Technical Report

Alice NANYANZI (alicenanyanzi@aims.ac.za)

April 4, 2018

Outline of the Thesis

1. Abstract
2. Introduction
   - Brief intro of networks
   - Say what you are going to do
   - why and how are you going to do the above
   - Context - compare to others
   - Structure of the thesis
3. Chapter 1: Review of networks
   - Similarity in networks
   - Voronnoi Tensselation
   - Delauny graphs
   - Harris corner detection
   - Image manipulation such as cropping, rotation, etc
   - Data analysis include PCA, etc
4. Chapter 2: Diffusion on networks
5. Chapter 3: Laplacian centrality of an edge
   - Laplacian centrality of a node
   - Motivation for edge centrality
6. Chapter 4: New horizon
   - Heat kernel Centrality
   - Communicability Centrality with k-hop (Do some toy models)
7. Conclusion

1 Review of Networks

1.1 Complex Systems and Complex Networks

Complex systems are very vital in our daily lives. They exist in fields such as social, economic, science, technology among others. During his interview with San Jose Mercury News in January 2000, Stephen Hawking referred to the 21st century as a century of complexity. Complex systems are composed of interconnected components,however, it has been observed that many complex systems display a behaviour phenomena (also known as emergent behaviour) that cannot be explained by any conventional analysis of the system's constituent parts (Casti, September 26, 2017). In otherwords, for one to understand the behaviour of a system, it is necessary for one to consider a holistic system-level view point. There are different approaches to study of complex systems for instance statistical description, empirical data analysis, simulations, analytical approach and network approach. In this work, we focus on the network approach in which we represent complex systems by complex networks whose nodes (vertices) and links(edges) represent the components and the interactions among components respectively. For example, a transportation system can be represented by network where nodes are cities or towns and the links are the roads, railways or flight routes. This is then followed by mathematical formulation of the problem, modelling and validation. An interesting early historical application of the network approach to the study of complex systems is the Käonigsberg bridge problem where Euler (Euler, 1976, 1953) solved the problem by reformulating problem in terms of a graph where vertices represent islands while edges represent the seven bridges joining any two islands. Work published by Leonhard Euler (Euler, 1976) is considered the genesis of the story of network theory. As the size of network increases from just graphs of tens or hundreds of nodes which could easily be analysed by direct use of eye so as to ascertain the structure of the network to complex networks consisting of million or billion of nodes which call for advanced analytic approach that involves development of statistical methods to quantify such large networks. The statistical methods aid in answering questions such as how many nodes or edges should be removed for the network to break down?, what is the shortest path length of the network?, and many others.

1.1.1 Characteristics of complex systems

1. Emergence

   Complex systems are emergent phenomena as they are the spontaneous outcome of the interactions among the many constituent units (Barrat et al., 2008). Complex systems, unlike engineering systems, they are not formed based on a blue print. They instead evolve to form emergent architecture and they exhibit unexpected properties and characteristics.

2. Self organisation

   Complex systems are characterised by self organising properties that are exhibited through the collective and unsupervised dynamics of the components of the system. The collective behaviour justifies the fact that one cannot understand the whole system and its dynamics by dismantling the system and studying each component in isolation. Taking the internet for example, there is no central authority responsible for controlling the addition of new computers(nodes) or connection lines (links) onto the network. Another example is the eco-system in which feeding pattern followed by animals is in no way controlled by any central agent.

3. No central organising mind

   In complex systems, it appears as though there exists a central control. However, in reality, control is spread over a decentralised structure that is to say system behaviour is due to the combination of the different components that make up the system (Casti, September 26, 2017). This property of complex systems makes them resilient as failure of some of the components may not critically affect the functioning of the whole system.

4. Uncertainty

Complex systems exhibit an unpredictable kind of behaviour. For instance, in a transportation system of a city, opening new free ways may result into an increase in traffic jam.

Other characteristics of complex systems include their evolving and adaptive nature. In addition, no conventional way can be used to describe complex systems.

1.1.2 Real-world Networks

In his work (Newman, 2003), Newman considered a loose categorisation of networks: social networks, communication networks, technological networks, and biological networks.

a. Social Networks

Networks considered as social networks are ones whose nodes correspond to people or groups of people while the edges represent the interactions or relationship between them (Jackson, 2010). For instance friendship networks such as facebook, twitter in which the interactions represent friendship ties among acquaintances, networks of intermarriages between families, social interaction networks which capture peoples' interactions through social activities or events, employee networks with companies, and many others. Some common networks that researchers have frequently experimented upon include: the Zachary karate network which consists of two communities centred at the administrator and instructor as a result of misunderstanding that prevailed with the karate club earlier on. The nodes in the network are the members of the club as the links represent interactions between members during non-club activities (Zachary, 1977). Other networks include the Dolphine network (Williams et al., 1993), terrorist networks (Magouirk et al., 2008) among others .

b. Information networks:

Information networks are also referred to as knowledge networks. Examples of networks under this category include: The world wide web which consists of billions of web pages as nodes that are linked together through links known as hyperlinks (Huberman, 2001). Another network categorised as information networks are citation networks that are composed of nodes which are articles while directed link between two nodes written as $i \longrightarrow j$ indicate article i cites article j.

c. Technological networks:

This category consist of networks made by man to aid in distribution or transfer of resources, services or commodities such as electricity, water, transportation services, and many others. Examples of such networks include the internet, transportation networks, power grids, to mention but a few (Faloutsos et al., 1999; Pagani and Aiello, 2013; Banavar et al., 1999).

d. Biological networks:

Biological networks exists in areas related human and processes that take place with in the human body, animals and their ways of survival, chemistry. Such networks are the human brain network, protein-protein interaction network, network of metabolic path ways, ecological networks (Estrada, 2011; Sporns et al., 2004; Schwikowski et al., 2000).

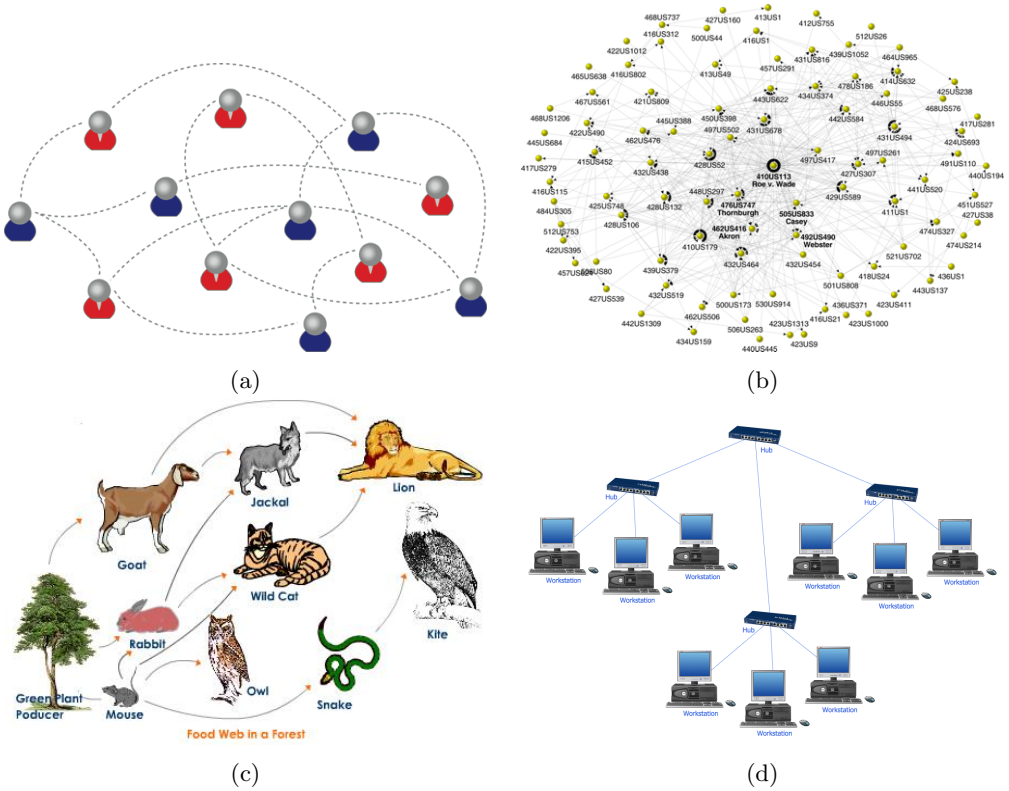


Figure 1: Networks in real world: (a) A social network. (b) A citation network. (c) A food web. (d) Computer network. Source: (Internet)

1.2 Paths,Walks and Connected Components of a Graph

Definition 1 (Walk). *A walk in a network is a series of edges (not necessarily distinct)*

$$(u_1, v_1), (u_2, v_2), \ldots, (u_l, v_l), \qquad \text{for which } v_i = u_{i+1} \ (i = 1, 2, \ldots, l-1).$$

A walk from v_1 to v_{l+1} is one of length l (Estrada et al., 2015).

A walk of length k is referred to as a k-walk. We can compute the number of k-walks between any pair of nodes in a simple undirected unweighed graph using the entries of $\mathbf{A}^k$ where $\mathbf{A}$ is the adjacency matrix whose entries a given by

$$\mathbf{A}_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Definition 2 (Path). *A path of length l is a walk of length l in which all the nodes and edges are distinct. A closed path is called a cycle (Estrada, 2011). For any pair of nodes v_i,v_j in a connected graph, there exists at least one path connecting v_i to v_j. The paths with minimum length are referred as shortest-paths.*
Definition 3 (Irreducible set of shortest paths). *An irreducible set of shortest paths of length l is the set $P_l = P_l(v_i, v_j), P_l(v_i, v_r), \cdots, P_l(v_s, v_t)$ in which the endpoints of every shortest-path $P_l(v_i, v_j)$ in the set are different. Each path in this set is referred to as an irreducible shortest-path (Estrada, 2012).*
Definition 4 (Connected component of a graph). *A component of an undirected graph is a subgraph in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the supergraph (Newman, 2010). A connected component is also referred to as a maximal connected subgraph of a graph.*
Definition 5 (Transition matrix of a random walk on a Graph). *It is the $n \times n$ matrix $\mathbf{P}_G$ whose entries are given by*

$$\mathbf{P}_{ij} = \frac{1}{d_i} \mathbf{A}_{ij}, \tag{2}$$

In otherwords, $\mathbf{P}_G = \mathbf{D}_G^{-1}\mathbf{A}$

1.3 Laplacian Matrix

The Laplacian matrix is one of matrices used to represent a graph or network. Recently, a number of researchers have been deeply involved in the study of the Laplacian matrix of a graph since this matrix has interesting spectral properties that provide more useful information about the structure of a graph as compared to other matrices such as the adjacency matrix. The Laplacian matrix has various applications that we will explore in subsequent sections.

1.4 Definitions and Properties of the Laplacian Matrix

Definition 6 (Laplacian Matrix)**.** *The Laplacian matrix of a network is the difference between the Degree matrix* $\mathbf{D}$ *and the Adjacency matrix* $\mathbf{A}$ *of a network. That is,*

$$\mathbf{L} = \mathbf{D} - \mathbf{A}. \tag{3}$$

Given a simple network $G = (V, E)$*, the entries of the Laplacian matrix* $\mathbf{L}(\mathbf{G})$ *are defined as*

$$L_{ij} = \begin{cases} k_{v_i} & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where k_{v_i} *denotes the degree of node* i *(Estrada, 2011).*

Alternatively, we can define the Laplacian matrix of a graph in terms of the vertex-edge incidence matrix $\mathbf{B}$*. That is,*

$$\mathbf{L} = \mathbf{B}\mathbf{B}^T, \tag{5}$$

where $\mathbf{B}^T$ *is the transpose of* $\mathbf{B}$ *(Estrada, 2011).*

The properties of the Laplacian matrix of a network are very useful in understanding the structure of a network. We therefore discuss some of these properties.

1. Real and symmetric matrix

   The entries of the Laplacian matrix are real numbers and are symmetric with respect to the main diagonal (Das, 2004). Thus, the spectrum is real.

2. Singular matrix

   The Laplacian matrix is a square matrix that is not invertible. Its determinant is equal to zero (Das, 2004).

3. Positive semi-definite

   A matrix is positive semi-definite if and only if all its eigenvalues are non-negative. For a given matrix $\mathbf{L}$, this property is denoted by $\mathbf{L} \geq 0$.

   Suppose the eigenvalues of a Laplacian matrix are: $\lambda_1, \lambda_2, \ldots, \lambda_n$, we show that $\lambda_i \geq 0 \ \forall \ i$.

   Proof. Let $\mathbf{v}_i$ be a normalised eigenvector of $\mathbf{L}$ with eigenvalue λ_i. Then, using Equation (5)

   $$\mathbf{v}_i^T\mathbf{B}^T\mathbf{B}\mathbf{v} \quad = \quad \mathbf{v}_i^T\mathbf{L}\mathbf{v}_i = \lambda_i\mathbf{v}_i^T\mathbf{v}_i = \lambda_i.$$

   Thus, any eigenvalue λ_i of the Laplacian matrix is equal to $(\mathbf{v}_i^T\mathbf{B}^T\mathbf{B}\mathbf{v}_i)$ which can be written as $(\mathbf{v}_i\mathbf{B})^T(\mathbf{B}\mathbf{v}_i)$. The quantity $(\mathbf{v}_i\mathbf{B})^T(\mathbf{B}\mathbf{v}_i)$ is the inner product of a real vector $(\mathbf{B}\mathbf{v}_i)$ with itself, that is, the sum of the squares of the real elements of the vector $(\mathbf{B}\mathbf{v}_i)$ and hence it is non-negative (Estrada et al., 2015). □

1.4.1 Spectrum of the Laplacian Matrix

As mentioned earlier, the spectrum of the Laplacian provides useful information about the structure of a network. The spectrum of the Laplacian matrix is the set of all its eigenvalues and their multiplicities (Estrada, 2011). Let $\lambda_1 < \lambda_2 < \ldots < \lambda_n$ be the distinct eigenvalues of $\mathbf{L}$ and let $m(\lambda_1), m(\lambda_2), \ldots, m(\lambda_n)$ be their multiplicities. Then, the spectrum of $\mathbf{L}$ is written as

$$Sp\mathbf{L} = \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ m(\lambda_1) & m(\lambda_2) & \ldots & m(\lambda_n) \end{pmatrix}. \tag{6}$$

We consider the non increasing order of the eigenvalues of $\mathbf{L}$: $\lambda_n \geq \lambda_{n-1} \geq \cdots \geq \lambda_2 \geq \lambda_1 = 0$. Some of the results associated with the spectrum of the Laplacian matrix include:

- The eigenvalues of $\mathbf{L}$ are bounded as $0 \leq \lambda_j \leq 2k_{max}$ and $\lambda_n \geq k_{max}$ (Estrada, 2011).
- The eigenvalue λ_1 is always equal to zero (Estrada, 2011). Atleast one eigenvalue of the Laplacian is 0.

  Proof. Consider the vector $v = (1/\sqrt{n}, \cdots, 1/\sqrt{n})$. We know that the ith entry of Lv is $\sum_{i\sim j} v_{(i)} - v_{(j)} = \sum_{i\sim j}(1/\sqrt{n} - 1/\sqrt{n}) = 0 = 0 \cdot v_{(i)}$. □

- The multiplicity of 0 as an eigenvalue of $\mathbf{L}$ is equal to the number of connected components in the network (Estrada, 2011).
- Every row sum and column sum of $\mathbf{L}$ is zero. Thus, the vector $\mathbf{v_1}$ of all ones is an eigenvector associated with $\lambda_1 = 0$, since $\mathbf{Lv_1} = \mathbf{0}$ (Das, 2004).
- A network is connected if its second smallest eigenvalue is nonzero. That is, $\lambda_2 > 0$ if and only if G is connected. The eigenvalue λ_2 is thus called the algebraic connectivity of a network, $a(G)$. The magnitude of this value depict how well connected the over all graph is. The algebraic connectivity has significant implications for properties such as clustering and synchronizability. The eigenvector corresponding to the eigenvalue λ_2 is called the Fiedler vector (Estrada et al., 2015).
- Let G be a graph with connected components $G_i (1 \leq i \leq s)$. Then the spectrum of G is the union of the spectra of G_i (and multiplicities are added) (Brouwer and Haemers, 2011).
- For a graph G, the sum of the eigenvalues, that is, the trace of L is twice the number of edges of G. Mathematically, $\sum_{i=1}^{n} \lambda_i = Tr(L) = 2E$. (Brouwer and Haemers, 2011)

Theorem 1 (Fiedler, 1975)**.** *Suppose $G = (V, E)$ is a connected network with graph Laplacian $\mathbf{L}$ whose second smallest eigenvalue is $\lambda_2 > 0$. Let x be the eigenvector associated with λ_2. Let $r \in \mathbb{R}$ and partition the nodes in V into two sets*

$$V_1 = \{i \in V | x_i \geq r\},\ V_2 = \{i \in V | x_i < r\}, \tag{7}$$

then the subgraphs of G induced by the sets V_1 and V_2 are connected (Estrada et al., 2015).

This result is useful for partitioning a network while ensuring that all the parts remain connected. This method of partitioning using eigenvalues is known as spectral clustering. For clusters of equal size, we choose r such that it is the median value of x.

Some analytic expressions for the spectra of different kinds of simple networks are:

- Star, S_n : $Sp(\mathbf{L}) = \{0\ 1^{n-2}\ n\}$.
- Complete, K_n : $Sp(\mathbf{L}) = \{0\ n^{n-1}\}$.
- Complete bipartite, $K_{m,n}$: $Sp(\mathbf{L}) = \{0\ m^{n-1}\ n^{m-1}\}$ (Estrada, 2011).

Theorem 2 (Kirchoff's Matrix-Tree Theorem)**.** *If G is a connected graph with Laplacian matrix $\mathbf{L}$, then the number of unique spanning trees of G is equal to the value of any cofactor of the matrix $\mathbf{L}$ (Harris et al., 2008).*

1.5 Robustness of Complex Systems

A complex systems is considered robust if it can with stand failures or perturbations that is to say a system can still perform as expected even in circumstances of failure of one or more components in the system. In other words, robustness intuitively deals with the existence of back-up possibilities. In a network, this cab be captured in the existence of alternative paths with in the network. Robustness of systems plays an important role in a number of fields for instance in Engineering, understanding robustness acts as a basis for designing communication, transportation systems, power grids that can perform basic operation despite failure of some system components. In biology,robustness explains why some mutations lead to diseases while others do not. For ecologists and environmental experts, robustness helps in predicting the failure of an ecosystem when faced with disruptive human behaviours. In general, study of system robustness aids in understanding system operation, improving system performance and designing of new robust systems. As mentioned earlier, the study of a network or graph underlying a complex system provides insights about the properties and characteristics of that systems. Thus, Barabasi (Barabási, 2016) highlighted the fact that networks play a vital role in robustness of complex systems which implies that exploring robustness of the network reflects that of the system.

1.6 Robustness measures in networks

According to Ellens (Ellens and Kooij, 2013), robustness of a network is its ability to perform well when subject to failures or attacks. The attacks take on two forms namely: random attacks and targeted attacks. However, in order to tell whether a particular network is robust, there is need for a measure that quantifies the robustness. In the past, various robustness measures have been put forward by researchers (Sydney et al., 2008). We explore some of the common measures of robustness in networks. We categorise the measures as follows:

1.6.1 Connectivity-based measures

Here we consider robustness measures that are based on the connectivity of the network. These include the classical connectivity κ, edge connectivity κ_e, and vertex connectivity κ_v. Firstly, the classical connectivity κ is a measure whose value $\kappa = 1$ for graphs in which there is a path between any pair of vertices that is connected graphs and $\kappa = 0$ for unconnected graphs that is graphs in which atleast one pair of vertices for which no path exists between them. Secondly, the edge connectivity κ_e and vertex connectivity κ_v are respectively the minimum number edges and vertices that need to be removed to disconnect the graph. The inequality $\kappa_v \leq \kappa_e \leq \delta_{min}$ holds for non-complete graphs, where δ_{min} is the minimum degree of vertices in a graph.

1.6.2 Distance-based measures

1. Diameter

   The diameter of a graph, denoted as D, is the maximum distance between pairs of nodes in the graph(Wang and Chen, 2003). The diameter of a graph $G = (V, E)$ is defined as

   $$D = \max_{i,j \in V} \{d_{ij}\},$$

   where d_{ij} is the shortest path between node i and j. Based on the diameter, a graph is considered more robust if it's diameter is shorter.

2. Average Path Length

   The average path length of a network is the average number of steps along the shortest paths for all possible pairs of network nodes. Let $G = (V, E)$ be a graph the average path length L_G is defined

by

$$L_G = \frac{1}{n(n-1)} \sum_{i,j \in V,\ i \neq j} d_{ij}, \tag{8}$$

where d_{ij} is the shortest path between node i and j and n is the total number of nodes in G (Wang and Chen, 2003). On comparing the average path length and diameter as measures of robustness, the former is considered a better measure as it strictly decreases on addition of edges which is not necessarily the case with the latter.

3. Efficiency

   We can observe that we cannot compute the robustness of disconnected graph based on the two distance-based measures discussed previously. However, this is a possibility when we adopt the efficiency measure. The efficiency of a graph, E is defined as

$$E = \frac{1}{n(n-1)} \sum_{i,j \in V, i \neq j} \frac{1}{d_{ij}}. \tag{9}$$

   It is important to note that these measures based on distance consider only shortest path distances which implies that other alternative paths are not put into consideration which is a disadvantage for that matter.

1.6.3 Spectral Graph measures

1. Algebraic connectivity

   Given the spectrum of the Laplacian matrix of a graph G in which the eignvalues are arranged in non-decreasing order: $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. The algebraic connectivity is the second smallest eigenvalue λ_2 of the Laplacian. It is the most common measure of robustness.The algebraic connectivity is equal to zero if and only if the graph is unconnected. The disadvantage of this measure is the fact that it does not necessarily capture the addition of edges to a graph that is to say the value of λ_2 does not strictly increase on edge addition.

2. Number of spanning trees

   A spanning tree is a subgraph containing $n-1$ edges and no cycles. According to the Kirchoff's Matrix-Tree Theorem, the number of unique spanning trees,ξ of graph G is equal to the value of any cofactor of the Laplacian (Harris et al., 2008). It is given by

$$\xi = \frac{1}{n} \prod_{i=2}^{n} \lambda_i. \tag{10}$$

   Baras and Hovareshti suggest the number of spanning trees as a global indicator of network robustness to edge removal. It has been proved that for p close to zero, the number of spanning tree gives similar results for robustness as the reliability polynomial (Baras and Hovareshti, 2009).

3. Effective resistance

   The effective graph resistance R, also called total effective resistance or Kirchhoff index, is defined as the sum of the effective resistances over all pairs of vertices. It can be expressed in terms of the non-zero eigenvalues of Laplacian as

$$R = \sum_{1 \leq i < j \leq n} R_{i,j} = n \sum_{i=2}^{n} \frac{1}{\lambda_i} \tag{11}$$

   Unlike the algebraic connectivity, the effective resistance involves not only one but all the non-zero eigenvalues of the Laplacian. It is for this result that any changes due to edge addition or removal are captured which makes the latter a better measure of robustness.

4. Natural Connectivity

   This spectral measure of robustness was put forward by Wu et al.(Wu et al.). It captures the core of robustness that is the capturing of redundancy of alternative paths. This is achieved by quantifying the weighted number of walks of all lengths in the graph. Closed walks are related to subgraphs of a graph for instance a closed walk of length $k = 3$ represents a triangle. The number of closed walks of all lengths is obtained following the principle used in the computing the subgraph centrality as in (Estrada, 2011) in which the shorter closed walks have more influence than their longer counterparts. The penalisation entails dividing the sum of closed walks of length k by the factorial of k. That is,

$$S = \sum_{k=0}^{\infty} \frac{n_k}{k!}, \tag{12}$$

   where n_k is the number of closed walks of length k. We also know that,

$$n_k = trace(\mathbf{A}^k) = \sum_{i=1}^{N} \lambda_i^k, \tag{13}$$

   where λ_i is the ith largest eigenvalue of $\mathbf{A}(G)$. Substituting for n_k in Eqn.12 gives

$$S = \sum_{k=0}^{\infty} \sum_{i=1}^{N} \frac{\lambda_i^k}{k!} = \sum_{i=1}^{N} \sum_{k=0}^{\infty} \frac{\lambda_i^k}{k!} = \sum_{i=1}^{N} e^{\lambda_i}. \tag{14}$$

   From Eqn. 14, we observe two facts. First, the weighted sum of closed walks can be obtained from the spectrum of the Adjacency matrix of a graph. second, the sum S will be a large number for large N and thus the need to scale S. The scaled version of S, termed as the 'average eigenvalue' and denoted by $\bar{\lambda}$ is given by

$$\bar{\lambda} = \ln\left(\frac{S}{N}\right) = \ln\left(\frac{\sum_{i=1}^{N} e^{\lambda_i}}{N}\right). \tag{15}$$

   Unlike the algebraic connectivity, the natural connectivity changes monotonically when edges are added or deleted which is one of the desired properties of a robustness measure.

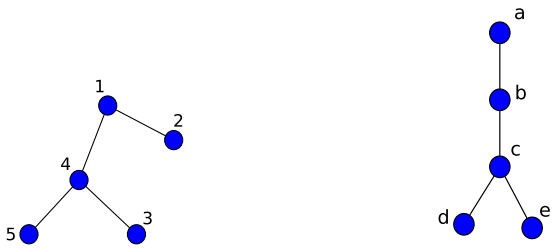
1.7 Graph Similarity

Graph similarity has a wide range of applications such as web searching, comparing chemical structures, synonym extraction, image clustering, among others (Zager and Verghese, 2008; Nikolić, 2012). The notion of graph similarity can be defined in many different ways based on the application for instance similarity of graphs can be based on whether graphs are identical copies of each other, how much the neighbourhood of a given node in one graph is similar to neighbourhood of a given node in the other graph, how much changes (node or edge deletion, redirection or addition) can be performed to one graph to obtain the other graph, among others Zager and Verghese (2008). In this work, however, we consider one of the most common notions of similarity which is isomorphism of graphs. Two graphs are isomorphic if there exists a bijective (one-to-one and onto) function between the sets of nodes such that two nodes are connected in one graph if and only if their images under the bijection are connected (Zager and Verghese, 2008). In otherwords, two graphs are isomorphic if they are structurally identical. The graphs in Fig. 2 are an example of isomorphic graphs.

1.8 Voronoi Diagrams and Delaunay Triangulation

1.8.1 Voronoi Diagrams

Voronoi diagrams (also called Voronoi tessellations, Voronoi decompositions, or Dirichlet tessellations) are important geometrical structures that are found almost everywhere in the world. They have a wide range of applications such as modeling of biological structures such as cells, study of growth patterns of



(a) (b)

Figure 2: Two isomorphic graphs

forests and forest canopies in ecology, tracing sources of infections in epidemics, in finding clear routes in autonomous robot navigations, among others. However, in later chapters, we will explore the application of Voronoi diagrams to image segmentation as explained in (Stoica, 2011).

Definition 7. *A Voronoi diagram is a special kind of decomposition of a metric space (or plane) into regions based on distances to a specified discrete set of objects in the space (usually denoted by a set of points normally referred to as seeds, sites or generators), according to the nearest-neighbor rule, such that each point is associated with the region of the plane closest to it. The regions are referred to as Voronoi cells. The Voronoi vertices are the vertices of a complex formed from a set of all Voronoi cells and their faces.In other words, a Voronoi vertex is the common boundary of 3 adjacent cells. The Voronoi edge, on the otherhand, is the common boundary of two adjacent cells.*

Let us consider a simple example of n ambulances placed at different spots of a city. The spots are considered as a subset of points denoted by $S = p_1, p_2, \cdots, p_n$. Let us assume that the distance between two points is given by the Euclidean distance function:

$$\ell_2 = d[(a_1, a_2), (b_1, b_2)] = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}. \tag{16}$$

Suppose other factors remain constant and that the nearest ambulance is contacted in cases of emergency, then the area to be served by an ambulance located at spot p_k is given by the region R_k around point p_k.

Before we go any further, we first discuss some of the common terminology used in Voronoi diagrams:

Definition 8 (Convex Set)**.** *In a Euclidean space, a convex region (or set) is a region where, for every pair of points within the region, every point on the straight line segment that joins the pair of points is also within the region.*
Definition 9 (Convex Hull)**.** *The convex hull (or convex envelope) of a set P of points in the Euclidean plane or in a Euclidean space is the smallest convex set that contains P.*
Definition 10 (Simplex)**.** *A k-simplex is a k-dimensional polytope (a geometric object with flat sides) which is the convex hull of its $k+1$ vertices.*
Definition 11 (Planar graph)**.** *A planar graph is a graph that can be drawn in the plane without any edges crossing. A planar graph drawn in this way divides the plane into regions bounded by the edges of the graph. These regions are called faces.*
Definition 12 (Dual Graph)**.** *The dual graph G^* of a plane graph G is a plane graph whose vertices correspond to the faces of G and whose edges correspond to the edges of G.*

Voronoi diagrams exhibit interesting properties which include:

- The closest pair of points corresponds to two adjacent cells in the Voronoi diagram.
- Assume the setting is the Euclidean plane and a group of different points are given. Then two points are adjacent on the convex hull if and only if their Voronoi cells share an infinitely long side.

- The Voronoi diagram on n points (or sites) is a planar graph with n faces and by Euler's formula for planar graphs, the number of Voronoi vertices and edges are atmost $2n-5$ and $3n-6$ respectively. The time complexity is $O(n \log n)$.
- Each point on an edge of the Voronoi diagram is equidistant from its two nearest neighbors p_i and p_j. Thus, there is a circle centered at such a point such that p_i and p_j lie on this circle, and no other site is interior to the circle.
- It follows that vertex at which three Voronoi cells $V(p_i)$, $V(p_j)$, and $V(p_k)$ intersect is equidistant from all sites. Thus it is the center of the circle passing through these sites, and this circle contains no other sites in its interior.
- If we assume that no four points are co-circular, then the vertices of the Voronoi diagram all have degree three.

1.8.2 Delaunay Triangulation

Let $P = p_1, p_2, \ldots, p_n \subset \mathbb{R}^2$ be a point set. A triangulation $\mathcal{T}$ of P is a maximal planar subdivision with vertex set P. Following the Empty circle property, $\mathcal{T}$ is a Delaunay triangulation of P if and only if the circumcircle of any triangle in $\mathcal{T}$ does not contain a point of P in its interior. Delaunay Triangulation is used in constructing Euclidean minimum Spanning Trees (EMST) used in solving the famous traveling salesman problem.

There are various methods of computing the Delaunay Triangulations namely: plane sweeping, iterative flipping from any other triangulation, randomized incremental construction, and conversion from Voronoi diagram. For this work, however, we will consider the Voronoi diagram based method. For a Euclidean space with point sites, the dual graph of the Voronoi diagram corresponds to the Delaunay Triangulation whose vertices are the point sites. Delaunay triangulations have interesting properties which include the following:

- Circum-circle property: The circum-circle of any triangle in the Delaunay triangulation is empty that is it contains no sites of P.
- Empty circle property: Two sites p_i and p_j are connected by an edge in the Delaunay triangulation, if and only if there is an empty circle passing through p_i and p_j.
- Closest pair property: The closest pair of sites in P are neighbors in the Delaunay triangulation.
- Given a point set P with n sites where there are h sites on the convex hull, in the plane, the Delaunay triangulation has $2n$–2–h triangles, and $3n$–3–h edges.
- The exterior face of the Delaunay triangulation is the convex hull of the point set.

1.9 The Harris Corner Point Detector

The detection of feature points in an image is vital in a number of tasks such as object tracking, 3D scene reconstruction from stereo image pairs and many other tasks in machine vision (Trajković and Hedley, 1998). The Harris corner point detector is one of the famous methods used in corner points detection. This method was introduced by Harris and Stephens (Harris and Stephens, 1988) as an improvement to the Moravec's corner detector (Moravec, 1979, 1980). In his method, Moravec considered corner points as the points where there is intensity variation in all directions. The Harris corner Detector in widely used because it is simple to compute, fast and most importantly, the corner points obtained based n this method are invariant in position to rotation, scale, illumination, and partially to affine intensity changes. On the other hand, Harris corner detector method is not invariant to image scaling, however, there are various ways of going about this problem as explained in (Stoica, 2011).

The Harris Corner Detector method determines the nature of the point by computing the average change of intensity in the image when shifting a small local window in the image by small amount in any direction. For instance for a flat region, all shifts of the window result into very small changes in intensity, for an

edge, a shift in the perpendicular results into a large change while for corner points, all shifts result into a large change in intensity. For a given image with intensities, I, a change due to a shift of a window w of size u, v by (x, y) is given by

$$E_{x,y} = \sum_{u,v} w_{u,v} |I_{x+u,y+v} - I_{u,v}|^2, \tag{17}$$

where E is the change due to the shift, x and y are the window's displacements in the x and y directions respectively, I is the intensity of image at a position (u, v), and w is the Gaussian window function $e^{-\frac{u^2+v^2}{2\sigma^2}}$.

The Taylors expansion of E gives:

$$E_{x,y} = \sum_{u,v} \{[I_x(u,v)x]^2 + [I_y(u,v)y]^2 + 2I_x(u,v)I_y(u,v)xy\}, \tag{18}$$

where $I_x = \partial I/\partial x$ and $I_y = \partial I/\partial y$. E is a close approximation of the local autocorrelation function given by

$$E(x,y) = \mathbb{M} \begin{bmatrix} x & y \end{bmatrix} \left(\sum_{u,v} w(u,v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \right) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \mathbb{M} \begin{bmatrix} x \\ y \end{bmatrix} \tag{19}$$

Matrix $\mathbb{M}$ describes the shape of the local autocorrelation function E at the origin. Let λ_1 and λ_2 be the eigenvalues of $\mathbb{M}$, according to (Harris and Stephens, 1988), the measure of corner and edge quality used for selecting core pixels known as the response function, denoted as R is given by

$$R = det(\mathbb{M}) - kTrace(\mathbb{M})^2, \tag{20}$$

where $det(\mathbb{M}) = \lambda_1\lambda_2$, $Trace(\mathbb{M}) = \lambda_1 + \lambda_2$ and k is an empirical constant such that $0.04 < k < 0.06$. A corner region with $R > 0$ is selected as a nominated corner pixel only if its response is an 8-way local maximum.

A clear step-by-step algorithm and Matlab code for the Harris corner Detector can be reviewed (Stoica, 2011).

1.10 Principal Component Analysis

Principal Component Analysis (PCA) is a very important and widely used statistical technique in various applications such as image compression, face recognition, pattern identification in high dimensional data, among others. It is such a powerful tool for data analysis. Before we dive into details of PCA, let us first define some of the statistical concepts that we will come across most often. Given a sample of size n of a population. We have the following definitions:

Definition 13 (Mean)**.** *The mean, $\bar{X}$, of a sample is the average of elements of the sample and is given by*

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{21}$$

Definition 14 (Standard Deviation)**.** *The standard deviation of a data set is a measure of how spread out the data is. It is average distance of a given point from the mean of the data set. The standard deviation, denoted by σ is given by*

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}} \tag{22}$$

Definition 15 (Variance)**.** *The variance, σ^2, is the square of the standard deviation. It is also a measure of spread of a data set. It is given by*

$$\sigma^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)} \tag{23}$$

It is important to note that both the standard deviation and variance measure the spread of data set that is 1-dimensional.

Definition 16 (Covariance)**.** *For data sets with more than* 1 *dimension, we use the covariance instead of variance (or standard deviation) measure the deviation from the mean amongst any pair of dimensions. The covariance of two dimensions x and y is given by*

$$cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)} \tag{24}$$

Definition 17 (Covariance Matrix)**.** *For a set of data with n dimensions, the covariance matrix is given by*

$$C^{n\times n} = (c_{i,j}, c_{i,j} = cov(Dim_i, Dim_j)), \tag{25}$$

where $C^{n\times n}$ is a matrix with n rows and n columns, Dim_i is the ith dimension.

For example, for a data set with 3 dimensions x,y, and z, we write the covariance matrix as

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix} \tag{26}$$

Given a set of data with n dimensions, we perform PCA on the data set following the steps below:

- Compute the mean for each dimension and then subtract the mean from the respective dimension for instance for dimensions x and y, subtract each x value, compute $x - \bar{x}$ and for each y value, compute $y - \bar{y}$. The data obtained after subtracting the mean is referred to as Normalised data.
- Calculate the covariance matrix as explained previously.
- Compute the eigenvalues and eigenvectors of the covariance matrix.
- For dimensional reduction and data compression, choose principal components which are the eigenvectors with large eigenvalues. The number of principal components corresponds to the new dimensionality.
- Feature vector formation. Having chosen which eigenvectors to keep, we construct the feature vector which is a matrix with the chosen eigenvectors as columns.
- New data set. We obtain the new data set (with reduced dimensions) by

$$NewData = FeatureVector^T \times NormalisedData^T \tag{27}$$

2 Diffusion on networks

Dynamical processes on networks aid in the modelling of processes that occur in real-world systems for example spread of diseases in a social group, transmission of signals in brain networks, spread of information in a social network, and many others. In this work, we explore the diffusion process is one of the most popular dynamic processes studied in literature.

According to (Newman, 2010), diffusion is, among others, the movement of substance from a region of high concentration to a region of low concentration. Such substance include heat, gas, and many more.

The diffusion process on networks is used in developing models that depict processes that occur in real-world systems. The modeling scheme used in this case involves: First, identifying each node of a network with a particular component or part of the system. Secondly, for each node i, a variable σ_i is introduced that characterises its dynamical state (Barrat et al., 2008). Examples of diffusion models that adopt the mentioned modeling scheme can be found in (Estrada et al., 2011; Kasprzak, 2012; López-Pintado, 2008).

In this work, we discuss the diffusion of heat on a network in which we explore the existing diffusion models, study the equilibrium behaviour, ascertain the impact of structure on rate of diffusion, and also look at the diffusion with long range interactions. In addition, we also discuss the heat kernel of a graph, its invariants such as the trace, zeta function, heat content for both the normal diffusion processes as well as diffusion with long-range interactions.

2.1 Heat Diffusion Model

Recently, various models have been developed to depict the process of heat diffusion on networks. Here, we consider a simple model on a simple graph as explained:

Let $G = (V, E)$ be a simple connected undirected graph with vertex set V and edge set E. Suppose we randomly select a few nodes (that is, sources) to which we assign specific amounts of heat as in vector ϕ_0. Let $\varepsilon \in [0, 1]$ be the heat diffusion coefficient that controls the rate of diffusion. When ε tends to 0, heat transfer among nodes becomes difficult and as a result, heat does not spread to each of the nodes with in the network. However, as ε tends to 1, heat spreads rapidly among nodes and thus, with out loss, heat is distributed to all nodes in the network.

At each time t, we obtain the quantities, $\phi_i(t)$, of heat at each node, i. The spread of heat is considered to occur along the edges connecting nodes, that is to say, through direct interactions.

The process of heat spread through out the network can therefore be modelled by

$$\frac{d\phi_i(t)}{dt} = \varepsilon \sum_j (\mathbf{A}_{ij} - \delta_{ij} k_i)\phi_j(t), \tag{28}$$

where $\mathbf{A}$ is the adjacency matrix, k_i is the degree of node i, and δ_{ij} is the Kronecker delta whose value is 1 if $i = j$ and 0 otherwise. In matrix-vector notation, we have

$$\frac{d\boldsymbol{\phi}(t)}{dt} = -\varepsilon \mathbf{L}\boldsymbol{\phi}(t), \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{29}$$

whose solution is

$$\boldsymbol{\phi}(t) = \boldsymbol{\phi}_0 \; e^{-\varepsilon t \mathbf{L}}. \tag{30}$$

Alternatively, the solution can be expressed as a linear combination of eigenvectors of the Laplacian matrix. That is

$$\boldsymbol{\phi}(t) = \sum_i \langle \boldsymbol{\phi}(0), \mathbf{v}_i \rangle \quad e^{-\varepsilon \lambda_i t} \mathbf{v}_i,$$

where λ_i, $\mathbf{v}_i$ are respectively the eigenvalues and corresponding eigenvectors of the Laplacian matrix and $\langle \boldsymbol{\phi}(0), \mathbf{v}_i \rangle$ is simply the projection of $\boldsymbol{\phi}(0)$ onto the set of eigenvectors (Anton and Rorres, 2007).

2.2 Equilibrium behaviour

We study the behaviour of the diffusion process and the quantities of heat at each of the nodes after an infinite period of time. For a simple undirected network, as t goes to infinity, we have

$$\lim_{t\to\infty} e^{-\varepsilon\lambda_i t} = \begin{cases} 0 & \text{if } \lambda_i > 0 \\ 1 & \text{if } \lambda_i = 0, \end{cases} \tag{31}$$

Asymptotically, the equilibrium state is completely determined by the kernel of $\mathbf{L}$. Since $\sum_j \mathbf{L}_{ij} = 0$, it is easy to see that $\mathbf{v^1} = \frac{1}{\sqrt{n}}[1, \cdots, 1]$, the eigenvector associated with $\lambda_i = 0$, is in the kernel of $\mathbf{L}$. We then have

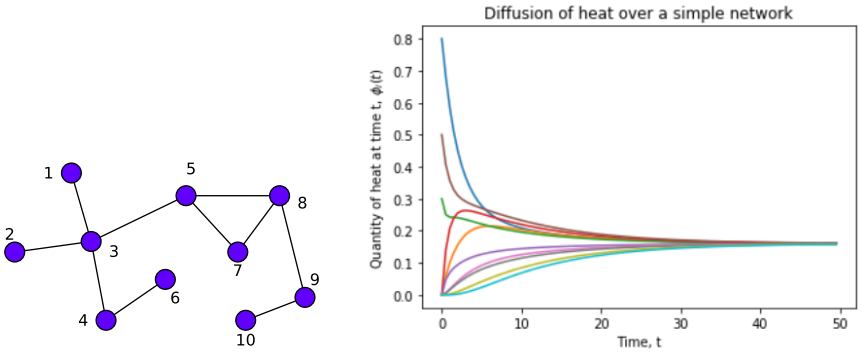
$$\lim_{t\to\infty} \boldsymbol{\phi}(t) = \langle \boldsymbol{\phi}(0), \mathbf{v^1} \rangle \mathbf{v^1}. \tag{32}$$

The quantity of heat $\phi_j(t)$ at any node j at time t is given by

$$\lim_{t\to\infty} \phi_j(t) = \frac{1}{n} \sum_{i=1}^{n} \phi_i(0). \tag{33}$$

At steady state, the quantity $\phi_i(t)$ at each of the nodes is the same, which is the average of the initial values at all of nodes. This is because, as expected, neighboring nodes in the network will exchange heat amongst each other until all nodes attain equal amounts of heat (i.e no difference in amount for any given pair of nodes).

For better understanding of the heat diffusion model, let us consider the following simple example.
Example 1. *Let us consider diffusion of heat over the network in Fig. 3(a). Suppose the quantity of heat at each node at time $t = 0$ is given by the vector $\boldsymbol{\phi}(0) = [0.3, 0.0, 0.8, 0.0, 0.5, 0.2, 0.0, 0.0, 0.0, 0.2]$, random values between 0 and 1. Let $\varepsilon = 0.05$. Fig. 3(b) illustrates how heat spreads over the network in Fig. 3(a).*



(a) (b)

Figure 3: (b) is an illustration of the diffusion process over the network in (a).

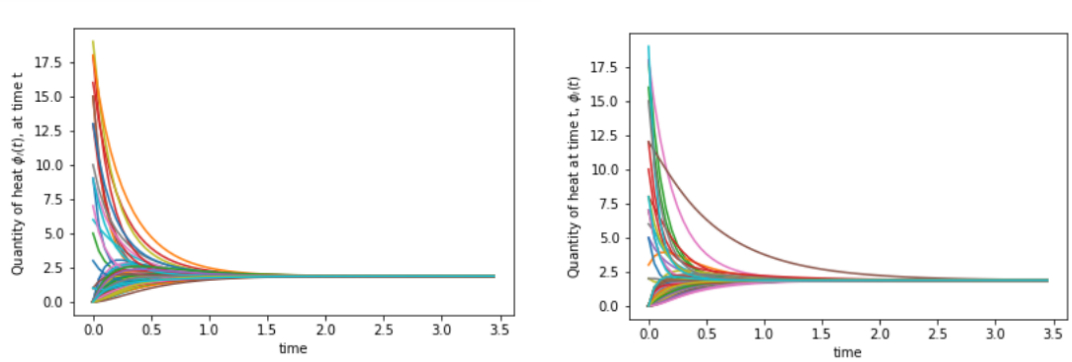
From Fig.3,we observe that at each time step t, nodes that initially have high amounts of heat (i.e $1, 3, 5, 6,$ *and* 10*) exchange heat with adjacent nodes that initially had none or little amounts of heat. The latter gain heat from the former and eventually all nodes in the network have relatively equal amounts of heat. This explains the fact that as time t increases, the quantity of heat $\boldsymbol{\phi}_j(t)$ at each node tends to the equilibrium value of* 0.2 *which is attained as t approaches* 50*.*

2.3 Impact of Network structure on the rate of diffusion

The structure of a network basically means the way in which nodes are connected in the network. For instance, in a regular network each node is connected to equal number of nodes, for a star network

one node is positioned in a way that all other nodes are connected to it. Since diffusion on networks occurs due to the interactions within neighbourhoods of nodes, it therefore implies that the topology of a network has a strong influence on the diffusion process.

We consider two networks with different structures: one is an Erdos-Renyi (ER) network that follows a Poisson degree distribution and the other Barabasi-Albert (BA) network in which connection of nodes follows scale free power-law distribution, that is to say, the probability of finding a node with degree k decreases as the negative power of k. We perform simulations of diffusion on these networks and the results are explained: Consider ER and BA networks with $n = 100$ and average degree $\bar{k} = 2.3$, we randomly select 20 nodes to which we assign random quantities (range of 0 to 20) of heat and then allow diffusion to occur. After every time step t, we compute the quantities at each node as depicted in Figure 4.



(a) (b)

Figure 4: Simulations for diffusion on networks following equation 29. To the left is the BA network and to the right is ER. Both networks have 100 nodes and average path length 2.3.

From Fig 4 we observe that diffusion occurs much faster in BA network than in ER network. In BA, equilibrium is reached after 1.5 time steps while in ER network equilibrium is reached much later after 2.5 time steps. This behaviour is attributed to the different structures of the two networks. The BA network is composed of highly connected nodes known as hubs that interact with a number of nodes at the same time thus fastening the diffusion process. On the other hand, the ER network is homogeneous which means there are no hubs and for that reason, diffusion occurs quite slower than in BA networks as evident in Fig. 4. Following these results, we can see that the structure of a network plays a key role in influencing the rate of diffusion and thus a determinant of how fast equilibrium is attained.

2.4 Influence of Heterogeneity on Diffusion over network

The heterogeneity of a network is the irregularity characterised by the existence of a nodes with degree significantly larger than the average degree of the network Estrada (2010); Albert and Barabási (2002); Newman (2003). The quantification of heterogeneity is one the areas where tremendous research has been on going and various measures have been introduced Estrada (2010). Here, we consider heterogeneity in scale free networks with $n = 1000$ and average degree $\bar{k} = 20$ by varying power exponent, γ. For different conductances x, we assign initial quantities of heat to each of the 200 nodes with highest degree. Figure 5 illustrates how the average quantities of heat of the selected initial diffusion nodes varies with time.

2.5 Impact of choice of Initial diffusion nodes on the diffusion process on networks

We consider a BA network and ER network, both of 100 nodes and average path length of 6. First case, we select 20 nodes in both networks with the highest degree centrality to which we assign specific

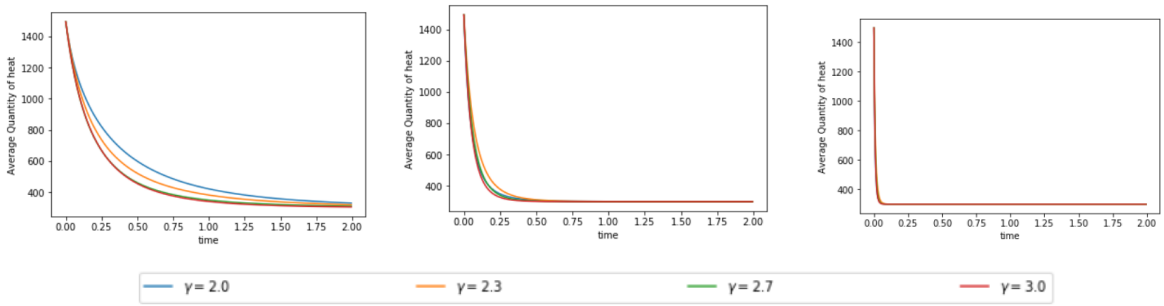


Figure 5: Plots of the average quantity of heat for 200 nodes with the highest degree centrality against time for 3 scale free networks having different values of the power exponent(2.0,2.3,2.7, and 3.0), n=1000 and average degree=6. The figures to the left, centre and right correspond to x values 0,0.1, and 0.3 respectively.

amounts of heat. Second case, we randomly select 5 nodes in both networks. At each time t, we measure the average quantity of heat at the 20 nodes and the results of the simulations are illustrated by Fig. 6.

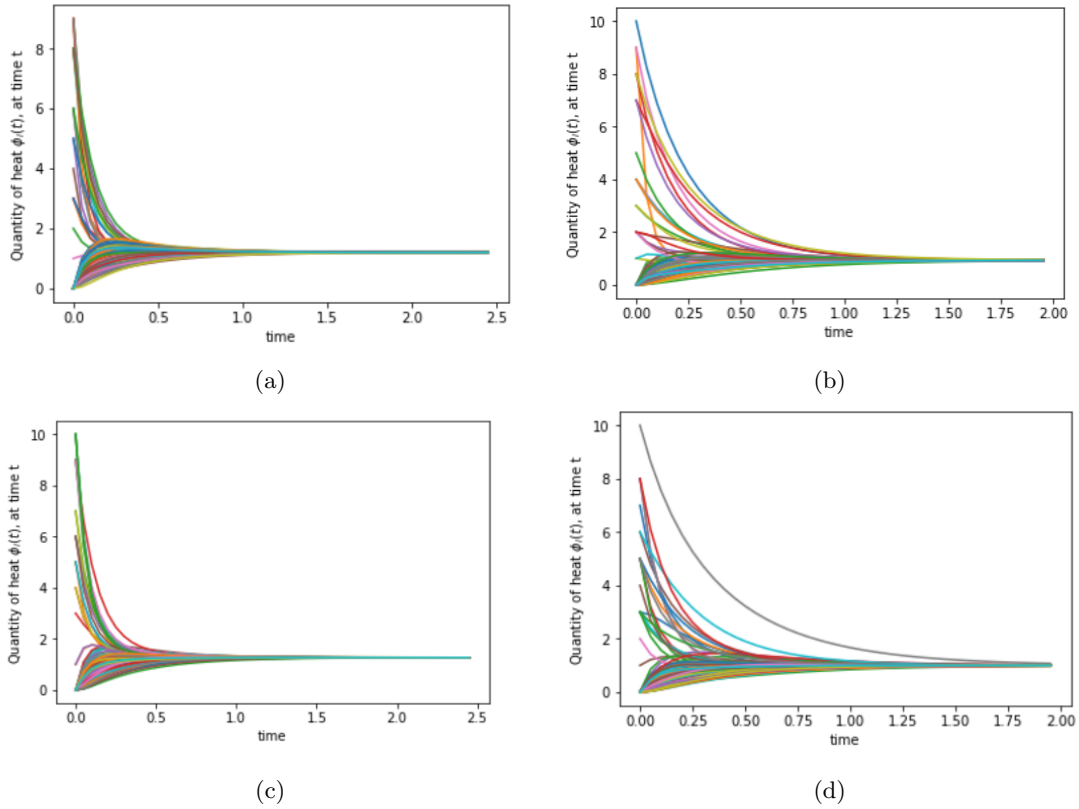


Figure 6: Results of the simulations for two networks. The top row corresponds to the BA network, (left) is illustration for which the 20 source nodes are ones with the highest degree and (right) is illustration for randomly selected source nodes. On the other hand, the bottom row is ER network for which (left) corresponds to diffusion for which the source nodes are the highest degree nodes and to figure to the right is one for which the source nodes are selected randomly.

From the simulations, we observe that when the source nodes (those from which diffusion kicks off) are chosen based on the degree, diffusion occurs much faster and equilibrium is attained quickly compared to when the source nodes are chosen randomly. We explain this observation based on the fact when the

highest degree nodes initiate the diffusion process, they quickly interact with their neigbouring nodes at the same time and since their neighbourhood is big in size, this results into fast spread of heat among the nodes in the network. On the otherhand, when source nodes are randomly selected, their is a possibility of including nodes with low degree among the selection, these low degree nodes are not fast agents of heat transfer due to their small neighbourhood which results into a relatively slow diffusion process.

2.6 Diffusion on Directed Networks

A directed graph, also known as a Digraph or directed network, is one in which all the edges are directed from one vertex to another.

There are various complex systems whose skeleton can be captured by directed networks. Examples include ecological networks, power grids, transportation networks, communication networks, metabolic networks, gene regulatory networks, citation networks among others. It is therefore paramount to study how dynamic processes such as diffusion, consensus, occur on such networks. There are various categories of directed networks which include:
Definition 18 (Weakly connected Digraph)**.** *A directed graph is called weakly connected if replacing all of its directed edges with undirected edges produces a connected (undirected) graph.*
Definition 19 (Strongly connected Digraph)**.** *A digraph is called strongly connected if and only if any two distinct nodes of the graph can be connected via a path that respects the orientation of the edges of the digraph (Saber and Murray, 2003).*

For a strongly connected digraph with atleast two distinct nodes and with no self loops, the diffusion process on this network can be modelled in a similar manner as its undirected counterpart by

$$\frac{d\boldsymbol{\phi}}{dt} = -C\mathbf{L}\boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0. \tag{34}$$

For undirected graph G, the graph Laplacian,L, is symmetric positive semi-definite. However, for directed graphs L is non-symmetric which implies that the diffusion on the former and latter graphs is not necessarily the same.
Definition 20 (Balanced Graphs)**.** *We say the node v_i of a digraph $G = (V, E)$ is balanced if and only if its in-degree and out-degree are equal, that is, $d_{out}(v_i) = d_{in}(v_i)$. A graph G is balanced if and only if all its nodes are balanced, i.e $\sum_j a_{ij} = \sum_j a_{ji}, \forall i$.*

2.7 Diffusion and Equilibrium behaviour in Directed Network

In order to understand the process of attainment of steady state in networks, we need to study the spectral properties of graph Laplacian. Let $G = (V, E)$ be a digraph with Laplacian $L(G)$ with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ in non-decreasing order.

2.7.1 Estimation of Eigenvalues of the Laplacian

Let d_{max} be the maximum node out-degree of G, then following from Gershgorin disk theorem, then all the eigenvalues of $L(G)$ are located in the following disk

$$D(G) = \{z \in \mathbb{C} : |z - d_{max}| \leq d_{max}\} \tag{35}$$

with centre at $z = d_{max} + 0j$ in the complex plane (Saber and Murray, 2003). Thus, for a strongly connected digraph G, L has a zero eigenvalue $\lambda_1 = 0$ and all the other non-trivial eigenvalues have non-negative real parts. Let us consider a strongly connected digraph $G = (V, E)$. Let $\boldsymbol{\phi}_0$ be the vector of quantities of heat at all nodes at $t = 0$, $C = 1$ be the diffusion coefficient. Similar to undirected case,the quantities of heat, $\boldsymbol{\phi}(t)$ at each node at a given time t is given by

$$\boldsymbol{\phi}(t) = \boldsymbol{\phi}_0 \, e^{-\mathbf{L}t}. \tag{36}$$

Theorem 3 (Limit Theorem for Exponential Matrices)**.** *Assume G is a strongly connected digraph with Laplacian $\mathbf{L}$ satisfying $\mathbf{Lv_r} = \mathbf{0}$, $\mathbf{v_l^T L} = \mathbf{0}$, and $\mathbf{v_l^T v_r} = 1$. Then*

$$R = \lim_{t \longrightarrow +\infty} exp(-Lt) = v_r v_l^T \in M_n, \tag{37}$$

where M_n denotes a set of square $n \times n$ matrices, v_r,and v_l^T denote the right and left eigenvalues of L associated with eigenvalue $\lambda_1 = 0$ (Saber and Murray, 2003).

From the theorem, we deduce that for a strongly connected digraph, equilibrium state can be attained and the quantity of heat at the nodes is given by

$$\lim_{t \longrightarrow \infty} = \boldsymbol{\phi}_0 \mathbf{v_r} \mathbf{v_l^T} \tag{38}$$

It is important to note that following Equation 38, any equilibrium value x^* can be attained such that $x_i^* = x_j^*$ for all i, j. This therefore motivates the search for which classes of digraphs attain equilibrium similar to that of undirected graphs where the value at each node is the average of the initial values at all nodes in the network.

Proposition 1. *Consider a directed network $G = (V, E)$ that is strongly connected. Then the digraph G globally attains average equilibrium if and only if $\mathbf{1}^T \mathbf{L} = 0$.*

2.7.2 Equilibrium state for Balanced Graphs

The proposition in (Saber and Murray, 2003) states that

Proposition 2. *Let $G = (V, E)$ be a digraph with an adjacency matrix $A = [a_{ij}]$ satisfying $a_{ii} = 0, \forall i$. Then, all the following statements are equivalent:*

i) G is balanced,

ii) $\mathbf{v_l} = \mathbf{1}$ is the left eigenvector of the Laplacian of G associated with the zero eigenvalue, that is, $\mathbf{1}^T \mathbf{L} = 0$.

iii) $\sum_{i=1}^{n} u_i = 0, \forall x \in \mathbb{R}^n$ with $u_i = \sum_{j \in N_i} a_{i,j}(x_j - x_i)$.

Since for a balanced digraph $\mathbf{v_l}$ is an all ones vector, it therefore follows from Proposition 2 that at equilibrium, the value at all nodes in a balanced graph is the average of the initial values at all nodes.

Example 2. *Let us consider two directed graphs, one is a balanced digraph and the other is not. We then assign initial quantities of heat to all nodes in the order 0 to 4 as in the vector $\phi_0 = [2, 0, 3, 0, 0]$ and set the diffusion coefficient, $C = 1$. We then obtain plots for diffusion on both graphs after a specific time t as shown in*

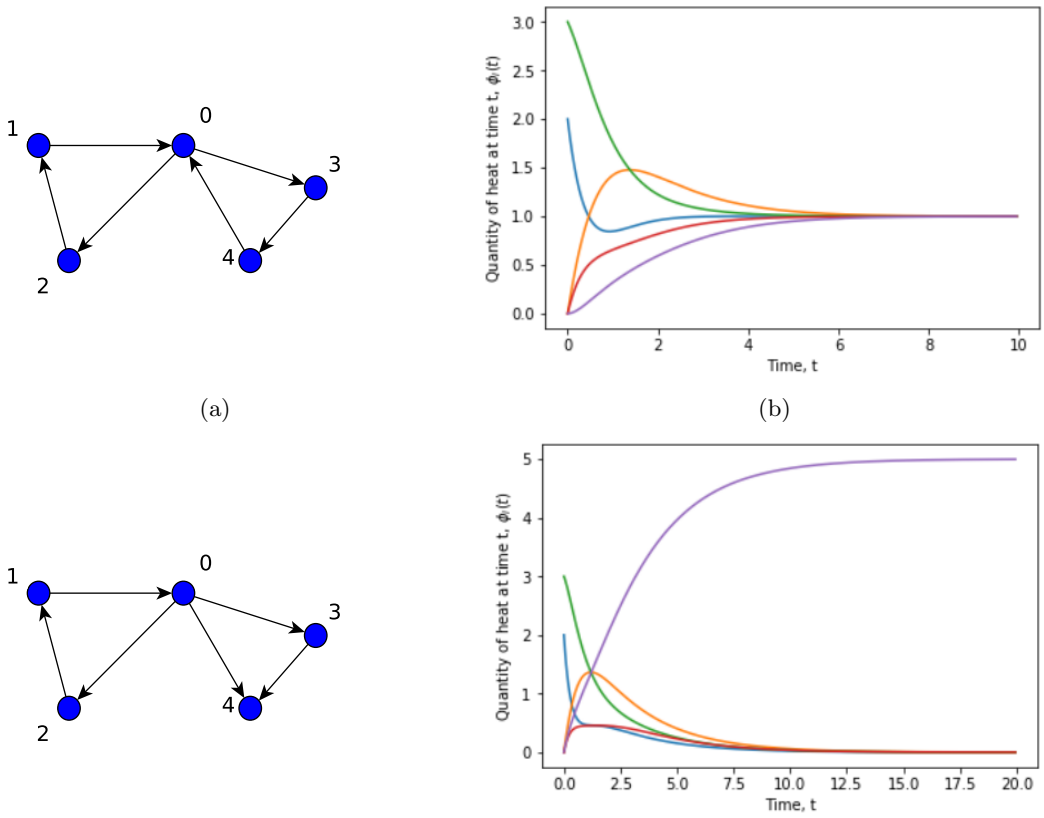


Figure 7: Diffusion over different categories of directed networks. (c) is an illustration of diffusion over weakly connected and unbalanced digraph in (a). (c) is an illustration of diffusion over strongly connected and balanced digraph (c).

For the balanced graph in Fig. 7a, its Laplacian matrix $\mathbf{L}$,v_r *and* v_l *are respectively:*

$$\mathbf{L} = \begin{pmatrix} 2 & 0 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad v_r = v_l = \begin{pmatrix} 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \end{pmatrix}$$

The values for v_r *and* v_l *satisfy Theorem 3 as well as Proposition 2 and thus, equilibrium is attained at* $x^* = 1.0$ *which is the average of initial values* x_0.
On the other hand, for the unbalanced graph in Fig. 7c, we have the following matrices

$$\mathbf{L} = \begin{pmatrix} 3 & 0 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad v_r = \begin{pmatrix} 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \\ 0.4472136 \end{pmatrix}, \text{ and } v_l = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.0 \\ 0.0 \\ 1.0 \end{pmatrix}$$

We observe that $Lv_r = 0$ *and* $v_l^T L = 0$. *However, the condition* $v_l^T v_r = 1$ *is not satisfied and therefore equilibrium cannot be attained as shown in Fig. 7. In addition, we observe that considering the structure, vertex* 4 *has only in coming edges which signifies that during the diffusion process, vertex* 4 *only receives heat from the immediate neighbours vertices* 0 *and* 3 *without giving out any due to lack of out going links. As a result, quantity of heat at vertex* 4 *keeps on increasing as shown in Fig. 7.*

2.8 Diffusion on network with long-range interactions

The 'classical' case considers diffusion over a network where a substance under consideration, say heat, flows along the edges of the network. However, long range interactions during diffusive processes on networks are evident in real world situations. Such interactions result into superdiffussion on networks which has been modelled by various models that include the random walks with Levy flights (RWLF), model based on fractional diffusion equation (FDE), and many more. In this work, we discuss an elegant model introduced by Estrada (Estrada et al., 2017b) which accounts for longrange interactions by use of k-path Laplacian matrices resulting into a generalised diffusion process on networks. To start with, let us understand what the k-path Laplacian matrices are.

2.8.1 k-path Laplacian matrices, $\mathbf{L}_k$

The k-path laplacian matrices are a natural generalisation of the combinatorial laplacian of a graph (Estrada, 2012). The motivation behind this generalisation is the idea of determining whether every node in a graph can be visited by means of a process that involves hopping from one node to another separated at a distance k. We can better understand the concept of k-path Laplacian through considering a polarisation process on a network, that is to say as, suppose a particle with a positive charge resides at a given node of simple graph $G = (V, E)$ and while at that node, it polarises all nodes at a distance d from it. Consequently, the particle's movement to another is such that it hops to any nearest non-positively charged node. While at the new node, the particle polarises neighbouring nodes in the same manner as before. As a result, the particle either hops to the nearest non-positive nodes or returns to the origin node as illustrated in Fig. 8 and Fig.9 for $d = 1$ and $d = 2$ respectively.

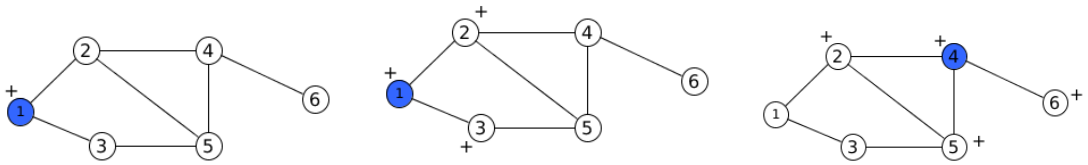


Figure 8: Illustration of how the polarisation analogy used as a motivation for the k-path Laplacian concept for networks. Starting with a positively charged particle at node 1 as shown in (**??**), taking $d = 1$, the particle polarises all its nearest neighbours at a distance d from it (that is nodes 2 and 3) as depicted in (b). The particle can therefore jump to the non-polarised nearest neighbours namely nodes 4 and 5 and 6(though node 6 is further compared to other two alternatives). Suppose the particle jumps to node 4, similar polarisation process as the particle polarises the new nearest neighbours. The particle either jumps to node 3 or returns to node 1 as shown in (c).

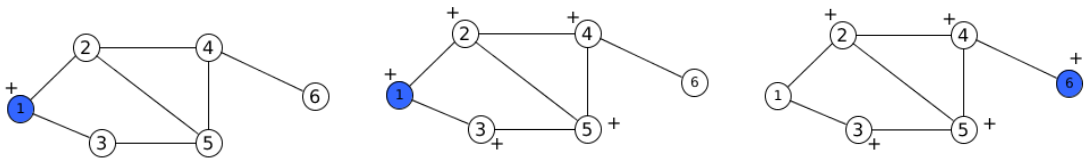


Figure 9: Illustration of how the a charged particle navigates the network taking jumps of length $d = 2$. As discussed before, a particle starting off at node 1 will polarise neighbouring nodes separated at not more than distance 2 from it (that is nodes 2,3, 5, and 4) as shown in (b). The particle then has only an option of jumping to the non-polarised node 6 after which a similar process occurs again as in (c).

As for the 'classical' case in which traversing the graph involves subsequent hops of length 1 at a time, terminolgy such as walk, path, and many more are defined. In the same way, for the generalised case in which hops of various length not exceeding the diameter of a graph are taken into account, we need to define terminology as well:

Definition 21 (k-hopping walk)**.** *A k-hopping walk of length l is any sequence of (not necessarily different) nodes $v_1, v_2, \cdots, v_l, v_{l+1}$ such that $d_{i,i+1} = k$ for each $i = 1, 2, \cdots, l$. In otherwords, this walk is referred to as a k-hopping walk from v_1 to v_{l+1} (Estrada, 2012).*

Definition 22 (k-path degree)**.** *The k-path degree $\delta_k(v_i)(k \leq d_{max})$ of a node v_i is the number of irreducible shortest-paths of length k having $v_i$as an endpoint (Estrada, 2012).*

Definition 23 (k-path Laplacian matrix)**.** *The k-path Laplacian matrix $L_k(k \leq d_{max})$ of a connected undirected graph $G = (V, E)$ is defined as the square symmetric $n \times n$ matrix whose entries are given by:*

$$L_k(ij) = \begin{cases} -1 & \text{if } d_{i,j} = k, \\ \delta_k(i) & \text{if } i = j, \\ 0 & \text{otherwise}, \end{cases} \tag{39}$$

where $d_{i,j}$ is the shortest path distance between nodes i and j, $delta_k(i)$ known as the k-path degree is the number of irreducible shortest paths of length k having node i as an endpoint.

Example 3. *Let us compute the k-path Laplacians for a simple graph G in Fig.10.*

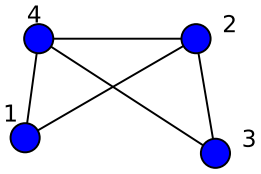


Figure 10: A Graph of size 4.

$$\mathbf{L_1(G)} = \begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}, \quad \mathbf{L_2(G)} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{L_3(G)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The concept of k-path Laplacians defined in Eqn 23 for finite undirected graphs was extended for connected and locally finite infinite graphs as follows: Consider $\Gamma = (V, E)$ to be an indirected finite or infinite graph with vertices V and edges E. We assume that Γ is connected and locally finite that is to say each vertex has only finitely many edges emanating from it. Let d be the distance metric on Γ, i.e. $d(v, w)$ is the length of the shortest path from v to w, and let $\delta_k(v)$ be the k-path degree of the vertex v, i.e.

$$\delta_k(v) := \#\{w \in V : d(v, w) = k\}. \tag{40}$$

Since γ is locally finite, $\delta_k(v)$ is finite for every $v \in V$. Denote by $C(V)$ the set of all complex-valued functions on V and by $C_0(V)$ the set of the complex-valued functions on V with finite support. Moreover, let $\ell^2(V)$ be the Hilbert space of square-summable functions on V with inner product

$$\langle f, g \rangle = \sum_{v \in V} f(v)\overline{g(v)}, \quad f, g \in \ell^2(V) \tag{41}$$

In $\ell^2(V)$ there is a standard orthonormal basis consisting of the vectors $e_v, v \in V$, where

$$e_v(w) := \begin{cases} 1 & \text{if w = v,} \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

Let $\mathbf{L}_k$ be the following mapping from $C(V)$ into itself:

$$(\mathbf{L}_k)(f) := \sum_{w \in V : d(v,w)=k} (f(v) - f(w)), \quad f \in C(V). \tag{43}$$

Definition 24 (k-hopping connected component)**.** *A k-hopping connected component in a graph $G = (V, E)$ is a subgraph $G' = (V', E')$, $V' \subset V, E' \subset E$, such that there is at least one k-hopping walk that visit every node $v_i \in V'$. As mentioned earlier, the motivation of the generalisation of graph Laplacian to find the solution of the problem of whether a given graph can be k-hopped. If not, how many k-hopping connected components exist?*

Theorem 4. *The number of k-hopping connected components in a connected undirected graph $G = (V, E)$ is given by $\eta_k(G) = m[\lambda_1(L_k) = 0]$ (Estrada, 2012).*

Proof. Let us consider a simple undirected graph G that is connected, in otherwords a graph that can be 1-hopped. Let $v_1, v_2, \cdots, v_n$ be an orthogonal basis of $\mathbb{R}^n$ such that $\mathbf{L}\mathbf{y}_j = \lambda_j(\mathbf{L})\mathbf{y}_j$. Using Rayleigh-Ritz Principle we obtain

$$\lambda_1(\mathbf{L}) = \min_{\mathbf{y} \in \mathbb{R}^n \ \{\mathbf{0}\}} \frac{\mathbf{y}^T \mathbf{L} \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \min_{\mathbf{y} \in \mathbb{R}^n \ \{\mathbf{0}\}} \frac{\sum_{p,q \in P(v_p, v_q)} (y_p - y_q)^2}{\sum_{p \in V} y_p^2} \tag{44}$$

Let $\mathbf{y}$ be an eigenvector of $\mathbf{L}$ with eigenvalue $\lambda_i(\mathbf{L}_k) = 0$. This implies that

$$\mathbf{y}^T \mathbf{L} \mathbf{y} = \frac{1}{2} \sum_{p,q \in P(v_p, v_q)} (y_p - y_q)^2, \tag{45}$$

Equation 45 holds if $y_p = y_q$ which happens if and only if the two vertices v_p and v_q are connected which implies that for all vertices connected by a path in the network, $\mathbf{y}$ should be a constant. which happens if and only if, $y_p = y_q$ for each pair of nodes which are connected by an edge. However, on assumption that G is connected, it implies that each pair of vertices of the connected component can be connected by a path and thus $y_p = y_q$ for all pair of vertices of the connected component. Consequently, for a one connected component, the vector $\mathbf{v} = \mathbf{0}$ is the only vector with corresponding to eigenvalue of 0. In otherwise, $\mathbf{v}$ is a constant vector spanning a dimensional space.

Let us now consider the case of n connected components. Suppose graph G is made of vertices arranged, without loss of generality, in way that its Laplacian matrix is organised as

$$\mathbf{L} = \begin{pmatrix} L_1 & 0 & \cdots & 0 \\ 0 & L_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_n \end{pmatrix}.$$

The spectrum of the block diagonal matrix $\mathbf{L}$ is given as the union of the spectra of blocks $\mathbf{L_i}$. The corresponding eigenvectors are given as those of $\mathbf{L_i}$ with zeroes at the positions of other blocks. Since $\mathbf{L_i}$ is a graph laplacian for a connected component and we know that each $\mathbf{L_i}$ has 0 as an eigenvalue with multiplicity 1, it therefore implies that the multiplicity of 0 as an eigenvalue of L corresponds to the number of connected components in the graph G (**?**).

Let us now consider a graph that can be k-hopped that is to say graph with only 1 k-connected component for $k > 1$. The graph Laplacian $\mathbf{L_k}$ is defined as in Equation 23. Following similar argument as the case of 1-hopped graph discussed before, we have

$$\mathbf{y}^T \mathbf{L_k} \mathbf{y} = \sum_{p,q \in P_k(v_p, v_q)} (y_p - y_q)^2 = 0, \tag{46}$$

which happens if and only if, $y_p = y_q$ for each pair of nodes which are connected by the shortest-path of length k. Now, let us assume that the graph is k-hopping connected. Then, because every pair of nodes is connected by paths of length ck, we have that $y_p = y_q \neq 0$ for every pair of nodes in the graph. Consequently, $\mathbf{y}$ is a constant vector spanning a one-dimensional space.

Now let $j > 1$ and let us assume that the graph has g k-hopping connected components $\mathbf{L_k^1}, \mathbf{L_k^2}, \cdots, \mathbf{L_k^g}$. Then, the k-Laplacian matrix can be written as:

$$\mathbf{L}_k = \begin{pmatrix} L_k^1 & 0 & \cdots & 0 \\ 0 & L_k^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_k^g \end{pmatrix}.$$

Following similar arguments as for the case of the k-hopping connected graph it can be seen that $\mathbf{L}_k$ has g orthogonal eigenvectors $\mathbf{y}_j$ of eigenvalue 0, such that $y_p = y_q \neq 0$ for each pair of nodes which are in the same k-connected of the graph and zero otherwise (Estrada, 2012).

□

Example 4. *Let us consider the graph, G in Fig. 10, since $d_{max} = 2$ we compute the 1-hopping and 2-hopping connected components of G.*

		no. of components	components
$\lambda_i(\mathbf{L}_1)$	**0.000** 2.000 4.000 4.000	1	1- 2- 3- 4
$\lambda_i(\mathbf{L}_2)$	**0.000** **0.000** **0.000** 2.000	3	1-3, 2, 4

2.8.2 The Generalised Laplacian Matrix

The generalised Laplacian matrix is obtained as a linear combination of the k-path Laplacian matrices and it is given by

$$L_G = \sum_{k=1}^{\Delta} c_k L_k \tag{47}$$

where $1 \leq \Delta \leq d_{max}$ and c_k are the coefficients (Estrada, 2012). The coefficients c_k play a crucial role in the generalisation of diffusion process on network and so determining the values of these coefficients is an important task. The values of c_k are expected to give more weight to shorter than to the longer range interactions. In (Estrada, 2012) Estrada proposed two approaches categorised as social and physical ways of influence.

2.8.3 Properties of the Generalised Laplacian Matrix

The generalised matrix L_G is real and symmetric which follows from the fact that L_G is a linear combination of real and symmetric k-path matrices. The generalised matrix is also a positive semi-definite matrix.

Proof. For any column vector $\mathbf{y}$

$$\mathbf{y}^T L_G \mathbf{y} = \mathbf{y}^T (c_1 L_1 + c_2 L_2 + \cdots + c_\delta L_\delta)\mathbf{y} = \mathbf{y}^T c_1 L_1 \mathbf{y} + \mathbf{y}^T c_2 L_2 \mathbf{y} + \cdots + \mathbf{y}^T c_\delta L_\delta \mathbf{y} \tag{48}$$

Since $\mathbf{y}^T c_k L_k \mathbf{y} \geq 0$ for $c_k > 0$ and $1 \leq k \leq \delta$ as in Eqn.45, then

$$\mathbf{y}^T L_G \mathbf{y} \geq 0 \tag{49}$$

□

2.8.4 Choice of Co-efficients, c_k

1) **Social Influence**

Here, we consider a social network where nodes represent the people with in a society and the links are the social relationship or ties among the people for instance friendship, family relations, collaboration, among others. In such networks, influence between two people connected to each other in the network can be accounted for. However, it is quite challenging to account for the indirect influence between two people that are not directly connected in the network. An approach introduced in (Estrada et al., 2011)considers that which on empirical evidence, the indirect influence or long range interactions among people can be thought of as a pre-conditioner for establishment of a new social tie. In otherwords, new social ties among humans are created as an investment in the future as justified by empirical evidence in (Estrada et al., 2011). it is quite obvious that two individuals that influence each other, probability is high that the two become friends compared to those that have no mutual influence. This process can be considered as an analogy in which the future value of money, in particular the future value of a growing annuity, is determined in quantitative finance but for this case we consider a transaction involving information instead of money. Suppose an individual A lends information to individual B whose present value is PVI at an interest rate r and for a time period t. The future value of the information FVI is given by

$$FVI = PVI(1+r)^t \tag{50}$$

Let us consider the process on a network where node v_1 lends information to node $v_(l+1)$. On assumption that information flows through the shortest path and considering a discrete time at every step, the information is transferred from v_1 to nearest neighbour v_2 at a value A and rate r. At v_2, the present value $PVI = A/(1+r)$. The information is enriched at v_2 at a growing rate of g and then transferred to v_3. The process is repeated as before and at each node information is enriched before transfer to the next node. Finally, the information at borrower node $v_(l+1)$ is $A(1+g)^{l-1}/(1+r)^l$. Thus, The cumulative present value of the information in this process is given by the sum of all the values at the nodes of the chain, that is,

$$PVI = A/(1+r) + A(1+g)/(1+r)^2 + \cdots + A(1+g)^(l-1)/(1+r)^l. \tag{51}$$

Suppose $g = r$ and $A = 1$ for the sake of simplicity, for a connected network with shortest distance between any pair of nodes denoted by $d_{i,j}$, the future value of information transmitted from i to j is:

$$FVI_{i,j} = d_{i,j}x^{di,j-1}, \tag{52}$$

where $x = 1 + r = 1 + g$. Thus, from the analogy, we can consider that the mutual influence between two nodes separated at distance k is given by the future value of the investment that the creation of a new link will represent to them.

For the social influence, we can define the coefficients in Eqn. 57 as $c_1 = 1$ and $c_{k\geq 2} = kx^{k-1}$, where $0 < x < 1/2$. The empirical parameter x also known as the conductance in (Estrada et al., 2011) controls the strength of interaction between nodes i and j separated at distance k. This implies that the strength of the casual contact between two nodes reduces with increase in social distance between them.

The generalised Laplacian matrix for which long-range interactions are accounted for by the social influence is given as

$$L_{G,x} = \begin{cases} \delta_{Gv_i} & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ -kx^{k-1} & \text{otherwise,} \end{cases} \tag{53}$$

where δ_{Gv_i} denotes the generalised degree. In modelling the spread of epidermic, Estrada in (Estrada et al., 2011) considered two types of contacts that is close contacts which are frequent interactions among individuals and casual or long range interactions are the non frequent encounters among individuals which facilitate the spread of infections.The latter which were considered as non-random where accounted by use of the social influence approach.

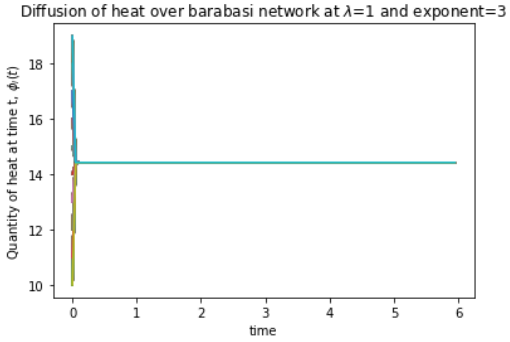
2) **Physical Influence**

It is observed in many man-made and naturally evolving systems that communication among the agents of the system follows a spatial decay as illustrated in sensor systems where sensors far away from the target display low noise-signal ratio as a result of attenuation (spatial decay) of signal energy, in earthquake incidences where the aftershocks follow a spatial decay, that is, areas further from the main shock are less affected compared to nearer areas. This spatial decay takes on the form $r^{-\alpha}$, where r is the distance from the main shock. Other examples of similar physical scenarios include the brain in which the interconnectivity certain neurons in mammalian neo- cortex decays exponentially with the intersomatic distance, and many others. According to Estrada (Estrada, 2012), the spatial decay can take on two forms namely:

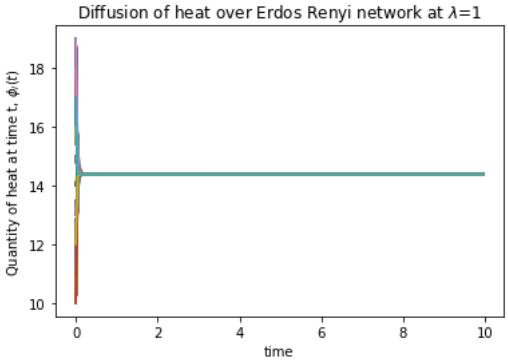
i) Exponential form For the exponential form, Eqn.57 attains a generalised form based on the Laplace-transformed k-Laplacian:

$$\tilde{\mathbf{L}}_{L,\lambda} = \mathbf{L} + \sum_{k=2}^{\infty} e^{-\lambda k} \mathbf{L}_k, \tag{54}$$

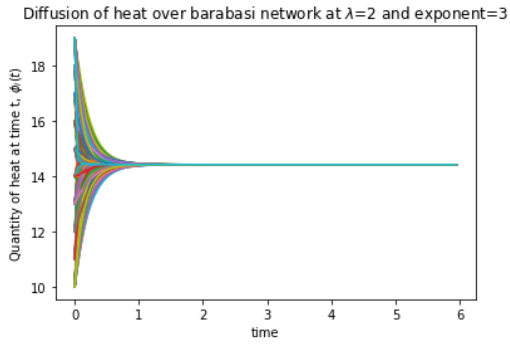
where $\lambda > 0$ is a parameter that depends on the specific situation to be modelled. Thus, the coefficients of Eqn.57 are $c_1 = 1$ and $c_{k\geq 2} = e^{-\lambda k}$.



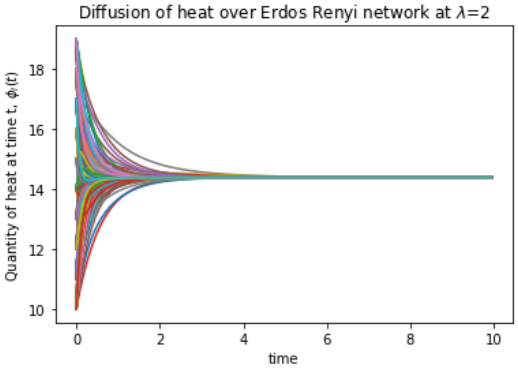
(a) $\lambda = 1$



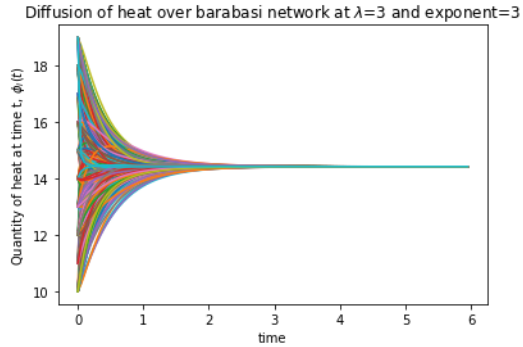
(b) $\lambda = 1$



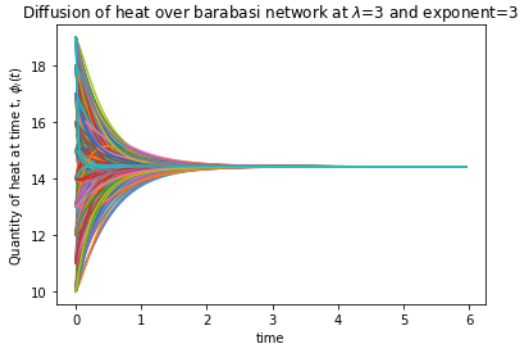
(c) $\lambda = 2$



(d) $\lambda = 2$



(e) $\lambda = 3$



(f) $\lambda = 3$

ii) **Power-law**

Here, we use the Mellin transform of k-Laplacian matrices to obtain the generalised equation:

$$\tilde{\mathbf{L}}_{M,s} = \sum_{k=1}^{\infty} k^{-s} \mathbf{L}_k, \tag{55}$$

where $\lambda > 0$ is a parameter that depends on the specific situation to be modelled and the coeffients of Eqn.57 are $c_k = k^{-s}$. In (Estrada et al., 2017b), it is shown that normal diffusion occurs only when $s > 3$. On the other hand, superdiffusion occurs when $1 < s < 3$ with superdiffusive exponent being $\kappa = \frac{2}{s-1}$, which leads to arbitrary values for $\kappa \in (1, \infty)$. The Mean Square Distance, $MSD \sim t^{\kappa}$. Normal diffusion is attained at $\kappa = 1$, superdiffusion is attained at $\kappa > 1$. However, a special type of diffusion known as Ballistic diffusion is characterised by the fact that

at small times the particles are not hindered yet by collisions and diffuse very fast. It is attained at $\kappa = 2$.

2.8.5 Generalised Diffusion Model

In this case, we consider diffusion on a graph where by both direct and long-range interactions are involved. One interesting study of long-range interactions is by Estrada on modelling epidermic spread in networks (Estrada et al., 2011). Here, the long range interactions are considered to be nonrandom and depend on the social distances between individuals in the social network. Recently, work in (Estrada, 2012) explains a method of generalisation of the diffusion process on a given graph based on the k-path Laplacian operators $\mathbf{L}_k$ in which we consider the fact that the diffusive particle at a given node can hop to not only its nearest neighbours (captured by the classical Laplacian operator $\mathbf{L}$) but to any other nodes in the graph with a probability that decays with the increase of the shortest path distance between the current node where the particle is residing to the one it will hop to. Thus, for a diffusive node hopping to other nodes separated at a distance k form its current location, such a diffusion process can be captured by replacing $\mathbf{L}$ in (29) by $\mathbf{L}_G$ in (43). That is

$$\frac{d\boldsymbol{\phi}}{dt} = -\varepsilon \mathbf{L}_G \boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{56}$$

where L_G is the generalised Laplacian matrix.

We then consider the generalised diffusion on graph where interactions are both short-range and long-range. The long-range interactions are accounted for by use of k-path Laplacian matrices. Following this generalisation, Eqn.56 then becomes

$$\frac{d\boldsymbol{\phi}}{dt} = -\varepsilon\Big(\sum_{k=1}^{\Delta} c_k \mathbf{L}_k\Big)\boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{57}$$

The Long Range Interaction Model (LRI-model) discussed in (Estrada et al., 2017b) considers a simple graph $G = (V, E)$ with diameter denoted by d_{max}. Its k-path Laplacian matrices can be transformed in various ways. However, in this work we consider two types of transforms as covered in (Estrada et al., 2017b) namely the Mellin and Laplace transforms. The Generalised Laplacian can be given by

$$\mathbf{L}_G = \tilde{L}_\tau = \begin{cases} \sum_{d=1}^{d_{max}} d^{-s} L_d & \text{for } \tau = Mell, s > 0 \\ L + \sum_{d=2}^{d_{max}} e^{-\lambda d} L_d & \text{for } \tau = Lapl, \lambda > 0. \end{cases} \tag{58}$$

Estrada, et.al (Estrada et al., 2017a) gave an insightful interpretation of the LRI-model in the following way. Let us consider a simple connected graph $G = (V, E)$ and the transformation: $f : G(V, E) \longrightarrow G' = (V, E', \phi, W)$, such that $E' = E \bigcup (p, q) \mid p, q \in V, (p, q) \notin E$ and $\phi : W \longrightarrow E'$ is a surjective mapping that assigns a weight to each of the elements of E'. The weights $w_{ij} \in W$ are given by the Mellin or Laplace transforms and they are specific for each graph, that is $w_{ij} = d_{ij}^{-s}$ or $w_{ij} = e^{-\lambda d_{ij}}$ for $d_{ij} > 1$, and $w_{ij} = 1$ for connected pairs of nodes. We this interpretation, it is important to note that when $s, \lambda \longrightarrow \infty$ the weighted graph $G' = (V, E', \phi, W)$ tends to the original graph $G(V, E)$. On the other hand, when $s, \lambda \longrightarrow 0$ the weighted graph $G' = (V, E', \phi, W)$ tends to the complete graph with N nodes ($N = |V|$), K_N.

2.9 The Heat Kernel

The heat kernel is very useful in a number of ways for instance in identifying communities in graph, in partitioning of graphs, as a pagerank of a graph, as a means of embedding a graph into a pattern space, among others (Chung, 2007, 2009; Kloster and Gleich, 2014). As discussed earlier, diffusion of heat on a graph can be modelled by the equation

$$\frac{d\phi}{dt} = -\mathbf{L}\phi, \tag{59}$$

where $\mathbf{L}$ is either the Laplacian matrix or it's normalised version.

The heat kernel is the fundamental solution to the diffusion equation (59). It is obtained by exponentiating the Laplacian eigensystem over time and it is given by

$$H_t = e^{-t\mathbf{L}} \tag{60}$$

It literally describes the flow of substance (heat) across edges (direct interactions) in the graph (Xiao et al., 2009). On applying spectral decomposition to Equation 60, we have

$$H_t = \mathbf{V}e^{(-t\mathbf{\Lambda})}\mathbf{V}^T = \sum_{i=0}^{n} e^{(-\lambda_i t)} v_i v_i^T \tag{61}$$

where λ_is are the eigenvalues of $\mathbf{L}$ in a non-decreasing order $0 = \lambda_1 \leq \lambda_2 \leq, \cdots \leq \lambda_n$ and v_i is the eigenvector corresponding to the eigenvalue λ_i (Anton and Rorres, 2007).

For a graph $G = (V, E)$, the heat kernel matrix of the graph is an $|V| \times |V|$ matrix whose entry for a pair of node u, v is given by

$$H_t(u, v) = \sum_{i=1}^{|V|} e^{-\lambda t} \phi_i(u)\phi_i(v) \tag{62}$$

When t tends to zero, the kernel behaviour can be obtained from the Taylor's expansion of Equation 60 which is

$$e^{-\mathbf{L}t} = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!}\mathbf{L}^k = \mathbf{I} - t\mathbf{L} + \frac{t^2\mathbf{L}^2}{2!} + \frac{t^3\mathbf{L}^3}{3!} + \cdots \tag{63}$$

Thus,

$$\lim_{t \longrightarrow 0} \left(e^{-t\mathbf{L}}\right) = \mathbf{I} - t\mathbf{L}. \tag{64}$$

It therefore implies that for t tending to zero, the heat kernel depends on the local connectivity structure of the graph. On the other hand, as t tends to infinity, following from Equation 61

$$\lim_{t \longrightarrow 0} \left(e^{-t\mathbf{L}}\right) = e^{(-\lambda_2 t)} v_2 v_2^T. \tag{65}$$

From Equation 65,its evident that for large t, the heat kernel behaviour is determined by the global structure of the graph.

2.9.1 The Heat Kernel Invariants

In this subsection, we discuss the invariants associated with heat kernel. These include the trace of the heat kernel, zeta function, derivative of zeta function at the origin and the heat content.

1 **Heat Kernel Trace** The trace of the heat kernel is the sum of the entries at the main diagonal of the heat kernel matrix. The trace of the heat kernel $Tr(H_t)$ is given by

$$Tr(H_t) = Tr(\mathbf{V}e^{-t\mathbf{\Lambda}}\mathbf{V}^T) = Tr(e^{-t\mathbf{\Lambda}}(\mathbf{V}^T\mathbf{V})) = Tr(e^{-t\mathbf{\Lambda}}) \tag{66}$$

Thus, the trace function, $Z(t)$, of the heat kernel is given by

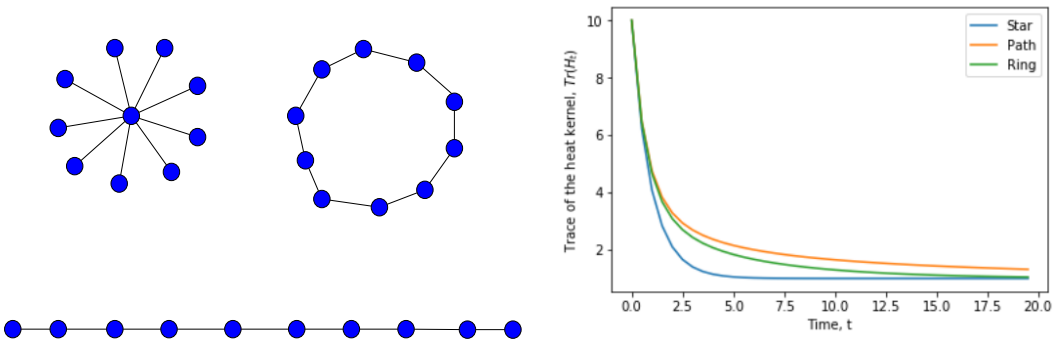
$$Z(t) = Tr(H_t) = \sum_{i=1}^{|V|} e^{-\lambda_i t}, \tag{67}$$

where λ_i is the eigenvalue of the normalised Laplacian matrix Xiao et al. (2009). From Equation 67, it is evident that the trace of the heat kernel is invariant to permutations. For a connected graph, Equation 67 can be written as

$$Z(t) = 1 + e^{-\lambda_2 t} + e^{-\lambda_3 t} + \cdots + e^{-\lambda_N t} \tag{68}$$

2.9.2 Heat kernel trace as a Graph Analysis Technique

According to Xiao Xiao et al. (2009), the trace of the heat kernel has a potential applicability to distinguishing graphs with different topologies based on the shape of the curves obtained by plots of the trace of the heat kernel as a function of time. Let us consider 3 simple graphs namely a star, path and 2-regular graph of size 10. Fig 12 shows the plot of heat kernel trace against time for the three graphs.



(a) (b)

Figure 12: (a) are the three graphs used for heat kernel analysis. (b) plot of the heat kernel trace against time for star (blue), path (orange) and regular(green) graphs.

From (b), we observe that since the 3 graphs have different topologies, the corresponding curves take on different shapes as well , that is to say, the curves are distinct. It is evident that since path and 2-regular graphs have almost similar topologies, the two curve corresponding the graphs are close to each other unlike for the star graph whose curve is quite separate and has a different (deeper trough). On comparing the plots for $s = 2$ and $s = 3$, we observe drastic shift in the curves for the former than the latter in comparison with the plot for the normal graph Laplacian in Fig.(b).

It is important to note that the use of the trace formula for characterisation of graphs is limited due to the fact that for each value of time, t, only a single scalar attribute is provided. This implies that the heat kernel trace function must be either sampled with time or a fixed time value must be selected.

2 Zeta function

The Zeta function associated with the Laplacian eigenvalues is obtained by exponentiating and summing the reciprocal of the non-zero Laplacian eigenvalues. It is thus defined by

$$\zeta(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s}. \tag{69}$$

Xiao Xiao et al. (2009) established a relationship between the zeta function and the heat kernel trace by use of the Mellin transform. For a function $f(t)$, it's Mellin transform is given by

$$F(s) = \int_0^\infty t^{s-1} f(t) dt. \tag{70}$$

Taking a function $f(t) = e^{-\lambda_i t}$, applying the Mellin transform gives

$$\lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} e^{-\lambda_i t} dt, \tag{71}$$

where $\Gamma(s)$ is the gamma function defined as

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt \tag{72}$$

. On summation for all non-zero eigenvalues of the Laplacian, Equation 71 becomes

$$\zeta(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \sum_{\lambda_i \neq 0} e^{-\lambda_i t} dt \tag{73}$$

The heat kernel trace can be also expressed as

$$Tr(h_t) = C + \sum_{\lambda_i \neq 0} e^{-\lambda_i t}, \tag{74}$$

where C is the multiplicity of zero eigenvalues of the Laplacian that is the number of connected components of a graph. Substituting Equation 74 into Equation 73 gives

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \{Tr(h_t) - C\} \, dt. \tag{75}$$

Thus the zeta function is related to the moments of the heat kernel trace. It is the moment generating function and thus a way of characterising the shape of the heat kernel trace.

3 **Derivative of Zeta Function at the Origin**

The derivative or slope of the zeta function at the origin is another characterisation of the heat kernel trace second to the zeta function which measures it's shape. Writing the Zeta function in terms of natural exponential we have

$$\zeta(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s} = \sum_{\lambda_i \neq 0} e^{-s \ln \lambda_i}. \tag{76}$$

Thus, the derivative is given by

$$\zeta'(s) = \sum_{\lambda_i \neq 0} \{-\ln \lambda_i\} e^{-s \ln \lambda_i} \tag{77}$$

so, the derivative at the origin is

$$\zeta'(s) = -\sum_{\lambda_i \neq 0} \ln \lambda_i \tag{78}$$

4 **Heat Content**

The heat content is defined as the sum of the entries of the heat kernel matrix of a graph. Its given by

$$Q(t) = \sum_{u \in V} \sum_{vi \in V} h_t(u, v) \tag{79}$$

Substituting for $h_t(u, v)$ gives

$$Q(t) = \sum_{p \in V} \sum_{qi \in V} \sum_{k=1}^{|V|} e^{(-\lambda_k t)} v_k(p) v_k(q), \tag{80}$$

which can be expanded into a polynomial in time as in (McDonald and Meyers, 2002)

$$Q(t) = \sum_{m=0}^{\infty} q_m t^m \tag{81}$$

2.10 Generalised Heat kernel

In the previous section, we have discussed the heat kernel for diffusion process that occurs through direct interactions between nearest neighbours in the network. In this section however, we consider the fact that in many observed real-world process, it is observed that interactions do not only occur among nearest neighbours but also among non-nearest which we term as the long range interactions. The heat equation

in this case is based the the concept of k-path Laplacian introduced by Estrada (Estrada, 2012) and is given by

$$\frac{d\boldsymbol{\phi}}{dt} = -\varepsilon L_G \boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{82}$$

where

$$L_G = \Big(\sum_{k=1}^{\Delta} c_k \mathbf{L}_k\Big) = c_1 L_1 + c_2 L_2 + \cdots + c_\Delta L_\Delta, \tag{83}$$

where $1 \leq \Delta \leq d_{max}$ and c_k are the coefficients. As discussed earlier, the coefficients c_k are chosen in a way that as distance k increases, the long range effect is weakened. Some of the common expressions for coefficients c_k are $c_1 = 1, c_{k\geq 2} = e^{-sk}$, $c_{k\geq 2} = k^{-s}$ which depict physical influence while $c_k = kx^{k-1}$ for the social influence.

The heat kernel H_{Gt} of the generalised diffusion process which we refer to as the generalised heat kernel, the fundamental solution to eqn.82 is given by

$$H_{Gt} = e^{-L_G t}, \tag{84}$$

which on expansion can be written as

$$H_{Gt} = e^{-t(c_1 L_1 + c_2 L_2 + \cdots + c_\Delta L_\Delta)}, \tag{85}$$

where $L_1, L_2, \cdots, L\Delta$ are the matrices, known as k-path Laplacian matrices, which are obtained by considering hops of length $k = 1, 2, \cdots, \Delta$ respectively. At $k = 1$, we recover the normal heat kernel in eqn.60.

2.11 Trace of the Generalised Heat Kernel

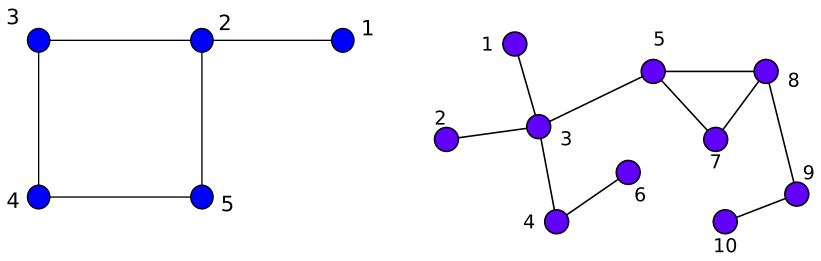
The trace of the generalised heat kernel is therefore given by

$$Z_G(t) = Tr(H_{Gt}) = \sum_{i=1}^{|V|} e^{-\lambda_{iG} t}. \tag{86}$$

Alternatively, Eqn. 56 can be written as

$$Z_G(t) = 1 + e^{-\lambda_{2G} t} + e^{-\lambda_{3G}} + \cdots + e^{-\lambda_{NG} t} \tag{87}$$

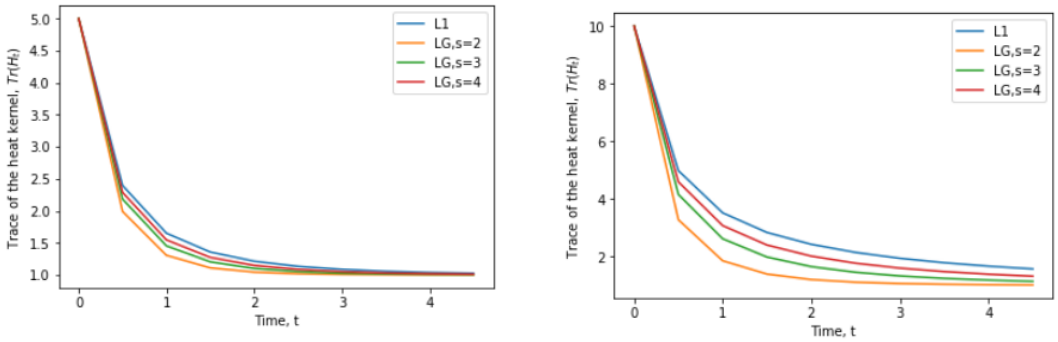
Let us consider two toy examples to illustrate the variation of the trace generalised heat kernel with time.



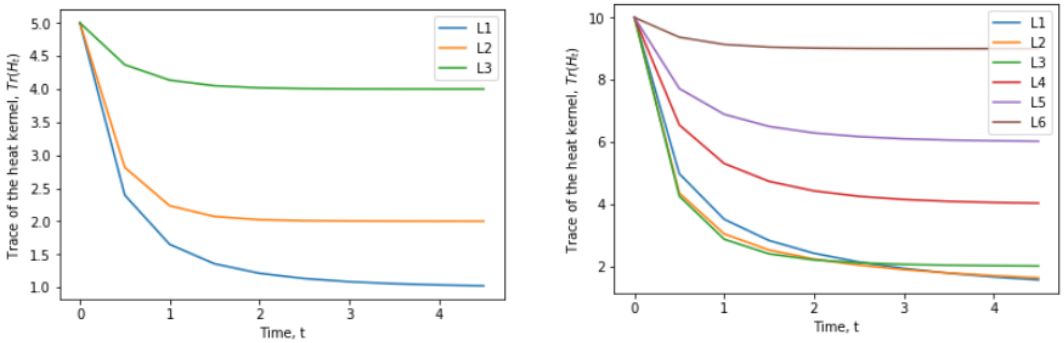
(a) (b)

Figure 13: (a) is a simple network of size 5 and (b) is a simple network of size 10

.



(a) (b)



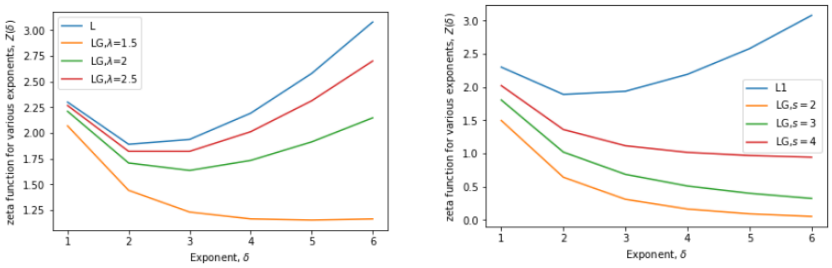
(a) (b)

Figure 15: Figures (a) and (b) are plots of trace function of the generalised Laplacian matrix against time for graphs (13a) and (13b) respectively. On the other hand, figures in the lower row that is (a) and (b) are plots for the trace function for k-path Laplacian matrices $(L_1, L_2, \cdots, L_\Delta)$ for graphs (13a) and (13b) respectively

The Mellin transforms of the k-path Laplacian matrices is one of ways used to account for long range interactions in networks. For the above simulations, we consider the exponents ranging from 2 to 4. In both cases, at $t = 0$, the value of the trace function is equal to the size of the eigenvalues as each exponential term, $e^{\lambda_{iG}t} = 1$. We also observe that as the exponent s increases, the corresponding curve approaches that of the combinatorial Laplacian. This is based on the fact that the strength of the longrange interactions are controlled by the exponent, s in such a way that the larger the value of s, the weaker the interactions. Interestingly, for $s = 2$ in both figures (a) and (b), the trace of the heat kernel has a deeper trough compared to others. This can be explained by the superdiffusion that is always manifested at $s = 2$ as discussed in the previous sections.

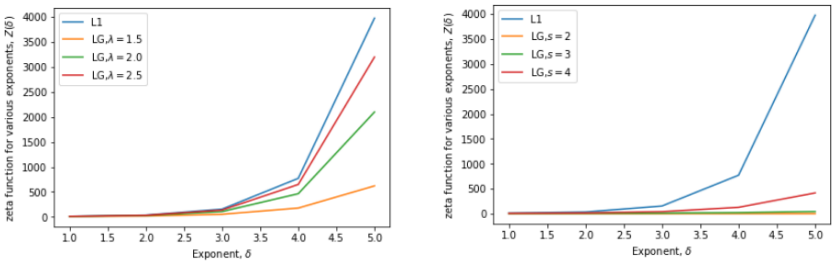
Let us then focus on the plots for the trace function for the individual k-path Laplacians against time. Figure (a) corresponds to the simple graph in (13a). We observe that there is a clear distinction for the trace functions of each of the 3-path Laplacian matrices namely L_1, L_2, and L_3. For $k = 1$ which is the combinatorial Laplacian, the curve has a deeper trough followed by $k = 2$ and lastly $k = 3$. This distinction in curves is possibly based on the change in eigenvalues for each of the k-path Laplacian following a decreasing of this format: $\lambda_i(L_1) \geq \lambda_i(L_2) \geq \lambda_i(L_3)$. On the other hand, however, Figure (b) shows a slightly different trend from what we observe in Figure (a) that is there is no clear distinction among the curves corresponding to different k-path Laplacian matrices. This is attributed to the fact that for some graphs, the eigenvalues for the k-path Laplacians do not necessarily follow a decreasing order as explained earlier.

2.12 Zeta function of the Generalised Zeta Function



(a) (b)

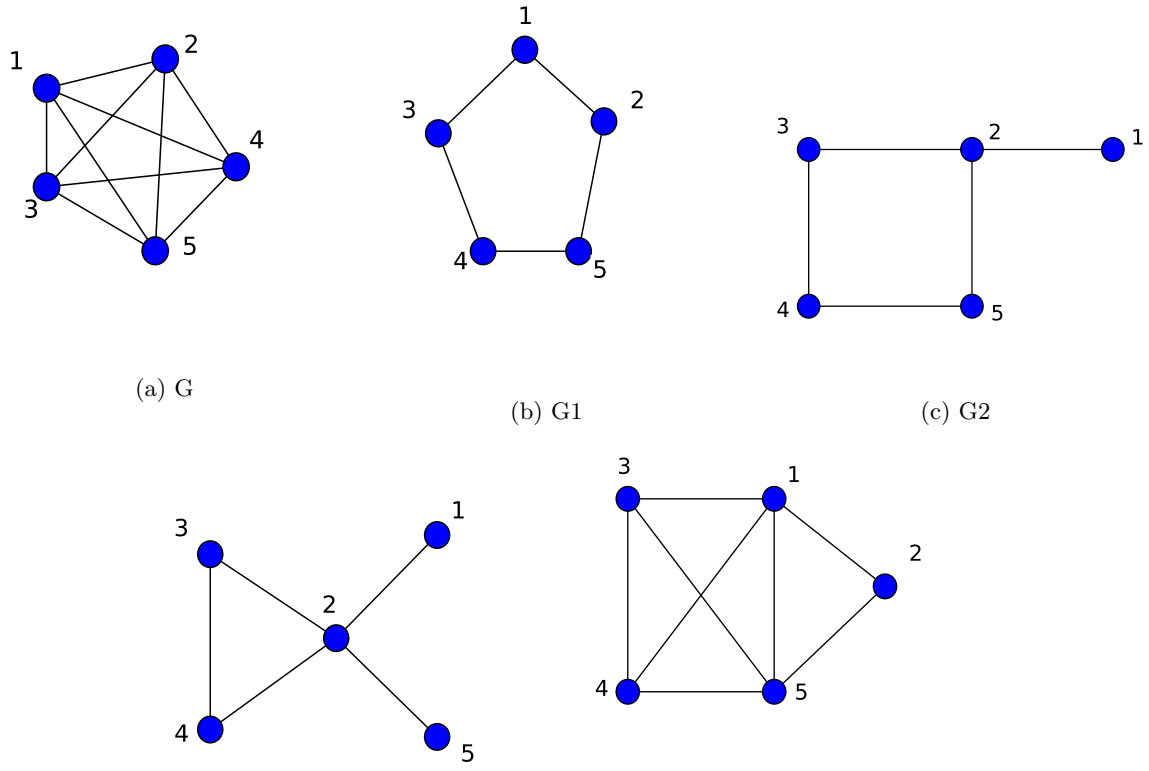
Figure 16: Illustration of the Zeta function of the graph in Fig. 13a against exponent δ. (a) corresponds to the Laplace transform of the graph Laplacian with $\lambda = 1.5, 2$, and 2.5. (b) corresponds to the Mellin transform of the graph Laplacian with $s = 2, 3$, and 4.



(a) (b)

Figure 17: Illustration of the Zeta function of the graph in Fig. 13b against exponent δ. (a) corresponds to the Laplace transform of the graph Laplacian with $\lambda = 1.5, 2$, and 2.5. (b) corresponds to the Mellin transform of the graph Laplacian with $s = 2, 3$, and 4.

2.13 Comparison with Complete Graph based on Trace of the diffusion kernel



(a) G

(b) G1

(c) G2

(d) G3

(e) G4

Figure 18: The five graphs of size 5. (c) is the complete graph,K_5 whose trace function is to be compared with that of the other 4 graphs.

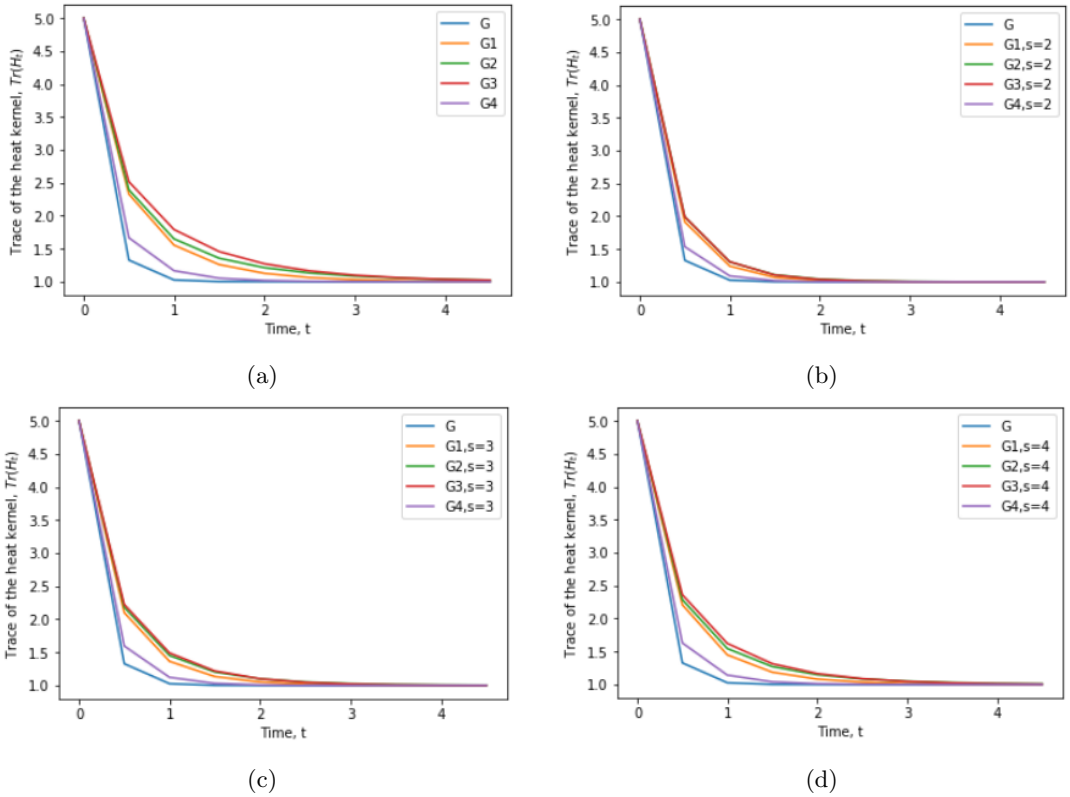


Figure 19c is a plot of the trace function of the graph Laplacian against time for a complete graph,G (blue) and other graphs $G1$(orange), $G2$(green), $G3$(red), and $G4$(purple). To the right, is a plot of the trace function of the Mellin transformed Laplacian matrix at $s = 3$ against time. We observe that the two plots follow a similar shape, however, plot in Figure 19d, the curves for graphs $G1$,$G2$, $G3$, and $G4$ get closer to the curves corresponding to the complete graph$G$. This can loosely be explained by the longrange interactions involved that result into a flow of information closer to that of a complete graph.

2.13.1 Simulations of diffusion on lattice

We consider a 2-dimensional discrete grid in which each point is connected to 8 of its nearest neighbours. Initially, we assign heat quantities to all the points on the grid and then we investigate how the diffusion process occurs and at each time t.

Let us take a 20 by 20 grid on which we assigned heat quantities of amounts 5, 7 and 10 to a few points and the rest are assigned zero. The diffusion on the lattice through direct interactions only as well as through both direct and indirect interactions by using the Social distance concept, Laplace and Mellin tranforms.

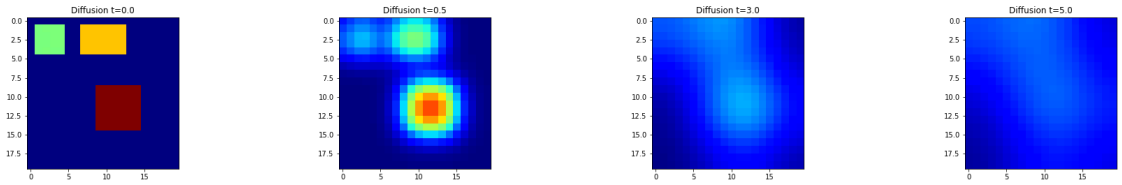
i) Diffusion on lattice through Direct interactions only

ii) Diffusion on lattice through both direct and indirect interactions. Considering a similar lattice with initial heat quantity assignments as in the directed case discussed before, we account for the diffusion process where interactions among nodes occurs through both direct and indirect interactions.

As discussed previously, long-range interactions can be accounted for by using various techniques as illustrated by the following illustrations.

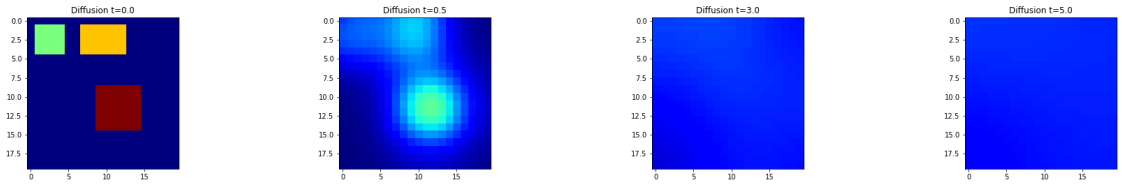
a) Long range interactions using social distance Taking $x = 0.1$ and $x = 0.2$.

At $t = 0$, diffusion on the grid starts off with 3 strong regions having high quantities of heat as observed from figures 20, and 22 which correspond to diffusion with direct interactions only

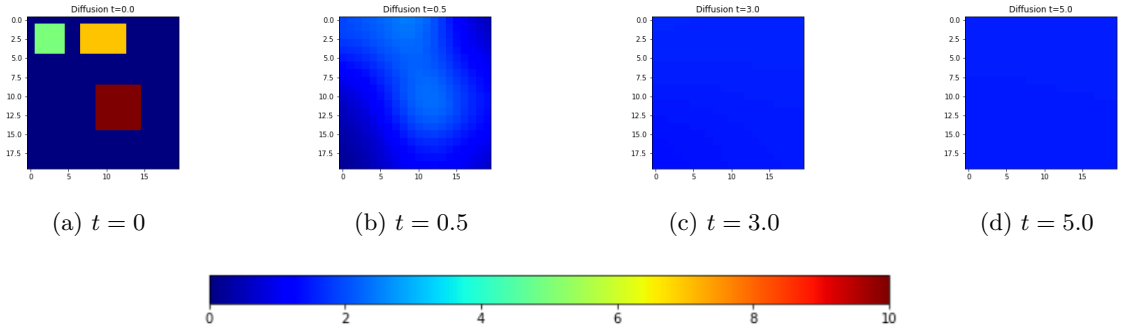


(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

Figure 20: Sample illustrations for progression of diffusion over a 20×20 lattice where specific regions are assigned particular heat quantities which in turn spread to other regions of low heat quantities through interactions along links in the network.



(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$



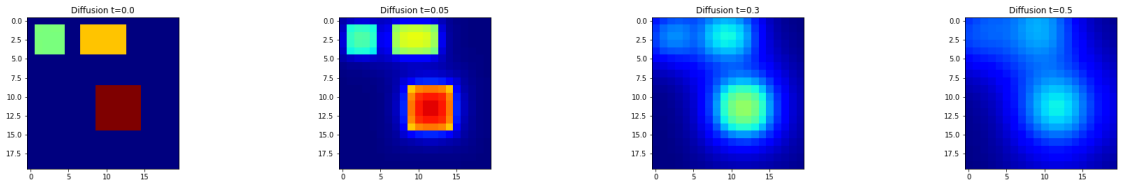
(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

Figure 22: Illustrations for diffusion over a grid with long-range interactions accounted for by the social distance technique. The upper row corresponds to diffusion with conductance $x = 0.1$ while lower row corresponds to results obtained $x = 0.2$. The intensity of heat follows a color grid where by red implies higher intensity followed by yellow and blue implies low heat intensities.

and diffusion with conductance $x = 0.1$ and 0.2 respectively. As the diffusion process continues, we see that at $t = 0.5$, strong heat points can still be spotted for direct interactions, relatively strong points in $x = 0.1$ and almost complete diffusion in $x = 0.2$. We can also observe that by $t = 2.0$, heat is uniformly distributed across the grid for $x = 0.1$ and $x = 0.2$. However, for the case of direct interactions (Fig.20d) diffusion is still ongoing and we can notice strong heat points at the centre of the grid. Following the sequences in the figures, we can conclude that as x (i.e increase in intensity of long range interactions), the diffusion process goes faster and equilibrium across the grid is reached faster as observed in the above simulations where for $x = 0.2$, $x = 0.1$ equilibrium is reached by $t = 3$ while for $x = 0$, equilibrium is not yet reached by then.

b) Longrange interactions using Laplace Transforms

c) Longrange interactions using Mellin Transforms of the k-Laplacian matrices. For the case of Mellin transforms, we observe that rate of diffusion is quite faster at $k = 2$(upper row) than at $k = 4$(lower row).This is due to the fact that at $k < 3$, super diffusion occurs, however normal diffusion occurs when $k > 3$ as shown in lower row of Fig.25.



(a) $t = 0$ (b) $t = 0.5$ (c) $t = 3.0$ (d) $t = 5.0$

Figure 23: Illustration for diffusion on the lattice with both direct and long range interactions. These longrange interactions are accounted for by Laplace transforms given by Equation 54 with $\lambda = 1$.

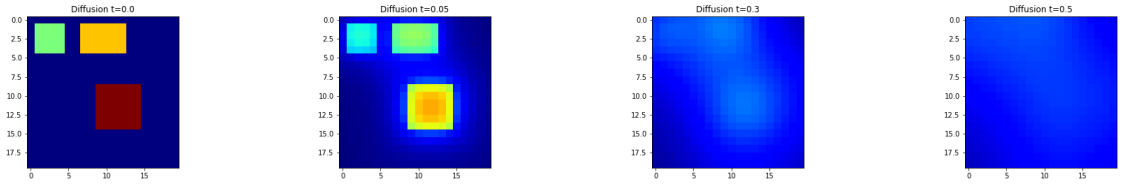


Figure 24

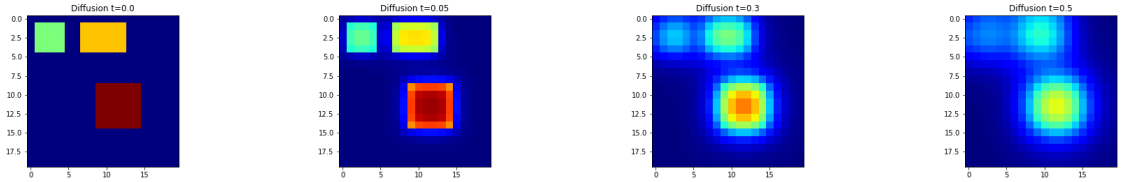


Figure 25: Illustrations of diffusion process on the lattice allowing for both direct and longrange interations. The latter are accounted for using Mellin transforms of the k-Laplacian matrices as in Equation. 55. The top row corresponds to the case for which $s = 2$ and the bottom row corresponds to case $s = 4$.

3 Image segmentation

Image segmentation is defined as the problem of localising regions of an image relative to content such as image homogeneity. There are various algorithms that perform image segmentation based on specific content localisation such as intensity, color, texture, etc. It is important to note a good image segmentation algorithm must perform fast computation and editing, produce arbitrary segmentation with enough interaction as well as producing intuitive images.

Graph theory concepts are often applied in developing a number of image segmentation algorithms such as graph cuts algorithms, the intelligent scissor algorithm, the random walker algorithm, among others.

3.1 Image segmentation based on Random Walks

A random walk (also known as drunkard's walk) is a stochastic process that describes a path that consists of a succession of random steps on some mathematical space. An example of a random walk on a graph can be described as in the following way: Suppose a walker starts off a given node i and then moves to one of the neighbouring nodes of i, say j, with a uniform probability. Once at j, the walker again moves to the next node following a similar process until the stopping node. the sequence of these randomly visited nodes is referred to as the random walk.

There are various types of random walks, however, for the work we consider lattice random walkers that is to say random walkers that occur on on the lattice.

Here, we study the application of random walker on 2-d lattice to the segmentation of images.

3.2 Random Walks and Electric potential

Let's consider a graph $G = (V, E)$ with two nodes, a and b, selected as traps with a score of 0 and 1 for hitting nodes a and b respectively. So for any other node x, the probability $p(x)$ of reaching node b given that the walker started at node x is given by

$$p(x) = \frac{1}{d_x} \sum_{y, y \in N(x)} p(y), \tag{88}$$

where $N(x)$ denotes the neighbours of x.

$$p(a) = 0 \tag{89}$$

$$p(b) = 1 \tag{90}$$

We then consider a graph $G = (V, E)$ as an electrical network which consists of connected wires. The vertices represent the intersections of the wires in the circuit such that v_x is the voltage at a given vertex $v \in V$. On the other hand, the edges represent the resistors along the wires such that for each edge $x, y \in E$ is a resistor of resistance $r_{x,y}$. Suppose we set the resistances $r_{x,y}$ for any pairs of nodes in graph to 1 Ohm and set the voltages of vertices a and b to 0 and 1 respectively that is

$$v(a) = 0 \tag{91}$$

$$v(b) = 1 \tag{92}$$

By Ohm's law the current flowing through vertices x and y is given by

$$i_{x,y} = v(x) - v(y) \quad \forall x \neq y \neq b, \quad x, y \in E \tag{93}$$

Applying Kirchoff's law, we have

$$\sum_{x \ y} i_{x,y} = 0, \quad \sum_{y,x \ y} v(x) - v(y) = 0, \tag{94}$$

which gives

$$v(x) = \frac{1}{d_x} \sum_{y, y \in N(x)} v(y) \tag{95}$$

We observe that system of equation of the probabilities in (88,89,90) and that of the voltages (91,92,95) follow the same law. That is to say they are both harmonic functions and satisfy the following properties

1. Mean value property Let us consider simpler example of one dimension case. Let A be the set of points $A = \{0, 1, 2, \cdots, N\}$ and let B be the set of boundary points $B = \{0, N\}$ and C be the set of interior points $C = \{1, 2, \cdots, N-1\}$. So the function $f(x)$ defined on A is harmonic if it satisfies the mean value property given as

$$f(x) = \frac{f(x+1) + f(x-1)}{2} \tag{96}$$

   From (88) and (95), we observe both functions p and v satisfy the mean value property in (96) since for any node other than a or b, the value of the function at that node is the average of the values of its neighbours.

   In addition, both functions have the same values at the boundary that is at nodes a and b.

2. Maximum principle A harmonic function $f(x)$ defined on A is said to satisfy the maximum principle if minimum and maximum values are attained only at the boundary. For functions v and p, we can observe from (89,90) and (91,92) respectively that the maximum principle holds.

3. Uniqueness Principle

   If $f(x)$ and $g(x)$ are harmonic functions on A such that $f(x) = g(x)$ on B, then $f(x) = g(x)$ for all x. Following from previous discussions, $v(x)$ and $p(x)$ follow the uniqueness principle which implies that the two functions are equal. This therefore implies that the probability of a random walker reaching a given labelled node y given that the walker started at node x is equal to the electric potential developed at node x when node y is set to a potential of 1 volt. It is this similarity between random walks and electric potentials on graphs that was applied in the development of image segmentation using random walker algorithm as we will explore in the next sections.

3.3 The Random walker Algorithm

3.3.1 Problem Formulation

First, we represent the image as a weighted undirected graph (discrete object) where each pixel is represented as a node and each node connected by weighted edges to either 4 or 8 of its nearest neighbours. The real-valued edge weights can be captured by various weighting functions. However, in this work we use a Gaussian weighting function which maps changes in pixel intensities to the edge weights as given by

$$w_{i,j} = exp(-\beta(g_i - g_j)^2), \tag{97}$$

where β is a free parameter, g_i is the intensity at pixel i. To capture colour,filter coefficients, texture or any other desired image features, we can modify (97). For instance, to capture colour, we replace $(g_i - g_j)^2$ with $\|g_i - g_j\|^2$ $\forall e_{i,j} \in E$ (Grady, 2006).

Having represented the image as a graph and given user-defined seeds that represent regions of belonging to the desired objects, we obtain an interactive image segmentation by assigning each unseeded node or pixel to the seed to which a random walker starting at the unseeded node first reaches a particular seed. The random walker is biased by the edge weights to avoid crossing sharp intensity gradients for respect of boundary objects.

Suppose we have an image that we would like to segment into K regions or objects that is to say a k-way image segmentation. First, seeds are selected by user to represent the desired k regions. The main task involves obtaining a k-tuples of probabilities with which a random walker starting at each unseeded node reaches the seeds. Each unseeded node is then assigned a label for the seed for which the largest probability was obtained. One approach to this problem would be simulation of the random walker process though unfortunately, the method would be infeasible for certain image segmentation problems of interest. Alternatively, work in (Grady, 2006) indicates that the probability with which a random walker first reaches a given seed can be found as a solution to the Combinatorial Dirichlet problem with boundary conditions at the location of the seed points with the seed in question set to 1 as the rest of the seeds are set to 0. With this approach, we can then analytically compute the desired probabilities as we will discuss in the next subsections.

3.3.2 Combinatorial Dirichlet Problem

Dirichlet problem is the problem of finding a harmonic function subject to its boundary conditions. A harmonic function is a function that satisfies the Laplace equation

$$\nabla^2 u = 0, \tag{98}$$

for a field u. The harmonic function that satisfies the boundary conditions minimises the Dirichlet integral

$$D[u] = \frac{1}{2}\int_{\Omega} |\nabla|^2 d\Omega, \tag{99}$$

for region Ω.
Let us consider an image represented as a graph G=(V,E) where nodes represent pixels while edges

represent the neighbourhood among the pixels. The Combinatorial Laplacian matrix of the graph is defined as

$$L_{ij} = \begin{cases} d_i & \text{if } i = j, \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise,} \end{cases} \tag{100}$$

where $L_{i,j}$ is indexed by vertices v_i and v_j. The combinatorial formulation of the Dirichlet integral is

$$D[x] = \frac{1}{2} x^T L x = \sum_{e_{ij} \in E} w_{ij}(x_i - x_j)^2. \tag{101}$$

A combinatorial harmonic is a function x that minimises 101.

3.3.3 Implementation of Algorithm

1. Computing probabilities:

   With Graph $G = (V, E)$ representing the image to be segmented, partition the vertex set into two such that is V_m and V_u which are the seeded/marked and unseeded vertices respectively such that $V_m \bigcup V_u = V$ and $V_m \bigcap V_u = \emptyset$. Without loss of generality, we assume that the nodes in L are arranged such that the seeded nodes are first followed by the unseeded ones. Then equation 101 can be written as

   $$D[x_u] = \frac{1}{2} \begin{bmatrix} x_m^T x_m^T \end{bmatrix} \begin{bmatrix} L_m & B \\ B^T & L_u \end{bmatrix} \begin{bmatrix} x_m \\ x_u \end{bmatrix} = \frac{1}{2}(x_m^T L_m x_m + x_u^T B^T x_m + x_u^T L_u x_u), \tag{102}$$

   where x_m and x_u correspond to the potentials at the seeded and unseeded nodes respectively. Differentiating $D[x]$ with respect to x_u and then finding the critical points gives the following system of linear equations with $|V_u|$ unknowns

   $$L_u x_u = -B^T x_m. \tag{103}$$

   Let x_i^s be the probability of a random walker starting at node v_i first reaches label s. Then for each label s the probabilities with each unlabelled nodes first reach s is given by

   $$L_u x^s = -B^T m^s, \tag{104}$$

   and for all labels, we have

   $$L_u X = -B^T M, \tag{105}$$

   where X has K columns taken by each x^s and M has columns given by each m^s.

2. Assigning labels to unseeded nodes Having computed the probabilities for each unseeded nodes, we assign a label to each node for which the highest probability was obtained.

3.4 Possible application of Long range interactions to Image Segmentation

We considered a possible application of long range interaction to segmentation of image based on random walks. This was implemented by replacing the Combinatorial Laplacian matrix L in equation 105 with the the k-path Laplacian matrices to account for long range jumps of the random walker. However, the output obtained using this method was not a better image segmentation due to the fact that the long-range jumps do not take into account the localisation of objects in an image which results into less intuitive image segmentation.

4 Laplacian Centrality of weighted Networks

The centrality of a node is a measure of how important or central a node is with in a network. A variety of centrality measures have been introduced for undirected unweighted networks based on various definitions of importance of a node. These include: degree, closeness, betweenness, eigenvector and subgraph centralities (Freeman, 1978; Estrada and Rodriguez-Velazquez, 2005). Standard centrality measures i.e degree, closeness, and betweenness were extended to weighted networks due to the fact that weighted networks provide more information about the network and therefore measures applied to these networks are of great importance (Newman, 2001; Barrat et al., 2004; Opsahl, 2009; Opsahl et al., 2010). These standard centrality measures give information on either the local environment of a node (i.e degree centrality) or the global position of a node in the network (i.e closeness, betweenness and subgraph centralities). This implies that information about the intermediate (between local and global) environment of a node cannot be captured by any of the standard centralities, yet, such information is very useful in the study of real-world networks. For instance, quantifying the relative importance of a particular actor in a social network. It is for this reason that a new type of centrality known as the Laplacian centrality was introduced (Qi et al., 2012).

With Laplacian centrality measure, the importance of a node is determined by the ability of the network to respond to the deactivation of the node from the network. In other words, it is a measure of the relative drop of Laplacian energy in the network due to the removal (or deactivation) of the node from the network. The drop of Laplacian energy with respect to node v is determined by the number of 2-walks that v participates in the network (Qi et al., 2012).

4.1 Laplacian energy of a network

Let $G = (V, E, W)$ be a simple undirected weighted network with the vertex set $V(G) = \{v_1, v_2, \cdots, v_n\}$, edge set E, where each edge $e = (v_i, v_j)$ is attached with a weight w_{ij}. If there is no edge between v_i and v_j, $w_{i,j} = 0$. In addition, $w_{i,i} = 0$ and $w_{i,j} = w_{j,i}$. We define

$$\mathbf{W(G)} = \begin{pmatrix} 0 & w_{1,2} & \dots & w_{1,n} \\ w_{2,1} & 0 & \dots & w_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{n,1} & w_{n,2} & \dots & 0 \end{pmatrix} \text{ and } \mathbf{X(G)} = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & x_n \end{pmatrix},$$

where x_i is the sum-weight of vertex v_i given by $x_i = \sum_{j=1}^{n} w_{i,j} = \sum_{u \in N(v_i)} w_{v_i,u}$, where $N(v_i)$ is the neighborhood of v_i.

Definition 25 (Weighted Laplacian matrix)**.** *The Laplacian matrix of a weighted network G is the matrix* $\mathbf{L(G)} = \mathbf{X(G)} - \mathbf{W(G)}$.

Definition 26 (Laplacian Energy of a network)**.** *Let $G = (V, E, W)$ be a weighted network on n vertices and $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of its Laplacian matrix. The Laplacian energy of G is defined as*

$$E_L(G) = \sum_{i=1}^{n} \lambda_i^2.$$

As networks become larger, computing eigenvalues of the Laplacian matrix becomes very hard. We therefore, use the entries of the Laplacian matrix rather than its eigenvalues to compute the Laplacian energy of a network as given by Theorem 5.

Theorem 5. *For any network $G = (V, E, W)$ on n vertices whose vertex sum-weights are $x_1, x_2, \dots, x_n$ respectively, we have*

$$E_L(G) = \sum_{i=1}^{n} x_i^2 + 2 \sum_{i<j} w_{i,j}^2. \tag{106}$$

Corollary 1. *If H is an arbitrary subgraph of a network G, then $E_L(H) \leq E_L(G)$. And equality holds if and only if $V(G) - V(H)$ is a set of isolated vertices.*

4.2 Laplacian centrality for a vertex

Definition 27. *If $G = (V, E, W)$ is a network with n vertices $\{v_1, v_2, \ldots, v_n\}$. Let G_i be the network obtained by deleting v_i from G. The Laplacian centrality is given by*

$$C_L(v_i, G) = \frac{(\Delta E)_i}{E_L(G)} = \frac{E_L(G) - E_L(G_i)}{E_L(G)} \tag{107}$$

For any vertex v, the denominator remains unchanged and from Corollary 1, we can tell that $E_L(G) - E_L(G_i)$ is non-negative. We then focus on obtaining the expression for $(\Delta E)_i$. In order to obtain the graph theoretical descriptions of Laplacian centrality, we will study the k-walks (discussed in Chapter 2) for the weighted graph, specifically, for $k = 2$. For better understanding of the weighted network concept, we represent a weighted network as an unweighted multigraph network by replacing each edge $e = (v_i, v_j)$ with $w_i j$ copies of multiedges. For instance, for a 2-walk $v_1 v_2 v_3$ in a weighted network, the number of 2-walks in its corresponding unweighted network is $w_{v_1,v_2} w_{v_2,v_3}$.

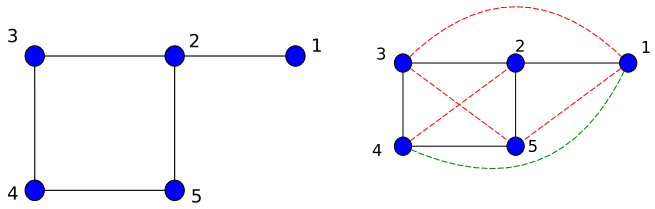
Theorem 6. *Let $G = (V, E, W)$ be a weighted network of n vertices $v_1, v_2, \ldots, v_n$. Let G_i be the network obtained by deleting vertex v from G, then the drop of Laplacian energy with respect to v_i is*

$$(\Delta E)_i = E_L(G) - E_L(G_i) = 4 \cdot NW_2^C(v_i) + 2 \cdot NW_2^E(v_i) + 2 \cdot NW_2^M(v_i). \tag{108}$$

where $NW_2^C(v_i)$, $NW_2^E(v_i)$, and $NW_2^M(v_i)$ are closed 2-walks containing vertex v_i, non-closed 2-walks with vertex v_i as one of the end points and non-closed 2-walks with vertex v_i as the middle point(Qi et al., 2012).

4.3 Laplacian Centrality of a node of a Graph with long-range interactions

In this section, we study the impact of long-range interactions in a graph on the Laplacian centrality of the nodes in the network. For example, let us consider a simple graph in Fig.26a. We account for the long range interactions (shown by the brocken red edges) by the mellin transform($c_k = k^{-s}$) of the k-Laplacian matrices of the graph for $s = 2$ as visible in Fig. 26b.



(a) (b)

Figure 26: (a) is a simple graph. (b) is the mellin transformation of the graph in (a) with $s = 2$.

Table 1: Laplacian Centralities of nodes with and without long range interactions

Node	Centrality without longrange	Centrality with longrange
1	0.2500	0.3175
2	0.6875	0.6712
3	0.5000	0.5185
4	0.4375	0.4702
5	0.5000	0.5185

From the table, we observe that accounting for longrange interactions results into a change in the laplacian centrality of the nodes in the graph. For instance, for nodes 1,3,4, and 5 the centralities increase while for node 2, its centrality decreases.

We can observe from Fig.1 that nodes with high degrees such as node 2 have less long-range links incident to them than their counter parts with lower degree (such as node 1). Since the Laplacian centrality of a given node is determined from the walks within the intermediate environment of the node that is the nearest and second nearest neighbourhood. As a result, removal of a given node from a graph with long-range interactions results into a drop in larger drop in energy compared to a scenario with no longrange interactions. However, despite a greater drop in energy, in some cases it is observed that for high degree nodes, the ratio of drop in energy on node removal to the total energy in a graph is lower than the case with no long range interactions as observed in the table for node 2.

4.4 Laplacian Centrality of an edge

Centrality measures in networks have proved to be relevant tools in network analysis. They are indicators of the 'importance' of nodes and edges in the networks. Though most work has been geared towards the study of importance of nodes (i.e degree, closeness, betweeness, subgraph, eigenvector, Laplacian centralities,etc.), interest in the study of edge centralities is now gaining ground with earlier works by Anthonisse (Anthonisse, 1971) and then following by prominent work by Girvan and Newman (Newman and Girvan, 2004). Some of the known edge centralities include edge degree, edge closeness (Ortiz Gaona et al., 2016), edge-betweenness which is calculated based on either shortest path distances or random walks as in (Newman and Girvan, 2004), k-path edge centrality Alahakoon et al. (2011), among others. The motivation for the introduction of edge centrality measures lies in real-world applications in a wide range of context such as community detection in networks (Newman and Girvan, 2004), identifying significant power lines, communication or transportation lines whose failure cause serious breakdown of the power, communication and transportation systems respectively (Ortiz Gaona et al., 2016), identification of strong relationships among people in social networks, etc.

Similar to laplacian centrality of a node, we define the Laplacian centrality of an edge as the drop in Laplacian energy when an edge is removed from a network. Let us consider an undirected weighted network $G = (V, E)$. The Laplacian energy of G is given by

$$E_L(G) = \sum_{i=1}^{n} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2, \tag{109}$$

where $x(v_i) = \sum_{j \in N(v_i)} w_{i,j}$.

On removing an arbitrary edge $e_{1,2}$, without loss of generality, assume $H = G - e_{1,2}$. Let $N(v_i)$ be the neighborhood of vertex v_i in G, $v_i \in e_{i,j}$ represent that edge $e_{i,j}$ is incident to vertex v_i in G, and $x'(v_i)$ be the corresponding sum-weight of the vertex v_i in H. We have:

$$x'(v_i) = \begin{cases} x(v_i) - w_{v_1,v_2} & \text{if } v_i \sim e_{1,2}, \\ x(v_i) & \text{otherwise.} \end{cases} \tag{110}$$

The Laplacian energy of the subgraph is given by

$$E_L(H) = \sum_{v_i \sim e_{1,2}} (x(v_i) - w_{v_1,v_2})^2 + \sum_{v_i \nsim e_{1,2}} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - 2 \cdot w_{v_1,v_2}^2 \tag{111}$$

Definition 28 (Laplacian centrality of an edge, $C_L(e)$)**.** *The Laplacian centrality of an edge e, in Graph G is given by*

$$C_L(e) = \frac{E_L(G) - E_L(H)}{E_L(G)} = \frac{\Delta E_L}{E_L(G)} \tag{112}$$

From (112), we observe the denominators remain the same in computing the Laplacian centrality for edges. This then directs our attention to obtaining the graph theoretical descriptions of the drop in the Laplacian energy when a given node is removed from the graph.

Following a similar procedure in the proof for Theorem 6, the drop in Laplacian energy which the difference between (109) and (111) is given by

$$\begin{aligned}
\Delta E &= \sum_{i=1}^{n} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - \Big(\sum_{v_i \sim e_{1,2}} (x(v_i) - w_{v_1,v_2}^2)^2 + \sum_{v_i \nsim e_{1,2}} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - 2\cdot w_{v_1,v_2}^2\Big) \\
&= \sum_{i=1}^{n} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - \Big(\sum_{v_i \sim e_{1,2}} (x^2(v_i) - 2x(v_i)\cdot w_{v_1,v_2} + w_{v_1,v_2}^2) + \sum_{v_i \nsim e_{1,2}} x^2(v_i) + 2\sum_{i<j} w_{i,j}^2 - 2\cdot w_{v_1,v_2}^2\Big) \\
&= 2\sum_{v_i \sim e_{ij}} x(v_i).w_{v_1,v_2} - \sum_{v_i \sim e_{ij}} w_{1,2}^2 + 2\cdot w_{v_1,v_2}^2 \\
&= 2\sum_{v_i \sim e_{i,j}} (x(v_i)\cdot w_{v_1,v_2}) - 2\cdot w_{v_1,v_2}^2 + 2\cdot w_{v_1,v_2}^2 \\
&= 2\sum_{v_i \sim e_{i,j}} x(v_i)\cdot w_{v_1,v_2} \\
&= 2\sum_{v_i \sim e_{i,j}} w_{v_1,v_2} \sum_{u \in N(v_i)} w_{v_i,u} \\
&= 2\cdot\sum_{v_1} w_{v_2,v_1} \sum_{u \in N(v_1); u \neq v_2} w_{v_1,u} + 2\cdot\sum_{v_2} w_{v_1,v_2} \sum_{u \in N(v_2); u \neq v_1} w_{v_2,u} + 2\cdot w_{v_1,v_2}\cdot w_{v_1,v_2} + 2\cdot w_{v_2,v_1}\cdot w_{v_2,v_1}
\end{aligned}$$

$$\Delta E = 2\cdot NW_2^U(v_1(E), v_2(M)) + 2\cdot NW_2^U(v_1(E), v_2(M)) + 4\cdot NW_2^C(v_1, v_2) \tag{113}$$

where

$NW_2^U(v_1(E), v_2(M))$ is the number of non-closed walks of length 2 with vertex v_2 as the middle point and vertex v_1 as an end point.

$NW_2^U(v_1(E), v_2(M))$ is the number of non-closed walks of length 2 with vertex v_1 as the middle point and vertex v_2 as an end point.

$NW_2^C(v_1, v_2)$ is the number of closed walks of length 2 between vertices v_1 and v_2.

From (113), we can easily tell the energy drop can be obtained by taking into account the immediate neighbourhood around the edge, that is, the nearest neighbours of the two nodes that are incident to the edge of interest.

Example 5. *Let us consider the weighted graph in Fig. 27. We compute the drop in Laplacian energy by First, we compute the Laplacian energy $E(G)$ of the graph as follows:*

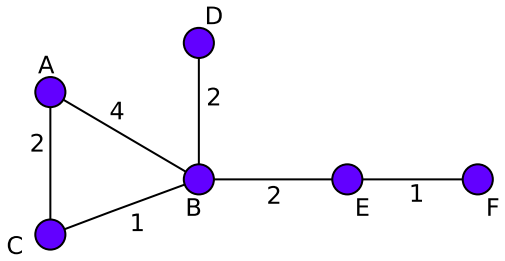


Figure 27

$$\begin{aligned}
E_L(G) &= \sum_{i=1}^{n} x_i^2 + 2\sum_{i<j} w_{i,j}^2 \\
&= (6^2 + 3^2 + 9^2 + 3^2 + 1^2 + 2^2) + 2(4^2 + 2^2 + 1^2 + 2^2 + 2 + 1^2) \\
&= 200
\end{aligned}$$

Table 2: Laplacian Centralities of edges

Edge(e)	$E(H); H = G - e$	$E(G) - E(H)$	$\Delta(E)$ by walks method in (113)
a,c	164	$200 - 164 = 36$	$2(2) + 2(8) + 4(4) = 36$
b,c	176	$200 - 176 = 24$	$2(2 + 2 + 4) + 2(2) + 4(1) = 24$
a,b	80	$200 - 80 = 120$	$2(8) + 2(4 + 8 + 8) + 2(16) = 120$
b,d	156	$200 - 156 = 44$	$2(0) + 2(4 + 8 + 2) + 4(4) = 44$
b,e	152	$200 - 158 = 88$	$2(2) + 2(4 + 8 + 2) + 4(4) = 48$
e,f	192	$200 - 192 = 8$	$2(0) + 2(2) + 4(1) = 8$

From Table 2, we observe the drop in energy computed by the difference between Laplacian energy of the graph G and that of the subgraph, H obtained on removing edge e is equal to that obtained using closed and non-closed walks in (113).

4.5 Laplacian Energy of Directed Networks

Though a lot of research has been inclined towards the study of undirected networks, however, directed networks are equally important as well. For instance, directed networks are used in the representation and study of structures of various real world networks such as communication networks, transportation networks, web-graphs among others.

The concept of Laplacian matrix and Laplacian energy of directed networks has been explored in various ways in Kissani and Mizoguchi (2010), Adiga and Smitha (2009).

Definition 29 (Skew Laplacian Energy of a simple, connected digraph)**.** *Adiga and Smitha (2009) Let G be a simple (n, m) digraph with vertex set $V(G) = \{v_1, v_2, ..., v_n\}$ and arc set $\Gamma(G) \in V(G) \times V(G)$. The skew-adjacency matrix of G is the $n \times n$ matrix $S(G) = [a_{ij}]$ where $a_{ij} = 1$ whenever $(v_i, v_j) \in E(G), a_{ij} = -1$ whenever $(v_j, v_i) \in \Gamma(G), a_{ij} = 0$otherwise. Let $D(G) = diag(d_1, d_2, d_3, \cdots, d_n)$ the diagonal matrix with vertex degrees $d_1, d_2, \cdots, d_n$ for vertices $v_1, v_2, \cdots, v_n$. The skew Laplacian energy of a digraph G is defined as*

$$E_{SL}(G) = \sum_{i=1}^{n} \lambda_i^2 \tag{114}$$

where n is the order of G and $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of the Laplacian energy $L(G) = D(G) - S(G)$ of the digraph G.

Definition 30 (Laplacian Energy of a directed graph)**.** *Kissani and Mizoguchi (2010) Let $A(G)$ be the adjacency matrix of a directed graph G whose entries are given as $A = (a_{ij})$, where $a_{ij} = 1$ whenever (v_i, v_j) is a directed edge and 0 otherwise. Let $D(G) = diag(d_1^{out}, d_2^{out}, \cdots, d_n^{out}$ be diagonal matrix with outdegree of the vertices $v_1, v_2, \cdots, v_n$. The Laplacian matrix $L(G) = D(G) - A(G)$. Then the Laplacian energy of G is defined as*

$$LE(G) = \sum_{i=1}^{n} \mu_i^2 \tag{115}$$

where n is the order of G and μ_i for $(i = 1, 2, \cdots, n)$ are the eigenvalues of $L(G)$.

From Definition 30, the following theorems hold:

Theorem 7. *Let G be a directed graph with vertex degrees $d_1^{out}, d_2^{out}, \cdots, d_n^{out}$. Then If G is a simple directed graph, then*

$$LE(G) = \sum_{i=1}^{n} (d_i^{out})^2 \tag{116}$$

If G is a symmetric directed graph (i.e a graph in which each edge is bidirected) then

$$LE(G) = \sum_{i=1}^{n} d_i^{out}(d_i^{out} + 1). \tag{117}$$

Theorem 8. *If G is a disconnected directed graph with components $G_1, G_2, \cdots, G_n$,*

$$LE(G) = \sum_{i=1}^{n} LE(G_i). \tag{118}$$

4.6 Laplacian Centrality of a node in directed network

Earlier on in this chapter, for simple undirected networks, we defined Laplacian Centrality of a vertex as the relative drop in Laplacian energy when the vertex is removed from the network. Similarly, we adopt the same definition for vertices in directed networks. Based on Definition 30, the drop in Laplacian energy of a simple directed graph $G = (V, E)$ due to the removal of vertex v is given by

$$\begin{aligned}
\Delta LE &= LE(G) - LE(G - v) \\
&= \sum_{i=1}^{n} (d_i^{out})^2 - \Big(\sum_{i \in N(v_{in})} (d_i^{out} - 1)^2 + \sum_{i \notin N(v_{in}); i \neq v} (d_i^{out})^2 \Big) \text{ where } N(v_{in}) = \{i \in N(v) | e_{iv} \in E\} \\
&= (d_v^{out})^2 + \sum_{i \in N(v_{in})} (2d_i^{out} - 1)
\end{aligned}$$

4.7 Edge Centrality based Edge reversal

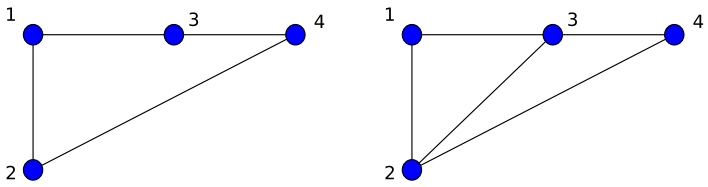
Instead of measuring the importance of an edge based on the drop in energy on edge removal as discussed for undirected networks, we consider the importance of an edge in a directed network as the relative change in energy on edge reversal.

4.8 Laplacian Energy as a fair measure of robustness of network

Apart of its application in identifying the most important nodes and edges that is the Laplacian centrality, the Laplacian energy of a network has been recently identified as a measure of robustness of network (Yang et al., 2016). Robustness of networks is defined as the response of a network on edge or node removal or addition. The study of robustness of networks plays a vital role in design of systems, understanding the performance and stability of systems such as ecological systems, technological systems, biological system among others. For instance, in air transportation sector, it is paramount to design air traffic networks that will ensure robustness when one or more flights are removed or added to the network thereby posing a need for an effective measure of robustness in such networks. The algebraic connectivity which is the second smallest eigenvalue (λ_2) of the Laplacian matrix of a network is one of the most common measure of robustness in networks (Jamakovic and Van Mieghem, 2008; Byrne et al., 2005). Unfortunately, the algebraic connectivity captures only the global information about the connectivity of a network which implies that some changes in network through addition or removal of edges may not be captured as the algebraic connectivity remains constant. However, work in (Yang et al., 2016) introduces the Laplacian energy of a network as a fair and effective measure of robustness compared to the algebraic connectivity. This is so because the Laplacian centrality captures the local information of the network thus a change in network structure by edge or node removal or addition is captured by a change in Laplacian energy. This therefore justifies Laplacian energy as a fair and effective measure of robustness compared to algebraic connectivity.

4.9 Comparison of Laplacian Energy and Algebraic Connectivity as measures of robustness of networks

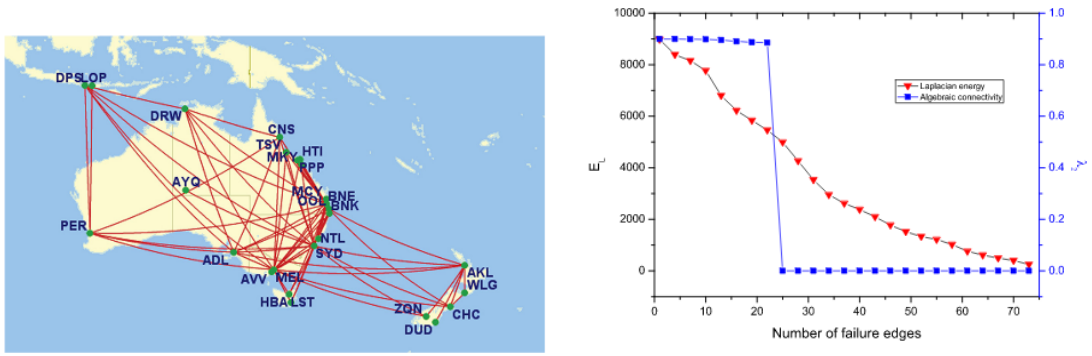
Let us consider a simple example in which we illustrate how the both the algebraic connectivity and Laplacian energy capture the changes in networks due to edge addition or removal.



(a) (b)

Figure 28: The simple network in (a) has algebraic connectivity $\lambda_2 = 2$ and Laplacian energy of 24. Adding a new edge,$e_{2,3}$ to form a network in (b) with algebraic connectivity $\lambda_2 = 2$ and Laplacian energy 36. We observe that on a change in network through edge addition, the algebraic connectivity remains constant while the Laplacian energy changes to reflect a change in network structure.

Further more, let us consider the comparison between the two measures on real work network that is the air traffic network of Jet-star Asia Airway among Indonesia, Australia, and New Zealand shown in (Fig. 29a).



(a) (b)

Figure 29: (a) is the air traffic network route map for Jetstar Asia Airway (Indonesia, Australia, and New Zealand) in 2015, (b) is a plot of Laplacian energy E_L and algebraic connectivity λ_2 against the number of randomly failed edges from the or Jetstar Asia Airway (Indonesia, Australia, and New Zealand). Source: (Yang et al., 2016)

From the Fig. (29b), we observe that values for both the laplacian energy and algebraic connectivity decreases with removal of edges from the network. However, on removal of 20 to 30 edges, the algebraic connectivity abruptly drops from 0.9 to close to 0 (that is in only one instance) which signifies a disconnection in the network that is, more than one connected component in the network. on the other hand though, the laplacian curve indicates gradual degradation of the network robustness on removal of 20 to 30 edges of the network. The ability of the Laplacian energy measure to capture the change from one connected component to more connected components in much more instances makes it an effective measure for network robustness over algebraic connectivity.

5 Graph Clustering

5.1 Principal Component Analysis (PCA) on objects

In performing PCA on a given number m of objects. we follow the following steps:

i) Select the objects from a database to which PCA is to be applied say m objects.

ii) Obtain the graph representation of each object using Delanauy triangulation with the harris corner detection technique. This results into graphs $G_1, G_2, \cdots, G_m$.

iii) Construct the feature vector $\mathbf{B}$ for each graph. For instance the vector can be obtained from the k leading co-efficients of the heat content polynomial that is $\mathbf{B} = (q_1, q_2, \cdots, q_k)^T$. Alternatively $\mathbf{B} = (l_1, l_2, \cdots, l_k)^T$ from the k leading Laplacian eigenvalues.

iv) Construct the matrix $\mathbf{S} = [\mathbf{B_1}|\mathbf{B_2}|\cdots|\mathbf{B_m}]$. The feature vectors form the columns of $\mathbf{S}$.

v) Compute the convariance matrix C, which is given by $C = \hat{\mathbf{S}}\hat{\mathbf{S}}^T$, where $\hat{\mathbf{S}}$ is matrix of normalised data which is computed by subtracting the mean of the feature vectors from each column of the matrix $\mathbf{S}$.

vi) Obtain the principal components direction by performing eigendecomposition on the covariance matrix C that is

$$C = \sum_{i=1}^{m} \lambda_i \mathbf{v}_i \mathbf{v}_i^T, \tag{119}$$

where λ_i are the eigenvalues and $\mathbf{v}_i$ are the eigenvectors. This is followed by selection of the first s leading eigenvectors termed as the principal components. The matrix $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_s)$. By selecting the principal components, we reduce the dimension of the data.

vii) Finally, we project individual graphs on to the co-ordinate eigenspace by $\mathcal{B} = \mathbf{U}\mathbf{B}_k$. Therefore, each graph G_k is represented by an s-component vector $\mathbf{B}_k$ in the eigenspace.

References

C. Adiga and M. Smitha. On the skew laplacian energy of a digraph. In *Int. Math. Forum*, volume 4, pages 1907–1914, 2009.

T. Alahakoon, R. Tripathi, N. Kourtellis, R. Simha, and A. Iamnitchi. K-path centrality: A new centrality measure in social networks. In *Proceedings of the 4th workshop on social network systems*, page 1. ACM, 2011.

R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

J. M. Anthonisse. The rush in a directed graph. *Stichting Mathematisch Centrum. Mathematische Besliskunde*, (BN 9/71):1–10, 1971.

H. Anton and C. Rorres. Elementary linear algebra, (2000). *Anton Textbook Inc, Ottawa*, 2007.

F. Aurenhammer and R. Klein. Voronoi diagrams. *Handbook of computational geometry*, 5:201–290, 2000.

F. Aurenhammer, R. Klein, and D.-T. Lee. *Voronoi diagrams and Delaunay triangulations*. World Scientific Publishing Company, 2013.

J. R. Banavar, A. Maritan, and A. Rinaldo. Size and form in efficient transportation networks. *Nature*, 399(6732):130, 1999.

A.-L. Barabási. *Network science*. Cambridge university press, 2016.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

J. S. Baras and P. Hovareshti. Efficient and robust communication topologies for distributed decision making in networked systems. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 3751–3756. IEEE, 2009.

A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11): 3747–3752, 2004.

A. Barrat, M. Barthelemy, and A. Vespignani. *Dynamical processes on complex networks*. Cambridge university press, 2008.

M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

A. E. Brouwer and W. H. Haemers. *Spectra of graphs*. Springer Science & Business Media, 2011.

R. Byrne, J. Feddema, and C. Abdallah. Algebraic connectivity and graph robustness. *Sandia National Laboratories, Albuquerque, New Mexico*, 87185, 2005.

J. L. Casti. MS Windows NT complexity, September 26, 2017. URL https://www.britannica.com/science/complexity-scientific-theory.

F. Chung. The heat kernel as the pagerank of a graph. *Proceedings of the National Academy of Sciences*, 104(50):19735–19740, 2007.

F. Chung. A local graph partitioning algorithm using heat kernel pagerank. *Internet Mathematics*, 6(3): 315–330, 2009.

K. C. Das. The laplacian spectrum of a graph. *Computers & Mathematics with Applications*, 48(5): 715–724, 2004.

P. De Meo, E. Ferrara, G. Fiumara, and A. Ricciardello. A novel measure of edge centrality in social networks. *Knowledge-based systems*, 30:136–150, 2012.

C. H. Edwards and D. E. Penney. *Differential equations and boundary value problems*, volume 2. Prentice Hall, 2004.

W. Ellens and R. E. Kooij. Graph measures and network robustness. *arXiv preprint arXiv:1311.5064*, 2013.

P. Erdos and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1): 17–60, 1960.

E. Estrada. Quantifying network heterogeneity. *Physical Review E*, 82(6):066102, 2010.

E. Estrada. *The structure of complex networks: theory and applications*. OUP Oxford, 2011.

E. Estrada. Path laplacian matrices: introduction and application to the analysis of consensus in networks. *Linear Algebra and its Applications*, 436(9):3373–3391, 2012.

E. Estrada. Introduction to complex networks: structure and dynamics. In *Evolutionary Equations with Applications in Natural Sciences*, pages 93–131. Springer, 2015.

E. Estrada and N. Hatano. Communicability in complex networks. *Physical Review E*, 77(3):036111, 2008.

E. Estrada and J. A. Rodriguez-Velazquez. Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103, 2005.

E. Estrada, F. Kalala-Mutombo, and A. Valverde-Colmeiro. Epidemic spreading in networks with non-random long-range interactions. *Physical Review E*, 84(3):036110, 2011.

E. Estrada, P. Knight, et al. *A first course in network theory*. Oxford University Press, USA, 2015.

E. Estrada, L. V. Gambuzza, and M. Frasca. Long-range interactions and network synchronization. *arXiv preprint arXiv:1704.01349*, 2017a.

E. Estrada, E. Hameed, N. Hatano, and M. Langer. Path laplacian operators and superdiffusive processes on graphs. i. one-dimensional case. *Linear Algebra and its Applications*, 523:307–334, 2017b.

L. Euler. Leonhard euler and the königsberg bridges. *Scientific American*, 189(1):66–70, 1953.

L. Euler. The solution of a problem relating to the geometry of position. 1976.

M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *ACM SIGCOMM computer communication review*, volume 29, pages 251–262. ACM, 1999.

L. C. Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.

J. Gower. A modified leverrier-faddeev algorithm for matrices with multiple eigenvalues. *Linear Algebra and its Applications*, 31:61–70, 1980.

L. Grady. Random walks for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 28(11):1768–1783, 2006.

I. Gribkovskaia, Ø. Halskau, and G. Laporte. The bridges of königsberg—a historical perspective. *Networks*, 49(3):199–203, 2007.

I. Gutman and O. E. Polansky. *Mathematical concepts in organic chemistry*. Springer Science & Business Media, 2012.

C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, pages 10–5244. Citeseer, 1988.

J. M. Harris, J. L. Hirst, and M. J. Mossinghoff. *Combinatorics and graph theory*, volume 2. Springer, 2008.

B. A. Huberman. The laws of the web, 2001.

Internet. Network. URL https://www.google.com/search?q=internet+network&client=ubuntu&hs=xZp&channel=fs&gl=za&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiZsrLOueTMAhVMCcAKHS_8AysQ_AUIBygB&biw=1215&bih=900#channel=fs&gl=za&tbm=isch&q=computer+network&imgrc=_. [Online; accessed 2016-03-29].

M. O. Jackson. *Social and economic networks*. Princeton university press, 2010.

A. Jamakovic and P. Van Mieghem. On the robustness of complex networks by using the algebraic connectivity. *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*, pages 183–194, 2008.

R. Kasprzak. Diffusion in networks. *Journal of Telecommunications and Information Technology*, pages 99–106, 2012.

P. Kissani and Y. Mizoguchi. Laplacian energy of directed graphs and minimizing maximum outdegree algorithms. *Kyushu University Institutional Repository*, 2010.

K. Kloster and D. F. Gleich. Heat kernel based community detection. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1386–1395. ACM, 2014.

R. Kondor and J.-P. Vert. Diffusion kernels. *kernel methods in computational biology*, pages 171–192, 2004.

J. Lafferty and G. Lebanon. Diffusion kernels on statistical manifolds. *Journal of Machine Learning Research*, 6(Jan):129–163, 2005.

D. López-Pintado. Diffusion in complex social networks. *Games and Economic Behavior*, 62(2):573–590, 2008.

H. Ma, H. Yang, M. R. Lyu, and I. King. Mining social networks using heat diffusion processes for marketing candidates selection. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 233–242. ACM, 2008.

J. Magouirk, S. Atran, and M. Sageman. Connecting terrorist networks. *Studies in Conflict & Terrorism*, 31(1):1–16, 2008.

P. McDonald and R. Meyers. Diffusions on graphs, poisson problems and spectral geometry. *Transactions of the American Mathematical Society*, 354(12):5111–5136, 2002.

S. Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

J. J. Molitierno. *Applications of combinatorial matrix theory to Laplacian matrices of graphs*. CRC Press, 2012.

H. P. Moravec. Visual mapping by a robot rover. In *Proceedings of the 6th international joint conference on Artificial intelligence-Volume 1*, pages 598–600. Morgan Kaufmann Publishers Inc., 1979.

H. P. Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. Technical report, STANFORD UNIV CA DEPT OF COMPUTER SCIENCE, 1980.

F. K. Mutombo. *Long-range interactions in complex networks*. PhD thesis, University of Strathclyde, 2012.

M. Newman. *Networks: an introduction*. OUP Oxford, 2010.

M. E. Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132, 2001.

M. E. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

M. E. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.

Nexus. Zachary's karate club [karate], 2012. URL http://nexus.igraph.org/api/dataset_info?id=1&format=html. [Online; accessed 2016-04-25].

M. Nikolić. Measuring similarity of graph nodes by neighbor matching. *Intelligent Data Analysis*, 16(6): 865–878, 2012.

T. Opsahl. *Structure and evolution of weighted networks*. PhD thesis, Queen Mary, University of London, 2009.

T. Opsahl and P. Panzarasa. Clustering in weighted networks. *Social networks*, 31(2):155–163, 2009.

T. Opsahl, F. Agneessens, and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

R. Ortiz Gaona, M. Postigo Boix, and J. L. Melus Moreno. Centrality metrics and line-graph to measure the importance of links in online social networks. *International Journal of New Technology and Research*, 2(12):20–26, 2016.

G. A. Pagani and M. Aiello. The power grid as a complex network: a survey. *Physica A: Statistical Mechanics and its Applications*, 392(11):2688–2700, 2013.

X. Qi, E. Fuller, Q. Wu, Y. Wu, and C.-Q. Zhang. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, 194:240–253, 2012.

X. Qi, R. D. Duval, K. Christensen, E. Fuller, A. Spahiu, Q. Wu, Y. Wu, W. Tang, C. Zhang, et al. Terrorist networks, network energy and node removal: a new measure of centrality based on laplacian energy. *Social Networking*, 2(01):19, 2013.

R. O. Saber and R. M. Murray. Agreement problems in networks with directed graphs and switching topology. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 4, pages 4126–4132. IEEE, 2003.

B. Schwikowski, P. Uetz, and S. Fields. A network of protein–protein interactions in yeast. *Nature biotechnology*, 18(12):1257, 2000.

J. Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.

L. I. Smith. A tutorial on principal components analysis. Technical report, 2002.

O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag. Organization, development and function of complex brain networks. *Trends in cognitive sciences*, 8(9):418–425, 2004.

A. R. Stoica. Delaunay diagram representations for use in image near-duplicate detection. *Senior project submittd tot he division of science, mathematics and computing of Bard College. New York*, 2011.

A. Sydney, C. Scoglio, P. Schumm, and R. E. Kooij. Elasticity: topological characterization of robustness in complex networks. In *Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Sytems*, page 19. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

D. Thanou, X. Dong, D. Kressner, and P. Frossard. Learning heat diffusion graphs. *IEEE Transactions on Signal and Information Processing over Networks*, 3(3):484–499, 2017.

M. Trajković and M. Hedley. Fast corner detection. *Image and vision computing*, 16(2):75–87, 1998.

X. F. Wang and G. Chen. Complex networks: small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*, 3(1):6–20, 2003.

D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

J. A. Williams, S. M. Dawson, and E. Slooten. The abundance and distribution of bottlenosed dolphins (tursiops truncatus) in doubtful sound, new zealand. *Canadian journal of zoology*, 71(10):2080–2088, 1993.

R. J. Wilson. *An introduction to graph theory*. Pearson Education India, 1970.

J. Wu, Y. Tan, H. Deng, Y. Li, B. Liu, and X. Lv. Spectral measure of robustness in complex networks, 2008. *arXiv preprint arXiv:0802.2564*.

B. Xiao, E. R. Hancock, and R. C. Wilson. Graph characteristics from the heat kernel trace. *Pattern Recognition*, 42(11):2589–2606, 2009.

C. Yang, J. Mao, and P. Wei. Air traffic network optimization via laplacian energy maximization. *Aerospace Science and Technology*, 49:26–33, 2016.

W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4):452–473, 1977.

L. A. Zager and G. C. Verghese. Graph similarity scoring and matching. *Applied mathematics letters*, 21(1):86–94, 2008.