

Characterizing Distances of Networks on the Tensor Manifold

Bipul Islam¹, Ji Liu¹, Romeil Sandhu¹

Stony Brook University, Stony Brook NY 08544, USA,
romeil.sandhu@stonybrook.edu

Abstract. At the core of understanding dynamical systems is the ability to maintain and control the systems behavior that includes notions of robustness, heterogeneity, and/or regime-shift detection. Recently, to explore such functional properties, a convenient representation has been to model such dynamical systems as a weighted graph consisting of a finite, but very large number of interacting agents. This said, there exists very limited relevant statistical theory that is able cope with real-life data, i.e., how does perform analysis and/or statistics over a “family” of networks as opposed to a specific network or network-to-network variation. Here, we are interested in the analysis of network families whereby each network represents a “point” on an underlying statistical manifold. To do so, we explore the Riemannian structure of the tensor manifold developed by Pennec previously applied to Diffusion Tensor Imaging (DTI) towards the problem of network analysis. In particular, while this note focuses on Pennec definition of “geodesics” amongst a family of networks, we show how it lays the foundation for future work for developing measures of network robustness for regime-shift detection. We conclude with experiments highlighting the proposed distance on synthetic networks and an application towards biological (stem-cell) systems.

Keywords: computational geometry, graph theory, control

1 Introduction

Notions of robustness, heterogeneity, and phase changes are ubiquitous concepts employed in understanding complex dynamical systems with a variety of applications [1–4]. This is seen in Figure 1A. While recent advancements in network analysis has arisen through the usage of spectral techniques [5], expander graphs [6], percolation theory [7], information theoretic approaches [2, 3], scale-free networks [4], and a myriad of graph measures reliant on the underlying discrete graph space (e.g., degree distribution, shortest path, centrality [8, 9]), such measures rarely incorporate the underlying dynamics in its construction. To this end, we have previously developed fundamental relationships between network functionality [2, 10] and certain topological and geometric properties of the corresponding graph [11, 12]. In particular, recent work of ours has shown the geometric notion of curvature (a measure of “flatness”) is positively correlated with a system’s robustness or its ability to adapt to dynamic changes

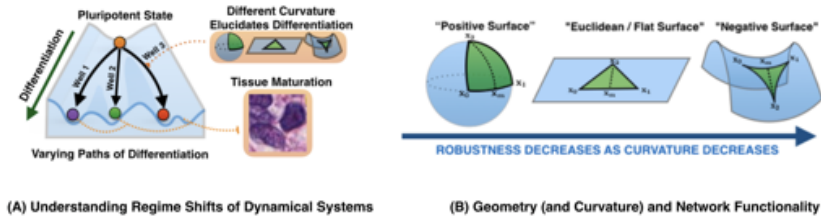


Fig. 1. This work focuses on the exploring a Riemannian framework to elucidate regime-shifts and bifurcation detection in dynamical systems which can include, but not limited to cellular differentiation and signaling promiscuity in biological systems.

[13, 14]. This can be seen in Figure 1B. Here, we are interested in expounding upon previous work [13] in order to exploit distances (and eventually statistics, robustness) over a family of networks. In particular, as vector-valued networks are becoming increasingly important, we believe such a framework may be of interest to broader complex network community. To do so, we are interested in studying geometry of networks as a dynamical system that evolves over time in which each given network or observation may represent a “point” (encoded by *some* matrix-based positive definite model) on an underlying Riemannian manifold. From this, we are interested in applying the framework of Penec [15, 16] to define geodesics such that distances between “points” (networks) in a given family can be measured as seen in Figure 2. This said, we note significant related work on employing statistical analysis for Riemannian manifolds has focused on directional statistics, statistics on shape spaces, as well as specific manifolds - we refer the reader [17–19] and references therein. While these methods are mostly extrinsic whereby one embeds the manifold in an ambient Euclidean space, we are interested in studying *intrinsic* Riemannian metrics that can be naturally applied to the network setting without restrictions on the network topology. As such, the remainder of the present paper is outlined as follows: In the next section, we provide preliminaries regarding the framework by Penec [15, 16] primarily used for DTI imaging. From this, we then show how this framework is a natural towards networks with a particular remarks on how we can begin on constructing robustness over network families. Section 4 presents preliminary results on synthetic and biological data and Section 5 concludes with future work.

2 Preliminaries

This section focuses on developing a measures of distance amongst networks in the context of developing measures of network robustness (future work).

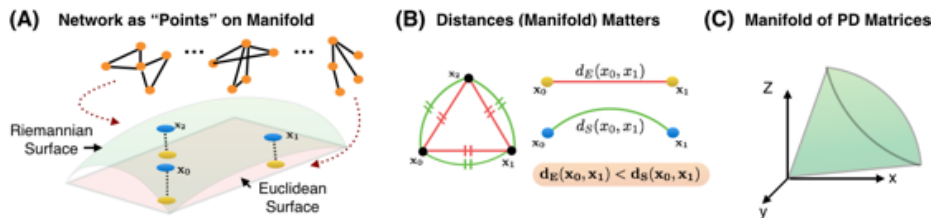


Fig. 2. (A) To understand global functionality, we treat networks as “points” on a manifold. (B) We require a framework for not only geodesic distance, but generalize distributions for network functionality [13]. (C) Manifold of Positive Definite Matrices

2.1 Geodesics on the Network Manifold

Let us consider a dynamical system that evolves over time and whose information is *encapsulated by a positive definite matrix (tensor)*. In this regard, it is well-known that the space of tensors is not a vector space, but instead forms a Riemannian manifold M [20]. In particular, we will leverage the theory of symmetric spaces which has been extensively studied since the seminal work of Nomizu [22]; a comprehensive work on tensor manifolds can be found in [21, 15]. Riemannian manifolds are endowed with a metric that smoothly assign to each point $\rho \in M$ an inner product on \mathcal{T}_ρ , the tangent space to M at ρ . More formally, let us denote \mathcal{Y} and \mathcal{Y}^+ and as the set of all Hermitian matrices and the cone of positive-definite matrices, respectively. We consider the following:

$$\Lambda_+ := \{\rho \in \mathcal{Y}^+ \mid \text{tr}(\rho) = 1\} \quad \mathcal{T}_\rho := \{\chi \in \mathcal{Y} \mid \text{tr}(\chi) = 0\} \quad (1)$$

where Λ_+ represents our space of networks and \mathcal{T}_ρ is the corresponding tangent space. From this, we let the observations of our dynamic system be recorded as $\{\rho_0, \rho_1, \rho_2, \dots, \rho_t\}$ and where each “point” (network) $\rho_i \in \mathcal{Y}^+$ is given by a tensor. As seen in Figure 2, we need a sensible notion of distance that is dependent on the underlying geometry. To do so, one defines the *exponential map* as the function that maps to each vector $\overrightarrow{\rho_0\rho_1} \in \mathcal{T}_{\rho_0}$, the point $\rho_1 \in \Lambda_+$ on the manifold M that is reached after unit time by the geodesic starting at ρ_0 with this tangent vector. The exponential map, herein denoted as $\exp_\rho : \mathcal{T}_\rho \mapsto M$ at point ρ is defined on the whole tangent space and for the particular space of tensors, this map is also one-to-one. Moreover, one can also define a unique inverse map denoted as the *logarithm map* $\text{Log}_{\rho_0} : M \mapsto \mathcal{T}_{\rho_0}$ that maps a point $\rho_1 \in M$ to the unique tangent vector $\chi \in \mathcal{T}_{\rho_0}$ at ρ_0 whose initial velocity is that of the unique geodesic γ with $\gamma(0) = \rho_0$ and $\gamma(1) = \rho_1$. Thus, the problem of computing distances amongst networks (encapsulated by a positive defense matrix) amounts to being able to properly computing geodesics on this space and for which, we must now consider a family of curves $\gamma(t)$ on the manifold and its speed vector $\dot{\gamma}(t)$. Then,

we seek the minimum distance (length) for any two points, e.g., ρ_0 and ρ_1 as

$$\mathcal{L}_{\rho_0}^{\rho_1} = \min_{\gamma} \int_{\rho_0}^{\rho_1} \|\dot{\gamma}(t)\|_{\gamma(t)} dt = \min_{\gamma} \int_{\rho_0}^{\rho_1} \left(\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle_{\gamma(t)} \right) \quad (2)$$

where $\gamma(0) = \rho_0$ and $\gamma(1) = \rho_1$. For this note, we relax the restriction upon trace one matrices for sake of clarity (without obfuscating the conceptual motivation).

2.2 Revisiting Riemannian Framework for Tensors

As highlighted, we seek to leverage Penneec [15] framework towards the network setting. While this framework has popularized most recently in the field of DTI [16, 23], the mechanics can be used to analyze network functionality with this note focusing on the computation of distances amongst such networks. In particular, the authors utilize a result from differential geometry regarding geodesics for invariant metrics on affine symmetric spaces [24] together with the Sylvester equation from control to propose an affine invariant metric whereby the unique geodesic curve from point ρ (and at the origin) on our tensor (network) manifold with a tangent vector χ can be shown as

$$\gamma(t)_{(\rho, \chi)} = \rho^{\frac{1}{2}} \exp(t\rho^{-\frac{1}{2}} \chi \rho^{-\frac{1}{2}}) \rho^{\frac{1}{2}} \quad \text{and} \quad \gamma(t)_{(I, \chi)} = \exp(t\chi) \quad (3)$$

where $\exp(\chi) = \sum_{k=0}^{+\infty} \frac{\chi^k}{k!}$ is the usual matrix exponential. From this, we have:

$$\begin{aligned} \frac{d\gamma(t)}{dt} &= \frac{d}{dt} \exp(t\chi) \\ &= \frac{d}{dt} \left(Q \begin{bmatrix} e^{t\chi_{1,1}} & 0 & \dots & 0 \\ 0 & e^{t\chi_{2,2}} & \dots & 0 \\ 0 & 0 & \dots & e^{t\chi_{n,n}} \end{bmatrix} Q^T \right) \\ &= Q \text{DIAG}(\chi_i \exp(t\chi_i)) Q^T \\ &= \exp(t\chi)^{\frac{1}{2}} \chi \exp(t\chi)^{\frac{1}{2}} \\ &= \gamma(t)_{(I, \chi)}^{\frac{1}{2}} \star \chi \end{aligned} \quad (4)$$

where the $A \star B = ABA^T$. Noting the geodesic $\gamma(t)_{(\rho, \chi)}$ between any two points $\rho_0, \rho_1 \in M$ where $\rho_1 = \exp_{\rho_0}(\chi)$ at $t = 1$, the *logarithm map* can be seen as

$$\begin{aligned} \gamma(1)_{(\rho_0, \chi)} &= \rho_1 = \rho_0^{\frac{1}{2}} \exp(\rho_0^{-\frac{1}{2}} \chi \rho_0^{-\frac{1}{2}}) \rho_0^{\frac{1}{2}} \\ \log(\rho_0^{-\frac{1}{2}} \rho_1 \rho_0^{-\frac{1}{2}}) &= \rho_0^{-\frac{1}{2}} \chi \rho_0^{-\frac{1}{2}} \\ \rho_0^{\frac{1}{2}} \log(\rho_0^{-\frac{1}{2}} \rho_1 \rho_0^{-\frac{1}{2}}) \rho_0^{\frac{1}{2}} &= \chi = \overrightarrow{\rho_0 \rho_1} = \text{Log}_{\rho_0}(\rho_1) \end{aligned} \quad (5)$$

From this, the distance between tensors can finally be computed as

$$\text{dist}^2(\rho_0, \rho_1) = \|\text{Log}_{\rho_0}(\rho_1)\|_{\rho_0}^2 = \|\log(\rho_0^{-\frac{1}{2}} \rho_1 \rho_0^{-\frac{1}{2}})\|_2^2. \quad (6)$$

We now have a natural distance that measures the length of the geodesic curve that connects any two ‘‘points’’ on our tensor (network) manifold. This is particularly important as it allows for characterization of functionality over a family of networks as we will highlight in the next sections.

2.3 Mean, Variance, and Gaussian Distributions

Given the above geodesic distance, one can formulate necessary statistics and most importantly form distributions. Specifically, with the above affine invariant metric, the unique mean $\bar{\rho}$ of N “points” $\{\rho_0, \rho_1, \rho_2, \dots, \rho_N\}$ on the tensor manifold can be computed via gradient descent:

$$\bar{\rho}_{t+1} = \bar{\rho}_t^{\frac{1}{2}} \exp \left(\frac{1}{N} \sum_{i=1}^N \log \left(\bar{\rho}_t^{-\frac{1}{2}} \rho_i \bar{\rho}_t^{-\frac{1}{2}} \right) \right) \bar{\rho}_t^{\frac{1}{2}} \quad (7)$$

It has been noted that although no closed form expression for the mean exists, the above converges towards a steady-state solution in a few iterations. As in the traditional Euclidean distance and similar to the Karcher or Frechet mean, the above mean is constructed on the premise that we seek a “point” that minimizes the sum of squared distance.e.e., $\rho = \sum_{i=1}^N \text{dist}^2(\rho, \rho_i)$. In a similar fashion, the variance can be computed as follows:

$$\Sigma = \frac{1}{N-1} \sum_{i=1}^N \text{Vec}_{\bar{\rho}}(\text{Log}_{\bar{\rho}}(\rho_i)) \text{Vec}_{\bar{\rho}}(\text{Log}_{\bar{\rho}}(\rho_i))^T \quad (8)$$

where $\text{Vec}_{\bar{\rho}}(\rho_i) = \text{Vec}_I(\log(\bar{\rho} \star \rho_i))$ and where the operation $\text{Vec}_I(A)$ is a vectorized projection of the independent coefficients of our tensor in the above formulation, i.e., $\text{Vec}_I(A) = (a_{1,1}, \sqrt{2}a_{1,2}, \dots, \sqrt{2}a_{1,n}, \dots, \sqrt{2}a_{n-1,n}, a_{n,n})$. From this, one obtains a generalized form of the Gaussian distribution on the tensor manifold:

$$N_{(\bar{\rho}, \Sigma)}(\rho_i) = k \exp \left(\frac{1}{2} \text{Vec}_{\bar{\rho}}(\text{Log}_{\bar{\rho}}(\rho_i))^T \Sigma^{-1} \text{Vec}_{\bar{\rho}}(\text{Log}_{\bar{\rho}}(\rho_i)) \right) \quad (9)$$

With the above, we can now begin to stylize how such a framework can be utilized towards characterizing network families.

3 Application To Network Characterization

This section presents how the above framework can be used to generalize notions of robustness to a family of networks. While the experiments focus on the fitness of “distance”, we provide further information to put this note in context.

3.1 Network Robustness From Scalar to Matrix Setting

As highlighted, our main underlying interest is to understand the functionality over network families. Recently, it has been shown that Ricci curvature (herein denoted “Ric”) from geometry is intrinsically connected to Boltzmann entropy H [25, 26] as well as functional robustness of networks R or the ability to maintain functionality in the presence of random fluctuations, i.e., $\Delta Ric \times \Delta R \geq 0$ [13, 14]. Given that geometry (curvature) of networks seemingly plays an integral role in understanding functionality, we require an appropriate discretization applicable

to a particular network. To this end, one can discretely defined Ricci Curvature (due to Ollivier [11]) between any two “points” x and y as:

$$\kappa(x, y) := 1 - \frac{W_1(\mu_x, \mu_y)}{d(x, y)} \quad (10)$$

This definition, motivated by coarse geometry, is applicable to the graph setting whereby the geodesic distance $d(x, y)$ is given by the hop metric and W_1 can be computed simply via linear programming. However, this measure is limited to local network-to-network variation as opposed to a family of networks. That is, the definition of equation (10) in the current setting has been used to analyze a given networks robustness through changes of curvature between any two end networks configurations. Therefore, our primary motivation is focused on developing a matrix-valued curvature measure such that we can analyze Ricci curvature over a family of networks in a more global setting. In particular, Ollivier’s definition is rooted in the concept of providing an “indicator” between negative, Euclidean “flat”, and positive curved spaces by measuring the ratio (distance) of the centers of a geodesic ball against that of the transport distance of ball’s distributions (via Wasserstein-1). Similarly, one can naturally expand this definition to the problem of a family of networks. To do so, we first need a geodesic distance $d_M(x, y)$ provided by equation (6) amongst networks along with the appropriate transport distance W_p . For the case of Euclidean geometry, one can define the transport distance between two Gaussians $\mu_1 = \mathcal{N}(m_1, C_1)$ and $\mu_2 = \mathcal{N}(m_2, C_2)$, where m_1 and $m_2 \in \mathbb{R}^n$ are the means and C_1 and $C_2 \in \mathbb{R}^{n \times n}$ are the covariances, as:

$$W_2(\mu_1, \mu_2)^2 = \|m_1 - m_2\|_2^2 + \text{trace}(C_1 + C_2 - 2(C_2^{1/2}C_1C_2^{1/2})^{1/2}). \quad (11)$$

Comparing this definition to the structure presented in Section 2, the difference lies with respect to the underlying geometry for which Pennek [15] provides the mathematical tools to generalize W_2 to our tensor manifold. While the generalization is on-going and future work, this puts experimental results of $d_M(x, y)$ in context; e.g., our framework interest is to naturally generalize to functionality.

3.2 Appropriate Matrix-Based Model: Graph Laplacian

Given our interest in network control, a natural network model for encapsulation is an approximated graph Laplacian whose structure provides a global representation of a given network with many useful properties that are intrinsically tied to system functionality. While a complete review is beyond the scope of this note (we refer the reader to [5]), a few areas that have garnered attention have included construction of expander graphs [6] to Cheegers inequality with increasing attention on connections between spectral graph theory and its respective geometry [27]. Mathematically, one can define the Laplacian operator for a specific network as $\Delta f(x) = f(x) - \sum_y f(y)\mu_x(y)$ with f being a real-valued function which coincides with the usual normalized graph Laplacian given by:

$$L = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \quad (12)$$

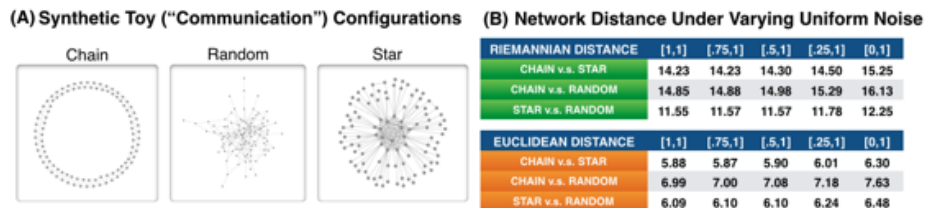


Fig. 3. (A) Graphical view of the three synthetic “toy” graphs that can be representative of varying types of communication networks, each with 200 nodes and 400 edges (i.e., only topology changes). (B) Distances of network-to-network under varying weights or “noise.” We see that the proposed framework is able to properly quantify difference as compared to Euclidean measure, i.e., the star-like network is “closest” to the random network as supported in [2, 13].

It is interesting to note in this connection that if $k \leq \kappa(x, y)$ is a lower bound for Ricci curvature defined in equation (10), then the eigenvalues of Laplacian is bounded $k \leq \lambda_2 \leq \dots \leq \lambda_N \leq 2 - k$ [27]. This relationship is important since $2 - \lambda_N$ measures the deviation of the graph from being bipartite, i.e., a graph whose vertices can be divided into two disjoint sets U and V such that every edge connects a vertex in U to one in V . As such ideas appear in resource allocation and control, the study of such structures in the time-varying setting motivate our focus on the Laplacian matrix. Lastly, while one can form a normalized Laplacian with unitary trace, we note that the Laplacian matrix by definition is a positive semi-definite matrix. Therefore, to enforce positive definiteness, we utilized an approximated Laplacian $\hat{L} := L + \epsilon_L I$ where $0 < \epsilon_L \ll 1$ such that $x^T \hat{L} x > 0$ for all nonzero $x \in \mathbb{R}^n$.

4 Results

We now present preliminary results to highlight potential for network analysis and set the foundation for future work in network robustness.

4.1 Toy Network Configurations

To provide initial intuition of the proposed framework, we constructed three toy “communication” based networks. As one can see in Figure 3, the networks are composed of 200 nodes and 400 edges with varying topology that represents a chain, a random, and a star-like network. Similar to our earlier work [13], we are interested in using the proposed framework to measure distances amongst these structures under varying “noise”. As highlighted by network curvature [13] and network entropy [2], one should expect from a communication perspective the star-like network is “closer” to a “random” network. That is, given that their functional properties such as robustness (a measure of “flatness”) are similar

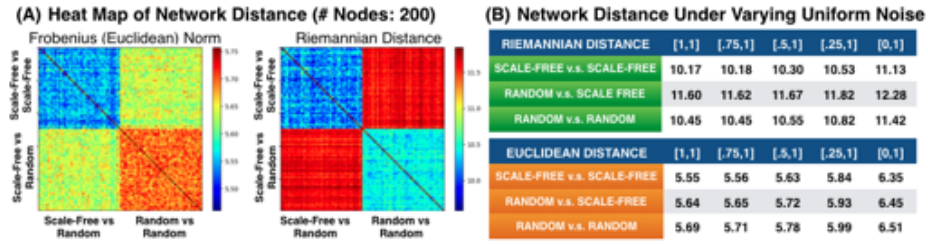


Fig. 4. Heat map of distances between scale-free and random networks measured by the proposed Riemannian metric to that of the Frobenius norm. (B) Average distance of network-to-network under varying weights or “noise.” We see that the proposed framework is able to properly quantify difference as compared to Euclidean measure. Note: Each network possess same number of nodes and edges, only topology changed.

compared towards the chain-like network, we should expect that such networks on the manifold lie closer to one another. This is exactly what is shown in Figure 3B. Moreover, as we move from an unweighted graph to a weighted graph in which weights are chosen randomly in a uniform manner, this relationship still holds. On the other hand, if we utilize the classical Euclidean distance (Frobenius norm) to measure network-to-network distance, we find that the chain-like network is in fact “closer” to the star-like network which is against intuition.

4.2 Scale-Free and Random Networks

This section presents clustering results and importance of accounting for the underlying geometry. In particular, we generate 50 scale-free networks [1] and 50 random (Erdos-Renyi) networks with each network composed 200 nodes and 1164 edges (differences are only in topology). We then compute pairwise distances amongst all potential network pairs based on the Euclidean distance, defined by the Frobenius norm, as well as the Riemannian distance, defined by (6). Figure 4A presents a “heat map” of the distances while Figure 4B presents average distance. As one can see, Euclidean distance fails to cluster networks as the distance between scale-free to random networks are “closer” than random to random networks. On the other hand, the framework proposed here is able to properly cluster the aforementioned networks which is an intuitive result given a basic understanding of geometry. However, this motivates this framework for further study on network robustness. We note if experiments are repeated and/or the number of nodes/edges are changed, the clustering result still holds.

4.3 Waddingtons Landscape: Cell Differentiation

In this section, we apply the proposed distance measure towards cellular differentiation as depicted in Figure 1. This epigenetic landscape provides a stylistic hierarchal view cellular differentiation, i.e., elevation associates to degrees of

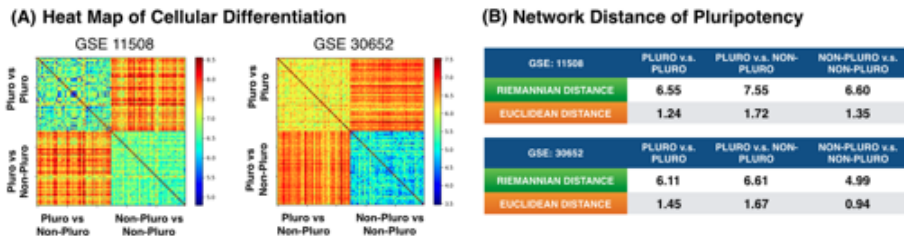


Fig. 5. Heat map of Riemannian distances between pluripotent and non-pluripotent networks under two public datasets. (B) Average numerical distance between the pluro and non-pluro networks. We see the proposed measure is able to classify such samples as well as the Euclidean measure. We note the topology here is fixed across each networks.

pluripotency or the ability for a cell to differentiate. At a high level, one can consider the intricate and complex processes of cellular differentiation as a network whereby such bifurcations relate to differentiated states [3]. To illustrate this, we accessed the Gene Expression Omnibus (GEO) of two public datasets, GSE11508 and GSE 30652, that possess 59 pluripotent / 160 non-pluripotent and 159 pluripotent / 32 non-pluripotent cell samples, respectively. Within each pluripotent samples, there contains tissue types that include human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs). This results in a potential mix of differentiation, i.e., iPSCs pluripotency may be much lower than hESCs yet are deemed pluripotent. Here, we are interested to examine how “far” apart differentiated samples are from their counterparts. As such, we construct a single pluripotent network similar to previous work [13] where the weights are the gene-to-gene correlation. This was done similarly for a single non-pluripotent network. Together, this was repeated to generate 50 pluripotent and non-pluripotent networks from each GSE11508 and GSE 30652. Figure 5 presents heat map results as well as the average network distance between such networks. As one can see, we see a marked increase in distance between pluro and non-pluro samples. We note while the Frobenius norm follows the same pattern, this is primarily due to the topology being fixed amongst samples as well as sparse interactions compared to work by Teschnedorff [3]. We provide this to highlight the tacit topological dependence on distance which requires further study that is on-going (e.g., accounting for signaling interactome hierarchy).

5 Conclusion

This note introduces a statistical geometric framework [15] previously applied to DTI Imaging [23] toward the problem of network analysis. To illustrate the importance of geometry in this setting, we presented preliminary results that show classical Euclidean distances are unable to properly differentiate between not only toy network configurations, but that of scale-free and random networks.

The results set the foundation for a more expansive study on global notions of robustness for which we will explore notions of network family curvatures.

References

1. A. Barabasi. "The Network Takeover," *Nature Physics*. 2012.
2. L. Demetrius and T. Manke. "Robustness and Network Evolution: Entropic principle." *Physica A: Statistical Mechanics and its Applications*. 2005.
3. J. West, G. Bianconi, S. Severini, and A. Teschendorff. "Differential Network Entropy Reveals Cancer System Hallmarks." *Nature (Scientific reports)* 2012.
4. A. Barabasi and R. Albert. "Emergence of Scaling Networks." *Science*. 1999.
5. R.K. Fan Chung. "Spectral Graph Theory." *American Mathematical Society*. 1997.
6. S. Hoory, N. Linial, and A. Wigderson. "Expander Graphs and Their Applications." *American Mathematical Society*. 2006.
7. D. Stauffer and A. Aharony. "Introduction to Percolation Theory." *Taylor and Francis*. 1994
8. S. Borgatti. "Centrality and Network Flow." *Social Networks*. (2005).
9. G. Ghoshal and A. Barabasi. "Ranking Stability and Super-Stable Nodes in Complex Networks." *Nature Communications*. 2011.
10. S. R. S. Varadhan. "Large Deviations and Applications." *SIAM*. 1984.
11. Y. Ollivier. "Ricci Curvature of Markov Chains on Metric spaces." *Journal of Functional Analysis*. 2009.
12. Y. Ollivier. "Ricci Curvature of Metric Spaces." *Math. Acad. Sci. Paris*. 2007.
13. R. Sandhu, T. Georgiou, E. Reznik, L. Zhu, I. Kolesov, Y. Senbabaoglu, and A. Tannenbaum. "Graph Curvature for Differentiating Cancer Networks." *Nature (Scientific Reports)*. 2015.
14. R. Sandhu, T. Georgiou, and A. Tannenbaum. "Ricci Curvature: An Economic Indicator for Market Fragility and Systemic Risk." *Science Advances*. 2016.
15. X. Pennec, P. Fillard, and N. Ayache. "A Riemannian Framework for Tensor Computing." *IJCV*. 2005.
16. V. Arsingy, P. Fillard, X. Pennec, and N. Ayache. "Log-Euclidean metrics for fast and simple calculus on diffusion tensors." *Magnetic Resonance in Medicine*. 2006.
17. P. Jupp and K. Mardia. "A unified view of the theory of directional statistics." *International Statistical Review*. 1989.
18. M. Moakher. "Means and averaging in the group of rotations." *SIAM Journal of Matrix Analysis and Applications*. 2002.
19. A. Edelman, T. Arias, and S. Smith. "The Geometry of Algorithms with Orthogonality Constraints." *SIAM Journal of Matrix Analysis and Applications*. 2002.
20. M. DoCarmo. *Riemannian Geometry*. Birkhauser. 1992.
21. P. Fletcher and S. Joshi. "Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors." *ECCV*. 2004.
22. K. Nomizu. "Invariant affine connections on homogeneous spaces." *American Journal of Math*. 1954.
23. Y. Rathi, A. Tannenbaum, O. Michailovich. "Segmenting Images on the Tensor Manifold." *CVPR*. 2007.
24. S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. 1978.
25. J. Lott and C. Villani. "Ricci Curvature for Metric-Measure Spaces via Optimal Transport." *Annals of Mathematics*. 2009.
26. K. Sturm. "On the geometry of metric measure spaces." *Acta Math*. 2006.
27. F. Bauer, J. Jost, and S. Liu. "Ollivier-Ricci Curvature and the Spectrum of the Normalized Graph Laplace Operator." *arXiv*. 2013.