Investigating Higher-Order Connectivity for Epidemic Spread

Colleen M. Farrelly*
Post Urban Ventures
Miami Florida, USA

Franck Kalala Mutombo†
University of Lubumbashi
Dept. of Math. and Comp. Sci.
Lubumbashi, DRC

February 16, 2025

Abstract

This study examines the role of higher-order connectivity in epidemic spread across networks and considers Hodge Laplacian traces to classify network types based on higher-order influence on epidemic severity. Application of simulation results to a social network of dolphins suggests real-world applicability of our main findings. Epidemic behavior is based on rich varieties of simplices included in the simplicial complex.

Higher-order connectivity, Epidemic spread, Hodge Laplacian, Simplicial complex, Epidemic severity, Higher-order structures

1 introduction

Graph objects play an increasingly central role in data science and machine learning [19, 8]. For instance, epidemic modeling, such as susceptibility-infected-remove (SIR) models, has elucidated spread on social and travel networks in recent COVID-19 and Ebola outbreaks [5, 18]; with the rise of emerging diseases, it is likely network-based epidemic models will play a role in future outbreak planning and response. However, current approaches to epidemic modeling on networks only consider pairwise interactions, such as two people who interact in each time frame or two towns connected by a transportation route. Recent advances in topological data analysis and topological deep learning, which include pairwise and higher-order structures of n-way interactions through simplicial complex extensions of graph data, have improved predictions in medicine [21, 20, 12] and captured data insights missed by graph-based methods [21, 10]. Some prior work has created simplicial complex extensions of SIR models to capture the role of higher-order interactions on social networks [4, 16]. However, these recent attempts do not explore spectral graph properties of the graph Laplacian and its higher order analogue, the Hodge Laplacian, that might correlate with epidemic potential on simplicial complex models. The spread capacity and flux measurements derived from Laplacians should relate the behavior of other differential equations on the graph/simplicial complex beyond simply heat flow equations. Our contributions include extending spectral analysis of Laplacians to measure potential epidemic impact on a variety of simplicial complexes, including a simulation study of random simplicial complexes and a real social network of dolphins. We hope that this will open the door to spectral analysis of simplicial complexes in relation to general spreading processes.

2 Method

2.1 SIR models on graphs

Epidemic models aim to capture the dynamics involved in the spreading of a disease, an idea, a computer virus, obesity or even product adoption [7]. They focus on addressing several key aspects of contagion dynamics, such as understanding how contagions spread within populations and determining whether a disease will escalate into an epidemic. These models also explore practical applications, including identifying the most effective individuals to vaccinate, predicting whether a YouTube video will go viral, and selecting the optimal target audience for marketing campaigns to maximize product penetration. Classic epidemiological models operate under the assumption of full mixing, where every individual has equal probability of interacting with any other individual in the population. However, relaxing this assumption by incorporating an underlying contact network results in more realistic models that account for the structure of interactions

*cfarrelly786@gmail.com
†mutombof@unilu.ac.cd

[2]. Among the various models available, this paper focuses on the Susceptible-Infected-Removed (SIR) model applied to networks. Without loss of generality if $S(t)$ represents the number of susceptible individuals at time t, $I(t)$ the number of infectious individuals, $R(t)$ the number of recovered individuals, β the transmission rate of infection, and γ the recovery rate, N the size of the network/total number of nodes, the dynamics of the SIR model on networks read:

$$\frac{ds}{dt} = -\beta < k > si, \quad \frac{di}{dt} = \beta < k > i - \gamma i, \quad \frac{dr}{dt} = \gamma i, \quad \text{with} \quad s + i + r = 1. \tag{2.1}$$

where $S/n = s$, $I/N = i$ and $r/N = R$. In the equations (2.2) we have assumed that individuals have approximately $< k >$ possibilities of contagion from neighbors, with $< k >$ being the average degree of the network. This fact is true when the corresponding network is homogeneous like a random network, Erdös-Rényi network, and lattices. Many real-world networks are characterized by highly heterogeneous topologies [6]. Recent empirical evidence [2] highlights the significance of this heterogeneity, demonstrating that many epidemiological networks exhibit heavy-tailed distributions. As a result, the average degree $< k >$ is no longer the primary determinant in such networks. Instead, fluctuations are expected to play a crucial role in shaping the properties and dynamics of epidemics. As a consequence of the topological fluctuations, we need a mathematical model that will take into account the degree of variability of nodes or vertices. The details about the mathematical models of epidemics on heterogeneous networks can be found in [2, 3]. In this paper we will be using the Laplacian and its high-order version to study spreading process using SIR model.

2.2 Simplicial Complex, Hodge Laplacian of a Graph and its trace

We give a pedestrian introduction to simplicial complexes without having the intention to go deeper in many details. Interested readers find technical details in [15]. A collection of n distinct vertices is called an n-simplex (or simply a simplex). A simplex is essentially a finite set of vertices, and its dimension is determined by the cardinality of the set. The orientation of a simplex is defined as an equivalence class of vertex orderings. An abstract simplicial complex is a collection of simplices with the property that if a simplex is included in the complex, every non-empty subset of that simplex is also included. The dimension of a complex is the highest cardinality among its simplices. An oriented complex is one in which an orientation is specified for all simplices, though the choice of orientations is arbitrary for our purposes. Now let $G = (V, E)$ be a simple undirected graph where V is a finite set of vertices and E is the set of edges. The set comprising all cliques of a graph is a simplicial complex called clique complex of the graph G, where a clique is a subset of vertices such that every pair of distinct vertices within the subset is connected by an edge. In other words, a clique in a graph G is an induced subgraph of G that is complete. The set of 1-clique consists of the set of vertices, the set of 2 cliques consist of edges and the set of 3 cliques consists of triangles. In this framework, a graph can be viewed as a 1-complex consisting of 0-simplices (vertices) and 1-simplices (edges). Extending this 1-complex by including cliques of the graph enriches it to a higher-dimensional complex. For example, adding all 3-cliques (triangles) to the graph results in a simplicial 2-complex. The Hodge Laplacian of a graph is a generalization of the graph Laplacian to high-order simplicial complexes. For a graph (a 1-simplicial complex) we have the 0-Hodge Laplacian and 1-Hodge Laplacian of a graph which can be explicitly defined in terms of matrices in the context of graphs. The 0-Hodge Laplacian ($\mathbf{\Delta}_0$) operates on $0-$cochains (functions on vertices) and is defined as:

$$\mathbf{\Delta}_0 = \mathbf{B}_0^T \mathbf{B}_0 \tag{2.2}$$

where,

$$(\mathbf{B}_0)_{e,v} = \begin{cases} +1, & \text{if vertex } v \text{ is the head of edge } e, \\ -1, & \text{if vertex } v \text{ is the tail of edge } e, \\ 0, & \text{otherwise.} \end{cases}$$

$\mathbf{\Delta}_0$ is equivalent to the graph Laplacian matrix, which captures the connectivity between vertices, that is, $\mathbf{\Delta}_0 = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the degree matrix and $\mathbf{D}$ the adjacency matrix of the graph. The 1-Hodge Laplacian ($\mathbf{\Delta}_1$) operates on $1-$cochains (functions on edges) and is defined as:

$$\mathbf{\Delta}_1 = \mathbf{B}_1 \mathbf{B}_1^T \tag{2.3}$$

where $\mathbf{B}_1\mathbf{B}_1^T$is the matrix that encodes relationships between edges based on the vertices they share. If the graph is extended to a simplicial complex with triangles (2-simplices), the 1-Hodge Laplacian becomes:

$$\mathbf{\Delta}_1 = \mathbf{B}_0 \mathbf{B}_0^T + \mathbf{B}_1^T \mathbf{B}_1 \tag{2.4}$$

where $\mathbf{B}_1$ is the incidence matrix for triangles (mapping edges to triangles). The first term $\mathbf{B}_0\mathbf{B}_0^T$ reflects edge relationships via vertices and the term $\mathbf{B}_1^T\mathbf{B}_1$ reflects edge relationships via triangles. These matrix definitions allow the Hodge Laplacian to be computed numerically and analyzed in applications such as spectral graph theory and topological data analysis. For advanced technical details one can visit [13]. The trace of the Hodge Laplacian is the sum of its eigenvalues.

We can find a close formula of the trace of the 0-Hodge Laplacian from (2.2), $tr(\mathbf{\Delta}_0) = tr(\mathbf{B}_0^T\mathbf{B}_0) = \sum_{v\in V} deg(v) = 2|E|$. Likewise, we have from (2.4) we have $tr(\mathbf{\Delta}_1) = tr(\mathbf{B}_0\mathbf{B}_0^T) + tr(\mathbf{B}_1^T\mathbf{B}_1) = 2|E| + 3|T|$, where $|E|$ and $|T|$ count respectively the number of edges and the number of triangle (2-simplices) in the graph. The trace of $\mathbf{\Delta}_1$ provides a measure of total connectivity between edges via both vertices and higher-order simplices (triangles). This is particularly important in applications like topological data analysis (TDA), where higher-order simplicial interactions are analyzed. In the next section we will use the trace of the Hodge Laplacian to study the spreading process in the framework of Susceptible Infected Remove (SIR) in a population of the animal social network, the Dolphin Network.

2.3 Simulation and datasets

We simulated random networks of 20 vertices with connectivity probabilities ranging from 0.1 to 0.9 using 0.1 increments. We then calculated the $0-$ and $1-$Hodge Laplacian traces and simulated an epidemic with a reproductive fraction of 0.4 (beta=0.1, gamma=0.25) and initial infection of one randomly picked vertex, representing an epidemic that is normally unlikely to spread widely through a population but may spread more than hypothesized in highly connected networks with higher-order structures. We ran our simulation 20 times and calculated mean and standard deviation for maximum number infected and our trace metrics. We then examined a real social network of dolphins [14] consisting of 62 dolphins in a connected network. We again calculated the 0-Hodge and 1-Hodge Laplacian trace of the network and ran 20 epidemic simulations using the same parameters as we used in the simulation study to examine the simplicial complex structures and start points of the epidemic to understand the impact of higher-order structure on real social networks.

3 Results

Our simulation study suggests three main groups of networks by trace and epidemic variability: sparse, dense, and in-between groups. Sparse graphs tend to have less severe epidemics and less complex simplicial complex structures (as expected); dense graphs tend to have more severe epidemics and high-order simplices in their simplicial complexes (as expected). However, the in-between group mixes sparse and dense connections, yielding a variety of simplicial structures that impact epidemic severity based on where the epidemic starts and where it initially spreads. Table 1 shows the Laplacian traces and epidemic severity (means across simulations with standard deviation in parentheses).

	Average Graph Trace	Average Simplex Trace	Average Number Infected
p=0.1	42 (10)	47 (14)	2 (1)
p=0.2	76 (10)	101 (21)	3 (2)
p=0.3	113 (11)	204 (40)	6 (4)
p=0.4	152 (16)	375 (79)	7 (3)
p=0.5	189 (13)	607 (104)	9 (4)
p=0.6	227 (11)	934 (104)	10 (3)
p=0.7	265 (11)	1433 (160)	11 (2)
p=0.8	305 (9)	2074 (179)	12 (3)
p=0.9	341 (11)	2818 (241)	13 (2)

Table 1: A summary of random graph simulation parameters and their average traces and number of infected individuals

The dolphin network falls into the in-between group, where lower-dimensional and high-dimensional simplices exist in the simplicial complex (Figure 1). As expected for an in-between network, the severity of the epidemic varied greatly depending on start point (from 3 to 19 total infected dolphins).

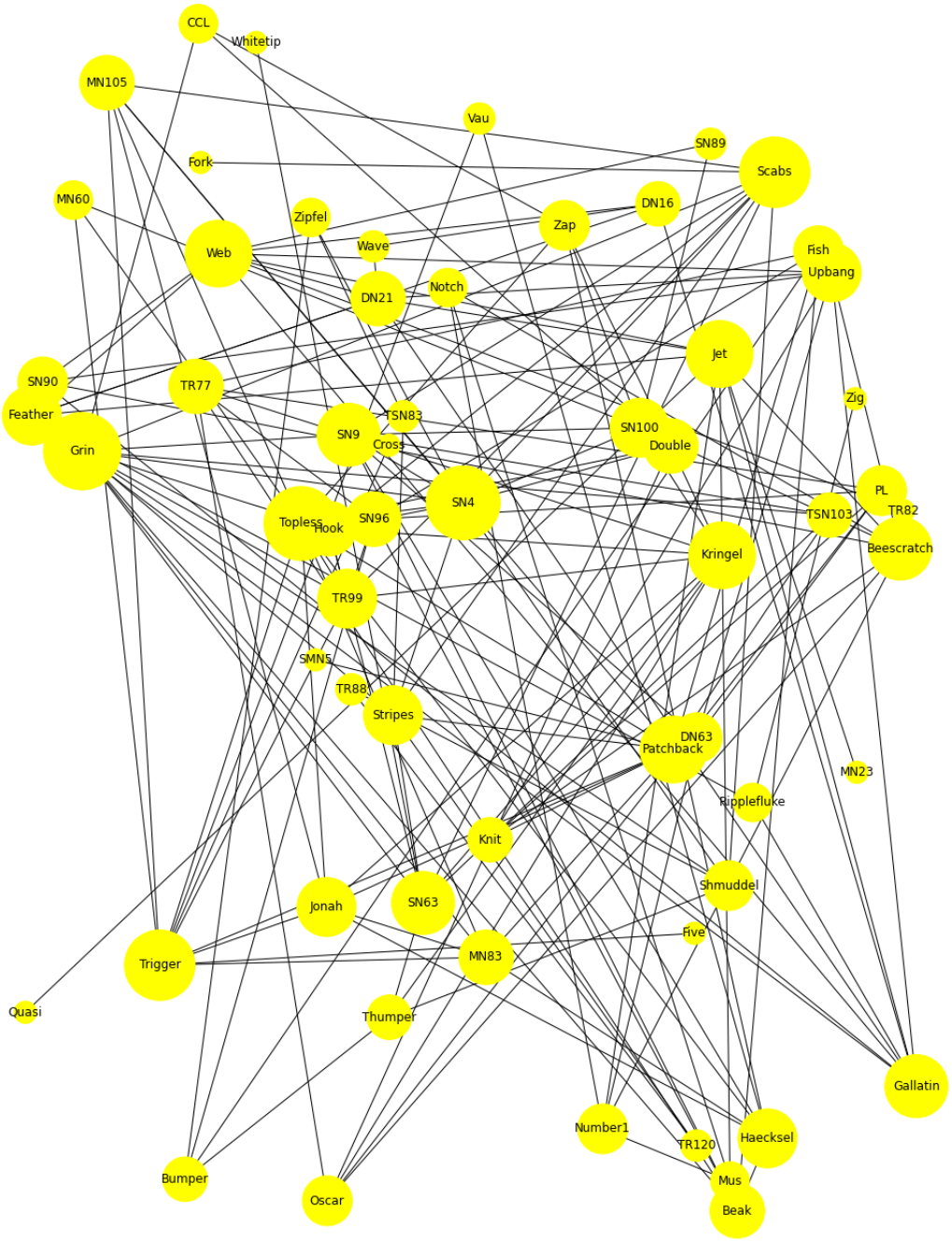

Figure 1: A plot of the dolphin network

4 Conclusion

Our results suggest in-between networks have highly variable epidemic behavior depending on higher-order structures and the exact start-point of an epidemic (sparse or dense regions). Given that Laplacian traces identify these types of graphs easily, we suggest classifying networks of interest with trace metrics prior to modeling such that researchers can examine higher-order structures of in-between networks with varying simplicial complex features.We hope to expand our analysis to other Hodge Laplacian spectral properties that might be relevant for epidemic behavior on higher-order networks (such as energy, algebraic connectivity, zeta function dervivatives, torsion) and consider persistent versions of Hodge Laplacian spectra for weighted graphs (such as those with edges based on frequency of interaction, hours spent in a location), as well as dynamic graphs that change connectivity over time. Recent work suggests the efficacy of such explorations in fields like neuroscience [1], single cell biology [17], and community detection [11]. Theoretical work on Hodge Laplacian spectra is ongoing and promises new tools and insights for network scientists [13].

References

[1] D. Anand and M. Chung. Hodge laplacian of brain networks. *IEEE Transactions On Medical Imaging*, 42:1563–1573, 2023.

[2] A. Barrat, M. Barthélemy, and A. Vespignani. *Dynamical Processes on Complex Networks.* Cambridge University Press, New York, NY, USA, 1st edition, 2008.

[3] M. Boguñà and R. Pastor-Satorras. Epidemic spreading in correlated complex networks. *Physical Review E*, 66: 047104, 2002.

[4] Y. Chen, Y. Gel, M. Marathe, and H. Poor. A simplicial epidemic model for covid-19 spread analysis. *Proceedings Of The National Academy Of Sciences*, 121:e2313171120, 2024.

[5] I. Cooper, A. Mondal, and C. Antonopoulos. A sir model assumption for the spread of covid-19 in different communities. *Chaos, Solitons & Fractals*, 139:110057, 2020.

[6] E. Estrada. Quantifying network heterogeneity. *Physical Review E*, 82:066102, 2010.

[7] E. Estrada, F. Kalala-Mutombo, and A. Valverde-Colmeiro. Epidemic spreading in networks with nonrandom long-range interactions. *Physical Review E*, 84(3 Pt 2):036110, 2011. doi: 10.1103/PhysRevE.84.036110.

[8] C. Farrelly and F. Mutombo. *Modern Graph Theory Algorithms with Python: Harness the power of graph algorithms and real-world network applications using Python.* Packt Publishing Ltd, 2024.

[9] V. Grande and M. Schaub. Disentangling the spectral properties of the hodge laplacian: not all small eigenvalues are equal. In *ICASSP 2024-2024 IEEE International Conference On Acoustics, Speech And Signal Processing (ICASSP)*, pages 9896–9900, 2024.

[10] M. Hajij, G. Zamzmi, T. Papamarkou, N. Miolane, A. Guzmàn-Sáenz, K. Ramamurthy, Others, and M. Schaub. Topological deep learning: Going beyond graph data. *arXiv preprint*, 2022. arXiv:2203.02702.

[11] S. Krishnagopal and G. Bianconi. Spectral detection of simplicial communities via hodge laplacians. *Physical Review E*, 104:064303, 2021.

[12] R. Levenson, Y. Singh, B. Rieck, Q. Hathaway, C. Farrelly, J. Rozenblit, Others, and D. Sarkar. Advancing precision medicine: algebraic topology and differential geometry in radiology and computational pathology. *Laboratory Investigation*, 104:102060, 2024.

[13] L. Lim. Hodge laplacians on graphs. *SIAM Review*, 62:685–715, 2020.

[14] D. Lusseau, K. Schneider, O. Boisseau, P. Haase, E. Slooten, and S. Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations: can geographic isolation explain this unique trait? *Behavioral Ecology And Sociobiology*, 54:396–405, 2003.

[15] James R. Munkres. *Elements of Algebraic Topology.* Westview Press, India, 1984.

[16] G. Palafox-Castillo and A. Berrones-Santos. Stochastic epidemic model on a simplicial complex. *Statistical Mechanics*, 2022.

[17] X. Qiu, Y. Zhang, J. Martin-Rufino, C. Weng, S. Hosseinzadeh, D. Yang, Others, and J. Weissman. Mapping transcriptomic vector fields of single cells. *Cell*, 185:690–711, 2022.

[18] A. Rizzo, B. Pedalino, and M. Porfiri. A network model for ebola spreading. *Journal Of Theoretical Biology*, 394: 212–222, 2016.

[19] F. Scarselli, M. Gori, A. Tsoi, M. Hagenbuchner, and G. Monfardini. The graph neural network model. *IEEE Transactions On Neural Networks*, 20:61–80, 2008.

[20] Y. Singh, C. Farrelly, Q. Hathaway, T. Leiner, J. Jagtap, G. Carlsson, and B. Erickson. Topological data analysis in medical imaging: current state of the art. *Insights Into Imaging*, 14:58, 2023.

[21] R. Wei, J. Wee, V. Laurent, and K. Xia. Hodge theory-based biomolecular data analysis. *Scientific Reports*, 12: 9699, 2022.