

# Persistent Homology Lower Bounds on High Order Network Distances

Weiyu Huang and Alejandro Ribeiro

**Abstract**—This paper presents methods to compare high order networks using persistent homology. High order networks are weighted complete hypergraphs collecting relationship functions between elements of tuples. They can be considered as generalizations of conventional networks where only relationship functions between pairs are defined. Valid metric distances between high order networks have been defined but they are inherently difficult to compute when the number of nodes is large. We relate high order networks to the filtrations of simplicial complexes and show that the difference between networks can be lower bounded by the difference between the homological features of their respective filtrations. Practical implications are explored by comparing the coauthorship networks of engineering and mathematics academic journals. The lower bounds succeed in discriminating engineering communities from mathematics communities and in differentiating engineering communities with different research interests.

## I. INTRODUCTION

We consider high order networks that describe relationships between elements of tuples and address the problem of computing an approximate value for their distance. While networks are more often defined as structures describing relationships between pairs of nodes [2], [3], there are many applications in which higher order relationships between triplets, quadruplets, and generic  $n$ -tuples are also important. For example, in coauthorship networks where relationships detail the number of joint publications by groups of scholars, apart from collaboration between pairs of authors, there is extra information to be gleaned from collaborations between triplets of authors, or even single author publications. The importance of expressing tuple relationships between groups of nodes in addition to pairs has been utilized in multiple domains including sensor networks [4], [5], cognitive learning [6], wireless networks [7], image ranking [8], object recognition [9], and social networks [10].

The problem of defining distances between networks is important even in the case of pairwise networks. However, while distances are not difficult to define, they are certainly difficult to compute. This has motivated the use of feature comparisons in which the difference between specific properties of the network is used as a tractable alternative. Examples of feature comparisons are clustering coefficients [11], neighborhood topologies [12], betweenness [13], motifs [14], wavelets [15], and graphlet degree distributions [16], [17]. While tractable, the use of features is application dependent, utilizes only a small portion of the information conveyed by the networks, and may yield conflicting comparative judgements – like two networks being close to a third but far from each other – because the triangle inequality is not necessarily valid. A proper distance between pairwise and high

order networks overcomes these drawbacks. This is the motivation for the definition of metrics that generalize the Gromov-Hausdorff distance between metric spaces [18], [19] to pairwise [20] and high order networks [21].

The metric distances between high order networks defined in [21] have been applied to compare networks with small number of nodes and have succeeded in identifying collaboration patterns of coauthorship networks. However, because they have to consider all possible node correspondences (Definition 2), network distances are difficult to compute when the number of nodes in the networks is large. The goal of this paper is to find lower bounds on the network distances that are computable in networks with large numbers of nodes. We achieve this using persistent homology. It has to be acknowledged that the computation of distance lower bounds suffers from some of the same problems associated with feature comparison. Still, distance lower bounds do have important properties such as the facts that a large lower bound entails a large distance and they can be used to estimate distance intervals because upper bounds are easy to determine using specific correspondences (see Definition 3).

The paper begins with a brief review of high order networks and high order network distances [21] (Section II). Concepts from persistence topology [22], [23] are then discussed and connections between high order networks and simplicial complexes as well as between tuple relationships and homological features are established (Section III). The difference between networks can therefore be estimated by the difference between the respective sets of homological features. We prove that the two families of network distances can be lower bounded using homological features (Theorems 1 and 2 in Section IV). We illustrate that all the information conveyed by non-trivial relationship functions of a network can be found in homological features (Proposition 4) and that the lower bounds established in Theorems 1 and 2 are tight. Presented are also examples where the differences between the homological features equate the metrics between networks. Since persistent homology can be computed efficiently for large networks [24], [25], we use these lower bounds to compare the coauthorship networks constructed from the publications of a number of academic journals from engineering and mathematics communities. The lower bounds succeed in discriminating engineering communities from mathematics communities and in differentiating the engineering communities of signal processing, automatic control, and wireless communications (Section V). The paper closes with concluding remarks (Section VI).

## II. HIGH ORDER NETWORKS

A network  $N_X^K$  of order  $K$  over the node space  $X$  is defined as a collection of  $K + 1$  relationship functions  $\{r_X^k : X^{k+1} \rightarrow \mathbb{R}_+\}_{k=0}^K$  from the space  $X^{k+1}$  of  $(k+1)$ -tuples to the nonnegative reals.  $N_X^K$  can be considered as a weighted complete hypergraph

Work in this paper is supported by NSF CCF-1217963 and AFOSR MURI FA9550-10-1-0567. The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, 200 South 33rd Street, Philadelphia, PA 19104. Email: {whuang, aribeiro}@seas.upenn.edu. Part of the results in this paper appeared in [1].

[26] whose weights for all hyperedges of elements of  $(k+1)$  tuples with  $0 \leq k \leq K$  are defined. The rank  $s(x_{0:k})$  of a given tuple  $x_{0:k} := (x_0, x_1, \dots, x_k) \in X^{k+1}$  is the number of unique elements in the tuple. We say that a tuple  $x_{0:k}$  is full rank if its rank equals its cardinality  $s(x_{0:k}) = |x_{0:k}| = k+1$ . For tuples that are not full rank, the relationship function between a tuple  $x_{0:k}$  is identical to the relationship functions of subtuples of  $x_{0:k}$  that have rank  $s(x_{0:k})$  since they contain the same information. This remark along with a symmetry property makes up the formal definition of high order networks.

**Definition 1**  $N_X^K = (X, r_X^0, r_X^1, \dots, r_X^K)$  is a  $K$ -order network if the following two properties hold:

**Symmetry.** For any  $0 \leq k \leq K$  and any point collections  $x_{0:k}$ , we have that  $r_X^k(x_{[0:k]}) = r_X^k(x_{0:k})$ , where  $x_{[0:k]} = ([x_0], [x_1], \dots, [x_k])$  is a reordering of  $x_{0:k}$ .

**Identity.** For any  $0 \leq k \leq K$  and tuple  $x_{0:k}$ , any of the subtuples  $x_{l_0:l_{\bar{k}}}$  with  $s(x_{0:k}) = s(x_{l_0:l_{\bar{k}}})$  satisfies  $r_X^k(x_{0:k}) = r_X^k(x_{l_0:l_{\bar{k}}})$ .

The set of all high order networks of order  $K$  is denoted as  $\mathcal{N}^K$ .

For point collections  $x_{0:k}$ , the values  $r_X^k(x_{0:k})$  of their  $k$ -order relationship functions are intended to represent a measure of similarity or dissimilarity for members of the group. Examples for the identity property include  $r_X^2(x, x) = r_X^1(x)$  and  $r_X^3(x', x, x') = r_X^2(x, x')$ . We assume that relationship values are normalized so that  $0 \leq r_X^k(x_{0:k}) \leq 1$  for all  $k$  and  $x_{0:k}$ . We say that two networks  $N_X^K$  and  $N_Y^K$  are  $k$ -isomorphic if there exists a bijection  $\pi : X \rightarrow Y$  such that for all  $x_{0:k} \in X^{k+1}$  we have

$$r_Y^k(\pi(x_{0:k})) = r_X^k(x_{0:k}), \quad (1)$$

where we use the shorthand notation  $r_Y^k(\pi(x_{0:k})) := r_Y^k(\pi(x_0), \pi(x_1), \dots, \pi(x_k))$ . Since the map  $\pi$  is bijective, (1) can only be satisfied when  $X$  is a permutation of  $Y$ . When networks  $N_X^K$  and  $N_Y^K$  are  $k$ -isomorphic we write  $N_X^K \cong_k N_Y^K$ . The space of  $K$ -order networks where  $k$ -isomorphic networks are represented by the same element is termed the set of  $K$ -networks modulo  $k$ -isomorphism and denoted by  $\mathcal{N}^K \text{ mod } \cong_k$ . For each  $0 \leq k \leq K$ , the space  $\mathcal{N}^K \text{ mod } \cong_k$  can be endowed with a pseudometric [21]. The definition of this distance requires introducing the notion of correspondence [27, Def. 7.3.17].

**Definition 2** A correspondence between two sets  $X$  and  $Y$  is a subset  $C \subseteq X \times Y$  such that  $\forall x \in X$ , there exists  $y \in Y$  such that  $(x, y) \in C$  and  $\forall y \in Y$  there exists  $x \in X$  such that  $(x, y) \in C$ . The set of all correspondences between  $X$  and  $Y$  is denoted as  $\mathcal{C}(X, Y)$ .

A correspondence in the sense of Definition 2 is a map between node sets  $X$  and  $Y$  so that every element of each set has at least one correspondent in the other set. Most importantly, this allows definition of correspondences between networks with different numbers of elements. We can now define the distance between two networks by selecting the correspondence that makes them most similar as we formally define next.

**Definition 3** Given networks  $N_X^K$  and  $N_Y^K$ , a correspondence  $C$  between the node spaces  $X$  and  $Y$ , and an integer  $0 \leq k \leq K$  define the  $k$ -order network difference with respect to  $C$  as

$$\Gamma_{X,Y}^k(C) := \max_{(x_{0:k}, y_{0:k}) \in C} |r_X^k(x_{0:k}) - r_Y^k(y_{0:k})|, \quad (2)$$

where the notation  $(x_{0:k}, y_{0:k})$  stands for  $(x_0, y_0), (x_1, y_1), \dots, (x_k, y_k)$ . The  $k$ -order network distance between networks  $N_X^K$  and  $N_Y^K$  is then defined as

$$d_{\mathcal{N}}^k(N_X^K, N_Y^K) := \min_{C \in \mathcal{C}(X,Y)} \{\Gamma_{X,Y}^k(C)\}. \quad (3)$$

For a given correspondence  $C \in \mathcal{C}(X, Y)$  the network difference  $\Gamma_{X,Y}^k(C)$  selects the maximum distance difference  $|r_X^k(x_{0:k}) - r_Y^k(y_{0:k})|$  among all pairs of correspondents – we compare  $r_X^k(x_{0:k})$  with  $r_Y^k(y_{0:k})$  when all the points  $x_l \in x_{0:k}$  and  $y_l \in y_{0:k}$  are correspondents. The distance in (3) is defined by selecting the correspondence that minimizes these maximal differences. The distance in Definition 3 is a pseudometric in  $\mathcal{N}^K \text{ mod } \cong_k$  [21]. For future reference, the notions of metric and pseudometric are formally stated next.

**Definition 4** Given a space  $\mathcal{S}$  and an isomorphism  $\cong$ , a function  $d : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$  is a metric in  $\mathcal{S} \text{ mod } \cong$  if for any  $a, b, c \in \mathcal{S}$  the function  $d$  satisfies:

- (i) **Nonnegativity.**  $d(a, b) \geq 0$ .
- (ii) **Symmetry.**  $d(a, b) = d(b, a)$ .
- (iii) **Identity.**  $d(a, b) = 0$  if and only if  $a \cong b$ .
- (iv) **Triangle inequality.**  $d(a, b) \leq d(a, c) + d(c, b)$ .

The function is a pseudometric in  $\mathcal{S} \text{ mod } \cong$  if for any  $a, b, c \in \mathcal{S}$  the function  $d$  satisfies (i), (ii), (iv), and

- (iii') **Relaxed identity.**  $d(a, b) = 0$  if  $a \cong b$ .

A metric  $d$  in  $\mathcal{S} \text{ mod } \cong$  gives a proper notion of distance. Since zero distances imply elements being isomorphic, the distance between elements reflects how far they are from being isomorphic. Pseudometrics are relaxed since elements not isomorphic may still have zero distances measured by the pseudometrics. Observe that since correspondences may be between networks with different number of elements,  $d_{\mathcal{N}}^k(N_X^K, N_Y^K)$  is defined when the node cardinalities  $|X|$  and  $|Y|$  are different.

We say that two  $K$ -order networks  $N_X^K$  and  $N_Y^K$  are isomorphic if there exists a bijection  $\pi : X \rightarrow Y$  such that (1) holds for all  $0 \leq k \leq K$  and  $x_{0:k} \in X^{k+1}$ . When networks  $N_X^K$  and  $N_Y^K$  are isomorphic we write  $N_X^K \cong N_Y^K$ . The space of  $K$ -order networks modulo isomorphism is denoted as  $\mathcal{N}^K \text{ mod } \cong$ . A family of pseudometrics measuring the difference over all order functions can be endowed to  $\mathcal{N}^K \text{ mod } \cong$  as we formally state next.

**Definition 5** Given networks  $N_X^K$  and  $N_Y^K$ , a correspondence  $C$  between the node spaces  $X$  and  $Y$ , and some vector  $p$ -norm  $\|\cdot\|_p$ , define the network difference with respect to  $C$  as

$$\|\Gamma_{X,Y}^K(C)\|_p := \left\| \left( \Gamma_{X,Y}^0(C), \Gamma_{X,Y}^1(C), \dots, \Gamma_{X,Y}^K(C) \right)^T \right\|_p, \quad (4)$$

where for each integer  $0 \leq k \leq K$ ,  $\Gamma_{X,Y}^k(C)$  is the  $k$ -order network difference with respect to  $C$  defined in (2). The  $p$ -norm network distance between  $N_X^K$  and  $N_Y^K$  is then defined as

$$d_{\mathcal{N},p}(N_X^K, N_Y^K) := \min_{C \in \mathcal{C}(X,Y)} \left\{ \|\Gamma_{X,Y}^K(C)\|_p \right\}. \quad (5)$$

In Definition 5 we compare relationship functions of all orders no higher than  $K$ . The norm  $\|\Gamma_{X,Y}^K(C)\|_p$  is assigned as the difference between  $N_X^K$  and  $N_Y^K$  measured by the correspondence  $C$ . The distance  $d_{\mathcal{N},p}(N_X^K, N_Y^K)$  is then the minimum of these

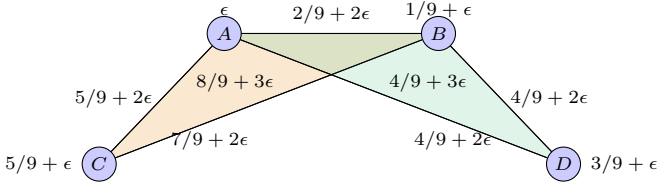


Fig. 1. Temporal dynamics for the formation of a research community. The  $k$ -order relationship in this 2-order dissimilarity network [cf. Definition 6] incorporates the dissimilarity term – the normalized time instant when members of  $(k+1)$ -tuple write their first joint paper – and the multiplication of  $\epsilon$  with the rank of the tuple. E.g.,  $A$  writes her first paper at time 0, and coauthors with  $B, D, C$  at times  $2/9, 4/9, 5/9$ , respectively. She also writes jointly with  $B$  and  $D$  at time  $4/9$ .

differences achieved by some correspondence. As was the case for Definition 3,  $d_{\mathcal{N},p}(N_X^K, N_Y^K)$  is defined even if the numbers of nodes in  $X$  and  $Y$  are different.

Observe that the  $k$ -order function  $r_X^k$  of a network  $N_X^K$  does not impose constraints on the  $l$ -order function  $r_X^l$  of the same network except the identity property. In practical situations, however, it is common to observe that adding nodes to a tuple results in either increasing or decreasing relationships between elements of the extended tuple. This motivates the consideration of dissimilarity networks and proximity networks that we undertake next.

#### A. Dissimilarity and Proximity Networks

In dissimilarity networks,  $r_X^k(x_{0:k})$  encodes a level of dissimilarity between elements. In this scenario it is reasonable that adding elements to a tuple makes the group more dissimilar and results in a higher value in the relationship function. This restriction makes up the formal definition that we introduce next.

**Definition 6** We say that the  $K$ -order network  $D_X^K$  is a dissimilarity network if for any  $0 \leq k \leq K$  and tuples  $x_{0:k} \in X^{k+1}$ , its relationship function is the summation of a dissimilarity term and the multiplication of its rank with a small constant  $\epsilon$ ,

$$r_X^k(x_{0:k}) = d_X^k(x_{0:k}) + \epsilon s(x_{0:k}). \quad (6)$$

The dissimilarity terms satisfy the order increasing property that for any  $1 \leq k \leq K$  and  $x_{0:k}$ ,

$$d_X^k(x_{0:k}) \geq d_X^{k-1}(x_{0:k-1}), \quad (7)$$

and the constant  $\epsilon$  is a strictly positive value that satisfies  $0 < \epsilon \leq 1 - (K+1)^{-1} \max_{\tilde{x}_{0:K} \in X^{K+1}} d_X^K(\tilde{x}_{0:K})$ . Denote the set of all dissimilarity networks of order  $K$  as  $\mathcal{D}^K$ .

In dissimilarity networks, adding a new element to a tuple makes the set *strictly more dissimilar* than it was. Figure 1 illustrates a dissimilarity network where dissimilarities mark the normalized time instants when members become coauthors. The rank correction term  $\epsilon s(x_{0:k})$  is a technical modification to distinguish between full rank (proper)  $k$ -tuples and rank deficient (degenerate) tuples. In practice it is set to a sufficiently small positive quantity compared to dissimilarities.

In proximity networks, the function  $r_X^k(x_{0:k})$  encodes a level of similarity or proximity between elements. Under this circumstance it is reasonable to assume that adding elements to a tuple makes the group less similar, resulting in a lower value in the relationship function. We formally introduce a definition next.

**Definition 7** We say that the  $K$ -order network  $P_X^K$  is a proximity network if for any  $0 \leq k \leq K$  and tuples  $x_{0:k} \in X^{k+1}$ , its

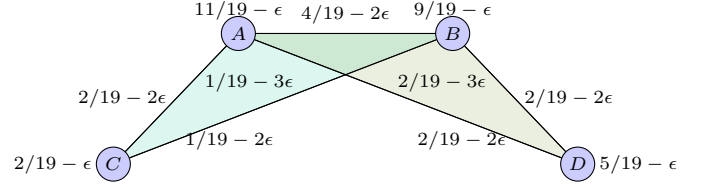


Fig. 2. Collaborations between authors in a research community. The  $k$ -order relationship in this 2-order network [cf. Definition 7] incorporates the proximity term – the number of publications between members of  $(k+1)$ -tuples normalized by the total number of papers – and the multiplication of  $-\epsilon$  with the rank of the tuple. E.g.,  $A$  writes 11 papers in total, and coauthors 4, 2, and 2 papers respectively with  $B, D$ , and  $C$ . She also writes 1 paper jointly with  $B$  and  $C$ .

relationship function is the summation of a proximity term and the multiplication of its rank with  $-\epsilon$ ,

$$r_X^k(x_{0:k}) = p_X^k(x_{0:k}) - \epsilon s(x_{0:k}), \quad (8)$$

The proximity terms satisfy the order increasing property that for any  $1 \leq k \leq K$  and  $x_{0:k}$ ,

$$p_X^k(x_{0:k}) \leq p_X^{k-1}(x_{0:k-1}), \quad (9)$$

and the constant  $\epsilon$  is a strictly positive value that satisfies  $0 < \epsilon \leq (K+1)^{-1} \min_{\tilde{x}_{0:K} \in X^{K+1}} p_X^K(\tilde{x}_{0:K})$ . Denote the set of all proximity networks of order  $K$  as  $\mathcal{P}^K$ .

In proximity networks, adding a new element to a tuple makes the set *strictly less similar* than it was. Figure 2 exemplifies a proximity network where proximities describe the normalized number of publications between members in a research community.

When the input networks to Definitions 3 and 5 are dissimilarity or proximity networks, we refer to the respective distances as the dissimilarity or proximity network distances. We state this formally in the following definition for future reference.

**Definition 8** Given dissimilarity networks  $D_X^K, D_Y^K \in \mathcal{D}^K$  we say that the  $k$ -order distance  $d_{\mathcal{D}}^k(D_X^K, D_Y^K) = d_{\mathcal{N}}^k(D_X^K, D_Y^K)$  of Definition 3 is the  $k$ -order dissimilarity network distance between  $D_X^K$  and  $D_Y^K$  and that  $p$ -norm distance  $d_{\mathcal{D},p}(D_X^K, D_Y^K) = d_{\mathcal{N},p}(D_X^K, D_Y^K)$  of Definition 5 is the  $p$ -norm dissimilarity network distance. The  $k$ -order proximity network distance  $d_{\mathcal{P}}^k(P_X^K, P_Y^K)$  and the  $p$ -norm proximity network distance  $d_{\mathcal{P},p}(P_X^K, P_Y^K)$  are defined similarly.

The restrictions to dissimilarities or proximities make the distances in Definition 8 well-defined *metrics* in the respective space [21]. Proximity and dissimilarity networks are related entities. We formalize this relationship through the introduction of dual networks in the following definition.

**Definition 9** Given a node space  $X$ , the  $K$ -order proximity and dissimilarity networks  $P_X^K = (X, \hat{p}_X^0, \hat{p}_X^1, \dots, \hat{p}_X^K)$  and  $D_X^K = (X, \hat{d}_X^0, \hat{d}_X^1, \dots, \hat{d}_X^K)$  are said duals if and only if  $\hat{p}_X^k(x_{0:k}) = 1 - \hat{d}_X^k(x_{0:k})$  for all orders  $0 \leq k \leq K$  and tuples  $x_{0:k}$ .

Given two networks, the distance between them is equal to the distance between their respective duals. We formally state this in the following proposition.

**Proposition 1** Consider proximity networks  $P_X^K$  and  $P_Y^K$  and their corresponding dual dissimilarity networks  $D_X^K$  and  $D_Y^K$ . The  $k$ -order proximity distances  $d_{\mathcal{P}}^k(P_X^K, P_Y^K)$  and  $k$ -order dissimilarity distances  $d_{\mathcal{D}}^k(D_X^K, D_Y^K)$  coincide for all  $0 \leq k \leq K$ .

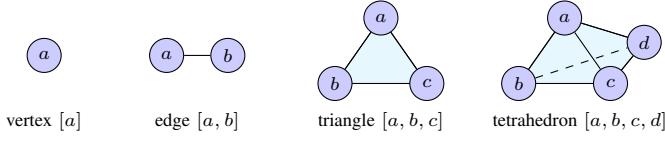


Fig. 3.  $k$ -simplices in  $\mathbb{R}^{k+1}$  for  $0 \leq k \leq 3$ .

Likewise, the  $p$ -norm proximity distance  $d_{\mathcal{P},p}(P_X^K, P_Y^K)$  and  $p$ -norm dissimilarity distance  $d_{\mathcal{D},p}(D_X^K, D_Y^K)$  coincide.

The metrics defined in Definition 8 provide us well-founded methods to compare high order networks. However, the combinatorial nature in searching for the optimal correspondence in (3) and (5) makes it impossible to find the exact solution when the number of nodes in the networks is large. This motivates the development of reasonable and computationally tractable lower bounds. These bounds will be obtained by relating dissimilarity networks to filtrations in computational homology [22], [23]. The mapping between high order networks and filtrations is discussed in the following section.

### III. RELATING DISSIMILARITY NETWORKS TO FILTRATIONS

In topology, given  $k + 1$  unique points  $x_{0:k}$ , it is common to consider them living in some  $\mathbb{R}^{k+1}$  space where the coordinates of point  $x_i$  are all zero except unity on the  $i$ -th axis. The  $k$ -simplex generated by the set of non-repeating points  $x_{0:k}$ ,  $\phi = [x_{0:k}]$ , is defined as the convex hull of the set of points,  $\text{conv}\{x_{0:k}\}$ . See Figure 3 for examples of  $k$ -simplex with  $0 \leq k \leq 3$ .

For all  $k \geq 1$ , removing a point  $x_l$  from the set  $x_{0:k}$  yields a set with  $k$  points that we denote as  $x_{0:\widehat{l}:k} := x_{0:k} \setminus x_l$ . Each of the  $k + 1$  convex hulls  $[x_{0:\widehat{l}:k}] = \text{conv}\{x_{0:k} \setminus x_l\}$  formed by removing the point  $x_l$  from the original set is a  $(k - 1)$ -simplex. A face of  $\phi$  is the convex hull of a non-empty subset of  $x_{0:k}$  and it is a proper face if the subset is not the entire set. For example, the set of proper faces for the 1-simplex  $[a, b]$  and the 2-simplex  $[a, b, c]$  in Figure 3 are  $\{[a], [b]\}$  and  $\{[a, b], [a, c], [b, c], [a], [b], [c]\}$ , respectively. Given the simplex  $\phi = [x_{0:k}]$ , the boundary  $\partial_k \phi$  of the simplex is the collection of all  $(k - 1)$ -dimensional faces considering orientations. Orientations are generalizations of directed edges in conventional pairwise graphs. For 1-simplices,  $[a, b]$  denotes the edges from  $a$  to  $b$  and  $[a, b] = -[b, a]$ . For the 2-simplex  $[a, b, c]$  in Figure 3, the orientation is either clockwise or counter-clockwise and  $[a, b, c] = [b, c, a] = [c, a, b] = -[a, c, b] = -[b, a, c] = -[c, b, a]$ . The boundary of a simplex  $\phi$  is written as

$$\partial_k \phi = \sum_{l=0}^k (-1)^l [x_{0:\widehat{l}:k}], \quad (10)$$

The  $(-1)^l$  ensures faces appear alternatively in the summation and therefore yields a well-defined orientation. Observe that since a 0-simplex  $\phi$  has no faces,  $\partial_0 \phi = 0$ . For the  $k$ -simplices in Figure 3,  $\partial_0[a] = 0$ ,  $\partial_1[a, b] = [b] - [a]$ ,  $\partial_2[a, b, c] = [b, c] - [a, c] + [a, b]$  and  $\partial_3[a, b, c, d] = [b, c, d] - [a, c, d] + [a, b, d] - [a, b, c]$ . A simplicial complex  $L$  is a finite collection of simplices such that every face of a simplex of  $L$  is also in  $L$  and the intersection of any two simplices is either empty or a shared face. See Figure 4 (a) for two connected bow ties as an example of a simplicial complex of dimension 2.

Simplicial complexes are high dimensional generalizations of conventional unweighted undirected graphs. The 1-skeleton of any

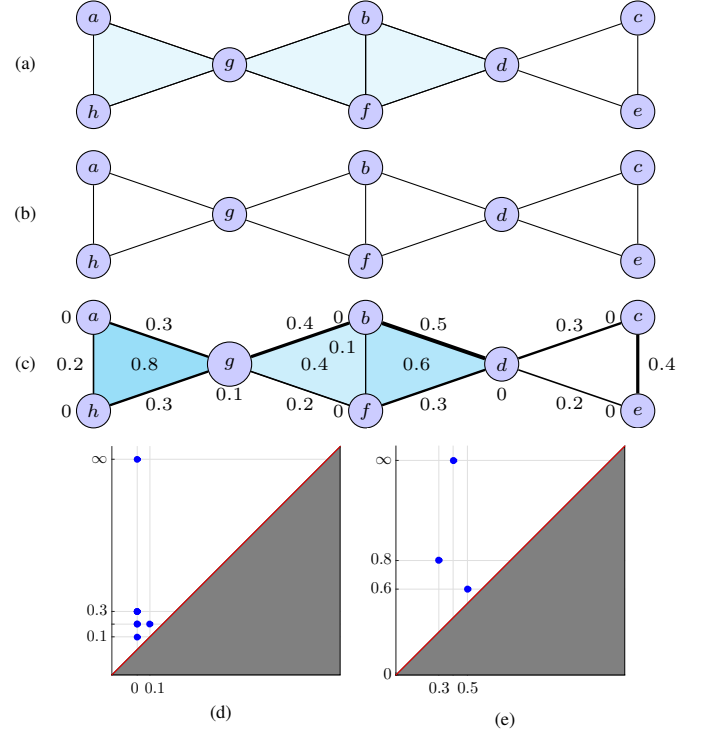


Fig. 4. (a): Two connected bow ties as an example of a simplicial complex  $L$  which consists of 8 0-simplices, 11 1-simplices, and 3 2-simplices. The intersection of the two 2-simplices  $[b, f, g]$  and  $[b, d, f]$  is a 1-simplex  $[b, f]$ , a shared face of the two 2-simplices. The intersection of the 2-simplices  $[a, g, h]$  and  $[b, f, g]$  is the 0-simplex  $[g]$ . (b): The 1-skeleton of  $L$ . Observe that it is a well-defined undirected unweighted graph. (c): A weighted high order network can be represented equivalently as a simplicial complex with weights. The weight of a simplex is the time instant the simplex appears in the simplicial complex. In the original network,  $r_X^1(a, g) = 0.3$ ; in the simplicial complex with weights, the 1-simplex  $[a, g]$  appears at time 0.3. (d): the zeroth-dimensional and (e): the first-dimensional persistence diagrams of the filtration induced by (c).

simplicial complex  $L$ , defined as the union of simplices in  $L$  with dimensions  $k \leq 1$ , is a well-defined simple graph (See Figure 4 (b)). A  $k$ -chain is a formal sum of  $k$ -simplices of  $L$ , denoted by  $\Phi = \sum_i \beta_i \phi_i$ , where each  $\phi_i$  is a  $k$ -simplex and each  $\beta_i$  is a coefficient. The  $k$ -chains together with the addition operation form the group of  $k$ -chains, denoted as  $\mathbf{C}_k(L)$ , or simply  $\mathbf{C}_k$ . For a  $k$ -chain with  $\Phi = \sum_i \beta_i \phi_i$ , its boundary is the sum of the boundaries of its simplices,  $\partial_k \Phi = \sum_i \beta_i (\partial_k \phi_i)$ . Hence,  $\partial_k$  maps a  $k$ -chain to a  $(k - 1)$ -chain,  $\partial_k : \mathbf{C}_k \rightarrow \mathbf{C}_{k-1}$ . The sequence of chain groups connected by boundary maps can be represented as

$$\cdots \xrightarrow{\partial_{k+2}} \mathbf{C}_{k-order+1} \xrightarrow{\partial_{k+1}} \mathbf{C}_k \xrightarrow{\partial_k} \mathbf{C}_{k-1} \xrightarrow{\partial_{k-1}} \cdots \quad (11)$$

For the connected bow ties in Figure 4 (a),  $\mathbf{C}_0 = \beta_1[a] + \beta_2[b] + \cdots + \beta_8[h]$ ,  $\mathbf{C}_2 = \beta'_1[a, g, h] + \beta'_2[b, g, f] + \beta'_3[b, d, f]$ . A  $k$ -cycle is a  $k$ -chain with empty boundary,  $\partial_k \Phi = 0$ . In the example,  $[a]$  is a 0-cycle and  $[a, g] + [g, h] - [a, h]$  is a 2-cycle.  $\mathbf{Z}_k$  denotes the group of  $k$ -cycles and is the kernel of the  $k$ -th boundary map,  $\mathbf{Z}_k = \ker \partial_k$ . Observe that any 0-chain is a 0-cycle, therefore  $\mathbf{Z}_0 = \mathbf{C}_0$ . A  $k$ -boundary is a  $k$ -chain that is the boundary of a  $(k + 1)$ -chain,  $\Phi = \partial_{k+1} \Psi$  for some  $\Psi \in \mathbf{C}_{k+1}$ . In the example,  $[g] - [h]$  is a 0-boundary since  $[g] - [h] = \partial_1[h, g]$  and  $[h, g]$  is a 1-chain. Similarly,  $[a, g] + [g, h] - [a, h]$  is a 2-boundary. Every boundary is also a cycle.  $\mathbf{B}_k$  denotes the group of  $k$ -boundaries and is the image of the  $(k + 1)$ -th boundary map,  $\mathbf{B}_k = \text{im } \partial_{k+1}$ . The  $k$ -th homology group is the  $k$ -th cycle group modulo the  $k$ -th boundary group,  $\mathbf{H}_k = \mathbf{Z}_k / \mathbf{B}_k$ . The homology groups considered in this paper are of the form  $\mathbf{H}_k \cong \sum_i \gamma_i \Phi_i$  where each  $\gamma_i \in \mathbb{R}$  denotes

a degree of freedom and  $\Phi_i = \sum_j \beta_j \phi_j$  with  $\beta_j \in \{-1, 1\}$  is a linear combination of simplices. We say that each chain  $\Phi_i$  represents a  $k$ -th dimensional homological feature.

We now connect computational topology with dissimilarity networks. Simplicial complexes can be considered as structures of high order networks, detailing the number and labels of vertices, edges, and higher dimensional counterparts. To incorporate relationship functions, we assign each simplex in the simplicial complex  $L$  a value between 0 and 1 denoting the time when this simplex appears. For any  $\alpha \in \mathbb{R}$ , we then define  $L_\alpha \subseteq L$  to be the collection of simplices appearing before or on time  $\alpha$ . If all faces of each simplex and intersections of any simplices in  $L_\alpha$  also appear before or on time  $\alpha$ ,  $L_\alpha$  is a well-defined simplicial complex and the nested sequence of  $\emptyset = L_{\alpha_0} \subseteq \dots \subseteq L_{\alpha_m} = L$  is defined as the filtration  $\mathcal{L}$  where  $0 = \alpha_0 < \dots < \alpha_m = 1$  is an ordered sequence of real numbers. From Definition 6, a well-defined filtration  $\mathcal{L}(D_X^K)$  can be induced from a given dissimilarity network  $D_X^K$  as we formally state next.

**Definition 10** Given a dissimilarity network  $D_X^K$ , its induced simplicial complex  $L(D_X^K)$  is defined on the same node space  $X$ . For full rank tuple  $x_{0:k}$ , the simplex  $[x_{0:k}]$  is in  $L(D_X^K)$  if the dissimilarity term  $d_X^k(x_{0:k})$  in the decomposition of  $r_X^k(x_{0:k})$  [cf. (6)] is not the maximum dissimilarity  $\max_{\tilde{x}_{0:k} \in X^{k+1}} d_X^K(\tilde{x}_{0:k})$ . The induced filtration  $\mathcal{L}(D_X^K)$  is the filtration where each simplex  $[x_{0:k}]$  in  $L(D_X^K)$  appears at  $r_X^k(x_{0:k})$ .

The reason that the appearance of a simplex in the simplicial complex depends on its dissimilarity function is because maximum dissimilarities between elements are expressed differently in dissimilarity networks and filtrations. In dissimilarity networks, the relationship  $r_X^k(x_{0:k})$  between elements of any tuple  $x_{0:k}$  with  $0 \leq k \leq K$  is defined. Elements are the most dissimilar when their dissimilarity term  $d_X^k(x_{0:k})$  is the maximum dissimilarity  $\max_{\tilde{x}_{0:k} \in X^{k+1}} d_X^K(\tilde{x}_{0:k})$  in the network. For filtrations, a simplex  $[x_{0:k}]$  whose elements are the most dissimilar would never appear. Equivalently speaking, the simplex  $[x_{0:k}]$  does not exist in the simplicial complex. In Definition 10, the induced filtration  $\mathcal{L}(D_X^K)$  ignores simplices whose elements are the most dissimilar. When a simplex  $[x_{0:k}]$  does appear in the simplicial complex, its appearance time is the corresponding relationship function  $r_X^k(x_{0:k})$ . We formally state that  $\mathcal{L}(D_X^K)$  is a well defined filtration in the following proposition.

**Proposition 2** The filtration  $\mathcal{L}(D_X^K)$  induced from a given dissimilarity network  $D_X^K$  is a well-defined filtration.

**Proof:** See Appendix A. ■

Proposition 2 also implies that  $L(D_X^K)$  is a well-defined simplicial complex. It establishes the connection between high order networks with filtrations. Tools developed in computational topology can therefore be used to analyze networked data. Figure 4 (c) presents an example of filtration in which the numbers adjacent to simplices denote the time when simplices appear.  $L_0$  consists of all vertices except  $[g]$  and  $L_{0.2}$  consists of all vertices union four 1-simplices  $[a, h]$ ,  $[b, f]$ ,  $[g, f]$ ,  $[d, e]$ .

Finally we present an intuitive idea of persistent homology. Consider the homological feature represented by  $\Phi$  that exists in the  $k$ -th homology group  $H_k(L_\alpha)$  for any  $\alpha$  satisfying  $\alpha_b \leq \alpha \leq \alpha_d$ . This feature starts to appear in the homology

group from time  $\alpha_b$  as a new independent non-trivial cycle. The feature diminishes at time  $\alpha_d$  since this cycle is trivialized by a boundary. This formation and elimination in the homology group of simplicial complexes in a filtration is defined as persistent homology. The interval  $[\alpha_b, \alpha_d]$  is named persistence interval for the corresponding homological feature and can also be represented as a point  $(\alpha_b, \alpha_d)$  in a two-dimensional plane. The collection of all such points for the  $k$ -th dimensional homological features is called the  $k$ -th dimensional persistence diagram and denoted as  $\mathcal{P}_k \mathcal{L}$ . For the example in Figure 4 (c), at time 0,  $L_0$  consists of all 0-simplices except  $[g]$ . Since every 0-simplex is a 0-cycle, there exist 7 zeroth-dimensional homological features. At time 0.1, the appearance of the 1-simplex  $[b, f]$  makes the 0-cycles  $[b]$  and  $[f]$  dependent and one zeroth-dimensional homological features dies, generating a zeroth-dimensional persistence interval  $[0, 0.1)$ . At the same time, a new zeroth-dimensional homological feature represented by  $[g]$  appears. As the filtration continues and more edges appear, all zeroth-dimensional homological features disappear except one denoting the entire connected component. For the first dimension, the homological feature represented by the cycle  $[a, h] + [h, g] + [g, a]$  appears at time 0.3 and is killed at time 0.8 by the appearance of the triangle  $[a, g, h]$ . At the end of the filtration, we have one zeroth-dimensional homological feature born at time 0 and one first-dimensional homological feature represented by  $[c, d] + [d, e] + [e, c]$  born at time 0.4. Figure 4 (d) and (e) plot the zeroth-dimensional and the first-dimensional persistence diagrams of the filtration induced from (c).

Persistence homologies can be computed fastly [24], [25]. For this reason, we aim to use them to lower bound network distances defined in Definition 8. We will focus on the analysis of dissimilarity networks. The analysis is then generalized to proximity networks as a result of the duality in Proposition 1. We start by finding bounds of  $\infty$ -norm dissimilarity network distances  $d_{\mathcal{D}, \infty}$  which we undertake in the next section.

#### IV. PERSISTENCE BOUNDS ON NETWORK DISTANCES

It has been shown that the persistent homology of filtrations on a simplicial complex is stable in the sense that similar filtrations yield similar persistence diagrams [28]. In future definitions, it would be important to consider the difference between persistence diagrams. We formalize this by introducing the notion of bottleneck distance between point sets as we specify next.

**Definition 11** The bottleneck distance  $d_B^\infty(Q, \tilde{Q})$  between two point sets  $Q$  and  $\tilde{Q}$  with same cardinality  $|Q| = |\tilde{Q}|$  in two dimensional space is defined as

$$d_B^\infty(Q, \tilde{Q}) = \min_{\gamma} \max_{q \in Q} \|q - \gamma(q)\|_\infty, \quad (12)$$

where  $\gamma$  ranges over all bijections from  $Q$  to  $\tilde{Q}$ .

For a given bijection  $\gamma$  between  $Q$  and  $\tilde{Q}$ , the infinity norm  $\|q - \gamma(q)\|_\infty$  between a pair of points  $q \in Q$  and  $\gamma(q) \in \tilde{Q}$  is assigned as the larger one among the  $x$ -axis coordinate difference and the  $y$ -axis coordinate difference between points  $q$  and  $\gamma(q)$ . The maximum over infinity norm  $\max_{q \in Q} \|q - \gamma(q)\|_\infty$  selects the maximum infinity norm  $\|q - \gamma(q)\|_\infty$  among all pairs in the bijection. The bottleneck distance in (12) is defined by selecting the bijection that minimizes these maximal infinity norms. Persistence diagrams are stable [28, Thm. 4.4] as we formally restate from a different perspective to favor our context.



**Lemma 1** *If two different filtrations  $\mathcal{L}$  and  $\tilde{\mathcal{L}}$  on a same simplicial complex yield finite dimensional persistence homologies and satisfy the strongly  $\delta$  interleaved condition, that is for any  $\alpha \geq 0$ ,*

$$L_\alpha \subseteq \tilde{L}_{\alpha+\delta} \subseteq L_{\alpha+2\delta}, \quad (13)$$

*then the bottleneck distance between the  $k$ -th dimensional persistence diagrams of  $\mathcal{L}$  and  $\tilde{\mathcal{L}}$  for any  $k$  is bounded by  $\delta$ , i.e.  $d_B^\infty(\mathcal{P}_k\mathcal{L}, \mathcal{P}_k\tilde{\mathcal{L}}) \leq \delta$ .*

Given two filtrations on a same simplicial complex, Lemma 1 guarantees that if the appearance time of any simplicies in two filtrations does not differ by  $\delta$ , the bottleneck distance between the persistence diagrams of the filtrations for any dimension is no greater than  $\delta$ .

**Remark 1** When the number of points in  $\mathcal{P}_k\mathcal{L}$  and  $\mathcal{P}_k\tilde{\mathcal{L}}$  is different, the bottleneck distance  $d_B^\infty(\mathcal{P}_k\mathcal{L}, \mathcal{P}_k\tilde{\mathcal{L}})$  is ill-defined. Instead, we may think that the filtrations have a number of homological features being trivialized at the same time they appear in the filtration, resulting in a number of points on the diagonal (the line with  $x$ -axis coordinate equating  $y$ -axis coordinate) in the corresponding persistence diagrams. To evaluate the difference between the persistence diagrams  $\mathcal{P}_k\mathcal{L}$  and  $\mathcal{P}_k\tilde{\mathcal{L}}$ , we can add a number of diagonal points to the persistence diagram that has fewer nodes. The points in the respective diagrams form two sets  $\{q_i\}$  and  $\{\tilde{q}_i\}$  of  $N$  points in  $\mathbb{R}^2$ . The bottleneck distance between  $\mathcal{P}_k\mathcal{L}$  and  $\mathcal{P}_k\tilde{\mathcal{L}}$  is then the optimal result of the Linear Bottleneck Assignment Problem (LBAP):  $\min_\pi \max_i c(q_i, \tilde{q}_{\pi(i)})$ , where  $\pi$  ranges over all permutations and the cost

$$c(q, \tilde{q}) = \min \left\{ \|q - \tilde{q}\|_\infty, \frac{1}{2} \max \{|q_x - \tilde{q}_y|, |q_x - \tilde{q}_y|\} \right\}. \quad (14)$$

In other words,  $c(q, \tilde{q})$  chooses the smaller one from the cost of matching  $q$  with  $\tilde{q}$  and the cost of matching both points with their respective closest point on the diagonal. Notice that  $c(q, \tilde{q}) = c(q, \tilde{q}')$  for any  $q$  whenever  $\tilde{q}, \tilde{q}'$  are on the diagonal. This makes the choice of the locations of the diagonal points added in the preceding section unsubstantial. The LBAP algorithm can be solved by applying the threshold algorithm described in [29, Algorithm 6.1] with running time  $O(N^{2.5})$  polynomial to the maximum number of points  $N$  in the persistence diagrams.

In this section we apply Lemma 1 to prove that the bottleneck distance between the persistence diagrams of the filtrations induced by two dissimilarity networks is a lower bound of their dissimilarity network distance. The result is then generalized to proximity networks. We begin with a formal statement.

**Theorem 1** *Let  $D_X^K$  and  $D_Y^K$  be two  $K$ -order dissimilarity networks. The bottleneck distance between the  $k$ -th dimensional persistence diagrams of the filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  is at most  $d_{\mathcal{D},\infty}(D_X^K, D_Y^K)$  for any  $0 \leq k \leq K$ , i.e.*

$$d_B^\infty(\mathcal{P}_k\mathcal{L}(D_X^K), \mathcal{P}_k\mathcal{L}(D_Y^K)) \leq d_{\mathcal{D},\infty}(D_X^K, D_Y^K). \quad (15)$$

To show Theorem 1, we need to transform filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  which are defined on two different simplicial complexes to two filtrations on a same simplicial complex. This enables us to consider if the collection of simplicies in two filtrations are  $\delta$ -nested as in (13), upon which Lemma 1 can be applied. We achieve this by introducing augmented networks which we describe formally in the following definition.

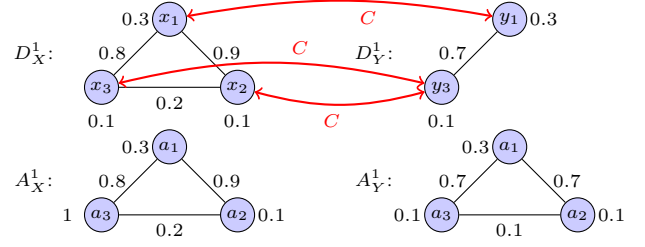


Fig. 5. An example to construct augmented networks.  $D_X^1$  and  $D_Y^1$  are 1-order dissimilarity networks. The augmented network constructed based on the correspondence  $C$  induces  $A_X^1 = D_X^1$ . An additional node  $a_2$  and two edges are augmented in  $A_Y^1$ .

**Definition 12** *Given two  $K$ -order dissimilarity networks  $D_X^K$  and  $D_Y^K$  and a correspondence  $C$  between their node sets  $X$  and  $Y$ , the augmented networks  $A_{X,C}^K$  and  $A_{Y,C}^K$  are a pair of  $K$ -order networks defined on the node space  $C$ . Each node  $a_i$  in  $A_{X,C}^K$  and  $A_{Y,C}^K$  represents a correspondent pair  $(x_{C_i}, y_{C_i})$  in  $C$ . Relationship functions for  $a_{0:k}$  with  $0 \leq k \leq K$  are defined as*

$$r_{A_X}^k(a_{0:k}) = r_X^k(x_{C_0:C_k}), \quad r_{A_Y}^k(a_{0:k}) = r_Y^k(y_{C_0:C_k}). \quad (16)$$

*Dissimilarities in the augmented networks are defined as*

$$d_{A_X}^k(a_{0:k}) = d_X^k(x_{C_0:C_k}), \quad d_{A_Y}^k(a_{0:k}) = d_Y^k(y_{C_0:C_k}). \quad (17)$$

*The  $X$ -rank  $s_X(a_{0:k})$  of a given tuple  $a_{0:k} \in A$  is the rank of the corresponding nodes  $x_{C_0:C_k}$  in  $D_X^K$ ,  $s_X(a_{0:k}) = s(x_{C_0:C_k})$ . The  $Y$ -rank is defined similarly,  $s_Y(a_{0:k}) = s(y_{C_0:C_k})$ .*

When the underlying correspondence is clear,  $C$  is omitted in the subscripts. For a given pair of dissimilarity networks  $D_X^K$  and  $D_Y^K$  and a correspondence  $C$  between  $X$  and  $Y$ , both the augmented networks  $A_X^K$  and  $A_Y^K$  have identical  $|C|$  nodes where  $|C|$  denotes the number of correspondence pairs in  $C$ . Each node  $a_i$  in both  $A_X^K$  and  $A_Y^K$  represents a correspondent pair  $(x_{C_i}, y_{C_i})$  in  $C$ . For each tuple  $a_{0:k}$ , its relationship  $r_{A_X}^k(a_{0:k})$  for the network  $A_X^K$  is the same as the relationship  $r_X^k(x_{C_0:C_k})$  between the tuple  $x_{C_0:C_k}$  in  $D_X^K$ . Dissimilarities  $d_{A_X}^k(a_{0:k})$  are the dissimilarity between the tuple  $x_{C_0:C_k}$ . From Definition 12 and the decompositions of relationship functions for dissimilarity networks [cf. (6)], relationship functions for augmented networks can be written alternatively as

$$r_{A_X}^k(a_{0:k}) = d_X^k(x_{C_0:C_k}) + \epsilon s(x_{C_0:C_k}). \quad (18)$$

Note that augmented networks are generally not dissimilarity networks. When a node  $x \in X$  appears twice in the correspondence, we would have two different nodes  $a \neq a'$  in the node space of  $A_X^K$  representing the same node  $x$  in the correspondence pairs. From (16) and the identity property of high order networks,  $r_{A_X}^1(a, a') = r_X^1(x, x) = r_X^0(x) = r_{A_X}^0(a)$ , which conflicts with the property of dissimilarity networks that adding a new element to a tuple makes the set strictly more dissimilar than it was. An example to construct augmented networks is illustrated in Figure 5. Well-defined filtrations can be induced from augmented networks similarly to Definition 10 as we formally state next.

**Definition 13** *Given an augmented  $A_X^K$ , its induced simplicial complex  $L(A_X^K)$  is defined on the same node space  $C$ . For point collection  $a_{0:k}$  with full rank tuple  $a_{0:k}$ , the simplex  $[a_{0:k}]$  is in  $L(A_X^K)$  if the dissimilarity term  $d_{A_X}^k(a_{0:k})$  is not the maximum dissimilarity  $\max_{\tilde{a}_{0:k} \in C^{k+1}} d_{A_X}^K(\tilde{a}_{0:k})$ . The induced filtration*

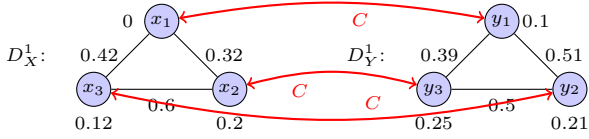


Fig. 6. An example where the bottleneck distance between the  $k$ -th dimensional persistence diagrams of the filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  is the same as their  $\infty$ -norm network distance for  $k \in \{0, 1\}$ . The optimal correspondence  $C$  yields  $d_{\mathcal{D}, \infty}(D_X^1, D_Y^1) = 0.1$ . The zeroth dimensional persistence intervals for  $\mathcal{L}(D_X^1)$  are  $[0, \infty)$ ,  $[0.12, 0.42)$ ,  $[0.2, 0.32)$  and for  $\mathcal{L}(D_Y^1)$  are  $[0.1, \infty)$ ,  $[0.21, 0.51)$ ,  $[0.25, 0.39)$ . The first dimensional persistence interval for  $\mathcal{L}(D_X^1)$  is  $[0.6, \infty)$  and for  $\mathcal{L}(D_Y^1)$  is  $[0.5, \infty)$ . The bottleneck distances between the zeroth as well as the first dimensional persistence diagrams of the filtrations induced from the two networks are 0.1.

$\mathcal{L}(A_X^K)$  is the filtration where each simplex  $[a_{0:k}]$  in  $L(A_X^K)$  appears at time  $r_{A_X}^k(a_{0:k})$ .

Augmented networks are valid  $K$ -order networks. They satisfy the symmetry and identity properties. The induced filtrations are well-founded filtrations and most importantly, preserve the homological features of the filtrations induced from the original dissimilarity networks. We state this in the next proposition.

**Proposition 3** Given a  $K$ -order dissimilarity network  $D_X^K$  and a correspondence  $C$  between  $X$  and the node space  $Y$ , the augmented network  $A_X^K$  constructed by Definition 12 is a valid  $K$ -order network. The filtration  $\mathcal{L}(A_X^K)$  induced by Definition 13 is a well-defined filtration and has identical persistence diagrams as  $\mathcal{L}(D_X^K)$ , i.e. for any  $0 \leq k \leq K$ ,

$$\mathcal{P}_k \mathcal{L}(A_X^K) = \mathcal{P}_k \mathcal{L}(D_X^K). \quad (19)$$

**Proof:** See Appendix B. ■

Finally we use Lemma 1 and Proposition 3 to prove Theorem 1.

**Proof of Theorem 1:** See Appendix B. ■

Theorem 1 provides theoretical justification that high order networks can be compared via the persistence diagrams of their induced filtrations. We emphasize that the lower bounds are tight. First, it follows from the definition of dissimilarity networks that all the  $k$ -order dissimilarities between full rank tuples of dissimilarity network  $D_X^K$  can be found in homological features. We formally state this in the following proposition.

**Proposition 4** Given a dissimilarity network  $D_X^K$ , any of its  $k$ -order dissimilarities between full rank tuples appear either in the death time of the  $(k-1)$ -th dimensional homological features or the birth time of the  $k$ -th dimensional homological features.

**Proof:** See Appendix C. ■

Proposition 4 guarantees that nothing about dissimilarities between full rank tuples is lost when we consider the persistence homologies of the induced filtrations. Moreover, there exist dissimilarity networks  $D_X^K$  and  $D_Y^K$  such that the bottleneck distance between the  $k$ -th dimensional persistence diagrams of the filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  is the same as their  $\infty$ -norm network distance for all orders  $0 \leq k \leq K$ . See Figure 6 for an illustration where  $d_B^\infty(\mathcal{P}_k \mathcal{L}(D_X^1), \mathcal{P}_k \mathcal{L}(D_Y^1)) = d_{\mathcal{D}, \infty}(D_X^1, D_Y^1) = 0.1$  for  $k \in \{0, 1\}$ .

**Remark 2** Theorem 1 establishes a relationship between  $d_{\mathcal{D}, \infty}$  and the bottleneck distance between the persistence diagrams of the filtrations. Using properties of  $p$ -norms, relationships between general  $p$ -norm network distance  $d_{\mathcal{D}, p}$  and the bottleneck distance between the persistence diagrams can also be constructed.

**Remark 3** In [30], persistence diagrams are used to compare point cloud data in three dimensional Euclidean space. It is shown that the bottleneck distance between the persistence diagrams of the Rips filtrations [31] constructed from two metric spaces is a lower bound to the Gromov-Hausdorff (GH) distance between the metric spaces. Theorem 1 can be considered as a generalization in which the filtrations induced from high order networks generalize the Rips filtrations constructed from the metric space. Similar properties of persistence homologies are established in a more universal setting.

The relationship between the dissimilarity network distance and the bottleneck distance can be generalized to proximity networks via the duality established in Proposition 1. We state this formally in the following corollary.

**Corollary 1** Consider two proximity networks  $P_X^K$  and  $P_Y^K$  and their corresponding dual dissimilarity networks  $D_X^K$  and  $D_Y^K$ . The bottleneck distance between the  $k$ -th dimensional persistence diagrams of the filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  is at most  $d_{\mathcal{P}, \infty}(P_X^K, P_Y^K)$  for any  $0 \leq k \leq K$ , i.e.

$$d_B^\infty(\mathcal{P}_k \mathcal{L}(D_X^K), \mathcal{P}_k \mathcal{L}(D_Y^K)) \leq d_{\mathcal{P}, \infty}(P_X^K, P_Y^K). \quad (20)$$

#### A. Persistence bounds on $k$ -order distances

Theorem 1 provides guarantee that the  $\infty$ -norm dissimilarity network distance  $d_{\mathcal{D}, \infty}$  can be tightly bounded by the bottleneck distance between the persistence diagrams. In this section we base on Theorem 1 to demonstrate that  $k$ -order dissimilarity network distance  $d_{\mathcal{D}}^k$  between relationship functions for a specific order  $k$  can also be tightly bounded using persistence homologies. We begin with a formal statement as next.

**Theorem 2** Given two dissimilarity networks  $D_X^K$  and  $D_Y^K$  and an integer  $1 \leq k \leq K$ , the bottleneck distance between the  $k'$ -th dimensional persistence diagrams of the filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  is at most  $d_{\mathcal{D}}^k(D_X^K, D_Y^K)$  for any  $0 \leq k' \leq k-1$ , i.e.

$$d_B^\infty(\mathcal{P}_{k'} \mathcal{L}(D_X^K), \mathcal{P}_{k'} \mathcal{L}(D_Y^K)) \leq d_{\mathcal{D}}^k(D_X^K, D_Y^K). \quad (21)$$

To show Theorem 2, first observe that for high order networks, all information conveyed by relationship functions  $r_X^{k'}$  is included in  $r_X^k$  for orders  $k' \leq k$ . This is because for any tuple  $x_{0:k'}$ , we can construct an extended tuple  $x_{0:k}$  by adding some nodes already included in  $x_{0:k'}$  such that their ranks are identical  $s(x_{0:k'}) = s(x_{0:k})$ . It then follows from the identity property of high order networks that  $r_X^{k'}(x_{0:k'}) = r_X^k(x_{0:k})$ . This also establishes a relationship for network distance between different orders as we present next.

**Proposition 5** Given high order networks  $N_X^K$  and  $N_Y^K$ , it follows that  $d_{\mathcal{N}}^{k'}(N_X^K, N_Y^K) \leq d_{\mathcal{N}}^k(N_X^K, N_Y^K)$  for  $0 \leq k' \leq k \leq K$ .

**Proof:** See Appendix C. ■

Proposition 5 implies that the  $k$ -order network distance increases as the order  $k$  becomes higher.  $k$ -order dissimilarity or

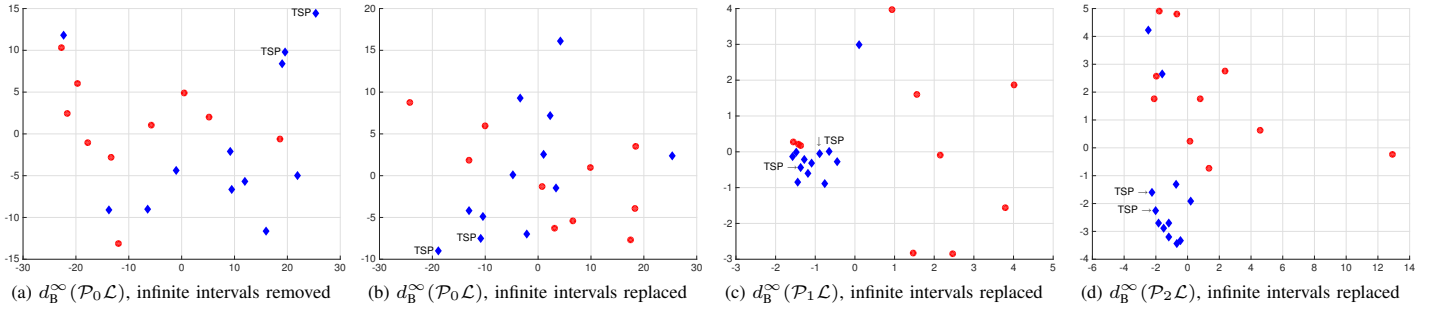


Fig. 7. Two dimensional Euclidean embeddings of the networks constructed from quinquennial publications in engineering and mathematics journals with respect to the network metric lower bounds  $d_B^\infty(\mathcal{P}_0\mathcal{L})$  with infinite intervals removed or replaced by finite intervals,  $d_B^\infty(\mathcal{P}_1\mathcal{L})$ , and  $d_B^\infty(\mathcal{P}_2\mathcal{L})$  with infinite intervals replaced. In the embeddings, red circles denote networks constructed from mathematics journals and blue diamonds represent networks from engineering journals. Networks constructed from publications of TSP are labeled.

proximity network distances inherit this property that  $d_{\mathcal{D}}^{k'} \leq d_{\mathcal{D}}^k$  and  $d_{\mathcal{P}}^{k'} \leq d_{\mathcal{P}}^k$  for any  $0 \leq k' \leq k \leq K$ . Proposition 5 also connects  $k$ -order network distances with  $\infty$ -norm network distance as we state in the next corollary.

**Corollary 2** *Given high order networks  $N_X^K$  and  $N_Y^K$ , it follows that  $d_{\mathcal{N},\infty}(N_X^K, N_Y^K) = d_{\mathcal{N}}^K(N_X^K, N_Y^K)$ .*

**Proof:** See Appendix C. ■

When the input networks are dissimilarity networks or proximity networks, we have that  $d_{\mathcal{D},\infty} = d_{\mathcal{D}}^K$  and  $d_{\mathcal{P},\infty} = d_{\mathcal{P}}^K$ . Finally we use Theorem 1 and Corollary 2 to prove Theorem 2.

**Proof of Theorem 2:** See Appendix C. ■

The requirement  $k' \leq k - 1$  in Theorem 2 comes from the fact that the  $k'$ -th dimensional persistence diagram conveys information of not only  $k'$ -th order but also  $(k' + 1)$ -th order relationship functions. We again emphasize that the lower bound described in Theorem 2 is tight and we can find dissimilarity networks  $D_X^K$  and  $D_Y^K$  such that  $d_{\mathcal{D}}^k(D_X^K, D_Y^K)$  equals the bottleneck distance between the  $k'$ -th dimensional persistence diagrams of the filtration  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  for any  $0 \leq k' \leq k - 1 \leq K$ . Similar to Theorem 1, the results in Theorem 2 can be generalized for proximity networks via duality as we present next.

**Corollary 3** *Consider two proximity networks  $P_X^K$  and  $P_Y^K$  and their corresponding dissimilarity networks  $D_X^K$  and  $D_Y^K$ . The bottleneck distance between the  $k'$ -th dimensional persistence diagrams of the filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(D_Y^K)$  is at most  $d_{\mathcal{P}}^k(P_X^K, P_Y^K)$  for any  $0 \leq k' \leq k - 1$ , i.e.*

$$d_B^\infty(\mathcal{P}_{k'}\mathcal{L}(D_X^K), \mathcal{P}_{k'}\mathcal{L}(D_Y^K)) \leq d_{\mathcal{P}}^k(P_X^K, P_Y^K). \quad (22)$$

## V. COMPARISON OF COAUTHORSHIP NETWORKS

We apply the lower bounds established in Corollaries 1 and 3 to compare second order coauthorship networks where relationship functions denote the number of publications of single authors, pairs of authors, and triplets. These coauthorship networks are proximity networks because they satisfy the order decreasing property in Definition 7. We consider publications in 5 journals from mathematics community: Computational Geometry (CG), Discrete Computational Geometry (DCG), Journal of Applied Probability (JAP), Journal of Mathematical Analysis and Applications (JMAA), SIAM Journal on Numerical Analysis (SJNA), and 6 journals from engineering community, all from IEEE: Signal Processing Magazine (SPM), Trans. Automatic Control (TAC), Trans. Pattern Analysis and Machine Intelligence (TPAMI), Trans.

Information Theory (TIT), Trans. Signal Processing (TSP), Trans. Wireless Communication (TWC). For each journal, we construct networks for the 2004-2008 and 2009-2013 quinquennia. For TAC, TSP, and TWC, we also construct networks for each individual year from 2004 to 2013. Lists of publications are queried from [32].

For each of these journals we consider all publications in the period of interest and construct proximity networks where the node space  $X$  is formed by all authors of the publications. Zeroth order proximities are defined as the total number of publications of each member of the network, first order proximities as the number of papers coauthored by pairs, and second order proximities as the number of papers coauthored by triplets. The positive constant  $\epsilon$  in the relationship functions [cf. (8)] is set sufficiently small compared to proximities. To make networks with different numbers of papers comparable we normalize all relationships by the total number of papers in the network. There are papers with more than three coauthors but we don't record proximities of order higher than 2. The persistence homologies of the induced filtrations of networks are computed using JavaPlex [33]. There exist persistence intervals that never diminish and these would dominate the bottleneck distance between persistence diagrams as in Definition 11. To solve this problem, we can either remove these persistence intervals or replace these intervals, e.g.  $[\alpha_b, \infty)$ , by persistence intervals  $[\alpha_b, 1)$  which are born at the same time but disappear at time 1 – the highest possible finite value. We note that the bottleneck distance between the persistence diagrams with infinite intervals either removed or replaced is still lower bounds of the network distances. By assuming that networks from the same community or constructed from the same journal have similar collaboration patterns, we show here that network metric lower bounds succeed in identifying these patterns and in distinguishing coauthorship networks from communities with different research interests.

### A. Engineering and mathematics communities

Figure 7 shows the two dimensional Euclidean embeddings [34] of the network metric lower bound  $d_B^\infty(\mathcal{P}_0\mathcal{L})$  with infinite persistence intervals removed or replaced by finite counterparts and of the lower bounds  $d_B^\infty(\mathcal{P}_1\mathcal{L})$  and  $d_B^\infty(\mathcal{P}_2\mathcal{L})$  with infinite intervals replaced. The 12 engineering networks (blue diamonds) separate clearly from the 10 mathematics networks (red circles) in all of the network metric lower bounds considered. An unsupervised classification with one linear boundary run across all four lower bounds would generate errors of 2 (9.09%) to 5 (22.73%) out of 22 networks. Euclidean embedding of the lower



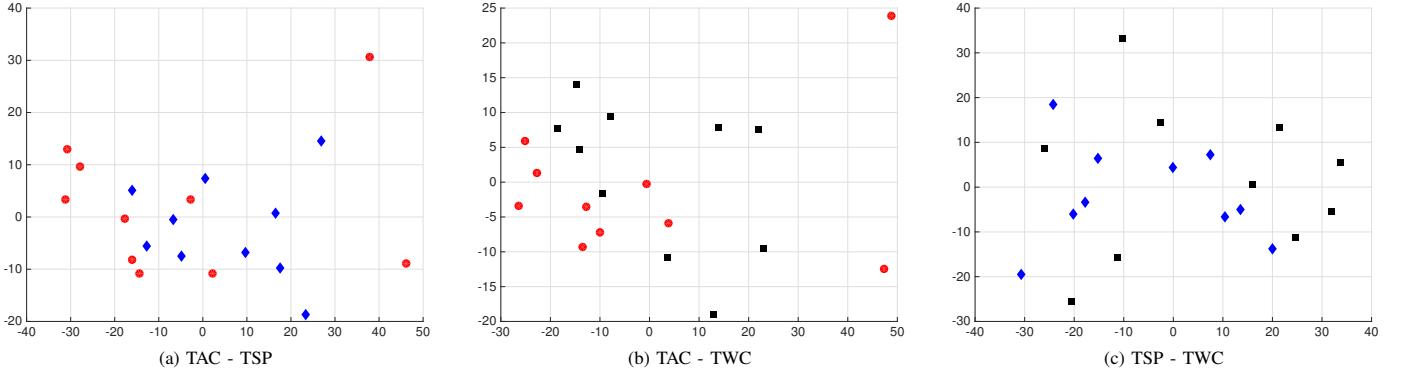


Fig. 8. Two dimensional Euclidean embeddings of the networks constructed from annual publications in TAC, TSP, and TWC with respect to the summation of the metric lower bounds  $d_B^\infty(\mathcal{P}_0\mathcal{L})$ ,  $d_B^\infty(\mathcal{P}_1\mathcal{L})$ , and  $d_B^\infty(\mathcal{P}_2\mathcal{L})$  with infinite intervals replaced by finite intervals. In the embeddings, red circles represent TAC, blue diamonds TSP, and black squares TWC.

bound  $d_B^\infty(\mathcal{P}_1\mathcal{L})$  with infinite intervals removed yields similar results. Most networks do not possess finite second dimensional persistence intervals and for this reason we do not consider the lower bound  $d_B^\infty(\mathcal{P}_2\mathcal{L})$  with infinite intervals removed.

We emphasize that networks constructed from the same journal with different quinquennia tend to be close to each other with respect to the metric lower bounds. As an illustration, the networks constructed from TSP with different quinquennia are marked in the embeddings and it is clear that their differences in persistence homologies are considerably low. Such scenarios are observed for several other journals as well. We also emphasize that bottleneck distances between the first and second dimensional persistence diagrams provide different insights compared to that based on the zeroth dimensional homological features. Indeed, we see clearer clustering patterns formed by networks constructed from engineering and mathematics journals. This can be viewed as an example to demonstrate the advantages of considering higher dimensional relationship functions besides the conventional pairwise ones.

### B. Engineering communities with different research interests

The network metric lower bounds succeed in distinguishing the different collaboration patterns in engineering and mathematics communities. We now illustrate that the lower bounds are also able to identify distinctive features of engineering communities with different research interests. To see this we consider the networks constructed from annual publications of TAC, TSP, and TWC.

Figure 8 shows the two dimensional Euclidean embeddings of the networks with respect to the summation of the network metric lower bounds  $d_B^\infty(\mathcal{P}_0\mathcal{L})$ ,  $d_B^\infty(\mathcal{P}_1\mathcal{L})$ , and  $d_B^\infty(\mathcal{P}_2\mathcal{L})$  with infinite intervals replaced by finite counterparts. The incorporation of the metric lower bounds  $d_B^\infty(\mathcal{P}_0\mathcal{L})$  with infinite intervals removed yields similar results. Most networks do not possess finite non-trivial first dimensional nor second dimensional persistence intervals and for this reason we do not consider lower bounds  $d_B^\infty(\mathcal{P}_1\mathcal{L})$  and  $d_B^\infty(\mathcal{P}_2\mathcal{L})$  with infinite intervals removed.

We expect more variations in annual networks because the time for averaging behavior is reduced. Besides, it is hard to argue that intrinsic and obvious differences exist in the collaboration patterns in automatic control, signal processing, and wireless communication communities. Still, networks constructed from the same journal but different annuals tend to be close to each other and form clustering structures. An unsupervised classification with one linear boundary run across the summation of lower bounds would generate 4 (20%) errors out of 20 networks in all three

classification problems considered. The less obvious clustering structure formed by networks from different journals in Figure 8 (c) compared to (a) and (b) also suggests that the collaboration patterns in research communities of signal processing and wireless communication are more similar compared to that of automatic control.

## VI. CONCLUSION

We establish connections between high order networks and simplicial complexes and use the differences between the induced homological features to evaluate the differences between networks. We justify that this is a lower bound to two families of valid metrics in the space of high order networks modulo permutation isomorphisms. These lower bounds succeed in distinguishing the collaboration patterns of engineering communities from mathematics communities and in discriminating engineering communities with different research interests.

### APPENDIX A

#### PROOF OF PROPOSITION 2

We first prove that the simplicial complex  $L(D_X^K)$  induced from a given dissimilarity network  $D_X^K$  is a well-defined simplicial complex. To that end, we demonstrate that (i) every face of a simplex of  $L(D_X^K)$  is also in  $L(D_X^K)$  and (ii) the intersection of any two simplices is either empty or a shared face. To see (i), suppose that simplex  $[x_{0:k}]$  is a simplex of  $L(D_X^K)$ . From Definition 10, we know that the dissimilarity between vertices of the simplex  $d_X^k(x_{0:k})$  is not the maximum dissimilarity in the network, i.e.

$$d_X^k(x_{0:k}) < \max_{\tilde{x}_{0:K} \in X^{K+1}} d_X^K(\tilde{x}_{0:K}). \quad (23)$$

For any faces of  $[x_{0:k}]$ , say  $[x_{0:\tilde{k}}]$  with  $\tilde{k} \leq k$ , from the order increasing property [cf. (7)] of dissimilarities in dissimilarity networks, we know that  $d_X^{\tilde{k}}(x_{0:\tilde{k}}) \leq d_X^k(x_{0:k})$ . Combining this with (23) yields  $d_X^{\tilde{k}}(x_{0:\tilde{k}}) < \max_{\tilde{x}_{0:K} \in X^{K+1}} d_X^K(\tilde{x}_{0:K})$ . Therefore, every face of a simplex of  $L(D_X^K)$  is also in  $L(D_X^K)$ . To see (ii), observe that simplicial complex  $L(D_X^K)$  is defined on the node space  $X$  and therefore the intersection of any two simplices must be the convex hull of some tuple  $x_{0:k} \in X$ . If the intersection is not empty, it must be a face for each of the two simplices. The analysis in proving (i) shows that the intersection also appears in  $L(D_X^K)$  and therefore it is a shared face.

Now we prove that  $\mathcal{L}(D_X^K)$  is a well-defined filtration. It suffices to demonstrate that for any  $\alpha$ , all faces of each simplex

and intersections of any simplices in  $L_\alpha$  also appear before or on time  $\alpha$ . Suppose simplex  $[x_{0:k}]$  appears before time  $\alpha$  for some chosen  $\alpha$ . From Definition 10, we know that  $r_X^k(x_{0:k}) \leq \alpha$ . For any faces of  $[x_{0:k}]$ , say  $[x_{0:\tilde{k}}]$  with  $\tilde{k} < k$ , the order increasing property implies  $d_X^k(x_{0:\tilde{k}}) \leq d_X^k(x_{0:k})$ . Moreover, the rank of the tuple  $x_{0:\tilde{k}}$  cannot be higher than the rank of  $x_{0:k}$ . Combining these with the decomposition of relationship functions of dissimilarity networks [cf. (6)] yields

$$r_X^{\tilde{k}}(x_{0:\tilde{k}}) \leq r_X^k(x_{0:k}) \leq \alpha, \quad (24)$$

which shows that the face  $[x_{0:\tilde{k}}]$  appears before time or on time  $\alpha$ . Similarly, it can be shown that the intersections of any simplices in  $L_\alpha$  appear before  $\alpha$ , concluding the proof.

## APPENDIX B

### PROOFS OF PROPOSITION 3 AND THEOREM 1

In the proof of Proposition 3 we use two lemmas. The first one establishes properties between relationships of tuples of different lengths in augmented networks.

**Lemma 2** *Relationship functions in an augmented network  $A_X^K$  as defined in Definition 12 satisfy the following properties:*

**Order increasing.**  $r_{A_X}^k(a_{0:k}) \geq r_{A_X}^{k-1}(a_{0:k-1})$  and  $d_{A_X}^k(a_{0:k}) \geq d_{A_X}^{k-1}(a_{0:k-1})$  for any  $a_{0:k} \in A$ .

**Core tuple.** For any  $a_{0:k}$ , if the  $X$ -rank of the tuple formed by deleting a node  $a_i$  is the same as the  $X$ -rank of the original tuple, i.e.  $s_X(a_{0:\hat{i}:k}) = s_X(a_{0:k})$ , deleting  $a_i$  results in the same relationship  $r_{A_X}^{k-1}(a_{0:\hat{i}:k}) = r_{A_X}^k(a_{0:k})$  and dissimilarity  $d_{A_X}^{k-1}(a_{0:\hat{i}:k}) = d_{A_X}^k(a_{0:k})$ .

**Proof:** To see the order increasing property, from (18) we have

$$\begin{aligned} r_{A_X}^k(a_{0:k}) &= d_X^k(x_{C_0:C_k}) + \epsilon s(x_{C_0:C_k}), \\ r_{A_X}^{k-1}(a_{0:k-1}) &= d_X^{k-1}(x_{C_0:C_{k-1}}) + \epsilon s(x_{C_0:C_{k-1}}). \end{aligned} \quad (25)$$

The order increasing property of dissimilarity network  $D_X^K$  implies  $d_X^k(x_{C_0:C_k}) \geq d_X^{k-1}(x_{C_0:C_{k-1}})$ . Dissimilarities in augmented networks are the same as dissimilarities in original networks. Therefore,  $d_{A_X}^k(a_{0:k}) \geq d_{A_X}^{k-1}(a_{0:k-1})$  illustrating the order increasing property for dissimilarities. Besides, removing a node  $x_{C_k}$  from the tuple  $x_{C_0:C_k}$  cannot make the rank  $s(x_{C_0:C_{k-1}})$  be greater than  $s(x_{C_0:C_k})$ . Combining these observations with (25) yields the order increasing property for relationship functions.

To see the core tuple property, the identity property of high order network  $D_X^K$  indicates  $r_X^{k-1}(x_{C_0:C_{\hat{i}}:C_k}) = r_X^k(x_{C_0:C_k})$ . Following the definition of relationships in augmented network  $A_X^K$  yields

$$r_{A_X}^{k-1}(a_{0:\hat{i}:k}) = r_X^{k-1}(x_{C_0:C_{\hat{i}}:C_k}) = r_X^k(x_{C_0:C_k}) = r_{A_X}^k(a_{0:k}). \quad (26)$$

Moreover, from the definition of  $X$ -rank,  $s_X(a_{0:\hat{i}:k}) = s_X(a_{0:k})$  implies  $s(x_{C_0:C_{\hat{i}}:C_k}) = s(x_{C_0:C_k})$ . Subtracting  $\epsilon s(x_{C_0:C_{\hat{i}}:C_k}) = \epsilon s(x_{C_0:C_k})$  from the left and right hand side of (26) and substituting into (25) implies  $d_{A_X}^{k-1}(a_{0:\hat{i}:k}) = d_{A_X}^k(a_{0:k})$ , completing the proof. ■

The second lemma is a statement about the boundary operation of simplices.

**Lemma 3** *For any simplex  $[a_{0:k}]$ ,  $\partial_{k-1}\partial_k[a_{0:k}] = 0$ .*

**Proof:** See [35, pp. 81]. ■

We now use Lemmas 2 and 3 to show Proposition 3.

**Proof of Proposition 3:** To prove Proposition 3, we prove the (i) symmetry property and (ii) identity property of high order networks [cf. Definition 1], (iii)  $\mathcal{L}(A_X^K)$  being a valid filtration, and (iv)  $\mathcal{L}(A_X^K)$  and  $\mathcal{L}(D_X^K)$  having identical persistent homology.

**Proof of the symmetry property:** For any point collections  $a_{0:k}$ , the symmetry property of dissimilarity network  $D_X^K$  and the definition of augmented networks imply  $r_{A_X}^k(a_{[0:k]}) = r_X^k(x_{[C_0:C_k]}) = r_X^k(x_{C_0:C_k}) = r_{A_X}^k(a_{0:k})$  for any reordering  $a_{[0:k]}$ . This shows the symmetry property of  $A_X^K$ . ■

**Proof of the identity property:** Given a point collection  $a_{0:k}$ , any of its subtuple  $a_{i_0:i_{\tilde{k}}}$  with  $s(a_{i_0:i_{\tilde{k}}}) = s(a_{0:k})$  would have the same elements in its node space  $C$  as the original tuple  $a_{0:k}$ . Therefore,  $x_{C_0:C_k}$  and  $x_{C_{i_0}:C_{i_{\tilde{k}}}}$  also have identical points in the node space  $X$  implying  $s(x_{C_0:C_k}) = s(x_{C_{i_0}:C_{i_{\tilde{k}}}})$ . The identity property of dissimilarity network  $D_X^K$  and the definition of augmented networks yield  $r_{A_X}^k(a_{0:k}) = r_X^k(x_{C_0:C_k}) = r_X^{\tilde{k}}(x_{C_{i_0}:C_{i_{\tilde{k}}}}) = r_{A_X}^{\tilde{k}}(a_{i_0:i_{\tilde{k}}})$ . This shows the identity property of  $A_X^K$ . ■

**Proof of  $\mathcal{L}(A_X^K)$  being a valid filtration:** The proof that  $L(A_X^K)$  induced from  $A_X^K$  is a valid simplicial complex is identical to the proof that  $L(D_X^K)$  induced from  $D_X^K$  is a valid simplicial complex as in Appendix A. This is because the order increasing property of dissimilarities in augmented networks also holds true as described in Lemma 2. That  $\mathcal{L}(A_X^K)$  is a well-defined filtration follows immediately from the order increasing property of relationships in augmented networks. ■

**Proof of  $\mathcal{L}(A_X^K)$  and  $\mathcal{L}(D_X^K)$  having identical persistence intervals:** To prove this statement we construct a sequence of nested correspondences

$$C_0 \subseteq C_1 \subseteq \dots \subseteq C_{|C|-|X|} = C. \quad (27)$$

In the sequence of correspondences,  $C_0$  is the largest subset of  $C$  where there does not exist a node  $x \in X$  with  $(x, y), (x, y') \in C_0$  for  $y, y' \in Y$  with  $y \neq y'$ . Since  $C_0$  is the largest subset satisfying this requirement, every node  $x \in X$  appears once and only once in some correspondences in  $C_0$ . For any  $0 \leq i \leq |C| - |X| - 1$ ,  $C_{i+1}$  is the correspondence formed by adding a new correspondence pair  $(x, y)$  for some  $y \in Y$  to the correspondence  $C_i$  such that the pair satisfies  $(x, y) \in C$  and  $(x, y) \notin C_i$ . We can perform such addition  $|C| - |X|$  times until every correspondence pair in  $C$  has been added and we recover the original correspondence  $C$ . An augmented network  $A_{X,C_i}^K$  can be defined for each correspondence  $C_i$  in the sequence (27), and this gives a sequence of augmented networks,

$$A_{X,C_0}^K, A_{X,C_1}^K, \dots, A_{X,C_{|C|-|X|}}^K = A_{X,C}^K = A_X^K. \quad (28)$$

Since every node  $x \in X$  appears once and only once in some correspondences in  $C_0$ , we can find a bijective mapping between node  $x \in X$  and node  $a = (x, y) \in C_0$ . It follows immediately from Definition 12 that  $D_X^K$  is isomorphic to  $A_{X,C_0}^K$ , and therefore filtrations  $\mathcal{L}(D_X^K)$  and  $\mathcal{L}(A_{X,C_0}^K)$  are identical. We now prove that for any  $0 \leq i \leq |C| - |X| - 1$ , induced filtrations  $\mathcal{L}(A_{X,C_i}^K)$  and  $\mathcal{L}(A_{X,C_{i+1}}^K)$  have identical persistence homologies.

Based on the constructions, there must exist two nodes  $a_j \neq a_{j'}$  in  $C_{i+1}$  with  $a_j \in C_i$  and  $a_{j'} \notin C_i$  such that  $a_j = (x, y)$

and  $a_{j'} = (x, y')$  for some common node  $x \in X$  and different nodes  $y \neq y'$  in  $Y$ .  $C_{i+1}$  has identical nodes as  $C_i$  except for the additional node  $a_{j'}$ . Persistence homologies of the filtrations  $\mathcal{L}(A_{X,C_i}^K)$  and  $\mathcal{L}(A_{X,C_{i+1}}^K)$  can only be different due to  $k$ -chains involving simplices  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$  for some  $k$  and some point collections  $a_{0:k-1}$  or  $a_{0:k-2}$ . In subsequent contexts, point collections  $a_{0:k-1}, a_{0:k-2}$  do not include  $a_j$  nor  $a_{j'}$ . Therefore, to prove that the induced filtrations  $\mathcal{L}(A_{X,C_i}^K)$  and  $\mathcal{L}(A_{X,C_{i+1}}^K)$  have identical persistence homologies, it suffices to demonstrate that for any  $\alpha$  and any  $k$ , the  $k$ -th homology group  $H_k(L_\alpha(A_{X,C_i}^K))$  and the  $(k-1)$ -th homology group  $H_{k-1}(L_\alpha(A_{X,C_i}^K))$  of the collection of simplices appearing before or on time  $\alpha$  will not be modified by  $k$ -chains involving  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$ .

We first prove that  $k$ -chains involving simplices  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$  will not create new non-trivial homological features compared to  $H_k(L_\alpha(A_{X,C_i}^K))$ . To see this, for any  $\alpha$ , let  $\Phi_0$  be a  $k$ -cycle  $\Phi_0 \in Z_k(L_\alpha(A_{X,C_{i+1}}^K))$  and involve simplices  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$ . Start with  $\tau = 0$  and conduct the following construction. If there is a simplex  $[a_{0:k-1,j'}]$  that is involved in  $\Phi_\tau$  for some tuple  $a_{0:k-1}$ , consider the simplex  $\Psi_\tau = \partial_{k+1}[a_{0:k-1,j,j'}]$ . Since the node in  $X$  represented by correspondent pairs  $a_j$  and  $a_{j'}$  are identical,  $s_X(a_{0:k-1,j,j'}) = s_X(a_{0:k-1,j'})$ . Following the core tuple property described in Lemma 2, we know that

$$d_{A_X, C_{i+1}}^{k+1}(a_{0:k-1,j,j'}) = d_{A_X, C_{i+1}}^k(a_{0:k-1,j'}). \quad (29)$$

$$r_{A_X, C_{i+1}}^{k+1}(a_{0:k-1,j,j'}) = r_{A_X, C_{i+1}}^k(a_{0:k-1,j'}). \quad (30)$$

Since  $\Phi_\tau \in Z_k(L_\alpha(A_{X,C_{i+1}}^K))$ ,  $[a_{0:k-1,j'}]$  is a simplex of  $L(A_{X,C_{i+1}}^K)$ . Definition 13 then implies  $d_{A_X, C_{i+1}}^{k-1}(a_{0:k-1,j,j'}) < \max_{\tilde{a}_{0:K} \in C^{K+1}} d_{A_X, C_{i+1}}^K(\tilde{a}_{0:K})$ . Substituting this inequality into (29) yields

$$d_{A_X, C_{i+1}}^{k+1}(a_{0:k-1,j,j'}) < \max_{\tilde{a}_{0:K} \in C^{K+1}} d_{A_X, C_{i+1}}^K(\tilde{a}_{0:K}), \quad (31)$$

which demonstrates that  $[a_{0:k-1,j,j'}]$  is also a simplex of  $L(A_{X,C_{i+1}}^K)$ . Similarly, from (30) and the fact that  $[a_{0:k-1,j'}]$  appears before or on time  $\alpha$ , we know that the simplex  $[a_{0:k-1,j,j'}]$  also appears before or on time  $\alpha$ . Combining these observations with the order increasing property from Lemma 2 indicates that every face of  $[a_{0:k-1,j,j'}]$  – every simplex in the chain  $\Psi_\tau$  – exists in  $L(A_{X,C_{i+1}}^K)$  and appears before or on time  $\alpha$ . Therefore,  $\Psi_\tau \in \mathbf{B}_k(L_\alpha(A_{X,C_{i+1}}^K))$ . Since every boundary is also a cycle,  $\Psi_\tau \in Z_k(L_\alpha(A_{X,C_{i+1}}^K))$ .

We form another  $k$ -chain

$$\Phi_{\tau+1} = \Phi_\tau - \hat{\beta}_\tau \Psi_\tau, \quad (32)$$

where the coefficient  $\hat{\beta}_\tau$  is chosen such that the simplex  $[a_{0:k-1,j'}]$  is not involved in the new chain  $\Phi_\tau$ . This is possible since  $[a_{0:k-1,j'}]$  appears in both  $\Psi_t$  and  $\Phi_t$ . Apart from the fact that  $[a_{0:k-1,j'}]$  does not involve in  $\Phi_{t+1}$ ,  $\Phi_{t+1}$  has two more important properties: (i) If  $\Phi_\tau \in Z_k(L_\alpha(A_{X,C_{i+1}}^K))$ ,  $\Phi_{\tau+1} \in Z_k(L_\alpha(A_{X,C_{i+1}}^K))$ . (ii)  $\Phi_{\tau+1}$  can not have any new simplices  $[a'_{0:k-1}]$  for some tuple  $a'_{0:k-1}$  compared to  $\Phi_\tau$ . (i) is because linear combination of cycles produce another cycle. To see (ii), it follows from (10) that

$$\Psi_\tau = \beta[a_{0:k-1,j}] + \beta'[a_{0:k-1,j'}] + \sum_{a''_{0:k-2}} \beta a''_{0:k-2}[a''_{0:k-2,j,j'}], \quad (33)$$

for some tuples  $a''_{0:k-2}$  and coefficients  $\beta, \beta', \beta a''_{0:k-2}$ . The only simplex  $[a'_{0:k-1,j'}]$  in  $\Psi_\tau$  for some tuple  $a'_{0:k-1}$  is the one

with  $a'_{0:k-1} = a_{0:k-1}$ . Hence,  $\Phi_{\tau+1}$  does not involve simplex  $[a'_{0:k-1,j'}]$  and does not introduce any new simplices  $[a'_{0:k-1,j'}]$  for any tuples  $a'_{0:k-1}$  compared to  $\Phi_\tau$ . There are only finite number of simplices  $[a_{0:k-1,j'}]$  and therefore after finite number of iterated constructions,  $\Phi_{t+1} = \Phi_0 - \sum_{\tau=0}^t \Psi_\tau$  will not involve simplices  $[a_{0:k-1,j'}]$  for any tuples  $a_{0:k-1}$ . We now argue that  $\Phi_{t+1}$  also does not involve simplices  $[a_{0:k-2,j,j'}]$  for any tuples  $a_{0:k-2}$ .  $\Phi_{t+1} \in Z_k(L_\alpha(A_{X,C_{i+1}}^K))$  implies  $\partial_k \Phi_{t+1} = 0$ . If  $\Phi_{t+1}$  involves simplex  $[a_{0:k-2,j,j'}]$  for some tuple  $a_{0:k-2}$ ,  $\partial_k[a_{0:k-2,j,j'}]$  involves the simplex  $[a_{0:k-2,j'}]$ . To cancel out this simplex, there is at least one simplices  $[a_{0:k-1,j'}]$  in  $\Phi_{t+1}$  for some tuples  $a_{0:k-1}$ . This contradicts the fact that  $\Phi_{t+1}$  does not involve simplices  $[a_{0:k-1,j'}]$  for any tuples  $a_{0:k-1}$ . Hence,  $\Phi_{t+1}$  does not involve simplices  $[a_{0:k-1,j'}]$  nor  $[a_{0:k-2,j,j'}]$  for any tuples  $a_{0:k-1}$ . This implies that it is a  $k$ -cycle in the simplicial complex induced from  $A_{X,C_i}^K$ , i.e.  $\Phi_{t+1} \in Z_k(L_\alpha(A_{X,C_i}^K))$ . Meanwhile,  $\Psi_\tau \in \mathbf{B}_k(L_\alpha(A_{X,C_{i+1}}^K))$  for any  $0 \leq \tau \leq t$  implies that  $\sum_{\tau=0}^t \Psi_\tau \in \mathbf{B}_k(L_\alpha(A_{X,C_{i+1}}^K))$ . Since

$$\Phi_0 - \Phi_{t+1} = \sum_{\tau=0}^t \Psi_\tau, \quad (34)$$

any new cycle  $\Phi_0$  in  $L_\alpha(A_{X,C_{i+1}}^K)$  generated by chains involving  $[a_{0:k-1,j'}]$  or  $[a_{0:k-2,j,j'}]$  are dependent of some cycle  $\Phi_{t+1}$  in  $L_\alpha(A_{X,C_i}^K)$  via boundaries  $\sum_{\tau=0}^t \Psi_\tau$ . Consequently,  $k$ -chains involving simplices  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$  will not create any new non-trivial homological features compared to  $H_k(L_\alpha(A_{X,C_i}^K))$ .

What remains to be shown is that  $k$ -chains involving simplices  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$  will not affect existing homological features  $H_{k-1}(L_\alpha(A_{X,C_i}^K))$ . Suppose  $\Phi_0$  is a  $k$ -chain which involves  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$  and trivializes a  $(k-1)$ -cycle  $\Phi'_0 \in Z_{k-1}(L_\alpha(A_{X,C_i}^K))$ . There are no simplices involving  $a_{j'}$  in the chain  $\partial_k \Phi_0 = \Phi'_0$ . We now prove that we can find another  $k$ -chain  $\Phi_{t+1}$  which does not involve  $[a_{0:k-1,j'}]$  nor  $[a_{k-2,j,j'}]$  such that  $\partial_k \Phi_{t+1} = \partial_k \Phi_0$  and every simplex of  $\Phi_{t+1}$  appears in  $L_\alpha(A_{X,C_i}^K)$ . Start with  $\tau = 0$  and conduct the following construction. If there is a simplex  $[a_{0:k-1,j'}]$  that is involved in  $\Phi_\tau$  for some tuple  $a_{0:k-1}$ , consider the simplex  $\Psi_\tau = \partial_{k+1}[a_{0:k-1,j,j'}]$ . Form a new  $k$ -chain  $\Phi_{\tau+1}$  as in (32). It has been shown that every simplex in  $\Psi_\tau$  appears in  $L_\alpha(A_{X,C_i}^K)$ . Therefore, every simplex in  $\Phi_{\tau+1}$  also appears in  $L_\alpha(A_{X,C_i}^K)$ . Moreover, from Lemma 3 we know that  $\partial_{k+1} \Psi_\tau = \partial_{k+1} \partial_k[a_{0:k-1,j,j'}] = 0$ , and consequently,

$$\partial_k \Phi_{\tau+1} = \partial_k (\Phi_\tau - \hat{\beta}_\tau \Psi_\tau) = \partial_k \Phi_\tau - \hat{\beta}_\tau \partial_{k+1} \partial_k[a_{0:k-1,j,j'}] = \partial_k \Phi_\tau. \quad (35)$$

It has also been shown that  $\Phi_{\tau+1}$  cannot have any new simplices  $[a'_{0:k-1,j'}]$  compared to  $\Phi_\tau$  with  $a'_{0:k-1} \neq a_{0:k-1}$ . There are only finite number of simplices and therefore after finite number of iterated constructions,  $\Phi_{t+1} = \Phi_0 - \sum_{\tau=0}^t \Psi_\tau$  will not involve simplices  $[a_{0:k-1,j'}]$  for any tuples  $a_{0:k-1}$ . We now argue that  $\Phi_{t+1}$  also does not involve simplices  $[a_{0:k-2,j,j'}]$  for any tuples  $a_{0:k-2}$ . There are no simplices involving node  $a_{j'}$  in the chain  $\partial_k \Phi_{t+1} = \partial_k \Phi_0 = \Phi'_0$ . If  $\Phi_{t+1}$  involves simplex  $[a_{0:k-2,j,j'}]$  for some tuple  $a_{0:k-2}$ ,  $\partial_k[a_{0:k-2,j,j'}]$  involves the simplex  $[a_{0:k-2,j'}]$ . To cancel out this simplex, there is at least one simplices  $[a_{0:k-1,j'}]$  in  $\Phi_{t+1}$  for some tuples  $a_{0:k-1}$ . This contradicts the fact that  $\Phi_{t+1}$  will not involve simplices  $[a_{0:k-1,j'}]$  for any tuples  $a_{0:k-1}$ . Hence,  $\Phi_{t+1}$  does not involve simplices  $[a_{0:k-1,j'}]$  nor  $[a_{0:k-2,j,j'}]$  for any tuples  $a_{0:k-1}$ . This implies that it is a  $k$ -chain in the simplicial complex induced from

$A_{X,C_i}^K$ , i.e.  $\Phi_{t+1} \in \mathbf{C}_k(L_\alpha(A_{X,C_i}^K))$ . This demonstrates that  $\Phi'_0$  is also trivialized by a chain in  $L_\alpha(A_{X,C_i}^K)$  and therefore  $k$ -chains involving simplices  $[a_{0:k-1,j'}]$  or  $[a_{k-2,j,j'}]$  will not affect existing homological features  $\mathbf{H}_{k-1}(L_\alpha(A_{X,C_i}^K))$ .

Since  $\mathcal{L}(A_{X,C_i}^K)$  and  $\mathcal{L}(A_{X,C_{i+1}}^K)$  have identical homological features for any  $0 \leq i \leq |C| - |X| - 1$  for the sequence of augmented networks in (28) and  $A_{X,C_0}^K \cong D_X^K$ ,  $\mathcal{L}(A_{X,C_0}^K)$  and  $\mathcal{L}(D_X^K)$  have identical persistence intervals. ■

Having demonstrated all four statements in Proposition 3, the global proof is completed. ■

Finally, we use Lemma 1 and Proposition 3 to show Theorem 1.

**Proof of Theorem 1:** Let  $\delta = d_{\mathcal{D},\infty}(D_X^K, D_Y^K)$ . From Definitions 5 and 8, there exists a correspondence  $C$  between  $X$  and  $Y$  such that  $|r_X^k(x_{0:k}) - r_Y^k(y_{0:k})| \leq \delta$  for any  $k$  and any pairs of correspondents  $(x_{0:k}, y_{0:k}) = (x_0, y_0), \dots, (x_k, y_k) \in C$ . Construct the pair of augmented networks  $A_X^K$  and  $A_Y^K$  defined in Definition 12 using this correspondence. It follows from (16) that given a tuple  $a_{0:k} \in C$  with each node  $a_i$  representing the correspondent pair  $(x_{C_i}, y_{C_i})$  in  $C$ ,

$$|r_{A_X}^k(a_{0:k}) - r_{A_Y}^k(a_{0:k})| = |r_X^k(x_{C_0:C_k}) - r_Y^k(y_{C_0:C_k})| \leq \delta. \quad (36)$$

Since  $A_X^K$  and  $A_Y^K$  have identical nodes, (36) implies that any simplices that appear at time  $\alpha$  in the induced filtration  $\mathcal{L}(A_X^K)$  [cf. Definition 13] will appear no earlier than  $\alpha - \delta$  and no later than  $\alpha + \delta$  in the induced filtration  $\mathcal{L}(A_Y^K)$ . Similar properties hold for any simplices that appear at time  $\alpha$  in the induced filtration  $\mathcal{L}(A_Y^K)$ . This shows that filtrations  $\mathcal{L}(A_X^K)$  and  $\mathcal{L}(A_Y^K)$  are strongly  $\delta$  interleaved [cf. (13)]. By Lemma 1, we know that for any  $k$ , the bottleneck distance between the  $k$ -th dimensional persistence diagrams of  $\mathcal{L}(A_X^K)$  and  $\mathcal{L}(A_Y^K)$  is bounded by  $d_{\mathcal{D},\infty}(D_X^K, D_Y^K)$ ,

$$d_{\mathbb{B}}^\infty(\mathcal{L}(A_X^K), \mathcal{L}(A_Y^K)) \leq d_{\mathcal{D},\infty}(D_X^K, D_Y^K). \quad (37)$$

Besides, Proposition 3 guarantees that the persistence diagram of  $\mathcal{L}(A_X^K)$  is the same as that of  $\mathcal{L}(D_X^K)$  and the persistence diagram of  $\mathcal{L}(A_Y^K)$  is the same as that of  $\mathcal{L}(D_Y^K)$ . Therefore,  $d_{\mathbb{B}}^\infty(\mathcal{L}(A_X^K), \mathcal{L}(A_Y^K)) = d_{\mathbb{B}}^\infty(\mathcal{L}(D_X^K), \mathcal{L}(D_Y^K))$ . Substituting this into (37) concludes the proof. ■

## APPENDIX C

### OTHER PROOFS IN SECTION IV

**Proof of Proposition 4:** In order to show that any  $k$ -order dissimilarity of a dissimilarity network  $D_X^K$  can be found in persistence homologies, observe that given any full rank  $(k+1)$ -tuples  $x_{0:k}$ , (24) indicates that the  $k$ -simplex  $\phi^k$  defined by the convex hull  $\text{conv}\{x_{0:k}\}$  appears strictly after any of its faces  $\text{conv}\{x_{0:\widehat{s}:k}\}$  in the filtration. Suppose  $\phi^k$  appears at time  $\alpha$  and denote  $\partial_k \phi^k = \sum_i \beta_i \psi_i^{k-1}$  with  $\beta_i$  the coefficients, then each  $\psi_i^{k-1}$  appears strictly before time  $\alpha$ .

Now suppose that the appearance of  $\phi^k$  trivializes a  $(k-1)$ -th dimensional homological feature. This means that  $\phi^k$  is the boundary to trivialize the  $(k-1)$ -th dimensional cycle  $\partial_k \phi^k$ . Since each face  $\psi_i^{k-1}$  of  $\phi^k$  appears strictly before time  $\alpha$ , the cycle  $\partial_k \phi^k$  results in a persistence interval. The death time of this persistence interval is  $\alpha$ , or equivalently, the time represented by the relationship  $r_X^k(x_{0:k})$ .

On the other hand, if the appearance of  $\phi^k$  does not trivialize a  $(k-1)$ -th dimensional homological feature, then the  $(k-1)$ -cycle  $\partial_k \phi^k$  is also in the boundary group  $\mathbf{B}_{k-1}(L_\alpha(D_X^K))$  of the collection of simplices appearing before or on time  $\alpha$ . This means that  $\partial_k \phi^k$  can be represented by a formal sum of the boundaries of some  $k$ -chains  $\Phi_i^k$ ,

$$\partial_k \phi^k = \sum_i \beta_i \partial_k \Phi_i^k \quad (38)$$

with coefficients  $\beta_i$  and  $k$ -chains  $\Phi_i^k$  appearing before or on time  $\alpha$ . By the definition of  $k$ -chains,  $\Phi_i^k = \sum_j \beta_j' \psi_j^k$  with coefficients  $\beta_j'$  and  $k$ -simplices  $\psi_j^k$  appearing before or on time  $\alpha$ . Therefore, (38) can be written as  $\partial_k \phi^k = \sum_j \beta_j' \partial_k \psi_j^k$ . Rearranging terms yields

$$\partial_k \left( \sum_j \beta_j' \psi_j^k - \phi^k \right) = 0. \quad (39)$$

This implies that  $\sum_i \beta_i' \psi_i^k - \phi^k$  is a  $k$ -cycle. There must be a new cycle formed since  $\phi^k$  just appears. The cycle cannot be trivialized immediately since any  $(k+1)$ -chain  $\Psi^{k+1}$  with  $\partial_{k+1} \Psi^{k+1} = \sum_i \beta_i' \psi_i^k - \phi^k$  would involve a simplex  $[x_{0:k,l}]$  for some node  $x_l$  with full rank tuple  $x_{0:k,l}$  where this simplex  $[x_{0:k,l}]$  appears strictly after  $\alpha$ . Therefore we have a  $k$ -th dimensional persistence interval with birth time  $\alpha$ , or equivalently, the time denoted by the relationship  $r_X^k(x_{0:k})$ . This concludes the proof. ■

**Proof of Proposition 5:** Given any  $0 \leq k' \leq k \leq K$ , it follows from (2) that for any correspondence  $C$  between  $X$  and  $Y$ ,

$$\Gamma_{X,Y}^k(C) = \max_{(x_{0:k}, y_{0:k}) \in C} |r_X^k(x_{0:k}) - r_Y^k(y_{0:k})|, \quad (40)$$

$$\Gamma_{X,Y}^{k'}(C) = \max_{(x_{0:k'}, y_{0:k'}) \in C} |r_X^{k'}(x_{0:k'}) - r_Y^{k'}(y_{0:k'})|. \quad (41)$$

For the  $(x_{0:k'}, y_{0:k'}) \in C$  achieving the maximum difference  $|r_X^{k'}(x_{0:k'}) - r_Y^{k'}(y_{0:k'})|$  in  $\Gamma_{X,Y}^{k'}(C)$ , we can construct another correspondent pair  $(x_{0:k}^*, y_{0:k}^*)$  such that  $x_{0:k}^*$  is a sub-tuple of  $x_{0:k}$  with same rank  $s(x_{0:k}^*) = s(x_{0:k})$  and  $y_{0:k}^*$  is a sub-tuple of  $y_{0:k}$  with  $s(y_{0:k}^*) = s(y_{0:k})$ . It follows from the identity property of high order networks that  $r_X^k(x_{0:k}^*) = r_X^k(x_{0:k})$  and  $r_Y^k(y_{0:k}^*) = r_Y^k(y_{0:k})$ . This implies that taking the maximum  $|r_X^k(x_{0:k}) - r_Y^k(y_{0:k})|$  over  $(x_{0:k}, y_{0:k}) \in C$  cannot yield a lower difference than  $|r_X^{k'}(x_{0:k}^*) - r_Y^{k'}(y_{0:k}^*)|$ , i.e.

$$\Gamma_{X,Y}^k(C) \geq |r_X^{k'}(x_{0:k}^*) - r_Y^{k'}(y_{0:k}^*)| = \Gamma_{X,Y}^{k'}(C). \quad (42)$$

Since (42) holds true for any correspondence  $C \in \mathcal{C}(X, Y)$ , the inequality must hold true when we take the minimum over all correspondences  $\mathcal{C}(X, Y)$  between  $X$  and  $Y$ ,

$$\min_{C \in \mathcal{C}(X, Y)} \Gamma_{X,Y}^k(C) \geq \min_{C' \in \mathcal{C}(X, Y)} \Gamma_{X,Y}^{k'}(C'). \quad (43)$$

Substituting the definition of  $k$ -order and  $k'$ -order network distances into (43) yields  $d_{\mathbb{D}}^k \geq d_{\mathbb{D}}^{k'}$  and concludes the proof. ■

**Proof of Corollary 2:** It follows from (4) that for any correspondence  $C$  between the node spaces  $X$  and  $Y$ , the difference between  $N_X^K$  and  $N_Y^K$  measured by  $C$  is

$$\|\Gamma_{X,Y}^K(C)\|_\infty = \max_{k=0,1,\dots,K} \{\Gamma_{X,Y}^k(C)\}. \quad (44)$$

From (42) we know that for any  $0 \leq k' \leq k \leq K$ ,  $\Gamma_{X,Y}^k(C) \geq \Gamma_{X,Y}^{k'}(C)$ , and therefore for any correspondence  $C$ ,



$\max_{k=0}^K \{\Gamma_{X,Y}^k(C)\} = \Gamma_{X,Y}^K(C)$ . Substituting this into (44) and taking a minimum over all correspondences  $\mathcal{C}(X, Y)$  concludes the proof. ■

Besides Theorem 1 and Corollary 2, the proof of Theorem 2 requires introducing the definition of truncated networks as we state next.

**Definition 14** Given  $K$ -order network  $N_X^K = (X, r_X^0, r_X^1, \dots, r_X^K)$ , its  $k$ -order truncated network  $N_X^k$  is defined as  $N_X^k = (X, r_X^0, r_X^1, \dots, r_X^k)$ .

A  $k$ -order truncated network  $N_X^k$  has the same node space as its parent network  $N_X^K$  and collects the lowest  $k+1$  order relationship functions of  $N_X^K$ . It follows immediately that  $N_X^k$  is a  $k$ -order network. We now use Theorem 1, Corollary 2, and Definition 14 to prove Theorem 2.

**Proof of Theorem 2:** Construct the  $k$ -order truncated networks  $D_X^k$  and  $D_Y^k$  from  $D_X^K$  and  $D_Y^K$ . It follows directly from Theorem 1 that for any  $0 \leq k' \leq k$ , the bottleneck distance of the persistence diagrams of  $\mathcal{P}_{k'}\mathcal{L}(D_X^k)$  and  $\mathcal{P}_{k'}\mathcal{L}(D_Y^k)$  is a lower bound of their  $\infty$ -norm network distance,

$$d_B^\infty(\mathcal{P}_{k'}\mathcal{L}(D_X^k), \mathcal{P}_{k'}\mathcal{L}(D_Y^k)) \leq d_{D,\infty}(D_X^k, D_Y^k). \quad (45)$$

Since  $D_X^k$  and  $D_Y^k$  are valid high order networks, Corollary 2 implies that  $d_{D,\infty}(D_X^k, D_Y^k) = d_D^k(D_X^k, D_Y^k)$ . Meanwhile,  $d_D^k(D_X^k, D_Y^k) = d_D^k(D_X^K, D_Y^K)$  follows from the facts that  $D_X^k$  and  $D_X^K$  have identical  $k$ -order relationship functions and  $D_Y^k$  and  $D_Y^K$  have same  $k$ -order relationship. Finally, for any  $k' \leq k-1$ , the  $k'$ -th dimensional persistence diagram of  $\mathcal{L}(D_X^k)$  is identical to that of  $\mathcal{L}(D_X^K)$  and the  $k'$ -th dimensional persistence diagram of  $\mathcal{L}(D_Y^k)$  is identical to that of  $\mathcal{L}(D_Y^K)$ , therefore

$$d_B^\infty(\mathcal{P}_{k'}\mathcal{L}(D_X^k), \mathcal{P}_{k'}\mathcal{L}(D_Y^k)) = d_B^\infty(\mathcal{P}_{k'}\mathcal{L}(D_X^K), \mathcal{P}_{k'}\mathcal{L}(D_Y^K)). \quad (46)$$

Combining these observations, the proof concludes. ■

## REFERENCES

- [1] W. Huang and A. Ribeiro, "Comparing High Order Networks by Persistence Homology," in *Proc. Global Conf. Signal Info. Process.*, vol. (submitted), Orlando FA, Dec 14-16 2015.
- [2] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Inc., 1993.
- [3] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, ser. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.
- [4] R. Ghrist and A. Muhammad, "Coverage and hole-detection in sensor networks via homology," in *International Symposium on Information Processing in Sensor Networks*, vol. 00, no. 1, 2005, pp. 254–260.
- [5] H. Chintakunta and H. Krim, "Divide and Conquer: Localizing Coverage Holes in Sensor Networks," in *2010 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Jun. 2010, pp. 1–8.
- [6] B. Zhang, "Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory," *Computational Intelligence Magazine, IEEE*, vol. 3, no. 3, pp. 49–63, 2008.
- [7] W. Ren, Q. Zhao, R. Ramanathan, J. Gao, A. Swami, A. Bar-Noy, M. P. Johnson, and P. Basu, "Broadcasting in multi-radio multi-channel wireless networks using simplicial complexes," in *Wireless Networks*, vol. 19, no. 6, Nov. 2012, pp. 1121–1133.
- [8] J. Xu and V. Singh, "Unified Hypergraph for Image Ranking in a Multimodal Context," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 2333–2336.
- [9] Y. Gao, M. Wang, D. Tao, R. Ji, and Q. Dai, "3-D object retrieval and recognition with hypergraph analysis," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4290–303, Sep. 2012.
- [10] A. Wilkerson, T. Moore, A. Swami, and H. Krim, "Simplifying the Homology Of Networks via Strong Collapses," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 5258–5262.
- [11] T. Wang and H. Krim, "Statistical classification of social networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 3977–3980.
- [12] R. Singh, J. Xu, and B. Berger, "Global alignment of multiple protein interaction networks with application to functional orthology detection," *Proceedings of the National Academy of Sciences*, vol. 105, no. 35, pp. 12 763–12 768, 2008.
- [13] L. Peng, L. Liu, S. Chen, and Q. Sheng, "A network comparison algorithm for predicting the conservative interaction regions in protein-protein interaction network," in *2010 IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, Sep. 2010, pp. 34–39.
- [14] S. Choobdar, P. Ribeiro, S. Bugla, and F. Silva, "Comparison of Co-authorship Networks across Scientific Fields Using Motifs," *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 147–152, Aug. 2012.
- [15] L. Yong, Z. Yan, and C. Lei, "Protein-protein interaction network comparison based on wavelet and principal component analysis," in *2010 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, 2010, pp. 430–437.
- [16] N. Pržulj, "Biological network comparison using graphlet degree distribution," *Bioinformatics*, vol. 23, no. 2, pp. e177–183, Jan. 2007.
- [17] T. Milenković and N. Pržulj, "Uncovering biological network function via graphlet degree signatures," *Cancer informatics*, vol. 6, p. 257, Jan. 2008.
- [18] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*. Birkha user Boston Inc., Boston, MA., 2007.
- [19] F. Memoli, "Gromov-Hausdorff distances in Euclidean spaces," in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2008, pp. 1–8.
- [20] G. Carlsson, F. Memoli, A. Ribeiro, and S. Segarra, "Axiomatic construction of hierarchical clustering in asymmetric networks," 2014. [Online]. Available: <http://arxiv.org/abs/1301.7724>
- [21] W. Huang and A. Ribeiro, "Metrics in the Space of High Order Networks," *Signal Processing, IEEE Transactions on*, vol. (revised), 2015.
- [22] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological Persistence and Simplification," in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, 2000, pp. 454–463.
- [23] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete Comput. Geo.*, vol. 33, pp. 249–274, 2005.
- [24] K. Mischaikow and V. Nanda, "Morse theory for filtrations and efficient computation of persistent homology," *Discrete & Computational Geometry*, vol. 50, no. 2, pp. 330–353, 2013.
- [25] S. Harker, K. Mischaikow, M. Mrozek, and V. Nanda, "Discrete Morse Theoretic Algorithms for Computing Homology of Complexes and Maps," *Foundations of Computational Mathematics*, vol. 14, no. 1, pp. 151–184, 2014.
- [26] C. Berge, *Graphs and Hypergraphs*. North-Holland Publishing Company, 1976.
- [27] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*. American Mathematical Society Providence, 2001, vol. 33.
- [28] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot, "Proximity of Persistence Modules and Their Diagrams," in *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry*, 2009, pp. 237–246.
- [29] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment problems*. Society for Industrial and Applied Mathematics, 2009.
- [30] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, and S. Y. Oudot, "Gromov-Hausdorff Stable Signatures for Shapes using Persistence," *Eurographics Symposium on Geometry Processing*, vol. 28, no. 5, pp. 1393–1403, 2009.
- [31] E. Carlsson, G. Carlsson, and V. D. Silva, "An algebraic topological method for feature identification," *Intl. J. Computational Geometry and Applications*, vol. 16, pp. 291–314, 2003.
- [32] "Engineering Village: the place to find answers to engineering questions." [Online]. Available: <http://www.engineeringvillage.com/search/quick.url>
- [33] A. Tausz, M. Vejdemo-Johansson, and H. Adams, "JavaPlex: A research software package for persistent (co)homology," in *Proceedings of ICMS 2014*, 2014, pp. 129–136.
- [34] M. A. A. Cox and T. F. Cox, "Multidimensional scaling," in *Handbook of Data Visualization*, ser. Springer Handbooks Comp.Statistics. Springer Berlin Heidelberg, 2008, pp. 315–347.
- [35] H. Edelsbrunner and J. L. Harer, *Computational Topology: An Introduction*. American Mathematical Soc., 2010.