sect. 6 we show that Dowker complexes are sensitive to asymmetry. Section 6.1 contains a family of asymmetric networks, the *cycle networks*, and a full characterization of their Dowker persistence diagrams. In Sect. 7 we provide details on an implementation of our methods. Finally, proofs of statements not contained in the main body of the paper are relegated to Appendix A, whereas details about the characterization results for Dowker persistence diagrams of cycle networks are given in Appendix B.

## 1.4 Notation

We will write  $\mathbb{K}$  to denote a field, which we will fix and use throughout the paper. We will write  $\mathbb{Z}_+$  and  $\mathbb{R}_+$  to denote the nonnegative integers and reals, respectively. The extended real numbers  $\mathbb{R} \cup \{\infty, -\infty\}$  will be denoted  $\mathbb{R}$ . The cardinality of a set  $X$  will be denoted  $\text{card}(X)$ . The collection of nonempty subsets of a set  $X$  will be denoted  $\text{pow}(X)$ . The natural numbers  $\{1, 2, 3, \dots\}$  will be denoted by  $\mathbb{N}$ . The dimension of a vector space  $V$  will be denoted  $\text{dim}(V)$ . The rank of a linear transformation  $f$  will be denoted  $\text{rank}(f)$ . An isomorphism between vector spaces  $V$  and  $W$  will be denoted  $V \cong W$ . A homotopy relation for two maps  $f, g : X \rightarrow Y$  between topological spaces will be denoted  $f \simeq g$ . Occasionally we will need to take about multisets, i.e. sets where elements can have multiplicity greater than 1. We will use square bracket notation  $[\dots]$  to denote multisets. Identity maps will be denoted by the notation  $\text{id}_\bullet$ . Given a simplicial complex  $\Sigma$ , we will often write  $V(\Sigma)$  to denote the vertex set of  $\Sigma$ . We will write  $\text{Bd}(\sigma)$  to denote the boundary of a simplex  $\sigma$ .

We say that a map  $F : C \rightarrow D$  between categories  $C$  and  $D$  is *functorial* whenever  $F$  satisfies the properties of being a functor between  $C$  and  $D$  (see Mac Lane 2013 for definitions of these terms). In this paper, the most common usage of “functorial” is in the following scenario. Suppose we have sets  $R, R'$  related by an inclusion  $R \subseteq R'$  along with geometric simplicial complexes  $A_R, B_R, A_{R'}, B_{R'}$  and homotopy equivalences  $f_R : A_R \rightarrow B_R, f_{R'} : A_{R'} \rightarrow B_{R'}$ . The assignment  $R \mapsto (A_R, B_R, f_R)$  is *functorial* if  $f_{R'} \circ \iota_A \simeq \iota_B \circ f_R$ . Here  $\iota_A : A_R \hookrightarrow A_{R'}$  and  $\iota_B : B_R \hookrightarrow B_{R'}$  are inclusions induced by the inclusion  $R \subseteq R'$ .

## 2 Background on persistent homology

We assume that the reader is familiar with terms and concepts related to simplicial homology, and refer to Munkres (1984) for details. Here we describe our choices of notation. Whenever we have a simplicial complex over a set  $X$  and a  $k$ -simplex  $\{x_0, x_1, \dots, x_k\}$ ,  $k \in \mathbb{Z}_+$ , we will assume that the simplex is *oriented* by the ordering  $x_0 < x_1 < \dots < x_k$ . We will write  $[x_0, x_1, \dots, x_k]$  to denote the equivalence class of the even permutations of this chosen ordering, and  $-[x_0, x_1, \dots, x_k]$  to denote the equivalence class of the odd permutations of this ordering. Given a simplicial



complex  $\Sigma$ , we will denote its geometric realization by  $|\Sigma|$ . The *weak topology* on  $|\Sigma|$  is defined by requiring that a subset  $A \subseteq |\Sigma|$  is closed if and only if  $A \cap |\sigma|$  is closed in  $|\sigma|$  for each  $\sigma \in \Sigma$ . A simplicial map  $f : \Sigma \rightarrow \Xi$  between two simplicial complexes induces a map  $|f| : |\Sigma| \rightarrow |\Xi|$  between the geometric realizations, defined as  $|f|(\sum_{v \in \Sigma} a_v v) := \sum_{v \in \Sigma} a_v f(v)$ . These induced maps satisfy the usual composition identity: given simplicial maps  $f : \Sigma \rightarrow \Xi$  and  $g : \Xi \rightarrow \Upsilon$ , we have  $|g \circ f| = |g| \circ |f|$ . To see this, observe the following:

$$|g \circ f| \left( \sum_{v \in \Sigma} a_v v \right) = \sum_{v \in \Sigma} a_v g(f(v)) = |g| \left( \sum_{v \in \Sigma} a_v f(v) \right) = |g| \circ |f| \left( \sum_{v \in \Sigma} a_v v \right). \tag{1}$$

A *filtration* of a simplicial complex  $\Sigma$  (also called a *filtered simplicial complex*) is defined to be a nested sequence  $\{\Sigma^\delta \subseteq \Sigma^{\delta'}\}_{\delta \leq \delta' \in \mathbb{R}}$  of simplicial complexes satisfying the condition that there exist  $\delta_I, \delta_F \in \mathbb{R}$  such that  $\Sigma^\delta = \emptyset$  for all  $\delta \leq \delta_I$ , and  $\Sigma^\delta = \Sigma$  for all  $\delta \geq \delta_F$ .

Fix a field  $\mathbb{K}$ . Given a finite simplicial complex  $\Sigma$  and a dimension  $k \in \mathbb{Z}_+$ , we will denote a *k-chain* in  $\Sigma$  as  $\sum_i a_i \sigma_i$ , where each  $a_i \in \mathbb{K}$  and each  $\sigma_i \in \Sigma$  is a *k-simplex*. We write  $C_k(\Sigma)$  or just  $C_k$  to denote the  $\mathbb{K}$ -vector space of all *k-chains*. We will write  $\partial_k$  to denote the associated *boundary map*  $\partial_k : C_k \rightarrow C_{k-1}$ :

$$\partial_k[x_0, \dots, x_k] := \sum_i (-1)^i [x_0, \dots, \hat{x}_i, \dots, x_k],$$

where  $\hat{x}_i$  denotes omission of  $x_i$  from the sequence.

We will write  $\mathcal{C} = (C_k, \partial_k)_{k \in \mathbb{Z}_+}$  to denote a *chain complex*, i.e. a sequence of vector spaces with boundary maps such that  $\partial_{k-1} \circ \partial_k = 0$ . Given a chain complex  $\mathcal{C}$  and any  $k \in \mathbb{Z}_+$ , the *kth homology of the chain complex*  $\mathcal{C}$  is denoted  $H_k(\mathcal{C}) := \ker(\partial_k) / \text{im}(\partial_{k+1})$ . The *kth Betti number* of  $\mathcal{C}$  is denoted  $\beta_k(\mathcal{C}) := \dim(H_k(\mathcal{C}))$ .

Given a simplicial map  $f$  between simplicial complexes, we write  $f_*$  to denote the induced chain map between the corresponding chain complexes (Munkres 1984, §1.12), and  $(f_k)_\#$  to denote the linear map on *kth homology vector spaces* induced for each  $k \in \mathbb{Z}_+$ .

The operations of passing from simplicial complexes and simplicial maps to chain complexes and induced chain maps, and then to homology vector spaces with induced linear maps, will be referred to as *passing to homology*. Recall the following useful fact, often referred to as *functoriality of homology* (Munkres 1984, Theorem 12.2): given a composition  $g \circ f$  of simplicial maps, we have

$$(g_k \circ f_k)_\# = (g_k)_\# \circ (f_k)_\# \quad \text{for each } k \in \mathbb{Z}_+. \tag{2}$$

A *persistence vector space* is defined to be a family of vector spaces  $\{U^\delta \xrightarrow{\mu_{\delta, \delta'}} U^{\delta'}\}_{\delta \leq \delta' \in \mathbb{R}}$  such that: (1)  $\mu_{\delta, \delta}$  is the identity for each  $\delta \in \mathbb{R}$ , and (2)  $\mu_{\delta, \delta''} = \mu_{\delta, \delta'} \circ \mu_{\delta', \delta''}$  for each  $\delta \leq \delta' \leq \delta'' \in \mathbb{R}$ . The persistence vector spaces that we consider in this work also satisfy the following conditions: (1)  $\dim(U^\delta) < \infty$  at each  $\delta \in \mathbb{R}$ , (2) there exist  $\delta_I, \delta_F \in \mathbb{R}$  such that all maps  $\mu_{\delta, \delta'}$  are isomorphisms for  $\delta, \delta' \geq \delta_F$  and for

$\delta, \delta' \leq \delta_I$ , and (3) there are only finitely many values of  $\delta \in \mathbb{R}$  such that  $U^{\delta-\varepsilon} \not\cong U^\delta$  for each  $\varepsilon > 0$ . Here  $\delta$  is referred to as a *resolution* parameter, and such a persistence vector space is described as being  $\mathbb{R}$ -indexed. The collection of all such persistence vector spaces is denoted  $\mathbf{PVec}(\mathbb{R})$ . Observe that by fixing  $k \in \mathbb{Z}_+$  and passing to the  $k$ th homology vector space at each step  $\Sigma^\delta$  of a filtered simplicial complex  $(\Sigma^\delta)_{\delta \in \mathbb{R}}$ , the functoriality of homology gives us the  $k$ th persistence vector space associated to  $(\Sigma^\delta)_{\delta \in \mathbb{R}}$ , denoted

$$\mathcal{H}_k(\Sigma) := \{H_k(\mathcal{C}^\delta) \xrightarrow{(\iota_{\delta, \delta'})^\#} H_k(\mathcal{C}^{\delta'})\}_{\delta \leq \delta' \in \mathbb{R}}.$$

The elements of  $\mathbf{PVec}(\mathbb{R})$  contain only a finite number of vector spaces, up to isomorphism. By the classification results in (Carlsson et al. 2005, §5.2), it is possible to associate a full invariant, called a *persistence barcode* or *persistence diagram*, to each element of  $\mathbf{PVec}(\mathbb{R})$ . This barcode is a multiset of *persistence intervals*, and is represented as a set of lines over a single axis. The barcode of a persistence vector space  $\mathcal{V}$  is denoted  $\mathbf{Pers}(\mathcal{V})$ . The intervals in  $\mathbf{Pers}(\mathcal{V})$  can be represented as the *persistence diagram of  $\mathcal{V}$* , which is as a multiset of points lying on or above the diagonal in  $\overline{\mathbb{R}}^2$ , counted with multiplicity. More specifically,

$$\text{Dgm}(\mathcal{V}) := [(\delta_i, \delta_{j+1}) \in \overline{\mathbb{R}}^2 : [\delta_i, \delta_{j+1}) \in \mathbf{Pers}(\mathcal{V})],$$

where the multiplicity of  $(\delta_i, \delta_{j+1}) \in \overline{\mathbb{R}}^2$  is given by the multiplicity of  $[\delta_i, \delta_{j+1}) \in \mathbf{Pers}(\mathcal{V})$ .

Persistence diagrams can be compared using the *bottleneck distance*, which we denote by  $d_B$ . Details about this distance, as well as the other material related to persistent homology, can be found in Chazal et al. (2016). Numerous other formulations of the material presented above can be found in Edelsbrunner et al. (2002, 2015), Zomorodian and Carlsson (2005), Carlsson and Silva (2010), Edelsbrunner and Harer (2010), Edelsbrunner and Morozov (2014) and Bauer and Lesnick (2014).

**Remark 5** Whenever we describe a persistence diagram as being *trivial*, we mean that either it is empty, or it does not have any off-diagonal points.

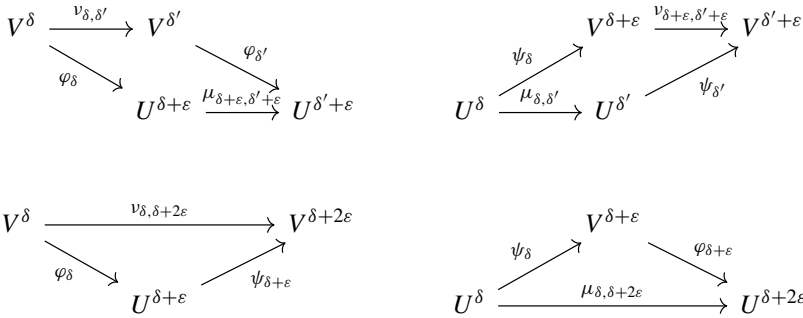
### 2.1 Interleaving distance and stability of persistence vector spaces

In what follows, we will consider  $\mathbb{R}$ -indexed persistence vector spaces  $\mathbf{PVec}(\mathbb{R})$ .

Given  $\varepsilon \geq 0$ , two  $\mathbb{R}$ -indexed persistence vector spaces  $\mathcal{V} = \{V^\delta \xrightarrow{\nu_{\delta, \delta'}} V^{\delta'}\}_{\delta \leq \delta'}$  and  $\mathcal{U} = \{U^\delta \xrightarrow{\mu_{\delta, \delta'}} U^{\delta'}\}_{\delta \leq \delta'}$  are said to be  $\varepsilon$ -interleaved (Chazal et al. 2009; Bauer and Lesnick 2014) if there exist two families of linear maps

$$\begin{aligned} \{\varphi_{\delta, \delta+\varepsilon} : V^\delta &\rightarrow U^{\delta+\varepsilon}\}_{\delta \in \mathbb{R}}, \\ \{\psi_{\delta, \delta+\varepsilon} : U^\delta &\rightarrow V^{\delta+\varepsilon}\}_{\delta \in \mathbb{R}} \end{aligned}$$

such that the following diagrams commute for all  $\delta \leq \delta' \in \mathbb{R}$ :



The purpose of introducing  $\epsilon$ -interleavings is to define a pseudometric on the collection of persistence vector spaces. The *interleaving distance* between two  $\mathbb{R}$ -indexed persistence vector spaces  $\mathcal{V}, \mathcal{U}$  is given by:

$$d_I(\mathcal{U}, \mathcal{V}) := \inf\{\epsilon \geq 0 : \mathcal{U} \text{ and } \mathcal{V} \text{ are } \epsilon\text{-interleaved}\}.$$

This definition induces a pseudometric on the collection of persistence vector spaces (Chazal et al. 2009; Bauer and Lesnick 2014). The interleaving distance is then related to the bottleneck distance as follows:

**Theorem 6** (Algebraic stability theorem, Chazal et al. 2009) *Let  $\mathcal{U}, \mathcal{V}$  be two  $\mathbb{R}$ -indexed persistence vector spaces. Then,*

$$d_B(\text{Dgm}(\mathcal{U}), \text{Dgm}(\mathcal{V})) \leq d_I(\mathcal{U}, \mathcal{V}).$$

Stability results are at the core of persistent homology, beginning with the classical bottleneck stability result in Cohen-Steiner et al. (2007). One of our key contributions is to use the Algebraic Stability Theorem stated above, along with Lemma Sect. 3 stated below, to prove stability results for methods of computing persistent homology of a network.

Before stating the following lemma, recall that two simplicial maps  $f, g : \Sigma \rightarrow \Xi$  are *contiguous* if for any simplex  $\sigma \in \Sigma$ ,  $f(\sigma) \cup g(\sigma)$  is a simplex of  $\Xi$ . Contiguous maps satisfy the following useful properties:

**Proposition 7** (Properties of contiguous maps) *Let  $f, g : \Sigma \rightarrow \Xi$  be two contiguous simplicial maps. Then,*

- (1)  $|f|, |g| : |\Sigma| \rightarrow |\Xi|$  are homotopic (Spanier 1994, §3.5), and
- (2) The chain maps induced by  $f$  and  $g$  are chain homotopic, and as a result, the induced maps  $f_{\#}$  and  $g_{\#}$  for homology are equal (Munkres 1984, Theorem 12.5).

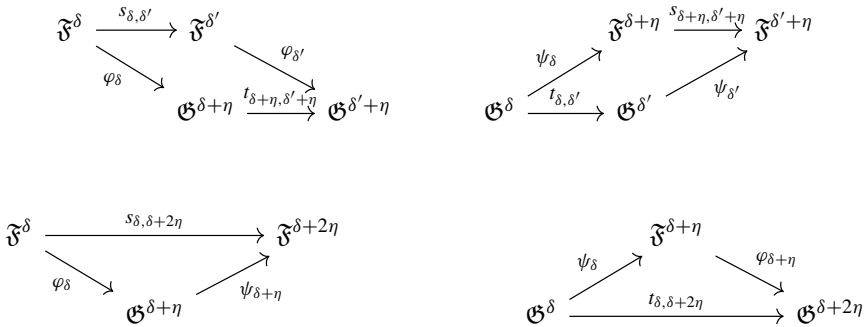
**Lemma 8** (Stability Lemma) *Let  $\mathfrak{F}, \mathfrak{G}$  be two filtered simplicial complexes written as*

$$\{\mathfrak{F}^{\delta} \xrightarrow{s_{\delta,\delta'}} \mathfrak{F}^{\delta'}\}_{\delta' \geq \delta \in \mathbb{R}} \text{ and } \{\mathfrak{G}^{\delta} \xrightarrow{t_{\delta,\delta'}} \mathfrak{G}^{\delta'}\}_{\delta' \geq \delta \in \mathbb{R}},$$

where  $s_{\delta, \delta'}$  and  $t_{\delta, \delta'}$  denote the natural inclusion maps. Suppose  $\eta \geq 0$  is such that there exist families of simplicial maps  $\{\varphi_\delta : \mathfrak{F}^\delta \rightarrow \mathfrak{G}^{\delta+\eta}\}_{\delta \in \mathbb{R}}$  and  $\{\psi_\delta : \mathfrak{G}^\delta \rightarrow \mathfrak{F}^{\delta+\eta}\}_{\delta \in \mathbb{R}}$  such that the following are satisfied for any  $\delta \leq \delta'$ :

- (1)  $t_{\delta+\eta, \delta'+\eta} \circ \varphi_\delta$  and  $\varphi_{\delta'}$  are contiguous
- (2)  $s_{\delta+\eta, \delta'+\eta} \circ \psi_\delta$  and  $\psi_{\delta'}$  are contiguous
- (3)  $\psi_{\delta+\eta} \circ \varphi_\delta$  and  $s_{\delta, \delta+2\eta}$  are contiguous
- (4)  $\varphi_{\delta+\eta} \circ \psi_\delta$  and  $t_{\delta, \delta+2\eta}$  are contiguous.

All the diagrams are as below:



For each  $k \in \mathbb{Z}_+$ , let  $\mathcal{H}_k(\mathfrak{F}), \mathcal{H}_k(\mathfrak{G})$  denote the  $k$ -dimensional persistence vector spaces associated to  $\mathfrak{F}$  and  $\mathfrak{G}$ . Then for each  $k \in \mathbb{Z}_+$ ,

$$d_B(\text{Dgm}_k(\mathcal{H}_k(\mathfrak{F})), \text{Dgm}_k(\mathcal{H}_k(\mathfrak{G}))) \leq d_I(\mathcal{H}_k(\mathfrak{F}), \mathcal{H}_k(\mathfrak{G})) \leq \eta.$$

### 3 Background on networks and our network distance

A network is a pair  $(X, \omega_X)$  where  $X$  is a finite set and  $\omega_X : X \times X \rightarrow \mathbb{R}$  is a weight function. Note that  $\omega_X$  need not satisfy the triangle inequality, any symmetry condition, or even the requirement that  $\omega_X(x, x) = 0$  for all  $x \in X$ . The weights are even allowed to be negative. The collection of all such networks is denoted  $\mathcal{N}$ .

When comparing networks, one needs a way to correlate points in one network with points in the other. To see how this can be done, let  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ . Let  $R$  be any nonempty relation between  $X$  and  $Y$ , i.e. a nonempty subset of  $X \times Y$ . The distortion of the relation  $R$  is given by:

$$\text{dis}(R) := \max_{(x, y), (x', y') \in R} |\omega_X(x, x') - \omega_Y(y, y')|.$$

A correspondence between  $X$  and  $Y$  is a relation  $R$  between  $X$  and  $Y$  such that  $\pi_X(R) = X$  and  $\pi_Y(R) = Y$ , where  $\pi_X : X \times Y \rightarrow X$  and  $\pi_Y : X \times Y \rightarrow Y$  denote the natural projections. The collection of all correspondences between  $X$  and  $Y$  will be denoted  $\mathcal{R}(X, Y)$ .

Following previous work in Carlsson et al. (2014); Chowdhury and Mémoli (2015, 2016) the *network distance*  $d_{\mathcal{N}} : \mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}_+$  is then defined as:

$$d_{\mathcal{N}}(X, Y) := \frac{1}{2} \min_{R \in \mathcal{R}} \text{dis}(R).$$

The  $d_{\mathcal{N}}$  as defined above is a pseudometric, and the networks at 0  $d_{\mathcal{N}}$ -distance have been completely characterized (Chowdhury and Mémoli 2015, 2017). Next we wish to prove the reformulation in Proposition 9. First we define the distortion of a map between two networks. Here a *map between networks*  $(X, \omega_X)$  and  $(Y, \omega_Y)$  is simply any function  $\varphi : X \rightarrow Y$ . Given any  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$  and a map  $\varphi : (X, \omega_X) \rightarrow (Y, \omega_Y)$ , the *distortion* of  $\varphi$  is defined as:

$$\text{dis}(\varphi) := \max_{x, x' \in X} |\omega_X(x, x') - \omega_Y(\varphi(x), \varphi(x'))|.$$

Next, given maps  $\varphi : (X, \omega_X) \rightarrow (Y, \omega_Y)$  and  $\psi : (Y, \omega_Y) \rightarrow (X, \omega_X)$ , we define two *co-distortion* terms:

$$C_{X,Y}(\varphi, \psi) := \max_{(x,y) \in X \times Y} |\omega_X(x, \psi(y)) - \omega_Y(\varphi(x), y)|,$$

$$C_{Y,X}(\psi, \varphi) := \max_{(y,x) \in Y \times X} |\omega_Y(y, \varphi(x)) - \omega_X(\psi(y), x)|.$$

**Proposition 9** *Let  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ . Then,*

$$d_{\mathcal{N}}(X, Y) = \frac{1}{2} \min\{\max(\text{dis}(\varphi), \text{dis}(\psi), C_{X,Y}(\varphi, \psi), C_{Y,X}(\psi, \varphi)) : \varphi : X \rightarrow Y, \psi : Y \rightarrow X \text{ any maps}\}.$$

**Remark 10** Proposition 9 is analogous to a result of Kalton and Ostrovskii (1999, Theorem 2.1) where—instead of  $d_{\mathcal{N}}$ —one has the Gromov–Hausdorff distance between metric spaces. We remark that when restricted to the special case of networks that are also metric spaces, the network distance  $d_{\mathcal{N}}$  agrees with the Gromov–Hausdorff distance. Details on the Gromov–Hausdorff distance can be found in Burago et al. (2001).

An important remark is that in the Kalton–Ostrovskii formulation, there is only one co-distortion term. When Proposition 9 is applied to metric spaces, the two co-distortion terms become equal by symmetry, and thus the Kalton–Ostrovskii formulation is recovered. But *a priori*, the lack of symmetry in the network setting requires us to consider both terms.

**Remark 11** In the following sections, we propose methods for computing persistent homology of networks, and prove that they are stable via Lemma 8. Note that similar results, valid in the setting of metric spaces, have appeared in Chazal et al. (2009) and Chazal et al. (2014). Whereas the proofs in Chazal et al. (2014) invoke an auxiliary construction of multivalued maps arising from correspondences, our proofs simply use the maps  $\varphi, \psi$  arising directly from the reformulation of  $d_{\mathcal{N}}$  (Proposition 9), thus streamlining the treatment.

When studying the effect of asymmetry on persistent homology, it will be useful to consider the network transformations that we define next.

**Definition 1** (*Symmetrization and transposition*) Define the *max-symmetrization* map  $\varepsilon : \mathcal{N} \rightarrow \mathcal{N}$  by  $(X, \omega_X) \mapsto (X, \widehat{\omega}_X)$ , where for any network  $(X, \omega_X)$ , we define  $\widehat{\omega}_X : X \times X \rightarrow \mathbb{R}$  as follows:

$$\widehat{\omega}_X(x, x') := \max(\omega_X(x, x'), \omega_X(x', x)), \text{ for } x, x' \in X.$$

Also define the *transposition* map  $\mathfrak{t} : \mathcal{N} \rightarrow \mathcal{N}$  by  $(X, \omega_X) \mapsto (X, \omega_X^\top)$ , where for any  $(X, \omega_X) \in \mathcal{N}$ , we define  $\omega_X^\top(x, x') := \omega_X(x', x)$  for  $x, x' \in X$ . For convenience, we denote  $X^\top := \mathfrak{t}(X)$  for any network  $X$ .

We are now ready to formulate our two methods for computing persistent homology of networks. The Rips filtration is the “workhorse” of persistent homology of metric spaces so it is natural to consider its generalization to general asymmetric networks.

#### 4 The Rips filtration of a network

Recall that for a metric space  $(X, d_X)$ , the *Rips complex* is defined for each  $\delta \geq 0$  as follows:

$$\mathfrak{R}_X^\delta := \{\sigma \in \text{pow}(X) : \text{diam}(\sigma) \leq \delta\}, \text{ where } \text{diam}(\sigma) := \max_{x, x' \in \sigma} d_X(x, x').$$

Following this definition, we define the Rips complex for a network  $(X, \omega_X)$  as follows:

$$\mathfrak{R}_X^\delta := \{\sigma \in \text{pow}(X) : \max_{x, x' \in \sigma} \omega_X(x, x') \leq \delta\}.$$

To any network  $(X, \omega_X)$ , we may associate the *Rips filtration*  $\{\mathfrak{R}_X^\delta \hookrightarrow \mathfrak{R}_X^{\delta'}\}_{\delta \leq \delta'}$ . We denote the  $k$ -dimensional persistence vector space associated to this filtration by  $\mathcal{H}_k^{\mathfrak{R}}(X)$ , and the corresponding persistence diagram by  $\text{Dgm}_k^{\mathfrak{R}}(X)$ . The Rips filtration is stable to small perturbations of the input data:

**Proposition 12** *Let  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ . Then  $d_B(\text{Dgm}_k^{\mathfrak{R}}(X), \text{Dgm}_k^{\mathfrak{R}}(Y)) \leq 2d_{\mathcal{N}}(X, Y)$ .*

We omit the proof because it is similar to that of Proposition 15, which we will prove in detail.

**Remark 13** The preceding proposition serves a dual purpose: (1) it shows that the Rips persistence diagram is robust to noise in input data, and (2) it shows that instead of computing the network distance between two networks, one can compute the bottleneck distance between their Rips persistence diagrams as a suitable proxy. The advantage to computing bottleneck distance is that it can be done in polynomial time

(see Efrat et al. 2001), whereas computing  $d_{\mathcal{N}}$  is NP-hard in general. This follows from the fact that the problem of computing  $d_{\mathcal{N}}$  includes the problem of computing the Gromov–Hausdorff distance between finite metric spaces, which is an NP-hard problem (Schmiedl 2015). We remark that the idea of computing Rips persistence diagrams to compare finite metric spaces first appeared in Chazal et al. (2009), and moreover, that Proposition 12 is an extension of Theorem 3.1 in Chazal et al. (2009).

The Rips filtration in the setting of symmetric networks has been used in Horak et al. (2009), Carstens and Horadam (2013), Giusti et al. (2015) and Petri et al. (2013), albeit without addressing stability results. To our knowledge, Proposition 12 is the first quantitative result justifying the constructions in these prior works.

**Remark 14** (Rips is insensitive to asymmetry) A critical weakness of the Rips complex construction is that it is not sensitive to asymmetry. To see this, recall the symmetrization map defined in Definition 1, and let  $(X, \omega_X) \in \mathcal{N}$ . Now for any  $\sigma \in \text{pow}(X)$ , we have  $\max_{x, x' \in \sigma} \omega_X(x, x') = \max_{x, x' \in \sigma} \widehat{\omega}_X(x, x')$ . It follows that for each  $\delta \geq 0$ , the Rips complexes of  $(X, \omega_X)$  and  $(X, \widehat{\omega}_X) = \mathfrak{s}(X, \omega_X)$  are equal, i.e.  $\mathfrak{R} = \mathfrak{R} \circ \mathfrak{s}$ . Thus the Rips persistence diagrams of the original and max-symmetrized networks are equal.

### 5 The Dowker filtration of a network

Given  $(X, \omega_X) \in \mathcal{N}$ , and for any  $\delta \in \mathbb{R}$ , consider the following relation on  $X$ :

$$R_{\delta, X} := \{(x, x') : \omega_X(x, x') \leq \delta\}. \tag{3}$$

Then  $R_{\delta, X} \subseteq X \times X$ , and  $R_{\delta_F, X} = X \times X$  for some sufficiently large  $\delta_F$ . Furthermore, for any  $\delta' \geq \delta$ , we have  $R_{\delta, X} \subseteq R_{\delta', X}$ . Using  $R_{\delta, X}$ , we build a simplicial complex  $\mathfrak{D}_{\delta}^{\text{si}}$  as follows:

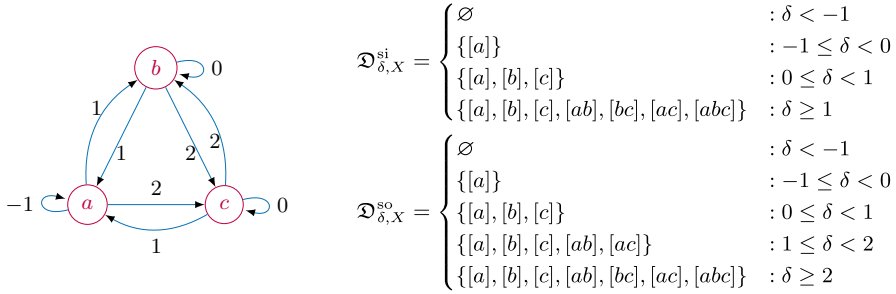
$$\mathfrak{D}_{\delta, X}^{\text{si}} := \{\sigma = [x_0, \dots, x_n] : \text{there exists } x' \in X \text{ such that } (x_i, x') \in R_{\delta, X} \text{ for each } x_i\}. \tag{4}$$

If  $\sigma \in \mathfrak{D}_{\delta, X}^{\text{si}}$ , it is clear that any face of  $\sigma$  also belongs to  $\mathfrak{D}_{\delta, X}^{\text{si}}$ . We call  $\mathfrak{D}_{\delta, X}^{\text{si}}$  the *Dowker  $\delta$ -sink simplicial complex* associated to  $X$ , and refer to  $x'$  as a  *$\delta$ -sink* for  $\sigma$  (where  $\sigma$  and  $x'$  should be clear from context).

Since  $R_{\delta, X}$  is an increasing sequence of sets, it follows that  $\mathfrak{D}_{\delta, X}^{\text{si}}$  is an increasing sequence of simplicial complexes. In particular, for  $\delta' \geq \delta$ , there is a natural inclusion map  $\mathfrak{D}_{\delta, X}^{\text{si}} \hookrightarrow \mathfrak{D}_{\delta', X}^{\text{si}}$ . We write  $\mathfrak{D}_X^{\text{si}}$  to denote the filtration  $\{\mathfrak{D}_{\delta, X}^{\text{si}} \hookrightarrow \mathfrak{D}_{\delta', X}^{\text{si}}\}_{\delta \leq \delta'}$  associated to  $X$ . We call this the *Dowker sink filtration on  $X$* . We will denote the  $k$ -dimensional persistence diagram arising from this filtration by  $\text{Dgm}_k^{\text{si}}(X)$ .

We define a dual construction as follows:

$$\mathfrak{D}_{\delta, X}^{\text{so}} := \{\sigma = [x_0, \dots, x_n] : \text{there exists } x' \in X \text{ such that } (x', x_i) \in R_{\delta, X} \text{ for each } x_i\}. \tag{5}$$



**Fig. 2** Computing the Dowker sink and source complexes of a network  $(X, \omega_X)$ . Observe that the sink and source complexes are different in the range  $1 \leq \delta < 2$

We call  $\mathfrak{D}_{\delta, X}^{\text{so}}$  the *Dowker  $\delta$ -source simplicial complex* associated to  $X$ . The filtration  $\{\mathfrak{D}_{\delta, X}^{\text{so}} \hookrightarrow \mathfrak{D}_{\delta', X}^{\text{so}}\}_{\delta \leq \delta'}$  associated to  $X$  is called the *Dowker source filtration*, denoted  $\mathfrak{D}_X^{\text{so}}$ . We denote the  $k$ -dimensional persistence diagram arising from this filtration by  $\text{Dgm}_k^{\text{so}}(X)$ . Notice that any construction using  $\mathfrak{D}_{\delta, X}^{\text{si}}$  can also be repeated using  $\mathfrak{D}_{\delta, X}^{\text{so}}$ , so we focus on the case of the sink complexes and restate results in terms of source complexes where necessary. In particular, we will prove in Sect. 5.1 that

$$\text{Dgm}_k^{\text{si}}(X) = \text{Dgm}_k^{\text{so}}(X) \text{ for any } k \in \mathbb{Z}_+,$$

so it makes sense to talk about “the” Dowker diagram associated to  $X$ .

The sink and source filtrations are not equal in general; this is illustrated in Fig. 2.

As in the case of the Rips filtration, both the Dowker sink and source filtrations are stable.

**Proposition 15** *Let  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ . Then  $d_B(\text{Dgm}_k^\bullet(X), \text{Dgm}_k^\bullet(Y)) \leq 2d_{\mathcal{N}}(X, Y)$ . Here  $\text{Dgm}^\bullet$  refers to either of  $\text{Dgm}^{\text{si}}$  and  $\text{Dgm}^{\text{so}}$ .*

**Proof of Proposition 15** Both cases are similar, so we just prove the result for  $\text{Dgm}^{\text{si}}$ . Let  $\eta = 2d_{\mathcal{N}}(X, Y)$ . Then by Proposition 9, there exist maps  $\varphi : X \rightarrow Y, \psi : Y \rightarrow X$  such that

$$\max(\text{dis}(\varphi), \text{dis}(\psi), C_{X, Y}(\varphi, \psi), C_{Y, X}(\psi, \varphi)) \leq \eta.$$

First we check that  $\varphi, \psi$  induce simplicial maps  $\varphi_\delta : \mathfrak{D}_{\delta, X}^{\text{si}} \rightarrow \mathfrak{D}_{\delta+\eta, Y}^{\text{si}}$  and  $\psi_\delta : \mathfrak{D}_{\delta, Y}^{\text{si}} \rightarrow \mathfrak{D}_{\delta+\eta, X}^{\text{si}}$  for each  $\delta \in \mathbb{R}$ .

Let  $\delta' \geq \delta \in \mathbb{R}$ . Let  $\sigma = [x_0, \dots, x_n] \in \mathfrak{D}_{\delta, X}^{\text{si}}$ . Then there exists  $x' \in X$  such that  $\omega_X(x_i, x') \leq \delta$  for each  $0 \leq i \leq n$ . Fix such an  $x'$ . Since  $\text{dis}(\varphi) \leq \eta$ , we have the following for each  $i$ :

$$|\omega_X(x_i, x') - \omega_Y(\varphi(x_i), \varphi(x'))| \leq \eta.$$

So  $\omega_Y(\varphi(x_i), \varphi(x')) \leq \omega_X(x_i, x') + \eta \leq \delta + \eta$  for each  $0 \leq i \leq n$ . Thus  $\varphi_\delta(\sigma) := \{\varphi(x_0), \dots, \varphi(x_n)\}$  is a simplex in  $\mathfrak{D}_{\delta+\eta, Y}^{\text{si}}$ . Thus the map on simplices  $\varphi_\delta$  induced by  $\varphi$  is simplicial for each  $\delta \in \mathbb{R}$ .



Similarly we check that the map  $\psi_\delta$  on simplices induced by  $\psi$  is simplicial. Now to prove the result, it will suffice to check the contiguity conditions in the statement of Lemma 8. Consider the following diagram:

$$\begin{array}{ccc} \mathfrak{D}_{\delta, X}^{\text{si}} & \xrightarrow{s_{\delta, \delta'}} & \mathfrak{D}_{\delta', X}^{\text{si}} \\ & \searrow \varphi_\delta & \searrow \varphi_{\delta'} \\ & & \mathfrak{D}_{\delta+\eta, Y}^{\text{si}} \xrightarrow{t_{\delta+\eta, \delta'+\eta}} \mathfrak{D}_{\delta'+\eta, Y}^{\text{si}} \end{array}$$

Here  $s_{\delta, \delta'}$  and  $t_{\delta+\eta, \delta'+\eta}$  are the inclusion maps. We claim that  $t_{\delta+\eta, \delta'+\eta} \circ \varphi_\delta$  and  $\varphi_{\delta'} \circ s_{\delta, \delta'}$  are contiguous simplicial maps. To see this, let  $\sigma \in \mathfrak{D}_{\delta, X}^{\text{si}}$ . Since  $s_{\delta, \delta'}$  is just the inclusion, it follows that  $t_{\delta+\eta, \delta'+\eta}(\varphi_\delta(\sigma)) \cup \varphi_{\delta'}(s_{\delta, \delta'}(\sigma)) = \varphi_\delta(\sigma)$ , which is a simplex in  $\mathfrak{D}_{\delta+\eta, Y}^{\text{si}}$  because  $\varphi_\delta$  is simplicial, and hence a simplex in  $\mathfrak{D}_{\delta'+\eta, Y}^{\text{si}}$  because the inclusion  $t_{\delta+\eta, \delta'+\eta}$  is simplicial. Thus  $t_{\delta+\eta, \delta'+\eta} \circ \varphi_\delta$  and  $\varphi_{\delta'} \circ s_{\delta, \delta'}$  are contiguous, and their induced linear maps for homology are equal. By a similar argument, we verify that  $s_{\delta+\eta, \delta'+\eta} \circ \psi_\delta$  and  $\psi_{\delta'} \circ t_{\delta, \delta'}$  are contiguous simplicial maps as well.

Next we check that the maps  $\psi_{\delta+\eta} \circ \varphi_\delta$  and  $s_{\delta, \delta+2\eta}$  in the figure below are contiguous.

$$\begin{array}{ccc} \mathfrak{D}_{\delta, X}^{\text{si}} & \xrightarrow{s_{\delta, \delta+2\eta}} & \mathfrak{D}_{\delta+2\eta, X}^{\text{si}} \\ & \searrow \varphi_\delta & \nearrow \psi_{\delta+\eta} \\ & & \mathfrak{D}_{\delta+\eta, Y}^{\text{si}} \end{array}$$

Let  $x_i \in \sigma$ . Note that for our fixed  $\sigma = [x_0, \dots, x_n] \in \mathfrak{D}_{\delta, X}^{\text{si}}$  and  $x'$ , we have:

$$\begin{aligned} |\omega_X(x_i, x') - \omega_X(\psi(\varphi(x_i)), \psi(\varphi(x')))| &\leq |\omega_X(x_i, x') - \omega_Y(\varphi(x_i), \varphi(x'))| \\ &\quad + |\omega_Y(\varphi(x_i), \varphi(x')) \\ &\quad - \omega_X(\psi(\varphi(x_i)), \psi(\varphi(x')))| \\ &\leq 2\eta. \end{aligned}$$

Thus we obtain  $\omega_X(\psi(\varphi(x_i)), \psi(\varphi(x')) \leq \omega_X(x_i, x') + 2\eta \leq \delta + 2\eta$ .

Since this holds for any  $x_i \in \sigma$ , it follows that  $\psi_{\delta+\eta}(\varphi_\delta(\sigma)) \in \mathfrak{D}_{\delta+2\eta, X}^{\text{si}}$ . We further claim that

$$\tau := \sigma \cup \psi_{\delta+\eta}(\varphi_\delta(\sigma)) = \{x_0, x_1, \dots, x_n, \psi(\varphi(x_0)), \dots, \psi(\varphi(x_n))\}$$

is a simplex in  $\mathfrak{D}_{\delta+2\eta, X}^{\text{si}}$ . Let  $0 \leq i \leq n$ . It suffices to show that  $\omega_X(x_i, \psi(\varphi(x')) \leq \delta + 2\eta$ .

Notice that from the reformulation of  $d_{\mathcal{N}}$  (Proposition 9), we have

$$C_{X, Y}(\varphi, \psi) = \max_{(x, y) \in X \times Y} |\omega_X(x, \psi(y)) - \omega_Y(\varphi(x), y)| \leq \eta.$$

Let  $y = \varphi(x')$ . Then  $|\omega_X(x_i, \psi(y)) - \omega_Y(\varphi(x_i), y)| \leq \eta$ . In particular,

$$\omega_X(x_i, \psi(\varphi(x'))) \leq \omega_Y(\varphi(x_i), \varphi(x')) + \eta \leq \omega_X(x_i, x') + 2\eta \leq \delta + 2\eta.$$

Since  $0 \leq i \leq n$  were arbitrary, it follows that  $\tau \in \mathcal{D}_{\delta+2\eta, X}^{\text{si}}$ . Thus  $\psi_{\delta+\eta} \circ \varphi_\delta$  and  $s_{\delta, \delta+2\eta}$  are contiguous. Similarly, we use the  $\text{dis}(\psi)$  and  $C_{Y, X}(\psi, \varphi)$  terms to verify that  $t_{\delta, \delta+2\eta}$  and  $\varphi_{\delta+\eta} \circ \psi_\delta$  are contiguous.

The result now follows by an application of Lemma 8. □

**Remark 16** The preceding proposition shows that the Dowker persistence diagram is robust to noise in input data, and that the bottleneck distance between Dowker persistence diagrams arising from two networks can be used as a proxy for computing the actual network distance. Note the analogy with Remark 13.

Both the Dowker and Rips filtrations are valid methods for computing persistent homology of networks, by virtue of their stability results (Propositions 12 and 15). However, we present the Dowker filtration as an appropriate method for capturing directionality information in directed networks. In Sect. 6 we discuss this particular feature of the Dowker filtration in full detail.

**Remark 17** (Symmetric networks) In the setting of symmetric networks, the Dowker sink and source simplicial filtrations coincide, and so we automatically obtain  $\text{Dgm}_k^{\text{so}}(X) = \text{Dgm}_k^{\text{si}}(X)$  for any  $k \in \mathbb{Z}_+$  and any  $(X, \omega_X) \in \mathcal{N}$ .

**Remark 18** (The metric space setting and relation to witness complexes) When restricted to the setting of metric spaces, the Dowker complex resembles a construction called the witness complex (De Silva and Carlsson 2004). In particular, a version of the Dowker complex for metric spaces, constructed in terms of *landmarks* and *witnesses*, was discussed in Chazal et al. (2014), along with stability results. When restricted to the special networks that are pseudo-metric spaces, our definitions and results agree with those presented in Chazal et al. (2014).

### 5.1 The functorial Dowker theorem and equivalence of diagrams

Let  $(X, \omega_X) \in \mathcal{N}$ , and let  $\delta \in \mathbb{R}$  be such that  $R_{\delta, X}$  is nonempty. By applying Dowker's theorem (Theorem 1) to the setting  $Y = X$ , we have  $H_k(\mathcal{D}_{\delta, X}^{\text{si}}) \cong H_k(\mathcal{D}_{\delta, X}^{\text{so}})$ , for any  $k \in \mathbb{Z}_+$ . We still have this equality in the case where  $R_{\delta, X}$  is empty, because then  $\mathcal{D}_{\delta, X}^{\text{si}}$  and  $\mathcal{D}_{\delta, X}^{\text{so}}$  are both empty. Thus we obtain:

**Corollary 19** *Let  $(X, \omega_X) \in \mathcal{N}$ ,  $\delta \in \mathbb{R}$ , and  $k \in \mathbb{Z}_+$ . Then,*

$$H_k(\mathcal{D}_{\delta, X}^{\text{si}}) \cong H_k(\mathcal{D}_{\delta, X}^{\text{so}}).$$

In the persistent setting, Theorem 1 and Corollary 19 suggest the following question:

*Given a network  $(X, \omega_X)$  and a fixed dimension  $k \in \mathbb{Z}_+$ , are the persistence diagrams of the Dowker sink and source filtrations of  $(X, \omega_X)$  necessarily equal?*

In what follows, we provide a positive answer to the question above. Our strategy is to use the Functorial Dowker Theorem (Theorem 3), for which we will provide a complete proof below. The Functorial Dowker Theorem implies equality between sink and source persistence diagrams.

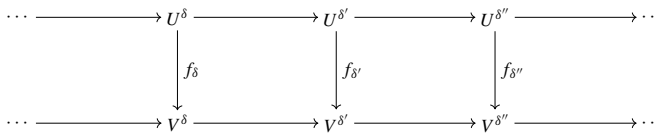
**Corollary 20** (Dowker duality) *Let  $(X, \omega_X) \in \mathcal{N}$ , and let  $k \in \mathbb{Z}_+$ . Then,*

$$\text{Dgm}_k^{\text{si}}(X) = \text{Dgm}_k^{\text{so}}(X).$$

*Thus we may call either of the diagrams above the  $k$ -dimensional Dowker diagram of  $X$ , denoted  $\text{Dgm}_k^{\mathcal{D}}(X)$ .*

Before proving the corollary, we state an  $\mathbb{R}$ -indexed variant of the Persistence Equivalence Theorem (Edelsbrunner and Harer 2010). This particular version follows from the *isometry theorem* (Bauer and Lesnick 2014), and we refer the reader to Chazal et al. (2016, Chapter 5) for an expanded presentation of this material.

**Theorem 21** (Persistence equivalence theorem) *Consider two persistence vector spaces  $\mathcal{U} = \{U^\delta \xrightarrow{\mu_{\delta, \delta'}} U^{\delta'}\}_{\delta \leq \delta' \in \mathbb{R}}$  and  $\mathcal{V} = \{V^\delta \xrightarrow{\nu_{\delta, \delta'}} V^{\delta'}\}_{\delta \leq \delta' \in \mathbb{R}}$  with connecting maps  $f_\delta : U^\delta \rightarrow V^\delta$ .*



*If the  $f_\delta$  are all isomorphisms and each square in the diagram above commutes, then:*

$$\text{Dgm}(\mathcal{U}) = \text{Dgm}(\mathcal{V}).$$

**Proof of Corollary 20** Let  $\delta \leq \delta' \in \mathbb{R}$ , and consider the relations  $R_{\delta, X} \subseteq R_{\delta', X} \subseteq X \times X$ . Suppose first that  $R_{\delta, X}$  and  $R_{\delta', X}$  are both nonempty. By applying Theorem 3, we obtain homotopy equivalences between the source and sink complexes that commute with the canonical inclusions up to homotopy. Passing to the  $k$ th homology level, we obtain persistence vector spaces that satisfy the commutativity properties of Theorem 21. The result follows from Theorem 21.

In the case where  $R_{\delta, X}$  and  $R_{\delta', X}$  are both empty, there is nothing to show because all the associated complexes are empty. Suppose  $R_{\delta, X}$  is empty, and  $R_{\delta', X}$  is nonempty. Then  $\mathcal{D}_{\delta, X}^{\text{si}}$  and  $\mathcal{D}_{\delta, X}^{\text{so}}$  are empty, so their inclusions into  $\mathcal{D}_{\delta', X}^{\text{si}}$  and  $\mathcal{D}_{\delta', X}^{\text{so}}$  induce zero maps upon passing to homology. Thus the commutativity of Theorem 21 is satisfied, and the result follows by Theorem 21. □

**The proof of the Functorial Dowker Theorem** It remains to prove Theorem 3. Because the proof involves numerous maps, we will adopt the notational convention of adding a subscript to a function to denote its codomain—e.g. we will write  $f_B$  to denote a function with codomain  $B$ .

First we recall the construction of a combinatorial barycentric subdivision (see Dowker 1952, §2; Lefschetz 1942, §4.7; Barmak 2011, Appendix A).

**Definition 2** (*Barycentric subdivisions*) For any simplicial complex  $\Sigma$ , one may construct a new simplicial complex  $\Sigma^{(1)}$ , called the *first barycentric subdivision*, as follows:

$$\Sigma^{(1)} := \{[\sigma_1, \sigma_2, \dots, \sigma_p] : \sigma_1 \subseteq \sigma_2 \subseteq \dots \subseteq \sigma_p, \text{ each } \sigma_i \in \Sigma\}.$$

Note that the vertices of  $\Sigma^{(1)}$  are the simplices of  $\Sigma$ , and the simplices of  $\Sigma^{(1)}$  are nested sequences of simplices of  $\Sigma$ . Furthermore, note that given any two simplicial complexes  $\Sigma, \Xi$  and a simplicial map  $f : \Sigma \rightarrow \Xi$ , there is a natural simplicial map  $f^{(1)} : \Sigma^{(1)} \rightarrow \Xi^{(1)}$  defined as:

$$f^{(1)}([\sigma_1, \dots, \sigma_p]) := [f(\sigma_1), \dots, f(\sigma_p)], \quad \sigma_1 \subseteq \sigma_2 \subseteq \dots, \sigma_p, \text{ each } \sigma_i \in \Sigma.$$

To see that this is simplicial, note that  $f(\sigma_i) \subseteq f(\sigma_j)$  whenever  $\sigma_i \subseteq \sigma_j$ . As a special case, observe that any inclusion map  $\iota : \Sigma \hookrightarrow \Xi$  induces an inclusion map  $\iota^{(1)} : \Sigma^{(1)} \hookrightarrow \Xi^{(1)}$ .

Given a simplex  $\sigma = [x_0, \dots, x_k]$  in a simplicial complex  $\Sigma$ , one defines the *barycenter* to be the point  $\mathcal{B}(\sigma) := \sum_{i=0}^k \frac{1}{k+1} x_i \in |\Sigma|$ . Then the spaces  $|\Sigma^{(1)}|$  and  $|\Sigma|$  can be identified via a homeomorphism  $\mathcal{E}_{|\Sigma|} : |\Sigma^{(1)}| \rightarrow |\Sigma|$  defined on vertices by  $\mathcal{E}_{|\Sigma|}(\sigma) := \mathcal{B}(\sigma)$  and extended linearly.

Details on the preceding list of definitions can be found in Munkres (1984, §2.14–15, 2.19), Spanier (1994, §3.3–4), and also Barmak (2011, Appendix A). The next proposition follows from the discussions in these references, and is a simple restatement of Barmak (2011, Proposition A.1.5). We provide a proof in the appendix for completeness.

**Proposition 22** (Simplicial approximation to  $\mathcal{E}_\bullet$ ) *Let  $\Sigma$  be a simplicial complex, and let  $\Phi : \Sigma^{(1)} \rightarrow \Sigma$  be a simplicial map such that  $\Phi(\sigma) \in \sigma$  for each  $\sigma \in \Sigma$ . Then  $|\Phi| \simeq \mathcal{E}_{|\Sigma|}$ .*

We now introduce some auxiliary constructions dating back to Dowker (1952) that use the setup stated in Theorem 3. For any nonempty relation  $R \subseteq X \times Y$ , one may define (Dowker 1952, §2) an associated map  $\Phi_{E_R} : E_R^{(1)} \rightarrow E_R$  as follows: first define  $\Phi_{E_R}$  on vertices of  $E_R^{(1)}$  by  $\Phi_{E_R}(\sigma) = s_\sigma$ , where  $s_\sigma$  is the least vertex of  $\sigma$  with respect to the total order. Next, for any simplex  $[\sigma_1, \dots, \sigma_k]$  of  $E_R^{(1)}$ , where  $\sigma_1 \subseteq \dots \subseteq \sigma_k$ , we have  $\Phi_{E_R}(\sigma_i) = s_{\sigma_i} \in \sigma_k$  for all  $1 \leq i \leq k$ . Thus  $[\Phi_{E_R}(\sigma_1), \dots, \Phi_{E_R}(\sigma_k)] = [s_{\sigma_1}, s_{\sigma_2}, \dots, s_{\sigma_k}]$  is a face of  $\sigma_k$ , hence a simplex of  $\Sigma$ . This defines  $\Phi_{E_R}$  as a simplicial map  $E_R^{(1)} \rightarrow E_R$ . This argument also shows that  $\Phi_{E_R}$  is order-reversing: if  $\sigma \subseteq \sigma'$ , then  $\Phi_{E_R}(\sigma) \geq \Phi_{E_R}(\sigma')$ .

**Remark 23** Applying Proposition 22 to the setup above, one sees that  $|\Phi_{E_R}| \simeq \mathcal{E}_{|E_R|}$ . After passing to a second barycentric subdivision  $E_R^{(2)}$  (obtained by taking a barycentric subdivision of  $E_R^{(1)}$ ) and obtaining a map  $\Phi_{E_R^{(1)}} : E_R^{(2)} \rightarrow E_R^{(1)}$ , one also has  $|\Phi_{E_R^{(1)}}| \simeq \mathcal{E}_{|E_R^{(1)}|}$ .

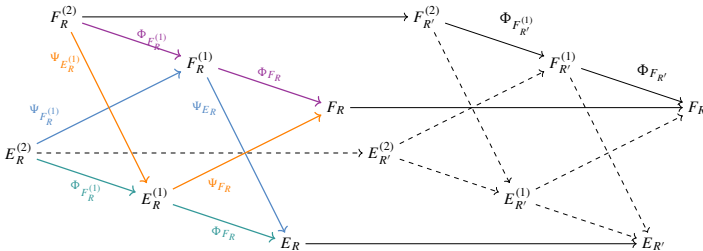
One also defines (Dowker 1952, §3) a simplicial map  $\Psi_{F_R} : E_R^{(1)} \rightarrow F_R$  as follows. Given a vertex  $\sigma = [x_0, \dots, x_k] \in E_R^{(1)}$ , one defines  $\Psi_{F_R}(\sigma) = y_\sigma$ , for some  $y_\sigma \in Y$  such that  $(x_i, y_\sigma) \in R$  for each  $i$ . To see why this vertex map is simplicial, let  $\sigma^{(1)} = [\sigma_0, \dots, \sigma_k]$  be a simplex in  $E_R^{(1)}$ . Let  $x \in \sigma_0$ . Then, because  $\sigma_0 \subseteq \sigma_1 \subseteq \dots \subseteq \sigma_k$ , we automatically have that  $(x, \Psi_{F_R}(\sigma_i)) \in R$ , for each  $i = 0, \dots, k$ . Thus  $\Psi_{F_R}(\sigma^{(1)})$  is a simplex in  $F_R$ . This definition involves a choice of  $y_\sigma$  when writing  $\Psi_{F_R}(\sigma) = y_\sigma$ , but all the maps resulting from such choices are contiguous (Dowker 1952, §3).

The preceding map induces a simplicial map  $\Psi_{F_R^{(1)}} : E_R^{(2)} \rightarrow F_R^{(1)}$  as follows. Given a vertex  $\tau^{(1)} = [\tau_0, \dots, \tau_k] \in E_R^{(2)}$ , i.e. a simplex in  $E_R^{(1)}$ , we define  $\Psi_{F_R^{(1)}}(\tau^{(1)}) := [\Psi_{F_R}(\tau_0), \dots, \Psi_{F_R}(\tau_k)]$ . Since  $\Psi_{F_R}$  is simplicial, this is a simplex in  $F_R$ , i.e. a vertex in  $F_R^{(1)}$ . Thus we have a vertex map  $\Psi_{F_R^{(1)}} : E_R^{(2)} \rightarrow F_R^{(1)}$ . To check that this map is simplicial, let  $\tau^{(2)} = [\tau_0^{(1)}, \dots, \tau_p^{(1)}]$  be a simplex in  $E_R^{(2)}$ . Then  $\tau_0^{(1)} \subseteq \tau_1^{(1)} \subseteq \dots \subseteq \tau_p^{(1)}$ , and because  $\Psi_{F_R}$  is simplicial, we automatically have

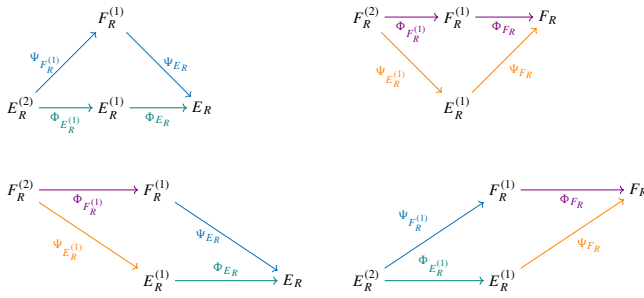
$$\Psi_{F_R}(\tau_0^{(1)}) \subseteq \Psi_{F_R}(\tau_1^{(1)}) \subseteq \dots \subseteq \Psi_{F_R}(\tau_p^{(1)}).$$

Thus  $\Psi_{F_R^{(1)}}(\tau^{(2)})$  is a simplex in  $F_R^{(1)}$ .

**Proof of Theorem 3** We write  $F_R^{(2)}$  to denote the barycentric subdivision of  $F_R^{(1)}$ , and obtain simplicial maps  $\Phi_{F_R^{(1)}} : F_R^{(2)} \rightarrow F_R^{(1)}$ ,  $\Phi_{F_R} : F_R^{(1)} \rightarrow F_R$ ,  $\Psi_{E_R^{(1)}} : F_R^{(2)} \rightarrow E_R^{(1)}$ , and  $\Psi_{F_R} : E_R^{(1)} \rightarrow F_R$  as above. Consider the following diagram:



We proceed by claiming contiguity of the following.



□

**Claim 1** *More specifically:*

- (1)  $\Phi_{E_R} \circ \Phi_{E_R^{(1)}}$  and  $\Psi_{E_R} \circ \Psi_{F_R^{(1)}}$  are contiguous.
- (2)  $\Phi_{F_R} \circ \Phi_{F_R^{(1)}}$  and  $\Psi_{F_R} \circ \Psi_{E_R^{(1)}}$  are contiguous.
- (3)  $\Psi_{E_R} \circ \Phi_{F_R^{(1)}}$  and  $\Phi_{E_R} \circ \Psi_{E_R^{(1)}}$  are contiguous.
- (4)  $\Psi_{F_R} \circ \Phi_{E_R^{(1)}}$  and  $\Phi_{F_R} \circ \Psi_{F_R^{(1)}}$  are contiguous.

Items (1) and (3) appear in the proof of Dowker’s theorem (Dowker 1952, Lemmas 5, 6), and it is easy to see that a symmetric argument shows Items (2) and (4). For completeness, we will verify these items in this paper, but defer this verification to the end of the proof.

By passing to the geometric realization and applying Proposition 7 and Remark 23, we obtain the following from Item (3) of Claim 1:

$$|\Psi_{E_R}| \circ |\Phi_{F_R^{(1)}}| \simeq |\Phi_{E_R}| \circ |\Psi_{E_R^{(1)}}|,$$

$$|\Psi_{E_R}| \circ \mathcal{E}_{|F_R^{(1)}|} \simeq \mathcal{E}_{|E_R|} \circ |\Psi_{E_R^{(1)}}|, \quad (\text{Remark 23})$$

$$\mathcal{E}_{|E_R|}^{-1} \circ |\Psi_{E_R}| \circ \mathcal{E}_{|F_R^{(1)}|} \simeq |\Psi_{E_R^{(1)}}|. \quad (\mathcal{E} \text{ is a homeomorphism, hence invertible})$$

Replacing this term in the expression for Item (2) of Claim 1, we obtain:

$$|\Psi_{F_R}| \circ |\Psi_{E_R^{(1)}}| \simeq |\Phi_{F_R}| \circ |\Phi_{F_R^{(1)}}| \simeq \mathcal{E}_{|F_R|} \circ \mathcal{E}_{|F_R^{(1)}|},$$

$$|\Psi_{F_R}| \circ \mathcal{E}_{|E_R|}^{-1} \circ |\Psi_{E_R}| \circ \mathcal{E}_{|F_R^{(1)}|} \simeq \mathcal{E}_{|F_R|} \circ \mathcal{E}_{|F_R^{(1)}|},$$

$$|\Psi_{F_R}| \circ \mathcal{E}_{|E_R|}^{-1} \circ |\Psi_{E_R}| \circ \mathcal{E}_{|F_R|}^{-1} \simeq \text{id}_{|F_R|}.$$

Similarly, we obtain the following from Item (4) of Claim 1:

$$|\Psi_{F_R}| \circ |\Phi_{E_R^{(1)}}| \simeq |\Phi_{F_R}| \circ |\Psi_{F_R^{(1)}}|, \text{ so } \mathcal{E}_{|F_R|}^{-1} \circ |\Psi_{F_R}| \circ \mathcal{E}_{|E_R^{(1)}|} \simeq |\Psi_{F_R^{(1)}}|.$$

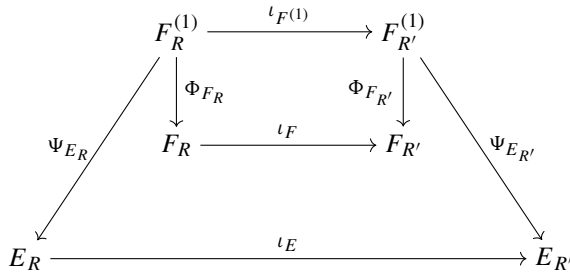
Replacing this term in the expression for Item (1) of Claim 1, we obtain:

$$|\Psi_{E_R}| \circ |\Psi_{F_R^{(1)}}| \simeq |\Phi_{E_R}| \circ |\Phi_{E_R^{(1)}}| \simeq \mathcal{E}_{|E_R|} \circ \mathcal{E}_{|E_R^{(1)}|},$$

$$\begin{aligned}
 |\Psi_{E_R}| \circ \mathcal{E}_{|F_R|}^{-1} \circ |\Psi_{F_R}| \circ \mathcal{E}_{|E_R^{(1)}|} &\simeq \mathcal{E}_{|E_R|} \circ \mathcal{E}_{|E_R^{(1)}|} \\
 |\Psi_{E_R}| \circ \mathcal{E}_{|F_R|}^{-1} \circ |\Psi_{F_R}| \circ \mathcal{E}_{|E_R|}^{-1} &\simeq \text{id}_{|E_R|}
 \end{aligned}$$

Define  $\Gamma_{|E_R|} : |F_R| \rightarrow |E_R|$  by  $\Gamma_{|E_R|} := |\Psi_{E_R}| \circ \mathcal{E}_{|F_R|}^{-1}$ . Then  $\Gamma_{|E_R|}$  is a homotopy equivalence, with inverse given by  $|\Psi_{F_R}| \circ \mathcal{E}_{|E_R|}^{-1}$ . This shows that  $|F_R| \simeq |E_R|$ , for any nonempty relation  $R \subseteq X \times Y$ .

Next we need to show that  $\Gamma_{|E_R|}$  commutes with the canonical inclusion. Consider the following diagram, where the  $\iota_\bullet$  maps denote the respective canonical inclusions (cf. Definition 2):



**Claim 2**  $\iota_E \circ \Psi_{E_R}$  and  $\Psi_{E_{R'}} \circ \iota_{F^{(1)}}$  are contiguous.

**Claim 3**  $\iota_F \circ \Phi_{F_R}$  and  $\Phi_{F_{R'}} \circ \iota_{F^{(1)}}$  are contiguous.

Suppose Claim 3 is true. Then, upon passing to geometric realizations, we have:

$$\begin{aligned}
 |\iota_F| \circ \mathcal{E}_{|F_R|} &\simeq |\iota_F| \circ |\Phi_{F_R}| \simeq |\Phi_{F_{R'}}| \circ |\iota_{F^{(1)}}| \simeq \mathcal{E}_{|F_{R'}|} \circ |\iota_{F^{(1)}}|, \\
 \mathcal{E}_{|F_{R'}|}^{-1} \circ |\iota_F| \circ \mathcal{E}_{|F_R|} &\simeq |\iota_{F^{(1)}}|.
 \end{aligned}$$

Suppose also that Claim 2 is true. Then we have:

$$\begin{aligned}
 |\Psi_{E_{R'}}| \circ |\iota_{F^{(1)}}| &\simeq |\iota_E| \circ |\Psi_{E_R}|, \\
 |\Psi_{E_{R'}}| \circ \mathcal{E}_{|F_{R'}|}^{-1} \circ |\iota_F| \circ \mathcal{E}_{|F_R|} &\simeq |\iota_E| \circ |\Psi_{E_R}|, \\
 |\Psi_{E_{R'}}| \circ \mathcal{E}_{|F_{R'}|}^{-1} \circ |\iota_F| &\simeq |\iota_E| \circ |\Psi_{E_R}| \circ \mathcal{E}_{|F_R|}^{-1}, \text{ i.e.} \\
 \Gamma_{|E_{R'}|} \circ |\iota_F| &\simeq |\iota_E| \circ \Gamma_{|E_R|}.
 \end{aligned}$$

This proves the theorem. It only remains to prove the various claims.

**Proof of Claim 1** In proving Claim 1, we supply the proofs of Items (2) and (4). These arguments are adapted from Dowker (1952, Lemmas 1, 5, and 6), where the proofs of Items (1) and (3) appeared.

For Item (2), let  $\tau^{(2)} = [\tau_0^{(1)}, \dots, \tau_k^{(1)}]$  be a simplex in  $F_R^{(2)}$ , where  $\tau_0^{(1)} \subseteq \dots \subseteq \tau_k^{(1)}$  is a chain of simplices in  $F_R^{(1)}$ . By the order-reversing property of the map  $\Phi_{F_R^{(1)}}$ , we have that  $\Phi_{F_R^{(1)}}(\tau_0^{(1)}) \supseteq \Phi_{F_R^{(1)}}(\tau_i^{(1)})$  for each  $i = 0, \dots, k$ . Define

$x := \Psi_{E_R}(\Phi_{F_R^{(1)}}(\tau_0^{(1)}))$ . Then  $(x, y) \in R$  for each  $y \in \Phi_{F_R^{(1)}}(\tau_0^{(1)})$ . But we also have  $(x, \Phi_{F_R}(\Phi_{F_R^{(1)}}(\tau_i^{(1)}))) \in R$  for each  $i = 0, \dots, k$ , because  $\Phi_{F_R}(\Phi_{F_R^{(1)}}(\tau_i^{(1)})) \in \Phi_{F_R^{(1)}}(\tau_i^{(1)}) \subseteq \Phi_{F_R}(\tau_0^{(1)})$  for each  $i = 0, \dots, k$ .

Next let  $0 \leq i \leq k$ . For each  $\tau \in \tau_i^{(1)}$ , we have  $\Psi_{E_R}(\tau) \in \Psi_{E_R^{(1)}}(\tau_i^{(1)})$  (by the definition of  $\Psi_{E_R^{(1)}}$ ). Because  $\Phi_{F_R^{(1)}}(\tau_0^{(1)}) \in \tau_0^{(1)} \subseteq \tau_i^{(1)}$ , we then have  $x = \Psi_{E_R}(\Phi_{F_R^{(1)}}(\tau_0^{(1)})) \in \Psi_{E_R^{(1)}}(\tau_i^{(1)})$ , which is a vertex of  $E_R^{(1)}$  or alternatively a simplex of  $E_R$ . But then, by definition of  $\Psi_{F_R}$ , we have that  $(x, \Psi_{F_R}(\Psi_{E_R^{(1)}}(\tau_i^{(1)}))) \in R$ . This holds for each  $0 \leq i \leq k$ . Since  $\tau^{(2)}$  was arbitrary, this shows that  $\Phi_{F_R} \circ \Phi_{F_R^{(1)}}$  and  $\Psi_{F_R} \circ \Psi_{E_R^{(1)}}$  are contiguous.

For Item (4), let  $\sigma^{(2)} = [\sigma_0^{(1)}, \dots, \sigma_k^{(1)}]$  be a simplex in  $E_R^{(2)}$ . Let  $0 \leq i \leq k$ . Then  $\sigma_0^{(1)} \subseteq \dots \subseteq \sigma_k^{(1)}$ , and  $\Phi_{E_R^{(1)}}(\sigma_i^{(1)}) \in \sigma_i^{(1)} \subseteq \sigma_k^{(1)}$ . So  $\Psi_{F_R}(\Phi_{E_R^{(1)}}(\sigma_i^{(1)})) \in \Psi_{F_R^{(1)}}(\sigma_k^{(1)})$ . On the other hand, we have  $\Psi_{F_R^{(1)}}(\sigma_i^{(1)}) \subseteq \Psi_{F_R^{(1)}}(\sigma_k^{(1)})$ . Then  $\Phi_{F_R}(\Psi_{F_R^{(1)}}(\sigma_i^{(1)})) \in \Psi_{F_R^{(1)}}(\sigma_i^{(1)}) \subseteq \Psi_{F_R^{(1)}}(\sigma_k^{(1)})$ . Since  $i$  was arbitrary, this shows that  $\Psi_{F_R} \circ \Phi_{E_R^{(1)}}$  and  $\Phi_{F_R} \circ \Psi_{F_R^{(1)}}$  both map the vertices of  $\sigma^{(2)}$  to the simplex  $\Psi_{F_R^{(1)}}(\sigma_k^{(1)})$ , hence are contiguous. This concludes the proof of the claim.  $\square$

**Proof of Claim 2** Let  $\tau^{(1)} = [\tau_0, \tau_1, \dots, \tau_k] \in F_R^{(1)}$ , where  $\tau_0 \subseteq \tau_1 \subseteq \dots \subseteq \tau_k$  is a chain of simplices in  $F_R$ . Then  $\iota_{F^{(1)}}(\tau^{(1)}) = \tau^{(1)}$ , and  $\Psi_{E_{R'}}(\tau^{(1)}) = [x_{\tau_0}, \dots, x_{\tau_k}]$ , for some choice of  $x_{\tau_i}$  terms. Also we have  $\iota_E \circ \Psi_{E_R}(\tau^{(1)}) = [x'_{\tau_0}, \dots, x'_{\tau_k}]$  for some other choice of  $x'_{\tau_i}$  terms. For contiguity, we need to show that

$$[x_{\tau_0}, \dots, x_{\tau_k}, x'_{\tau_0}, \dots, x'_{\tau_k}] \in E_{R'}.$$

But this is easy to see: letting  $y \in \tau_0$ , we have  $\{(x_{\tau_0}, y), \dots, (x_{\tau_k}, y), (x'_{\tau_0}, y), \dots, (x'_{\tau_k}, y)\} \subseteq R$ . Since  $\tau^{(1)}$  was arbitrary, it follows that we have contiguity.  $\square$

**Proof of Claim 3** Let  $\tau^{(1)} = [\tau_0, \tau_1, \dots, \tau_k] \in F_R^{(1)}$ , where  $\tau_0 \subseteq \tau_1 \subseteq \dots \subseteq \tau_k$  is a chain of simplices in  $F_R$ . Then  $\Phi_{F_R}(\tau_i) \in \tau_k$  for each  $0 \leq i \leq k$ . Thus  $\iota_F \circ \Phi_{F_R}(\tau^{(1)})$  is a face of  $\tau_k$ . Similarly,  $\Phi_{F_{R'}} \circ \iota_{F^{(1)}}(\tau^{(1)})$  is also a face of  $\tau_k$ . Since  $\tau^{(1)}$  was an arbitrary simplex of  $F_R^{(1)}$ , it follows that  $\iota_F \circ \Phi_{F_R}$  and  $\Phi_{F_{R'}} \circ \iota_{F^{(1)}}$  are contiguous.  $\square$

### 5.2 The equivalence between the finite FDT and the simplicial FNTs

In this section, we present our answer to Question 1. We begin with a weaker formulation of Theorem 3 and some simplicial Functorial Nerve Theorems.

**Theorem 24** (The finite FDT) *Let  $X, Y$  be two totally ordered sets, and without loss of generality, suppose  $X$  is finite. Let  $R \subseteq R' \subseteq X \times Y$  be two nonempty relations, and*



let  $E_R, F_R, E_{R'}, F_{R'}$  be their associated simplicial complexes (as in Theorem 3). Then there exist homotopy equivalences  $\Gamma_{|E_R|} : |F_R| \rightarrow |E_R|$  and  $\Gamma_{|E_{R'}|} : |F_{R'}| \rightarrow |E_{R'}|$  that commute up to homotopy with the canonical inclusions.

The finite FDT (Theorem 24) is an immediate consequence of the general FDT (Theorem 3).

**Definition 3** Let  $\mathcal{A} = \{A_i\}_{i \in I}$  be a family of nonempty sets indexed by  $I$ . The nerve of  $\mathcal{A}$  is the simplicial complex  $\mathcal{N}(\mathcal{A}) := \{\sigma \in \text{pow}(I) : \sigma \text{ is finite, nonempty, and } \cap_{i \in \sigma} A_i \neq \emptyset\}$ .

**Definition 4** (Covers of simplices and subcomplexes) Let  $\Sigma$  be a simplicial complex. Then a collection of subcomplexes  $\mathcal{A}_\Sigma = \{\Sigma_i\}_{i \in I}$  is said to be a cover of subcomplexes for  $\Sigma$  if  $\Sigma = \cup_{i \in I} \Sigma_i$ . Furthermore,  $\mathcal{A}_\Sigma$  is said to be a cover of simplices if each  $\Sigma_i \in \mathcal{A}_\Sigma$  has the property that  $\Sigma_i = \text{pow}(V(\Sigma_i))$ . In this case, each  $\Sigma_i$  has precisely one top-dimensional simplex, consisting of the vertex set  $V(\Sigma_i)$ .

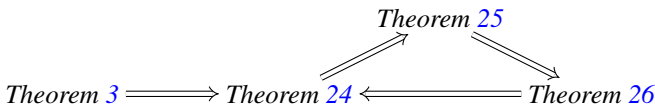
We present two simplicial formulations of the Functorial Nerve Theorem that turn out to be equivalent; the statements differ in that one is about covers of simplices and the other is about covers of subcomplexes.

**Theorem 25** (Functorial nerve I) Let  $\Sigma \subseteq \Sigma'$  be two simplicial complexes, and let  $\mathcal{A}_\Sigma = \{\Sigma_i\}_{i \in I}, \mathcal{A}_{\Sigma'} = \{\Sigma'_i\}_{i \in I'}$  be finite covers of simplices for  $\Sigma$  and  $\Sigma'$  such that  $I \subseteq I'$  and  $\Sigma_i \subseteq \Sigma'_i$  for each  $i \in I$ . In particular,  $\text{card}(I') < \infty$ . Suppose that for each finite subset  $\sigma \subseteq I'$ , the intersection  $\cap_{i \in \sigma} \Sigma'_i$  is either empty or contractible (and likewise for  $\cap_{i \in \sigma} \Sigma_i$ ). Then  $|\Sigma| \simeq |\mathcal{N}(\mathcal{A}_\Sigma)|$  and  $|\Sigma'| \simeq |\mathcal{N}(\mathcal{A}_{\Sigma'})|$ , via maps that commute up to homotopy with the canonical inclusions.

**Theorem 26** (Functorial nerve II) The statement of Theorem 25 holds even if  $\mathcal{A}_\Sigma$  and  $\mathcal{A}_{\Sigma'}$  are covers of subcomplexes. Explicitly, the statement is as follows. Let  $\Sigma \subseteq \Sigma'$  be two simplicial complexes, and let  $\mathcal{A}_\Sigma = \{\Sigma_i\}_{i \in I}, \mathcal{A}_{\Sigma'} = \{\Sigma'_i\}_{i \in I'}$  be finite covers of subcomplexes for  $\Sigma$  and  $\Sigma'$  such that  $I \subseteq I'$  and  $\Sigma_i \subseteq \Sigma'_i$  for each  $i \in I$ . In particular,  $\text{card}(I') < \infty$ . Suppose that for each finite subset  $\sigma \subseteq I'$ , the intersection  $\cap_{i \in \sigma} \Sigma'_i$  is either empty or contractible (and likewise for  $\cap_{i \in \sigma} \Sigma_i$ ). Then  $|\Sigma| \simeq |\mathcal{N}(\mathcal{A}_\Sigma)|$  and  $|\Sigma'| \simeq |\mathcal{N}(\mathcal{A}_{\Sigma'})|$ , via maps that commute up to homotopy with the canonical inclusions.

The following result summarizes our answer to Question 1.

**Theorem 27** (Equivalence) The finite FDT, the FNT I, and the FNT II are all equivalent. Moreover, all of these results are implied by the FDT, as below:



We present the proof of Theorem 27 over the course of the next few subsections.

**Remark 28** By virtue of Theorem 27, we will write simplicial FNT to mean either of the FNT I or FNT II.

**Theorem 24 implies Theorem 25**

**Proof of Theorem 25** Let  $V, V'$  denote the vertex sets of  $\Sigma, \Sigma'$ , respectively. We define the relations  $R \subseteq V \times I$  and  $R' \subseteq V' \times I'$  as follows:  $(v, i) \in R \iff v \in \Sigma_i$  and  $(v', i') \in R' \iff v' \in \Sigma'_i$ . Then  $R \subseteq R'$ , the set  $I'$  is finite by assumption, and so we are in the setting of the finite FDT (Theorem 24) (perhaps invoking the Axiom of Choice to obtain the total order on  $V'$ ). It suffices to show that  $E_R = \Sigma, E_{R'} = \Sigma', F_R = \mathcal{N}(\mathcal{A}_\Sigma)$ , and  $F_{R'} = \mathcal{N}(\mathcal{A}_{\Sigma'})$ , where  $E_R, E_{R'}, F_R, F_{R'}$  are as defined in Theorem 3.

First we claim the  $E_R = \Sigma$ . By the definitions of  $R$  and  $E_R$ , we have  $E_R = \{\sigma \subseteq V : \exists i \in I, (v, i) \in R \forall v \in \sigma\} = \{\sigma \subseteq V : \exists i \in I, v \in \Sigma_i \forall v \in \sigma\}$ . Let  $\sigma \in E_R$ , and let  $i \in I$  be such that  $v \in \Sigma_i$  for all  $v \in \sigma$ . Then  $\sigma \subseteq V(\Sigma_i)$ , and since  $\Sigma_i = \text{pow}(V(\Sigma_i))$  by the assumption about covers of simplices, we have  $\sigma \in \Sigma_i \subseteq \Sigma$ . Thus  $E_R \subseteq \Sigma$ . Conversely, let  $\sigma \in \Sigma$ . Then  $\sigma \in \Sigma_i$  for some  $i$ . Thus for all  $v \in \sigma$ , we have  $(v, i) \in R$ . It follows that  $\sigma \in E_R$ . This shows  $E_R = \Sigma$ . The proof that  $E_{R'} = \Sigma'$  is analogous.

Next we claim that  $F_R = \mathcal{N}(\mathcal{A}_\Sigma)$ . By the definition of  $F_R$ , we have  $F_R = \{\tau \subseteq I : \exists v \in V, (v, i) \in R \forall i \in \tau\}$ . Let  $\tau \in F_R$ , and let  $v \in V$  be such that  $(v, i) \in R$  for each  $i \in \tau$ . Then  $\cap_{i \in \tau} \Sigma_i \neq \emptyset$ , and so  $\tau \in \mathcal{N}(\mathcal{A}_\Sigma)$ . Conversely, let  $\tau \in \mathcal{N}(\mathcal{A}_\Sigma)$ . Then  $\cap_{i \in \tau} \Sigma_i \neq \emptyset$ , so there exists  $v \in V$  such that  $v \in \Sigma_i$  for each  $i \in \tau$ . Thus  $\sigma \in F_R$ . This shows  $F_R = \mathcal{N}(\mathcal{A}_\Sigma)$ . The case for  $R'$  is analogous.

An application of Theorem 24 now completes the proof. □

**Theorem 26 implies Theorem 24**

**Proof** Let  $X$  and  $Y$  be two sets, and suppose  $X$  is finite. Let  $R \subseteq R' \subseteq X \times Y$  be two relations. Consider the simplicial complexes  $E_R, F_R, E_{R'}, F_{R'}$  as defined in Theorem 3. Let  $V_R := V(E_R)$ . For each  $x \in V_R$ , define  $A_x := \{\tau \in F_R : (x, y) \in R \text{ for all } y \in \tau\}$ . Then  $A_x$  is a subcomplex of  $F_R$ . Furthermore,  $\cup_{x \in V_R} A_x = F_R$ . To see this, let  $\tau \in F_R$ . Then there exists  $x \in X$  such that  $(x, y) \in R$  for all  $y \in \tau$ , and so  $\tau \in A_x$ .

Let  $\mathcal{A} := \{A_x : x \in V_R\}$ . We have seen that  $\mathcal{A}$  is a cover of subcomplexes for  $F_R$ . It is finite because the indexing set  $V_R$  is a subset of  $X$ , which is finite by assumption. Next we claim that  $\mathcal{N}(\mathcal{A}) = E_R$ . Let  $\sigma \in E_R$ . Then there exists  $y \in Y$  such that  $(x, y) \in R$  for all  $x \in \sigma$ . Thus  $\cap_{x \in \sigma} A_x \neq \emptyset$ , and so  $\sigma \in \mathcal{N}(\mathcal{A})$ . Conversely, let  $\sigma \in \mathcal{N}(\mathcal{A})$ . Then  $\cap_{x \in \sigma} A_x \neq \emptyset$ , and so there exists  $y \in Y$  such that  $(x, y) \in R$  for all  $x \in \sigma$ . Thus  $\sigma \in E_R$ .

Next we check that nonempty finite intersections of elements in  $\mathcal{A}$  are contractible. Let  $\sigma \in \mathcal{N}(\mathcal{A}) = E_R$ . Let  $V_\sigma := \cap_{x \in \sigma} V(A_x) \subseteq V(F_R)$ . We claim that  $\cap_{x \in \sigma} A_x = \text{pow}(V_\sigma)$ , i.e. that the intersection is a full simplex in  $F_R$ , hence contractible. The inclusion  $\cap_{x \in \sigma} A_x \subseteq \text{pow}(V_\sigma)$  is clear, so we show the reverse inclusion. Let  $\tau \in \text{pow}(V_\sigma)$ , and let  $y \in \tau$ . Then  $y \in \cap_{x \in \sigma} A_x$ , so  $(x, y) \in R$  for each  $x \in \sigma$ . This holds for each  $y \in \tau$ , so it follows that  $\tau \in \cap_{x \in \sigma} A_x$ . Thus  $\cap_{x \in \sigma} A_x = \text{pow}(V_\sigma)$ . We remark that this also shows that  $\mathcal{A}$  is a cover of simplices for  $F_R$ .

Now for each  $x \in V(E_{R'})$ , define  $A'_x := \{\tau \in F_{R'} : (x, y) \in R' \text{ for all } y \in \tau\}$ . Also define  $\mathcal{A}' := \{A'_x : x \in V(E_{R'})\}$ . The same argument shows that  $\mathcal{A}'$  is a finite cover of subcomplexes (in particular, a cover of simplices) for  $F_{R'}$  with all finite intersections either empty or contractible, and that  $E_{R'} = \mathcal{N}(\mathcal{A}')$ . An application of Theorem 26 now shows that  $|E_R| \simeq |F_R|$  and  $|E_{R'}| \simeq |F_{R'}|$ , via maps that commute up to homotopy with the inclusions  $|E_R| \hookrightarrow |E_{R'}|$  and  $|F_R| \hookrightarrow |F_{R'}|$ .  $\square$

**Theorem 25 implies Theorem 26**

We lead with some remarks about the ideas involved in this proof. Theorem 26 is a functorial statement in the sense that it is about an arbitrary inclusion  $\Sigma \subseteq \Sigma'$ . Restricting the statement to just  $\Sigma$  would lead to a non-functorial statement. A proof of this non-functorial statement, via a non-functorial analogue of Theorem 25, can be obtained using techniques presented in Björner et al. (1985) (see also Kozlov 2007, Theorem 15.24). We strengthen these techniques to our functorial setting and thus obtain a proof of Theorem 26 via Theorem 25.

We first present a lemma related to barycentric subdivisions and several lemmas about gluings and homotopy equivalences. These will be used in proving Theorem 26.

**Definition 5 (Induced subcomplex)** Let  $\Sigma$  be a simplicial complex, and let  $\Delta$  be a subcomplex. Then  $\Delta$  is an *induced subcomplex* if  $\Delta = \Sigma \cap \text{pow}(V(\Delta))$ .

**Lemma 29** *Let  $\Sigma$  be a simplicial complex, and let  $\Delta$  be a subcomplex. Then  $\Delta^{(1)}$  is an induced subcomplex of  $\Sigma^{(1)}$ , i.e.  $\Delta^{(1)} = \Sigma^{(1)} \cap \text{pow}(V(\Delta^{(1)}))$ .*

**Proof** Let  $\sigma$  be a simplex of  $\Delta^{(1)}$ . Then  $\sigma$  belongs to  $\Sigma^{(1)}$ , and also to the full simplex  $\text{pow}(V(\Delta^{(1)}))$ . Thus  $\Delta^{(1)} \subseteq \Sigma^{(1)} \cap \text{pow}(V(\Delta^{(1)}))$ . Conversely, let  $\sigma \in \Sigma^{(1)} \cap \text{pow}(V(\Delta^{(1)}))$ . Since  $\sigma \in \Sigma^{(1)}$ , we can write  $\sigma = [\tau_0, \dots, \tau_k]$ , where  $\tau_0 \subseteq \dots \subseteq \tau_k$ . Since  $\sigma \in \text{pow}(V(\Delta^{(1)}))$  and the vertices of  $\Delta^{(1)}$  are simplices of  $\Delta$ , we also know that each  $\tau_i$  is a simplex of  $\Delta$ . Thus  $\sigma \in \Delta^{(1)}$ . The equality follows.  $\square$

**Lemma 30 (Carrier Lemma, Björner et al. 1985 §4)** *Let  $X$  be a topological space, and let  $\Sigma$  be a simplicial complex. Also let  $f, g : X \rightarrow |\Sigma|$  be any two continuous maps such that  $f(x), g(x)$  belong to the same simplex of  $|\Sigma|$ , for any  $x \in X$ . Then  $f \simeq g$ .*

**Lemma 31 (Gluing Lemma, see Lemmas 4.2, 4.7, 4.9, Björner et al. 1985)** *Let  $\Sigma$  be a simplicial complex, and let  $U \subseteq V(\Sigma)$ . Suppose  $|\Sigma \cap \text{pow}(U)|$  is contractible. Then there exists a homotopy equivalence  $\varphi : |\Sigma \cup \text{pow}(U)| \rightarrow |\Sigma|$ .*

The Gluing and Carrier Lemmas presented above are classical. We provide full details for the Gluing lemma inside the proof of the following functorial generalization of Lemma 31.

**Lemma 32 (Functorial Gluing Lemma)** *Let  $\Sigma \subseteq \Sigma'$  be two simplicial complexes. Also let  $U \subseteq V(\Sigma)$  and  $U' \subseteq V(\Sigma')$  be such that  $U \subseteq U'$ . Suppose  $|\Sigma \cap \text{pow}(U)|$  and  $|\Sigma' \cap \text{pow}(U')|$  are contractible. Then,*

- (1) There exists a homotopy equivalence  $\varphi : |\Sigma \cup \text{pow}(U)| \rightarrow |\Sigma|$  such that  $\varphi(x)$  and  $\text{id}_{|\Sigma \cup \text{pow}(U)|}(x)$  belong to the same simplex of  $|\Sigma \cup \text{pow}(U)|$  for each  $x \in |\Sigma \cup \text{pow}(U)|$ . Furthermore, the homotopy inverse is given by the inclusion  $\iota : |\Sigma| \hookrightarrow |\Sigma \cup \text{pow}(U)|$ .
- (2) Given a homotopy equivalence  $\varphi : |\Sigma \cup \text{pow}(U)| \rightarrow |\Sigma|$  as above, there exists a homotopy equivalence  $\varphi' : |\Sigma' \cup \text{pow}(U')| \rightarrow |\Sigma'|$  such that  $\varphi'|_{|\Sigma \cup \text{pow}(U)|} = \varphi$ , and  $\varphi'(x)$  and  $\text{id}_{|\Sigma' \cup \text{pow}(U')|}(x)$  belong to the same simplex of  $|\Sigma' \cup \text{pow}(U')|$  for each  $x \in |\Sigma' \cup \text{pow}(U')|$ . Furthermore, the homotopy inverse is given by the inclusion  $\iota' : |\Sigma'| \hookrightarrow |\Sigma' \cup \text{pow}(U')|$ .

**Proof of Lemma 32** The proof uses this fact: any continuous map of an  $n$ -sphere  $\mathbb{S}^n$  into a contractible space  $Y$  can be continuously extended to a mapping of the  $(n + 1)$ -disk  $\mathbb{D}^{n+1}$  into  $Y$ , where  $\mathbb{D}^{n+1}$  has  $\mathbb{S}^n$  as its boundary (Spanier 1994, p. 27). First we define  $\varphi$ . On  $|\Sigma|$ , define  $\varphi$  to be the identity. Next let  $\sigma$  be a minimal simplex in  $|\text{pow}(U) \setminus \Sigma|$ . By minimality, the boundary of  $\sigma$  (denoted  $\text{Bd}(\sigma)$ ) belongs to  $|\Sigma \cap \text{pow}(U)|$ , and  $|\Sigma|$  in particular. Thus  $\varphi$  is defined on  $\text{Bd}(\sigma)$ , which is an  $n$ -sphere for some  $n \geq 0$ . Furthermore,  $\varphi$  maps  $\text{Bd}(\sigma)$  into the contractible space  $|\Sigma \cap \text{pow}(U)|$ . Then we use the aforementioned fact to extend  $\varphi$  continuously to all of  $\sigma$  so that  $\varphi$  maps  $\sigma$  into  $|\Sigma \cap \text{pow}(U)|$ . Furthermore, both  $\text{id}_{|\Sigma \cup \text{pow}(U)|}(\sigma) = \sigma$  and  $\varphi(\sigma)$  belong to the simplex  $|\text{pow}(U)|$ . By iterating this procedure, we obtain a retraction  $\varphi : |\Sigma \cup \text{pow}(U)| \rightarrow |\Sigma|$  such that  $\varphi(x)$  and  $x$  belong to the same simplex in  $|\Sigma \cup \text{pow}(U)|$ , for each  $x \in |\Sigma \cup \text{pow}(U)|$ . Thus  $\varphi$  is homotopic to  $\text{id}_{|\Sigma \cup \text{pow}(U)|}$  by Lemma 30. Thus we have a homotopy equivalence:

$$\text{id}_{|\Sigma|} = \varphi \circ \iota, \quad \iota \circ \varphi \simeq \text{id}_{|\Sigma \cup \text{pow}(U)|}. \quad (\text{here } \iota := \iota_{|\Sigma| \hookrightarrow |\Sigma \cup \text{pow}(U)|})$$

For the second part of the proof, suppose that a homotopy equivalence  $\varphi : |\Sigma \cup \text{pow}(U)| \rightarrow |\Sigma|$  as above is provided. We need to extend  $\varphi$  to obtain  $\varphi'$ . Define  $\varphi'$  to be equal to  $\varphi$  on  $|\Sigma \cup \text{pow}(U)|$ , and equal to the identity on  $G := |\Sigma'| \setminus |\Sigma \cup \text{pow}(U)|$ . Let  $\sigma$  be a minimal simplex in  $|\text{pow}(U') \setminus G|$ . Then by minimality,  $\text{Bd}(\sigma)$  belongs to  $|\Sigma' \cap \text{pow}(U')|$ . As before, we have  $\varphi'$  mapping  $\text{Bd}(\sigma)$  into the contractible space  $|\Sigma' \cap \text{pow}(U')|$ , and we extend  $\varphi'$  continuously to a map of  $\sigma$  into  $|\Sigma' \cap \text{pow}(U')|$ . Once again,  $\text{id}_{|\Sigma' \cup \text{pow}(U')|}(x)$  and  $\varphi'(x)$  belong to the same simplex  $|\text{pow}(U')|$ , for all  $x \in \sigma$ . Iterating this procedure gives a continuous map  $\varphi' : |\Sigma' \cup \text{pow}(U')| \rightarrow |\Sigma'|$ . This map is not necessarily a retraction, because there may be a simplex  $\sigma \in |\Sigma \cup \text{pow}(U)| \cap |\Sigma'|$  on which  $\varphi'$  is not the identity. However, it still holds that  $\varphi'$  is continuous, and that  $x, \varphi'(x)$  get mapped to the same simplex for each  $x \in |\Sigma' \cup \text{pow}(U')|$ . Thus Lemma 30 still applies to show that  $\varphi'$  is homotopic to  $\text{id}_{|\Sigma' \cup \text{pow}(U')|}$ .

We write  $\iota'$  to denote the inclusion  $\iota' : |\Sigma'| \hookrightarrow |\Sigma' \cup \text{pow}(U')|$ . By the preceding work, we have  $\iota' \circ \varphi' \simeq \text{id}_{|\Sigma' \cup \text{pow}(U')|}$ . Next let  $x \in |\Sigma'|$ . Then either  $x \in |\Sigma'| \cap |\Sigma \cup \text{pow}(U)|$ , or  $x \in G$ . In the first case, we know that  $\varphi'(x) = \varphi(x)$  and  $\text{id}_{|\Sigma'|}(x) = \text{id}_{|\Sigma \cup \text{pow}(U)|}(x)$  belong to the same simplex of  $|\Sigma \cup \text{pow}(U)|$  by the assumption on  $\varphi$ . In the second case, we know that  $\varphi'(x) = x = \text{id}_{|\Sigma'|}(x)$ . Thus for any  $x \in |\Sigma'|$ , we know that  $\varphi'(x)$  and  $\text{id}_{|\Sigma'|}(x)$  belong to the same simplex in  $|\Sigma' \cup \text{pow}(U')|$ . By Lemma 30, we then have  $\varphi'|_{|\Sigma'|} \simeq \text{id}_{|\Sigma'|}$ . Thus  $\varphi' \circ \iota' \simeq \text{id}_{|\Sigma'|}$ . This shows that  $\varphi'$  is the necessary homotopy equivalence.  $\square$

Now we present the proof of Theorem 26.

**Notation** Let  $I$  be an ordered set. For any subset  $J \subseteq I$ , we write  $(J)$  to denote the sequence  $(j_1, j_2, j_3, \dots)$ , where the ordering is inherited from the ordering on  $I$ .

**Proof of Theorem 26** The first step is to functorially deform  $\mathcal{A}_\Sigma$  and  $\mathcal{A}_{\Sigma'}$  into covers of simplices while still preserving all associated homotopy types. Then we will be able to apply Theorem 25. We can assume by Lemma 29 that each subcomplex  $\Sigma_i$  is induced, and likewise for each  $\Sigma'_i$ . We start by fixing an enumeration  $I' = \{l_1, l_2, \dots\}$ . Thus  $I'$  becomes an ordered set.

**Passing to covers of simplices** We now define some inductive constructions. In what follows, we will define complexes denoted  $\Sigma^\bullet, \Sigma'^\bullet$  obtained by “filling in”  $\Sigma$  and  $\Sigma'$  while preserving homotopy equivalence, as well as covers of these larger complexes denoted  $\Sigma_{\star, \bullet}, \Sigma'_{\star, \bullet}$ . First define:

$$\Sigma^{(l_1)} := \begin{cases} \Sigma \cup \text{pow}(V(\Sigma_{l_1})) & : \text{if } l_1 \in I \\ \Sigma & : \text{otherwise.} \end{cases}$$

$$\Sigma'^{(l_1)} := \Sigma' \cup \text{pow}(V(\Sigma'_{l_1})).$$

Next, for all  $i \in I$ , define

$$\Sigma_{i, (l_1)} := \begin{cases} \Sigma_i \cup \text{pow}(V(\Sigma_i) \cap V(\Sigma_{l_1})) & : \text{if } l_1 \in I \\ \Sigma_i & : \text{otherwise.} \end{cases}$$

And for all  $i \in I'$ , define

$$\Sigma'_{i, (l_1)} := \Sigma'_i \cup \text{pow}(V(\Sigma'_i) \cap V(\Sigma'_{l_1})).$$

Now by induction, suppose  $\Sigma^{(l_1, \dots, l_n)}$  and  $\Sigma_{i, (l_1, \dots, l_n)}$  are defined for all  $i \in I$ . Also suppose  $\Sigma'^{(l_1, \dots, l_n)}$  and  $\Sigma'_{i, (l_1, \dots, l_n)}$  are defined for all  $i \in I'$ . Then we define:

$$\Sigma^{(l_1, \dots, l_n, l_{n+1})} := \begin{cases} \Sigma^{(l_1, \dots, l_n)} \cup \text{pow}(V(\Sigma_{l_{n+1}, (l_1, \dots, l_n)})) & : \text{if } l_{n+1} \in I \\ \Sigma^{(l_1, \dots, l_n)} & : \text{otherwise.} \end{cases}$$

$$\Sigma'^{(l_1, \dots, l_n, l_{n+1})} := \Sigma'^{(l_1, \dots, l_n)} \cup \text{pow}(V(\Sigma'_{l_{n+1}, (l_1, \dots, l_n)})).$$

For all  $i \in I$ , we have

$$\Sigma_{i, (l_1, l_2, \dots, l_{n+1})} := \begin{cases} \Sigma_{i, (l_1, l_2, \dots, l_n)} \cup \text{pow}(V(\Sigma_{i, (l_1, l_2, \dots, l_n)} \cap V(\Sigma_{l_{n+1}, (l_1, l_2, \dots, l_n)}))) & : \text{if } l_{n+1} \in I \\ \Sigma_{i, (l_1, l_2, \dots, l_n)} & : \text{otherwise.} \end{cases}$$

And for all  $i \in I'$ , we have

$$\Sigma'_{i, (l_1, l_2, \dots, l_{n+1})} := \Sigma'_{i, (l_1, l_2, \dots, l_n)} \cup \text{pow}(V(\Sigma'_{i, (l_1, l_2, \dots, l_n)} \cap V(\Sigma'_{l_{n+1}, (l_1, l_2, \dots, l_n)}))).$$

Finally, for any  $n \leq \text{card}(I')$ , we define  $\mathcal{A}_{\Sigma, (l_1, \dots, l_n)} := \{\Sigma_{i, (l_1, \dots, l_n)} : i \in I\}$  and  $\mathcal{A}_{\Sigma', (l_1, \dots, l_n)} := \{\Sigma'_{i, (l_1, \dots, l_n)} : i \in I'\}$ . We will show that these are covers of  $\Sigma^{(l_1, l_2, \dots, l_n)}$  and  $\Sigma'^{(l_1, l_2, \dots, l_n)}$ , respectively.

The next step is to prove by induction that for any  $n \leq \text{card}(I')$ , we have  $|\Sigma| \simeq |\Sigma^{(l_1, l_2, \dots, l_n)}|$  and  $|\Sigma'| \simeq |\Sigma'^{(l_1, l_2, \dots, l_n)}|$ , that  $\mathcal{N}(\mathcal{A}_{\Sigma}) = \mathcal{N}(\mathcal{A}_{\Sigma, (l_1, l_2, \dots, l_n)})$  and  $\mathcal{N}(\mathcal{A}_{\Sigma'}) = \mathcal{N}(\mathcal{A}_{\Sigma', (l_1, l_2, \dots, l_n)})$ , and that nonempty finite intersections of the new covers  $\mathcal{A}_{\Sigma, (l_1, l_2, \dots, l_n)}$ ,  $\mathcal{A}_{\Sigma', (l_1, l_2, \dots, l_n)}$  remain contractible. For the base case  $n = 0$ , we have  $\Sigma = \Sigma^{( )}$ ,  $\Sigma' = \Sigma'^{( )}$ . Thus the base case is true by assumption. We present the inductive step next. □

**Claim 4** For this claim, let  $\bullet$  denote  $l_1, \dots, l_n$ , where  $0 < n < \text{card}(I')$ . Define  $l := l_{n+1}$ . Suppose the following is true:

- (1) The collections  $\mathcal{A}_{\Sigma, (\bullet)}$  and  $\mathcal{A}_{\Sigma', (\bullet)}$  are covers of  $\Sigma^{(\bullet)}$  and  $\Sigma'^{(\bullet)}$ .
- (2) The nerves of the coverings are unchanged:  $\mathcal{N}(\mathcal{A}_{\Sigma}) = \mathcal{N}(\mathcal{A}_{\Sigma, (\bullet)})$  and  $\mathcal{N}(\mathcal{A}_{\Sigma'}) = \mathcal{N}(\mathcal{A}_{\Sigma', (\bullet)})$ .
- (3) Each of the subcomplexes  $\Sigma_{i, (\bullet)}$ ,  $i \in I$ , and  $\Sigma'_{j, (\bullet)}$ ,  $j \in I'$  is induced in  $\Sigma^{(\bullet)}$  and  $\Sigma'^{(\bullet)}$ , respectively.
- (4) Let  $\sigma \subseteq I$ . If  $\bigcap_{i \in \sigma} \Sigma_{i, (\bullet)}$  is nonempty, then it is contractible. Similarly, let  $\tau \subseteq I'$ . If  $\bigcap_{i \in \tau} \Sigma'_{i, (\bullet)}$  is nonempty, then it is contractible.
- (5) We have homotopy equivalences  $|\Sigma| \simeq |\Sigma^{(\bullet)}|$  and  $|\Sigma'| \simeq |\Sigma'^{(\bullet)}|$  via maps that commute with the canonical inclusions.

Then the preceding statements are true for  $\Sigma^{(\bullet, l)}$ ,  $\Sigma'^{(\bullet, l)}$ ,  $\mathcal{A}_{\Sigma, (\bullet, l)}$ , and  $\mathcal{A}_{\Sigma', (\bullet, l)}$  as well.

**Proof** For the first claim, we have  $\Sigma^{(\bullet, l)} = \Sigma^{(\bullet)} \cup \text{pow}(V(\Sigma_{l, (\bullet)})) \subseteq \bigcup_{i \in I} \Sigma_{i, (\bullet, l)}$ . For the inclusion, we used the inductive assumption that  $\Sigma^{(\bullet)} = \bigcup_{i \in I} \Sigma_{i, (\bullet)}$ . Similarly,  $\Sigma'^{(\bullet, l)} \subseteq \bigcup_{i \in I'} \Sigma'_{i, (\bullet, l)}$ .

For the second claim, let  $i \in I$ . Then  $V(\Sigma_{i, (l_1)}) = V(\Sigma_i)$ , and in particular, we have  $V(\Sigma_{i, (\bullet, l)}) = V(\Sigma_{i, (\bullet)}) = V(\Sigma_i)$ . Next observe that for any  $\sigma \subseteq I$ , the intersection

$$\begin{aligned} \bigcap_{i \in \sigma} \Sigma_i \neq \emptyset &\iff \bigcap_{i \in \sigma} V(\Sigma_i) \neq \emptyset \iff \bigcap_{i \in \sigma} V(\Sigma_{i, (\bullet, l)}) \\ &\neq \emptyset \iff \bigcap_{i \in \sigma} \Sigma_{i, (\bullet, l)} \neq \emptyset. \end{aligned}$$

Thus  $\mathcal{N}(\mathcal{A}_{\Sigma}) = \mathcal{N}(\mathcal{A}_{\Sigma, (\bullet)}) = \mathcal{N}(\mathcal{A}_{\Sigma, (\bullet, l)})$ , and similarly  $\mathcal{N}(\mathcal{A}_{\Sigma'}) = \mathcal{N}(\mathcal{A}_{\Sigma', (\bullet)}) = \mathcal{N}(\mathcal{A}_{\Sigma', (\bullet, l)})$ .

For the third claim, again let  $i \in I$ . If  $l \notin I$ , then  $\Sigma_{i, (\bullet, l)} = \Sigma_{i, (\bullet)}$ , so we are done by the inductive assumption. Suppose  $l \in I$ . Since  $\Sigma_{i, (\bullet)}$  is induced by the inductive assumption, we have:

$$\begin{aligned} \Sigma_{i, (\bullet, l)} &= \Sigma_{i, (\bullet)} \cup (\text{pow}(V(\Sigma_{i, (\bullet)}) \cap V(\Sigma_{l, (\bullet)}))) \\ &= (\Sigma^{(\bullet)} \cap \text{pow}(V(\Sigma_{i, (\bullet)}))) \cup (\text{pow}(V(\Sigma_{i, (\bullet)})) \cap \text{pow}(V(\Sigma_{l, (\bullet)}))) \\ &= (\Sigma^{(\bullet)} \cup \text{pow}(V(\Sigma_{l, (\bullet)}))) \cap \text{pow}(V(\Sigma_{i, (\bullet)})) \\ &= \Sigma^{(\bullet, l)} \cap \text{pow}(V(\Sigma_{i, (\bullet)})) = \Sigma^{(l)} \cap \text{pow}(V(\Sigma_{i, (\bullet, l)})). \end{aligned}$$

Thus  $\Sigma_{i,(\bullet,l)}$  is induced. The same argument holds for the  $I'$  case.

For the fourth claim, let  $\sigma \subseteq I$ , and suppose  $\bigcap_{i \in \sigma} \Sigma_{i,(\bullet,l)}$  is nonempty. By the previous claim, each  $\Sigma_{i,(\bullet,l)}$  is induced. Thus we write:

$$\begin{aligned} \bigcap_{i \in \sigma} \Sigma_{i,(\bullet,l)} &= \Sigma^{(\bullet,l)} \cap \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet,l)})) \\ &= \Sigma^{(\bullet,l)} \cap \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)})) \\ &= (\Sigma^{(\bullet)} \cup \text{pow}(V(\Sigma_{l,(\bullet)}))) \cap \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)})) \\ &= (\bigcap_{i \in \sigma} (\Sigma^{(\bullet)} \cap \text{pow}(V(\Sigma_{i,(\bullet)})))) \\ &\quad \cup \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)}) \cap V(\Sigma_{l,(\bullet)})) \\ &= (\bigcap_{i \in \sigma} \Sigma_{i,(\bullet)}) \cup \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)}) \cap V(\Sigma_{l,(\bullet)})). \end{aligned}$$

For convenience, define  $A := (\bigcap_{i \in \sigma} \Sigma_{i,(\bullet)})$  and  $B := \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)}) \cap V(\Sigma_{l,(\bullet)}))$ . Then  $|A|$  is contractible by inductive assumption, and  $|B|$  is a full simplex, hence contractible. Also,  $A \cap B$  has the form

$$\begin{aligned} &(\bigcap_{i \in \sigma} (\Sigma^{(\bullet)} \cap \text{pow}(V(\Sigma_{i,(\bullet)})))) \cap \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)}) \cap V(\Sigma_{l,(\bullet)})) \\ &= \Sigma^{(\bullet)} \cap \text{pow}(\bigcap_{i \in \sigma} V(\Sigma_{i,(\bullet)}) \cap V(\Sigma_{l,(\bullet)})) \\ &= \bigcap_{i \in \sigma} \Sigma_{i,(\bullet)} \cap \Sigma_{l,(\bullet)}, \end{aligned}$$

and the latter is contractible by inductive assumption. Thus by Lemma 31, we have  $|A \cup B|$  contractible. This proves the claim for the case  $\sigma \subseteq I$ . The case  $\tau \subseteq I'$  is similar.

Now we proceed to the final claim. Since  $\Sigma_{l,(\bullet)}$  is induced, we have  $\Sigma_{l,(\bullet)} = \Sigma^{(\bullet)} \cap \text{pow}(V(\Sigma_{l,(\bullet)}))$ . By the contractibility assumption, we know that  $|\Sigma_{l,(\bullet)}|$  is contractible. Also we know that  $|\Sigma'_{l,(\bullet)}| = |\Sigma'^{(\bullet)} \cap \text{pow}(V(\Sigma'_{l,(\bullet)}))|$  is contractible. By assumption we also have  $V(\Sigma_{l,(\bullet)}) \subseteq V(\Sigma'_{l,(\bullet)})$ . Thus by Lemma 32, we obtain homotopy equivalences  $\Phi_l : |\Sigma^{(\bullet,l)}| \rightarrow |\Sigma^{(\bullet)}|$  and  $\Phi'_l : |\Sigma'^{(\bullet,l)}| \rightarrow |\Sigma'^{(\bullet)}|$  such that  $\Phi'_l$  extends  $\Phi_l$ . Furthermore, the homotopy inverses of  $\Phi_l$  and  $\Phi'_l$  are just the inclusions  $|\Sigma^{(\bullet)}| \hookrightarrow |\Sigma^{(\bullet,l)}|$  and  $|\Sigma'^{(\bullet)}| \hookrightarrow |\Sigma'^{(\bullet,l)}|$ .

Now let  $\iota : |\Sigma^{(\bullet)}| \rightarrow |\Sigma'^{(\bullet)}|$  and  $\iota_l : |\Sigma^{(\bullet,l)}| \rightarrow |\Sigma'^{(\bullet,l)}|$  denote the canonical inclusions. We wish to show the equality  $\Phi'_l \circ \iota_l = \iota \circ \Phi_l$ . Let  $x \in |\Sigma^{(\bullet,l)}|$ . Because  $\Phi'_l$  extends  $\Phi_l$  (this is why we needed the *functorial* gluing lemma), we have

$$\Phi'_l(\iota_l(x)) = \Phi'_l(x) = \Phi_l(x) = \iota(\Phi_l(x)).$$

Since  $x \in |\Sigma^{(\bullet,l)}|$  was arbitrary, the equality follows immediately. By the inductive assumption, we already have homotopy equivalences  $|\Sigma^{(\bullet)}| \rightarrow |\Sigma|$  and  $|\Sigma'^{(\bullet)}| \rightarrow |\Sigma'|$  that commute with the canonical inclusions. Composing these maps with  $\Phi_l$  and  $\Phi'_l$  completes the proof of the claim.  $\square$

By the preceding work, we replace the subcomplexes  $\Sigma_l, \Sigma'_l$  by full simplices of the form  $\Sigma_{l,(\bullet,l)}, \Sigma'_{l,(\bullet,l)}$ . In this process, the nerves remain unchanged and the complexes  $\Sigma, \Sigma'$  are replaced by homotopy equivalent complexes  $\Sigma^{(\bullet,l)}, \Sigma'^{(\bullet,l)}$ . Furthermore,

this process is functorial—the homotopy equivalences commute with the canonical inclusions  $\Sigma \hookrightarrow \Sigma^{(\bullet, I)}$  and  $\Sigma' \hookrightarrow \Sigma'^{(\bullet, I)}$ .

Repeating the inductive process in Claim 4 for all the finitely many  $l \in I$  yields a simplicial complex  $\Sigma^{(I)}$  along with a cover of simplices  $\mathcal{A}_{\Sigma, (I)}$ . We also perform the same procedure for all  $l \in I' \setminus I$  (this does not affect  $\Sigma^{(I)}$ ) to obtain a simplicial complex  $\Sigma'^{(I')}$  along with a cover of simplices  $\mathcal{A}_{\Sigma', (I')}$ . Furthermore,  $\Sigma^{(I)}$  and  $\Sigma'^{(I')}$  are related to  $\Sigma$  and  $\Sigma'$  by a finite sequence of homotopy equivalences that commute with the canonical inclusions. Also, we have  $\mathcal{N}(\mathcal{A}_\Sigma) = \mathcal{N}(\mathcal{A}_{\Sigma, (I)})$  and  $\mathcal{N}(\mathcal{A}_{\Sigma'}) = \mathcal{N}(\mathcal{A}_{\Sigma', (I')})$ . Thus we obtain the following picture:

$$\begin{array}{ccccccc}
 |\Sigma| & \xleftarrow{\simeq} & |\Sigma^{(I)}| & \xleftarrow{\simeq} & \dots & \xleftarrow{\simeq} & |\Sigma^{(I')}| & \xrightarrow{\simeq} & |\mathcal{N}(\mathcal{A}_{\Sigma, (I)})| & \xlongequal{\quad} & |\mathcal{N}(\mathcal{A}_\Sigma)| \\
 \downarrow \iota & & \downarrow \iota_{(I)} & & & & \downarrow \iota_{(I')} & & \downarrow \iota_{\mathcal{N}, (I)} & & \downarrow \iota_{\mathcal{N}} \\
 |\Sigma'| & \xleftarrow{\simeq} & |\Sigma'^{(I')}| & \xleftarrow{\simeq} & \dots & \xleftarrow{\simeq} & |\Sigma^{(I')}| & \xrightarrow{\simeq} & |\mathcal{N}(\mathcal{A}_{\Sigma', (I')})| & \xlongequal{\quad} & |\mathcal{N}(\mathcal{A}_{\Sigma'})|
 \end{array}$$

By applying Theorem 25 to the block consisting of  $|\Sigma^{(I)}|$ ,  $|\Sigma'^{(I')}|$ ,  $|\mathcal{N}(\mathcal{A}_{\Sigma, (I)})|$  and  $|\mathcal{N}(\mathcal{A}_{\Sigma', (I')})|$ , we obtain a square that commutes up to homotopy. Then by composing the homotopy equivalences constructed above, we obtain a square consisting of  $|\Sigma|$ ,  $|\Sigma'|$ ,  $|\mathcal{N}(\mathcal{A}_\Sigma)|$ , and  $|\mathcal{N}(\mathcal{A}_{\Sigma'})|$  that commutes up to homotopy. Thus we obtain homotopy equivalences  $|\Sigma| \simeq |\mathcal{N}(\mathcal{A}_\Sigma)|$  and  $|\Sigma'| \simeq |\mathcal{N}(\mathcal{A}_{\Sigma'})|$  via maps that commute up to homotopy with the canonical inclusions.  $\square$

## 6 Dowker persistence diagrams and asymmetry

From the very definition of the Rips complex at any given resolution, one can see that the Rips complex is blind to asymmetry in the input data (Remark 14). In this section, we argue that either of the Dowker source and sink complexes is sensitive to asymmetry. Thus when analyzing datasets containing asymmetric information, one may wish to use the Dowker filtration instead of the Rips filtration. In particular, this property suggests that the Dowker persistence diagram is a stronger invariant for directed networks than the Rips persistence diagram.

In this section, we consider a family of examples, called *cycle networks*, for which the Dowker persistence diagrams capture meaningful structure, whereas the Rips persistence diagrams do not.

We then probe the question “What happens to the Dowker or Rips persistence diagram of a network upon reversal of one (or more) edges?” Intuitively, if either of these persistence diagrams captures asymmetry, we would see a change in the diagram after applying this reversal operation to an edge.

### 6.1 Cycle networks

For each  $n \in \mathbb{N}$ , let  $(X_n, E_n, W_{E_n})$  denote the weighted graph with vertex set  $X_n := \{x_1, x_2, \dots, x_n\}$ , edge set  $E_n := \{(x_1, x_2), (x_2, x_3), \dots, (x_{n-1}, x_n), (x_n, x_1)\}$ , and edge weights  $W_{E_n} : E_n \rightarrow \mathbb{R}$  given by writing  $W_{E_n}(e) = 1$  for each  $e \in E_n$ . Next let  $\omega_{G_n} : X_n \times X_n \rightarrow \mathbb{R}$  denote the shortest path distance induced on  $X_n \times X_n$  by  $W_{E_n}$ . Then we write  $G_n := (X_n, \omega_{G_n})$  to denote the network with node set  $X_n$  and



weights given by  $\omega_{G_n}$ . Note that  $\omega_{G_n}(x, x) = 0$  for each  $x \in X_n$ . See Fig. 3 for some examples.

We say that  $G_n$  is the *cycle network of length  $n$* . Cycle networks are highly asymmetric because for every consecutive pair of nodes  $(x_i, x_{i+1})$  in a graph  $G_n$ , where  $1 \leq i \pmod{n} \leq n$ , we have  $\omega_{G_n}(x_i, x_{i+1}) = 1$ , whereas  $\omega_{G_n}(x_{i+1}, x_i) = \text{diam}(G_n) = n - 1$ , which is much larger than 1 when  $n$  is large.

To provide further evidence that Dowker persistence is sensitive to asymmetry, we computed both the Rips and Dowker persistence diagrams, in dimensions 0 and 1, of cycle networks  $G_n$ , for values of  $n$  between 3 and 6. Computations were carried out using `Javaplex` in Matlab with  $\mathbb{Z}_2$  coefficients. The results are presented in Fig. 3. Based on our computations, we were able to conjecture and prove the result in Theorem 33, which gives a precise characterization of the 1-dimensional Dowker persistence diagram of a cycle network  $G_n$ , for any  $n$ . Furthermore, the 1-dimensional Dowker persistence barcode for any  $G_n$  contains only one persistent interval, which agrees with our intuition that there is only one nontrivial loop in  $G_n$ . On the other hand, for large  $n$ , the 1-dimensional Rips persistence barcodes contain more than one persistent interval. This can be seen in the Rips persistence barcode of  $G_6$ , presented in Fig. 3. Moreover, for  $n = 3, 4$ , the 1-dimensional Rips persistence barcode does not contain any persistent interval at all. This suggests that Dowker persistence diagrams/barcodes are an appropriate method for analyzing cycle networks, and perhaps asymmetric networks in general.

**Notation** In the remainder of this section, we will prove results involving Dowker sink complexes of the cycle networks  $G_n$  and associated vector spaces at a range of resolutions  $\delta$ . For convenience, we will write  $\mathfrak{D}_\delta^{\text{si}} := \mathfrak{D}_{\delta, G_n}^{\text{si}}$  (where  $n$  will be fixed) and  $C_k^\delta := C_k(\mathfrak{D}_\delta^{\text{si}})$ , the  $k$ -chain vector space associated to  $\mathfrak{D}_\delta^{\text{si}}$  for each  $k \in \mathbb{Z}_+$ . For each  $k \in \mathbb{Z}_+$ , the boundary map from  $C_k^\delta$  to  $C_{k-1}^\delta$  will be denoted  $\partial_k^\delta$ . Whenever we write  $x_i$  to denote a vertex of  $G_n$ , the subscript  $i$  should be understood as  $i \pmod{n}$ . We write  $e_i$  to denote the 1-simplex  $[x_i, x_{i+1}]$  for each  $1 \leq i \leq n$ , where  $x_{n+1}$  is understood to be  $x_1$ . Given an element  $\gamma \in \ker(\partial_k^\delta) \subseteq C_k^\delta$ , we will write  $\langle \gamma \rangle_\delta$  to denote its equivalence class in the quotient vector space  $\ker(\partial_k^\delta) / \text{im}(\partial_k^\delta)$ . We will refer to the operation of taking this quotient as *passing to homology*.

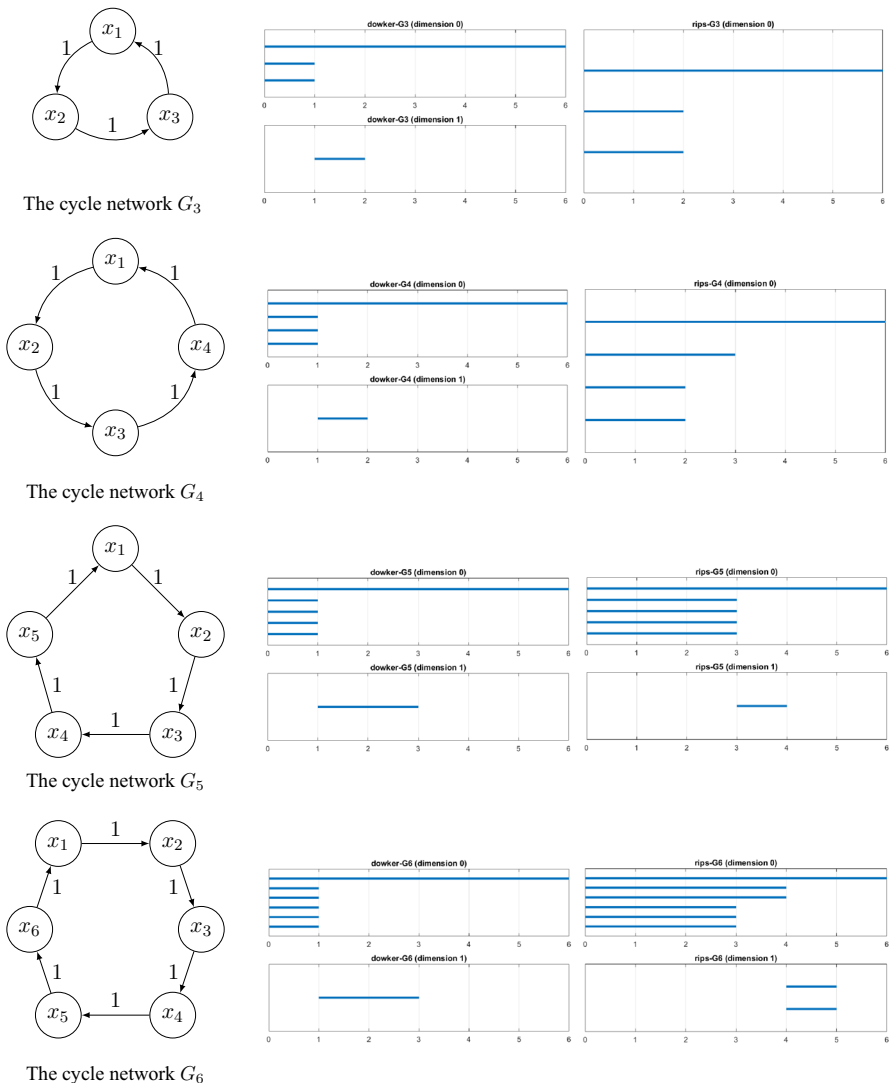
The following theorem contains the characterization result for 1-dimensional Dowker persistence diagrams of cycle networks.

**Theorem 33** *Let  $G_n = (X_n, \omega_{G_n})$  be a cycle network for some  $n \in \mathbb{N}$ ,  $n \geq 3$ . Then we obtain:*

$$\text{Dgm}_1^{\mathfrak{D}}(G_n) = \left\{ (1, \lceil n/2 \rceil) \in \mathbb{R}^2 \right\}.$$

Thus  $\text{Dgm}_1^{\mathfrak{D}}(G_n)$  consists of precisely the point  $(1, \lceil n/2 \rceil) \in \mathbb{R}^2$  with multiplicity 1.

**Proof of Theorem 33** The proof occurs in three stages: first we show that a 1-cycle appears at  $\delta = 1$ , next we show that this 1-cycle does not become a boundary until  $\delta = \lceil n/2 \rceil$ , and finally that any other 1-cycle belongs to the same equivalence class upon passing to homology (this shows that the single point in the persistence diagram has multiplicity 1).



**Fig. 3** The first column contains illustrations of cycle networks  $G_3, G_4, G_5$  and  $G_6$ . The second column contains the corresponding Dowker persistence barcodes, in dimensions 0 and 1. Note that the persistent intervals in the 1-dimensional barcodes agree with the result in Theorem 33. The third column contains the Rips persistence barcodes of each of the cycle networks. Note that for  $n = 3, 4$ , there are no persistent intervals in dimension 1. On the other hand, for  $n = 6$ , there are two persistent intervals in dimension 1

Note that for  $\delta < 1$ , there are no 1-simplices in  $\mathcal{D}_\delta^{\text{si}}$ , and so  $H_1(\mathcal{D}_\delta^{\text{si}})$  is trivial. Suppose  $1 \leq \delta < 2$ . □

**Claim 5** *There are no 2-simplices in  $\mathcal{D}_\delta^{\text{si}}$  for  $1 \leq \delta < 2$ .*

**Proof** To see this, let  $x_i, x_j, x_k$  be any three distinct vertices in  $X_n$ . Assume towards a contradiction that there exists  $x \in X_n$  such that  $(x_i, x), (x_j, x), (x_k, x) \in R_{\delta, X_n}$ ,

where  $R_{\delta, x_n}$  is as given by Eq. 3. Thus  $\omega_{G_n}(x_i, x) \in \{0, 1\}$ , so either  $x = x_i$  or  $x = x_{i+1}$ . Similarly we get that  $x = x_j$  or  $x = x_{j+1}$ , and that  $x = x_k$  or  $x = x_{k+1}$ . But this is a contradiction, since  $x_i, x_j, x_k$  are all distinct.  $\square$

By the claim, there are no 2-simplices in  $\mathfrak{D}_\delta^{\text{si}}$ , so  $\text{im}(\partial_2^\delta)$  is trivial and the only 1-chains are linear combinations of  $e_i$  terms. Next, we define:

$$v_n := e_1 + e_2 + \dots + e_n = [x_1, x_2] + [x_2, x_3] + \dots + [x_n, x_1].$$

Note that  $v_n \in C_1^\delta$  for all  $\delta \geq 1$ . We also have  $\partial_1^\delta(v_n) = 0$ , for any  $\delta \geq 1$ . In other words,  $v_n$  is a 1-cycle for any  $\delta \geq 1$ .

**Claim 6** *Let  $1 \leq \delta < 2$ . Then  $v_n$  generates  $\ker(\partial_1^\delta) \subseteq C_1^\delta$ .*

**Proof** The only 1-simplices in  $\mathfrak{D}_\delta^{\text{si}}$  are of the form  $e_i$ , for  $1 \leq i \leq n$ . So it suffices to show that any linear combination of the  $e_i$  terms is a multiple of  $v_n$ . Let  $u = \sum_{i=1}^n a_i e_i \in \ker(\partial_1^\delta)$ , for some  $a_1, \dots, a_n \in \mathbb{K}$ . Then,

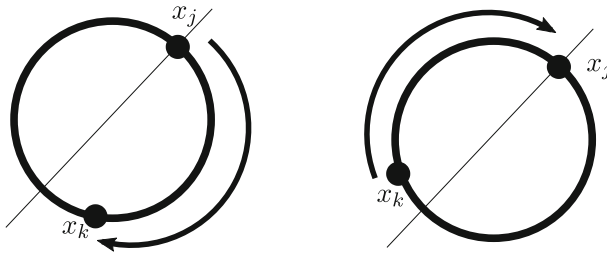
$$\begin{aligned} 0 = \partial_1^\delta(u) &= \sum_{i=1}^n a_i \partial_1^\delta(e_i) = \sum_{i=1}^n a_i ([x_{i+1}] - [x_i]) \\ &= \sum_{i=1}^n (a_{i-1} - a_i) [x_i], \quad \text{where } x_0 \text{ is understood to be } x_n. \end{aligned}$$

Since all the  $[x_i]$  are linearly independent, it follows that  $a_1 = a_2 = \dots = a_n$ . Thus it follows that  $u$  is a constant multiple of  $v_n$ . This proves the claim.  $\square$

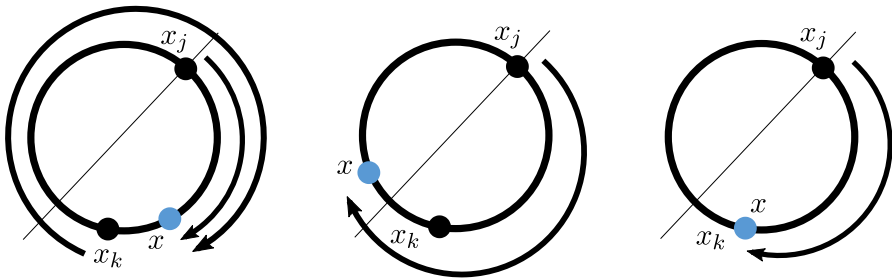
By the two preceding claims, it follows that  $\{\langle v_n \rangle_\delta\}$  is a basis for  $H_1(\mathfrak{D}_\delta^{\text{si}})$ , for  $\delta \in [1, 2)$ . More specifically,  $\langle v_n \rangle_\delta$  is a cycle that appears at  $\delta = 1$  and does not become a boundary until at least  $\delta = 2$ , and any other cycle in  $C_1^\delta$ , for  $\delta \in [1, 2)$ , is in the linear span of  $v_n$ . Next, suppose  $\delta \geq 2$ . Note that this allows the appearance of cycles that are not in the span of  $v_n$ . In the next claim, we show that upon passing to homology, the equivalence class of any such cycle coincides with that of  $v_n$ . This will show that there can be at most one nontrivial element in  $\text{Dgm}_1^{\text{si}}(G_n)$ .

**Claim 7** *Let  $\delta \geq 2$ , and let  $y = \sum_{i=1}^p a_i \sigma_i \in \ker(\partial_1^\delta)$  for some  $p \in \mathbb{N}$ , some  $a_1, \dots, a_p \in \mathbb{K}$ , and some  $\sigma_1, \dots, \sigma_p \in \mathfrak{D}_\delta^{\text{si}}$ . Then there exists a choice of coefficients  $(b_i)_{i=1}^n \in \mathbb{K}^n$  such that  $z = \sum_{i=1}^n b_i e_i \in \ker(\partial_1^\delta)$  and  $y - z \in \text{im}(\partial_2^\delta)$ . Moreover, we obtain  $\langle y \rangle_\delta = \langle z \rangle_\delta = \langle v_n \rangle_\delta$  upon passing to homology.*

**Proof** To see this, fix  $\sigma_i \in \mathfrak{D}_\delta^{\text{si}}$ , and write  $\sigma_i = [x_j, x_k]$  for some  $1 \leq j, k \leq n$ . If  $k = j + 1$  (resp.  $k = j - 1$ ), then we already have  $\sigma_i = e_j$  (resp.  $\sigma_i = e_k$ ), so there is nothing more to show. Assume  $k \notin \{j + 1, j - 1\}$ . Since  $\omega_{G_n}(x_j, x_k) + \omega_{G_n}(x_k, x_j) = n$ , we have two cases: (1)  $\omega_{G_n}(x_j, x_k) \leq n/2$ , or (2)  $\omega_{G_n}(x_k, x_j) < n/2$ . In the first case, we have  $k = j + l$  for some integer  $l \in [2, n/2]$  (all numbers are taken modulo  $n$ ). In the second case,  $j = k + l$  for some integer  $l \in [2, n/2)$  (also modulo  $n$ ).



**Fig. 4** Given two points  $x_j, x_k \in X_n$ , we have either  $\omega_{G_n}(x_j, x_k) \leq n/2$ , or  $\omega_{G_n}(x_k, x_j) < n/2$ . To see this, note that  $\omega_{G_n}(x, x') + \omega_{G_n}(x', x) = n$  for any  $x \neq x' \in X_n$



**Fig. 5** Three possible locations for a  $\delta$ -sink  $x$  of a simplex  $[x_j, x_k]$ , assuming that  $\omega_{G_n}(x_j, x_k) \leq n/2$ . For the figure on the left, note that  $\omega_{G_n}(x_k, x) \geq n/2 \geq \omega_{G_n}(x_j, x_k)$ . For the figure in the middle, note that  $\omega_{G_n}(x_j, x) \geq \omega_{G_n}(x_j, x_k)$ . Finally, for the figure on the right, where  $x = x_k$ , note that  $\omega_{G_n}(x_j, x) = \omega_{G_n}(x_j, x_k)$  and  $\omega_{G_n}(x_k, x) = 0$

The situation is illustrated in Fig. 4. Both cases are similar, so we only prove the case  $\omega_{G_n}(x_j, x_k) \leq n/2$ .

Recall that any  $\delta$ -sink  $x \in X_n$  for  $[x_j, x_k]$  satisfies  $\max(\omega_{G_n}(x_j, x), \omega_{G_n}(x_k, x)) \leq \delta$ , by the  $\delta$ -sink condition (Eq. 4). Also note that such a  $\delta$ -sink  $x$  satisfies

$$\max(\omega_{G_n}(x_j, x), \omega_{G_n}(x_k, x)) \geq \omega_{G_n}(x_j, x_k),$$

as can be seen from Fig. 5. So whenever some  $x \in X_n$  is a  $\delta$ -sink for  $[x_j, x_k]$ , we have  $x_k$  as a valid  $\delta$ -sink for  $[x_j, x_k]$ . Since  $[x_j, x_k] \in \mathcal{D}_\delta^{\text{si}}$ , it must have a  $\delta$ -sink  $x \in X_n$ . Thus  $x_k$  is a valid  $\delta$ -sink for  $[x_j, x_k]$ . Next let  $l \in [2, n/2]$  be an integer such that  $k = j + l$  (modulo  $n$ ). Notice that:

$$\begin{aligned} 0 &= \omega_{G_n}(x_k, x_k) = \omega_{G_n}(x_{j+l}, x_k) < \omega_{G_n}(x_{j+l-1}, x_k) \\ &< \dots < \omega_{G_n}(x_{j+1}, x_k) < \omega_{G_n}(x_j, x_k) \leq \delta. \end{aligned}$$

Then observe that:

$$[x_j, x_{j+1}, x_k], [x_{j+1}, x_{j+2}, x_k], \dots, [x_{k-2}, x_{k-1}, x_k] \in \mathcal{D}_\delta^{\text{si}},$$

since  $x_k$  is a  $\delta$ -sink for all these 2-simplices. One then verifies the following:

$$\begin{aligned} & \partial_2^\delta ([x_j, x_{j+1}, x_k] + [x_{j+1}, x_{j+2}, x_k] + \cdots + [x_{k-2}, x_{k-1}, x_k]) \\ &= \partial_2^\delta \left( \sum_{q=0}^{k-j-2} [x_{j+q}, x_{j+q+1}, x_k] \right) \\ &= \sum_{q=0}^{k-j-2} [x_{j+q+1}, x_k] - \sum_{q=0}^{k-j-2} [x_{j+q}, x_k] + \sum_{q=0}^{k-j-2} [x_{j+q}, x_{j+q+1}] \\ &= \sum_{q=0}^{k-j-2} [x_{j+q+1}, x_k] - [x_j, x_k] - \sum_{q=0}^{k-j-3} [x_{j+q+1}, x_k] + \sum_{q=0}^{k-j-2} [x_{j+q}, x_{j+q+1}] \\ &= [x_j, x_{j+1}] + [x_{j+1}, x_{j+2}] + \cdots + [x_{k-1}, x_k] - [x_j, x_k] \\ &= e_j + e_{j+1} + \cdots + e_{k-1} - \sigma_i. \end{aligned}$$

Thus  $a_i(e_j + e_{j+1} + \cdots + e_{k-1}) - a_i\sigma_i \in \text{im}(\partial_2^\delta)$ . Repeating this process for all  $\sigma_i$ ,  $i \in \{1, \dots, p\}$ , we obtain the coefficients  $(b_i)_{i=1}^n$  such that  $\sum_{i=1}^p a_i\sigma_i - \sum_{i=1}^n b_i e_i \in \text{im}(\partial_2^\delta)$ . Let  $z = \sum_{i=1}^n b_i e_i$ . Then  $y - z \in \text{im}(\partial_2^\delta)$ . Moreover, since  $\partial_1^\delta \circ \partial_2^\delta = 0$ , it follows that  $\partial_1^\delta(y) - \partial_1^\delta(z) = 0$ , so  $z \in \ker(\partial_1^\delta)$ .

Finally, note that an argument analogous to that of Claim 6 shows that  $b_1 = b_2 = \cdots = b_n$ . Hence it follows that  $z$  is a multiple of  $v_n$ . Thus  $\langle z \rangle_\delta = \langle v_n \rangle_\delta$ . This proves the claim.  $\square$

By Claims 6 and 7, it follows that  $H_1(\mathcal{Q}_\delta^{\text{si}})$  is generated by  $\langle v_n \rangle_\delta$  for all  $\delta \geq 1$ , so  $\dim(H_1(\mathcal{Q}_\delta^{\text{si}})) \leq 1$  for all  $\delta \geq 1$ . It remains to show that  $\langle v_n \rangle_\delta$  does not become trivial until  $\delta = \lceil n/2 \rceil$ .

The cases  $n = 3, 4$  can now be completed quickly, so we focus on these simpler situations first. For either of  $n = 3, 4$ , we have  $\lceil n/2 \rceil = 2$ . Suppose  $\delta = 2$  and  $n = 3$ . Then we have  $[x_1, x_2, x_3] \in \mathcal{Q}_\delta^{\text{si}}$  because  $\text{diam}(G_n) = 2$  and any of  $x_1, x_2, x_3$  can be a 2-sink for  $[x_1, x_2, x_3]$ . Then,

$$\partial_2^\delta([x_1, x_2, x_3]) = [x_2, x_3] - [x_1, x_3] + [x_1, x_2] = e_1 + e_2 + e_3 = v_3.$$

Recall that by Claim 5,  $v_3 \notin \text{im}(\partial_1^\delta)$  for any  $\delta < 2$ . Thus by Claim 6 and the preceding equation,  $v_3$  generates  $\ker(\partial_1^\delta)$  for  $1 \leq \delta < 2$ , and becomes a boundary for precisely  $\delta \geq 2$ . Thus  $\text{Dgm}_1^{\text{si}}(G_3) = \{(1, 2)\}$ . Next, suppose  $\delta = 2$  and  $n = 4$ . Then we have  $[x_1, x_2, x_3], [x_1, x_3, x_4] \in \mathcal{Q}_\delta^{\text{si}}$  with  $x_3, x_1$  as 2-sinks, respectively. By a direct computation, we then have:

$$\partial_2^\delta([x_1, x_2, x_3] + [x_1, x_3, x_4]) = e_1 + e_2 + e_3 + e_4 = v_4.$$

By following the same argument as for the case  $n = 3$ , we see that  $\text{Dgm}_1^{\text{si}}(G_4) = \{(1, 2)\}$ .

In the sequel, we assume that  $n > 4$ . Recall that it remains to show that  $\langle v_n \rangle_\delta$  does not become trivial until  $\delta = \lceil n/2 \rceil$ , and that  $\langle v_n \rangle_\delta = 0$  for all  $\delta \geq \lceil n/2 \rceil$ . We

have already shown that  $\langle v_n \rangle_\delta$  is not trivial for  $\delta \in [1, 2)$ . We proceed by defining the following:

$$\gamma_n := [x_1, x_2, x_3] + [x_1, x_3, x_4] + \cdots + [x_1, x_{n-1}, x_n] = \sum_{i=1}^{n-2} [x_1, x_{i+1}, x_{i+2}].$$

**Claim 8** For each  $\delta \geq \lceil n/2 \rceil$ , we have  $\gamma_n \in C_2^\delta$  and  $\partial_2^\delta(\gamma_n) = v_n$ . In particular,  $\langle v_n \rangle_\delta = 0$  for all such  $\delta$ .

**Proof** Let  $\delta \geq \lceil n/2 \rceil$ . Notice that

$$\omega_{G_n}(x_{\lceil n/2 \rceil+1}, x_1) = n - \omega_{G_n}(x_1, x_{\lceil n/2 \rceil+1}) = n - \lceil n/2 \rceil \leq n/2 \leq \lceil n/2 \rceil \leq \delta,$$

so  $\omega_{G_n}(x_i, x_1) \leq \delta$  for each  $i \in \{\lceil n/2 \rceil + 1, \lceil n/2 \rceil + 2, \dots, n\}$ . Then for each  $i \in \{\lceil n/2 \rceil + 1, \lceil n/2 \rceil + 2, \dots, n - 1\}$ , we have  $[x_i, x_{i+1}, x_1] \in \mathfrak{D}_\delta^{\text{si}}$ , with  $x_1$  as a  $\delta$ -sink.

Also notice that for each  $i \in \{1, \dots, \lceil n/2 \rceil\}$ ,

$$\omega_{G_n}(x_i, x_{\lceil n/2 \rceil+1}) \leq \omega_{G_n}(x_1, x_{\lceil n/2 \rceil+1}) = \lceil n/2 \rceil \leq \delta,$$

so  $\omega_{G_n}(x_i, x_{\lceil n/2 \rceil+1}) \leq \delta$ . Thus for any  $i \in \{2, \dots, \lceil n/2 \rceil\}$ , we have  $[x_1, x_i, x_{i+1}] \in \mathfrak{D}_\delta^{\text{si}}$ , with  $x_{\lceil n/2 \rceil+1}$  as a  $\delta$ -sink.

Combining the two preceding observations, we see that for any  $i \in \{2, \dots, n - 2\}$ , we have  $[x_1, x_{i+1}, x_{i+2}] \in \mathfrak{D}_\delta^{\text{si}}$ . It follows that  $\gamma_n \in C_2^\delta$ .

Next we observe the following:

$$\begin{aligned} \partial_2^\delta(\gamma_n) &= \partial_2^\delta \left( \sum_{i=1}^{n-2} [x_1, x_{i+1}, x_{i+2}] \right) \\ &= \sum_{i=1}^{n-2} [x_{i+1}, x_{i+2}] - \sum_{i=1}^{n-2} [x_1, x_{i+2}] + \sum_{i=1}^{n-2} [x_1, x_{i+1}] \\ &= \sum_{i=1}^{n-2} [x_{i+1}, x_{i+2}] - \sum_{i=1}^{n-2} [x_1, x_{i+2}] + [x_1, x_2] + \sum_{i=2}^{n-2} [x_1, x_{i+1}] \\ &= \sum_{i=1}^{n-2} [x_{i+1}, x_{i+2}] + [x_1, x_2] - [x_1, x_n] = v_n. \end{aligned}$$

It follows that for any  $\delta \geq \lceil n/2 \rceil$ , we have  $v_n \in \text{im}(\partial_2^\delta)$ , and so  $\langle v_n \rangle_\delta = 0$  for each such  $\delta$ . □

**Claim 9** There does not exist  $\delta \in [2, \lceil n/2 \rceil)$  such that  $\langle v_n \rangle_\delta$  is trivial.

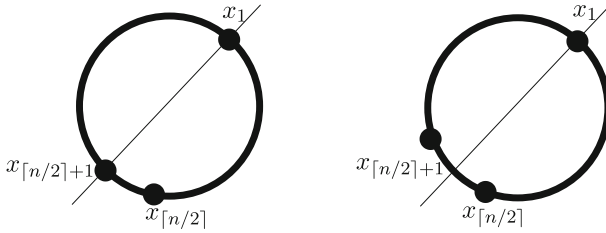


Fig. 6 Placement of  $x_{[n/2]}$  and  $x_{[n/2]+1}$ , depending on whether  $n$  is even or not

**Proof** Let  $2 \leq \delta < [n/2]$ . As a first step, we wish to show that  $\gamma_n \notin C_2^\delta$ . For this step, it suffices to show that the 2-simplex  $\sigma := [x_1, x_{[n/2]}, x_{[n/2]+1}]$  does not belong to  $\mathfrak{D}_\delta^{si}$ . The placement of  $x_{[n/2]}$  and  $x_{[n/2]+1}$  is illustrated in Fig. 6.

By an argument similar to that used in Fig. 5, one verifies that there exists a  $\delta$ -sink for  $\sigma$  if and only if at least one of  $x_1, x_{[n/2]}, x_{[n/2]+1}$  is a  $\delta$ -sink for  $\sigma$ . But note the following:

$$\begin{aligned} \omega_{G_n}(x_{[n/2]}, x_1) &= n - ([n/2] - 1) = \begin{cases} n/2 + 1 & : n \text{ even} \\ [n/2] & : n \text{ odd} \end{cases} \\ &\geq [n/2] > \delta, \end{aligned}$$

so  $x_1$  cannot be a  $\delta$ -sink for  $\sigma$ . Similarly we note that  $\omega_{G_n}(x_{[n/2]+1}, x_{[n/2]}) = n > \delta$  and  $\omega_{G_n}(x_1, x_{[n/2]+1}) = [n/2] > \delta$ , so neither  $x_{[n/2]}$  nor  $x_{[n/2]+1}$  can be  $\delta$ -sinks for  $\sigma$ . Thus  $\sigma \notin \mathfrak{D}_\delta^{si}$ , and so  $\gamma_n \notin C_2^\delta$ .

Suppose there exists  $\gamma' \in C_2^\delta$  such that  $\partial_2^\delta(\gamma') = v_n$ . Since  $[x_1, x_2]$  is a summand of  $v_n$ , we must have  $a_j[x_1, x_2, x_j]$  as a summand of  $\gamma'$ , for some coefficient  $a_j$  and some  $3 \leq j \leq n$ . First suppose that  $x_j$  is a sink for  $[x_1, x_2, x_j]$ . We claim that  $\gamma'$  is homologous to a chain containing  $[x_1, x_2, x_3]$  as a summand. If  $j = 3$ , then we are done, so suppose  $j > 3$ . Then we also know that  $[x_1, x_2, x_3, x_j]$  is a 3-simplex in  $\mathfrak{D}_\delta^{si}$ . Let  $\gamma''$  be the chain obtained from  $\gamma'$  by replacing all  $[x_1, x_2, x_j]$  terms where  $j > 3$  with terms of the form  $[x_1, x_2, x_3] - [x_2, x_3, x_j] + [x_1, x_3, x_j]$ . Since  $\partial_3^\delta([x_1, x_2, x_3, x_j]) = [x_2, x_3, x_j] - [x_1, x_3, x_j] + [x_1, x_2, x_j] - [x_1, x_2, x_3]$  and  $\partial_2^\delta \circ \partial_3^\delta = 0$ , we know that  $\partial_2^\delta(\gamma'') = \partial_2^\delta(\gamma') = v_n$ .

Now  $\partial_2^\delta([x_1, x_2, x_3])$  contributes an  $[x_1, x_3]$  summand which does not appear in  $v_n$ , so it must be cancelled by some other terms in  $\gamma''$ . Thus there must exist another summand  $[x_1, x_3, x_k]$  in  $\gamma''$ , where  $k > 3$ . We repeat the preceding argument to update  $\gamma''$  to a chain homologous to  $\gamma'$  containing both  $[x_1, x_2, x_3]$  and  $[x_1, x_3, x_4]$  as summands, but no summands of the form  $[x_1, x_3, x_j]$  for  $j > 4$ . We describe one more step of this iteration. Since  $[x_1, x_4]$  does not appear in  $v_n$ , we must have some  $[x_1, x_4, x_j]$  terms in  $\gamma''$  such that after applying  $\partial_2^\delta$ , the  $[x_1, x_4]$  terms from  $[x_1, x_3, x_4]$  and  $[x_1, x_4, x_j]$  are cancelled. We previously collected and removed all  $[x_1, x_2, x_4]$  terms, so we must have  $j > 4$ . Thus the iteration forces us to have “bigger and bigger” triangles. Proceeding in this way, we obtain a chain homologous to  $\gamma'$  that contains  $[x_1, x_{[n/2]}, x_{[n/2]+1}]$  as a summand. But this is a contradiction to what we have shown previously, i.e. that  $[x_1, x_{[n/2]}, x_{[n/2]+1}]$  is not a simplex in  $\mathfrak{D}_\delta^{si}$ .

In the case where  $x_j$  is not a sink for  $[x_1, x_2, x_j]$ , we must have  $x_2$  as a sink instead. Using similar reasoning as above, we replace  $\gamma'$  in this instance by a homologous chain containing  $[x_n, x_1, x_2]$  as a summand. Since  $[x_n, x_2]$  is not a summand of  $v_n$ , we obtain another homologous chain containing  $[x_{n-1}, x_n, x_1]$  as a summand, then a homologous chain containing  $[x_{n-1}, x_n, x_1]$ ,  $[x_{n-2}, x_{n-1}, x_1]$  as summands, and so on until we again obtain a homologous chain containing  $[x_1, x_{\lceil n/2 \rceil}, x_{\lceil n/2 \rceil + 1}]$  as a summand. Once again, this is a contradiction. This proves the claim.

Thus we have shown that  $v_n$  is a nontrivial cycle that appears at  $\delta = 1$ , and becomes a boundary at exactly  $\delta = \lceil n/2 \rceil$ . Furthermore, we have shown that upon passing to homology, the equivalence classes of all cycles coincide with that of  $v_n$ . Thus there is only one off-diagonal point  $(1, \lceil n/2 \rceil)$  on the 1-dimensional persistence diagram, which appears with multiplicity one. This concludes the proof.  $\square$

**Remark 34** From our experimental results (see Fig. 3), it appears that the 1-dimensional Rips persistence diagram of a cycle network does not admit a characterization as simple as that given by Theorem 33 for the 1-dimensional Dowker persistence diagram. Moreover, the Rips complexes  $\mathfrak{R}_{G_n}^\delta$ ,  $\delta \in \mathbb{R}$ ,  $n \in \mathbb{N}$  correspond to certain types of *independence complexes* that appear independently in the literature, and whose homotopy types remain open (Engström 2009, Question 5.3). On a related note, we point the reader to Adamaszek and Adams (2017) for a complete characterization of the homotopy types of Rips complexes of points on the circle (equipped with the restriction of the arc length metric).

To elaborate on the connection to Engström (2009), we write  $H_n^k$  to denote the undirected graph with vertex set  $\{1, \dots, n\}$ , and edges given by pairs  $(i, j)$  where  $1 \leq i < j \leq n$  and either  $j - i < k$  or  $(n + i) - j < k$ . Next we write  $\text{Ind}(H_n^k)$  to denote the *independence complex* of  $H_n^k$ , which is the simplicial complex consisting of subsets  $\sigma \subseteq \{1, 2, \dots, n\}$  such that no two elements of  $\sigma$  are connected by an edge in  $H_n^k$ . Then we have  $\text{Ind}(H_n^k) = \mathfrak{R}_{G_n}^{n-k}$  for each  $k, n \in \mathbb{N}$  such that  $k < n$ . To gain intuition for this equality, fix a basepoint 1, and consider the values of  $j \in \mathbb{N}$  for which the simplex  $[1, j]$  belongs to  $\text{Ind}(H_n^k)$  and to  $\mathfrak{R}_{G_n}^{n-k}$ , respectively. In either case, we have  $k + 1 \leq j \leq n - k + 1$ . Using the rotational symmetry of the points, one then obtains the remaining 1-simplices. Rips complexes are determined by their 1-skeleton, so this suffices to construct  $\mathfrak{R}_{G_n}^{n-k}$ . Analogously,  $\text{Ind}(H_n^k)$  is determined by the edges in  $H_n^k$ , and hence also by its 1-skeleton. In Engström (2009, Question 5.3), the author writes that the homotopy type of  $\text{Ind}(H_n^k)$  is still unsolved. Characterizing the persistence diagrams  $\text{Dgm}_k^{\mathfrak{R}}(G_n)$  thus seems to be a useful future step, both in providing a computational suggestion for the homotopy type of  $\text{Ind}(H_n^k)$ , and also in providing a valuable example in the study of persistence of directed networks.

**Remark 35** Theorem 33 has the following implication for data analysis: nontrivial 1-dimensional homology in the Dowker persistence diagram of an asymmetric network suggests the presence of directed cycles in the underlying data. Of course, it is not necessarily true that nontrivial 1-dimensional persistence can occur *only* in the presence of a directed circle.

**Remark 36** Our motivation for studying cycle networks is that they constitute directed analogues of circles, and we were interested in seeing if the 1-dimensional Dowker



persistence diagram would be able to capture this analogy. Theorem 33 shows that this is indeed the case: we get a single nontrivial 1-dimensional persistence interval, which is what we would expect when computing the persistent homology of a circle in the metric space setting. We further studied the 2-dimensional Dowker persistence diagrams of cycle networks. Our computational examples, some of which are illustrated in Fig. 7, enabled us to conjecture:

**Conjecture 1** *Let  $n \in \mathbb{N}$ ,  $n \geq 3$ , and let  $G_n$  be a cycle network. If  $n$  is odd, then  $\text{Dgm}_2^{\mathfrak{D}}(G_n)$  is trivial. If  $n$  is even, then  $\text{Dgm}_2^{\mathfrak{D}}(G_n) = [(\frac{n}{2}, \frac{n}{2} + 1) \in \mathbb{R}^2]$ , and the multiplicity of this point is  $\frac{n}{2} - 1$ .*

This computationally motivated conjecture is in fact true; moreover, we have a full characterization of the persistence diagram of a cycle network across all dimensions  $k \in \mathbb{Z}_+$ . This characterization relies on results in Adamaszek and Adams (2017) and Adamaszek et al. (2016), and is stated for even and odd dimensions below:

**Theorem 37** (Even dimension) *Fix  $n \in \mathbb{N}$ ,  $n \geq 3$ . If  $l \in \mathbb{N}$  is such that  $n$  is divisible by  $(l + 1)$ , and  $k := \frac{nl}{l+1}$  is such that  $0 \leq k \leq n - 2$ , then  $\text{Dgm}_{2l}^{\mathfrak{D}}(G_n)$  consists of precisely the point  $(\frac{nl}{l+1}, \frac{nl}{l+1} + 1)$  with multiplicity  $\frac{n}{l+1} - 1$ . If  $l$  or  $k$  do not satisfy the conditions above, then  $\text{Dgm}_{2l}^{\mathfrak{D}}(G_n)$  is trivial.*

As a special case, Theorem 37 proves Conjecture 1 by setting  $l = 1$ . If  $n$  is odd, then it is not divisible by  $(l + 1) = 2$ , and so  $\text{Dgm}_2^{\mathfrak{D}}(G_n)$  is trivial. If  $n$  is even, then it is divisible by  $(l + 1) = 2$ , and  $\frac{nl}{l+1} = \frac{n}{2} \leq n - 2$  because  $n$  is at least 4. Thus  $\text{Dgm}_2^{\mathfrak{D}}(G_n)$  consists of the point  $(\frac{n}{2}, \frac{n}{2} + 1)$  with multiplicity  $\frac{n}{2} - 1$ .

**Theorem 38** (Odd dimension) *Fix  $n \in \mathbb{N}$ ,  $n \geq 3$ . Then for  $l \in \mathbb{N}$ , define  $M_l := \{m \in \mathbb{N} : \frac{nl}{l+1} < m < \frac{n(l+1)}{l+2}\}$ . If  $M_l$  is empty, then  $\text{Dgm}_{2l+1}^{\mathfrak{D}}(G_n)$  is trivial. Otherwise, we have:*

$$\text{Dgm}_{2l+1}^{\mathfrak{D}}(G_n) = \left\{ \left( a_l, \left\lceil \frac{n(l+1)}{l+2} \right\rceil \right) \right\},$$

where  $a_l := \min \{m \in M_l\}$ . We use set notation (instead of multisets) to mean that the multiplicity is 1.

In particular, for  $l = 0$ , we have  $\frac{nl}{l+1} = 0$  and  $\frac{n(l+1)}{l+2} = \frac{n}{2} \geq 3/2$ , so  $1 \in M_l$ . Thus we have  $\text{Dgm}_1^{\mathfrak{D}}(G_n) = \left\{ \left( 1, \left\lceil \frac{n}{2} \right\rceil \right) \right\}$ , and so Theorem 38 recovers Theorem 33 as a special case. However, whereas the proof of Theorem 33 is elementary and pedagogical (it relies on intuitive observations about the structure of a cycle network), the proofs of Theorems 37 and 38 use sophisticated machinery developed across Adamaszek and Adams (2017) and Adamaszek et al. (2016). We provide details for Theorem 33 in the body of the paper, and relegate full details of Theorems 37 and 38 to Appendix B.

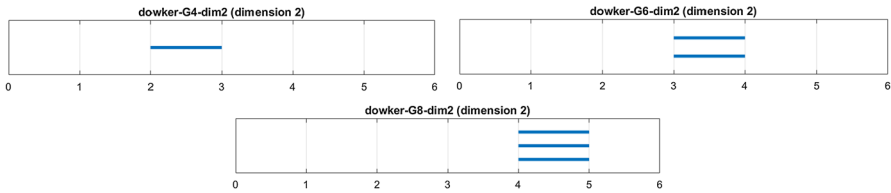


Fig. 7 Sample 2-dimensional Dowker persistence barcodes for cycle networks  $G_4, G_6, G_8$ . In our experiments, 2-dimensional Dowker persistence barcodes for  $G_n$  were always empty for  $n$  odd

### 6.2 Sensitivity to network transformations

We first make the following:

**Definition 6 (Pair swaps)** Let  $(X, \omega_X) \in \mathcal{N}$  be a network. For any  $z, z' \in X$ , define the  $(z, z')$ -swap of  $(X, \omega_X)$  to be the network  $S_X(z, z') := (X^{z, z'}, \omega_X^{z, z'})$  defined as follows:

$$X^{z, z'} := X,$$

$$\text{For any } x, x' \in X^{z, z'}, \quad \omega_X^{z, z'}(x, x') := \begin{cases} \omega_X(x', x) & : x = z, x' = z' \\ \omega_X(x', x) & : x' = z, x = z' \\ \omega_X(x, x') & : \text{otherwise.} \end{cases}$$

We then pose the following question:

*Given a network  $(X, \omega_X)$  and an  $(x, x')$ -swap  $S_X(x, x')$  for some  $x, x' \in X$ , how do the Rips or Dowker persistence diagrams of  $S_X(x, x')$  differ from those of  $(X, \omega_X)$ ?*

This situation is illustrated in Fig. 8. Example 42 shows an example where the Dowker persistence diagram captures the variation in a network that occurs after a pair swap, whereas the Rips persistence diagram fails to capture this difference. Furthermore, Remark 40 shows that Rips persistence diagrams always fail to do so.

We also consider the extreme situation where all the directions of the edges of a network are reversed, i.e. the network obtained by applying the pair swap operation to each pair of nodes. We would intuitively expect that the persistence diagrams would not change. The following discussion shows that the Rips and Dowker persistence diagrams are invariant under taking the transpose of a network.

**Proposition 39** Recall the transposition map  $\mathfrak{t}$  and the shorthand notation  $X^\top = \mathfrak{t}(X)$  from Definition 1. Let  $k \in \mathbb{Z}_+$ . Then  $\text{Dgm}_k^{\text{si}}(X) = \text{Dgm}_k^{\text{so}}(X^\top)$ , and therefore  $\text{Dgm}_k^{\text{D}}(X) = \text{Dgm}_k^{\text{D}}(X^\top)$  by Theorem 3.

**Remark 40 (Pair swaps and their effect)** Let  $(X, \omega_X) \in \mathcal{N}$ , let  $z, z' \in X$ , and let  $\sigma \in \text{pow}(X)$ . Then we have:

$$\max_{x, x' \in \sigma} \omega_X(x, x') = \max_{x, x' \in \sigma} \omega_X^{z, z'}(x, x').$$

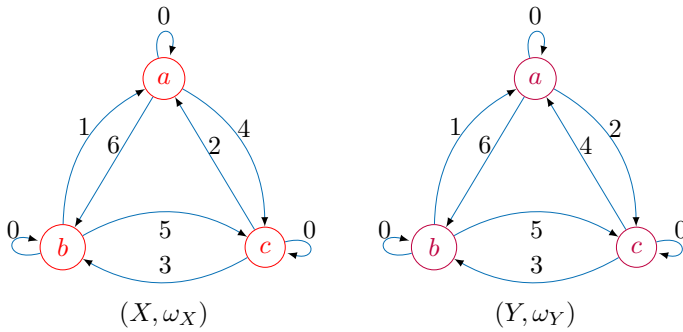


Fig. 8  $(Y, \omega_Y)$  is the  $(a, c)$ -swap of  $(X, \omega_X)$

Using this observation, one then repeats the arguments used in the proof of Proposition 39 to show that:

$$\text{Dgm}_k^{\text{Ri}}(X) = \text{Dgm}_k^{\text{Ri}}(S_X(z, z')), \text{ for each } k \in \mathbb{Z}_+.$$

This encodes the intuitive fact that Rips persistence diagrams are blind to pair swaps. Moreover, successively applying the pair swap operation over all pairs produces the transpose of the original network, and so it follows that  $\text{Dgm}_k^{\text{Ri}}(X) = \text{Dgm}_k^{\text{Ri}}(X^\top)$ .

On the other hand,  $k$ -dimensional Dowker persistence diagrams are not necessarily invariant to pair swaps when  $k \geq 1$ . Indeed, Example 42 below constructs a space  $X$  for which there exist points  $z, z' \in X$  such that

$$\text{Dgm}_1^{\text{Do}}(X) \neq \text{Dgm}_1^{\text{Do}}(S_X(z, z')).$$

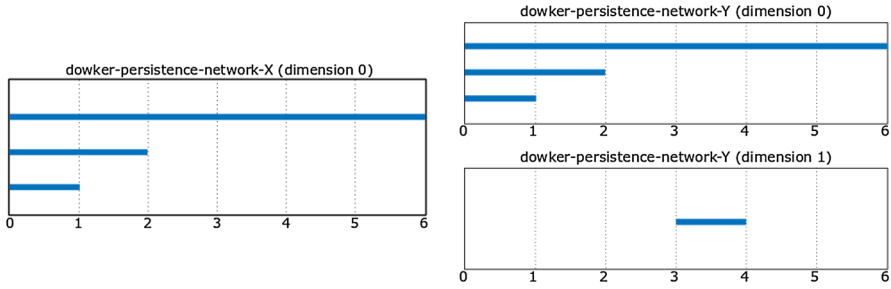
However, 0-dimensional Dowker persistence diagrams are still invariant to pair swaps:

**Proposition 41** *Let  $(X, \omega_X) \in \mathcal{N}$ , let  $z, z'$  be any two points in  $Z$ , and let  $\sigma \in \text{pow}(X)$ . Then we have:*

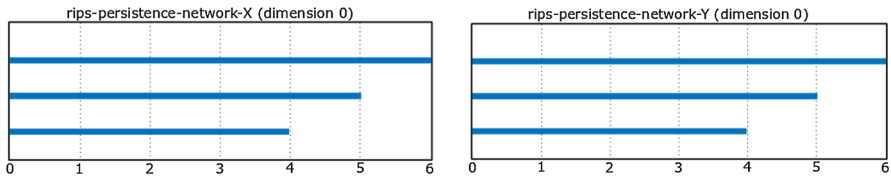
$$\text{Dgm}_0^{\text{Do}}(X) = \text{Dgm}_0^{\text{Do}}(S_X(z, z')).$$

**Example 42** Consider the three node dissimilarity networks  $(X, \omega_X)$  and  $(Y, \omega_Y)$  in Fig. 8. Note that  $(Y, \omega_Y)$  coincides with  $S_X(a, c)$ . We present both the Dowker and Rips persistence barcodes obtained from these networks. Note that the Dowker persistence barcode is sensitive to the difference between  $(X, \omega_X)$  and  $(Y, \omega_Y)$ , whereas the Rips barcode is blind to this difference. We refer the reader to Sect. 7 for details on how we compute these barcodes (Figs. 9, 10).

To show how the Dowker complex is constructed, we also list the Dowker sink complexes of the networks in Fig. 8, and also the corresponding homology dimensions



**Fig. 9** Dowker persistence barcodes of networks  $(X, \omega_X)$  and  $(Y, \omega_Y)$  from Fig. 8



**Fig. 10** Rips persistence barcodes of networks  $(X, \omega_X)$  and  $(Y, \omega_Y)$  from Fig. 8. Note that the Rips diagrams indicate no persistent homology in dimensions higher than 0, in contrast with the Dowker diagrams in Fig. 9

across a range of resolutions. Note that when we write  $[a, b](a)$ , we mean that  $a$  is a sink corresponding to the simplex  $[a, b]$ .

$$\begin{aligned}
 \mathfrak{D}_{0,X}^{\text{si}} &= \{[a], [b], [c]\} & \dim(H_1(\mathfrak{D}_{0,X}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{1,X}^{\text{si}} &= \{[a], [b], [c], [a, b](a)\} & \dim(H_1(\mathfrak{D}_{1,X}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{2,X}^{\text{si}} &= \{[a], [b], [c], [a, b](a), [a, c](a), [b, c](a), [a, b, c](a)\} & \dim(H_1(\mathfrak{D}_{2,X}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{3,X}^{\text{si}} &= \{[a], [b], [c], [a, b](a), [a, c](a), [b, c](a), [a, b, c](a)\} & \dim(H_1(\mathfrak{D}_{3,X}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{0,Y}^{\text{si}} &= \{[a], [b], [c]\} & \dim(H_1(\mathfrak{D}_{0,Y}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{1,Y}^{\text{si}} &= \{[a], [b], [c], [a, b](a)\} & \dim(H_1(\mathfrak{D}_{1,Y}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{2,Y}^{\text{si}} &= \{[a], [b], [c], [a, b](a), [a, c](c)\} & \dim(H_1(\mathfrak{D}_{2,Y}^{\text{si}})) &= 0 \\
 \mathfrak{D}_{3,Y}^{\text{si}} &= \{[a], [b], [c], [a, b](a), [a, c](c), [b, c](b)\} & \dim(H_1(\mathfrak{D}_{3,Y}^{\text{si}})) &= 1 \\
 \mathfrak{D}_{4,Y}^{\text{si}} &= \{[a], [b], [c], [a, b](a), [a, c](a), [b, c](a), [a, b, c](a)\} & \dim(H_1(\mathfrak{D}_{4,Y}^{\text{si}})) &= 0
 \end{aligned}$$

Note that for  $\delta \in [3, 4)$ ,  $\dim(H_1(\mathfrak{D}_{\delta,Y}^{\text{si}})) = 1$ , whereas  $\dim(H_1(\mathfrak{D}_{\delta,X}^{\text{si}})) = 0$  for each  $\delta \in \mathbb{R}$ .

Based on the discussion in Remark 40, Proposition 41, and Example 42, we conclude the following:

**Moral:** *Unlike Rips persistence diagrams, Dowker persistence diagrams are truly sensitive to asymmetry.*

We summarize some of these results:

**Theorem 43** Recall the symmetrization and transposition maps  $\mathfrak{s}$  and  $\mathfrak{t}$  from Definition 1. Then:

- (1)  $\mathfrak{R} \circ \mathfrak{s} = \mathfrak{R}$ ,
- (2)  $\mathfrak{D}^{\text{so}} \circ \mathfrak{t} = \mathfrak{D}^{\text{si}}$ , and
- (3)  $\mathfrak{D}^{\text{si}} \circ \mathfrak{t} = \mathfrak{D}^{\text{so}}$ .

Also, there exist  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$  such that  $(\mathfrak{D}^{\text{si}} \circ \mathfrak{s})(X) \neq \mathfrak{D}^{\text{si}}(X)$ , and  $(\mathfrak{D}^{\text{so}} \circ \mathfrak{s})(Y) \neq \mathfrak{D}^{\text{so}}(Y)$ .

**Proof** These follow from Example 42, Remark 14, and Proposition 39. □

## 7 Implementation and an experiment on network classification

In this section, we present the results of an experiment where we applied our methods to perform a classification task on a database of networks. All persistent homology computations were carried out using the `Javaplex` package for Matlab. A full description of `Javaplex` can be found in Tausz et al. (2011). We used  $\mathbb{K} = \mathbb{Z}_2$  as the field of coefficients for all our persistence computations. The dataset and software used for our computations are available as part of the `PersNet` software package on <https://research.math.osu.edu/networks/Datasets.html>. A version of our simulated hippocampal networks experiment has appeared in Chowdhury and Mémoli (2016).

All networks in the following experiment were normalized to have weights in the range  $[0, 1]$ . The Dowker sink filtration was passed into `Javaplex` to obtain 0 and 1-dimensional Dowker persistence barcodes.

### 7.1 Simulated hippocampal networks

In the neuroscience literature, it has been shown that as an animal explores a given environment or arena, specific “place cells” in the hippocampus show increased activity at specific spatial regions, called “place fields” (O’Keefe and Dostrovsky 1971). Each place cell shows a spike in activity when the animal enters the place field linked to this place cell, accompanied by a drop in activity as the animal moves far away from this place field. To understand how the brain processes this data, a natural question to ask is the following: Is the time series data of the place cell activity, referred to as “spike trains”F, enough to detect the structure of the arena?

Approaches based on homology (Curto and Itskov 2008) and persistent homology (Dabaghian et al. 2012) have shown positive results in this direction. In Dabaghian et al. (2012), the authors simulated the trajectory of a rat in an arena containing “holes.” A simplicial complex was then built as follows: whenever  $n + 1$  place cells with overlapping place fields fired together, an  $n$ -simplex was added. This yielded a filtered simplicial complex indexed by a time parameter. By computing persistence, it was then shown that the number of persistent bars in the 1-dimensional barcode of this filtered simplicial complex would accurately represent the number of holes in the arena.

We repeated this experiment with the following change in methodology: we simulated the movement of an animal, and corresponding hippocampal activity, in arenas

with a variety of obstacles. We then induced a directed network from each set of hippocampal activity data, and computed the associated 1-dimensional Dowker persistence diagrams. We were interested in seeing if the bottleneck distances between diagrams arising from similar arenas would differ significantly from the bottleneck distance between diagrams arising from different arenas. To further exemplify our methods, we repeated our analysis after computing the 1-dimensional Rips persistence diagrams from the hippocampal activity networks.

In our experiment, there were five arenas. The first was a square of side length  $L = 10$ , with four circular “holes” or “forbidden zones” of radius  $0.2L$  that the trajectory could not intersect. The other four arenas were those obtained by removing the forbidden zones one at a time. In what follows, we refer to the arenas of each type as *4-hole*, *3-hole*, *2-hole*, *1-hole*, and *0-hole arenas*. For each arena, a random-walk trajectory of 5000 steps was generated, where the animal could move along a square grid with 20 points in each direction. The grid was obtained as a discretization of the box  $[0, L] \times [0, L]$ , and each step had length  $0.05L$ . The animal could move in each direction with equal probability. If one or more of these moves took the animal outside the arena (a disallowed move), then the probabilities were redistributed uniformly among the allowed moves. Each trajectory was tested to ensure that it covered the entire arena, excluding the forbidden zones. Formally, we write the time steps as a set  $T := \{1, 2, \dots, 5000\}$ , and denote the trajectory as a map  $\text{traj} : T \rightarrow [0, L]^2$ .

For each of the five arenas, 20 trials were conducted, producing a total of 100 trials. For each trial  $l_k$ , an integer  $n_k$  was chosen uniformly at random from the interval  $[150, 200]$ . Then  $n_k$  place fields of radius  $0.05L$  were scattered uniformly at random inside the corresponding arena for each  $l_k$ . An illustration of the place field distribution is provided in Fig. 11. A spike on a place field was recorded whenever the trajectory would intersect it. So for each  $1 \leq i \leq n_k$ , the spiking pattern of cell  $x_i$ , corresponding to place field  $\text{PF}_i$ , was recorded via a function  $r_i : T \rightarrow \{0, 1\}$  given by:

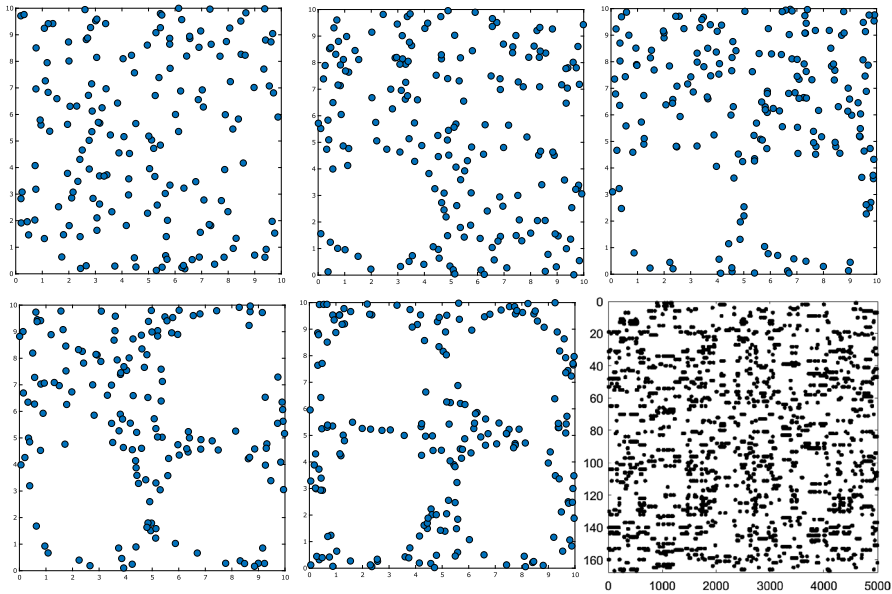
$$r_i(t) = \begin{cases} 1 & \text{: if traj}(t) \text{ intersects PF}_i, \\ 0 & \text{: otherwise} \end{cases} \quad t \in T.$$

The matrix corresponding to  $r_i$  is called the *raster* of cell  $x_i$ . A sample raster is illustrated in Fig. 11. For each trial  $l_k$ , the corresponding network  $(X_k, \omega_{X_k})$  was constructed as follows:  $X_k$  consisted of  $n_k$  nodes representing place fields, and for each  $1 \leq i, j \leq n_k$ , the weight  $\omega_{X_k}(x_i, x_j)$  was given by:

$$\omega_{X_k}(x_i, x_j) := 1 - \frac{N_{i,j}(5)}{\sum_{i=1}^{n_k} N_{i,j}(5)},$$

where  $N_{i,j}(5) = \text{card}(\{(s, t) \in T^2 : t \in [2, 5000], t - s \in [1, 5],$   
 $r_j(t) = 1, r_i(s) = 1\})$ .

In words,  $N_{i,j}(5)$  counts the pairs of times  $(s, t)$ ,  $s < t$ , such that cell  $x_j$  spikes (at a time  $t$ ) after cell  $x_i$  spikes (at a time  $s$ ), and the delay between the two spikes is fewer than 5 time steps. The idea is that if cell  $x_j$  frequently fires within a short span of time



**Fig. 11** Bottom right: Sample place cell spiking pattern matrix. The x-axis corresponds to the number of time steps, and the y-axis corresponds to the number of place cells. Black dots represent spikes. Clockwise from bottom middle: Sample distribution of place field centers in 4, 3, 0, 1, and 2-hole arenas

after cell  $x_i$  fires, then place fields  $PF_i$  and  $PF_j$  are likely to be in close proximity to each other. The column sum of the matrix corresponding to  $\omega_{X_k}$  is normalized to 1, and so  $\omega_{X_k}^\top$  can be interpreted as the transition matrix of a Markov process.

Next, we computed the 1-dimensional Dowker persistence diagrams of each of the 100 networks. Note that  $\text{Dgm}_1^{\mathcal{D}}(\omega_X) = \text{Dgm}_1^{\mathcal{D}}(\omega_X^\top)$  by Proposition 39, so we are actually obtaining the 1-dimensional Dowker persistence diagrams of transition matrices of Markov processes. We then computed a  $100 \times 100$  matrix consisting of the bottleneck distances between all the 1-dimensional persistence diagrams. The single linkage dendrogram generated from this bottleneck distance matrix is shown in Fig. 12. The labels are in the format  $\text{env} - \langle \text{nh} \rangle - \langle \text{nn} \rangle$ , where  $\text{nh}$  is the number of holes in the arena/environment, and  $\text{nn}$  is the number of place fields. Note that with some exceptions, networks corresponding to the same arena are clustered together. We conclude that the Dowker persistence diagram succeeded in capturing the intrinsic differences between the five classes of networks arising from the five different arenas, even when the networks had different sizes.

We then computed the Rips persistence diagrams of each network, and computed the  $100 \times 100$  bottleneck distance matrix associated to the collection of 1-dimensional diagrams. The single linkage dendrogram generated from this matrix is given in Fig. 13. Notice that the Rips dendrogram does not do a satisfactory job of classifying arenas correctly.

**Remark 44** We note that an alternative method of comparing the networks obtained from our simulations would have been to compute the pairwise network distances, and

plot the results in a dendrogram. But  $d_{\mathcal{N}}$  is NP-hard to compute—this follows from the fact that computing  $d_{\mathcal{N}}$  includes the problem of computing Gromov–Hausdorff distance between finite metric spaces, which is NP-hard (Schmiedl 2015). So instead, we are computing the bottleneck distances between 1-dimensional Dowker persistence diagrams, as suggested by Remark 16.

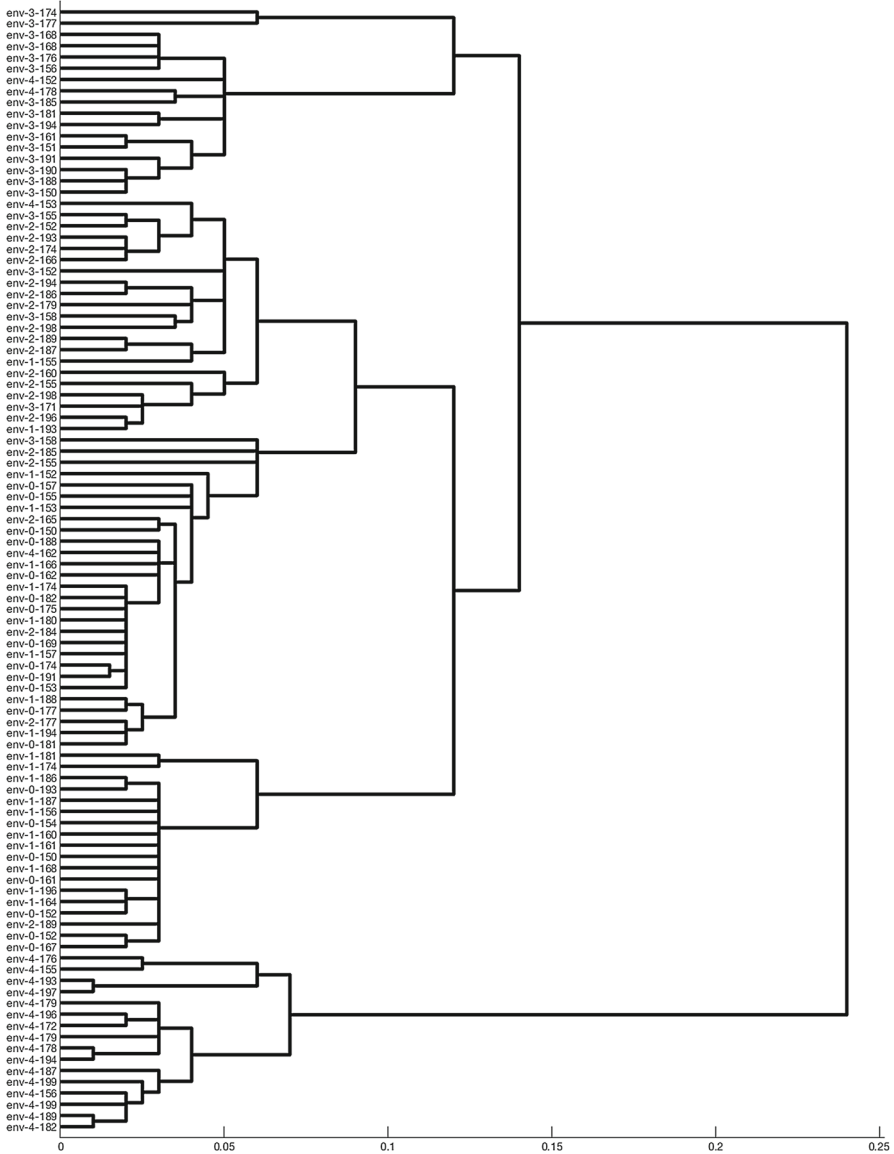
**Remark 45** It is possible to compare the current approach with the one taken in Dabaghian et al. (2012) on a common dataset. We performed this comparison in Chowdhury et al. (2017) for a similar experiment, but with a stochastic firing model for the place cells. Interestingly, it turns out that the network approach with Dowker persistence performs better than the approach in Dabaghian et al. (2012), as indicated by computing 1-nearest neighbor classification error rates on the bottleneck distance matrices. A possible explanation is that preprocessing the spiking data into a network automatically incorporates a form of error correction, where the errors consist of stochastic firing between cells that are non-adjacent. On the other hand, such errors are allowed to accumulate over time in the approach taken in Dabaghian et al. (2012). For an alternative error-correction approach, see Chowdhury et al. (2017).

## 8 Discussion

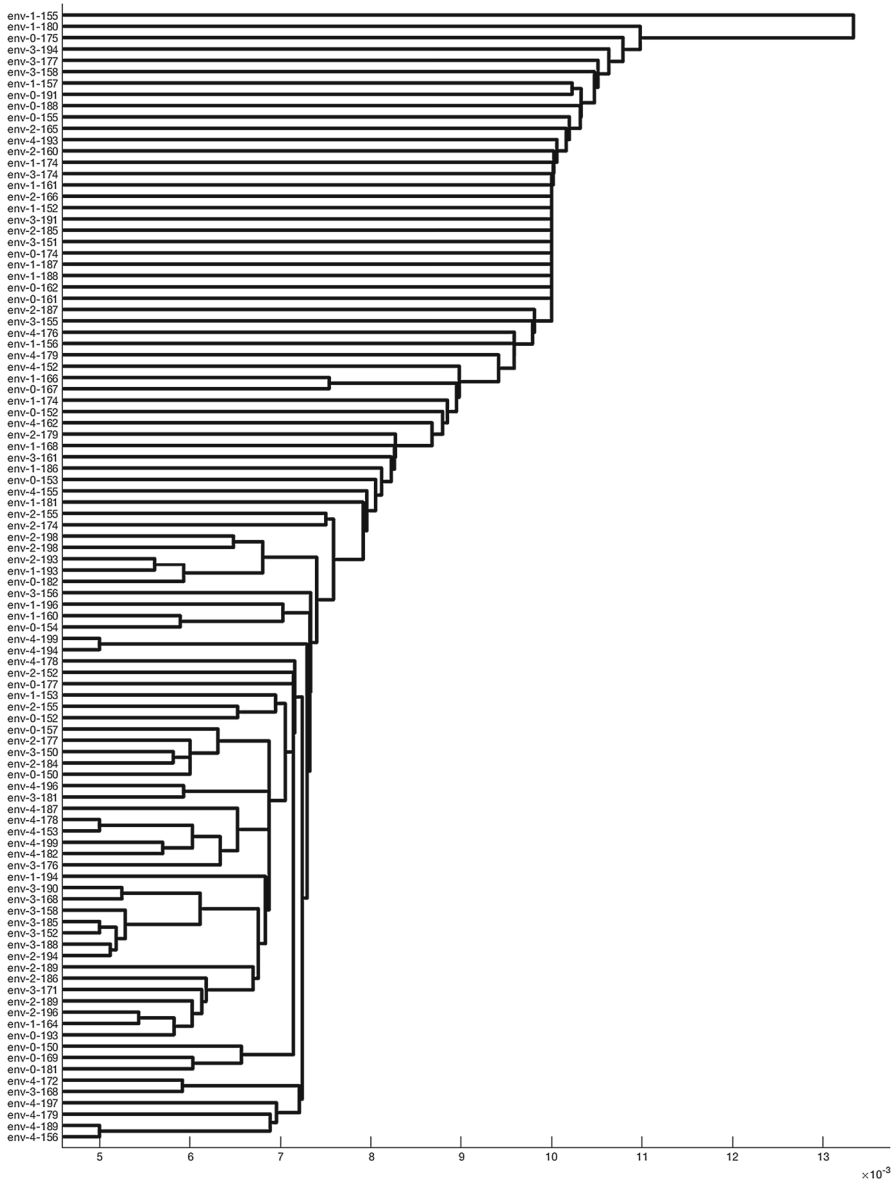
We provided a complete description of the Rips and Dowker persistence diagrams of general networks. The stability results we have provided give quantitative guarantees on the robustness of these persistence diagrams. As a building block, we proved a functorial generalization of Dowker’s theorem, which also yields an independent proof of a folklore strengthening of Dowker’s theorem. We have provided numerous examples suggesting that Dowker persistence diagrams are an appropriate method for analyzing general asymmetric networks. For a particular class of such examples, the family of cycle networks, we have fully characterized their Dowker persistence diagrams in all dimensions. Finally, we have implemented our methods for a classification task on a database of networks, and provided interpretations for our results.

We believe that the story of “persistent homology of asymmetric networks” has more aspects to be uncovered. Of particular interest to us is the analysis of alternative methods of producing simplicial complexes from asymmetric networks, for example, the *directed flag complex* construction of Dłotko et al. (2016). Yet another interesting extension to the non-metric framework has appeared in Edelsbrunner and Wagner (2017), in the context of computing generalized Čech and Rips complexes for Bregman divergences. We remark that a persistent homology framework for the directed flag complex has been proposed by Turner (2016), but the computational aspects of this construction have not been addressed in the current literature. Another approach for computing persistence diagrams from asymmetric networks, which bypasses the construction of any simplicial complex and operates directly at the chain level is given in Chowdhury and Mémoli (2018). Some other interesting questions relate to cycle networks: for example, we would like to obtain a characterization of the Rips persistence diagrams of cycle networks for any dimension  $k \geq 1$ . Finally, it is important to devise more efficient implementations for the Dowker complexes we present here. It is likely





**Fig. 12** Single linkage dendrogram corresponding to the distance matrix obtained by computing bottleneck distances between 1-dimensional Dowker persistence diagrams of our database of hippocampal networks (Sect. 7.1). Note that the 4, 3, and 2-hole arenas are well separated into clusters at threshold 0.1



**Fig. 13** Single linkage dendrogram corresponding to the distance matrix obtained by computing bottleneck distances between 1-dimensional Rips persistence diagrams of our database of hippocampal networks (Sect. 7.1). Notice that the hierarchical clustering fails to capture the correct arena types

that ideas from the literature on efficient construction of Čech complexes (Dantchev and Ivriřsimtzis 2012; Edelsbrunner and Wagner 2017) will be helpful in this regard.

**Acknowledgements** This work was supported by NSF Grants IIS-1422400 and CCF-1526513. We thank Pascal Wild and Zhengchao Wan for pointing out errors on an early preprint, and also Guilherme Vituri, Osman Okutan, and Tim Porter for useful discussions. We are especially thankful to Henry Adams for numerous helpful observations and suggestions, especially regarding the material in Appendix B, and for suggesting the proof strategy for Theorem 49.

## Appendix A: Proofs

**Proof of Lemma 8** The first inequality holds by the Algebraic Stability Theorem. For the second inequality, note that the contiguous simplicial maps in the diagrams above induce chain maps between the corresponding chain complexes, and these in turn induce equal linear maps at the level of homology vector spaces. To be more precise, first consider the maps  $t_{\delta+\eta, \delta'+\eta} \circ \varphi_\delta$  and  $\varphi_{\delta'} \circ s_{\delta, \delta'}$ . These simplicial maps induce linear maps  $(t_{\delta+\eta, \delta'+\eta} \circ \varphi_\delta)_\#$ ,  $(\varphi_{\delta'} \circ s_{\delta, \delta'})_\# : H_k(\mathfrak{F}^\delta) \rightarrow H_k(\mathfrak{G}^{\delta'+\eta})$ . Because the simplicial maps are assumed to be contiguous, we have:

$$(t_{\delta+\eta, \delta'+\eta} \circ \varphi_\delta)_\# = (\varphi_{\delta'} \circ s_{\delta, \delta'})_\#.$$

By invoking functoriality of homology, we then have:

$$(t_{\delta+\eta, \delta'+\eta})_\# \circ (\varphi_\delta)_\# = (\varphi_{\delta'})_\# \circ (s_{\delta, \delta'})_\#.$$

Analogous results hold for the other pairs of contiguous maps. Thus we obtain commutative diagrams upon passing to homology, and so  $\mathcal{H}_k(\mathfrak{F})$ ,  $\mathcal{H}_k(\mathfrak{G})$  are  $\eta$ -interleaved for each  $k \in \mathbb{Z}_+$ . Thus we get:

$$d_I(\mathcal{H}_k(\mathfrak{F}), \mathcal{H}_k(\mathfrak{G})) \leq \eta.$$

□

**Proof of Proposition 9** First we show that:

$$d_{\mathcal{N}}(X, Y) \geq \frac{1}{2} \inf \{ \max(\text{dis}(\varphi), \text{dis}(\psi)), C_{X,Y}(\varphi, \psi), C_{Y,X}(\psi, \varphi) : \varphi : X \rightarrow Y, \psi : Y \rightarrow X \text{ any maps} \}.$$

Let  $\eta = d_{\mathcal{N}}(X, Y)$ , and let  $R$  be a correspondence such that  $\text{dis}(R) = 2\eta$ . We define maps  $\varphi : X \rightarrow Y$  and  $\psi : Y \rightarrow X$  as follows: for each  $x \in X$ , set  $\varphi(x) = y$  for some  $y$  such that  $(x, y) \in R$ . Similarly, for each  $y \in Y$ , set  $\psi(y) = x$  for some  $x$  such that  $(x, y) \in R$ . Thus for any  $x \in X$ ,  $y \in Y$ , we obtain  $|\omega_X(x, \psi(y)) - \omega_Y(\varphi(x), y)| \leq 2\eta$  and  $|\omega_X(\psi(y), x) - \omega_Y(y, \varphi(x))| \leq 2\eta$ . So we have both  $C_{X,Y}(\varphi, \psi) \leq 2\eta$  and  $C_{Y,X}(\psi, \varphi) \leq 2\eta$ . Also for any  $x, x' \in X$ , we have  $(x, \varphi(x)), (x', \varphi(x')) \in R$ . Thus we also have

$$|\omega_X(x, x') - \omega_Y(\varphi(x), \varphi(x'))| \leq 2\eta.$$

So  $\text{dis}(\varphi) \leq 2\eta$  and similarly  $\text{dis}(\psi) \leq 2\eta$ . This proves the “ $\geq$ ” case.

Next we wish to show:

$$d_{\mathcal{N}}(X, Y) \leq \frac{1}{2} \inf\{\max(\text{dis}(\varphi), \text{dis}(\psi), C_{X,Y}(\varphi, \psi), C_{Y,X}(\psi, \varphi)) : \varphi : X \rightarrow Y, \psi : Y \rightarrow X \text{ any maps}\}.$$

Suppose  $\varphi, \psi$  are given, and  $\frac{1}{2} \max(\text{dis}(\varphi), \text{dis}(\psi), C_{X,Y}(\varphi, \psi), C_{Y,X}(\psi, \varphi)) = \eta$ .

Let  $R_X = \{(x, \varphi(x)) : x \in X\}$  and let  $R_Y = \{(\psi(y), y) : y \in Y\}$ . Then  $R = R_X \cup R_Y$  is a correspondence. We wish to show that for any  $z = (a, b), z' = (a', b') \in R$ ,

$$|\omega_X(a, a') - \omega_Y(b, b')| \leq 2\eta.$$

This will show that  $\text{dis}(R) \leq 2\eta$ , and so  $d_{\mathcal{N}}(X, Y) \leq \eta$ .

To see this, let  $z, z' \in R$ . Note that there are four cases: (1)  $z, z' \in R_X$ , (2)  $z, z' \in R_Y$ , (3)  $z \in R_X, z' \in R_Y$ , and (4)  $z \in R_Y, z' \in R_X$ . In the first two cases, the desired inequality follows because  $\text{dis}(\varphi), \text{dis}(\psi) \leq 2\eta$ . The inequality follows in cases (3) and (4) because  $C_{X,Y}(\varphi, \psi) \leq 2\eta$  and  $C_{Y,X}(\psi, \varphi) \leq 2\eta$ , respectively. Thus  $d_{\mathcal{N}}(X, Y) \leq \eta$ .  $\square$

**Proof of Proposition 22** It suffices to show that  $\Phi$  is a simplicial approximation to  $\mathcal{E}_{|\Sigma|}$ , i.e. whenever  $\mathcal{E}_{|\Sigma|}(x) \in |\sigma|$  for some vertex  $x \in |\Sigma^{(1)}|$  and some simplex  $\sigma \in |\Sigma|$ , we also have  $|\Phi|(x) \in |\sigma|$  (Spanier 1994, §3.4). Here  $|\sigma|$  denotes the *closed simplex* of  $\sigma$ ; for any simplex  $\sigma = [v_0, \dots, v_k]$ , this is the collection of formal convex combinations  $\sum_{i=0}^k a_i v_i$  with  $a_i \geq 0$  for each  $0 \leq i \leq k$  and  $\sum_{i=0}^k a_i = 1$ .

Let  $x = \sum_{i=0}^k a_i \sigma_i$  be a vertex in  $|\Sigma^{(1)}|$ , with each  $a_i > 0$ . Then we have  $\mathcal{E}_{|\Sigma|}(x) = \sum_{i=0}^k a_i \mathcal{B}(\sigma_i) = \sum_{i=0}^k a_i \sum_{v \in \sigma_i} v / \text{card}(\sigma_i)$ , a vertex in  $|\sigma_k|$ .

Also we have  $|\Phi|(x) = \sum_{i=0}^k a_i \Phi(\sigma_i)$ , a vertex in  $|\sigma_k|$ . Thus  $\Phi$  is a simplicial approximation to  $\mathcal{E}_{|\Sigma|}$ , and so we have  $|\Phi| \simeq \mathcal{E}_{|\Sigma|}$ .  $\square$

**Proof of Proposition 39** Let  $\delta \in \mathbb{R}$ . We first claim that  $\mathfrak{D}_{\delta}^{\text{si}}(X) = \mathfrak{D}_{\delta}^{\text{so}}(X^{\top})$ . Let  $\sigma \in \mathfrak{D}_{\delta}^{\text{si}}(X)$ . Then there exists  $x'$  such that  $\omega_X(x, x') \leq \delta$  for any  $x \in \sigma$ . Thus  $\omega_{X^{\top}}(x', x) \leq \delta$ . So  $\sigma \in \mathfrak{D}_{\delta}^{\text{so}}(X^{\top})$ . A similar argument shows the reverse containment. This proves our claim. Thus for  $\delta \leq \delta' \leq \delta''$ , we obtain the following diagram:

$$\begin{array}{ccccccc} \mathfrak{D}_{\delta}^{\text{si}}(X) & \longrightarrow & \mathfrak{D}_{\delta'}^{\text{si}}(X) & \longrightarrow & \mathfrak{D}_{\delta''}^{\text{si}}(X) & \longrightarrow & \dots \\ \parallel & & \parallel & & \parallel & & \\ \mathfrak{D}_{\delta}^{\text{so}}(X^{\top}) & \longrightarrow & \mathfrak{D}_{\delta'}^{\text{so}}(X^{\top}) & \longrightarrow & \mathfrak{D}_{\delta''}^{\text{so}}(X^{\top}) & \longrightarrow & \dots \end{array}$$

Since the maps  $\mathfrak{D}_{\delta}^{\text{si}} \rightarrow \mathfrak{D}_{\delta'}^{\text{si}}, \mathfrak{D}_{\delta}^{\text{so}} \rightarrow \mathfrak{D}_{\delta'}^{\text{so}}$  for  $\delta' \geq \delta$  are all inclusion maps, it follows that the diagrams commute. Thus at the homology level, we obtain, via functoriality of homology, a commutative diagram of vector spaces where the intervening vertical maps are isomorphisms. By the Persistence Equivalence Theorem (21), the diagrams  $\text{Dgm}_k^{\text{si}}(X)$  and  $\text{Dgm}_k^{\text{so}}(X^{\top})$  are equal. By invoking Corollary 20, we obtain  $\text{Dgm}_k^{\mathfrak{D}}(X) = \text{Dgm}_k^{\mathfrak{D}}(X^{\top})$ .  $\square$

**Proof of Proposition 41** Let  $\delta \in \mathbb{R}$ . For notational convenience, we write, for each  $k \in \mathbb{Z}_+$ ,

$$\begin{aligned} \mathfrak{D}_\delta^{\text{si}} &:= \mathfrak{D}_{\delta, X}^{\text{si}} & C_k^\delta &:= C_k(\mathfrak{D}_{\delta, X}^{\text{si}}) & \partial_k^\delta &:= \partial_k^\delta : C_k^\delta \rightarrow C_{k-1}^\delta \\ \mathfrak{D}_{\delta, S}^{\text{si}} &:= \mathfrak{D}_{\delta, S_X(z, z')}^{\text{si}} & C_k^{\delta, S} &:= C_k(\mathfrak{D}_{\delta, S_X(z, z')}^{\text{si}}) & \partial_k^{\delta, S} &:= \partial_k^{\delta, S} : C_k^{\delta, S} \rightarrow C_{k-1}^{\delta, S} \end{aligned}$$

First note that pair swaps do not affect the entry of 0-simplices into the Dowker filtration. More precisely, for any  $x \in X$ , we can unpack the definition of  $R_{\delta, X}$  (Eq. 3) to obtain:

$$[x] \in \mathfrak{D}_\delta^{\text{si}} \iff \omega_X(x, x) \leq \delta \iff \omega_X^{z, z'}(x, x) \leq \delta \iff [x] \in \mathfrak{D}_{\delta, S}^{\text{si}}.$$

Thus for any  $\delta \in \mathbb{R}$ , we have  $C_0^\delta = C_0^{\delta, S}$ . Since all 0-chains are automatically 0-cycles, we have  $\ker(\partial_0^\delta) = \ker(\partial_0^{\delta, S})$ .

Next we wish to show that  $\text{im}(\partial_1^\delta) = \text{im}(\partial_1^{\delta, S})$  for each  $\delta \in \mathbb{R}$ . Let  $\gamma \in C_1^\delta$ . We first need to show the forward inclusion, i.e. that  $\partial_1^\delta(\gamma) \in \text{im}(\partial_1^{\delta, S})$ . It suffices to show this for the case that  $\gamma$  is a single 1-simplex  $[x, x'] \in \mathfrak{D}_\delta^{\text{si}}$ ; the case where  $\gamma$  is a linear combination of 1-simplices will then follow by linearity. Let  $\gamma = [x, x'] \in \mathfrak{D}_\delta^{\text{si}}$  for  $x, x' \in X$ . Then we have the following possibilities:

- (1)  $x'' \in X \setminus \{z, z'\}$  is a  $\delta$ -sink for  $[x, x']$ .
- (2)  $z$  (or  $z'$ ) is the only  $\delta$ -sink for  $[x, x']$ , and  $x, x' \notin \{z, z'\}$ .
- (3)  $z$  (or  $z'$ ) is the only  $\delta$ -sink for  $[x, x']$ , and either  $x$  or  $x'$  belongs to  $\{z, z'\}$ .
- (4)  $z$  (or  $z'$ ) is the only  $\delta$ -sink for  $[x, x']$ , and both  $x, x'$  belong to  $\{z, z'\}$ .

In cases (1), (2), and (4), the  $(z, z')$ -pair swap has no effect on  $[x, x']$ , in the sense that we still have  $[x, x'] \in \mathfrak{D}_{\delta, S}^{\text{si}}$ . So  $[x'] - [x] = \partial_1^\delta(\gamma) = \partial_1^{\delta, S}(\gamma) \in \text{im}(\partial_1^{\delta, S})$ . Next consider case (3), and assume for notational convenience that  $[x, x'] = [z, x']$  and  $z'$  is the only  $\delta$ -sink for  $[z, x']$ . By the definition of a  $\delta$ -sink, we have  $\bar{\omega}_X(z, z') \leq \delta$  and  $\bar{\omega}_X(x', z') \leq \delta$ . Notice that we also have:

$$[z, z'], [z', x'] \in \mathfrak{D}_\delta^{\text{si}}, \text{ with } z' \text{ as a } \delta\text{-sink.}$$

After the  $(z, z')$ -pair swap, we still have  $\bar{\omega}_X^{z, z'}(x', z') \leq \delta$ , but possibly  $\bar{\omega}_X^{z, z'}(z, z') > \delta$ . So it might be the case that  $[z, x'] \notin \mathfrak{D}_{\delta, S}^{\text{si}}$ . However, we now have:

$$\begin{aligned} [z', x'] &\in \mathfrak{D}_{\delta, S}^{\text{si}}, \text{ with } z' \text{ as a } \delta\text{-sink, and} \\ [z, z'] &\in \mathfrak{D}_{\delta, S}^{\text{si}}, \text{ with } z \text{ as a } \delta\text{-sink.} \end{aligned}$$

Then we have:

$$\begin{aligned} \partial_1^\delta(\gamma) &= \partial_1^\delta([z, x']) = x' - z = z' - z + x' - z' \\ &= \partial_1^\delta([z, z']) + \partial_1^\delta([z', x']) \\ &= \partial_1^{\delta, S}([z, z']) + \partial_1^{\delta, S}([z', x']) \in \text{im}(\partial_1^{\delta, S}), \end{aligned}$$

where the last equality is defined because we have checked that  $[z, z'], [z', x'] \in \mathcal{D}_{\delta, S}^{\text{si}}$ . Thus  $\text{im}(\partial_1^\delta) \subseteq \text{im}(\partial_1^{\delta, S})$ , and the reverse inclusion follows by a similar argument.

Since  $\delta \in \mathbb{R}$  was arbitrary, this shows that  $\text{im}(\partial_1^\delta) = \text{im}(\partial_1^{\delta, S})$  for each  $\delta \in \mathbb{R}$ . Previously we had  $\ker(\partial_0^\delta) = \ker(\partial_0^{\delta, S})$  for each  $\delta \in \mathbb{R}$ . It then follows that  $H_0(\mathcal{D}_\delta^{\text{si}}) = H_0(\mathcal{D}_{\delta, S}^{\text{si}})$  for each  $\delta \in \mathbb{R}$ .

Next let  $\delta' \geq \delta \in \mathbb{R}$ , and for any  $k \in \mathbb{Z}_+$ , let  $f_k^{\delta, \delta'} : C_k^\delta \rightarrow C_k^{\delta'}$ ,  $g_k^{\delta, \delta'} : C_k^{\delta, S} \rightarrow C_k^{\delta', S}$  denote the chain maps induced by the inclusions  $\mathcal{D}_\delta^{\text{si}} \hookrightarrow \mathcal{D}_{\delta'}^{\text{si}}$ ,  $\mathcal{D}_{\delta, S}^{\text{si}} \hookrightarrow \mathcal{D}_{\delta', S}^{\text{si}}$ . Since  $\mathcal{D}_\delta^{\text{si}}$  and  $\mathcal{D}_{\delta, S}^{\text{si}}$  have the same 0-simplices at each  $\delta \in \mathbb{R}$ , we know that  $f_0^{\delta, \delta'} \equiv g_0^{\delta, \delta'}$ .

Let  $\gamma \in \ker(\partial_0^\delta) = \ker(\partial_0^{\delta, S})$ , and let  $\gamma + \text{im}(\partial_1^\delta) \in H_0(\mathcal{D}_\delta^{\text{si}})$ . Then observe that

$$\begin{aligned} (f_0^{\delta, \delta'})_\#(\gamma + \text{im}(\partial_1^\delta)) &= f_0^{\delta, \delta'}(\gamma) + \text{im}(\partial_1^{\delta'}) \quad (f_0^{\delta, \delta'} \text{ is a chain map}) \\ &= g_0^{\delta, \delta'}(\gamma) + \text{im}(\partial_1^{\delta'}) \quad (f_0^{\delta, \delta'} \equiv g_0^{\delta, \delta'}) \\ &= g_0^{\delta, \delta'}(\gamma) + \text{im}(\partial_1^{\delta', S}) \quad (\text{im}(\partial_1^{\delta'}) = \text{im}(\partial_1^{\delta', S})) \\ &= (g_0^{\delta, \delta'})_\#(\gamma + \text{im}(\partial_1^{\delta, S})). \quad (g_0^{\delta, \delta'} \text{ is a chain map}) \end{aligned}$$

Thus  $(f_0^{\delta, \delta'})_\# = (g_0^{\delta, \delta'})_\#$  for each  $\delta' \geq \delta \in \mathbb{R}$ . Since we also have  $H_0(\mathcal{D}_\delta^{\text{si}}) = H_0(\mathcal{D}_{\delta, S}^{\text{si}})$  for each  $\delta \in \mathbb{R}$ , we now apply the Persistence Equivalence Theorem (Theorem 21) to conclude the proof.  $\square$

## Appendix B: Higher dimensional Dowker persistence diagrams of cycle networks

The contents of this section rely on results in Adamaszek and Adams (2017) and Adamaszek et al. (2016). We introduce some minimalistic versions of definitions from the referenced papers to use in this section. The reader should refer to these papers for the original definitions.

Given a metric space  $(M, d_M)$  and  $m \in M$ , we will write  $\overline{B(m, \varepsilon)}$  to denote a closed  $\varepsilon$ -ball centered at  $m$ , for any  $\varepsilon > 0$ . For a subset  $X \subseteq M$  and some  $\varepsilon > 0$ , the Čech complex of  $X$  at resolution  $\varepsilon$  is defined to be the following simplicial complex:

$$\check{C}(X, \varepsilon) := \left\{ \sigma \subseteq X : \bigcap_{x \in \sigma} \overline{B(x, \varepsilon)} \neq \emptyset \right\}.$$

In the setting of metric spaces, the Čech complex coincides with the Dowker source and sink complexes. We will be interested in the special case where the underlying metric space is the circle. We write  $S^1$  to denote the circle with unit circumference. Next, for any  $n \in \mathbb{N}$ , we write  $\mathbb{X}_n := \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\}$  to denote the collection of  $n$  equally spaced points on  $S^1$  with the restriction of the arc length metric on  $S^1$ . Also let  $G_n$  denote the  $n$ -node cycle network with vertex set  $\mathbb{X}_n$  (in contrast with  $\mathbb{X}_n$ , here  $G_n$  is equipped with the asymmetric weights defined in Sect. 6.1). The connection

between  $\mathbb{X}_n$  and Dowker complexes of the cycle networks  $G_n$  is highlighted by the following observation:

**Proposition 46** *Let  $n \in \mathbb{N}$ . Then for any  $\delta \in [0, 1]$ , we have  $\check{C}(\mathbb{X}_n, \frac{\delta}{2}) = \mathfrak{D}_{n\delta, G_n}^{\text{si}}$ .*

The scaling factor arises because  $G_n$  has diameter  $\sim n$ , whereas  $\mathbb{X}_n \subseteq S^1$  has diameter  $\sim 1/2$ . This proposition provides a pedagogical step which helps us transport results from the setting of Adamaszek and Adams (2017) and Adamaszek et al. (2016) to that of the current paper.

**Proof** For  $\delta = 0$ , both the Čech and Dowker complexes consist of the  $n$  vertices, and are equal. Similarly for  $\delta = 1$ , both  $\check{C}(\mathbb{X}_n, 1)$  and  $\mathfrak{D}_{n, G_n}^{\text{si}}$  are equal to the  $(n - 1)$ -simplex.

Now suppose  $\delta \in (0, 1)$ . Let  $\sigma \in \mathfrak{D}_{n\delta, G_n}^{\text{si}}$ . Then  $\sigma$  is of the form  $[\frac{k}{n}, \frac{k+1}{n}, \dots, \frac{k+n\delta\lfloor \frac{\delta}{2} \rfloor}{n}]$  for some integer  $0 \leq k \leq n - 1$ , where the  $n\delta$ -sink is  $\frac{k+n\delta\lfloor \frac{\delta}{2} \rfloor}{n}$  and all the numerators are taken modulo  $n$ . We claim that  $\sigma \in \check{C}(\mathbb{X}_n, \frac{\delta}{2})$ . To see this, observe that  $d_{S^1}(\frac{k}{n}, \frac{k+n\delta\lfloor \frac{\delta}{2} \rfloor}{n}) \leq \delta$ , and so  $\overline{B(\frac{k}{n}, \frac{\delta}{2})} \cap \overline{B(\frac{k+n\delta\lfloor \frac{\delta}{2} \rfloor}{n}, \frac{\delta}{2})} \neq \emptyset$ . Then we have  $\sigma \in \bigcap_{i=0}^{n\delta} \overline{B(\frac{k+i}{n}, \frac{\delta}{2})}$ , and so  $\sigma \in \check{C}(\mathbb{X}_n, \frac{\delta}{2})$ .

Now let  $\sigma \in \check{C}(\mathbb{X}_n, \frac{\delta}{2})$ . Then  $\sigma$  is of the form  $[\frac{k}{n}, \frac{k+1}{n}, \dots, \frac{k+j}{n}]$  for some integer  $0 \leq k \leq n - 1$ , where  $j$  is an integer such that  $\frac{j}{n} \leq \delta$ . In this case, we have  $\sigma = \mathbb{X}_n \cap_{i=0}^j \overline{B(\frac{k+i}{n}, \delta)}$ . Then in  $G_n$ , after applying the scaling factor  $n$ , we have  $\sigma \in \mathfrak{D}_{n\delta, G_n}^{\text{si}}$ , with  $\frac{k+j}{n}$  as an  $n\delta$ -sink in  $G_n$ . This shows equality of the two simplicial complexes.  $\square$

**Theorem 47** (Theorem 3.5, Adamaszek et al. 2016) *Fix  $n \in \mathbb{N}$ , and let  $0 \leq k \leq n - 2$  be an integer. Then,*

$$\check{C}(\mathbb{X}_n, \frac{k}{2n}) \simeq \begin{cases} \bigvee^{n-k-1} S^{2l} & \text{if } \frac{k}{n} = \frac{l}{l+1}, \\ S^{2l+1} & \text{or if } \frac{l}{l+1} < \frac{k}{n} < \frac{l+1}{l+2}, \end{cases}$$

for some  $l \in \mathbb{Z}_+$ . Here  $\bigvee$  denotes the wedge sum, and  $\simeq$  denotes homotopy equivalence.

**Theorem 37** (Even dimension) *Fix  $n \in \mathbb{N}$ ,  $n \geq 3$ . If  $l \in \mathbb{N}$  is such that  $n$  is divisible by  $(l + 1)$ , and  $k := \frac{nl}{l+1}$  is such that  $0 \leq k \leq n - 2$ , then  $\text{Dgm}_{2l}^{\text{D}}(G_n)$  consists of precisely the point  $(\frac{nl}{l+1}, \frac{nl}{l+1} + 1)$  with multiplicity  $\frac{n}{l+1} - 1$ . If  $l$  or  $k$  do not satisfy the conditions above, then  $\text{Dgm}_{2l}^{\text{D}}(G_n)$  is trivial.*

**Proof of Theorem 37** Let  $l \in \mathbb{N}$  be such that  $(l + 1)$  divides  $n$  and  $0 \leq k \leq n - 2$ . Then  $\mathfrak{D}_{k, G_n}^{\text{si}} = \check{C}(\mathbb{X}_n, \frac{k}{2n})$  has the homotopy type of a wedge sum of  $(n - k - 1)$  copies of  $S^{2l}$ , by Theorem 47. Here the equality follows from Proposition 46. Notice that  $n - k - 1 = \frac{n}{l+1} - 1$ . Furthermore, by another application of Theorem 47, it is always possible to choose  $\varepsilon > 0$  small enough so that  $\mathfrak{D}_{k-\varepsilon, G_n}^{\text{si}} = \check{C}(\mathbb{X}_n, \frac{k-\varepsilon}{2n})$

and  $\mathfrak{D}_{k+\varepsilon, G_n}^{\text{si}} = \check{\mathbf{C}}(\mathbb{X}_n, \frac{k+\varepsilon}{2n})$  have the homotopy types of odd-dimensional spheres. Thus the inclusions  $\mathfrak{D}_{k-\varepsilon, G_n}^{\text{si}} \subseteq \mathfrak{D}_{k, G_n}^{\text{si}} \subseteq \mathfrak{D}_{k+\varepsilon, G_n}^{\text{si}}$  induce zero maps upon passing to homology. It follows that  $\text{Dgm}_{2l}^{\mathfrak{D}}(G_n)$  consists of the point  $(\frac{nl}{l+1}, \frac{nl}{l+1} + 1)$  with multiplicity  $\frac{n}{l+1} - 1$ .

If  $l \in \mathbb{N}$  does not satisfy the condition described above, then there does not exist an integer  $1 \leq j \leq n - 2$  such that  $j/n = l/(l + 1)$ . So for each  $1 \leq j \leq n - 2$ ,  $\mathfrak{D}_{j, G_n}^{\text{si}} = \check{\mathbf{C}}(\mathbb{X}_n, \frac{j}{2n})$  has the homotopy type of an odd-dimensional sphere by Theorem 47, and thus does not contribute to  $\text{Dgm}_{2l}^{\mathfrak{D}}(G_n)$ . If  $l$  satisfies the condition but  $k \geq n - 1$ , then  $\check{\mathbf{C}}(\mathbb{X}_n, \frac{k}{2n})$  is just the  $(n - 1)$ -simplex, hence contractible.  $\square$

Theorem 37 gives a characterization of the even dimensional Dowker persistence diagrams of cycle networks. The most interesting case occurs when considering the 2-dimensional diagrams: we see that cycle networks of an even number of nodes have an interesting barcode, even if the bars are all short-lived. For dimensions 4, 6, 8, and beyond, there are fewer and fewer cycle networks with nontrivial barcodes (in the sense that only cycle networks with number of nodes equal to a multiple of 4, 6, 8, and so on have nontrivial barcodes). For a complete picture, it is necessary to look at odd-dimensional persistence diagrams. This is made possible by the next set of constructions.

We have already recalled the definition of a Rips complex of a metric space. To facilitate the assessment of the connection to Adamaszek and Adams (2017), we temporarily adopt the notation  $\mathbf{VR}(X, \varepsilon)$  to denote the Vietoris–Rips complex of a metric space  $(X, d_X)$  at resolution  $\varepsilon > 0$ , i.e. the simplicial complex  $\{\sigma \subseteq X : \text{diam}(\sigma) \leq \varepsilon\}$ .

**Theorem 48** (Theorem 9.3, Proposition 9.5, Adamaszek and Adams 2017) *Let  $0 < r < \frac{1}{2}$ . Then there exists a map  $T_r : \text{pow}(S^1) \rightarrow \text{pow}(S^1)$  and a map  $\pi_r : S^1 \rightarrow S^1$  such that there is an induced homotopy equivalence*

$$\mathbf{VR}(T_r(X), \frac{2r}{1+2r}) \xrightarrow{\cong} \check{\mathbf{C}}(X, r).$$

*Next suppose  $X \subseteq S^1$  and let  $0 < r \leq r' < \frac{1}{2}$ . Then there exists a map  $\eta : S^1 \rightarrow S^1$  such that the following diagram commutes:*

$$\begin{array}{ccc} \mathbf{VR}(T_r(X), \frac{2r}{1+2r}) & \xrightarrow{\eta} & \mathbf{VR}(T_{r'}(X), \frac{2r'}{1+2r'}) \\ \pi_r \downarrow \simeq & & \pi_{r'} \downarrow \simeq \\ \check{\mathbf{C}}(X, r) & \xrightarrow{\subseteq} & \check{\mathbf{C}}(X, r') \end{array}$$

**Theorem 49** *Consider the setup of Theorem 48. If  $\check{\mathbf{C}}(X, r)$  and  $\check{\mathbf{C}}(X, r')$  are homotopy equivalent, then the inclusion map between them is a homotopy equivalence.*

Before providing the proof, we show how it implies Theorem 38.

**Theorem 38** (Odd dimension) *Fix  $n \in \mathbb{N}$ ,  $n \geq 3$ . Then for  $l \in \mathbb{N}$ , define  $M_l := \{m \in \mathbb{N} : \frac{nl}{l+1} < m < \frac{n(l+1)}{l+2}\}$ . If  $M_l$  is empty, then  $\text{Dgm}_{2l+1}^{\mathfrak{D}}(G_n)$  is trivial. Otherwise, we have:*



$$\text{Dgm}_{2l+1}^{\mathfrak{D}}(G_n) = \left\{ \left( a_l, \left\lceil \frac{n(l+1)}{l+2} \right\rceil \right) \right\},$$

where  $a_l := \min \{m \in M_l\}$ . We use set notation (instead of multisets) to mean that the multiplicity is 1.

**Proof of Theorem 38** By Proposition 46 and Theorem 47, we know that  $\mathfrak{D}_{k,G_n}^{\text{si}} = \check{C}(\mathbb{X}_n, \frac{k}{2n}) \simeq S^1$  for integers  $0 < k < \frac{n}{2}$ . Let  $b \in \mathbb{N}$  be the greatest integer less than  $n/2$ . Then by Theorem 49, we know that each inclusion map in the following chain is a homotopy equivalence:

$$\mathfrak{D}_{1,G_n}^{\text{si}} \subseteq \dots \subseteq \mathfrak{D}_{b,G_n}^{\text{si}} = \mathfrak{D}_{\lceil n/2 \rceil, G_n}^{\text{si}}.$$

It follows that  $\text{Dgm}_1^{\mathfrak{D}}(G_n) = \left\{ \left( 1, \left\lceil \frac{n}{2} \right\rceil \right) \right\}$ . The notation in the last equality means that  $\mathfrak{D}_{b,G_n}^{\text{si}} = \mathfrak{D}_{\delta,G_n}^{\text{si}}$  for all  $\delta \in [b, b + 1)$ , where  $b + 1 = \lceil n/2 \rceil$ .

In the more general case, let  $l \in \mathbb{N}$  and let  $M_l$  be as in the statement of the result. Suppose first that  $M_l$  is empty. Then by Proposition 46 and Theorem 47, we know that  $\mathfrak{D}_{k,G_n}^{\text{si}}$  has the homotopy type of a wedge of even-dimensional spheres or an odd-dimensional sphere of dimension strictly different from  $(2l + 1)$ , for any choice of integer  $k$ . Thus  $\text{Dgm}_{2l+1}^{\mathfrak{D}}(G_n)$  is trivial.

Next suppose  $M_l$  is nonempty. By another application of Proposition 46 and Theorem 47, we know that  $\mathfrak{D}_{k,G_n}^{\text{si}} = \check{C}(\mathbb{X}_n, \frac{k}{2n}) \simeq S^{2l+1}$  for integers  $\frac{nl}{l+1} < k < \frac{n(l+1)}{l+2}$ . Write  $a_l := \min \{m \in M_l\}$  and  $b_l := \max \{m \in M_l\}$ . Then by Theorem 49, we know that each inclusion map in the following chain is a homotopy equivalence:

$$\mathfrak{D}_{a_l,G_n}^{\text{si}} \subseteq \dots \subseteq \mathfrak{D}_{b_l,G_n}^{\text{si}} = \mathfrak{D}_{\lceil n(l+1)/(l+2) \rceil, G_n}^{\text{si}}.$$

It follows that  $\text{Dgm}_{2l+1}^{\mathfrak{D}}(G_n) = \left\{ \left( a_l, \left\lceil \frac{n(l+1)}{l+2} \right\rceil \right) \right\}$ . □

It remains to provide a proof of Theorem 49. For this, we need some additional machinery.

**Cyclic maps and winding fractions** We introduce some more terms from Adamaszek and Adams (2017), but for efficiency, we try to minimize the scope of the definitions to only what is needed for our purpose. Recall that we write  $S^1$  to denote the circle with unit circumference. Thus we naturally identify any  $x \in S^1$  with a point in  $[0, 1)$ . We fix a choice of  $0 \in S^1$ , and for any  $x, x' \in S^1$ , the length of a clockwise arc from  $x$  to  $x'$  is denoted by  $\overrightarrow{d}_{S^1}(x, x')$ . Then, for any finite subset  $X \subseteq S^1$  and any  $r \in (0, 1/2)$ , the *directed Vietoris–Rips graph*  $\overrightarrow{\text{VR}}(X, r)$  is defined to be the graph with vertex set  $X$  and edge set  $\{(x, x') : 0 < \overrightarrow{d}_{S^1}(x, x') < r\}$ . Next, let  $\overrightarrow{G}$  be a Vietoris–Rips graph such that the vertices are enumerated as  $x_0, x_1, \dots, x_{n-1}$ , according to the *clockwise* order in which they appear. A *cyclic map* between  $\overrightarrow{G}$  and a Vietoris–Rips graph  $\overrightarrow{H}$  is a map of vertices  $f$  such that for each edge  $(x, x') \in \overrightarrow{G}$ , we have either  $f(x) = f(x')$ , or  $(f(x), f(x')) \in \overrightarrow{H}$ , and  $\sum_{i=0}^{n-1} \overrightarrow{d}_{S^1}(f(x_i), f(x_{i+1})) = 1$ . Here  $x_n := x_0$ .

Next, the *winding fraction* of a Vietoris–Rips graph  $\vec{G}$  with vertex set  $V(\vec{G})$  is defined to be the infimum of numbers  $\frac{k}{n}$  such that there is an order-preserving map  $V(\vec{G}) \rightarrow \mathbb{Z}/n\mathbb{Z}$  such that each edge is mapped to a pair of numbers at most  $k$  apart. A key property of the winding fraction, denoted  $\text{wf}$ , is that if there is a cyclic map between Vietoris–Rips graphs  $\vec{G} \rightarrow \vec{H}$ , then  $\text{wf}(\vec{G}) \leq \text{wf}(\vec{H})$ .

**Theorem 50** (Corollary 4.5, Proposition 4.9, Adamaszek and Adams 2017) *Let  $X \subseteq S^1$  be a finite set and let  $0 < r < \frac{1}{2}$ . Then,*

$$\mathbf{VR}(X, r) \simeq \begin{cases} S^{2l+1} & : \frac{l}{2l+1} < \text{wf}(\overrightarrow{\mathbf{VR}}(X, r)) < \frac{l+1}{2l+3} \text{ for some } l \in \mathbb{Z}_+, \\ \bigvee^j S^{2l} & : \text{wf}(\overrightarrow{\mathbf{VR}}(X, r)) = \frac{l}{2l+1}, \text{ for some } j \in \mathbb{N}. \end{cases}$$

Next let  $X' \subseteq S^1$  be another finite set, and let  $r \leq r' < \frac{1}{2}$ . Suppose  $f : \overrightarrow{\mathbf{VR}}(X, r) \rightarrow \overrightarrow{\mathbf{VR}}(X', r')$  is a cyclic map between Vietoris–Rips graphs and  $\frac{l}{2l+1} < \text{wf}(\overrightarrow{\mathbf{VR}}(X, r)) \leq \text{wf}(\overrightarrow{\mathbf{VR}}(X', r')) < \frac{l+1}{2l+3}$ . Then  $f$  induces a homotopy equivalence between  $\mathbf{VR}(X, r)$  and  $\mathbf{VR}(X', r')$ .

We now have the ingredients for a proof of Theorem 49.

**Proof of Theorem 49** Since the maps  $\pi_r$  and  $\pi_{r'}$  induce homotopy equivalences, it follows that

$$\mathbf{VR}(T_r(X), \frac{2r}{1+2r}) \simeq \mathbf{VR}(T_{r'}(X), \frac{2r'}{1+2r'}).$$

By the characterization result in Theorem 50, we know that there exists  $l \in \mathbb{Z}_+$  such that

$$\frac{l}{2l+1} < \text{wf}(\overrightarrow{\mathbf{VR}}(T_r(X), \frac{2r}{1+2r})) \leq \text{wf}(\overrightarrow{\mathbf{VR}}(T_{r'}(X), \frac{2r'}{1+2r'})) < \frac{l+1}{2l+3}.$$

The map  $\eta$  in Theorem 48 appears in (Adamaszek and Adams 2017, Proposition 9.5) through an explicit construction. Moreover, it is shown that  $\eta$  induces a cyclic map  $\text{wf}(\overrightarrow{\mathbf{VR}}(T_r(X), \frac{2r}{1+2r})) \rightarrow \text{wf}(\overrightarrow{\mathbf{VR}}(T_{r'}(X), \frac{2r'}{1+2r'}))$ . Thus by Theorem 50,  $\eta$  induces a homotopy equivalence between  $\mathbf{VR}(T_r(X), \frac{2r}{1+2r})$  and  $\mathbf{VR}(T_{r'}(X), \frac{2r'}{1+2r'})$ . Finally, the commutativity of the diagram in Theorem 48 shows that the inclusion  $\check{C}(X, r) \subseteq \check{C}(X, r')$  induces a homotopy equivalence.  $\square$

**Remark 51** The analogue of Theorem 49 for Čech complexes appears as Proposition 4.9 of Adamaszek and Adams (2017) for Vietoris–Rips complexes. We prove Theorem 49 by connecting Čech and Vietoris–Rips complexes using Proposition 9.5 of Adamaszek and Adams (2017). However, as remarked in §9 of Adamaszek and Adams (2017), one could prove Theorem 49 directly using a parallel theory of winding fractions for Čech complexes.

## References

- Adamaszek, M., Adams, H.: The Vietoris–Rips complexes of a circle. *Pac. J. Math.* **290**(1), 1–40 (2017)
- Adamaszek, M., Adams, H., Frick, F., Peterson, C., Previte-Johnson, C.: Nerve complexes of circular arcs. *Discrete Comput. Geom.* **56**(2), 251–273 (2016)
- Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A.: Systemic risk and stability in financial networks. *Am. Econ. Rev.* **105**(2), 564–608 (2015)
- Atkin, R.H.: From cohomology in physics to q-connectivity in social science. *Int. J. Man Mach. Stud.* **4**(2), 139–167 (1972)
- Atkin, R.: *Mathematical Structure in Human Affairs*. Crane, Russak, New York (1975)
- Barmak, J.A.: *Algebraic Topology of Finite Topological Spaces and Applications*, vol. 2032. Springer, Belrin (2011)
- Burago, D., Burago, Y., Ivanov, S.: *A Course in Metric Geometry*, Volume 33 of AMS Graduate Studies in Mathematics. American Mathematical Society, Providence (2001)
- Burghelca, D., Dey, T.K.: Topological persistence for circle-valued maps. *Discrete Comput. Geom.* **50**(1), 69–98 (2013)
- Björner, A.: Topological methods. *Handb. Comb.* **2**, 1819–1872 (1995)
- Björner, A., Korte, B., Lovász, L.: Homotopy properties of greedoids. *Adv. Appl. Math.* **6**(4), 447–494 (1985)
- Bauer, U., Lesnick, M.: Induced matchings of barcodes and the algebraic stability of persistence. In: *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, p. 355. ACM, London (2014)
- Barabási, A.-L., Oltvai, Z.N.: Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* **5**(2), 101–113 (2004)
- Carlsson, G.: Topology and data. *Bull. Am. Math. Soc.* **46**(2), 255–308 (2009)
- Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L.J., Oudot, S.Y.: Proximity of persistence modules and their diagrams. In: *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, pp. 237–246. ACM, London (2009)
- Chazal, F., Cohen-Steiner, D., Guibas, L.J., Mémoli, F., Oudot, S.Y.: Gromov–Hausdorff stable signatures for shapes using persistence. In: *Computer Graphics Forum*, vol. 28, pp. 1393–1403. Wiley Online Library, New York (2009)
- Chowdhury, S., Dai, B., Mémoli, F.: The importance of forgetting: limiting memory improves recovery of topological characteristics from neural data (2017). arXiv preprint [arXiv:1710.11279](https://arxiv.org/abs/1710.11279)
- Chowdhury, S., Dai, B., Mémoli, F.: Topology of stimulus space via directed network persistent homology. *Cosyne Abstracts 2017* (2017)
- Carlsson, G., De Silva, V.: Zigzag persistence. *Found. Comput. Math.* **10**(4), 367–405 (2010)
- Chazal, F., De Silva, V., Glisse, M., Oudot, S.: *The Structure and Stability of Persistence Modules*. Springer, Berlin (2016)
- Chazal, F., De Silva, V., Oudot, S.: Persistence stability for geometric complexes. *Geom. Dedic.* **173**(1), 193–214 (2014)
- Carstens, C.J., Horadam, K.J.: Persistent homology of collaboration networks. *Math. Problems Eng.* **2013**, 815035 (2013)
- Curto, C., Itskov, V.: Cell groups reveal structure of stimulus space. *PLoS Comput. Biol.* **4**(10), e1000205 (2008)
- Carlsson, G., Mémoli, F.: Persistent clustering and a theorem of J. Kleinberg (2008). arXiv preprint [arXiv:0808.2241](https://arxiv.org/abs/0808.2241)
- Carlsson, G., Mémoli, F.: Characterization, stability and convergence of hierarchical clustering methods. *J. Mach. Learn. Res.* **11**, 1425–1470 (2010)
- Chowdhury, S., Mémoli, F.: Metric structures on networks and applications. In: *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1470–1472 (2015)
- Chowdhury, S., Mémoli, F.: Distances between directed networks and applications. In: *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6420–6424. IEEE, Washington (2016)
- Chowdhury, S., Mémoli, F.: Persistent homology of directed networks. In: *2016 50th Asilomar Conference on Signals, Systems and Computers*, pp. 77–81. IEEE, Washington (2016)
- Chowdhury, S., Mémoli, F.: Distances and isomorphism between networks and the stability of network invariants (2017). arXiv preprint [arXiv:1708.04727](https://arxiv.org/abs/1708.04727)

- Chowdhury, S., Mémoli, F.: Persistent path homology of directed networks. In: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1152–1169. SIAM, Philadelphia (2018)
- Carlsson, G.E., Mémoli, F., Ribeiro, A., Segarra, S.: Hierarchical quasi-clustering methods for asymmetric networks. In: Proceedings of the 31th International Conference on Machine Learning, pp. 352–360 (2014)
- Chazal, F., Oudot, S.Y.: Towards persistence-based reconstruction in Euclidean spaces. In: Proceedings of the Thirtieth Annual Symposium on Computational Geometry, pp. 232–241. ACM, New York (2008)
- Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. *Discrete Comput. Geom.* **37**(1), 103–120 (2007)
- Collins, A., Zomorodian, A., Carlsson, G., Guibas, L.J.: A barcode shape descriptor for curve point cloud data. *Comput. Graph.* **28**(6), 881–894 (2004)
- Carlsson, G., Zomorodian, A., Collins, A., Guibas, L.J.: Persistence barcodes for shapes. *Int. J. Shape Model.* **11**(02), 149–187 (2005)
- Dey, T.K., Fan, F., Wang, Y.: Computing topological persistence for simplicial maps. In: Proceedings of the Thirtieth Annual Symposium on Computational Geometry, pp. 345. ACM, London (2014)
- Diotko, P., Hess, K., Levi, R., Nolte, M., Reimann, M., Scolamiero, M., Turner, K., Muller, E., Markram, H.: Topological analysis of the connectome of digital reconstructions of neural microcircuits (2016). arXiv preprint [arXiv:1601.01580](https://arxiv.org/abs/1601.01580)
- Dantchev, S., Ivrişimtzis, I.: Efficient construction of the Čech complex. *Comput. Graph.* **36**(6), 708–713 (2012)
- Dabaghian, Y., Mémoli, F., Frank, L., Carlsson, G.: A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Comput. Biol.* **8**(8), e1002581 (2012)
- Dowker, C.H.: Homology groups of relations. *Ann. Math.* **56**, 84–95 (1952)
- De Silva, V., Carlsson, G.: Topological estimation using witness complexes. In: Proceedings of the Symposium on Point-Based Graphics, pp. 157–166 (2004)
- Elliott, M., Golub, B., Jackson, M.O.: Financial networks and contagion. *Am Econ Rev* **104**(10), 3115–3153 (2014)
- Edelsbrunner, H., Harer, J.: Persistent homology—a survey. *Contemp Math* **453**, 257–282 (2008)
- Edelsbrunner, H., Harer, J.: *Computational Topology: An Introduction*. American Mathematical Soc, Providence (2010)
- Efrat, A., Itai, A., Katz, M.J.: Geometry helps in bottleneck matching and related problems. *Algorithmica* **31**(1), 1–28 (2001)
- Edelsbrunner, H., Jabłoński, G., Mrozek, M.: The persistent homology of a self-map. *Found. Comput. Math.* **15**(5), 1213–1244 (2015)
- Easley, D., Kleinberg, J.: *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York (2010)
- Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. *Discrete Comput. Geom.* **28**(4), 511–533 (2002)
- Edelsbrunner, H., Morozov, D.: Persistent homology: theory and practice. In: European Congress of Mathematics Kraków, 2–7 July, 2012, pp. 31–50 (2014)
- Engström, A.: Complexes of directed trees and independence complexes. *Discrete Math.* **309**(10), 3299–3309 (2009)
- Edelsbrunner, H., Wagner, H.: *Topological data analysis with Bregman divergences* (2017)
- Frosini, P., Landi, C.: Size theory as a topological tool for computer vision. *Pattern Recogn. Image Anal.* **9**(4), 596–603 (1999)
- Frosini, P.: Measuring shapes by size functions. In: *Intelligent Robots and Computer Vision X: Algorithms and Techniques*, pp. 122–133. International Society for Optics and Photonics (1992)
- Ghrist, R.: Barcodes: the persistent topology of data. *Bull. Am. Math. Soc.* **45**(1), 61–75 (2008)
- Ghrist, R.: *Elementary Applied Topology*. Createspace, Scotts Valley (2014)
- Giusti, C., Pastalkova, E., Curto, C., Itskov, V.: Clique topology reveals intrinsic geometric structure in neural correlations. *Proc. Natl. Acad. Sci.* **112**(44), 13455–13460 (2015)
- Horak, D., Maletić, S., Rajković, M.: Persistent homology of complex networks. *J. Stat. Mech. Theory Exp.* **2009**(03), P03034 (2009)

- Hiraoka, Y., Nakamura, T., Hirata, A., Escobar, E.G., Matsue, K., Nishiura, Y.: Hierarchical structures of amorphous solids characterized by persistent homology. *Proc. Natl. Acad. Sci.* **113**(26), 7035–7040 (2016)
- Huson, D.H., Rupp, R., Scornavacca, C.: *Phylogenetic Networks: Concepts, Algorithms and Applications*. Cambridge University Press, Cambridge (2010)
- Johnson, J.: *Hypernetworks in the Science of Complex Systems*, vol. 3. World Scientific, Singapore (2013)
- Khalid, A., Kim, B.S., Chung, M.K., Ye, J.C., Jeon, D.: Tracing the evolution of multi-scale functional networks in a mouse model of depression using persistent brain network homology. *NeuroImage* **101**, 351–363 (2014)
- Kumar, R., Novak, J., Tomkins, A.: Structure and evolution of online social networks. In: *Link Mining: Models, Algorithms, and Applications*, pp. 337–357. Springer, Berlin (2010)
- Kalton, N.J., Ostrovskii, M.I.: Distances between Banach spaces. In: Fliess, M. (ed.) *Forum Mathematicum*, vol. 11, pp. 17–48. Walter de Gruyter, Berlin (1999)
- Kozlov, D.: *Combinatorial Algebraic Topology*, vol. 21. Springer, Berlin (2007)
- Krishnamoorthy, B., Provan, S., Tropsha, A.: A topological characterization of protein structure. In: Pardalos, P.M., Boginski, V.L., Alkis, V. (eds.) *Data Mining in Biomedicine*, pp. 431–455. Springer, Berlin (2007)
- Lee, H., Chung, M.K., Kang, H., Kim, B.-N., Lee, D.S.: Computing the shape of brain networks using graph filtration and Gromov–Hausdorff metric. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*, pp. 302–309. Springer, Berlin (2011)
- Lefschetz, S.: *Algebraic Topology*, vol. 27. American Mathematical Soc, Providence (1942)
- Masys, A.J.: *Networks and Network Analysis for Defence and Security*. Springer, Berlin (2014)
- Mac Lane, S.: *Categories for the Working Mathematician*, vol. 5. Springer, Berlin (2013)
- Munkres, J.R.: *Elements of Algebraic Topology*, vol. 7. Addison-Wesley, Reading (1984)
- Newman, M.E.J.: The structure and function of complex networks. *SIAM Rev.* **45**(2), 167–256 (2003)
- Dowker's theorem. <https://ncatlab.org/nlab/show/Dowker%27s+theorem>. Accessed 24 Apr 2017
- O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34**(1), 171–175 (1971)
- Pessoa, L.: Understanding brain networks and brain organization. *Phys. Life Rev.* **11**(3), 400–435 (2014)
- Petri, G., Scolamiero, M., Donato, I., Vaccarino, F.: Topological strata of weighted complex networks. *PLoS ONE* **8**(6), e66506 (2013)
- Robins, V.: Towards computing homology from finite approximations. *Topol. Proc.* **24**, 503–532 (1999)
- Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* **52**(3), 1059–1069 (2010)
- Schmiedl, F.: Shape matching and mesh segmentation. Dissertation, Technische Universität München, München (2015)
- Sporns, O., Kötter, R.: Motifs in brain networks. *PLoS Biol.* **2**(11), e369 (2004)
- Spanier, E.H.: *Algebraic Topology*, vol. 55. Springer, Berlin (1994)
- Sporns, O.: *Networks of the Brain*. MIT Press, Cambridge (2011)
- Sporns, O.: *Discovering the Human Connectome*. MIT Press, Cambridge (2012)
- Turner, K.: Generalizations of the Rips filtration for quasi-metric spaces with persistent homology stability results (2016). arXiv preprint [arXiv:1608.00365](https://arxiv.org/abs/1608.00365)
- Tausz, A., Vejdemo-Johansson, M., Adams, H.: Javaplex: a research software package for persistent (co)homology (2011)
- Weinberger, S.: What is... persistent homology? *Not. AMS* **58**(1), 36–39 (2011)
- Xia, K., Wei, G.-W.: Persistent homology analysis of protein structure, flexibility, and folding. *Int. J. Numer. Methods Biomed. Eng.* **30**(8), 814–844 (2014)
- Zomorodian, A., Carlsson, G.: Computing persistent homology. *Discrete Comput. Geom.* **33**(2), 249–274 (2005)