

# Persistent Homology of Directed Networks

Samir Chowdhury and Facundo Mémoli

**Abstract**—While persistent homology has been successfully used to provide topological summaries of point cloud data, the question of computing persistent homology of graphs or networks remains unclear. In particular, the existing literature does not provide a treatment of persistent homology for directed networks that is sensitive to asymmetry. We study a method for constructing simplicial complexes from weighted, directed networks that captures directionality information, and we are able to prove that the persistent homology of such complexes is stable with respect to a certain notion of network distance. We illustrate our construction on a database of simulated hippocampal networks.

## I. INTRODUCTION

When faced with the task of analyzing a complex network, a common approach is to extract certain invariants of the network, and infer properties of the original network from the properties of these invariants. In the most general setting, a network on  $n$  nodes is an  $n \times n$  matrix with real-valued entries. The invariants that one can extract from a network depends on the additional conditions satisfied by the matrix. In the simple case of an undirected weighted network on  $n$  nodes, which can be viewed as a symmetric  $n \times n$  matrix with real-valued entries, the spectral theorem guarantees a set of real eigenvalues that can yield information about the dynamics of the network. An even simpler situation is that of an  $n$ -point metric space, for which the adjacency matrix would have zeros on the diagonal, and would also satisfy metric properties such as symmetry and the triangle inequality. In this simpler setting, one mainstream approach is to extract an invariant known in the applied mathematics literature as a *persistence diagram*—a set of intervals that represents the topology of the space, parametrized by resolution. This diagram is obtained by first converting the metric data into a family of topological objects called *simplicial complexes*, organizing them by an inclusion hierarchy called a *filtration*, and then applying the theory of *persistent homology* (PH) to this filtration. For the sake of brevity, we simply assert that PH has been successfully applied to numerous problems in the sciences, and point the reader to [1]–[3] for accounts of these applications written by the pioneers of the theory. For an application to signal processing, see [4].

The success of persistent homology in the setting of finite metric spaces leads us to believe that persistence diagrams would be a powerful invariant in the general network setting as well, if only one could find an appropriate adaptation that does not require any metric assumptions. Some past approaches that have claimed to perform persistent homology

on networks actually assume that the underlying data satisfies metric properties, thus working in the simplified setting where PH is fully understood [5], [6]. A more general approach is followed in [7]–[10], where the input data sets are required to be symmetric matrices, and the main construction involves a *Rips complex* (also called a *weighted clique complex*), which we describe for now as an object that is sensitive to the underlying graph structure of the network. However, there are two challenges facing this method: (1) directed networks are not treated in a way that is sensitive to asymmetry, and (2) the clique complex, which can be an  $n$ -dimensional object for  $n$ -point networks, is built up using only 1-dimensional data and thus discards additional data which may be of importance to the network.

In this paper, we study a construction that is known in the finite metric space setting as the *Dowker complex* [11]. We extend both the Dowker complex and clique complex construction to the setting of directed networks with real-valued weights, without assuming any metric properties. Following the line of work in [12], where stability of persistence diagrams arising from finite metric spaces was established, we formulate similar stability results for persistence diagrams arising from taking the Rips or Dowker filtrations of a network. We also show that unlike the Rips filtration, the Dowker filtration is sensitive to asymmetry.

We test our method on a database of simulated hippocampal networks, obtained by modeling the brain activity of an animal that travels through five types of arenas having different numbers of obstacles. We follow a model used in [13], [14], in which a simplicial complex representation of the animal’s spatial environment was built by tracking which groups of cells in the hippocampus “spiked” together. We extend this model to take into account both the *frequency* with which cell groups fire together, and also the *causality relation* between cell groups that fire one after another. We record this data as a Markov chain that we interpret as a network, and by applying the Dowker filtration construction, we are able to recover the number of obstacles contained in each arena. In particular, by using our stability results, we are able to show that the networks which arise from the same environment have a higher tendency to cluster together than networks arising from different environments.

## II. BACKGROUND ON PERSISTENT HOMOLOGY

The following is a high-level overview of persistent homology; details can be found in [15]. Given a finite set  $X$  with cardinality  $|X|$ , a *simplicial complex* is a subset  $K_X$  of the power set  $\text{Pow}(X)$  such that whenever  $\sigma \in K_X$ , we also have  $\tau \in K_X$  for any subset  $\tau \subseteq \sigma$ . The elements  $\sigma \in K_X$  are

Samir Chowdhury is a Ph.D. student with the Department of Mathematics at The Ohio State University. Facundo Mémoli is with the Departments of Mathematics and Computer Science and Engineering at The Ohio State University. Email at chowdhury.57@osu.edu, memoli@math.osu.edu.

called *simplices*, and any subset  $\tau \subseteq \sigma$  is called a *face*. More specifically, simplices consisting of  $n + 1$  points (for  $n \geq 0$ ) are called  $n$ -simplices. To a simplicial complex  $K_X$ , one may associate a vector space  $H_k(K_X)$  for each dimension  $k \leq |X|$ , obtained by a method called *computing homology with field coefficients*.

Persistent homology enters the picture in the following way: for a finite set of points, one induces a family of simplicial complexes  $K_X^\varepsilon$  for a range of values of  $\varepsilon \in \mathbb{R}$ . The simplicial complexes are produced in a way such that  $K_X^\varepsilon \subseteq K_X^{\varepsilon'}$  for any  $\varepsilon' \geq \varepsilon$ . The nested family of simplicial complexes, along with the inclusion maps, is called a *filtration*, and is denoted by  $\{K_X^\varepsilon \hookrightarrow K_X^{\varepsilon'}\}_{\varepsilon \leq \varepsilon'}$ . For a given dimension  $k$ , computing homology with field coefficients for each of these simplicial complexes then yields a sequence of vector spaces with linear transformations mapping between them, denoted  $\{H_k(K_X^\varepsilon) \rightarrow H_k(K_X^{\varepsilon'})\}_{\varepsilon \leq \varepsilon'}$ . In practical applications, one usually takes coefficients in the field  $\mathbb{Z}_2$ . The dimension of each such vector space is called its *Betti- $k$  number*. Each vector space is generated by *homology classes* that correspond to topological features of the underlying simplicial complex. The crux of persistent homology is to track the birth and death of these topological features in  $K_X^\varepsilon$  for different values of  $\varepsilon$ . The lifespan of each homology class is an interval, and a *persistence barcode* is simply the collection of all such intervals associated to the space. This collection is visualized as a stack of intervals, hence the name ‘barcode.’

An alternate, equivalent representation of the information in a persistence barcode is a *persistence diagram*, which consists of the birth and death times of features as pairs  $(b, d)$  in the extended plane  $\overline{\mathbb{R}}^2$ , counted with multiplicity. Because  $d \geq b$ , these points lie above the diagonal, and for a technical reason, the persistence diagram is defined to include the diagonal with infinite multiplicity. This permits the definition of the *bottleneck distance* between two persistence diagrams  $D_k(X)$  and  $D_k(Y)$  in dimension  $k$ :

$$d_B(D_k(X), D_k(Y)) := \min \left\{ \max_{x \in D_k(X)} \|x - \varphi(x)\|_\infty : \varphi : D_k(X) \rightarrow D_k(Y) \text{ a bijection} \right\}.$$

It is worth noting that the Betti numbers of a topological space have an intuitive meaning: Betti-0 is the number of connected components, Betti-1 gives the number of 1-dimensional holes, and Betti-2 gives the number of 2-dimensional voids.

We computed persistence using the Javaplex package for Matlab; a full description of Javaplex can be found in [16].

### III. NETWORKS AND NETWORK DISTANCES

A *network* is a finite set  $X$  together with an *edge weight function*  $\omega_X : X \times X \rightarrow \mathbb{R}$ . The collection of all networks will be denoted  $\mathcal{N}$ . Note that we do not assume any metric properties in defining a network. Given a finite set  $X$  and two edge weight functions  $\omega_X, \omega'_X$  defined on it, we can use the  $\ell^\infty$  distance as a measure of network similarity between  $(X, \omega_X)$  and  $(X, \omega'_X)$ :

$$\|\omega_X - \omega'_X\|_{\ell^\infty(X \times X)} := \max_{x, x' \in X} |\omega_X(x, x') - \omega'_X(x, x')|.$$

A slight generalization of the  $\ell^\infty$  distance is required when dealing with networks having different sizes: Given two sets  $X$  and  $Y$ , we need to decide how to match up points of  $X$  with points of  $Y$ . Any such matching will yield a subset  $R \subseteq X \times Y$  such that  $\pi_X(R) = X$  and  $\pi_Y(R) = Y$ , where  $\pi_X$  and  $\pi_Y$  are the projection maps from  $X \times Y$  to  $X$  and  $Y$ , respectively. The formal term for such a matching is a *correspondence*. The *distortion* of a correspondence is defined as follows:

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

We denote the set of all correspondences between  $X$  and  $Y$  by  $\mathcal{R}(X, Y)$ . Note that the set  $X \times Y$  is always a correspondence, so  $\mathcal{R}(X, Y)$  is always nonempty. We then minimize the distortion over all correspondences, and define the *network distance* as follows:

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

The motivation for defining the network distance in this way is that when restricted to the case of comparing finite metric spaces  $(X, d_X), (Y, d_Y)$ , the network distance is exactly the Gromov-Hausdorff distance that was used in [12], [17] to prove stability results for persistence diagrams. Thus the stability results in our paper can be viewed as direct extensions of results in the existing literature. The network distance as defined above has previously been used to study network motifs and other invariants [18]–[20].

### IV. THE RIPS COMPLEX OF A NETWORK

For a metric space  $(X, d_X)$ , the diameter of a subset  $\sigma \subseteq X$  is defined as  $\text{diam}(\sigma) := \max_{x, x' \in \sigma} d_X(x, x')$ . The *Rips complex* of a metric space  $(X, d_X)$  is then defined for each  $\delta \in \mathbb{R}$  as  $\mathfrak{R}_X^\delta := \{\sigma \in \text{Pow}(X) : \text{diam}(\sigma) \leq \delta\}$  [21].

Since diameter is a metric concept, it does not apply directly to general networks. An analogous notion for networks can be defined as follows: for any  $(X, \omega_X) \in \mathcal{N}$ , we define the *weight of a subset* as a map  $\text{wgt}_X : \text{Pow}(X) \rightarrow \mathbb{R}$  given by:

$$\text{wgt}_X(\sigma) := \max_{x, x' \in \sigma} \omega_X(x, x').$$

The weight of a two-node network is illustrated in Figure 1. Observe that for a metric space, the weight of a subset is exactly its diameter. Next we can define the *Rips complex of a network* as before:

$$\mathfrak{R}_X^\delta = \{\sigma \in \text{Pow}(X) : \text{wgt}_X(\sigma) \leq \delta\}, \text{ for } (X, \omega_X) \in \mathcal{N}.$$

The Rips complex as defined above yields a valid simplicial complex on a network for each parameter  $\delta \in \mathbb{R}$ . Thus to any network  $(X, \omega_X)$ , we may associate the *Rips filtration*  $\{\mathfrak{R}_X^\delta \hookrightarrow \mathfrak{R}_X^{\delta'}\}_{\delta \leq \delta'}$ . For each  $k \in \mathbb{Z}_{\geq 0}$ , we denote the  $k$ -dimensional persistence diagram by  $D_k^{\mathfrak{R}}(X)$ . The Rips filtration of a network is stable to small perturbations of the input data; this is the content of the next proposition.

**Proposition 1.** *Let  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ . Then we have:*

$$d_B(D_k^{\mathfrak{R}}(X), D_k^{\mathfrak{R}}(Y)) \leq 2d_{\mathcal{N}}(X, Y).$$

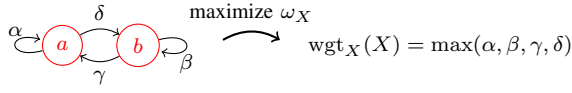


Fig. 1: Computing the weight of a network  $(X, \omega_X)$ .

The preceding proposition gives a way to compare two different networks by computing their persistence diagrams. Thus computing persistence diagrams becomes a well-defined method for comparing networks, in the same spirit as that of computing clustering coefficients, degree distributions, or betweenness centrality. The idea of computing persistence diagrams to compare finite metric spaces appeared in [12].

The Rips complex approach is undoubtedly valuable; see [9], [10], [22] for applications of the Rips complex to *symmetric* networks. However, it is open to several criticisms. First, the Rips complex of a network is blind to directed edge weights, because for any two-point subset as in Figure 1, the  $\text{wgt}$  function simply picks out the maximum edge weight, regardless of direction. Thus Rips is not an appropriate method for analyzing networks where directedness contains important structural information. Second, the Rips complex does not absorb information in dimensions higher than one: in order for any subset of points to span a simplex, the only requirement is that the pairwise edge weights in the subset are below a certain threshold. One might instead desire an  $n$ -simplex to form only if the  $n+1$  points together satisfy a common property. Finally, simplices in a Rips complex are not formed with respect to any “central authority.” This could be undesirable in, for example, a small-world network, where one would desire simplices to be formed with respect to particular “hub” nodes.

We propose using a more powerful structure, called the *Dowker filtration* (named after CH Dowker [23]), which addresses each of the concerns listed above.

## V. THE DOWKER FILTRATION OF A NETWORK

Given a network  $(X, \omega_X) \in \mathcal{N}$  and  $\delta \in \mathbb{R}$ , let  $R_{\delta, X} \subseteq X \times X$  be defined as follows:

$$R_{\delta, X} := \{(x, x') : \bar{\omega}_X(x, x') \leq \delta\}.$$

Observe that there exist  $\delta_0, \delta_1 \in \mathbb{R}$  such that  $R_{\delta, X} = \emptyset$  for any  $\delta \leq \delta_0$ , and  $R_{\delta, X} = X \times X$  for any  $\delta \geq \delta_1$ . Furthermore, given  $\delta \leq \delta'$ , we have  $R_{\delta, X} \subseteq R_{\delta', X}$ .

Using  $R_{\delta, X}$ , we build a simplicial complex  $\mathfrak{D}_{\delta, X}^{\text{si}}$  as follows:

$$\mathfrak{D}_{\delta, X}^{\text{si}} := \{\sigma \in \text{Pow}(X) : \text{there exists } p \in X \text{ such that } (x, p) \in R_{\delta, X} \text{ for each } x \in \sigma\}.$$

Thus an  $n$ -simplex  $\sigma = [x_1, \dots, x_n]$  belongs to  $\mathfrak{D}_{\delta, X}^{\text{si}}$  whenever there exists a distinguished point  $p \in X$  such that  $\bar{\omega}_X(x_i, p) \leq \delta$  for each  $x_i \in \sigma$ . 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 complex* associated to  $X$ , and refer to  $p$  as a *sink* for  $\sigma$  (where  $\sigma$  and  $p$  should be clear from context).

Since  $\{R_{\delta, X}\}_{\delta}$  is an increasing (i.e. nested) sequence of sets, it follows that  $\mathfrak{D}_{\delta, X}^{\text{si}}$  is an increasing sequence of simplicial complexes. In particular, given  $\delta \leq \delta'$ , 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  $X$* . For each  $k \in \mathbb{Z}_{\geq 0}$ , we denote the  $k$ -dimensional persistence diagram arising from this filtration by  $D_k^{\text{si}}(X)$ .

Note that we can define a dual construction as follows:

$$\mathfrak{D}_{\delta, X}^{\text{so}} := \{\sigma \in \text{Pow}(X) : \text{there exists } p \in X \text{ such that } (p, x) \in R_{\delta, X} \text{ for each } x \in \sigma\}.$$

Thus an  $n$ -simplex  $\sigma = [x_1, \dots, x_n]$  belongs to  $\mathfrak{D}_{\delta, X}^{\text{so}}$  whenever there exists  $p \in X$  such that  $\bar{\omega}_X(p, x_i) \leq \delta$  for each  $x_i \in \sigma$ . We call  $\mathfrak{D}_{\delta, X}^{\text{so}}$  the *Dowker  $\delta$ -source complex* associated to  $X$ , and refer to the distinguished point  $p$  as a *source* for  $\sigma$ . 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 associated  $k$ -dimensional persistence diagram by  $D_k^{\text{so}}(X)$ .

Two relevant observations that we can already make are as follows: (1) Dowker complexes capture higher-dimensional structure, because all the  $n+1$  points of an  $n$ -simplex satisfy a common property on the edge weights (relative to a source or sink node), and (2) Nodes with numerous low-weight edges pointing inward (or outward) are crucial in the formation of the complex, suggesting that it would appropriately capture the effect of hub nodes in a small-world network.

In the setting of finite metric spaces, the Dowker complex is related to constructions known as the Čech and witness complexes [21]. Stability results analogous to our Proposition 3 can be formulated in the metric space setting using well-known notions of distance between metric spaces [17]. Our contribution lies in clarifying the Dowker complex for networks, and providing stability results using a notion of network distance that has been applied in other constructions [18]–[20].

### A. Dowker’s Theorem and Equivalence of Diagrams

As an application of a result by Dowker [23], we get the following result: for any network  $(X, \omega_X) \in \mathcal{N}$ , any  $\delta \in \mathbb{R}$ , and any dimension  $k \in \mathbb{Z}_{\geq 0}$ , we have  $H_k(\mathfrak{D}_{\delta, X}^{\text{si}}) = H_k(\mathfrak{D}_{\delta, X}^{\text{so}})$ . In particular, by proving a certain extension of Dowker’s theorem and then applying the Persistence Equivalence Theorem [15], we are able to prove the following duality result:

**Theorem 2.** *For any  $k \in \mathbb{Z}_{\geq 0}$  and  $(X, \omega_X) \in \mathcal{N}$ , we have:*

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

Thus we may call either of the two diagrams above the  *$k$ -dimensional Dowker diagram of  $X$* , denoted  $D_k^{\mathfrak{D}}(X)$ .

As in the case of the Rips filtration, the Dowker filtration enjoys the following stability property:

**Proposition 3.** *Let  $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ . Then we have:*

$$d_{\text{B}}(D_k^{\mathfrak{D}}(X), D_k^{\mathfrak{D}}(Y)) \leq 2d_{\mathcal{N}}(X, Y).$$

As in the case of the Rips filtration, the preceding proposition gives a method of comparing two networks (of different

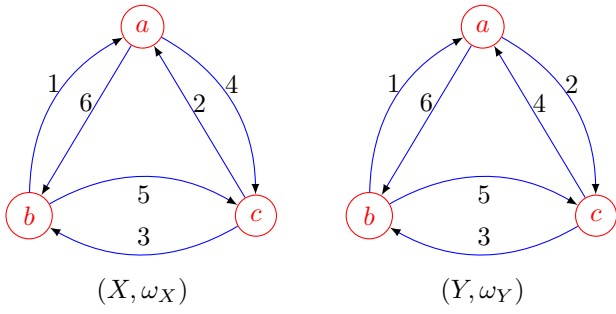


Fig. 2:  $Y$  is similar to  $X$ , but with one edge flipped.

sizes) by comparing their persistence diagrams. Thus the bottleneck distance, which is a polynomial-time computation [24], becomes a proxy for computing the network distance.

### B. Dowker Filtrations are Sensitive to Directionality

The directedness of a network  $(X, \omega_X)$  is implicitly captured in the construction of its source and sink Dowker complexes. Additionally, we provide an explicit example showing that Dowker filtrations are sensitive to directionality. Consider the networks  $(X, \omega_X)$  and  $(Y, \omega_Y)$  in Figure 2. Observe that  $(Y, \omega_Y)$  is obtained from  $(X, \omega_X)$  by swapping the edge weights between nodes  $a$  and  $c$ . We argue the following: (1) the Rips persistence diagrams of the two networks are the same, and (2) the Dowker persistence diagrams of the two networks are different.

To see (1), one can first verify that  $\mathfrak{R}_X^\delta = \mathfrak{R}_Y^\delta$  for each  $\delta \in \mathbb{R}$ , by noting that  $\text{wgt}_X(\{a, c\}) = \text{wgt}_Y(\{a, c\}) = 4$ . One can then show, using the Persistence Equivalence Theorem [15], that  $D_k^{\mathfrak{R}}(X) = D_k^{\mathfrak{R}}(Y)$  for each dimension  $k$ .

To show that the Dowker diagrams of  $X$  and  $Y$  are different, we simply compute the persistence barcodes using Javaplex, and present the results in Figure 3. Note that there are no persistent 1-dimensional intervals for  $X$ , whereas there is a single 1-dimensional interval, persisting for  $\delta \in [3, 4]$ , for  $Y$ .

## VI. APPLICATION: SIMULATED HIPPOCAMPAL NETWORKS

In the neuroscience literature, it has been shown that as an animal explores a given environment, specific “place cells” in the hippocampus show increased activity at specific spatial regions, called “place fields” [25]. 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 goes elsewhere. 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”, enough to detect the structure of the environment?

Approaches based on homology [13] and persistent homology [14] have shown positive results in this direction. In [14], 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. Then it was

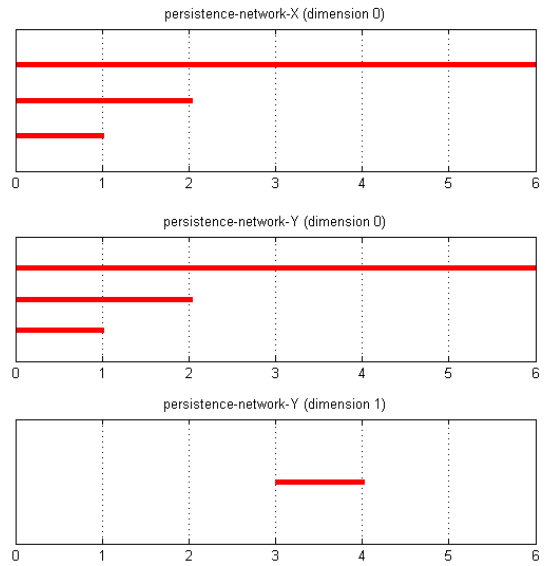


Fig. 3: Dowker persistence barcodes of networks  $(X, \omega_X)$  and  $(Y, \omega_Y)$  from Figure 2.

shown that the Betti-1 number of this simplicial complex would accurately represent the number of holes in the arena.

We repeated this experiment with the following change in methodology: viewing the trajectory as a Markov chain, we induced a directed network and computed the associated Dowker sink filtrations. The directedness corresponds to the natural intuition that an animal might have a preferred or familiar path for entering certain regions. After computing the Dowker sink filtrations associated to these directed networks, we compared the bottleneck distances between the resulting 1-dimensional persistence diagrams.

In our experiment, there were five environments. 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 environments were those obtained by removing the forbidden zones one at a time. In what follows, we refer to the environments of each type as *4-hole*, *3-hole*, *2-hole*, *1-hole*, and *0-hole environments*. For each environment, 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 environment (a disallowed move), then the probabilities were redistributed uniformly among the allowed moves. Each trajectory was tested to ensure that it covered the entire environment, excluding the forbidden zones.

For each of the five environments, 20 trials were conducted, producing a total of 100 trials. For each trial  $t_k$ , 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 environment for each  $t_k$ .

A spike on a place field was recorded whenever the trajectory would intersect it. For each trial  $t_k$ , the corresponding network  $(X, \omega_X)$  was constructed as follows:  $X$  consisted of  $n_k$  nodes representing place fields, and for each  $1 \leq i, j \leq n_k$ , the weight  $\omega_X(x_i, x_j)$  was given by  $\omega_X(x_i, x_j) = 1 - \frac{N_{i,j}(5)}{N_j}$ , where

$$N_{i,j}(5) = \# \text{ times cell } x_j \text{ spiked in a window of 5} \\ \text{time units after cell } x_i \text{ spiked,}$$

and  $N_j = \text{total \# times cell } x_j \text{ spiked at any time.}$

Next, we computed the Dowker persistence diagrams of each of the 100 networks. We then computed a  $100 \times 100$  matrix consisting of the bottleneck distances between all the 1-dimensional diagrams. The single linkage dendrogram generated from this bottleneck distance matrix is shown in Figure 4. The labels are in the format `env-<nh>-<nn>`, where `nh` is the number of holes in the environment, and `nn` is the number of place fields. Note that with some exceptions, networks corresponding to the same environment are clustered together. We conclude that the Dowker filtration succeeded in capturing the intrinsic differences between the five classes of networks arising from the five different environments, even when the networks had different sizes.

At the moment we are working on expanding the database of networks on which to test and evaluate the Dowker filtration.

#### REFERENCES

- [1] R. Ghrist, "Barcodes: the persistent topology of data," *Bulletin of the American Mathematical Society*, vol. 45, no. 1, pp. 61–75, 2008.
- [2] H. Edelsbrunner and J. Harer, "Persistent homology—a survey," *Contemporary mathematics*, vol. 453, pp. 257–282, 2008.
- [3] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society*, vol. 46, no. 2, pp. 255–308, 2009.
- [4] S. Emrani, T. Gentimis, and H. Krim, "Persistent homology of delay embeddings and its application to wheeze detection," *Signal Processing Letters, IEEE*, vol. 21, no. 4, pp. 459–463, 2014.
- [5] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, and D. S. Lee, "Computing the shape of brain networks using graph filtration and gromov-hausdorff metric," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*. Springer, 2011, pp. 302–309.
- [6] A. Khalid, B. S. Kim, M. K. Chung, J. C. Ye, and D. Jeon, "Tracing the evolution of multi-scale functional networks in a mouse model of depression using persistent brain network homology," *NeuroImage*, vol. 101, pp. 351–363, 2014.
- [7] D. Horak, S. Maletić, and M. Rajković, "Persistent homology of complex networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, no. 03, p. P03034, 2009.
- [8] C. Carstens and K. Horadam, "Persistent homology of collaboration networks," *Mathematical Problems in Engineering*, vol. 2013, 2013.
- [9] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, "Clique topology reveals intrinsic geometric structure in neural correlations," *Proceedings of the National Academy of Sciences*, vol. 112, no. 44, pp. 13455–13460, 2015.
- [10] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino, "Topological strata of weighted complex networks," *PLoS one*, vol. 8, no. 6, p. e66506, 2013.
- [11] R. Ghrist, *Elementary applied topology*. Createspace, 2014.
- [12] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, and S. Y. Oudot, "Gromov-hausdorff stable signatures for shapes using persistence," in *Computer Graphics Forum*, vol. 28, no. 5. Wiley Online Library, 2009, pp. 1393–1403.
- [13] C. Curto and V. Itskov, "Cell groups reveal structure of stimulus space," *PLoS Computational Biology*, vol. 4, no. 10, p. e1000205, 2008.
- [14] Y. Dabaghian, F. Mémoli, L. Frank, and G. Carlsson, "A topological paradigm for hippocampal spatial map formation using persistent homology," *PLoS Comput Biol*, vol. 8, no. 8, p. e1002581, 2012.

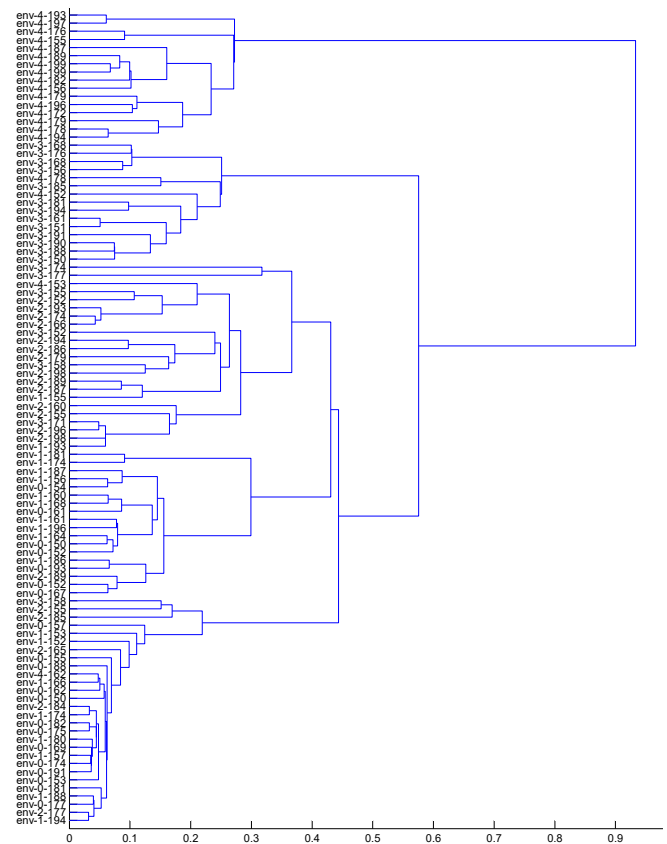


Fig. 4: Single linkage dendrogram corresponding to database of hippocampal networks. The y-axis ticks have the format `env-<nh>-<nn>`, where `nh` is the number of holes in the arena and `nn` is the number of place fields in the arena.

- [15] H. Edelsbrunner and J. Harer, *Computational topology: an introduction*. American Mathematical Soc., 2010.
- [16] A. Tausz, M. Vejdemo-Johansson, and H. Adams, "Javaplex: A research software package for persistent (co) homology."
- [17] F. Chazal, V. De Silva, and S. Oudot, "Persistence stability for geometric complexes," *Geometriae Dedicata*, vol. 173, no. 1, pp. 193–214, 2014.
- [18] G. E. Carlsson, F. Mémoli, A. Ribeiro, and S. Segarra, "Hierarchical quasi-clustering methods for asymmetric networks," in *Proceedings of the 31th International Conference on Machine Learning, ICML 2014*.
- [19] S. Chowdhury and F. Mémoli, "Metric structures on networks and applications," in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2015, pp. 1470–1472.
- [20] —, "Distances between directed networks and applications," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, to appear.
- [21] V. De Silva and G. Carlsson, "Topological estimation using witness complexes," *Proc. Sympos. Point-Based Graphics*, pp. 157–166, 2004.
- [22] W. Huang and A. Ribeiro, "Comparing high order networks by persistence homology," in *Proc. Global Conf. Signal Info. Process., vol.(submitted), Orlando FA, 2015*.
- [23] C. H. Dowker, "Homology groups of relations," *Annals of mathematics*, pp. 84–95, 1952.
- [24] A. Efrat, A. Itai, and M. J. Katz, "Geometry helps in bottleneck matching and related problems," *Algorithmica*, vol. 31, no. 1, pp. 1–28, 2001.
- [25] J. O’Keefe and J. Dostrovsky, "The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat," *Brain research*, vol. 34, no. 1, pp. 171–175, 1971.