DISTANCES AND ISOMORPHISM BETWEEN NETWORKS AND THE STABILITY OF NETWORK INVARIANTS

SAMIR CHOWDHURY AND FACUNDO MÉMOLI

ABSTRACT. We develop the theoretical foundations of a network distance that has recently been applied to various subfields of topological data analysis, namely persistent homology and hierarchical clustering. While this network distance has previously appeared in the context of finite networks, we extend the setting to that of compact networks. The main challenge in this new setting is the lack of an easy notion of sampling from compact networks; we solve this problem in the process of obtaining our results. The generality of our setting means that we automatically establish results for exotic objects such as directed metric spaces and Finsler manifolds. We identify readily computable network invariants and establish their quantitative stability under this network distance. We also discuss the computational complexity involved in precisely computing this distance, and develop easily-computable lower bounds by using the identified invariants. By constructing a wide range of explicit examples, we show that these lower bounds are effective in distinguishing between networks. Finally, we provide a simple algorithm that computes a lower bound on the distance between two networks in polynomial time and illustrate our metric and invariant constructions on a database of random networks and a database of simulated hippocampal networks.

CONTENTS

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, 100 MATH TOWER, 231 WEST 18TH AVENUE, COLUMBUS, OH 43210. PHONE: (614) 292-4975, FAX: (614) 292-1479

E-mail address: chowdhury.57@osu.edu, memoli@math.osu.edu.

Date: April 20, 2018.

2010 *Mathematics Subject Classification.* 54EXX.

Key words and phrases. Networks, motifs, isomorphism, networks invariants, distance between networks.

1. Introduction

Networks which show the relationships within and between complex systems are key tools in a variety of current research areas. In the domain of bioinformatics, networks have been used to represent molecular activity [83], metabolic pathways [77], functional relations between enzyme clusters [69] and genetic regulation [85, 64]. Networks have been used as a natural tool for representing brain anatomy and function [90, 91, 92, 80, 73]. Network-based methods have also appeared in data mining [98], where the goal is to extract patterns or substructures that appear with higher frequency than in a randomized network [21, 46, 102]. Other examples of networks include social networks [54, 27], information networks such as the World Wide Web [31, 68, 51], and technological networks such as the electric power grid [100, 56]. For a more comprehensive list of the literature on complex networks, consult [66].

A recent development in network analysis is the application of methods from topological data analysis, e.g. persistent homology, to network data. Although this connection is being developed rapidly [86, 87, 75, 76, 38, 37, 25, 59, 17], there is much that remains to be explored. It is important to point out that some authors have treated networks as metric spaces [55, 49] or symmetric objects such as weighted graphs [13, 38, 76], both of which are well-understood mathematical objects. In contrast, the authors of [17, 16] treated a network as a generalization of a metric space, i.e. a finite set X equipped with an arbitrary function $\omega_X : X \times X \to \mathbb{R}$, and provided a formal treatment of persistent homology in this general setting. A key theoretical contribution of [17] was a set of proofs showing *stability* of persistent homology methods applied to networks, which is a necessary step in making such methods amenable to data analysis. The essential ingredient in the discussion regarding stability was a definition of a *network distance* $d_{\mathcal{N}}$, which in turn had appeared in earlier applications of topological data analysis methods (namely hierarchical clustering) to networks [12, 11], and more recently in [88] and [19] (in proving stability for an alternative persistent homology method for network data).

From a theoretical perspective, there still exists a gap in our understanding of this network distance $d_{\mathcal{N}}$, which is structurally based on the Gromov-Hausdorff distance [41, 42] between metric spaces. Beyond its origins in metric geometry [7], the Gromov-Hausdorff distance between metric spaces has found applications in the context of shape and data analysis [62, 60, 61, 10, 15]. Thus one expects that an appropriate modification of the Gromov-Hausdorff distance would be a valuable tool in network analysis. The literature on $d_{\mathcal{N}}$ listed above suggests that this is indeed the case.

However, while $d_{\mathcal{N}}$ has found successful theoretical applications, there are important foundational questions about $d_{\mathcal{N}}$ that need to be answered, such as the following:

Question 1. *What is the "continuous limit" of a finite (discrete) network, and how does one extend $d_{\mathcal{N}}$ to this setting?*

In this paper we provide an answer to the question above. We view a network as a first countable topological space X equipped with a continuous weight function $\omega_X : X \times X \to \mathbb{R}$. This definition immediately subsumes both metric spaces and the finite networks (understood to have the discrete topology) described earlier. We state most of our results for *compact networks*, which are networks satisfying the additional constraint that the underlying space is compact.

Our rationale for using the term *network* to refer to the object described above arises from the finite case: a n-point network, for $n \in \mathbb{N}$, is simply an $n \times n$ matrix of real numbers, which can be viewed as the adjacency matrix of a directed, weighted graph with self-loops. Directed, weighted graphs are of central interest in the graph theory literature, and our treatment of $d_{\mathcal{N}}$ should be viewed as a non-combinatorial approach towards producing a similarity measure on such objects.

Proceeding further with the connection to graph theory, we consider the connection between the standard notion of graph isomorphism, which we call *strong isomorphism*, and the notion of being at $d_{\mathcal{N}}$-distance zero. Interestingly enough, for two networks (X, ω_X) and (Y, ω_Y), strong isomorphism implies $d_{\mathcal{N}}(X, Y) = 0$, but the reverse implication is not necessarily true. This leads to the following question:

Question 2. *What is the "correct" notion of network isomorphism, in relation to $d_{\mathcal{N}}$?*

A final question of practical importance is the following:

Question 3. *How does one approximate $d_{\mathcal{N}}$, given that the Gromov-Hausdorff distance is known* [61] *to be an NP-hard computation?*

1.1. **Contributions and organization of the paper.** We begin in §2 with our model for networks: finite, compact, and general. We do not impose any condition on the edge weights involved, apart from requiring that they be real numbers. In particular, we do not assume symmetry or the triangle inequality. In this section, we introduce the notion of *weak isomorphism* as an answer to Question (2) posed above. We also provide a simple interpretation of weak isomorphism in the setting of finite networks, and an explanation for why this simple interpretation is highly non-trivial in the compact case.

In §2 we also define the network distance $d_{\mathcal{N}}$, as well as a modified distance that we denote by $\widehat{d}_{\mathcal{N}}$. We develop intuition for $d_{\mathcal{N}}$ through numerous computations of $d_{\mathcal{N}}$ for small networks. The details of these computations suggest that computing $d_{\mathcal{N}}$ is, in general, a difficult problem. This leads to the discussion of *network invariants* in 4, which are demonstrated to provide *lower bounds* on the $d_{\mathcal{N}}$-distance between two networks and thus constitute one part of our answer to Question (3). We complete our answer to this question in §5, where we show that in general, $d_{\mathcal{N}}$ is NP-hard to compute, and so a better approach is to use the lower bounds as proxies for computing the actual network distance. Moreover, we provide algorithmic details about the computation of one of our strongest lower bounds and exemplify its use for classifying networks in a database.

In §3, we turn our attention to Question (1). We present compact networks as the continuous analogues of finite networks, obtained in the limit from a process of taking finer and finer samples of finite networks. We then give a precise characterization of weak isomorphism in the setting of compact networks.

After answering some of the foundational questions about $d_{\mathcal{N}}$ that we have raised above, we devote §6 to demonstrating some surprising connections between our notion of $d_{\mathcal{N}}$ and the notion of cut distance that is available in the graph theory literature [57]. Roughly speaking, the cut distance is an ℓ^1 notion, whereas $d_{\mathcal{N}}$ is an ℓ^∞ notion. We show that under the appropriate interpretation in the realm of compact metric spaces, these distances are equivalent.

We conclude in §7 with a discussion of the paper. Proofs not contained in the main text are relegated to §A.

To aid researchers who may be interested in using our methods, we have released a Matlab software package called `PersNet` to accompany this paper. This package may be downloaded from `https://github.com/fmemoli/PersNet`.

1.2. **Related literature.** Finding a suitable metric for network similarity is a central aim in network analysis. Classical approaches in this direction involve the edit distance and the maximum/minimum common sub/supergraph problem [14, 47, 39, 40]. Similarity measures on graphs with the same nodes have been studied in [53], and within the framework of shuffled graph classification in [96, 97]. Metrics generated by random walks and graph kernels are discussed in [52, 89, 95]. More recent approaches using graph edit distance are discussed in [36, 9, 30, 103, 58]. Spectral methods for graph matching are described in [45, 44]. A method using probability densities to approximate networks embedded in hyperbolic space is described in [1]. The cut metric has been used to study convergent sequences and limits of graphs by Lovász, Borgs, Chayes and collaborators [57, 4], with extensions developed by Diaconis and Janson [24]. In particular, the cut metric bears the most similarity to the network distance that we propose, although these notions of distance have independent roots in the mathematical literature. The main object of interest in the cut metric literature is a *graphon*, which is the limiting object of a convergent sequence of graphs. Graphons are useful in applications because they form a family of very general random graph models—see [29] for an application to hierarchical clustering.

Whereas the networks we consider are very general, many of the papers mentioned above make one or more of the simplifying assumptions that the networks involved have the same number of nodes, are undirected, or are unweighted. For example, in the case of undirected networks, a popular approach is to obtain eigenvectors and eigenvalues of the graph Laplacian via the spectral theorem [20]. However, because the spectral theorem does not apply for non-symmetric matrices, studying eigenvalues or eigenvectors may not be the best approach for directed networks. For applications, assuming that a network is undirected or unweighted results in a loss of structure when the original data set contains asymmetric relations that may be interpreted as directed edge weights. This is often the case when studying biological networks [26, 101, 79].

1.3. **Notation and basic terminology.** We will denote the cardinality of any set S by $\operatorname{card}(S)$. For any set S we denote by $F(S)$ the collection of all finite subsets of S. For a topological space X, we write $\mathcal{C}(X)$ to denote the closed subsets of X. A recurring notion is that of the Hausdorff distance between two subsets of a metric space. For a given metric space (Z, d_Z), the Hausdorff distance between two nonempty subsets $A, B \subseteq Z$ is given by:

$$d_{\mathrm{H}}^Z(A,B) = \max\left\{\sup_{a\in A}\inf_{b\in B} d_Z(a,b), \sup_{b\in B}\inf_{a\in A} d_Z(a,b)\right\}.$$

In particular, we have $d_{\mathrm{H}}^Z(A,B)=0$ if and only if $\overline{A}=\overline{B}$ [7, Proposition 7.3.3]. We will denote the non-negative reals by $\mathbb{R}_+$. The all-ones matrix of size $n\times n$ will be denoted $\mathbb{1}_{n\times n}$. Given a function $f : X \to Y$ between two sets X and Y, the image of f will be denoted $\operatorname{im}(f)$ or $f(X)$. We use square brackets to denote preimages, e.g. if $B \subseteq Y$, then $f^{-1}[B]$ denotes the preimage of B under f. Given a topological space X and a subset $A \subset X$, we will write $\overline{A}$ to denote the closure of A. The nonempty elements of the power set of a set S will be denoted $\mathcal{P}(S)$.

We recall some basic definitions. Let X be a topological space. Then $x \in X$ is a *limit point of a sequence* $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ if each open set containing x also contains all but finitely many terms of $(x_n)_{n\in\mathbb{N}}$. The sequence is then said to *converge* to x. A point $x \in X$ is a *limit point of a set* $A \subseteq X$ if every open set containing x contains a point of A distinct from x.

2. Networks, isomorphism, and network distances

For real-world applications, the object of interest is often the collection of all finite networks, which we denote by $\mathcal{FN}$. Formally, one writes:

$$\mathcal{FN} := \{(X, \omega_X) : X \text{ a finite set}, \omega_X : X \times X \to \mathbb{R} \text{ any map}\}.$$

However, in order to build a satisfactory theoretical foundation, one also needs to develop a formalism for infinite networks. Thus we proceed with the following definition.

Definition 1 (Networks)**.** Let X be a first countable topological space, and let ω_X be a continuous function from $X \times X$ (endowed with the product topology) to $\mathbb{R}$. By a *network*, we will mean a pair (X, ω_X). We will denote the collection of all networks by $\mathcal{N}$.

Notice in particular that $\mathcal{N}$ includes metric spaces (they are first countable, and the distance function is continuous) as well as spaces that are quasi-metric or directed (no symmetry), pseudometric (no nondegeneracy), semimetric (no triangle inequality), or all of the above. Recall that a space is first countable if each point in the space has a countable local basis (see [93, p. 7] for more details). First countability is a technical condition guaranteeing that when the underlying topological space of a network is compact, it is also sequentially compact.

Given a network (X, ω_X), we will refer to the points of X as *nodes* and ω_X as the *weight function* of X. Pairs of nodes will be referred to as *edges*. Given a nonempty subset $A \subset X$, we will refer to $(A, \omega_X|_{A\times A})$ as the *sub-network* of X induced by A. For notational convenience, we will often write $X \in \mathcal{N}$ to mean $(X, \omega_X) \in \mathcal{N}$.

Recall that any finite set X can be equipped with the discrete topology, and any map $\omega_X : X \times X \to \mathbb{R}$ is continuous with respect to the discrete topology. Thus the elements of $\mathcal{FN}$ trivially fit into the framework of $\mathcal{N}$. Throughout the paper, we will always understand finite networks to be equipped with the discrete topology.

While we are interested in $\mathcal{FN}$ for practical applications, a key ingredient of our theoretical framework is the collection of *compact networks*. We define these to be the networks (X, ω_X) satisfying the additional constraint that X is compact. The collection of compact networks is denoted $\mathcal{CN}$. Specifically, we write:

$$\mathcal{CN} := \{(X, \omega_X) : X \text{ compact, first countable topological space}, \omega_X : X \times X \to \mathbb{R} \text{ continuous}\}.$$

Compact networks are of special practical interest because they can be finitely approximated, in a manner that we will make precise in §3.1. Real world networks that are amenable to computational tasks are necessarily finite, so whenever possible, we will state our results for compact networks. Occasionally we will provide examples of noncompact networks to illustrate interesting theoretical points.

A natural question in understanding the structure of $\mathcal{N}$ would be: which elements of $\mathcal{N}$ are equivalent? A suitable answer to this question requires us to develop notions of *isomorphism* that

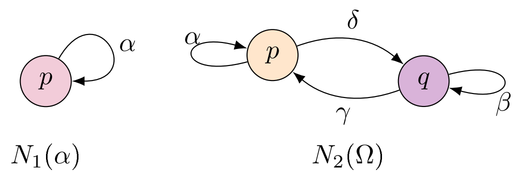

FIGURE 1. Networks over one and two nodes with their weight functions.

show various degrees of restrictiveness. These notions of isomorphism form a recurrent theme throughout this paper.

We first develop the notion of *strong isomorphism* of networks. The definition follows below.

Definition 2 (Weight preserving maps)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$. A map $\varphi : X \to Y$ is *weight preserving* if:

$$\omega_X(x, x') = \omega_Y(\varphi(x), \varphi(x')) \text{ for all } x, x' \in X.$$

Definition 3 (Strong isomorphism)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$. To say (X, ω_X) and (Y, ω_Y) are *strongly isomorphic* means that there exists a weight preserving bijection $\varphi : X \to Y$. We will denote a strong isomorphism between networks by $X \cong^s Y$. Note that this notion is exactly the usual notion of isomorphism between weighted graphs.

Strongly isomorphic networks formalize the idea that the information contained in a network should be preserved when we relabel the nodes in a compatible way.

Example 1. Networks with one or two nodes will be very instructive in providing examples and counterexamples, so we introduce them now with some special terminology.

- A network with one node p can be specified by $\alpha \in \mathbb{R}$, and we denote this by $N_1(\alpha)$. We have $N_1(\alpha) \cong^s N_1(\alpha')$ if and only if $\alpha = \alpha'$.
- A network with two nodes will be denoted by $N_2(\Omega)$, where $\Omega = \left(\begin{smallmatrix} \alpha & \delta \\ \gamma & \beta \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$. Given $\Omega, \Omega' \in \mathbb{R}^{2\times 2}$, $N_2(\Omega) \cong^s N_2(\Omega')$ if and only if there exists a permutation matrix P of size 2×2 such that $\Omega' = P\,\Omega\,P^T$.
- Any k-by-k matrix $\Sigma \in \mathbb{R}^{k\times k}$ induces a network on k nodes, which we refer to as $N_k(\Sigma)$. Notice that $N_k(\Sigma) \cong^s N_\ell(\Sigma')$ if and only if $k = \ell$ and there exists a permutation matrix P of size k such that $\Sigma' = P\,\Sigma\,P^T$.

Having defined a notion of isomorphism between networks, the next goal is to present the network distance $d_{\mathcal{N}}$ that is the central focus of this paper, and verify that $d_{\mathcal{N}}$ is compatible with strong isomorphism. We remind the reader that restricted formulations of this network distance have appeared in earlier applications of *hierarchical clustering* [12, 11] and *persistent homology* [16, 17, 19] methods to network data, and our overarching goal in this paper is to provide a theoretical foundation for this useful notion of network distance. In our presentation, we use a formulation of $d_{\mathcal{N}}$ that is more general than any other version available in the existing literature. As such, we proceed pedagogically and motivate the definition of $d_{\mathcal{N}}$ by tracing its roots in the metric space literature.

One strategy for defining a notion of distance between networks would be to take a well-understood notion of distance between metric spaces and extend it to all networks. The network distance $d_{\mathcal{N}}$ arises by following this strategy and extending the well-known Gromov-Hausdorff distance d_{GH} between compact metric spaces [41, 7, 74]. The definition of d_{GH} is rooted in the *Hausdorff distance* $d_{\text{H}}^{\bullet}$ between closed subsets of a metric space.

Definition 4. Given metric spaces (X, d_X) and (Y, d_Y), the *Gromov-Hausdorff distance* between them is defined as:

$$d_{\mathrm{GH}}((X, d_X), (Y, d_Y)) := \inf \big\{ d_{\mathrm{H}}^Z(\varphi(X), \psi(Y)) : Z \text{ a metric space}, \\ \varphi : X \to Z,\ \psi : Y \to Z \text{ isometric embeddings} \big\}.$$

The Gromov-Hausdorff distance dates back to at least the early 1980s [41], and it satisfies numerous desirable properties. It is a valid metric on the collection of isometry classes of compact metric spaces, is complete, admits many precompact families, and has well-understood notions of convergence [7, Chapter 7]. Moreover, it has found real-world applications in the *shape matching* [62, 63] and persistent homology literature [15], and its computational aspects have been studied as well [61]. As such, it is a strong candidate for use in defining a network distance.

Unfortunately, the formulation of d_{GH} above is heavily dependent on a metric space structure, and the notion of Hausdorff distance may not make sense in the setting of networks. So d_{GH} as defined above cannot be directly extended to a network distance. However, it turns out that there is a reformulation of d_{GH} that utilizes the language of *correspondences* [48, 7]. We present this construction next, and note that the resulting network distance $d_{\mathcal{N}}$ will agree with d_{GH} when restricted to metric spaces.

Definition 5 (Correspondence)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$. A *correspondence between X and Y* is a relation $R \subseteq X \times Y$ such that $\pi_X(R) = X$ and $\pi_Y(R) = Y$, where π_X and π_Y are the canonical projections of $X \times Y$ onto X and Y, respectively. The collection of all correspondences between X and Y will be denoted $\mathscr{R}(X, Y)$, abbreviated to $\mathscr{R}$ when the context is clear.

Example 2 (1-point correspondence)**.** Let X be a set, and let $\{p\}$ be the set with one point. Then there is a unique correspondence $R = \{(x, p) : x \in X\}$ between X and $\{p\}$.

Example 3 (Diagonal correspondence)**.** Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ be two enumerated sets with the same cardinality. A useful correspondence is the *diagonal correspondence*, defined as $\Delta := \{(x_i, y_i) : 1 \leqslant i \leqslant n\}$. When X and Y are infinite sets with the same cardinality, and $\varphi : X \to Y$ is a given bijection, then we write the diagonal correspondence as $\Delta := \{(x, \varphi(x)) : x \in X\}$.

Definition 6 (Distortion of a correspondence)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$ and let $R \in \mathscr{R}(X, Y)$. The *distortion* of R is given by:

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

Remark 4 (Composition of correspondences)**.** Let $(X, \omega_X),\ (Y, \omega_Y),\ (Z, \omega_Z) \in \mathcal{N}$, and let $R \in \mathscr{R}(X, Y),\ S \in \mathscr{R}(Y, Z)$. Then we define:

$$R \circ S := \{(x, z) \in X \times Z \mid \exists y, (x, y) \in R, (y, z) \in S\}.$$

In the proof of Theorem 12, we verify that $R \circ S \in \mathscr{R}(X, Z)$, and that $\operatorname{dis}(R \circ S) \leqslant \operatorname{dis}(R) + \operatorname{dis}(S)$.

Definition 7 (The first network distance)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{N}$. We define the *network distance* between X and Y as follows:

$$d_{\mathcal{N}}((X, \omega_X), (Y, \omega_Y)) := \tfrac{1}{2} \inf_{R \in \mathscr{R}} \operatorname{dis}(R).$$

When the context is clear, we will often write $d_{\mathcal{N}}(X, Y)$ to denote $d_{\mathcal{N}}((X, \omega_X), (Y, \omega_Y))$. We define the collection of *optimal correspondences* $\mathscr{R}^{\mathrm{opt}}$ between X and Y to be the collection

$\{R \in \mathscr{R}(X,Y) : \mathrm{dis}(R) = 2d_{\mathcal{N}}(X,Y)\}$. This set is always nonempty when $X, Y \in \mathcal{FN}$, but may be empty in general (see Example 9).

Remark 5. The intuition behind the preceding definition of network distance may be better understood by examining the case of a finite network. Given a finite set X and two edge weight functions ω_X, ω_X' defined on it, we can use the ℓ^∞ distance as a measure of network similarity between (X, ω_X) and (X, ω_X'):

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

A generalization of the ℓ^∞ distance is required when dealing with networks having different sizes: Given two sets X and Y, we need to decide how to match up points of X with points of Y. Any such matching will yield a subset $R \subseteq X \times Y$ such that $\pi_X(R) = X$ and $\pi_Y(R) = Y$, where π_X and π_Y are the projection maps from $X \times Y$ to X and Y, respectively. This is precisely a correspondence, as defined above. A valid notion of network similarity may then be obtained as the distortion incurred by choosing an optimal correspondence—this is precisely the idea behind the definition of the network distance above.

Remark 6. Some simple but important remarks are the following:

(1) When restricted to metric spaces, $d_{\mathcal{N}}$ agrees with d_{GH}. This can be seen from the reformulation of d_{GH} in terms of correspondences [7, Theorem 7.3.25], [48]. Whereas d_{GH} vanishes only on pairs of isometric compact metric spaces (which are Type I weakly isomorphic as networks), $d_{\mathcal{N}}$ vanishes on pairs of Type II weakly isomorphic networks. This comment will be elucidated in the proof of Theorem 12.

(2) Given $X, Y \in \mathcal{FN}$, the network distance reduces to the following:

$$d_{\mathcal{N}}(X,Y) = \frac{1}{2} \min_{R \in \mathscr{R}} \max_{(x,y),(x',y') \in R} |\omega_X(x,x') - \omega_Y(y,y')|.$$

Moreover, there is always at least one optimal correspondence R^{opt} for which $d_{\mathcal{N}}(X,Y)$ is achieved; this is a consequence of considering finite networks.

(3) For any $X, Y \in \mathcal{CN}$, we have $\mathscr{R}(X,Y) \neq \varnothing$, and $d_{\mathcal{N}}(X,Y)$ is always bounded. Indeed, $X \times Y$ is always a valid correspondence between X and Y. So we have:

$$d_{\mathcal{N}}(X,Y) \leqslant \frac{1}{2}\mathrm{dis}(X \times Y) \leqslant \frac{1}{2}\left(\sup_{x,x'} \left|\omega_X(x,x')\right| + \sup_{y,y' \in Y} \left|\omega_Y(y,y')\right|\right) < \infty.$$

Example 7. Now we give some examples to illustrate the preceding definitions.

- For $\alpha, \alpha' \in \mathbb{R}$ consider two networks with one node each: $N_1(\alpha) = (\{p\}, \alpha)$ and $N_1(\alpha') = (\{p'\}, \alpha')$. By Example 2 there is a unique correspondence $R = \{(p,p')\}$ between these two networks, so that $\mathrm{dis}(R) = |\alpha - \alpha'|$ and as a result $d_{\mathcal{N}}(N_1(\alpha), N_1(\alpha')) = \frac{1}{2}|\alpha - \alpha'|$.
- Let $(X, \omega_X) \in \mathcal{FN}$ be any network and for $\alpha \in \mathbb{R}$ let $N_1(\alpha) = (\{p\}, \alpha)$. Then $R = \{(x,p),\, x \in X\}$ is the unique correspondence between X and $\{p\}$, so that

$$d_{\mathcal{N}}(X, N_1(\alpha)) = \frac{1}{2} \max_{x,x' \in X} \left|\omega_X(x,x') - \alpha\right|.$$

We are now ready to make our first attempt at answering Question (2): we test whether $d_{\mathcal{N}}$ is compatible with strong isomorphism. Given two strongly isomorphic networks, i.e. networks $(X, \omega_X), (Y, \omega_Y)$ and a weight preserving bijection $\varphi : X \to Y$, it is easy to use the diagonal correspondence (Example 3) to verify that $d_{\mathcal{N}}(X,Y) = 0$. However, it is easy to see that the reverse implication is not true in general. Using the one-point correspondence (Example 2), one can see

Strong isomorphism: ϕ_X, ϕ_Y injective and surjective

$$\begin{tikzcd} & Z \arrow[dl, "\phi_X"'] \arrow[dr, "\phi_Y"] & \\ X & \cong^s & Y \end{tikzcd} \qquad \begin{tikzcd} & Z \arrow[dl, "\phi_X"'] \arrow[dr, "\phi_Y"] & \\ X & \cong^w_{\mathrm{I}} & Y \end{tikzcd}$$

Type I weak isomorphism: ϕ_X, ϕ_Y only surjective

FIGURE 2. Relaxing the requirements on the maps of this "tripod structure" is a natural way to weaken the notion of strong isomorphism.

that $d_{\mathcal{N}}(N_1(1), N_2(\mathbb{1}_{2\times 2})) = 0$. Here $\mathbb{1}_{n\times n}$ denotes the all-ones matrix of size $n \times n$ for any $n \in \mathbb{N}$. However, these two networks are not strongly isomorphic, because they do not even have the same cardinality. Thus to answer Question (2), we need to search for a different, perhaps weaker notion of isomorphism.

To proceed in this direction, first notice that a strong isomorphism between two networks (X, ω_X) and (Y, ω_Y), given by a bijection $f : X \to Y$, is equivalent to the following condition: there exists a set Z and bijective maps $\varphi_X : Z \to X, \varphi_Y : Z \to Y$ such that $\omega_X(\varphi_X(z), \varphi_X(z')) = \omega_Y(\varphi_Y(z), \varphi_Y(z'))$ for each $z, z' \in Z$. To see this, simply let $Z = \{(x, f(x)) : x \in X\}$ and let φ_X, φ_Y be the projection maps on the first and second coordinates, respectively. Based on this observation, we make the next definition.

Definition 8. Let (X, ω_X) and $(Y, \omega_Y) \in \mathcal{N}$. We define X and Y to be *Type I weakly isomorphic*, denoted $X \cong^w_{\mathrm{I}} Y$, if there exists a set Z and surjective maps $\varphi_X : Z \to X$ and $\varphi_Y : Z \to Y$ such that $\omega_X(\varphi_X(z), \varphi_X(z')) = \omega_Y(\varphi_Y(z), \varphi_Y(z'))$ for each $z, z' \in Z$.

Notice that Type I weak isomorphism is in fact a *relaxation* of the notion of strong isomorphism. Indeed, if in addition to being surjective, we require the maps ϕ_X and ϕ_Y to be injective, then the strong notion of isomorphism is recovered. In this case, the map $\phi_Y \circ \phi_X^{-1} : X \to Y$ would be a weight preserving bijection between the networks X and Y. The relaxation of strong isomorphism to a Type I weak isomorphism is illustrated in Figure 2. Also observe that the relaxation is *strict*. For example, the networks $X = N_1(1)$ and $Y = N_2(\mathbb{1}_{2\times 2})$, are weakly but not strongly isomorphic via the map that sends both nodes of Y to the single node of X.

We remark that when dealing with infinite networks, it will turn out that an even weaker notion of isomorphism is required. We define this weakening next.

Definition 9. Let (X, ω_X) and $(Y, \omega_Y) \in \mathcal{N}$. We define X and Y to be *Type II weakly isomorphic*, denoted $X \cong^w_{\mathrm{II}} Y$, if for each $\varepsilon > 0$, there exists a set Z_ε and surjective maps $\phi^\varepsilon_X : Z_\varepsilon \to X$ and $\phi^\varepsilon_Y : Z_\varepsilon \to Y$ such that

$$|\omega_X(\phi^\varepsilon_X(z), \phi^\varepsilon_X(z')) - \omega_Y(\phi^\varepsilon_Y(z), \phi^\varepsilon_Y(z'))| < \varepsilon \text{ for all } z, z' \in Z_\varepsilon. \tag{1}$$

Remark 8 (Type I isomorphism is stronger than Type II)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{CN}$ and suppose $\varphi : X \to Y$ is a surjective map such that $\omega_X(x, x') = \omega_Y(\varphi(x'), \varphi(x'))$ for all $x, x' \in X$. Then X and Y are Type I weakly isomorphic and hence Type II weakly isomorphic, i.e. $X \cong^w_{\mathrm{II}} Y$. This result follows from Definition 8 by: (1) choosing $Z = X$, and (2) letting ϕ_X be the identity map, and (3) letting $\phi_Y = \varphi$. The converse implication, i.e. that Type I weak isomorphism implies the existence of a surjective map as above, is not true: an example is shown in Figure 3.

Example 9 (Infinite networks without optimal correspondences)**.** The following example illustrates the reason we had to develop multiple notions of weak isomorphism. The key idea is that the infimum in Definition 7 is not necessarily obtained when X and Y are infinite networks. To see this, let (X, ω_X) denote $[0,1]$ equipped with the Euclidean distance, and let (Y, ω_Y) denote $\mathbb{Q} \cap [0,1]$ with the restriction of the Euclidean distance. Since the closure of Y in $[0,1]$ is just X, the Hausdorff distance between X and Y is zero (recall that given $A, B \subseteq \mathbb{R}$, we have $d_{\mathrm{H}}^{\mathbb{R}}(A,B) = 0$ if and only if $\overline{A} = \overline{B}$ [7, Proposition 7.3.3]). It follows from the definition of d_{GH} (Definition 4) and the equivalence of $d_{\mathcal{N}}$ and d_{GH} on metric spaces (Remark 6) that $d_{\mathcal{N}}(X,Y) = 0$.

However, one cannot define an optimal correspondence between X and Y. To see this, assume towards a contradiction that R^{opt} is such an optimal correspondence, i.e. $\mathrm{dis}(R^{\mathrm{opt}}) = 0$. For each $x \in X$, there exists $y_x \in Y$ such that $(x, y_x) \in R^{\mathrm{opt}}$. By making a choice of $y_x \in Y$ for each $x \in X$, define a map $f : X \to Y$ given by $x \mapsto y_x$. Then $d_X(x,x') = d_Y(f(x), f(x'))$ for each $x, x' \in X$. Thus f is an isometric embedding from X into itself (note that $Y \subseteq X$). But $X = [0,1]$ is compact, and an isometric embedding from a compact metric space into itself must be surjective [7, Theorem 1.6.14]. This is a contradiction, because $f(X) \subseteq Y \neq X$.

We observe that $d_{\mathcal{N}}(X,Y) = 0$ and so X and Y are weakly isomorphic of Type II, but not of Type I. To see this, assume towards a contradiction that X and Y are Type I weakly isomorphic. Let Z be a set with surjective maps $\varphi_X : Z \to X$ and $\varphi_Y : Z \to Y$ satisfying $\omega_X \circ (\varphi_X, \varphi_X) = \omega_Y \circ (\varphi_Y, \varphi_Y)$. Then $\{(\varphi_X(z), \varphi_Y(z)) : z \in Z\}$ is an optimal correspondence. This is a contradiction by the previous reasoning.

Recall that our motivation for introducing notions of isomorphism on $\mathcal{N}$ was to determine which networks deserve to be considered equivalent. It is easy to see that strong isomorphism induces an equivalence class on $\mathcal{N}$. The same is true for both types of weak isomorphism, and we record this result in the following proposition.

Proposition 10. *Weak isomorphism of Types I and II both induce equivalence relations on $\mathcal{N}$.*

In the setting of $\mathcal{FN}$, it is not difficult to show that the two types of weak isomorphism coincide. This is the content of the next proposition. By virtue of this result, there is no ambiguity in dropping the "Type I/II" modifier when saying that two finite networks are weakly isomorphic.

Proposition 11. *Let $X, Y \in \mathcal{FN}$ be finite networks. Then X and Y are Type I weakly isomorphic if and only if they are Type II weakly isomorphic.*

Type I weak isomorphisms will play a vital role in the content of this paper, but for now, we focus on Type II weak isomorphism. The next theorem justifies calling $d_{\mathcal{N}}$ a network *distance*, and shows that $d_{\mathcal{N}}$ is compatible with Type II weak isomorphism.

Theorem 12. *$d_{\mathcal{N}}$ is a metric on $\mathcal{N}$ modulo Type II weak isomorphism.*

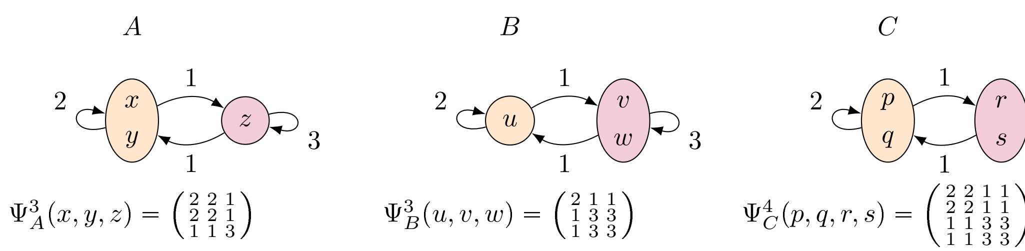

FIGURE 3. Note that Remark 8 *does not* fully characterize weak isomorphism, even for finite networks: All three networks above, with the given weight matrices, are Type I weakly isomorphic since C maps surjectively onto A and B. But there are no surjective, weight preserving maps $A \to B$ or $B \to A$.

The proof is in Appendix A. For finite networks, we immediately obtain:

The restriction of $d_{\mathcal{N}}$ to $\mathcal{FN}$ yields a metric modulo Type I weak isomorphism.

The proof of Proposition 11 will follow from the proof of Theorem 12. In fact, an even stronger result is true: weak isomorphism of Types I and II coincide for compact networks as well. We present the statement below, and dedicate Section 3.2 to its proof.

Theorem 13 (Weak isomorphism in $\mathcal{CN}$)**.** *Let $X, Y \in \mathcal{CN}$. Then X and Y are Type II weakly isomorphic if and only if X and Y are Type I weakly isomorphic, i.e. there exists a set V and surjections $\varphi_X : V \to X,\ \varphi_Y : V \to Y$ such that:*

$$\omega_X(\varphi_X(v), \varphi_X(v')) = \omega_Y(\varphi_Y(v), \varphi_Y(v')) \quad \textit{for all } v, v' \in V.$$

We end the current subsection with the following definition, which will be used heavily in §3.

Definition 10 (ε-approximations)**.** Let $\varepsilon > 0$. A network $(X, \omega_X) \in \mathcal{N}$ is said to be *ε-approximable* by $(Y, \omega_Y) \in \mathcal{N}$ if $d_{\mathcal{N}}(X, Y) < \varepsilon$. In this case, Y is said to be an *ε-approximation* of X. Typically, we will be interested in the case where X is infinite and Y is finite, i.e. in ε-approximating infinite networks by finite networks.

2.1. **The second network distance.** Even though the definition of $d_{\mathcal{N}}$ is very general, in some restricted settings it may be convenient to consider a network distance that is easier to formulate. For example, in computational purposes it suffices to assume that we are computing distances between finite networks. Also, a reduction in computational cost is obtained if we restrict ourselves to computing distortions of bijections instead of general correspondences. The next definition arises from such considerations.

Definition 11 (The second network distance)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{FN}$ be such that $\mathrm{card}(X) = \mathrm{card}(Y)$. Then define:

$$\widehat{d}_{\mathcal{N}}(X, Y) := \frac{1}{2} \inf_{\varphi} \sup_{x, x' \in X} \left|\omega_X(x, x') - \omega_Y\left(\varphi(x), \varphi(x')\right)\right|,$$

where $\varphi : X \to Y$ ranges over all bijections from X to Y.

Notice that $\widehat{d}_{\mathcal{N}}(X, Y) = 0$ if and only if $X \cong^s Y$. Also, $\widehat{d}_{\mathcal{N}}$ satisfies symmetry and triangle inequality. It turns out via Example 14 that $d_{\mathcal{N}}$ and $\widehat{d}_{\mathcal{N}}$ agree on networks over two nodes. However, the two notions do not agree in general. In particular, a minimal example where $d_{\mathcal{N}} \neq \widehat{d}_{\mathcal{N}}$ occurs for three node networks, as we show in Remark 15.

Example 14 (Networks with two nodes)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{FN}$ where $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$. Then we claim $d_{\mathcal{N}}(X, Y) = \widehat{d}_{\mathcal{N}}(X, Y)$. Furthermore, if $X = N_2\left(\left(\begin{smallmatrix} \alpha & \delta \\ \beta & \gamma \end{smallmatrix}\right)\right)$ and $Y = N_2\left(\left(\begin{smallmatrix} \alpha' & \delta' \\ \beta' & \gamma' \end{smallmatrix}\right)\right)$, then we have the explicit formula:

$$\begin{aligned} d_{\mathcal{N}}(X, Y) &= \frac{1}{2} \min\left(\Gamma_1, \Gamma_2\right), \text{ where} \\ \Gamma_1 &= \max\left(|\alpha - \alpha'|, |\beta - \beta'|, |\delta - \delta'|, |\gamma - \gamma'|\right), \\ \Gamma_2 &= \max\left(|\alpha - \gamma'|, |\gamma - \alpha'|, |\delta - \beta'|, |\beta - \delta'|\right). \end{aligned}$$

Details for this calculation are in §A.

Remark 15 (A three-node example where $d_{\mathcal{N}} \neq \widehat{d}_{\mathcal{N}}$)**.** Assume (X, ω_X) and (Y, ω_Y) are two networks with the same cardinality. Then

$$d_{\mathcal{N}}(X, Y) \leqslant \widehat{d}_{\mathcal{N}}(X, Y).$$

The inequality holds because each bijection induces a correspondence, and we are minimizing over all correspondences to obtain $d_{\mathcal{N}}$. However, the inequality may be strict, as demonstrated by the following example. Let $X = \{x_1, \ldots, x_3\}$ and let $Y = \{y_1, \ldots, y_3\}$. Define $\omega_X(x_1, x_1) = \omega_X(x_3, x_3) = \omega_X(x_1, x_3) = 1, \omega_X = 0$ elsewhere, and define $\omega_Y(y_3, y_3) = 1, \omega_Y = 0$ elsewhere. In terms of matrices, $X = N_3(\Sigma_X)$ and $Y = N_3(\Sigma_Y)$, where

$$\Sigma_X = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \Sigma_Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Define $\Gamma(x, x', y, y') = |\omega_X(x, x') - \omega_Y(y, y')|$ for $x, x' \in X$, $y, y' \in Y$. Let φ be any bijection. Then we have:

$$\begin{aligned} \max_{x,x' \in X} \Gamma(x, x', \varphi(x), \varphi(x')) &= \max\{\Gamma(x_1, x_3, \varphi(x_1), \varphi(x_3)), \Gamma(x_1, x_1, \varphi(x_1), \varphi(x_1)), \\ &\qquad \Gamma(x_3, x_3, \varphi(x_3), \varphi(x_3)), \Gamma(\varphi^{-1}(y_3), \varphi^{-1}(y_3), y_3, y_3)\} \\ &= 1. \end{aligned}$$

So $\widehat{d}_{\mathcal{N}}(X, Y) = \frac{1}{2}$. On the other hand, consider the correspondence

$$R = \{(x_1, y_3), (x_2, y_2), (x_3, y_3), (x_2, y_1)\}.$$

Then $\max_{(x,y),(x'y') \in R} |\omega_X(x, x') - \omega_Y(y, y')| = 0$. Thus $d_{\mathcal{N}}(X, Y) = 0 < \widehat{d}_{\mathcal{N}}(X, Y)$.

Example 16 (Networks with three nodes)**.** Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{FN}$, where we write $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2, y_3\}$. Because we do not necessarily have $d_{\mathcal{N}} = \widehat{d}_{\mathcal{N}}$ on three node networks by Remark 15, the computation of $d_{\mathcal{N}}$ becomes more difficult than in the two node case presented in Example 14. A certain reduction is still possible, which we present next. Consider the following list $\mathcal{L}$ of matrices representing correspondences, where a 1 in position (i, j) means that (x_i, y_j) belongs to the correspondence.

$$\begin{array}{c}
\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & & 1 \\ & 1 & \end{pmatrix} \begin{pmatrix} 1 & & \\ 1 & & \\ & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \\ & & 1 \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & 1 \\ 1 & & \end{pmatrix} \\ \hline
\begin{pmatrix} & 1 & \\ 1 & & \\ & & 1 \end{pmatrix} \begin{pmatrix} & 1 & \\ & & 1 \\ 1 & & \end{pmatrix} \begin{pmatrix} & 1 & \\ & 1 & \\ 1 & & 1 \end{pmatrix} \begin{pmatrix} & 1 & 1 \\ 1 & & \\ 1 & & \end{pmatrix} \begin{pmatrix} & 1 & \\ 1 & & 1 \\ & 1 & \end{pmatrix} \\ \hline
\begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix} \begin{pmatrix} & & 1 \\ 1 & & \\ & 1 & \end{pmatrix} \begin{pmatrix} & & 1 \\ & & 1 \\ 1 & 1 & \end{pmatrix} \begin{pmatrix} 1 & & 1 \\ & 1 & \\ & 1 & \end{pmatrix} \begin{pmatrix} & & 1 \\ 1 & 1 & \\ & & 1 \end{pmatrix}
\end{array}$$

Now let $R \in \mathscr{R}(X, Y)$ be any correspondence. Then R contains a correspondence $S \in \mathscr{R}(X, Y)$ such that the matrix form of S is listed in $\mathcal{L}$. Thus $\operatorname{dis}(R) \geqslant \operatorname{dis}(S)$, since we are maximizing over a larger set. It follows that $d_{\mathcal{N}}(X, Y)$ is obtained by taking $\arg\min \frac{1}{2} \operatorname{dis}(S)$ over all correspondences $S \in \mathscr{R}(X, Y)$ with matrix forms listed in $\mathcal{L}$.

For an example of this calculation, let S denote the correspondence $\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$ represented by the matrix $\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$. Then $\operatorname{dis}(S)$ is the maximum among the following:

$$\begin{array}{lll}
|\omega_X(x_1, x_1) - \omega_Y(y_1, y_1)| & |\omega_X(x_1, x_2) - \omega_Y(y_1, y_2)| & |\omega_X(x_1, x_3) - \omega_Y(y_1, y_3)| \\
|\omega_X(x_2, x_1) - \omega_Y(y_2, y_1)| & |\omega_X(x_2, x_2) - \omega_Y(y_2, y_2)| & |\omega_X(x_2, x_3) - \omega_Y(y_2, y_3)| \\
|\omega_X(x_3, x_1) - \omega_Y(y_3, y_1)| & |\omega_X(x_3, x_2) - \omega_Y(y_3, y_2)| & |\omega_X(x_3, x_3) - \omega_Y(y_3, y_3)|.
\end{array}$$

The following proposition provides an explicit connection between $d_{\mathcal{N}}$ and $\widehat{d}_{\mathcal{N}}$. An illustration is also provided in Figure 4.

Proposition 17. *Let* $(X, \omega_X), (Y, \omega_Y) \in \mathcal{FN}$. *Then,*

$$d_{\mathcal{N}}(X,Y) = \inf\left\{\widehat{d}_{\mathcal{N}}(X',Y') : X', Y' \in \mathcal{FN},\ X' \cong^w_{\mathrm{I}} X,\ Y' \cong^w_{\mathrm{I}} Y, \text{ and } \operatorname{card}(X') = \operatorname{card}(Y')\right\}.$$

Remark 18 (Computational aspects of $d_{\mathcal{N}}$ and $\widehat{d}_{\mathcal{N}}$)**.** Even though $\widehat{d}_{\mathcal{N}}$ has a simpler formulation than $d_{\mathcal{N}}$, computing $\widehat{d}_{\mathcal{N}}$ still turns out to be an NP-hard problem, as we discuss in §5. Moreover, we show in Theorem 47 that computing $d_{\mathcal{N}}$ is at least as hard as computing $\widehat{d}_{\mathcal{N}}$.

Instead of trying to compute $d_{\mathcal{N}}$, we will focus on finding network invariants that can be computed easily. This is the content of §4. For each of these invariants, we will prove a stability result to demonstrate its validity as a proxy for $d_{\mathcal{N}}$.

2.2. **Special families: dissimilarity networks and directed metric spaces.** The second network distance $\widehat{d}_{\mathcal{N}}$ that we introduced in the previous section turned out to be compatible with strong isomorphism. Interestingly, by narrowing down the domain of $d_{\mathcal{N}}$ to the setting of compact dissimilarity networks, we obtain a subfamily of $\mathcal{N}$ where $d_{\mathcal{N}}$ is compatible with strong isomorphism. A *dissimilarity* network is a network (X, A_X) where A_X is a map from $X \times X$ to $\mathbb{R}_+$, and $A_X(x,x') = 0$ if and only if $x = x'$. Neither symmetry nor triangle inequality is assumed. We denote the collection of all such networks as $\mathcal{FN}^{\mathrm{dis}}$, $\mathcal{CN}^{\mathrm{dis}}$, and $\mathcal{N}^{\mathrm{dis}}$ for the finite, compact, and general settings, respectively.

Theorem 19 ([11])**.** *The restriction of* $d_{\mathcal{N}}$ *to* $\mathcal{FN}^{\mathrm{dis}}$ *is a metric modulo strong isomorphism.*

The expression for $d_{\mathcal{N}}$ was used in the context of $\mathcal{FN}^{\mathrm{dis}}$ in [11, 12] to study the stability properties of hierarchical clustering methods on metric spaces and directed dissimilarity networks. That setting is considerably simpler than the situation in this paper, because in general we allow $d_{\mathcal{N}}(X,Y) = 0$ for $X, Y \in \mathcal{N}$ even when X and Y are not strongly isomorphic. In Theorem 23 below, we provide an extension of Theorem 19 to a class of compact dissimilarity networks that contains all finite dissimilarity networks.

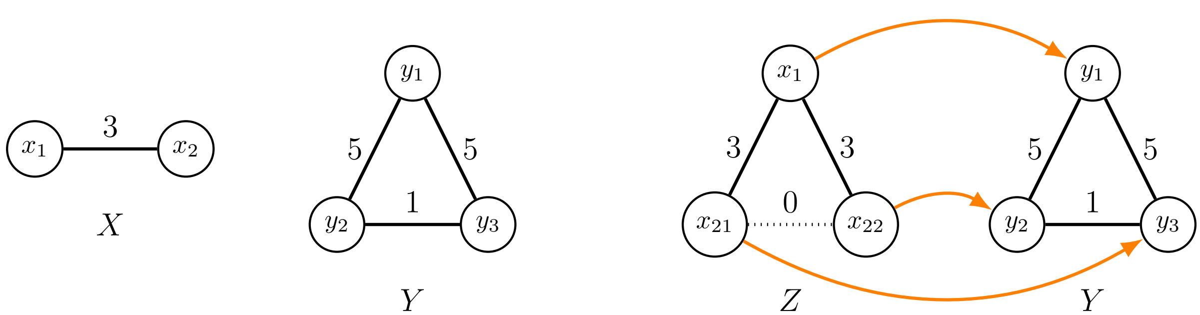

FIGURE 4. The two networks on the left have different cardinalities, but computing correspondences shows that $d_{\mathcal{N}}(X,Y) = 1$. Similarly one computes $d_{\mathcal{N}}(X,Z) = 0$, and thus $d_{\mathcal{N}}(Y,Z) = 0$ by triangle inequality. On the other hand, the bijection given by the red arrows shows $\widehat{d}_{\mathcal{N}}(Y,Z) = 1$. Applying Proposition 17 then recovers $d_{\mathcal{N}}(X,Y) = 1$.

Example 20. Finite metric spaces and finite ultrametric spaces constitute obvious examples of dissimilarity networks. Recall that, in an ultrametric space (X, d_X), we have the strong triangle inequality $d_X(x, x') \leqslant \max\{d_X(x, x''), d_X(x'', x')\}$ for all $x, x', x'' \in X$. More interesting classes of dissimilarity networks arise by relaxing the symmetry and triangle inequality conditions of metric spaces.

Definition 12 (Finite reversibility and Ψ-controllability)**.** The *reversibility* ρ_X of a dissimilarity network (X, A_X) is defined to be the following quantity:

$$\rho_X := \sup_{x \neq x' \in X} \frac{A_X(x, x')}{A_X(x', x)}.$$

(X, A_X) is said to have *finite reversibility* if $\rho_X < \infty$. Notice that $\rho_X \geqslant 1$, with equality if and only if A_X is symmetric.

Next let $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous function such that $\Psi(0, 0) = 0$. A dissimilarity network (X, A_X) is said to be Ψ*-controlled* if we have

$$A_X(x, x') \leqslant \Psi\big(A_X(x, x''), A_X(x', x'')\big) \text{ for all } x, x', x'' \in X.$$

This condition automatically encodes a notion of reversibility:

$$\begin{aligned} A_X(x, x') &\leqslant \Psi\big(A_X(x, x), A_X(x', x)\big) &= \Psi\big(0, A_X(x', x)\big), \\ A_X(x', x) &\leqslant \Psi\big(A_X(x', x'), A_X(x, x')\big) &= \Psi\big(0, A_X(x, x')\big). \end{aligned}$$

In the sequel, whenever we write "$(X, A_X) \in \mathcal{N}^{\text{dis}}$ is Ψ-controlled" without explicit reference to a map Ψ, we mean that there exists a function $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfying the conditions above.

Remark 21. Any finite dissimilarity network is finitely reversible and Ψ-controllable. For example, Ψ can be taken to be a bump function that vanishes outside $\mathbb{R}^2_+ \backslash U$—where U is some open set containing $\operatorname{im}(A_X)$ and excluding $(0, 0)$—and constant at $\max_{x, x' \in X} A_X(x, x')$ on U.

Dissimilarity networks satisfying the symmetry condition, but not the triangle inequality, have a long history dating back to Fréchet's thesis [33] and continuing with work by Pitcher and Chittenden [78], Niemytzki [67], Galvin and Shore [34, 35], and many others, as summarized in [43]. One of the interesting directions in this line of work was the development of a "local triangle inequality" and related metrization theorems [67], which has been continued more recently in [99].

Dissimilarity networks satisfying the triangle inequality, but not symmetry, include the special class of objects called *directed metric spaces*, which we define below.

Definition 13. Let (X, A_X) be a dissimilarity network. Given any $x \in X$ and $r \in \mathbb{R}_+$, the *forward-open ball* of radius r centered at x is

$$B^+(x, r) := \{x' \in X : A_X(x, x') < r\}.$$

The *forward-open topology induced by* A_X is the topology on X generated by the collection $\{B^+(x, r) : x \in X,\ r > 0\}$. The idea of forward open balls is prevalent in the study of Finsler geometry; see [2, p. 149] for details.

Definition 14 (Directed metric spaces)**.** A *directed metric space* or *quasi-metric space* is a dissimilarity network (X, ν_X) such that X is equipped with the forward-open topology induced by ν_X and $\nu_X : X \times X \to \mathbb{R}_+$ satisfies:

$$\nu_X(x, x'') \leqslant \nu_X(x, x') + \nu_X(x', x'') \text{ for all } x, x', x'' \in X.$$

The function ν_X is called a directed metric or quasi-metric on X. Notice that compact directed metric spaces constitute a subfamily of $\mathcal{CN}^{\text{dis}}$.

Directed metric spaces with finite reversibility were studied in [84], and constitute important examples of networks that are strictly non-metric. More specifically, the authors of [84] extended notions of Hausdorff distance and Gromov-Hausdorff distance to the setting of directed metric spaces with finite reversibility, and our network distance $d_{\mathcal{N}}$ subsumes this theory while extending it to even more general settings.

Remark 22 (Finsler metrics)**.** An interesting class of directed metric spaces arises from studying Finsler manifolds. A *Finsler manifold* (M, F) is a smooth, connected manifold M equipped with an asymmetric norm F (called a *Finsler function*) defined on each tangent space of M [2]. A Finsler function induces a directed metric $d_F : M \times M \to \mathbb{R}_+$ as follows: for each $x, x' \in M$,

$$d_F(x, x') := \inf \left\{ \int_a^b F(\gamma(t), \dot{\gamma}(t))\, dt : \gamma : [a, b] \to M \text{ a smooth curve joining } x \text{ and } x' \right\}.$$

Finsler metric spaces have received interest in the applied literature. In [81], the authors prove that Finsler metric spaces with *reversible geodesics* (i.e. the reverse curve $\gamma'(t) := \gamma(1 - t)$ of any geodesic $\gamma : [0, 1] \to M$ is also a geodesic) is a *weighted quasi-metric* [81, p. 2]. Such objects have been shown to be essential in biological sequence comparison [94].

We end this section with a strengthening of Theorem 19 to the setting of compact networks. Recall that the interesting part of Theorem 19 was to show that $d_{\mathcal{N}}(X, Y) = 0 \implies X \cong^s Y$; we generalize this result to a certain class of *compact* dissimilarity networks.

Theorem 23. *Let $(X, A_X), (Y, A_Y) \in \mathcal{CN}^{dis}$ be equipped with the forward-open topologies induced by A_X and A_Y, respectively. Suppose also that at least one of the two networks is Ψ-controlled. Then $d_{\mathcal{N}}(X, Y) = 0 \implies X \cong^s Y$.*

Remark 24 (Generalizations of Theorem 19)**.** For any finite dissimilarity network (X, A_X), the discrete topology is precisely the topology induced by A_X. We have already stated before that finite dissimilarity networks trivially satisfy finite reversibility and triangulability. It folows that Theorem 23 is a bona fide generalization of Theorem 19.

We will present a further generalization in the setting of compact dissimilarity networks—where the restrictions on topology and weights are further relaxed—in a forthcoming publication.

2.3. **Two families of examples: the directed circles.** In this section, we explicitly construct an infinite network in $\mathcal{N}^{\text{dis}}$, and a family of infinite networks in $\mathcal{CN}^{\text{dis}}$.

2.3.1. *The general directed circle.* First we construct an asymmetric network in $\mathcal{N}^{\text{dis}}$. To motivate this construction, recall from the classification of topological 1-manifolds that *any connected, closed topological 1-manifold is homeomorphic to the circle* $\mathbb{S}^1$. So as a first construction of a quasi-metric space, it is reasonable to adopt $\mathbb{S}^1$ as our model and endow it with a quasi-metric weight function.

First define the set $\vec{\mathbb{S}}^1 := \left\{ e^{i\theta} \in \mathbb{C} : \theta \in [0, 2\pi) \right\}$. For any $\alpha, \beta \in [0, 2\pi)$, define $\vec{d}(\alpha, \beta) := \beta - \alpha \bmod 2\pi$, with the convention $\vec{d}(\alpha, \beta) \in [0, 2\pi)$. Then $\vec{d}(\alpha, \beta)$ is the counterclockwise geodesic distance along the unit circle from $e^{i\alpha}$ to $e^{i\beta}$. As such, it satisfies the triangle inequality and vanishes on a pair $(e^{i\theta_1}, e^{i\theta_2})$ if and only if $\theta_1 = \theta_2$. Next for each $e^{i\theta_1}, e^{i\theta_2} \in \vec{\mathbb{S}}^1$, define

$$\omega_{\vec{\mathbb{S}}^1}(e^{i\theta_1}, e^{i\theta_2}) := \vec{d}(\theta_1, \theta_2).$$

To finish the construction, we specify $\vec{\mathbb{S}}^1$ to have the discrete topology. Clearly this is first countable and makes $\omega_{\vec{\mathbb{S}}^1}$ continuous, but the resulting network will not be compact. Hence it is natural to ask if there exists a coarser topology that we can place on $\vec{\mathbb{S}}^1$.

We claim that a coarser topology does not work to make $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1})$ fit the framework of $\mathcal{N}$. To see why, let $\alpha \in [0, 2\pi)$. Suppose $\omega_{\vec{\mathbb{S}}^1}$ is continuous with respect to some topology on $\vec{\mathbb{S}}^1$, to be determined. Fix $0 < \varepsilon \ll 2\pi$, and define $V := \omega_{\vec{\mathbb{S}}^1}^{-1}[(-\varepsilon, \varepsilon)]$. Then V is open in the product topology, and in particular contains $(e^{i\alpha}, e^{i\alpha})$. Since V is a union of open rectangles, there exists an open set $U \subseteq \vec{\mathbb{S}}^1$ such that $(e^{i\alpha}, e^{i\alpha}) \in U \times U \subseteq V$. Suppose towards a contradiction that $U \neq \{e^{i\alpha}\}$. Then there exists $e^{i\beta} \in U$, for some $\beta \neq \alpha$. Then $\omega_{\vec{\mathbb{S}}^1}(e^{i\alpha}, e^{i\beta}) \in (0, \varepsilon)$. But by the definition of $\omega_{\vec{\mathbb{S}}^1}$, we must have $\omega_{\vec{\mathbb{S}}^1}(e^{i\beta}, e^{i\alpha}) \in [2\pi - \varepsilon, 2\pi)$, which is a contradiction to having $\omega_{\vec{\mathbb{S}}^1}(U, U) \subseteq (-\varepsilon, \varepsilon)$.

Definition 15. We define the *directed unit circle* to be $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1})$ with the discrete topology.

This asymmetric network provides us with concrete examples of ε-approximations (Definition 10), for any $\varepsilon > 0$. To see this, fix any $n \in \mathbb{N}$, and consider the *directed circle network on n nodes* $(\vec{\mathbb{S}}^1_n, \omega_{\vec{\mathbb{S}}^1_n})$ obtained by writing

$$\vec{\mathbb{S}}^1_n := \left\{ e^{\frac{2\pi i k}{n}} \in \mathbb{C} : k \in \{0, 1, \ldots, n-1\} \right\},$$

and defining $\omega_{\vec{\mathbb{S}}^1_n}$ to be the restriction of $\omega_{\vec{\mathbb{S}}^1}$ on this set. An illustration of $\vec{\mathbb{S}}^1$ and $\vec{\mathbb{S}}^1_n$ for $n = 6$ is provided in Figure 5.

Theorem 25. *As $n \to \infty$, the sequence of finite dissimilarity networks $\vec{\mathbb{S}}^1_n$ limits to the dissimilarity network $\vec{\mathbb{S}}^1$ in the sense of $d_{\mathcal{N}}$.*

Proof of Theorem 25. Let $\varepsilon > 0$, and let $n \in \mathbb{N}$ be such that $2\pi/n < \varepsilon$. It suffices to show that $d_{\mathcal{N}}(\vec{\mathbb{S}}^1, \vec{\mathbb{S}}^1_n) < \varepsilon$. Define a correspondence between $\vec{\mathbb{S}}^1$ and $\vec{\mathbb{S}}^1_n$ as follows:

$$R := \left\{ (e^{i\theta}, e^{\frac{2\pi i k}{n}}) : \theta \in (\tfrac{2\pi i(k-1)}{n}, \tfrac{2\pi i k}{n}],\ k \in \{0, 1, 2, \ldots, n-1\} \right\}.$$

Here each point of $\vec{\mathbb{S}}^1$ is placed in correspondence with the closest point on $\vec{\mathbb{S}}^1_n$ obtained by traveling counterclockwise on $\vec{\mathbb{S}}^1$. Next let $0 \leqslant \theta_1 \leqslant \theta_2 < 2\pi$, and let $j, k \in \{0, 1, \ldots, n-1\}$ be such that $\theta_1 \in (\frac{2\pi i(j-1)}{n}, \frac{2\pi i j}{n}]$ and $\theta_2 \in (\frac{2\pi i(k-1)}{n}, \frac{2\pi i k}{n}]$. Then we have:

$$\begin{aligned} |\omega_{\vec{\mathbb{S}}^1}(e^{i\theta_1}, e^{i\theta_2}) - \omega_{\vec{\mathbb{S}}^1}(e^{\frac{2\pi i j}{n}}, e^{\frac{2\pi i k}{n}})| &= |\theta_2 - \theta_1 - \tfrac{2\pi i k}{n} + \tfrac{2\pi i j}{n}| \\ &\leqslant |\tfrac{2\pi i k}{n} - \theta_2| + |\tfrac{2\pi i j}{n} - \theta_1| < 4\pi/n < 2\varepsilon. \end{aligned}$$

Similarly we have:

$$|\omega_{\vec{\mathbb{S}}^1}(e^{i\theta_2}, e^{i\theta_1}) - \omega_{\vec{\mathbb{S}}^1}(e^{\frac{2\pi i k}{n}}, e^{\frac{2\pi i j}{n}})| = |2\pi - \theta_2 + \theta_1 - 2\pi + \tfrac{2\pi i k}{n} - \tfrac{2\pi i j}{n}| < 2\varepsilon.$$

It follows that $\operatorname{dis}(R) < 2\varepsilon$. Thus $d_{\mathcal{N}}(\vec{\mathbb{S}}^1, \vec{\mathbb{S}}^1_n) < \varepsilon$. The theorem follows. $\square$

2.3.2. *The directed circles with finite reversibility.* Now we define a family of directed circles parametrized by reversibility. Unlike the construction in §2.3.1, these directed networks belong to the family $\mathcal{CN}^{\text{dis}}$. An illustration is provided in Figure 5.

Recall from §2.3.1 that for $\alpha, \beta \in [0, 2\pi)$, we wrote $\vec{d}(\alpha, \beta)$ to denote the counterclockwise geodesic distance along the unit circle from $e^{i\alpha}$ to $e^{i\beta}$. Fix $\rho \geqslant 1$. For each $e^{i\theta_1}, e^{i\theta_2} \in \vec{\mathbb{S}}^1$, define

$$\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\theta_1}, e^{i\theta_2}) := \min\left(\vec{d}(\theta_1, \theta_2), \rho \vec{d}(\theta_2, \theta_1) \right).$$

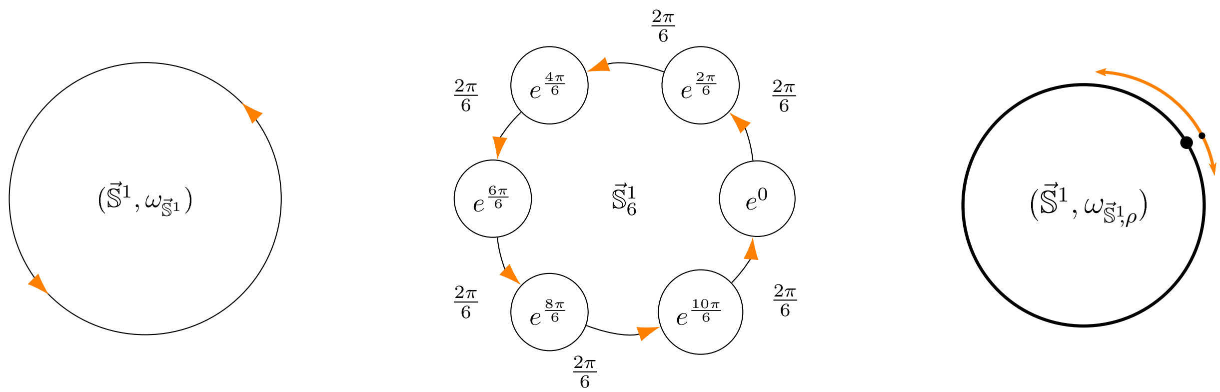

FIGURE 5. The directed circle $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1})$, the directed circle on 6 nodes $(\vec{\mathbb{S}}^1_6, \omega_{\vec{\mathbb{S}}^1_6})$, and the directed circle with reversibility ρ, for some $\rho \in [1, \infty)$. Traveling in a clockwise direction is possibly only in the directed circle with reversibility ρ, but this incurs a penalty modulated by ρ.

In particular, $\omega_{\vec{\mathbb{S}}^1,\rho}$ has reversibility ρ (cf. Definition 12).

Finally, we equip $\vec{\mathbb{S}}^1$ with the standard subspace topology generated by the open balls in $\mathbb{C}$. In this case, $\vec{\mathbb{S}}^1$ is compact and first countable. It remains to check that $\omega_{\vec{\mathbb{S}}^1,\rho}$ is continuous.

Proposition 26. $\omega_{\vec{\mathbb{S}}^1,\rho} : \vec{\mathbb{S}}^1 \times \vec{\mathbb{S}}^1 \to \mathbb{R}$ *is continuous.*

Proof of Proposition 26. It suffices to show that the preimages of basic open sets under $\omega_{\vec{\mathbb{S}}^1,\rho}$ are open. Let (a, b) be an open interval in $\mathbb{R}$. Let $(e^{i\alpha}, e^{i\beta}) \in \omega^{-1}_{\vec{\mathbb{S}}^1,\rho}[(a, b)]$, where $\alpha, \beta \in [0, 2\pi)$. There are three cases: (1) $\alpha < \beta$, (2) $\beta < \alpha$, or (3) $\alpha = \beta$.

Suppose first that $\alpha < \beta$. There are two subcases: either $\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\alpha}, e^{i\beta}) = \vec{d}(\alpha, \beta)$, or $= \rho\vec{d}(\beta, \alpha)$. Fix $r > 0$ to be determined later, but small enough so that $B(\alpha, r) \cap B(\beta, r) = \varnothing$. Let $\gamma \in B(\alpha, r)$, $\delta \in B(\beta, r)$. Then $\vec{d}(\gamma, \delta) \in B(\vec{d}(\alpha, \beta), 2r)$. Also,

$$\left|\rho\vec{d}(\gamma, \delta) - \rho\vec{d}(\alpha, \beta)\right| = \rho\left|\vec{d}(\gamma, \delta) - \vec{d}(\alpha, \beta)\right| < 2r\rho.$$

Now r can be made arbitrarily small, so that for any $\gamma \in B(\alpha, r)$ and any $\delta \in B(\delta, r)$, we have $\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\gamma}, e^{i\delta}) \in (a, b)$. It follows that $(e^{i\alpha}, e^{i\beta})$ is contained in an open set contained inside $\omega^{-1}_{\vec{\mathbb{S}}^1,\rho}[(a, b)]$. An analogous proof shows this to be true for the $\beta < \alpha$ case.

Next suppose $\alpha = \beta$. Fix $0 < r < b/(2\rho)$. We need to show $\omega_{\vec{\mathbb{S}}^1,\rho}(B(\alpha, r), B(\alpha, r)) \subseteq (a, b)$. Note that $0 \in (a, b)$. Let $\gamma, \delta \in B(\alpha, r)$. There are three subcases. If $\gamma = \delta$, then $\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\gamma}, e^{i\delta}) = 0 \in (a, b)$. If $\vec{d}(\gamma, \delta) < 2r$, then $\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\gamma}, e^{i\delta}) < 2r < b$. Finally, suppose $\vec{d}(\gamma, \delta) \geqslant 2r$. Then we must have $\vec{d}(\delta, \gamma) < 2r$, so $\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\gamma}, e^{i\delta}) \leqslant \rho\vec{d}(\delta, \gamma) < 2r\rho < b$. Thus for any $\gamma, \delta \in B(\alpha, r)$, we have $\omega_{\vec{\mathbb{S}}^1,\rho}(e^{i\gamma}, e^{i\delta}) \in (a, b)$.

It follows that $\omega^{-1}_{\vec{\mathbb{S}}^1,\rho}[(a, b)]$ is open. This proves the claim. □

We summarize the preceding observations in the following:

Definition 16. Let $\rho \in [1, \infty)$. We define the *directed unit circle with reversibility* ρ to be $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1,\rho})$. This is a compact, asymmetric network in $\mathcal{CN}^{\text{dis}}$.

Remark 27. Just as in Theorem 25, we can finitely approximate $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1,\rho})$ by endowing uniformly distributed points on $\vec{\mathbb{S}}^1$ with the restriction of $\omega_{\vec{\mathbb{S}}^1,\rho}$. We do not repeat the details here.

Remark 28 (Directed circle with finite reversibility—forward-open topology version)**.** Instead of using the subspace topology generated by the standard topology on $\mathbb{C}$, we can also endow $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1,\rho})$ with the forward-open topology generated by $\omega_{\vec{\mathbb{S}}^1,\rho}$. The open balls in this topology are precisely the open balls in the subspace topology induced by the standard topology, the only adjustment being the "center" of each ball. The directed metric space $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1,\rho})$ equipped with the forward-open topology is another example of a compact, asymmetric network in $\mathcal{CN}^{\text{dis}}$.

3. The case of compact networks

In this section, we characterize the compact networks in the metric space $(\mathcal{N}/\cong^w_{\text{II}}, d_{\mathcal{N}})$, where $d_{\mathcal{N}} : \mathcal{N}/\cong^w_{\text{II}} \times \mathcal{N}/\cong^w_{\text{II}} \to \mathbb{R}_+$ is defined (abusing notation) as follows:

$$d_{\mathcal{N}}([X],[Y]) := d_{\mathcal{N}}(X,Y), \quad \text{for each } [X],[Y] \in \mathcal{N}/\cong^w_{\text{II}} .$$

To check that $d_{\mathcal{N}}$ is well-defined on $[X],[Y] \in \mathcal{N}/\cong^w_{\text{II}}$, let $X' \in [X], Y' \in [Y]$. Then:

$$d_{\mathcal{N}}([X'],[Y']) = d_{\mathcal{N}}(X',Y') = d_{\mathcal{N}}(X,Y) = d_{\mathcal{N}}([X],[Y]),$$

where the second-to-last equality follows from the triangle inequality and the observation that $d_{\mathcal{N}}(X,X') = d_{\mathcal{N}}(Y,Y') = 0$.

Our main result is that the two types of weak isomorphism coincide in the setting of compact networks. As a stepping stone towards proving this result, we explore the notion of "sampling" finite networks from compact networks.

3.1. Compact networks and finite sampling.

In this section, we prove that any compact network admits an approximation by a finite network up to arbitrary precision, in the sense of $d_{\mathcal{N}}$.

Example 29 (Some compact and noncompact networks)**.** In §2.3.1, we constructed an example of a noncompact, asymmetric network. In Section 2.3.2, we constructed a family of compact, asymmetric networks: the directed circles with reversibility $\rho \in [1,\infty)$. We also remarked that $(\vec{\mathbb{S}}^1, \omega_{\vec{\mathbb{S}}^1,\rho})$ can be "approximated" up to arbitrary precision by picking n equidistant points on $\vec{\mathbb{S}}^1$ and equipping this collection with the restriction of $\omega_{\vec{\mathbb{S}}^1,\rho}$ (also see Theorem 25). We view this process as "sampling" finite networks from a compact network. In the next result, we present this sampling process as a theorem that applies to *any* compact network.

Theorem 30 (Sampling from a compact network)**.** *Let (X,ω_X) be a compact network. Then for any $\varepsilon > 0$, we can choose a finite subset $X' \subseteq X$ such that*

$$d_{\mathcal{N}}((X,\omega_X),(X',\omega_X|_{X'\times X'})) < \varepsilon.$$

Remark 31. When considering a compact metric space (X,d_X), the preceding theorem relates to the well-known notion of taking finite ε-nets in a metric space. Recall that for $\varepsilon > 0$, a subset $S \subseteq X$ is an ε-net if for any point $x \in X$, we have $B(x,\varepsilon) \cap S \neq \varnothing$. Such an ε-net satisfies the nice property that $d_{\text{GH}}(X,S) < \varepsilon$ [7, 7.3.11]. In particular, one can find a finite ε-net of (X,d_X) for any $\varepsilon > 0$ by compactness.

Observe that we do not make quantitative estimates on the cardinality of the ε-approximation produced in Theorem 30. In the setting of compact metric spaces, the size of an ε-net relates to the rich theory of metric entropy developed by Kolmogorov and Tihomirov [28, Chapter 17].

By virtue of Theorem 30, one can always approximate a compact network up to any given precision. The next theorem implies that a sampled network limits to the underlying compact network as the sample gets more and more dense.

Theorem 32 (Limit of dense sampling)**.** *Let (X, ω_X) be a compact network, and let $S = \{s_1, s_2, \ldots\}$ be a countable dense subset of X with a fixed enumeration. For each $n \in \mathbb{N}$, let X_n be the finite network with node set $\{s_1, \ldots, s_n\}$ and weight function $\omega_X|_{X_n \times X_n}$. Then we have:*

$$d_{\mathcal{N}}(X, X_n) \downarrow 0 \text{ as } n \to \infty.$$

3.1.1. *Proofs of result in §3.1.*

Proof of Theorem 30. The idea is to find a cover of X by open sets $G_1, \ldots, G_q$ and representatives $x_i \in G_i$ for each $1 \leqslant i \leqslant q$ such that whenever we have $(x, x') \in G_i \times G_j$, we know by continuity of ω_X that $|\omega_X(x, x') - \omega_X(x_i, x_j)| < \varepsilon$. Then we define a correspondence that associates each $x \in G_i$ to x_i, for $1 \leqslant i \leqslant q$. Such a correspondence has distortion bounded above by ε.

Let $\varepsilon > 0$. Let $\mathscr{B}$ be a base for the topology on X.

Let $\{B(r, \varepsilon/4) : r \in \mathbb{R}\}$ be an open cover for $\mathbb{R}$. Then by continuity of ω_X, we get that

$$\left\{\omega_X^{-1}[B(r, \varepsilon/4)] : r \in \mathbb{R}\right\}$$

is an open cover for $X \times X$. Each open set in this cover can be written as a union of open rectangles $U \times V$, for $U, V \in \mathscr{B}$. Thus the following set is an open cover of $X \times X$:

$$\mathscr{U} := \left\{U \times V : U, V \in \mathscr{B},\ U \times V \subseteq \omega_X^{-1}[B(r, \varepsilon/4)],\ r \in \mathbb{R}\right\}.$$

Claim 1. There exists a finite open cover $\mathscr{G} = \{G_1, \ldots, G_q\}$ of X such that for any $1 \leqslant i, j \leqslant q$, we have $G_i \times G_j \subseteq U \times V$ for some $U \times V \in \mathscr{U}$.

Proof of Claim 1. The proof of the claim proceeds by a repeated application of the Tube Lemma [65, Lemma 26.8]. Since $X \times X$ is compact, we take a finite subcover:

$$\mathscr{U}^f := \{U_1 \times V_1, \ldots, U_n \times V_n\}, \text{ for some } n \in \mathbb{N}.$$

Let $x \in X$. Then we define:

$$\mathscr{U}_x^f := \{U \times V \in \mathscr{U}^f : x \in U\},$$

and write

$$\mathscr{U}_x^f = \left\{U_{i_1}^x \times V_{i_1}^x, \ldots, U_{i_{m(x)}}^x \times V_{i_{m(x)}}^x\right\}.$$

Here $m(x)$ is an integer depending on x, and $\left\{i_1, \ldots, i_{m(x)}\right\}$ is a subset of $\{1, \ldots, n\}$.

Since $\mathscr{U}^f$ is an open cover of $X \times X$, we know that $\mathscr{U}_x^f$ is an open cover of $\{x\} \times X$. Next define:

$$A_x := \bigcap_{k=1}^{m(x)} U_{i_k}^x.$$

Then A_x is open and contains x. In the literature [65, p. 167], the set $A_x \times X$ is called a *tube* around $\{x\} \times X$. Notice that $A_x \times X \subseteq \mathscr{U}_x^f$. Since x was arbitrary in the preceding construction, we define $\mathscr{U}_x^f$ and A_x for each $x \in X$. Then note that $\{A_x : x \in X\}$ is an open cover of X. Using compactness of X, we choose $\{s_1, \ldots, s_p\} \subseteq X,\ p \in \mathbb{N}$, such that $\left\{A_{s_1}, \ldots, A_{s_p}\right\}$ is a finite subcover of X.

Once again let $x \in X$, and let $\mathscr{U}_x^f$ and A_x be defined as above. Define the following:

$$B_x := \left\{A_x \times V_{i_k}^x : 1 \leqslant k \leqslant m(x)\right\}.$$

Since $x \in A_x$ and $X \subseteq \cup_{k=1}^{m(x)} V_{i_k}^x$, it follows that B_x is a cover of $\{x\} \times X$. Furthermore, since $\{A_{s_1}, \ldots, A_{s_p}\}$ is a cover of X, it follows that the finite collection $\{B_{s_1}, \ldots, B_{s_p}\}$ is a cover of $X \times X$.

Let $z \in X$. Since $X \subseteq \cup_{k=1}^{m(x)} V_{i_k}^x$, we pick $V_{i_k}^x$ for $1 \leqslant k \leqslant m(x)$ such that $z \in V_{i_k}^x$. Since x was arbitrary, such a choice exists for each $x \in X$. Therefore, we define:

$$C_z := \{V \in \mathscr{B} : z \in V,\ A_{s_i} \times V \in B_{s_i} \text{ for some } 1 \leqslant i \leqslant p\}.$$

Since each B_{s_i} is finite and there are finitely many B_{s_i}, we know that C_z is a finite collection. Next define:

$$D_z := \bigcap_{V \in C_z} V.$$

Then D_z is open and contains z. Notice that $X \times D_z$ is a tube around $X \times \{z\}$. Next, using the fact that $\{A_{s_i} : 1 \leqslant i \leqslant p\}$ is an open cover of X, pick $A_{s_{i(z)}}$ such that $z \in A_{s_{i(z)}}$. Here $1 \leqslant i(z) \leqslant p$ is some integer depending on z. Then define

$$G_z := D_z \cap A_{s_{i(z)}}.$$

Then G_z is open and contains z. Since z was arbitrary, we define G_z for each $z \in X$. Then $\{G_z : z \in X\}$ is an open cover of X, and we take a finite subcover:

$$\mathscr{G} := \{G_1, \ldots, G_q\},\ q \in \mathbb{N}.$$

Finally, we need to show that for any choice of $1 \leqslant i, j \leqslant q$, we have $G_i \times G_j \subseteq U \times V$ for some $U \times V \in \mathscr{U}$. Let $1 \leqslant i, j \leqslant q$. Note that we can write $G_i = G_w$ and $G_j = G_y$ for some $w, y \in X$. By the definition of G_w, we then have the following for some index $i(w)$ depending on w:

$$G_w \subseteq A_{s_{i(w)}} \subseteq U^{s_{i(w)}} \text{ for some } U^{s_{i(w)}} \times V^{s_{i(w)}} \in \mathscr{U}^f_{s_{i(w)}},\ 1 \leqslant i(w) \leqslant p.$$

Note that the second containment holds by definition of $A_{s_{i(w)}}$. Since $\mathscr{U}^f_{s_{i(w)}}$ is a cover of $\{s_{i(w)}\} \times X$, we choose $V^{s_{i(w)}}$ to contain y. Then observe that $A_{s_{i(w)}} \times V^{s_{i(w)}} \in B_{s_{i(w)}}$. Then $V^{s_{i(w)}} \in C_y$, and so we have:

$$G_y \subseteq D_y \subseteq V^{s_{i(w)}}.$$

It follows that $G_i \times G_j = G_w \times G_y \subseteq U^{x_{i(w)}} \times V^{x_{i(w)}} \in \mathscr{U}$. ■

Now we fix $\mathscr{G} = \{G_1, \ldots, G_q\}$ as in Claim 1. Before defining X', we perform a disjointification step. Define:

$$\widetilde{G}_1 := G_1,\ \widetilde{G}_2 := G_2 \backslash \widetilde{G}_1,\ \widetilde{G}_3 := G_3 \backslash (\widetilde{G}_1 \cup \widetilde{G}_2),\ \ldots,\ \widetilde{G}_q := G_q \backslash \left(\cup_{k=1}^{q-1} \widetilde{G}_k \right).$$

Finally we define X' as follows: pick a representative $x_i \in \widetilde{G}_i$ for each $1 \leqslant i \leqslant q$. Let $X' = \{x_i : 1 \leqslant i \leqslant q\}$. Define a correspondence between X and X' as follows:

$$R := \left\{ (x, x_i) : x \in \widetilde{G}_i, 1 \leqslant i \leqslant q \right\}.$$

Let $(x, x_i), (x', x_j) \in R$. Then we have $(x, x'), (x_i, x_j) \in \widetilde{G}_i \times \widetilde{G}_j \subseteq G_i \times G_j$. By the preceding work, we know that $G_i \times G_j \subseteq U \times V$, for some $U \times V \in \mathscr{U}$. Therefore $\omega_X(x, x'), \omega_X(x_i, x_j) \in B(r, \varepsilon/4)$ for some $r \in \mathbb{R}$. It follows that:

$$|\omega_X(x, x') - \omega_X(x_i, x_j)| < \varepsilon/2.$$

Since $(x, x_i), (x', x_j) \in R$ were arbitrary, we have $\operatorname{dis}(R) < \varepsilon/2$. Hence $d_{\mathcal{N}}(X, X') < \varepsilon$. □

Proof of Theorem 32. The first part of this proof is similar to that of Theorem 30. Let $\varepsilon > 0$. Let $\mathscr{B}$ be a base for the topology on X. Then $\{\omega_X^{-1}[B(r,\varepsilon/8)] : r \in \mathbb{R}\}$ is an open cover for $X \times X$. Each open set in this cover can be written as a union of open rectangles $U \times V$, for $U, V \in \mathscr{B}$. Thus the following set is an open cover of $X \times X$:

$$\mathscr{U} := \left\{U \times V : U, V \in \mathscr{B},\ U \times V \subseteq \omega_X^{-1}[B(r,\varepsilon/8)],\ r \in \mathbb{R}\right\}.$$

By applying Claim 1 from the proof of Theorem 30, we obtain a finite open cover $\mathscr{G} = \{G_1, \ldots, G_q\}$ of X such that for any $1 \leqslant i, j \leqslant q$, we have $G_i \times G_j \subseteq U \times V$ for some $U \times V \in \mathscr{U}$. For convenience, we assume that each G_i is nonempty.

Now let $1 \leqslant i \leqslant q$. Then $G_i \cap S \neq \varnothing$, because S is dense in X. Choose $p(i) \in \mathbb{N}$ such that $s_{p(i)} \in G_i$. We repeat this process for each $1 \leqslant i \leqslant q$, and then define

$$n := \max\{p(1), p(2), \ldots, p(q)\}.$$

Now define X_n to be the network with node set $\{s_1, s_2, \ldots, s_n\}$ and weight function given by the appropriate restriction of ω_X. Also define S_n to be the network with node set $\{s_{p(1)}, s_{p(2)}, \ldots, s_{p(q)}\}$ and weight function given by the restriction of ω_X.

Claim 2. Let A be a subset of X equipped with the weight function $\omega_X|_{A\times A}$. Then $d_{\mathcal{N}}(S_n, A) < \varepsilon/2$.

Proof of Claim 2. We begin with $\mathscr{G} = \{G_1, \ldots, G_q\}$. Notice that each G_i contains $s_{p(i)}$. To avoid ambiguity in our construction, we will need to ensure that G_i does not contain $s_{p(j)}$ for $i \neq j$. So our first step is to obtain a cover of A by disjoint sets while ensuring that each $s_{p(i)} \in S_n$ belongs to exactly one element of the new cover. We define:

$$G_1^* := G_1 \backslash S_n,\ G_2^* := G_2 \backslash S_n,\ G_3^* := G_3 \backslash S_n,\ \ldots,\ G_q^* := G_q \backslash S_n, \text{ and}$$

$$\widetilde{G}_1 := G_1^* \cup \{s_{p(1)}\},\ \widetilde{G}_2 := (G_2^* \backslash \widetilde{G}_1) \cup \{s_{p(2)}\},\ \widetilde{G}_3 := \left(G_3^* \backslash (\widetilde{G}_1 \cup \widetilde{G}_2)\right) \cup \{s_{p(3)}\},\ \ldots,$$
$$\widetilde{G}_q := \left(G_q^* \backslash \left(\cup_{k=1}^{q-1} \widetilde{G}_k\right)\right) \cup \{s_{p(q)}\}.$$

Notice that $\left\{\widetilde{G}_i : 1 \leqslant i \leqslant q\right\}$ is a cover for A, and for each $1 \leqslant i \leqslant q$, $\widetilde{G}_i$ contains $s_{p(j)}$ if and only if $i = j$. Now we define a correspondence between A and S_n as follows:

$$R := \left\{(x, s_{p(i)}) : x \in A \cap \widetilde{G}_i,\ 1 \leqslant i \leqslant q\right\}.$$

Next let $(x, s_{p(i)}), (x', s_{p(j)}) \in R$. Then we have $(x, x'), (s_{p(i)}, s_{p(j)}) \in \widetilde{G}_i \times \widetilde{G}_j \subseteq G_i \times G_j \subseteq U \times V$ for some $U \times V \in \mathscr{U}$. Therefore $\omega_X(x, x')$ and $\omega_X(s_{p(i)}, s_{p(j)})$ both belong to $B(r, \varepsilon/8)$ for some $r \in \mathbb{R}$. Thus we have:

$$|\omega_X(x, x') - \omega_X(s_{p(i)}, s_{p(j)})| < \varepsilon/4.$$

It follows that $\operatorname{dis}(R) < \varepsilon/4$, and so $d_{\mathcal{N}}(A, S_n) < \varepsilon/2$. ■

Finally, we note that $d_{\mathcal{N}}(X, X_n) \leqslant d_{\mathcal{N}}(X, S_n) + d_{\mathcal{N}}(S_n, X_n) < \varepsilon/2 + \varepsilon/2 = \varepsilon$, by Claim 2. Since $\varepsilon > 0$ was arbitrary, it follows that $d_{\mathcal{N}}(X, X_n) \to 0$.

For the final statement in the theorem, let $m \geqslant n$ and observe that $S_n \subseteq X_n \subseteq X_m$. Thus whenever we have $d_{\mathcal{N}}(X, X_n) < \varepsilon$, we also have $d_{\mathcal{N}}(X, X_m) < \varepsilon$. It follows that:

$$d_{\mathcal{N}}(X, X_m) \leqslant d_{\mathcal{N}}(X, X_n) \text{ for any } m, n \in \mathbb{N},\ m \geqslant n. \qquad \square$$

3.2. **Compact networks and weak isomorphism.** By Theorem 12, $d_{\mathcal{N}}$ is a proper metric on $\mathcal{N}$ modulo Type II weak isomorphism, which is equivalent to Type I weak isomorphism when restricted to $\mathcal{FN}$. The comparison between $\mathbb{Q} \cap [0,1]$ and $[0,1]$ in Example 9 shows that in general, these two notions of weak isomorphism are not equivalent. This leads to the following natural question: when restricted to $\mathcal{CN}$, are we still able to recover equivalence between Type I and Type II weak isomorphism?

In the following theorem, we provide a positive answer to this question.

Theorem 13 (Weak isomorphism in $\mathcal{CN}$)**.** *Let $X, Y \in \mathcal{CN}$. Then X and Y are Type II weakly isomorphic if and only if X and Y are Type I weakly isomorphic, i.e. there exists a set V and surjections $\varphi_X : V \to X,\ \varphi_Y : V \to Y$ such that:*

$$\omega_X(\varphi_X(v), \varphi_X(v')) = \omega_Y(\varphi_Y(v), \varphi_Y(v')) \quad \textit{for all } v, v' \in V.$$

Proof of Theorem 13. By the definition of $\cong^w_{\mathrm{I}}$, it is clear that if $X \cong^w_{\mathrm{I}} Y$, then $d_{\mathcal{N}}(X,Y) = 0$, i.e. $X \cong^w_{\mathrm{II}} Y$ (cf. Theorem 12).

Conversely, suppose $d_{\mathcal{N}}(X,Y) = 0$. Our strategy is to obtain a set $Z \subseteq X \times Y$ with canonical projection maps $\pi_X : Z \to X, \pi_Y : Z \to Y$ and surjections $\psi_X : X \to \pi_X(Z), \psi_Y : Y \to \pi_Y(Z)$ as in the following diagram:

$$\begin{tikzcd}
& X \arrow[dl, "\mathrm{id}_X"'] \arrow[dr, "\psi_X"] & & Z \arrow[dl, "\pi_X"'] \arrow[dr, "\pi_Y"] & & Y \arrow[dl, "\psi_Y"'] \arrow[dr, "\mathrm{id}_Y"] & \\
X & \cong^w_{\mathrm{I}} & \pi_X(Z) & \cong^w_{\mathrm{I}} & \pi_Y(Z) & \cong^w_{\mathrm{I}} & Y
\end{tikzcd}$$

Furthermore, we will require:

$$\omega_X(\pi_X(z), \pi_X(z')) = \omega_Y(\pi_Y(z), \pi_Y(z')) \quad \text{for all } z, z' \in Z, \tag{2}$$

$$\omega_X(x, x') = \omega_X(\psi_X(x), \psi_X(x')) \quad \text{for all } x, x' \in X, \tag{3}$$

$$\omega_Y(y, y') = \omega_Y(\psi_Y(y), \psi_Y(y')) \quad \text{for all } y, y' \in Y. \tag{4}$$

As a consequence, we will obtain a chain of Type I weak isomorphisms

$$X \cong^w_{\mathrm{I}} \pi_X(Z) \cong^w_{\mathrm{I}} \pi_Y(Z) \cong^w_{\mathrm{I}} Y.$$

Since Type I weak isomorphism is an equivalence relation (Proposition 10), it will follow that X and Y are Type I weakly isomorphic.

By applying Theorem 30, we choose sequences of finite subnetworks $\{X_n \subseteq X : n \in \mathbb{N}\}$ and $\{Y_n \subseteq Y : n \in \mathbb{N}\}$ such that $d_{\mathcal{N}}(X_n, X) < 1/n$ and $d_{\mathcal{N}}(Y_n, Y) < 1/n$ for each $n \in \mathbb{N}$. By the triangle inequality, $d_{\mathcal{N}}(X_n, Y_n) < 2/n$ for each n.

For each $n \in \mathbb{N}$, let $T_n \in \mathscr{R}(X_n, X), P_n \in \mathscr{R}(Y, Y_n)$ be such that $\operatorname{dis}(T_n) < 2/n$ and $\operatorname{dis}(P_n) < 2/n$. Define $\alpha_n := 4/n - \operatorname{dis}(T_n) - \operatorname{dis}(P_n)$, and notice that $\alpha_n \to 0$ as $n \to \infty$. Since $d_{\mathcal{N}}(X,Y) = 0$ by assumption, for each $n \in \mathbb{N}$ we let $S_n \in \mathscr{R}(X,Y)$ be such that $\operatorname{dis}(S_n) < \alpha_n$. Then,

$$\operatorname{dis}(T_n \circ S_n \circ P_n) \leqslant \operatorname{dis}(T_n) + \operatorname{dis}(S_n) + \operatorname{dis}(P_n) < 4/n. \qquad \text{(cf. Remark 4)}$$

Then for each $n \in \mathbb{N}$, we define $R_n := T_n \circ S_n \circ P_n \in \mathscr{R}(X_n, Y_n)$. By Remark 4, we know that R_n has the following expression:

$$R_n = \{(x_n, y_n) \in X_n \times Y_n : \text{ there exist } \tilde{x} \in X,\ \tilde{y} \in Y \text{ such that } (x_n, \tilde{x}) \in T_n, \\ (\tilde{x}, \tilde{y}) \in S_n,\ (\tilde{y}, y_n) \in P_n\}.$$

Next define:

$$\mathcal{S} := \left\{ (\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}} \in (X \times Y)^{\mathbb{N}} : (\tilde{x}_n, \tilde{y}_n) \in S_n \text{ for each } n \in \mathbb{N} \right\}.$$

Since X, Y are first countable and compact, the product $X \times Y$ is also first countable and compact, hence sequentially compact. Any sequence in a sequentially compact space has a convergent subsequence, so for convenience, we replace each sequence in $\mathcal{S}$ by a convergent subsequence. Next define:

$$Z := \{(x, y) \in X \times Y : (x, y) \text{ a limit point of some } (\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}} \in \mathcal{S}\}.$$

Claim 3. Z is a closed subspace of $X \times Y$. Hence it is compact and sequentially compact.

The second statement in the claim follows from the first: assuming that Z is a closed subspace of the compact space $X \times Y$, we obtain that Z is compact. Any subspace of a first countable space is first countable, so Z is also first countable. Next, observe that $\pi_X(Z)$ equipped with the subspace topology is compact, because it is a continuous image of a compact space. It is also first countable because it is a subspace of the first countable space X. Furthermore, the restriction of ω_X to $\pi_X(Z)$ is continuous. Thus $\pi_X(Z)$ equipped with the restriction of ω_X is a compact network, and by similar reasoning, we get that $\pi_Y(Z)$ equipped with the restriction of ω_Y is also a compact network.

Proof of Claim 3. We will show that $Z \subseteq X \times Y$ contains all its limit points. Let $(x, y) \in X \times Y$ be a limit point of Z. Let $\{U_n \subseteq X : n \in \mathbb{N}, (x, y) \in U_n\}$ be a countable neighborhood base of (x, y). For each $n \in \mathbb{N}$, the finite intersection $V_n := \cap_{i=1}^{n} U_i$ is an open neighborhood of (x, y), and thus contains a point $(x_n, y_n) \in Z$ that is distinct from (x, y) (by the definition of a limit point). Pick such an (x_n, y_n) for each $n \in \mathbb{N}$. Then $(x_n, y_n)_{n\in\mathbb{N}}$ is a sequence in Z converging to (x, y) such that $(x_n, y_n) \in V_n$ for each $n \in \mathbb{N}$.

For each $n \in \mathbb{N}$, note that because $(x_n, y_n) \in Z$ and V_n is an open neighborhood of (x_n, y_n), there exists a sequence in $\mathcal{S}$ converging to (x_n, y_n) for which all but finitely many terms are contained in V_n. So for each $n \in \mathbb{N}$, let $(\tilde{x}_n, \tilde{y}_n) \in S_n$ be such that $(\tilde{x}_n, \tilde{y}_n) \in V_n$. Then the sequence $(\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}} \in \mathcal{S}$ converges to (x, y). Thus $(x, y) \in Z$. Since (x, y) was an arbitrary limit point of Z, it follows that Z is closed. ■

Proof of Equation 2. We now prove Equation 2. Let $z = (x, y),\ z' = (x', y') \in Z$, and let $(\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}}, (\tilde{x}'_n, \tilde{y}'_n)_{n\in\mathbb{N}}$ be elements of $\mathcal{S}$ that converge to $(x, y), (x', y')$ respectively. We wish to show $|\omega_X(x, x') - \omega_Y(y, y')| = 0$. Let $\varepsilon > 0$, and observe that:

$$\begin{aligned} &|\omega_X(x, x') - \omega_Y(y, y')| \\ &= |\omega_X(x, x') - \omega_X(\tilde{x}_n, \tilde{x}'_n) + \omega_X(\tilde{x}_n, \tilde{x}'_n) - \omega_Y(\tilde{y}_n, \tilde{y}'_n) + \omega_Y(\tilde{y}_n, \tilde{y}'_n) - \omega_Y(y, y')| \\ &\leqslant |\omega_X(x, x') - \omega_X(\tilde{x}_n, \tilde{x}'_n)| + |\omega_X(\tilde{x}_n, \tilde{x}'_n) - \omega_Y(\tilde{y}_n, \tilde{y}'_n)| + |\omega_Y(\tilde{y}_n, \tilde{y}'_n) - \omega_Y(y, y')|. \end{aligned}$$

Claim 4. Suppose we are given sequences $(\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}}, (\tilde{x}'_n, \tilde{y}'_n)_{n\in\mathbb{N}}$ in Z converging to (x, y) and (x', y') in Z, respectively. Then there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have:

$$|\omega_X(x, x') - \omega_X(\tilde{x}_n, \tilde{x}'_n)| < \varepsilon/4, \qquad |\omega_Y(\tilde{y}_n, \tilde{y}'_n) - \omega_Y(y, y')| < \varepsilon/4.$$

Proof of Claim 4. Write $a := \omega_X(x, x'), b := \omega_Y(y, y')$. Since $\omega_X,\ \omega_Y$ are continuous, we know that $\omega_X^{-1}[B(a, \varepsilon/4)]$ and $\omega_Y^{-1}[B(b, \varepsilon/4)]$ are open neighborhoods of (x, x') and (y, y'). Since each open set in the product space $X \times X$ is a union of open rectangles of the form $A \times A'$ for A, A' open subsets of X, we choose an open set $A \times A' \subseteq \omega_X^{-1}[B(a, \varepsilon/4)]$ such that $(x, x') \in A \times A'$. Similarly, we choose an open set $B \times B' \subseteq \omega_Y^{-1}[B(b, \varepsilon/4)]$ such that $(y, y') \in B \times B'$. Then $A \times B,\ A' \times B'$ are open neighborhoods of $(x, y),\ (x', y')$ respectively. Since $(\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}}$ and

$(\tilde{x}'_n, \tilde{y}'_n)_{n\in\mathbb{N}}$ converge to (x, y) and (x', y'), respectively, we choose $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $(\tilde{x}_n, \tilde{y}_n) \in A \times B$ and $(\tilde{x}'_n, \tilde{y}'_n) \in A' \times B'$. The claim now follows. ∎

Now choose $N \in \mathbb{N}$ such that the property in Claim 4 is satisfied, as well as the additional property that $8/N < \varepsilon/4$. Then for any $n \geqslant N$, we have:

$$|\omega_X(x, x') - \omega_Y(y, y')| \leqslant \varepsilon/4 + |\omega_X(\tilde{x}_n, \tilde{x}'_n) - \omega_Y(\tilde{y}_n, \tilde{y}'_n)| + \varepsilon/4.$$

Separately note that for each $n \in \mathbb{N}$, having $(\tilde{x}_n, \tilde{y}_n), (\tilde{x}'_n, \tilde{y}'_n) \in S_n$ implies that there exist (x_n, y_n) and $(x'_n, y'_n) \in R_n$ such that $(x_n, \tilde{x}_n), (x'_n, \tilde{x}'_n) \in T_n$ and $(\tilde{y}_n, y_n), (\tilde{y}'_n, y'_n) \in P_n$. Thus we can bound the middle term above as follows:

$$\begin{aligned}
&|\omega_X(\tilde{x}_n, \tilde{x}'_n) - \omega_Y(\tilde{y}_n, \tilde{y}'_n)| \\
&= |\omega_X(\tilde{x}_n, \tilde{x}'_n) - \omega_X(x_n, x'_n) + \omega_X(x_n, x'_n) - \omega_Y(y_n, y'_n) + \omega_Y(y_n, y'_n) - \omega_Y(\tilde{y}_n, \tilde{y}'_n)| \\
&\leqslant |\omega_X(\tilde{x}_n, \tilde{x}'_n) - \omega_X(x_n, x'_n)| + |\omega_X(x_n, x'_n) - \omega_Y(y_n, y'_n)| + |\omega_Y(y_n, y'_n) - \omega_Y(\tilde{y}_n, \tilde{y}'_n)| \\
&\leqslant \operatorname{dis}(T_n) + \operatorname{dis}(R_n) + \operatorname{dis}(P_n) < 8/n \leqslant 8/N < \varepsilon/4.
\end{aligned}$$

The preceding calculations show that:

$$|\omega_X(x, x') - \omega_Y(y, y')| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that $\omega_X(x, x') = \omega_Y(y, y')$. This proves Equation 2.

It remains to define surjective maps $\psi_X : X \to \pi_X(Z), \psi_Y : Y \to \pi_Y(Z)$ and to verify Equations 3 and 4. Both cases are similar, so we only show the details of constructing ψ_X and verifying Equation 3.

Construction of ψ_X. Let $x \in X$. Suppose first that $x \in \pi_X(Z)$. Then we simply define $\psi_X(x) = x$. We also make the following observation, to be used later: for each $n \in \mathbb{N}$, letting $y \in Y$ be such that $(x, y) \in S_n$, there exists $x_n \in X_n$ and $y_n \in Y_n$ such that $(x_n, x) \in T_n$ and $(y, y_n) \in P_n$.

Next suppose $x \in X \backslash \pi_X(Z)$. For each $n \in \mathbb{N}$, let $x_n \in X_n$ be such that $(x_n, x) \in T_n$, and let $\tilde{x}_n \in X$ be such that $(x_n, \tilde{x}_n) \in T_n$. Also for each $n \in \mathbb{N}$, let $\tilde{y}_n \in Y$ be such that $(\tilde{x}_n, \tilde{y}_n) \in S_n$. Then for each $n \in \mathbb{N}$, let $y_n \in Y_n$ be such that $(\tilde{y}_n, y_n) \in P_n$. Then by sequential compactness of $X \times Y$, the sequence $(\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}}$ has a convergent subsequence which belongs to $\mathcal{S}$ and converges to a point $(\tilde{x}, \tilde{y}) \in Z$. In particular, we obtain a sequence $(\tilde{x}_n)_{n\in\mathbb{N}}$ converging to a point $\tilde{x}$, such that (x_n, x) and $(x_n, \tilde{x}_n) \in T_n$ for each $n \in \mathbb{N}$. Define $\psi_X(x) = \tilde{x}$.

Since $x \in X$ was arbitrary, this construction defines $\psi_X : X \to \pi_X(Z)$. Note that ψ_X is simply the identity on $\pi_X(Z)$, hence is surjective.

Proof of Equation 3. Now we verify Equation 3. Let $\varepsilon > 0$. There are three cases to check:

Case 1: $x, x' \in \pi_X(Z)$: In this case, we have:

$$|\omega_X(x, x') - \omega_X(\psi_X(x), \psi_X(x'))| = \omega_X(x, x') - \omega_X(x, x') = 0.$$

Case 2: $x, x' \in X \backslash \pi_X(Z)$: By continuity of ω_X, we obtain an open neighborhood $U := \omega_X^{-1}[B(\omega_X(\psi_X(x), \psi_X(x')), \varepsilon/2)]$ of (x, x'). By the definition of ψ_X on $X \backslash \pi_X(Z)$, we obtain sequences $(\tilde{x}_n, \tilde{y}_n)_{n\in\mathbb{N}}$ and $(\tilde{x}'_n, \tilde{y}'_n)_{n\in\mathbb{N}}$ in $\mathcal{S}$ converging to $(\psi_X(x), \tilde{y})$ and $(\psi_X(x'), \tilde{y}')$ for some $\tilde{y}, \tilde{y}' \in Y$. By applying Claim 4, we obtain $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $(\tilde{x}_n, \tilde{x}'_n) \in U$. Note that we also obtain sequences $(x_n)_{n\in\mathbb{N}}$ and $(x'_n)_{n\in\mathbb{N}}$ such that $(x_n, x), (x_n, \tilde{x}_n) \in T_n$ and $(x'_n, x'), (x'_n, \tilde{x}'_n) \in T_n$. Choose N large enough so that it satisfies the property above and also that $4/N < \varepsilon/2$. Then for any $n \geqslant N$,

$$\begin{aligned}
&|\omega_X(x,x') - \omega_X(\psi_X(x),\psi_X(x'))| \\
&= |\omega_X(x,x') - \omega_X(x_n,x'_n) + \omega_X(x_n,x'_n) - \omega_X(\tilde{x}_n,\tilde{x}'_n) + \omega_X(\tilde{x}_n,\tilde{x}'_n) - \omega_X(\psi_X(x),\psi_X(x'))| \\
&\leqslant \operatorname{dis}(T_n) + \operatorname{dis}(T_n) + \varepsilon/2 < 4/n + \varepsilon/2 \leqslant 4/N + \varepsilon/2 < \varepsilon.
\end{aligned}$$

Case 3: $x \in \pi_X(Z), x' \in X\backslash\pi_X(Z)$**:** By the definition of ψ_X on $X\backslash\pi_X(Z)$, we obtain: (1) a sequence $(\tilde{x}'_n)_{n\in\mathbb{N}}$ converging to $\psi_X(x')$, and (2) another sequence $(x'_n)_{n\in\mathbb{N}}$ such that (x'_n, x') and $(x'_n, \tilde{x}'_n)$ both belong to T_n, for each $n \in \mathbb{N}$. By the definition of ψ_X on $\pi_X(Z)$, we obtain a sequence $(x_n)_{n\in\mathbb{N}}$ such that $(x_n, x) \in T_n$ for each $n \in \mathbb{N}$.

Let $U := \omega_X^{-1}[B(\omega_X(x,\psi_X(x')),\varepsilon/2)]$. Since $(\tilde{x}'_n)_{n\in\mathbb{N}}$ converges to $\psi_X(x')$, we know that all but finitely many terms of the sequence $(x,\tilde{x}'_n)_{n\in\mathbb{N}}$ belong to U. So we choose N large enough so that for each $n \geqslant N$, we have:

$$\begin{aligned}
&|\omega_X(x,x') - \omega_X(x,\psi_X(x'))| \\
&= |\omega_X(x,x') - \omega_X(x_n,x'_n) + \omega_X(x_n,x'_n) - \omega_X(x,\tilde{x}'_n) + \omega_X(x,\tilde{x}'_n) - \omega_X(x,\psi_X(x'))| \\
&\leqslant \operatorname{dis}(T_n) + \operatorname{dis}(T_n) + \varepsilon/2 < 4/n + \varepsilon/2 \leqslant 4/N + \varepsilon/2 < \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, Equation 3 follows. The construction of ψ_Y and proof for Equation 4 are similar. This concludes the proof of the theorem. ☐

As a consequence of Theorem 13, we see that weak isomorphisms of Types I and II coincide in the setting of $\mathcal{CN}$. Thus we recover a desirable notion of equivalence in the setting of compact networks.

4. Invariants of networks

At this point, we have computed $d_\mathcal{N}$ between several examples of networks, as in Example 14 and Remark 15. We also asserted in Remark 18 that $d_\mathcal{N}$ is in general difficult to compute. The solution we propose is to compute *quantitatively stable invariants* of networks, and compare the invariants instead of comparing the networks directly. In this section, we restrict our attention to computing invariants of compact networks, which satisfy the useful property that the images of the weight functions are compact.

Intuitively, the invariants that we associate to two strongly isomorphic networks should be the same. We define an $\mathbb{R}$-invariant of networks to be a map $\iota : \mathcal{CN} \to \mathbb{R}$ such that for any $X, Y \in \mathcal{CN}$, if $X \cong^s Y$ then $\iota(X) = \iota(Y)$. Any $\mathbb{R}$-invariant is an example of a *pseudometric* (and in particular, a *metric*) *space valued invariant*, which we define next. Recall that a pseudometric space (V, d_V) is a metric space where we allow $d_V(v, v') = 0$ even if $v \neq v'$.

Definition 17. Let (V, d_V) be any metric or pseudometric space. A V*-valued invariant* is any map $\iota : \mathcal{CN} \to V$ such that $\iota(X,\omega_X) = \iota(Y,\omega_Y)$ whenever $X \cong^s Y$.

Recall that $\mathcal{P}(\mathbb{R})$, the nonempty elements of the power set of $\mathbb{R}$, is a pseudometric space when endowed with the Hausdorff distance [7, Proposition 7.3.3].

In what follows, we will construct several maps and claim that they are pseudometric space valued invariants; this claim will be substantiated in Proposition 38. We will eventually prove that our proposed invariants are *quantitatively stable*. This notion is made precise in §4.1.

Example 33. Define the *diameter* map to be the map

$$\operatorname{diam} : \mathcal{CN} \to \mathbb{R} \text{ given by } (X,\omega_X) \mapsto \max_{x,x' \in X} |\omega_X(x,x')|.$$

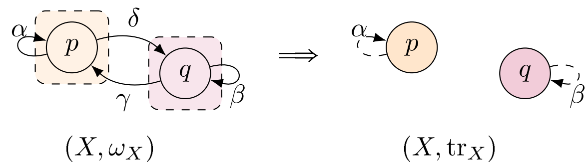

FIGURE 6. The trace map erases data between pairs of nodes.

Then diam is an $\mathbb{R}$-invariant. Observe that the maximum is achieved for $(X, \omega_X) \in \mathcal{CN}$ because X (hence $X \times X$) is compact and $\omega_X : X \times X \to \mathbb{R}$ is continuous. An application of diam to Example 7 gives an upper bound on $d_{\mathcal{N}}(X, Y)$ for $X, Y \in \mathcal{CN}$ in the following way:

$$d_{\mathcal{N}}(X, Y) \leqslant d_{\mathcal{N}}(X, N_1(0)) + d_{\mathcal{N}}(N_1(0), Y) = \tfrac{1}{2}(\mathrm{diam}(X) + \mathrm{diam}(Y)) \text{ for any } X, Y \in \mathcal{CN}.$$

Example 34. Define the *spectrum* map

$$\mathrm{spec} : \mathcal{CN} \to \mathcal{P}(\mathbb{R}) \text{ by } (X, \omega_X) \mapsto \{\omega_X(x, x') : x, x' \in X\}.$$

The spectrum also has two local variants. Define the out-local spectrum of X by $x \mapsto \mathrm{spec}_X^{\mathrm{out}}(x) := \{\omega_X(x, x'),\ x' \in X\}$. Notice that $\mathrm{spec}(X) = \bigcup_{x \in X} \mathrm{spec}_X^{\mathrm{out}}(x)$ for any network X, thus justifying the claim that this construction localizes spec. Similarly, we define the in-spectrum of X as the map $x \mapsto \mathrm{spec}_X^{\mathrm{in}}(x) := \{\omega_X(x', x) : x' \in X\}$. Notice that one still has $\mathrm{spec}(X) = \bigcup_{x \in X} \mathrm{spec}_X^{\mathrm{in}}(x)$ for any network X. Finally, we observe that the two local versions of spec do not necessarily coincide in an asymmetric network.

The spectrum is closely related to the *multisets* used by Boutin and Kemper [5] to produce invariants of weighted undirected graphs. For an undirected graph G, they considered the collection of all subgraphs with three nodes, along with the edge weights for each subgraph (compare to our notion of spectrum). Then they proved that the distribution of edge weights of these subgraphs is an invariant when G belongs to a certain class of graphs.

Example 35. Define the *trace* map $\mathrm{tr} : \mathcal{CN} \to \mathcal{P}(\mathbb{R})$ by $(X, \omega_X) \mapsto \mathrm{tr}(X) := \{\omega_X(x, x) : x \in X\}$. This also defines an associated map $x \mapsto \mathrm{tr}_X(x) := \omega_X(x, x)$. An example is provided in Figure 6: in this case, we have $(X, \mathrm{tr}_X) = (\{p, q\}, (\alpha, \beta))$.

Example 36 (The out and in maps)**.** Let $(X, \omega_X) \in \mathcal{CN}$, and let $x \in X$. Now define $\mathrm{out} : \mathcal{CN} \to \mathcal{P}(\mathbb{R})$ and $\mathrm{in} : \mathcal{CN} \to \mathcal{P}(\mathbb{R})$ by

$$\mathrm{out}(X) = \left\{\max_{x' \in X} |\omega_X(x, x')| : x \in X\right\} \qquad \text{for all } (X, \omega_X) \in \mathcal{CN}$$
$$\mathrm{in}(X) = \left\{\max_{x' \in X} |\omega_X(x', x)| : x \in X\right\} \qquad \text{for all } (X, \omega_X) \in \mathcal{CN}.$$

For each $x \in X$, $\max_{x' \in X} |\omega_X(x, x')|$ and $\max_{x' \in X} |\omega_X(x', x)|$ are achieved because $\{x\} \times X$ and $X \times \{x\}$ are compact. We also define the associated maps out_X and in_X by writing, for any $(X, \omega_X) \in \mathcal{CN}$ and any $x \in X$,

$$\mathrm{out}_X(x) = \max_{x' \in X} |\omega_X(x, x')| \qquad \mathrm{in}_X(x) = \max_{x' \in X} |\omega_X(x', x)|.$$

To see how these maps operate on a network, let $X = \{p, q, r\}$ and consider the weight matrix $\Sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 5 \end{smallmatrix}\right)$. The network corresponding to this matrix is shown in Figure 7. We ascertain the

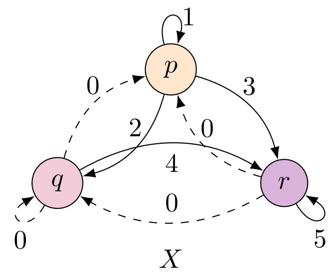

FIGURE 7. The out map applied to each node yields the greatest weight of an arrow *leaving* the node, and the in map returns the greatest weight *entering* the node.

following directly from the matrix:

$$\begin{aligned} \mathrm{out}_X(p) &= 3 \qquad & \mathrm{in}_X(p) &= 1 \\ \mathrm{out}_X(q) &= 4 \qquad & \mathrm{in}_X(q) &= 2 \\ \mathrm{out}_X(r) &= 5 \qquad & \mathrm{in}_X(r) &= 5. \end{aligned}$$

So the out map returns the maximum (absolute) value in each row, and the in map pulls out the maximum (absolute) value in each column of the weight matrix. As in the preceding example, we may use the Hausdorff distance to compare the images of networks under the out and in maps.

Constructions similar to out and in have been used by Jon Kleinberg to study the problem of searching the World Wide Web for user-specified queries [50]. In Kleinberg's model, for a search query σ, *hubs* are pages that point to highly σ-relevant pages (compare to $\mathrm{out}_\bullet$), and *authorities* are pages that are pointed to by pages that have a high σ-relevance (compare to $\mathrm{in}_\bullet$). Good hubs determine good authorities, and good authorities turn out to be good search results.

Example 37 ($\mathrm{min\text{-}out}$ and $\mathrm{min\text{-}in}$)**.** Define the maps $\mathrm{m}^{\mathrm{out}} : \mathcal{CN} \to \mathbb{R}$ and $\mathrm{m}^{\mathrm{in}} : \mathcal{CN} \to \mathbb{R}$ by

$$\begin{aligned} \mathrm{m}^{\mathrm{out}}((X,\omega_X)) &= \min_{x\in X} \mathrm{out}_X(x) \qquad \text{for all } (X,\omega_X) \in \mathcal{CN} \\ \mathrm{m}^{\mathrm{in}}((X,\omega_X)) &= \min_{x\in X} \mathrm{in}_X(x) \qquad \text{for all } (X,\omega_X) \in \mathcal{CN}. \end{aligned}$$

Then both m^{in} and $\mathrm{m}^{\mathrm{out}}$ are $\mathbb{R}$-invariants. We take the minimum when defining $\mathrm{m}^{\mathrm{out}}, \mathrm{m}^{\mathrm{in}}$ because for any network (X,ω_X), we have $\max_{x\in X} \mathrm{out}_X(x) = \max_{x\in X} \mathrm{in}_X(x) = \mathrm{diam}(X)$. Also observe that the minima are achieved above because X is compact.

Proposition 38. *The maps* $\mathrm{out}, \mathrm{in}, \mathrm{tr}, \mathrm{spec}$, *and* $\mathrm{spec}^\bullet$ *are* $\mathcal{P}(\mathbb{R})$*-invariants. Similarly,* $\mathrm{diam}, \mathrm{m}^{\mathrm{out}}$, *and* m^{in} *are* $\mathbb{R}$*-invariants.*

We introduce some notation before presenting the next invariant. For a sequence $(x_i)_{i=1}^n$ of nodes in a network X, we will denote the associated weight matrix by $(\!(\omega_X(x_i,x_j))\!)_{i,j=1}^n$. Entry (i,j) of this matrix is simply $\omega_X(x_i,x_j)$.

Definition 18 (Motif sets)**.** For each $n \in \mathbb{N}$ and each $X \in \mathcal{CN}$, define $\Psi_X^n : X^n \to \mathbb{R}^{n\times n}$ to be the map $(x_1,\cdots,x_n) \mapsto (\!(\omega_X(x_i,x_j))\!)_{i,j=1}^n$. Note that Ψ_X^n is simply a map that sends each sequence of length n to its corresponding weight matrix. Let $\mathcal{C}(\mathbb{R}^{n\times n})$ denote the closed subsets of $\mathbb{R}^{n\times n}$. Then let $\mathrm{M}_n : \mathcal{CN} \to \mathcal{C}(\mathbb{R}^{n\times n})$ denote the map defined by

$$(X,\omega_X) \mapsto \{\Psi_X^n(x_1,\ldots,x_n) : x_1,\ldots,x_n \in X\}.$$

We refer to $\mathrm{M}_n(X)$ as the n-motif set of X. Notice that the image of M_n is closed in $\mathbb{R}^{n\times n}$ because each coordinate is the continuous image of the compact set $X \times X$ under ω_X, hence the image of M_n is compact in $\mathbb{R}^{n\times n}$ and hence closed.

Notice that for $X \in \mathcal{FN}$ and for a fixed $n \in \mathbb{N}$, the set $\mathrm{M}_n(X)$ is a *finite* subset of $\mathbb{R}^{n\times n}$. The interpretation is that $\mathrm{M}_n(X)$ is a bag containing all the motifs of X that one can form by looking at all subnetworks of size n (with repetitions).

Example 39. For the networks from Example 1, we have $\mathrm{M}_1(N_2(\Omega)) = \{\alpha, \beta\}$ and

$$\mathrm{M}_2(N_2(\Omega)) = \left\{ \left(\begin{smallmatrix} \alpha & \alpha \\ \alpha & \alpha \end{smallmatrix}\right), \left(\begin{smallmatrix} \beta & \beta \\ \beta & \beta \end{smallmatrix}\right), \left(\begin{smallmatrix} \alpha & \delta \\ \gamma & \beta \end{smallmatrix}\right), \left(\begin{smallmatrix} \beta & \gamma \\ \delta & \alpha \end{smallmatrix}\right) \right\}, \qquad \mathrm{M}_2(N_1(\alpha)) = \left\{\left(\begin{smallmatrix} \alpha & \alpha \\ \alpha & \alpha \end{smallmatrix}\right)\right\}.$$

Remark 40. Our definition of motif sets is inspired by a definition made by Gromov, termed "curvature classes," in the context of compact metric spaces [42, 3.27].

Definition 19 (Motif sets are metric space valued invariants)**.** Our use of motif sets is motivated by the following observation, which appeared in [61, Section 5]. For any $n \in \mathbb{N}$, let $\mathcal{C}(\mathbb{R}^{n\times n})$ denote the set of closed subsets of $\mathbb{R}^{n\times n}$. Under the Hausdorff distance induced by the ℓ^∞ metric on $\mathbb{R}^{n\times n}$, this set becomes a valid metric space [7, Proposition 7.3.3]. The motif sets defined in Definition 18 define a metric space valued invariant as follows: for each $n \in \mathbb{N}$, let $\mathrm{M}_n : \mathcal{CN} \to \mathcal{C}(\mathbb{R}^{n\times n})$ be the map $X \mapsto \mathrm{M}_n(X)$. We call this the *motif set invariant*. So for $(X, \omega_X), (Y, \omega_Y) \in \mathcal{CN}$, for each $n \in \mathbb{N}$, we let $(Z, d_Z) = (\mathbb{R}^{n\times n}, \ell^\infty)$ and consider the following distance between the n-motif sets of X and Y:

$$d_n(\mathrm{M}_n(X), \mathrm{M}_n(Y)) := d_{\mathrm{H}}^Z(\mathrm{M}_n(X), \mathrm{M}_n(Y)).$$

Since d_{H} is a proper distance between closed subsets, $d_n(\mathrm{M}_n(X), \mathrm{M}_n(Y)) = 0$ if and only if $\mathrm{M}_n(X) = \mathrm{M}_n(Y)$.

4.1. **Quantitative stability of invariants of networks.** Let (V, d_V) be a given pseudometric space. The V-valued invariant $\iota : \mathcal{CN} \to V$ is said to be *quantitatively stable* if there exists a constant $L > 0$ such that

$$d_V\big(\iota(X), \iota(Y)\big) \leqslant L \cdot d_{\mathcal{N}_1}(X, Y)$$

for all networks X and Y. The least constant L such that the above holds for all $X, Y \in \mathcal{CN}$ is the Lipschitz constant of ι and is denoted $\mathbf{L}(\iota)$.

Note that by identifying a non-constant quantitatively stable V-valued invariant ι, we immediately obtain a lower bound for the $d_{\mathcal{N}}$ distance between any two compact networks (X, ω_X) and (Y, ω_Y). Furthermore, given a finite family $\iota_\alpha : \mathcal{CN} \to V$, $\alpha \in A$, of non-constant quantitatively stable invariants, we may obtain the following lower bound for the distance between compact networks X and Y:

$$d_{\mathcal{N}}(X, Y) \geqslant \Big(\max_{\alpha\in A} \mathbf{L}(\iota_\alpha)\Big)^{-1} \max_{\alpha \in A} d_V(\iota_\alpha(X), \iota_\alpha(Y)).$$

It is often the case that computing $d_V(\iota(X), \iota(Y))$ is substantially simpler than computing the $d_{\mathcal{N}}$ distance between X and Y (which leads to a possibly NP-hard problem). The invariants described in the previous section are quantitatively stable.

Proposition 41. *The invariants* $\mathrm{diam}, \mathrm{tr}, \mathrm{out}, \mathrm{in}, \mathrm{m}^{\mathrm{out}}$*, and* m^{in} *are quantitatively stable, with Lipschitz constant* $\mathbf{L} = 2$.

Example 42. Proposition 41 provides simple lower bounds for the $d_{\mathcal{N}}$ distance between compact networks. One application is the following: for all networks X and Y, we have $d_{\mathcal{N}}(X, Y) \geqslant \frac{1}{2}\big|\,\mathrm{diam}(X) - \mathrm{diam}(Y)\big|$. For example, for the networks $X = N_2(\left(\begin{smallmatrix} 1 & 5 \\ 2 & 4 \end{smallmatrix}\right))$ and $Y = N_k(\mathbb{1}_{k\times k})$ (the

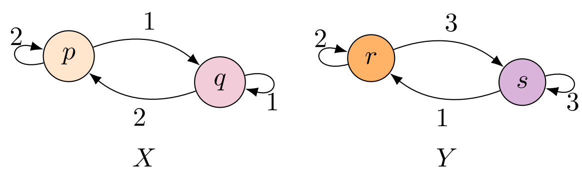

FIGURE 8. Lower-bounding $d_{\mathcal{N}}$ by using global spectra (cf. Example 44).

all-ones matrix—also see Example 1 to recall the N_2 and N_k notation) we have $d_{\mathcal{N}}(X,Y) \geqslant \frac{1}{2}|5-1| = 2$, for all $k \in \mathbb{N}$. For another example, consider the weight matrices

$$\Sigma := \begin{pmatrix} 0 & 5 & 2 \\ 3 & 1 & 4 \\ 1 & 4 & 3 \end{pmatrix} \text{ and } \Sigma' := \begin{pmatrix} 3 & 4 & 2 \\ 3 & 1 & 5 \\ 3 & 3 & 4 \end{pmatrix}.$$

Let $X = N_3(\Sigma)$ and $Y = N_3(\Sigma')$. By comparing the diagonals, we can easily see that $X \not\cong^s Y$, but let's see how the invariants we proposed can help. Note that $\mathrm{diam}(X) = \mathrm{diam}(Y) = 5$, so the lower bound provided by diameter ($\frac{1}{2}|5-5| = 0$) does not help in telling the networks apart. However, $\mathrm{tr}(X) = \{0,1,3\}$ and $\mathrm{tr}(Y) = \{3,1,4\}$, and Proposition 41 then yields

$$d_{\mathcal{N}}(X,Y) \geqslant \frac{1}{2} d_{\mathrm{H}}^{\mathbb{R}}(\{0,1,3\},\{1,3,4\}) = \frac{1}{2}.$$

Consider now the out and in maps. Note that one has $\mathrm{out}(X) = \{5,4\}$, $\mathrm{out}(Y) = \{4,5\}$, $\mathrm{in}(X) = \{3,5,4\}$, and $\mathrm{in}(Y) = \{3,4,5\}$. Then $d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{out}(X),\mathrm{out}(Y)) = 0$, and $d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{in}(X),\mathrm{in}(Y)) = 0$. Thus in both cases, we obtain $d_{\mathcal{N}}(X,Y) \geqslant 0$. So in this particular example, the out and in maps are not useful for obtaining a lower bound to $d_{\mathcal{N}}(X,Y)$ via Proposition 41.

Now we state a proposition regarding the stability of global and local spectrum invariants. These will be of particular interest for computational purposes as we explain in §5.

Proposition 43. *Let* $\mathrm{spec}^{\bullet}$ *refer to either the* out *or* in *version of local spectrum. Then, for all* $(X,\omega_X),(Y,\omega_Y) \in \mathcal{CN}$ *we have*

$$\begin{aligned} d_{\mathcal{N}}(X,Y) &\geqslant \frac{1}{2} \inf_{R \in \mathscr{R}} \sup_{(x,y) \in R} d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X^{\bullet}(x), \mathrm{spec}_Y^{\bullet}(y)) \\ &\geqslant \frac{1}{2} d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}(X), \mathrm{spec}(Y)). \end{aligned}$$

As a corollary, we get $\mathbf{L}(\mathrm{spec}^{\bullet}) = \mathbf{L}(\mathrm{spec}) = 2$.

Example 44 (An application of Proposition 43)**.** Consider the networks in Figure 8. By Proposition 43, we may calculate a lower bound for $d_{\mathcal{N}}(X,Y)$ by simply computing the Hausdorff distance between $\mathrm{spec}(X)$ and $\mathrm{spec}(Y)$, and dividing by 2. In this example, $\mathrm{spec}(X) = \{1,2\}$ and $\mathrm{spec}(Y) = \{1,2,3\}$. Thus $d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}(X),\mathrm{spec}(Y)) = 1$, and $d_{\mathcal{N}}(X,Y) \geqslant \frac{1}{2}$.

Computing the lower bound involving local spectra requires solving a bottleneck linear assignment problem over the set of all correspondences between X and Y. This can be solved in polynomial time; details are provided in §5. The second lower bound stipulates computing the Hausdorff distance on $\mathbb{R}$ between the (global) spectra of X and Y – a computation which can be carried out in (smaller) polynomial time as well.

To conclude this section, we state a theorem asserting that motif sets form a family of quantitatively stable invariants.

Theorem 45. *For each $n \in \mathbb{N}$, M_n is a stable invariant with $\mathbf{L}(\mathrm{M}_n) = 2$.*

Remark 46. While motif sets are of interest as a quantitatively stable network invariant, they also feature in an interesting reconstruction result on metric spaces. In [42, Section $3.27\frac{1}{2}$], Gromov proved that two compact metric spaces are isometric if and only if their motif sets are equal. It turns out that this reconstruction theorem can be strengthened to hold even in the setting of networks, and the proof utilizes Theorem 45. We will release this result in a future publication.

5. Computational aspects

In this section we first discuss some algorithmic details on how to compute the lower bounds for $d_{\mathcal{N}}$ involving local spectra and then present some computational examples. All networks in this section are assumed to be finite. Our software and datasets are available on `https://github.com/fmemoli/PersNet` as part of the `PersNet` software package.

5.1. The complexity of computing $d_{\mathcal{N}}$.

By Remark 15 and Proposition 17 we know that in the setting of finite networks, it is possible to obtain an upper bound on $d_{\mathcal{N}}$, in the case $\mathrm{card}(X) = \mathrm{card}(Y)$, by using $\widehat{d}_{\mathcal{N}}$. Solving for $\widehat{d}_{\mathcal{N}}(X,Y)$ reduces to minimizing the function $\max_{x,x' \in X} f(\varphi)$ over all bijections φ from X to Y. Here $f(\varphi) := \max_{x,x'} |\omega_X(x,x') - \omega_Y(\varphi(x),\varphi(x'))|$. However, this is an instance of an NP-hard problem known as the *quadratic bottleneck assignment problem* [72]. The structure of the optimization problem induced by $d_{\mathcal{N}}$ is very similar to that of $\widehat{d}_{\mathcal{N}}$, so it seems plausible that computing $d_{\mathcal{N}}$ would be NP-hard as well. This intuition is confirmed in Theorem 47. We remark that similar results were obtained for the Gromov-Hausdorff distance by F. Schmiedl in his PhD thesis [82].

Theorem 47. *Computing $d_{\mathcal{N}}$ is NP-hard.*

Proof. To obtain a contradiction, assume that $d_{\mathcal{N}}$ is not NP-hard. Let $X, Y \in \mathcal{FN}^{\mathrm{dis}}$ such that $\mathrm{card}(X) = \mathrm{card}(Y)$. We write $\mathscr{R}(X,Y) = R_B \sqcup R_N$, where R_B consists of correspondences for which the projections π_X, π_Y are injective, thus inducing a bijection between X and Y, and $R_N = \mathscr{R}(X,Y) \backslash R_B$. Note that for any $R \in R_N$, there exist x, x', y such that $(x,y), (x',y) \in R$, or there exist x, y, y' such that $(x,y), (x,y') \in R$. Define $\Psi : \mathbb{R} \to \mathbb{R}$ by:

$$\Psi(\zeta) = \begin{cases} \zeta + C & : \zeta \neq 0 \\ 0 & : \zeta = 0 \end{cases}, \text{ where}$$
$$C = \max_{R \in \mathscr{R}(X,Y)} \mathrm{dis}(R) + 1.$$

For convenience, we will write $\Psi(X), \Psi(Y)$ to mean $(X, \Psi \circ \omega_X)$ and $(Y, \Psi \circ \omega_Y)$ respectively. We will also write:

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

Consider the problem of computing $d_{\mathcal{N}}(\Psi(X), \Psi(Y))$. First observe that for any $R \in R_B$, we have $\mathrm{dis}(R) = \mathrm{dis}_\Psi(R)$. To see this, let $R \in R_B$. Let $(x,y), (x',y') \in R$, and note that $x \neq x', y \neq y'$. Then:

$$|\Psi(\omega_X(x,x')) - \Psi(\omega_Y(y,y'))| = |\omega_X(x,x') + C - \omega_Y(y,y') - C| = |\omega_X(x,x') - \omega_Y(y,y')|.$$

Since $(x,y), (x',y')$ were arbitrary, it follows that $\mathrm{dis}(R) = \mathrm{dis}_\Psi(R)$. This holds for all $R \in R_B$.

On the other hand, let $R \in R_N$. By a previous observation, we assume that there exist x, x', y such that $(x, y), (x', y) \in R$. For such a pair, we have:

$$|\Psi(\omega_X(x, x')) - \Psi(\omega_Y(y, y))| = |\omega_X(x, x') + C - 0| \geqslant \max_{S \in \mathscr{R}(X,Y)} \operatorname{dis}(S) + 1.$$

It follows that $\operatorname{dis}_\Psi(R) > \operatorname{dis}_\Psi(S)$, for any $S \in R_B$. Hence:

$$\begin{aligned} d_{\mathcal{N}}(\Psi(X), \Psi(Y)) &= \frac{1}{2} \min_{R \in \mathscr{R}(X,Y)} \operatorname{dis}_\Psi(R) \\ &= \frac{1}{2} \min_{R \in R_B} \operatorname{dis}_\Psi(R) \\ &= \frac{1}{2} \min_{R \in R_B} \operatorname{dis}(R) \\ &= \frac{1}{2} \min_{\varphi} \operatorname{dis}(\varphi), \text{ where } \varphi \text{ ranges over bijections } X \to Y \\ &= \widehat{d}_{\mathcal{N}}(X, Y). \end{aligned}$$

It is known (see Remark 48 below) that computing $\widehat{d}_{\mathcal{N}}$ is NP-hard. But the preceding calculation shows that $\widehat{d}_{\mathcal{N}}$ can be computed through $d_{\mathcal{N}}$, which, by assumption, is not NP-hard. This is a contradiction. Hence $d_{\mathcal{N}}$ is NP-hard. □

Remark 48. We can be more precise about why computing $\widehat{d}_{\mathcal{N}}$ is a case of the QBAP. Let $X = \{x_1, \ldots, x_n\}$ and let $Y = \{y_1, \ldots, y_n\}$. Let Π denote the set of all $n \times n$ permutation matrices. Note that any $\pi \in \Pi$ can be written as $\pi = ((\pi_{ij}))_{i,j=1}^n$, where each $\pi_{ij} \in \{0, 1\}$. Then $\sum_j \pi_{ij} = 1$ for any i, and $\sum_i \pi_{ij} = 1$ for any j. Computing $\widehat{d}_{\mathcal{N}}$ now becomes:

$$\widehat{d}_{\mathcal{N}}(X, Y) = \frac{1}{2} \min_{\pi \in \Pi} \max_{1 \leqslant i,k,j,l, \leqslant n} \Gamma_{ijkl}\pi_{ij}\pi_{kl}, \text{ where } \Gamma_{ikjl} = |\omega_X(x_i, x_k) - \omega_Y(y_j, y_l)|.$$

This is just the QBAP, which is known to be NP-hard [8].

5.2. **Lower bounds and an algorithm for computing minimum matchings.** Lower bounds for $d_{\mathcal{N}}$ involving the comparison of local spectra of two networks such as those in Proposition 43 require computing the minimum of a functional $J(R) := \max_{(x,y) \in R} C(x, y)$ where $C : X \times Y \to \mathbb{R}_+$ is a given *cost* function and R ranges in $\mathscr{R}(X, Y)$. This is an instance of a bottleneck linear assignment problem (or LBAP) [8]. We remark that the current instance differs from the standard formulation in that one is now optimizing over correspondences and not over permutations. Hence the standard algorithms need to be modified.

Assume $n = \operatorname{card}(X)$ and $m = \operatorname{card}(Y)$. In this section we adopt matrix notation and regard R as a matrix $((r_{i,j})) \in \{0, 1\}^{n \times m}$. The condition $R \in \mathscr{R}(X, Y)$ then requires that $\sum_i r_{i,j} \geqslant 1$ for all j and $\sum_j r_{i,j} \geqslant 1$ for all i. We denote by $C = ((c_{i,j})) \in \mathbb{R}_+^{n \times m}$ the matrix representation of the cost function C described above. With the goal of identifying a suitable algorithm, the key observation is that the optimal value $\min_{R \in \mathscr{R}} J(R)$ must coincide with a value realized in the matrix C.

An algorithm with complexity $O(n^2 \times m^2)$ is the one in Algorithm 1 (we give it in Matlab pseudo-code). The algorithm belongs to the family of *thresholding algorithms* for solving matching problems over permutations, see [8]. Notice that R is a binary matrix and that procedure **Test-Correspondence** has complexity $O(n \times m)$. In the worst case, the matrix C has $n \times m$ distinct entries, and the **while** loop will need to exhaustively test them all, hence the claimed complexity of $O(n^2 \times m^2)$. Even though a more efficient version (with complexity $O((n \times m) \log(n \times m))$ can be obtained by using a bisection strategy on the range of possible values contained in the matrix C (in a manner similar to what is described for the case of permutations in [8]), here for clarity we limit our presentation to the version detailed above.

Algorithm 1 MinMax matching

```
procedure MINMAXMATCH(C)
    v = sort(unique(C(:)));
    k = 1;
    while ~ done do
        c = v(k);
        R = (C <= c);
        done = TESTCORRESPONDENCE(R);
        k = k + 1;
    return c
procedure TESTCORRESPONDENCE(R)
    done = prod(sum(R))*prod(sum(R')) > 0;
    return done
```

5.3. **Computational example: randomly generated networks.** As a first application of our ideas we generated a database of weighted directed networks with different numbers of "communities" and different total cardinalities using the software provided by [32]. Using this software, we generated 35 random networks as follows: 5 networks with 5 communities and 200 nodes each (class `c5-n200`), 5 networks with 5 communities and 100 nodes each (class `c5-n100`), 5 networks with 4 communities and 128 nodes each (class `c4-n128`), 5 networks with 2 communities and 20 nodes each (class `c2-n20`), 5 networks with 1 community and 50 nodes each (class `c1-n50`), and 10 networks with 1 community and 128 nodes each (class `c1-n128`). In order to make the comparison more realistic, as a preprocessing step we divided all the weights in each network by the diameter of the network. In this manner, discriminating between networks requires differentiating their structure and not just the scale of the weights. Note that the (random) weights produced by the software [32] are all non-negative.

Using a Matlab implementation of Algorithm 1 we computed a 35×35 matrix of values corresponding to a lower bound based simultaneously on both the in and out local spectra. This strengthening of Proposition 43 is stated below.

Proposition 49. *For all $X, Y \in \mathcal{FN}$,*

$$d_{\mathcal{N}}(X,Y) \geqslant \frac{1}{2} \min_{R \in \mathscr{R}} \max_{(x,y) \in R} C(x,y), \text{ where}$$

$$C(x,y) = \max\left(d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X^{\mathrm{in}}(x), \mathrm{spec}_Y^{\mathrm{in}}(y)), d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X^{\mathrm{out}}(x), \mathrm{spec}_Y^{\mathrm{out}}(y))\right).$$

This bound follows from Proposition 43 by the discussion at the beginning of §4.1.

The results are shown in the form of the lower bound matrix and its single linkage dendrogram in Figures 9 and 10, respectively. Notice that the labels in the dendrogram permit ascertaining the quality of the classification provided by the local spectra bound. With only very few exceptions, networks with similar structure (same number of communities) were clustered together regardless of their cardinality. Notice furthermore how networks with 4 and 5 communities merge together before merging with networks with 1 and 2 communities, and vice versa. For comparison, we provide details about the performance of the global spectra lower bound on the same database in Figures 10

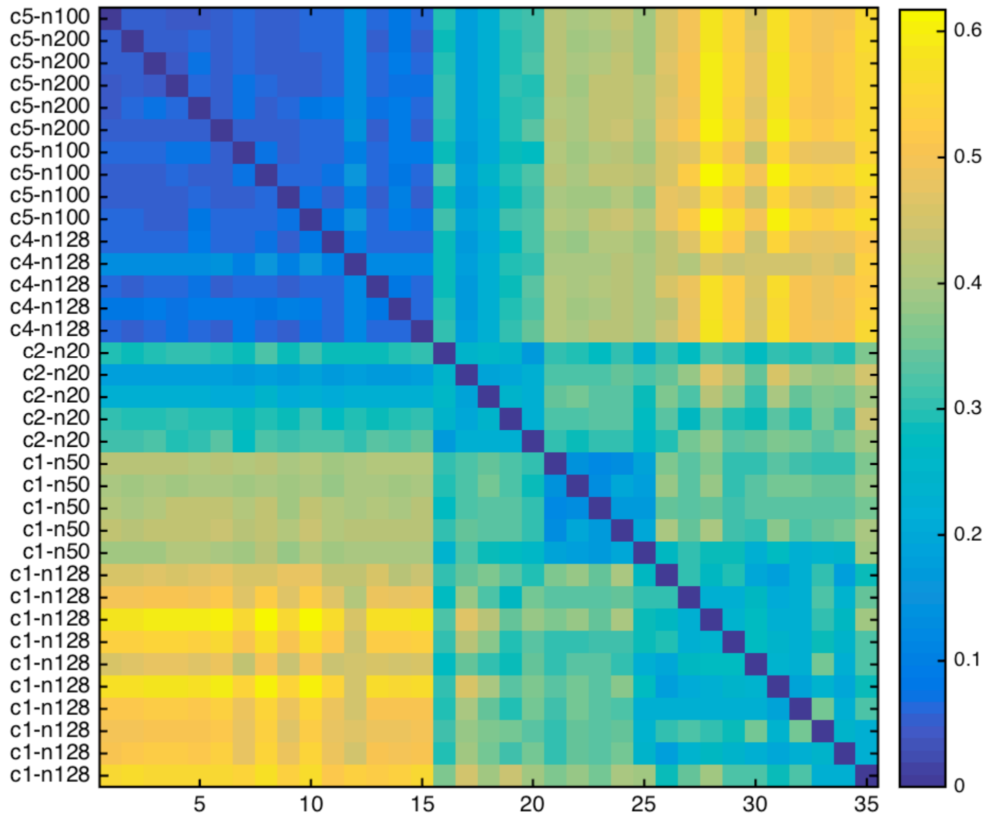

FIGURE 9. Lower bound matrix arising from matching local spectra on the database of community networks. The labels indicate the number of communities and the total number of nodes. Results correspond to using local spectra as described in Proposition 49.

and 12. The results are clearly inferior to those produced by the local version, as predicted by the inequality in Proposition 43.

5.4. **Computational example: simulated hippocampal networks.** A natural observation about humans is that as they navigate an environment, they produce "mental maps" which enable them to recall the features of the environment at a later time. This is also true for other animals with higher cognitive function. In the neuroscience literature, it is accepted that the hippocampus in an animal's brain is responsible for producing a mental map of its physical environment [3, 6]. More specifically, it has been shown that as a rat explores a given environment, specific physical regions ("place fields") become linked to specific "place cells" in the hippocampus [70, 71]. Each place cell shows a spike in neuronal activity when the rat enters its place field, accompanied by a drop in activity as the rat goes elsewhere. In order to understand how the brain processes this data, a natural question to ask is the following: Is the time series data of the place cell activity, often referred to as "spike trains", enough to recover the structure of the environment?

Approaches based on homology [22] and persistent homology [23] have shown that the preceding question admits a positive answer. We were interested in determining if, instead of computing homology groups, we could represent the time series data as networks, and then apply our invariants

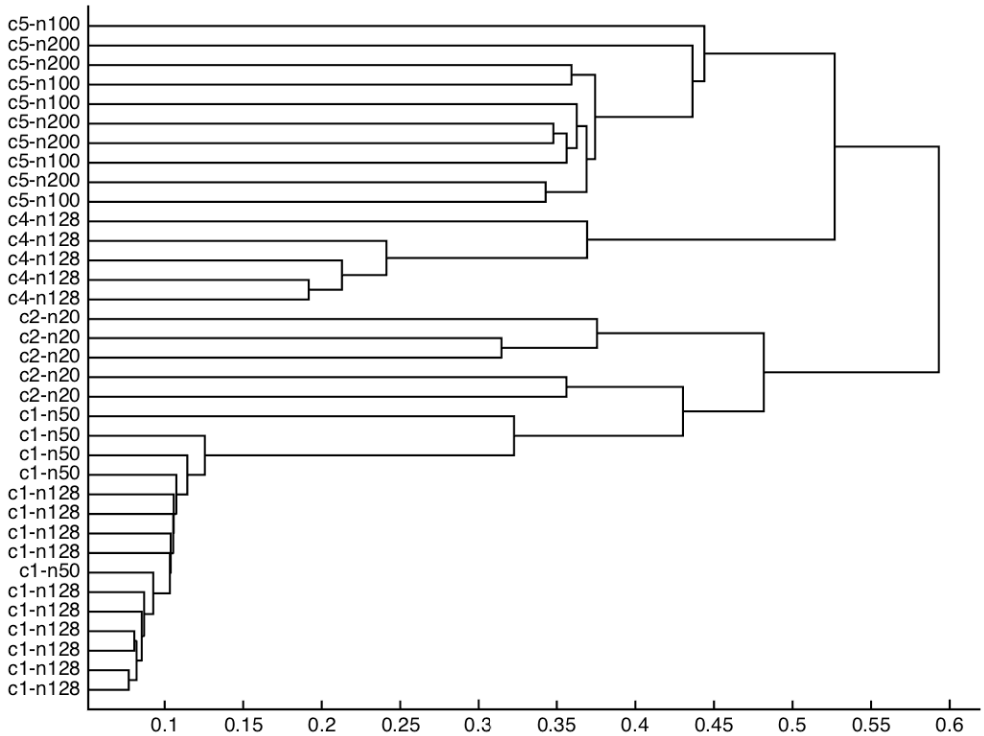

FIGURE 10. Single linkage dendrogram corresponding to the database of community networks. The labels indicate the number of communities and the total number of nodes. Results correspond to using local spectra as described in Proposition 49.

to distinguish between different environments. Our preliminary results on simulated hippocampal data indicate that such may be the case.

In our experiment, there were two environments: (1) a square of side length L, and (2) a square of side length L, with a disk of radius $0.33L$ removed from the center. In what follows, we refer to the environments of the second type as *1-hole environments*, and those of the first type as *0-hole environments*. For each environment, a random-walk trajectory of 5000 steps was generated, where the agent could move above, below, left, or right with equal probability. If one or more of these moves took the agent outside the environment (a disallowed move), then the probabilities were redistributed uniformly among the allowed moves. The length of each step in the trajectory was $0.1L$.

In the first set of 20 trials for each environment, 200 place fields of radius $0.1L$ were scattered uniformly at random. In the next two sets, the place field radii were changed to $0.2L$ and $0.05L$. This produced a total of 60 trials for each environment. For each trial, the corresponding network (X, ω_X) was constructed as follows: X consisted of 200 place cells, and for each $1 \leqslant i, j \leqslant 200$, the weight $\omega_X(x_i, x_j)$ was given by:

$$\omega_X(x_i, x_j) = 1 - \frac{\text{\# times cell } x_j \text{ spiked in a window of five time units after cell } x_i \text{ spiked}}{\text{\# times cell } x_j \text{ spiked}}.$$

The results of applying the local spectra lower bound are shown in Figures 14, 15 and 16. The labels `env-0, env-1` correspond to 0 and 1-hole environments, respectively. Note that with some exceptions, networks corresponding to the same environment are clustered together, regardless of place field radius. Finally, in Figure 13 we present the single linkage dendrogram obtained from

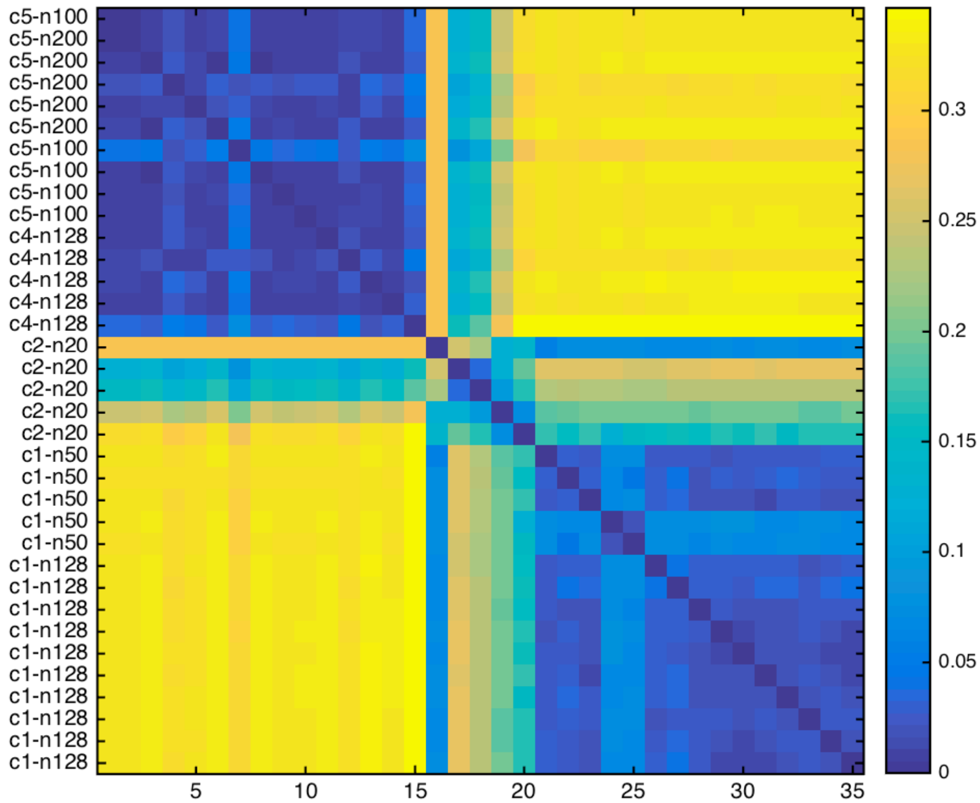

FIGURE 11. Lower bound matrix arising from matching global spectra on the database of community networks. The labels indicate the number of communities and the total number of nodes.

comparing all 120 networks together. In light of these results, we are interested in seeing how these methods can applied to other time series data arising from biology.

As a final remark, we note that it is possible to obtain better clustering on the hippocampal network dataset by using $d_{\mathcal{N}}$-invariants that arise from *persistent homology*. We refer the reader to [17] for details.

6. COMPARISON OF $d_{\mathcal{N}}$ WITH THE CUT METRIC ON GRAPHS

In our work throughout this paper, we have developed the theoretical framework of a certain notion of network distance that has proven to be useful for applying methods from the topological data analysis literature to network data. In each of these applications, networks were modeled as generalizations of metric spaces, and so the appropriate notion of network distance turned out to be a generalization of the Gromov-Hausdorff distance between compact metric spaces. However, an alternative viewpoint would be to model networks as weighted, directed graphs. From this perspective, a well-known metric on the space of all graphs is the *cut metric* [57, 4]. In particular, it is known that the completion of the space of all graphs with respect to the cut metric is compact [57, p. 149]. An analogous result is known for the network distance, and we will establish it in a forthcoming publication. It turns out that there are other structural similarities between the cut metric and the network distance. In this section, we will develop an interpretation of an ℓ^∞

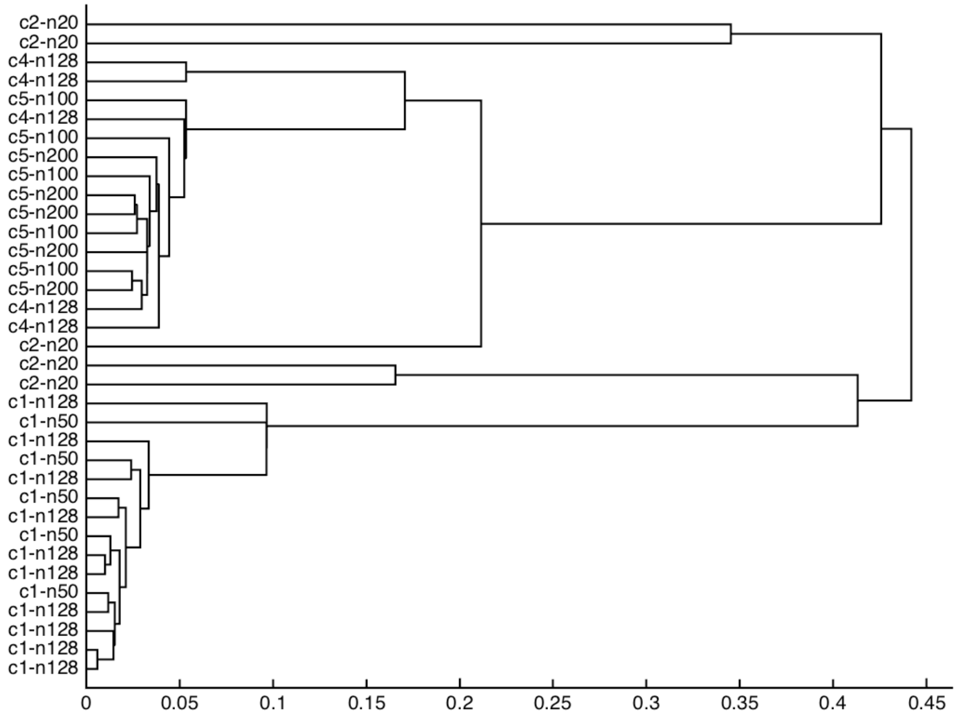

FIGURE 12. Single linkage dendrogram corresponding to the database of community networks. The labels indicate the number of communities and the total number of nodes. Results correspond to using global spectra as signatures.

version of the cut metric in the setting of compact metric spaces, and show that it agrees with the Gromov-Hausdorff distance in this setting.

6.1. **The cut distance between finite graphs.** Let $G = (V, E)$ denote a vertex-weighted and edge-weighted graph on a vertex set $V = \{1, 2, \ldots, n\}$. Let α_i denote the weight of node i, with the assumption that each $\alpha_i \geqslant 0$, and $\sum_i \alpha_i = 1$. Let $\beta_{ij} \in \mathbb{R}$ denote the weight of edge ij. For any $S, T \subseteq V$, define:

$$e_G(S,T) := \sum_{s \in S, t \in T} \alpha_s \alpha_t \beta_{st}.$$

Note for future reference that one may regard e_G as a function from $\mathcal{P}(V) \times \mathcal{P}(V)$ into $\mathbb{R}$, where $\mathcal{P}(S)$ denotes the nonempty elements of the power set of a set S.

Let A be an $n \times n$ matrix of real numbers. Some classical norms include the ℓ^1 norm $\|A\|_1 = n^{-2} \sum_{i,j=1}^n |A_{ij}|$, the ℓ_2 norm $\|A\|_2 = (n^{-2} \sum_{i,j=1}^n |A_{ij}|^2)^{1/2}$, and the ℓ^∞ norm $\|A\|_\infty = \max_{i,j} |A_{ij}|$. Note that the n^{-2} term is included for normalization.

The *cut norm of* A is defined as

$$\|A\|_\square := \frac{1}{n^2} \max_{S,T \subseteq \{1,\ldots,n\}} \Big| \sum_{i \in S, j \in T} A_{ij} \Big|.$$

The *cut metric* or *cut distance* is defined between weighted graphs on the same node set $V = \{1, 2, \ldots, n\}$ as

$$\delta_\square(G, G') := \max_{S,T \subseteq V} |e_G(S,T) - e_{G'}(S,T)|.$$

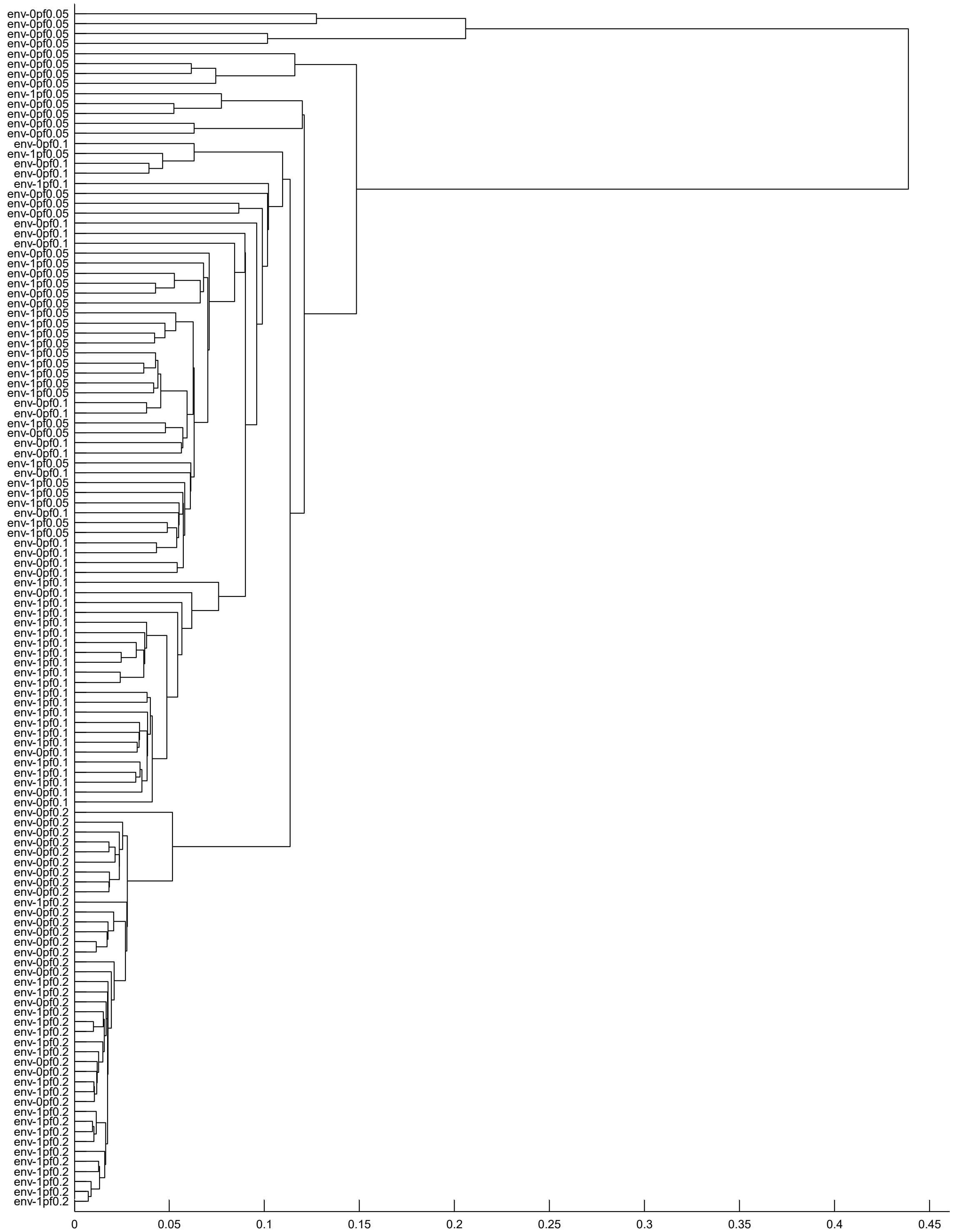

FIGURE 13. Single linkage dendrogram corresponding to 120 hippocampal networks of place field radii $0.05L, 0.1L,$ and $0.2L$. Results are based on the local spectrum lower bound of Proposition 49.

Next we consider weighted graphs with different numbers of nodes. Let G, G' be graphs on n and m nodes respectively, with node weights $(\alpha_i)_{i=1}^n, (\alpha'_k)_{k=1}^m$ and edge weights $(\beta_{ij})_{i,j=1}^n, (\beta'_{kl})_{k,l=1}^m$, respectively. A *fractional overlay* is a non-negative $n \times m$ matrix W such that

$$\sum_{i=1}^n W_{ik} = \alpha'_k \text{ for } 1 \leqslant k \leqslant m, \text{ and } \sum_{k=1}^m W_{ik} = \alpha_i \text{ for } 1 \leqslant i \leqslant n.$$

Define $\mathcal{W}(G, G')$ to be the set of all fractional overlays between G and G'. Let $W \in \mathcal{W}(G, G')$. Consider the graphs $G(W), G'(W)$ on the node set $\{(i,k) : i \leqslant n, k \leqslant m, i, k \in \mathbb{N}\}$ defined in the following way: node (i,k) carries weight W_{ik} in both $G(W), G'(W)$, edge $((i,k),(j,l))$ carries weight β_{ij} in $G(W)$ and β'_{kl} in $G'(W)$. Then the cut distance becomes

$$d_\square(G, G') := \min_{W \in \mathcal{W}(G,G')} \delta_\square(G(X), G'(X)). \tag{$\star$}$$

6.2. **The cut distance and the Gromov-Hausdorff distance.** In our interpretation, a fractional overlay is analogous to a *correspondence*, as defined in §2. We define correspondences between networks, but a similar definition can be made for metric spaces, and in the case of finite metric spaces, a correspondence can be regarded as a binary matrix. Since correspondences are used to define the Gromov-Hausdorff distance between compact metric spaces, and our definition of network distance is motivated by GH distance, we would like to reinterpret the cut distance in the setting of compact metric spaces. Our goal is to show that in this setting, a certain analogue of the cut distance agrees with the GH distance.

For any compact metric space (X, d_X), let $\mathcal{P}(X)$ denote the nonempty elements of the power set of X, and let e_X be any $\mathbb{R}_+$-valued function defined on $\mathcal{P}(X) \times \mathcal{P}(X)$. In analogy with the definition of e_G for graphs, one would like e_X to absorb information about the metric d_X on X.

Given a subset $R \subseteq X \times Y$, let π_1 and π_2 denote the canonical projections to the X and Y coordinates, respectively. Let Ξ denote the map that takes a compact metric space (X, d_X) and returns a function $e_X : \mathcal{P}(X) \times \mathcal{P}(X) \to \mathbb{R}$. We impose the following two conditions on the assignment $d_X \mapsto e_X$ induced by the map Ξ:

(1) For all $x, x' \in X$,
$$\Xi(d_X)(\{x\}, \{x'\}) = d_X(x, x'). \tag{$\#_1$}$$
Thus $e_X = \Xi(d_X)$ and d_X agree on singleton sets.

(2) For all $T, S \subseteq X \times Y$,
$$|\Xi(d_X)(\pi_1(T), \pi_1(S)) - \Xi(d_Y)(\pi_2(T), \pi_2(S))| \leqslant \max_{t \in T, s \in S} |d_X(\pi_1(t), \pi_1(s)) - d_Y(\pi_2(t), \pi_2(s))|. \tag{$\#_2$}$$

The latter of the two conditions above can be viewed as a continuity condition.

Example 50. Some natural candidates for the assignment $d_X \mapsto e_X$ which satisfy the two conditions above are:

- Ξ^{H} such that $e_X(A, B) = d_{\mathrm{H}}^X(A, B)$, the Hausdorff distance,
- $\Xi^{\max}$ such that $e_X(A, B) = \sup_{a \in A, b \in B} d_X(a, b)$,
- $\Xi^{\min}$ such that $e_X(A, B) = \inf_{a \in A, b \in B} d_X(a, b)$.

It is clear that all these satisfy $(\#_1)$. In Proposition 52 we prove that they satisfy condition $(\#_2)$.

An analogue of the cut distance ($\star$) in the setting of compact metric spaces is the following.

Definition 20 (An analogue of the cut distance for compact metric spaces)**.** Let Ξ be any map satisfying ($\#_1$) and ($\#_2$). Then for compact metric spaces X and Y, define

$$d_\square^\Xi(X,Y) := \frac{1}{2}\inf_R \sup_{S,T\subseteq R} |\Xi(d_X)(\pi_1(T),\pi_1(S)) - \Xi(d_Y)(\pi_2(T),\pi_2(S))|,$$

where R ranges over correspondences between X and Y, and π_1, π_2 are the canonical projections from $X \times Y$ onto X and Y respectively. We include the coefficient 2^{-1} to make comparison with d_{GH} simpler. We claim that this interpretation of $d_\square^\Xi$ is identical to d_{GH}.

Note that one definition of the Gromov-Hausdorff distance for compact metric spaces [7, Theorem 7.3.25] is the following:

$$d_{\mathrm{GH}}(X,Y) = \frac{1}{2}\inf_{R\in\mathscr{R}} \sup_{(x,y),(x',y')\in R} |d_X(x,x') - d_Y(y,y')|.$$

We also make the following definitions for *distortion*:

$$\begin{aligned}\mathrm{dis}_{\mathrm{GH}}(R) &= \sup_{(x,y),(x',y')\in R} |d_X(x,x') - d_Y(y,y')|,\\ \mathrm{dis}_\square^\Xi(R) &= \sup_{S,T\subseteq R} |\Xi(d_X)(\pi_1 T,\pi_1 S) - \Xi(d_Y)(\pi_2 T,\pi_2 S)|.\end{aligned}$$

Proposition 51. *For all compact metric spaces X, Y and any assignment Ξ satisfying ($\#_1$) and ($\#_2$) above, one has $d_{GH}(X,Y) = d_\square^\Xi(X,Y)$.*

Proof of Proposition 51. Write $e_X = \Xi(d_X)$ and $e_Y = \Xi(d_Y)$. Let $R \in \mathscr{R}(X,Y)$. In computing $\mathrm{dis}_{\mathrm{GH}}^\Xi(R)$, we take the supremum over all subsets of R, including singletons. Since e_X (resp. e_Y) agrees with d_X (resp. d_Y) on singletons, it follows that $\mathrm{dis}_{\mathrm{GH}}(R) \leqslant \mathrm{dis}_\square^\Xi(R)$. Thus $d_{\mathrm{GH}} \leqslant d_\square^\Xi$.

We now need to show $d_\square^\Xi \leqslant d_{\mathrm{GH}}$.

Let $\eta > 0$ such that $d_{\mathrm{GH}}(X,Y) < \eta$. Then, by one of the characterizations of d_{GH} [7, Chapter 7], there exists a joint-metric δ defined on $X \sqcup Y$ and a correspondence R such that $\delta(x,y) < \eta$ for all $(x,y) \in R$. In particular, δ agrees with d_X and d_Y when restricted to the appropriate spaces. Now we have

$$\begin{aligned}\mathrm{dis}_\square^\Xi(R) &= \sup_{T,S\subseteq R} |e_X(\pi_1(T),\pi_1(S)) - e_Y(\pi_2(T),\pi_2(S))|\\ &\leqslant \sup_{T,S\subseteq R}\sup_{t\in T, s\in S} |d_X(\pi_1(t),\pi_1(s)) - d_Y(\pi_2(t),\pi_2(s))|\\ &= \sup_{T,S\subseteq R}\sup_{t\in T, s\in S} |\delta(\pi_1(t),\pi_1(s)) - \delta(\pi_2(t),\pi_2(s))|\end{aligned}$$

By the triangle inequality, we have the following for any $t \in T$ and $s \in S$:

$$\begin{aligned}|\delta(\pi_1(t),\pi_1(s)) - \delta(\pi_2(t),\pi_2(s))| &\leqslant |\delta(\pi_1(t),\pi_1(s)) - \delta(\pi_1(s),\pi_2(t))|\\ &\qquad + |\delta(\pi_1(s),\pi_2(t)) - \delta(\pi_2(t),\pi_2(s))|\\ &\leqslant |\delta(\pi_1(t),\pi_2(t))| + |\delta(\pi_1(s),\pi_2(s))|\\ &< \eta + \eta = 2\eta. \text{ Thus we conclude}\\ \mathrm{dis}_\square^\Xi(R) &\leqslant 2\eta.\end{aligned}$$

This shows $d_\square^\Xi(X,Y) \leqslant \eta$, and so $d_\square^\Xi(X,Y) \leqslant d_{\mathrm{GH}}(X,Y)$. So we conclude $d_{\mathrm{GH}}(X,Y) = d_\square^\Xi(X,Y)$, under the assumptions made in this discussion. $\square$

Proposition 52. *Each of the maps $\Xi^{\min}$, $\Xi^{\max}$, and Ξ^{H} satisfies condition ($\#_2$).*

Proof. We only give details for Ξ^{H}. The argument for the other cases is similar. Let X, Y be compact metric spaces, and let $T, S \subseteq X \times Y$. First we recall the Hausdorff distance between two closed subsets $E, F \subseteq X$:

$$d_{\mathrm{H}}^X(E,F) = \max\left\{\sup_{e\in E}\inf_{f\in F} d_X(e,f), \sup_{f\in F}\inf_{e\in E} d_X(e,f)\right\}.$$

So d_{H}^X between two sets in X is written as a max of two numbers a and b, and we have the general result

$$|\max(a,b) - \max(a',b')| \leqslant \max(|a-a'|, |b-b'|).$$

Another general result about "calculating" with $\sup$ is that $|\sup f - \sup g| \leqslant \sup |f-g|$ for real valued functions f and g. Both these properties are consequences of the triangle inequality, and we use them here:

$$\begin{aligned}
|d_{\mathrm{H}}^X(\pi_1(T), \pi_1(S)) - d_{\mathrm{H}}^Y(\pi_2(T), \pi_2(S))| &= |\max(a,b) - \max(a',b')| \\
&\leqslant \max(|a-a'|, |b-b'|), \text{ where} \\
a &= \sup_{t\in T}\inf_{s\in S} d_X(\pi_1(t), \pi_1(s)) \\
a' &= \sup_{t\in T}\inf_{s\in S} d_Y(\pi_2(t), \pi_2(s)) \\
b &= \sup_{s\in S}\inf_{t\in T} d_X(\pi_1(t), \pi_1(s)) \\
b' &= \sup_{s\in S}\inf_{t\in T} d_Y(\pi_2(t), \pi_2(s))
\end{aligned}$$

We consider only one of the terms $|a-a'|$; the other term can be treated similarly.

$$\begin{aligned}
|a-a'| &= |\sup_{t\in T}\inf_{s\in S} d_X(\pi_1(t), \pi_1(s)) - \sup_{t\in T}\inf_{s\in S} d_Y(\pi_2(t), \pi_2(s))| \\
&\leqslant \sup_{t\in T} |\inf_{s\in S} d_X(\pi_1(t), \pi_1(s)) - \inf_{s\in S} d_Y(\pi_2(t), \pi_2(s))| \\
&\leqslant \sup_{t\in T, s\in S} |d_X(\pi_1(t), \pi_1(s)) - d_Y(\pi_2(t), \pi_2(s))|.
\end{aligned}$$

The same bound holds for $|b-b'|$. Thus d_{H} satisfies condition ($\#_2$), as claimed. $\square$

7. Discussion

We introduced a model for the space of all networks, and defined two notions of weak isomorphism (Types I and II) between any two networks. We saw immediately that the two types coincide for finite networks, and observed that it is non-trivial to verify that they coincide for compact networks. We proposed a notion of distance compatible with Type II weak isomorphism and verified that our definition actually does induce a (pseudo)metric. In the latter part of the paper, we proved sampling theorems by which one may approximate compact networks, and verified that weak isomorphism of Types I and II coincide for compact networks. To motivate the sampling theorem, we first constructed an explicit example of a compact, asymmetric network—the directed circle—and also provided an explicit family of finite dissimilarity networks that approximate the directed circle up to arbitrary precision.

We constructed multiple quantitatively stable invariants, with examples illustrating their behavior, and quantified their stability. We also returned to the question of computing $d_{\mathcal{N}}(X,Y)$ explicitly, and proved that one way is to compute the related quantity $\widehat{d}_{\mathcal{N}}$. Later in the paper, we observed that this is an instance of the quadratic bottleneck assignment problem, which is NP-hard, and proved that computing $d_{\mathcal{N}}$ is NP-hard as well.

We also compared $d_{\mathcal{N}}$ with an ℓ^{∞} version of the cut metric on graphs. The cut metric is a notion of distance between graphs which has been used for studying the convergence of sequences of graphs to continuous objects. We carried out this comparison in the context of compact metric spaces where $d_{\mathcal{N}}$ boils down to the Gromov-Hausdorff distance and proved that the ℓ^{∞} version of the cut metric agrees with $d_{\mathcal{N}}$ for a large family of parameters.

Finally, we provided an algorithm of complexity $O(n^2 \times m^2)$ that uses local spectra to compute a lower bound for $d_{\mathcal{N}}$, and illustrated our constructions on: (1) a database of random networks with different numbers of nodes and varying community structures, and (2) simulated hippocampal networks with the same number of nodes, where the weights on the networks are assumed to capture a certain notion of shape in the data. We also provided Matlab code and datasets for applying our methods as part of the `PersNet` software package (`https://github.com/fmemoli/PersNet`).

Further results. This paper is the first in a series of two papers (the other being [18]) laying out the theoretical foundations of the network distance $d_{\mathcal{N}}$. We briefly advertise some of the other results regarding $d_{\mathcal{N}}$ that we have released in [18]. It turns out that the space of weak isomorphism classes of compact networks is complete, i.e. any Cauchy sequence of compact networks converges to a compact network. The space $\mathcal{N}$ admits rich classes of precompact families, which gives us an idea about its topological complexity. Moreover, this space is geodesic, and geodesics between two points need not be of a unique form. The motif sets that we established as invariants return to feature in a very interesting reconstruction theorem: two compact networks (satisfying some additional assumptions) are weakly isomorphic if and only if they have the same n-motif sets, for all $n \in \mathbb{N}$. Finally, we close the arc that began with strong isomorphism in the following theorem: two compact networks (satisfying some additional assumptions) are weakly isomorphic if and only if they contain some essential substructures that are strongly isomorphic.

Acknowledgments. This work is partially supported by the National Science Foundation under grants CCF-1526513, IIS-1422400, and DMS-1547357. Facundo Mémoli acknowledges support from the Mathematical Biosciences Institute at The Ohio State University.

APPENDIX A. PROOFS

Proof of Proposition 10. The case for Type I weak isomorphism is similar to that of Type II, so we omit it. For Type II weak isomorphism, the reflexive and symmetric properties are easy to see, so we only provide details for verifying transitivity. Let $A, B, C \in \mathcal{N}$ be such that $A \cong^w_{\text{II}} B$ and $B \cong^w_{\text{II}} C$. Let $\varepsilon > 0$, and let P, S be sets with surjective maps $\varphi_A : P \to A$, $\varphi_B : P \to B$, $\psi_B : S \to B$, $\psi_C : S \to C$ such that:

$$|\omega_A(\varphi_A(p), \varphi_A(p')) - \omega_B(\varphi_B(p), \varphi_B(p'))| < \varepsilon/2 \quad \text{for each } p, p' \in P, \text{ and}$$
$$|\omega_B(\psi_B(s), \psi_B(s')) - \omega_C(\psi_C(s), \psi_C(s'))| < \varepsilon/2 \quad \text{for each } s, s' \in S.$$

Next define $T := \{(p, s) \in P \times S : \varphi_B(p) = \psi_B(s)\}$.

Claim 5. The projection maps $\pi_P : T \to P$ and $\pi_S : T \to S$ are surjective.

Proof. Let $p \in P$. Then $\varphi_B(p) \in B$, and since $\psi_B : S \to B$ is surjective, there exists $s \in S$ such that $\psi_B(s) = \varphi_B(p)$. Thus $(p, s) \in T$, and $\pi_P(p, s) = p$. This suffices to show that $\pi_P : T \to P$ is a surjection. The case for $\pi_S : T \to S$ is similar. ■

It follows from the preceding claim that $\varphi_A \circ \pi_P : T \to A$ and $\psi_C \circ \pi_S : T \to C$ are surjective. Next let $(p, s), (p', s') \in T$. Then,

$$\begin{aligned}
&|\omega_A(\varphi_A(\pi_P(p, s)), \varphi_A(\pi_P(p', s'))) - \omega_C(\psi_C(\pi_S(p, s)), \psi_C(\pi_S(p', s')))| \\
&= |\omega_A(\varphi_A(p), \varphi_A(p')) - \omega_C(\psi_C(s), \psi_C(s'))| \\
&= |\omega_A(\varphi_A(p), \varphi_A(p')) - \omega_B(\varphi_B(p), \varphi_B(p')) + \omega_B(\varphi_B(p), \varphi_B(p')) - \omega_C(\psi_C(s), \psi_C(s'))| \\
&= |\omega_A(\varphi_A(p), \varphi_A(p')) - \omega_B(\varphi_B(p), \varphi_B(p')) + \omega_B(\psi_B(s), \psi_B(s')) - \omega_C(\psi_C(s), \psi_C(s'))| \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, it follows that $A \cong^w_{\text{II}} C$. □

Proof of Theorem 12. It is clear that $d_{\mathcal{N}}(X, Y) \geqslant 0$. To show $d_{\mathcal{N}}(X, X) = 0$, consider the correspondence $R = \{(x, x) : x \in X\}$. Then for any $(x, x), (x', x') \in R$, we have $|\omega_X(x, x') - \omega_X(x, x')| = 0$. Thus $\operatorname{dis}(R) = 0$ and $d_{\mathcal{N}}(X, X) = 0$.

Next we show symmetry, i.e. $d_{\mathcal{N}}(X, Y) \leqslant d_{\mathcal{N}}(Y, X)$ and $d_{\mathcal{N}}(Y, X) \leqslant d_{\mathcal{N}}(X, Y)$. The two cases are similar, so we just show the second inequality. Let $\eta > d_{\mathcal{N}}(X, Y)$. Let $R \in \mathscr{R}(X, Y)$ be such that $\operatorname{dis}(R) < 2\eta$. Then define $\tilde{R} = \{(y, x) : (x, y) \in R\}$. Note that $\tilde{R} \in \mathscr{R}(Y, X)$. We have:

$$\begin{aligned}
\operatorname{dis}(\tilde{R}) &= \sup_{(y,x),(y',x') \in \tilde{R}} |\omega_Y(y, y') - \omega_X(x, x')| \\
&= \sup_{(x,y),(x',y') \in R} |\omega_Y(y, y') - \omega_X(x, x')| \\
&= \sup_{(x,y),(x',y') \in R} |\omega_X(x, x') - \omega_Y(y, y')| = \operatorname{dis}(R).
\end{aligned}$$

So $\operatorname{dis}(R) = \operatorname{dis}(\tilde{R})$. Then $d_{\mathcal{N}}(Y, X) = \frac{1}{2} \inf_{S \in \mathscr{R}(Y,X)} \operatorname{dis}(S) \leqslant \frac{1}{2} \operatorname{dis}(\tilde{R}) < \eta$. This shows $d_{\mathcal{N}}(Y, X) \leqslant d_{\mathcal{N}}(X, Y)$. The reverse inequality follows by a similar argument.

Next we prove the triangle inequality. Let $R \in \mathscr{R}(X, Y)$, $S \in \mathscr{R}(Y, Z)$, and let

$$R \circ S = \{(x, z) \in X \times Z \mid \exists y, (x, y) \in R, (y, z) \in S\}$$

First we claim that $R \circ S \in \mathscr{R}(X, Z)$. This is equivalent to checking that for each $x \in X$, there exists z such that $(x, z) \in R \circ S$, and for each $z \in Z$, there exists x such that $(x, z) \in R \circ S$. The proofs of these two conditions are similar, so we just prove the former. Let $x \in X$. Let $y \in Y$ be such that $(x, y) \in R$. Then there exists $z \in Z$ such that $(y, z) \in S$. Then $(x, z) \in R \circ S$.

Next we claim that $\operatorname{dis}(R \circ S) \leqslant \operatorname{dis}(R) + \operatorname{dis}(S)$. Let $(x, z), (x', z') \in R \circ S$. Let $y \in Y$ be such that $(x, y) \in R$ and $(y, z) \in S$. Let $y' \in Y$ be such that $(x', y') \in R$, $(y', z') \in S$. Then we have:

$$\begin{aligned} |\omega_X(x, x') - \omega_Z(z, z')| &= |\omega_X(x, x') - \omega_Y(y, y') + \omega_Y(y, y') - \omega_Z(z, z')| \\ &\leqslant |\omega_X(x, x') - \omega_Y(y, y')| + |\omega_Y(y, y') - \omega_Z(z, z')| \\ &\leqslant \operatorname{dis}(R) + \operatorname{dis}(S). \end{aligned}$$

This holds for any $(x, z), (x', z') \in R \circ S$, and proves the claim.

Now let $\eta_1 > d_{\mathcal{N}}(X, Y)$, let $\eta_2 > d_{\mathcal{N}}(Y, Z$, and let $R \in \mathscr{R}(X, Y)$, $S \in \mathscr{R}(Y, Z)$ be such that $\operatorname{dis}(R) < 2\eta_1$ and $\operatorname{dis}(S) < 2\eta_2$. Then we have:

$$d_{\mathcal{N}}(X, Z) \leqslant \tfrac{1}{2} \operatorname{dis}(R \circ S) \leqslant \tfrac{1}{2} \operatorname{dis}(R) + \tfrac{1}{2} \operatorname{dis}(S) < 2\eta_1 + 2\eta_2.$$

This shows that $d_{\mathcal{N}}(X, Z) \leqslant d_{\mathcal{N}}(X, Y) + d_{\mathcal{N}}(Y, Z)$, and proves the triangle inequality.

Finally, we claim that $X \cong^w_{\mathrm{II}} Y$ if and only if $d_{\mathcal{N}}(X, Y) = 0$. Suppose $d_{\mathcal{N}}(X, Y) = 0$. Let $\varepsilon > 0$, and let $R(\varepsilon) \in \mathscr{R}(X, Y)$ be such that $\operatorname{dis}(R(\varepsilon)) < \varepsilon$. Then for any $z = (x, y), z' = (x', y') \in R(\varepsilon)$, we have $|\omega_X(x, x') - \omega_Y(y, y')| < \varepsilon$. But this is equivalent to writing $|\omega_X(\pi_X(z), \pi_X(z')) - \omega_Y(\pi_Y(z), \pi_Y(z'))| < \varepsilon$, where $\pi_X : R(\varepsilon) \to X$ and $\pi_Y : R(\varepsilon) \to Y$ are the canonical projection maps. This holds for each $\varepsilon > 0$. Thus $X \cong^w_{\mathrm{II}} Y$.

Conversely, suppose $X \cong^w_{\mathrm{II}} Y$, and for each $\varepsilon > 0$ let $Z(\varepsilon)$ be a set with surjective maps $\phi^\varepsilon_X : Z(\varepsilon) \to X$, $\phi^\varepsilon_Y : Z \to Y$ such that $|\omega_X(\phi_X(z), \phi_X(z')) - \omega_Y(\phi_Y(z), \phi_Y(z'))| < \varepsilon$ for all $z, z' \in Z(\varepsilon)$. For each $\varepsilon > 0$, let $R(\varepsilon) = \{(\phi^\varepsilon_X(z), \phi^\varepsilon_Y(z)) : z \in Z(\varepsilon)\}$. Then $R(\varepsilon) \in \mathscr{R}(X, Y)$ for each $\varepsilon > 0$, and $\operatorname{dis}(R(\varepsilon)) = \sup_{z,z' \in Z} |\omega_X(\phi_X(z), \phi_X(z')) - \omega_Y(\phi_Y(z), \phi_Y(z'))| < \varepsilon$.

We conclude that $d_{\mathcal{N}}(X, Y) = 0$. Thus $d_{\mathcal{N}}$ is a metric modulo Type II weak isomorphism. □

Proof of Example 14. We start with some notation: for $x, x' \in X$, $y, y' \in Y$, let

$$\Gamma(x, x', y, y') = |\omega_X(x, x') - \omega_Y(y, y')|.$$

Let $\varphi : X \to Y$ be a bijection. Note that $R_\varphi := \{(x, \varphi(x)) : x \in X\}$ is a correspondence, and this holds for any bijection (actually any surjection) φ. Since we minimize over all correspondences for $d_{\mathcal{N}}$, we conclude $d_{\mathcal{N}}(X, Y) \leqslant \widehat{d}_{\mathcal{N}}(X, Y)$.

For the reverse inequality, we represent all the elements of $\mathscr{R}(X, Y)$ as 2-by-2 binary matrices R, where a 1 in position ij means $(x_i, y_j) \in R$. Denote the matrix representation of each $R \in \mathscr{R}(X, Y)$ by mat(R), and the collection of such matrices as mat$(\mathscr{R})$. Then we have:

$$\operatorname{mat}(\mathscr{R}) = \{\left(\begin{smallmatrix} 1 & a \\ b & 1 \end{smallmatrix}\right) : a, b \in \{0, 1\}\} \cup \{\left(\begin{smallmatrix} a & 1 \\ 1 & b \end{smallmatrix}\right) : a, b \in \{0, 1\}\}$$

Let $A = \{(x_1, y_1), (x_2, y_2)\}$ (in matrix notation, this is $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$) and let $B = \{(x_1, y_2), (x_2, y_1)\}$ (in matrix notation, this is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$). Let $R \in \mathscr{R}(X, Y)$. Note that either $A \subseteq R$ or $B \subseteq R$. Suppose that $A \subseteq R$. Then we have:

$$\max_{(x,y),(x',y') \in A} \Gamma(x, x', y, y') \leqslant \max_{(x,y),(x',y') \in R} \Gamma(x, x', y, y')$$

Let $\Omega(A)$ denote the quantity on the left hand side. A similar result holds in the case $B \subseteq R$:

$$\max_{(x,y),(x',y') \in B} \Gamma(x, x', y, y') \leqslant \max_{(x,y),(x',y') \in R} \Gamma(x, x', y, y')$$

Let $\Omega(B)$ denote the quantity on the left hand side. Since either $A \subseteq R$ or $B \subseteq R$, we have

$$\min \{\Omega(A), \Omega(B)\} \leqslant \min_{R \in \mathscr{R}} \max_{(x,y),(x',y') \in R} \Gamma(x, x', y, y')$$

We may identify A with the bijection given by $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$. Similarly we may identify B with the bijection sending $x_1 \mapsto y_2$, $x_2 \mapsto y_1$. Thus we have

$$\min_{\varphi} \max_{x,x' \in X} \Gamma(x,x',\varphi(x),\varphi(x')) \leqslant \min_{R \in \mathscr{R}} \max_{(x,y),(x',y') \in R} \Gamma(x,x',y,y').$$

So we have $\widehat{d}_{\mathcal{N}}(X,Y) \leqslant d_{\mathcal{N}}(X,Y)$. Thus $\widehat{d}_{\mathcal{N}} = d_{\mathcal{N}}$.

Next, let $\{p,q\}$ and $\{p',q'\}$ denote the vertex sets of X and Y. Consider the bijection φ given by $p \mapsto p'$, $q \mapsto q'$ and the bijection ψ given by $p \mapsto q'$, $q \mapsto p'$. Note that the weight matrix is determined by setting $\omega_X(p,p) = \alpha$, $\omega_X(p,q) = \delta$, $\omega_X(q,p) = \beta$, and $\omega_X(q,q) = \gamma$, and similarly for Y. Then we get $\operatorname{dis}(\varphi) = \max(|\alpha-\alpha'|, |\beta-\beta'|, |\gamma-\gamma'|, |\delta-\delta'|)$ and $\operatorname{dis}(\psi) = \max((|\alpha-\gamma'|, |\gamma-\alpha'|, |\delta-\beta'|, |\beta-\delta'|)$. The formula follows immediately. $\square$

Proof of Proposition 17. We begin with an observation. Given $X, Y \in \mathcal{FN}$, let $X', Y' \in \mathcal{FN}$ be such that $X \cong^w_{\text{II}} X'$, $Y \cong^w_{\text{II}} Y'$, and $\operatorname{card}(X') = \operatorname{card}(Y')$. Then we have:

$$d_{\mathcal{N}}(X,Y) \leqslant d_{\mathcal{N}}(X,X') + d_{\mathcal{N}}(X',Y') + d_{\mathcal{N}}(Y',Y) = d_{\mathcal{N}}(X',Y') \leqslant \widehat{d}_{\mathcal{N}}(X',Y'),$$

where the last inequality follows from Remark 15.

Next let $\eta > d_{\mathcal{N}}(X,Y)$, and let $R \in \mathscr{R}(X,Y)$ be such that $\operatorname{dis}(R) < 2\eta$. We wish to find networks X' and Y' such that $\widehat{d}_{\mathcal{N}}(X',Y') < \eta$. Write $Z = X \times Y$, and write $f : Z \to X$ and $g : Z \to Y$ to denote the (surjective) projection maps $(x,y) \mapsto x$ and $(x,y) \mapsto y$. Notice that we may write $R = \{(f(z), g(z)) : z \in R \subseteq Z\}$. In particular, by the definition of a correspondence, the restrictions of f, g to R are still surjective.

Define two weight functions $f^*\omega, g^*\omega : R \times R \to \mathbb{R}$ by $f^*\omega(z,z') = \omega_X(f(z), f(z'))$ and $g^*\omega(z,z') = \omega_Y(g(z), g(z'))$. Let $(U, \omega_U) = (R, f^*\omega)$ and let $(V, \omega_V) = (R, g^*\omega)$. Note that $d_{\mathcal{N}}(X,U) = 0$ by Remark 8, because $\operatorname{card}(U) \geqslant \operatorname{card}(X)$ and for all $z, z' \in U$, we have $\omega_U(z,z') = f^*\omega(z,z') = \omega_X(f(z), f(z'))$ for the surjective map f. Similarly $d_{\mathcal{N}}(Y,V) = 0$.

Next let $\varphi : U \to V$ be the bijection $z \mapsto z$. Then we have:

$$\begin{aligned}
\sup_{z,z' \in U} |\omega_U(z,z') - \omega_V(\varphi(z),\varphi(z'))| &= \sup_{z,z' \in U} |\omega_U(z,z') - \omega_V(z,z')| \\
&= \sup_{z,z' \in R} |\omega_X(f(z),f(z')) - \omega_Y(g(z),g(z'))| \\
&= \sup_{(x,y),(x',y') \in R} |\omega_X(x,x') - \omega_Y(y,y')| \\
&= \operatorname{dis}(R). \text{ In particular,} \\
\inf_{\varphi : U \to V \text{ bijection}} \operatorname{dis}(\varphi) &\leqslant \operatorname{dis}(R).
\end{aligned}$$

So there exist networks U, V with the same node set (and thus the same cardinality) such that $\widehat{d}_{\mathcal{N}}(U,V) \leqslant \frac{1}{2}\operatorname{dis}(R) < \eta$. We have already shown that $d_{\mathcal{N}}(X,Y) \leqslant \widehat{d}_{\mathcal{N}}(U,V)$. Since $\eta > d_{\mathcal{N}}(X,Y)$ was arbitrary, it follows that we have:

$$d_{\mathcal{N}}(X,Y) = \inf\left\{\widehat{d}_{\mathcal{N}}(X',Y') : X' \cong^w_{\text{II}} X, Y' \cong^w_{\text{II}} Y, \text{ and } \operatorname{card}(X') = \operatorname{card}(Y')\right\}. \qquad \square$$

Proof of Theorem 23. By Theorem 13, we know that X and Y are Type I weakly isomorphic. So there exists a set V with surjections $\varphi_X : V \to X$, $\varphi_Y : V \to Y$ such that $A_X(\varphi_X(v), \varphi_X(v')) = A_Y(\varphi_Y(v), \varphi_Y(v'))$ for all $v, v' \in V$. Thus we obtain (not necessarily unique) maps $f : X \to Y$ and $g : Y \to X$ that are weight-preserving. Hence the composition $g \circ f : X \to X$ is a weight-preserving map. Without loss of generality, assume that X is Ψ-controlled. Recall that this means that there exists a continuous function $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ such that $\Psi(0,0) = 0$ and $A_X(x,x') \leqslant \Psi(A_X(x,x''), A_X(x',x''))$ for all $x, x', x'' \in X$.

It is known that an isometric embedding from a compact metric space into itself must be bijective [7, Theorem 1.6.14]. We now prove a similar result using the assumptions of our theorem. Let $h : X \to X$ be a weight-preserving map. By the assumption of a dissimilarity network, we know that f, g and h are injective.

We check that h is continuous, using the assumptions about the topology on X. Let $V \subseteq h(X)$ be open. Define $U := h^{-1}[V]$. We claim that U is open. Let $x \in U$, and consider $h(x) \in V$. Since V is open and the forward balls form a base for the topology, we pick $\varepsilon > 0$ such that $B^+(h(x), \varepsilon) \subseteq V$. Now let $x' \in B^+(x, \varepsilon)$. Then $A_X(h(x), h(x')) = A_X(x, x') < \varepsilon$, so $h(x') \in B^+(h(x), \varepsilon) \subseteq V$. Hence $x' \in U$. It follows that $B^+(x, \varepsilon) \subseteq U$. Hence U is open, and h is continuous.

Next we check that X is Hausdorff, using the Ψ_X-controllability assumption. Let $x, x' \in X$, where $x \neq x'$. Using continuity of Ψ_X, let $\varepsilon > 0$ be such that $\Psi([0,\varepsilon), [0,\varepsilon)) \subseteq [0, A_X(x,x'))$. We wish to show that $B^+(x,\varepsilon) \cap B^+(x',\varepsilon) = \varnothing$. Towards a contradiction, suppose this is not the case and let $z \in B^+(x,\varepsilon) \cap B^+(x',\varepsilon)$. But then $A_X(x,x') \leqslant \Psi\big(A_X(x,z), A_X(x',z)\big) < A_X(x,x')$, a contradiction. It follows that X is Hausdorff.

Now $h(X)$ is compact, being the continuous image of a compact space, and it is closed in X because it is a compact subset of a Hausdorff space.

Finally we show that h is surjective. Towards a contradiction, suppose that the open set $X \backslash h(X)$ is nonempty, and let $x \in X \backslash h(X)$. Using the topology assumption on X, pick $\varepsilon > 0$ such that $B^+(x,\varepsilon) \subseteq X \backslash h(X)$. Define $x_0 := x$, and $x_n := h(x_{n-1})$ for each $n \in \mathbb{N}$. Then for each $n \in \mathbb{N}$, we have $A_X(x_0, x_n) \geqslant \varepsilon$. Since h is weight-preserving, we also have $A_X(x_k, x_{k+n}) \geqslant \varepsilon$ for all $k, n \in \mathbb{N}$. Since X is sequentially compact, the sequence $(x_k)_{k \geqslant 0}$ has a convergent subsequence $(x_{k_j})_{j \in \mathbb{N}}$ that limits to some $z \in X$. Thus $B^+(z, r)$ contains all but finitely many terms of this sequence, for any $r > 0$. Now for any $m, n \in \mathbb{N}$ we observe:

$$\begin{aligned} A_X(x_m, x_n) \leqslant \Psi\big(A_X(x_m, z), A_X(x_n, z)\big), \text{ where } A_X(x_m, z) &\leqslant \Psi\big(A_X(x_m, x_m), A_X(z, x_m)\big) \\ &= \Psi\big(0, A_X(z, x_m)\big), \\ \text{and similarly } A_X(x_n, z) &\leqslant \Psi\big(0, A_X(z, x_n)\big). \end{aligned}$$

Since Ψ is continuous and vanishes at $(0,0)$, we choose $\delta > 0$ such that $\Psi([0,\delta), [0,\delta)) \subseteq [0,\varepsilon)$. We also choose $\eta > 0$ such that $\Psi(0, [0,\eta)) \subseteq [0,\delta)$. Since $B^+(z,\eta)$ contains all but finitely many terms of the sequence $(x_{k_j})_{j \geqslant \mathbb{N}}$, we pick $N \in \mathbb{N}$ so that $x_{k_m} \in B^+(z,\eta)$, for all $m \geqslant N$. Let $m, n \geqslant N$. Then $A_X(z, x_{k_n}) < \eta$ and $A_X(z, x_{k_m}) < \eta$. Thus $A_X(x_{k_n}, z) \leqslant \Psi(0, A_X(z, x_{k_n})) < \delta$ and $A_X(x_{k_m}, z) \leqslant \Psi(0, A_X(z, x_{k_m})) < \delta$. It follows that $A_X(x_{k_m}, x_{k_n}) < \varepsilon$.

But this is a contradiction to what we have shown before. Thus h is surjective, hence bijective. Since h was an arbitrary weight-preserving map from X into itself, the same result holds for $g \circ f : X \to X$. This shows that g is surjective. It follows that $X \cong^s Y$. $\square$

Proof of Proposition 38. All of these cases are easy to check, so we will just record the proof for spec. Suppose (X, ω_X) and (Y, ω_Y) are strongly isomorphic via φ. Let $x \in X$. Let $\alpha \in \mathrm{spec}_X(x)$. Then there exists $x' \in X$ such that $\alpha = \omega_X(x, x')$. But since $\omega_X(x,x') = \omega_Y(\varphi(x), \varphi(x'))$, we also have $\alpha \in \mathrm{spec}_Y(\varphi(x))$. Thus $\mathrm{spec}_X(x) \subseteq \mathrm{spec}_X(\varphi(x))$. The reverse containment is similar. Thus for any $x \in X$, $\mathrm{spec}_X(x) = \mathrm{spec}_Y(\varphi(x))$. Since $\mathrm{spec}(X) = \bigcup_{x \in X} \mathrm{spec}_X(x)$, it follows that $\mathrm{spec}(X) = \mathrm{spec}(Y)$. $\square$

Lemma 53. *Let $(X, \omega_X), (Y, \omega_Y) \in \mathcal{CN}$. Let f represent any of the maps* tr, out, *and* in, *and let f_X (resp. f_Y) represent the corresponding map* $\mathrm{tr}_X, \mathrm{out}_X, \mathrm{in}_X$ *(resp.* $\mathrm{tr}_Y, \mathrm{out}_Y, \mathrm{in}_Y$*). Then we* *obtain:*

$$d_{\mathrm{H}}^{\mathbb{R}}(f(X), f(Y)) = \inf_{R \in \mathscr{R}(X,Y)} \sup_{(x,y)\in R} \big|f_X(x) - f_Y(y)\big|.$$

Proof of Lemma 53. Observe that $f(X) = \{f_X(x) : x \in X\} = f_X(X)$, so we need to show

$$d_{\mathrm{H}}^{\mathbb{R}}(f_X(X), f_Y(Y)) = \inf_{R \in \mathscr{R}(X,Y)} \sup_{(x,y)\in R} \big|f_X(x) - f_Y(y)\big|.$$

Recall that by the definition of Hausdorff distance on $\mathbb{R}$, we have

$$d_{\mathrm{H}}^{\mathbb{R}}(f_X(X), f_Y(Y)) = \max\Big\{ \sup_{x\in X} \inf_{y \in Y} |f_X(x) - f_Y(y)|, \sup_{y\in Y} \inf_{x\in X} |f_X(x) - f_Y(y)|\Big\}.$$

Let $a \in X$ and let $R \in \mathscr{R}(X,Y)$. Then there exists $b \in Y$ such that $(a,b) \in R$. Then we have:

$$\begin{aligned} |f_X(a) - f_Y(b)| &\leqslant \sup_{(x,y)\in R} |f_X(x) - f_Y(y)|, \text{ and so} \\ \inf_{b\in Y} |f_X(a) - f_Y(b)| &\leqslant \sup_{(x,y)\in R} |f_X(x) - f_Y(y)|. \end{aligned}$$

This holds for all $a \in X$. Then,

$$\sup_{a\in X} \inf_{b\in Y} |f_X(a) - f_Y(b)| \leqslant \sup_{(x,y)\in R} |f_X(x) - f_Y(y)|.$$

This holds for all $R \in \mathscr{R}(X,Y)$. So we have

$$\sup_{a\in X} \inf_{b\in Y} |f_X(a) - f_Y(b)| \leqslant \inf_{R\in\mathscr{R}} \sup_{(x,y)\in R} |f_X(x) - f_Y(y)|.$$

By a similar argument, we also have

$$\begin{aligned} \sup_{b\in Y} \inf_{a\in X} |f_X(a) - f_Y(b)| &\leqslant \inf_{R\in\mathscr{R}} \sup_{(x,y)\in R} |f_X(x) - f_Y(y)|. \\ \text{Thus } d_{\mathrm{H}}^{\mathbb{R}}(f_X(X), f_Y(Y)) &\leqslant \inf_{R\in\mathscr{R}} \sup_{(x,y)\in R} |f_X(x) - f_Y(y)|. \end{aligned}$$

Now we show the reverse inequality. Let $x \in X$, and let $\eta > d_{\mathrm{H}}^{\mathbb{R}}(f_X(X), f_Y(Y))$. Then there exists $y \in Y$ such that $|f_X(x) - f_Y(y)| < \eta$. Define $\varphi(x) = y$, and extend φ to all of X in this way. Let $y \in Y$. Then there exists $x \in X$ such that $|f_X(x) - f_Y(y)| < \eta$. Define $\psi(y) = x$, and extend ψ to all of Y in this way. Let $R = \{(x, \varphi(x)) : x \in X\} \cup \{(\psi(y), y) : y \in Y\}$. Then for each $(a,b) \in R$, we have $|f_X(a) - f_Y(b)| < \eta$. Thus we have $\inf_{R\in\mathscr{R}} \sup_{(x,y)\in R} |f_X(x) - f_Y(y)| < \eta$. Since $\eta > d_{\mathrm{H}}^{\mathbb{R}}(f_X(X), f_Y(Y))$ was arbitrary, it follows that

$$\inf_{R\in\mathscr{R}(X,Y)} \sup_{(x,y)\in R} |f_X(x) - f_Y(y)| \leqslant d_{\mathrm{H}}^{\mathbb{R}}(f_X(X), f_Y(Y)). \qquad \square$$

Proof of Proposition 41. Let $\eta > d_{\mathcal{N}}(X,Y)$. We break this proof into three parts.

The diam **case.** Recall that diam is an $\mathbb{R}$-valued invariant, so we wish to show $|\mathrm{diam}(X) - \mathrm{diam}(Y)| \leqslant 2d_{\mathcal{N}}(X,Y)$. Let $R \in \mathscr{R}(X,Y)$ be such that for any $(a,b),(a',b') \in R$, we have $|\omega_X(a,a') - \omega_Y(b,b')| < 2\eta$.

Let $x, x' \in X$ such that $|\omega_X(x,x')| = \operatorname{diam}(X)$, and let y, y' be such that $(x,y),(x',y') \in R$. Then we have:

$$\begin{aligned} |\omega_X(x,x') - \omega_Y(y,y')| &< 2\eta \\ |\omega_X(x,x') - \omega_Y(y,y')| + |\omega_Y(y,y')| &< 2\eta + |\omega_Y(y,y')| \\ |\omega_X(x,x')| &< \operatorname{diam}(Y) + 2\eta. \\ \text{Thus } \operatorname{diam}(X) &< \operatorname{diam}(Y) + 2\eta \end{aligned}$$

Similarly, we get $\operatorname{diam}(Y) < \operatorname{diam}(X) + 2\eta$. It follows that $|\operatorname{diam}(X) - \operatorname{diam}(Y)| < 2\eta$. Since $\eta > d_{\mathcal{N}}(X,Y)$ was arbitrary, it follows that:

$$|\operatorname{diam}(X) - \operatorname{diam}(Y)| \leqslant 2d_{\mathcal{N}}(X,Y).$$

For tightness, consider the networks $X = N_1(1)$ and $Y = N_1(2)$. By Example 7, we have that $d_{\mathcal{N}}(X,Y) = \frac{1}{2}$. On the other hand, $\operatorname{diam}(X) = 1$ and $\operatorname{diam}(Y) = 2$ so that $|\operatorname{diam}(X) - \operatorname{diam}(Y)| = 1 = 2d_{\mathcal{N}}(X,Y)$.

The cases tr**,** out**, and** in**.** First we show $\mathbf{L}(\operatorname{tr}) = 2$. By Lemma 53, it suffices to show

$$\inf_{R\in\mathscr{R}(X,Y)} \sup_{(x,y)\in R} |\operatorname{tr}_X(x) - \operatorname{tr}_Y(y)| < 2\eta.$$

Let $R \in \mathscr{R}(X,Y)$ be such that for any $(a,b),(a',b') \in R$, we have $|\omega_X(a,a') - \omega_Y(b,b')| < 2\eta$. Then we obtain $|\omega_X(a,a) - \omega_Y(b,b)| < 2\eta$. Thus $|\operatorname{tr}_X(a) - \operatorname{tr}_Y(b)| < 2\eta$. Since $(a,b) \in R$ was arbitrary, it follows that $\sup_{(a,b)\in R} |\operatorname{tr}_X(a) - \operatorname{tr}_X(b)| < 2\eta$. It follows that $\inf_{R\in\mathscr{R}} \sup_{(a,b)\in R} |\operatorname{tr}_X(a) - \operatorname{tr}_X(b)| < 2\eta$. The result now follows because $\eta > d_{\mathcal{N}}(X,Y)$ was arbitrary. The proofs for out and in are similar, so we just show the former. By Lemma 53, it suffices to show

$$\inf_{R\in\mathscr{R}(X,Y)} \sup_{(x,y)\in R} |\operatorname{out}_X(x) - \operatorname{out}_Y(y)| < 2\eta.$$

Recall that $\operatorname{out}_X(x) = \max_{x'\in X} |\omega_X(x,x')|$. Let $R \in \mathscr{R}(X,Y)$ be such that $|\omega_X(x,x') - \omega_Y(y,y')| < 2\eta$ for any $(x,y),(x',y') \in R$. By triangle inequality, it follows that $|\omega_X(x,x')| < |\omega_Y(y,y')| + 2\eta$. In particular, for $(x',y') \in R$ such that $|\omega_X(x,x')| = \operatorname{out}_X(x)$, we have $\operatorname{out}_X(x) < |\omega_Y(y,y')| + 2\eta$. Hence $\operatorname{out}_X(x) < \operatorname{out}_Y(y) + 2\eta$. Similarly, $\operatorname{out}_Y(y) < \operatorname{out}_X(x) + 2\eta$. Thus we have $|\operatorname{out}_X(x) - \operatorname{out}_Y(y)| < 2\eta$. This holds for all $(x,y) \in R$, so we have:

$$\sup_{(x,y)\in R} |\operatorname{out}_X(x) - \operatorname{out}_Y(y)| < 2\eta.$$

Minimizing over all correspondences, we get:

$$\inf_{R\in\mathscr{R}} \sup_{(a,b)\in R} |\operatorname{out}_X(a) - \operatorname{out}_Y(b)| < 2\eta.$$

The result follows because $\eta > d_{\mathcal{N}}(X,Y)$ was arbitrary.

Finally, we need to show that our bounds for the Lipschitz constant are tight. Let $X = N_1(1)$ and let $Y = N_1(2)$. Then $d_{\mathcal{N}}(X,Y) = \frac{1}{2}$. We also have $d_{\mathrm{H}}^{\mathbb{R}}(\operatorname{tr}(X), \operatorname{tr}(Y)) = |1-2| = 1$, and similarly $d_{\mathrm{H}}^{\mathbb{R}}(\operatorname{out}(X), \operatorname{out}(Y)) = d_{\mathrm{H}}^{\mathbb{R}}(\operatorname{in}(X), \operatorname{in}(Y)) = 1$.

The cases $\mathrm{m}^{\mathrm{out}}$ **and** m^{in}**.** The two cases are similar, so let's just prove $\mathbf{L}(\mathrm{m}^{\mathrm{out}}) = 2$. Since $\mathrm{m}^{\mathrm{out}}$ is an $\mathbb{R}$-invariant, we wish to show $|\mathrm{m}^{\mathrm{out}}(X) - \mathrm{m}^{\mathrm{out}}(Y)| < 2\eta$. It suffices to show:

$$|\mathrm{m}^{\mathrm{out}}(X) - \mathrm{m}^{\mathrm{out}}(Y)| \leqslant d_{\mathrm{H}}^{\mathbb{R}}(\operatorname{out}(X), \operatorname{out}(Y)),$$

because we have already shown

$$d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{out}(X),\mathrm{out}(Y)) = \inf_{R\in\mathscr{R}(X,Y)} \sup_{(x,y)\in R} |\mathrm{out}_X(x) - \mathrm{out}_Y(y)| < 2\eta.$$

Here we have used Lemma 53 for the first equality above.

Let $\varepsilon > d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{out}(X),\mathrm{out}(Y))$. Then for any $x \in X$, there exists $y \in Y$ such that:

$$|\mathrm{out}_X(x) - \mathrm{out}_Y(y)| < \varepsilon.$$

Let $a \in X$ be such that $\mathrm{m}^{\mathrm{out}}(X) = \mathrm{out}_X(a)$. Then we have:

$$|\mathrm{out}_X(a) - \mathrm{out}_Y(y)| < \varepsilon,$$

for some $y \in Y$. In particular, we have:

$$\mathrm{m}^{\mathrm{out}}(Y) \leqslant \mathrm{out}_Y(y) < \varepsilon + \mathrm{out}_X(a) = \varepsilon + \mathrm{m}^{\mathrm{out}}(X).$$

Similarly, we obtain:

$$\mathrm{m}^{\mathrm{out}}(X) < \varepsilon + \mathrm{m}^{\mathrm{out}}(Y).$$

Thus we have $|\mathrm{m}^{\mathrm{out}}(X) - \mathrm{m}^{\mathrm{out}}(Y)| < \varepsilon$. Since $\varepsilon > d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{out}(X),\mathrm{out}(Y))$ was arbitrary, we have:

$$|\mathrm{m}^{\mathrm{out}}(X) - \mathrm{m}^{\mathrm{out}}(Y)| \leqslant d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{out}(X),\mathrm{out}(Y)).$$

The inequality now follows by Lemma 53 and our proof in the case of the out map.

For tightness, note that $|\mathrm{m}^{\mathrm{out}}(N_1(1)) - \mathrm{m}^{\mathrm{out}}(N_1(2))| = |1-2| = 1 = 2\cdot\frac{1}{2} = 2d_{\mathcal{N}}(N_1(1), N_1(2))$. The same example works for the m^{in} case. ☐

Proof of Proposition 43. (First inequality.) Let $X, Y \in \mathcal{CN}$ and let $\eta > d_{\mathcal{N}}(X,Y)$. Let $R \in \mathscr{R}(X,Y)$ be such that $\sup_{(x,y),(x',y')\in R} |\omega_X(x,x') - \omega_Y(y,y')| < 2\eta$. Let $(x,y) \in R$, and let $\alpha \in \mathrm{spec}_X(x)$. Then there exists $x' \in X$ such that $\omega_X(x,x') = \alpha$. Let $y' \in Y$ be such that $(x',y') \in R$. Let $\beta = \omega_Y(y,y')$. Note $\beta \in \mathrm{spec}_Y(y)$. Also note that $|\alpha - \beta| < 2\eta$. By a symmetric argument, for each $\beta \in \mathrm{spec}_Y(y)$, there exists $\alpha \in \mathrm{spec}_X(x)$ such that $|\alpha - \beta| < 2\eta$. So $d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X(x),\mathrm{spec}_Y(y)) < 2\eta$. This is true for any $(x,y) \in R$, and so we have $\sup_{(x,y)\in R} d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X(x),\mathrm{spec}_Y(y)) \leqslant 2\eta$. Then we have:

$$\inf_{R\in\mathscr{R}} \sup_{(x,y)\in R} d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X(x),\mathrm{spec}_Y(y)) \leqslant 2\eta.$$

Since $\eta > d_{\mathcal{N}}(X,Y)$ was arbitrary, the first inequality follows.

(Second inequality.) Let $R \in \mathscr{R}(X,Y)$. Let $\eta(R) = \sup_{(x,y)\in R} d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X(x),\mathrm{spec}_Y(y))$. Let $\alpha \in \mathrm{spec}(X)$. Then $\alpha \in \mathrm{spec}_X(x)$ for some $x \in X$. Let $y \in Y$ such that $(x,y) \in R$. Then there exists $\beta \in \mathrm{spec}_Y(y)$ such that $|\alpha-\beta| \leqslant d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X(x),\mathrm{spec}_Y(y))$, and in particular, $|\alpha-\beta| \leqslant \eta(R)$. In other words, for each $\alpha \in \mathrm{spec}(X)$, there exists $\beta \in \mathrm{spec}(Y)$ such that $|\alpha - \beta| \leqslant \eta(R)$. By a symmetric argument, for each $\beta \in \mathrm{spec}(Y)$, there exists $\alpha \in \mathrm{spec}(X)$ such that $|\alpha - \beta| \leqslant \eta(R)$. Thus $d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}(X),\mathrm{spec}(Y)) \leqslant \eta(R)$. This holds for any $R \in \mathscr{R}$. Thus we have

$$d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}(X),\mathrm{spec}(Y)) \leqslant \inf_{R\in\mathscr{R}} \sup_{(x,y)\in R} d_{\mathrm{H}}^{\mathbb{R}}(\mathrm{spec}_X(x),\mathrm{spec}_Y(y)).$$

This proves the second inequality. ☐

Proof of Theorem 45. Let $n \in \mathbb{N}$. We wish to show $d_n(\mathrm{M}_n(X),\mathrm{M}_n(Y)) \leqslant 2d_{\mathcal{N}}(X,Y)$. Let $R \in \mathscr{R}(X,Y)$. Let $(x_i) \in X^n$, and let $(y_i) \in Y^n$ be such that for each i, we have $(x_i,y_i) \in R$. Then for all $j,k \in \{1,\ldots,n\}$, $|\omega_X(x_i,x_j) - \omega_Y(y_i,y_j)| \leqslant \mathrm{dis}(R)$.

Thus $\inf_{(y_i)\in Y^n} |\omega_X(x_i,x_j) - \omega_Y(y_i,y_j)| \leqslant \operatorname{dis}(R)$. This is true for any $(x_i) \in X^n$. Thus we get:

$$\sup_{(x_i)\in X^n} \inf_{(y_i)\in Y^n} |\omega_X(x_i,x_j) - \omega_Y(y_i,y_j)| \leqslant \operatorname{dis}(R).$$

By a symmetric argument, we get $\sup_{(y_i)\in Y^n} \inf_{(x_i)\in X^n} |\omega_X(x_i,x_j) - \omega_Y(y_i,y_j)| \leqslant \operatorname{dis}(R)$. Thus $d_n(\mathrm{M}_n(X), \mathrm{M}_n(Y)) \leqslant \operatorname{dis}(R)$. This holds for any $R \in \mathscr{R}(X,Y)$. Thus $d_n(\mathrm{M}_n(X), \mathrm{M}_n(Y)) \leqslant \inf_{R\in\mathscr{R}(X,Y)} \operatorname{dis}(R) = 2d_{\mathcal{N}}(X,Y)$.

For tightness, let $X = N_1(1)$ and let $Y = N_1(2)$. Then $d_{\mathcal{N}}(X,Y) = \frac{1}{2}$, so we wish to show $d_n(\mathrm{M}_n(X), \mathrm{M}_n(Y)) = 1$ for each $n \in \mathbb{N}$. Let $n \in \mathbb{N}$. Let $\mathbb{1}_{n\times n}$ denote the $n \times n$ matrix with 1 in each entry. Then $\mathrm{M}_n(X) = \{\mathbb{1}_{n\times n}\}$ and $\mathrm{M}_n(Y) = \{2 \cdot \mathbb{1}_{n\times n}\}$. Thus $d_n(\mathrm{M}_n(X), \mathrm{M}_n(Y)) = 1$. Since n was arbitrary, we conclude that equality holds for each $n \in \mathbb{N}$. $\square$

Appendix B. Additional results on simulated hippocampal networks

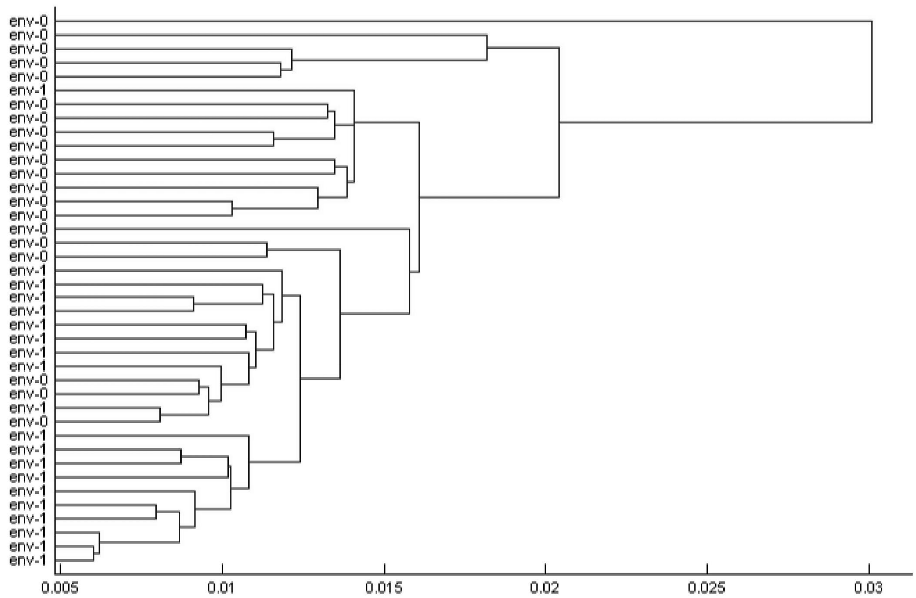

FIGURE 14. Single linkage dendrogram based on local spectrum lower bound of Proposition 49 corresponding to hippocampal networks with place field radius $0.2L$.

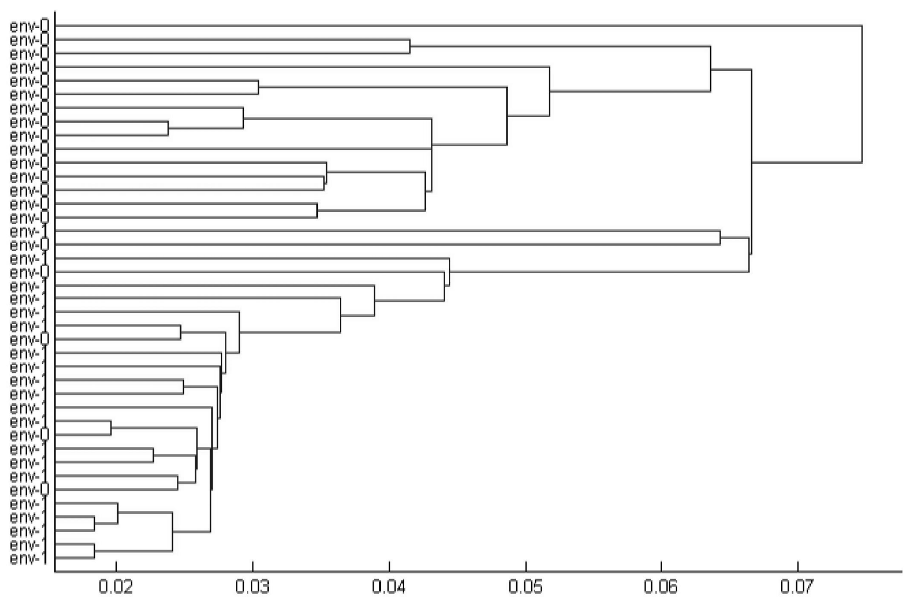

FIGURE 15. Single linkage dendrogram based on local spectrum lower bound of Proposition 49 corresponding to hippocampal networks with place field radius $0.1L$.

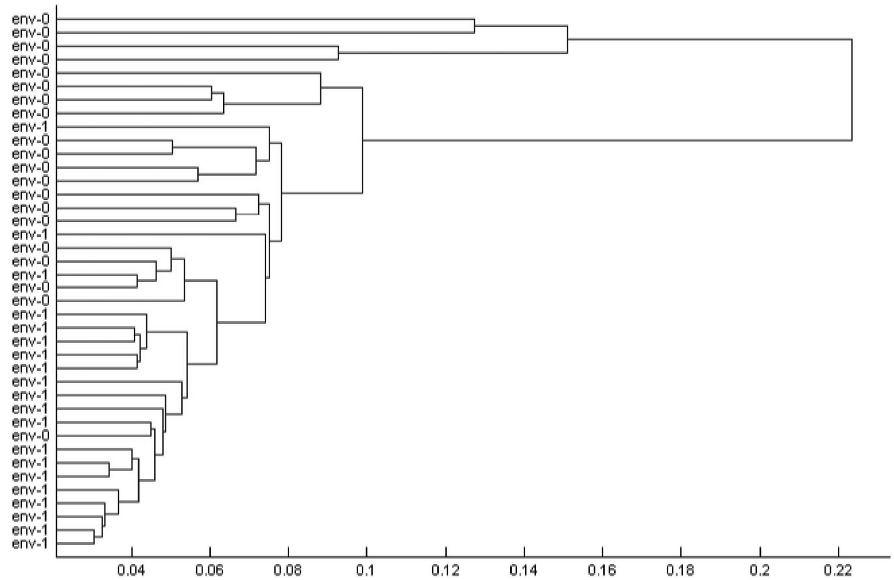

FIGURE 16. Single linkage dendrogram based on local spectrum lower bound of Proposition 49 corresponding to hippocampal networks with place field radius $0.05L$.

REFERENCES

[1] Dena Asta and Cosma Rohilla Shalizi. Geometric network comparison. *arXiv preprint arXiv:1411.1350*, 2014.

[2] David Bao, S-S Chern, and Zhongmin Shen. *An introduction to Riemann-Finsler geometry*, volume 200. Springer Science & Business Media, 2012.

[3] Phillip J Best, Aaron M White, and Ali Minai. Spatial processing in the brain: the activity of hippocampal place cells. *Annual review of neuroscience*, 24(1):459–486, 2001.

[4] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.

[5] Mireille Boutin and Gregor Kemper. Lossless representation of graphs using distributions. *arXiv preprint arXiv:0710.1870*, 2007.

[6] Emery N Brown, Loren M Frank, Dengda Tang, Michael C Quirk, and Matthew A Wilson. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *The Journal of Neuroscience*, 18(18):7411–7425, 1998.

[7] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*, volume 33 of *AMS Graduate Studies in Math.* American Mathematical Society, 2001.

[8] Rainer E Burkard, Mauro Dell'Amico, and Silvano Martello. *Assignment Problems*. SIAM, 2009.

[9] Bin Cao, Ying Li, and Jianwei Yin. Measuring similarity between graphs based on the Levenshtein distance. *Appl. Math*, 7(1L):169–175, 2013.

[10] Gunnar Carlsson and Facundo Mémoli. Characterization, stability and convergence of hierarchical clustering methods. *Journal of Machine Learning Research*, 11:1425–1470, 2010.

[11] Gunnar Carlsson, Facundo Mémoli, Alejandro Ribeiro, and Santiago Segarra. Axiomatic construction of hierarchical clustering in asymmetric networks. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 5219–5223. IEEE, 2013.

[12] Gunnar Carlsson, Facundo Mémoli, Alejandro Ribeiro, and Santiago Segarra. Hierarchical quasi-clustering methods for asymmetric networks. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014*, 2014.

[13] CJ Carstens and KJ Horadam. Persistent homology of collaboration networks. *Mathematical Problems in Engineering*, 2013, 2013.

[14] Gary Chartrand, Farrokh Saba, and Hung Bin Zou. Edge rotations and distance between graphs. *Časopis pro pěstování matematiky*, 110(1):87–91, 1985.

[15] Frédéric Chazal, David Cohen-Steiner, Leonidas J Guibas, Facundo Mémoli, and Steve Y Oudot. Gromov-Hausdorff stable signatures for shapes using persistence. In *Computer Graphics Forum*, volume 28, pages 1393–1403. Wiley Online Library, 2009.

[16] Samir Chowdhury and Facundo Mémoli. Persistent homology of directed networks. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 77–81. IEEE, 2016.

[17] Samir Chowdhury and Facundo Mémoli. A functorial Dowker theorem and persistent homology of asymmetric networks. *arXiv preprint arXiv:1608.05432*, 2018.

[18] Samir Chowdhury and Facundo Mémoli. The metric space of networks. *arXiv preprint arXiv:1804.02820*, 2018.

[19] Samir Chowdhury and Facundo Mémoli. Persistent path homology of directed networks. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1152–1169. SIAM, 2018.

[20] Fan RK Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

[21] Diane J Cook and Lawrence B Holder. Substructure discovery using minimum description length and background knowledge. *Journal of Artificial Intelligence Research*, pages 231–255, 1994.

[22] Carina Curto and Vladimir Itskov. Cell groups reveal structure of stimulus space. *PLoS Computational Biology*, 4(10), 2008.

[23] Yuri Dabaghian, Facundo Mémoli, Loren Frank, and Gunnar Carlsson. A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Comput Biol*, 8(8), 2012.

[24] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. Technical report, Department of Mathematics, Uppsala University, 2007.

[25] Pawel Dlotko, Kathryn Hess, Levi Ran, Henry Markram, Eilif Muller, Max Nolte, Michael Reimann, and Martina Scolamiero. Topological analysis of the connectome of digital reconstructions of neural microcircuits. 2016.

[26] Gaëlle Doucet, Mikaël Naveau, Laurent Petit, Nicolas Delcroix, Laure Zago, Fabrice Crivello, Gael Jobard, Nathalie Tzourio-Mazoyer, Bernard Mazoyer, Emmanuel Mellet, et al. Brain activity at rest: a multiscale hierarchical functional organization. *Journal of neurophysiology*, 105(6):2753–2763, 2011.

[27] David Easley and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.

[28] Gerald A Edgar. Classics on fractals. 1993.

[29] Justin Eldridge, Mikhail Belkin, and Yusu Wang. Graphons, mergeons, and so on! In *Advances in Neural Information Processing Systems*, pages 2307–2315, 2016.

[30] Mirtha-Lina Fernández and Gabriel Valiente. A graph distance metric combining maximum common subgraph and minimum common supergraph. *Pattern Recognition Letters*, 22(6):753–758, 2001.

[31] Dennis Fetterly, Mark Manasse, Marc Najork, and Janet Wiener. A large-scale study of the evolution of web pages. In *Proceedings of the 12th international conference on World Wide Web*, pages 669–678. ACM, 2003.

[32] Santo Fortunato. Benchmark graphs to test community detection algorithms. https://sites.google.com/site/santofortunato/inthepress2.

[33] Maurice Fréchet. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 22(1):1–72, 1906.

[34] Fred Galvin and Samuel Shore. Completeness in semimetric spaces. *Pacific Journal of Mathematics*, 113(1):67–75, 1984.

[35] Fred Galvin and Samuel Shore. Distance functions and topologies. *The American Mathematical Monthly*, 98(7):620–623, 1991.

[36] Xinbo Gao, Bing Xiao, Dacheng Tao, and Xuelong Li. A survey of graph edit distance. *Pattern Analysis and applications*, 13(1):113–129, 2010.

[37] Chad Giusti, Robert Ghrist, and Danielle S Bassett. Two's company, three (or more) is a simplex: Algebraic-topological tools for understanding higher-order structure in neural data. *Journal of Computational Neuroscience*, 41(1):doi–10, 2016.

[38] Chad Giusti, Eva Pastalkova, Carina Curto, and Vladimir Itskov. Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112(44):13455–13460, 2015.

[39] Wayne Goddard and Henda C Swart. Distances between graphs under edge operations. *Discrete Mathematics*, 161(1):121–132, 1996.

[40] Ronald L Graham. A similarity measure for graphs: reflections on a theme of Ulam. *From Cardinals to Chaos. Reflections on the Life and Legacy of Stanislaw Ulam Cambridge University Press, Cambridge*, pages 114–121, 1989.

[41] Mikhail Gromov. Structures métriques pour les variétés Riemanniennes. *Textes Mathématiques [Mathematical Texts]*, 1, 1981.

[42] Mikhail Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999.

[43] Gary Gruenhage. Generalized metric spaces. *Handbook of set-theoretic topology*, pages 423–501, 1984.

[44] Nan Hu and Leonidas Guibas. Spectral descriptors for graph matching. *arXiv preprint arXiv:1304.1572*, 2013.

[45] Nan Hu, Raif Rustamov, and Leonidas Guibas. Stable and informative spectral signatures for graph matching. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 2313–2320. IEEE, 2014.

[46] Akihiro Inokuchi, Takashi Washio, and Hiroshi Motoda. Complete mining of frequent patterns from graphs: Mining graph data. *Machine Learning*, 50(3):321–354, 2003.

[47] Mark Johnson. An ordering of some metrics defined on the space of graphs. *Czechoslovak Mathematical Journal*, 37(1):75–85, 1987.

[48] Nigel J Kalton and Mikhail I Ostrovskii. Distances between banach spaces. In *Forum Mathematicum*, volume 11, pages 17–48. Walter de Gruyter, 1999.

[49] Arshi Khalid, Byung Sun Kim, Moo K Chung, Jong Chul Ye, and Daejong Jeon. Tracing the evolution of multi-scale functional networks in a mouse model of depression using persistent brain network homology. *NeuroImage*, 101:351–363, 2014.

[50] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.

[51] Jon M Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S Tomkins. The web as a graph: measurements, models, and methods. In *Computing and combinatorics*, pages 1–17. Springer, 1999.

[52] Risi Kondor and John Lafferty. Diffusion kernels on graphs and other discrete input spaces. In *ICML*, volume 2, pages 315–322, 2002.

[53] Danai Koutra, Joshua T Vogelstein, and Christos Faloutsos. Deltacon: A principled massive-graph similarity function. In *Proceedings of the 2013 SIAM International Conference on Data Mining*, pages 162–170. SIAM, 2013.

[54] Ravi Kumar, Jasmine Novak, and Andrew Tomkins. Structure and evolution of online social networks. In *Link mining: models, algorithms, and applications*, pages 337–357. Springer, 2010.

[55] Hyekyoung Lee, Moo K Chung, Hyejin Kang, Boong-Nyun Kim, and Dong Soo Lee. Computing the shape of brain networks using graph filtration and Gromov-Hausdorff metric. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 302–309. Springer, 2011.

[56] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187. ACM, 2005.

[57] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[58] Owen Macindoe and Whitman Richards. Graph comparison using fine structure analysis. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 193–200. IEEE, 2010.

[59] Paolo Masulli and Alessandro EP Villa. The topology of the directed clique complex as a network invariant. *SpringerPlus*, 5(1):1–12, 2016.

[60] Facundo Mémoli. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011.

[61] Facundo Mémoli. Some properties of Gromov–Hausdorff distances. *Discrete & Computational Geometry*, pages 1–25, 2012. 10.1007/s00454-012-9406-8.

[62] Facundo Mémoli and Guillermo Sapiro. Comparing point clouds. In *SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 32–40, New York, NY, USA, 2004. ACM.

[63] Facundo Mémoli and Guillermo Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5(3):313–347, 2005.

[64] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[65] James R Munkres. *Topology*. Prentice Hall, 2000.

[66] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
[67] VW Niemytzki. On the "third axiom of metric space". *Transactions of the American Mathematical Society*, 29(3):507–513, 1927.
[68] Alexandros Ntoulas, Junghoo Cho, and Christopher Olston. What's new on the web?: the evolution of the web from a search engine perspective. In *Proceedings of the 13th international conference on World Wide Web*, pages 1–12. ACM, 2004.
[69] Hiroyuki Ogata, Wataru Fujibuchi, Susumu Goto, and Minoru Kanehisa. A heuristic graph comparison algorithm and its application to detect functionally related enzyme clusters. *Nucleic acids research*, 28(20):4021–4028, 2000.
[70] John O'Keefe and Jonathan Dostrovsky. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 34(1):171–175, 1971.
[71] John O'Keefe and Lynn Nadel. *The hippocampus as a cognitive map*, volume 3. Clarendon Press Oxford, 1978.
[72] Panos M Pardalos and Henry Wolkowicz, editors. *Quadratic assignment and related problems*. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 16. American Mathematical Society, Providence, RI, 1994.
[73] Luiz Pessoa. Understanding brain networks and brain organization. *Physics of life reviews*, 11(3):400–435, 2014.
[74] Peter Petersen. *Riemannian geometry*, volume 171. Springer Science & Business Media, 2006.
[75] Giovanni Petri, Paul Expert, Federico Turkheimer, Robin Carhart-Harris, David Nutt, Peter Hellyer, and Francesco Vaccarino. Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11(101):20140873, 2014.
[76] Giovanni Petri, Martina Scolamiero, Irene Donato, and Francesco Vaccarino. Topological strata of weighted complex networks. *PloS one*, 8(6):e66506, 2013.
[77] Ron Y Pinter, Oleg Rokhlenko, Esti Yeger-Lotem, and Michal Ziv-Ukelson. Alignment of metabolic pathways. *Bioinformatics*, 21(16):3401–3408, 2005.
[78] Arthur Dunn Pitcher and Edward Wilson Chittenden. On the foundations of the calcul fonctionnel of Fréchet. *Transactions of the American Mathematical Society*, 19(1):66–78, 1918.
[79] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.
[80] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069, 2010.
[81] Sorin V Sabau, Kazuhiro Shibuya, and Hideo Shimada. Metric structures associated to Finsler metrics. *arXiv preprint arXiv:1305.5880*, 2013.
[82] Felix Schmiedl. *Shape matching and mesh segmentation: mathematical analysis, algorithms and an application in automated manufacturing*. PhD thesis, München, Technische Universität München, Diss., 2015, 2015.
[83] Roded Sharan and Trey Ideker. Modeling cellular machinery through biological network comparison. *Nature biotechnology*, 24(4):427–433, 2006.
[84] Yi-Bing Shen and Wei Zhao. Gromov pre-compactness theorems for nonreversible Finsler manifolds. *Differential Geometry and its Applications*, 28(5):565–581, 2010.
[85] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the transcriptional regulation network of Escherichia coli. *Nature genetics*, 31(1):64–68, 2002.
[86] Ann Sizemore, Chad Giusti, and Danielle S Bassett. Classification of weighted networks through mesoscale homological features. *Journal of Complex Networks*, 5(2):245–273, 2016.
[87] Ann E Sizemore, Chad Giusti, Ari Kahn, Jean M Vettel, Richard F Betzel, and Danielle S Bassett. Cliques and cavities in the human connectome. *Journal of computational neuroscience*, 44(1):115–145, 2018.
[88] Zane Smith, Samir Chowdhury, and Facundo Mémoli. Hierarchical representations of network data with optimal distortion bounds. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 1834–1838. IEEE, 2016.
[89] Alexander J Smola and Risi Kondor. Kernels and regularization on graphs. In *Learning theory and kernel machines*, pages 144–158. Springer, 2003.
[90] Olaf Sporns. *Networks of the Brain*. MIT press, 2011.
[91] Olaf Sporns. *Discovering the human connectome*. MIT press, 2012.
[92] Olaf Sporns and Rolf Kötter. Motifs in brain networks. *PLoS biology*, 2(11):e369, 2004.
[93] Lynn Arthur Steen and J Arthur Seebach. *Counterexamples in topology*, volume 18. Springer, 1978.

[94] Aleksandar Stojmirović and Yi-Kuo Yu. Geometric aspects of biological sequence comparison. *Journal of Computational Biology*, 16(4):579–610, 2009.

[95] S V N Vishwanathan, Nicol N Schraudolph, Risi Kondor, and Karsten M Borgwardt. Graph kernels. *The Journal of Machine Learning Research*, 11:1201–1242, 2010.

[96] Joshua T Vogelstein and Carey E Priebe. Shuffled graph classification: Theory and connectome applications. *arXiv preprint arXiv:1112.5506*, 2011.

[97] Joshua T Vogelstein, William Gray Roncal, R Jacob Vogelstein, and Carey E Priebe. Graph classification using signal-subgraphs: Applications in statistical connectomics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(7):1539–1551, 2013.

[98] Takashi Washio and Hiroshi Motoda. State of the art of graph-based data mining. *Acm Sigkdd Explorations Newsletter*, 5(1):59–68, 2003.

[99] Pawel Waszkiewicz. The local triangle axiom in topology and domain theory. *Applied General Topology*, 4(1):47–70, 2013.

[100] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[101] BT Thomas Yeo, Fenna M Krienen, Jorge Sepulcre, Mert R Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L Roffman, Jordan W Smoller, Lilla Zöllei, Jonathan R Polimeni, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 106(3):1125–1165, 2011.

[102] Kenichi Yoshida, Hiroshi Motoda, and Nitin Indurkhya. Graph-based induction as a unified learning framework. *Applied Intelligence*, 4(3):297–316, 1994.

[103] Laura A Zager and George C Verghese. Graph similarity scoring and matching. *Applied mathematics letters*, 21(1):86–94, 2008.