Heat Kernel in Multi-hop heat diffusion on Graph

Alice Nanyanzi

March 15, 2018

Abstract

The heat kernel of a graph is the fundamental solution to the heat diffusion equation associated with Laplacian of a graph. The heat kernel can also be viewed as a description of flow of information (or heat in this case) over the edges of the graph with time. It is computed by exponentiating the Laplacian eigen-system with time. In this paper, we extend this concept by considering diffusion on the graph along direct edges as well as along indirect edges which we refer to as long-range interactions. The heat kernel resulting as a fundamental solution to the mentioned type of diffusion is what we refer to as the generalised heat kernel. We account for the long-range interactions in two ways namely; the Mellin and Laplace transforms of the Laplacian. We then perform experiments on the generalised heat kernel which is then followed by explanation of results.

1 Introduction

Dynamics on networks such as diffusion, consensus, synchronisation, etc. have recently attracted attention of researchers. This is because of their applicability in various areas such as modelling of epidermic spread, diffusion of information in social networks, among others.

It has been observed in modelling of dynamic processes on networks that the substance in consideration say information, heat, disease, rumour, and many more do spread that the flow is not only along direct edges of the networks but also through indirect interactions. Estrada explored ways of accounting for the indirect or long-range interactions by use of the social distance analogy which was associated with constraints in selection of the conductance parameter x (Estrada et al., 2011). In his subsequent work, elegant approaches were put

forward based on the Mellin and Laplace transforms of the graph Laplacian (Estrada, 2012; Estrada et al., 2017a).

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, etc. (Newman, 2010).

Diffusion process over networks is one of the methods used in developing simple models that depict the spread of infections in a population, dissemination of information over social network for instance social network marketing, spread of heat over a conductor, among others. A number of models based on diffusion process have been developed and documented in (Estrada et al., 2011; Kasprzak, 2012; López-Pintado, 2008).

The diffusion kernel which is the fundamental solution of the heat equation is quite an important aspect in the study of diffusion on networks. It has been used as a means of characterising of networks, in graph clustering, among others (Xiao et al., 2009). The heat kernel in these cases is based on diffusion that occurs along the edges of a network. In this work, however, we explore the heat kernel for which diffusion occurs through both short-range (over the edges) and long-range interactions. We call this the generalised heat kernel.

2   Heat Diffusion Model

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. With the heat diffusion coefficient, $\varepsilon \in [0, 1]$ which 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.

Consider a simple undirected network. Suppose we have a quantity ϕ_i of a substance on a vertex i at time t. Suppose the substance moves along edges from node j to an adjacent node i at a rate $C(\phi_j - \phi_i)$, where C is the diffusion constant $(C > 0)$. The amount of substance moving from j to i in a small time interval dt is $C(\phi_j - \phi_i)\, dt$. Then the rate at which substance ϕ_i changes is given by

$$\frac{d\phi_i}{dt} \;=\; \varepsilon \sum_j A_{ij}(\phi_j - \phi_i) = \varepsilon \sum_j A_{ij}\phi_j - \varepsilon\phi_i \sum_j A_{ij} = \varepsilon \sum_j A_{ij}\phi_j - \varepsilon\phi_i k_i.$$

Thus,

$$\frac{d\phi_i}{dt} = \varepsilon \sum_j (A_{ij} - \delta_{ij}k_i)\phi_j, \tag{1}$$

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 (Newman, 2010; Ma et al., 2008; López-Pintado, 2008).

In matrix-vector notation, we have

$$\frac{d\boldsymbol{\phi}}{dt} = -\varepsilon \mathbf{L}\boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{2}$$

whose solution is

$$\boldsymbol{\phi}(t) = \boldsymbol{\phi}_0 \; e^{-\varepsilon \mathbf{L}t}. \tag{3}$$

Alternatively, the solution can be expressed as a linear combination of eigenvectors of the Laplacian matrix (Edwards and Penney, 2004). That is

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

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. The Laplacian matrix is given by

$$L_{ij} = \begin{cases} k_i & \text{if } i = j, \\ -1 & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where k_i is the degree of node i.

2.1 The Heat Kernel

As discussed earlier, diffusion of heat on a graph can be modelled by the equation

$$\frac{d\phi}{dt} = -L\phi, \tag{5}$$

where L is either the Laplacian matrix or the normalised Laplacian Matrix.

The heat kernel is the fundamental solution to the diffusion equation (5). It is obtained by exponentiating the Laplacian eigensystem and it is given by

$$H_t = e^{(-Lt)} \tag{6}$$

3

It literally describes the flow of substance (heat) across edges (direct interactions) in the graph (Xiao et al., 2009).

When t tends to zero, the heat kernel behaviour can be obtained from the Taylor's expansion of (6) which is

$$e^{-\mathbf{L}t} = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!}\mathbf{L}^k = \mathbf{I} - \mathbf{L}t + \frac{\mathbf{L}^2 t^2}{2!} - \frac{\mathbf{L}^3 t^3}{3!} + \cdots \tag{7}$$

Thus,

$$\lim_{t \longrightarrow 0}\left(e^{-\mathbf{L}t}\right) = \mathbf{I} - \mathbf{L}t. \tag{8}$$

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, we apply the spectral decomposition of Equation 6 which is

$$H_t = \mathbf{V}e^{(-\Lambda t)}\mathbf{V}^T \tag{9}$$

where Λ is the diagonal matrix of eigenvalues λ_i of the Laplacian matrix and V is a matrix with eigenvectors of the Laplacian matrix as columns.

$$H_t = \sum_{i=0}^{n} e^{(-\lambda_i t)} v_i v_i^T \tag{10}$$

where λ_is are the eigenvalues of L in a non-decreasing order $0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ and v_i is the eigenvector corresponding to the eigenvalue λ_i. Thus,

$$\lim_{t \longrightarrow \infty}\left(e^{-\mathbf{L}t}\right) = e^{(-\lambda_2 t)} v_2 v_2^T. \tag{11}$$

From Equation 11, its evident that for large t, the heat kernel behaviour is determined by the global structure of the graph.

Heat Kernel Trace

The trace of the heat kernel $Tr(H_t)$ is given by

$$Tr(H_t) = Tr(\mathbf{V}e^{(-\Lambda t)}\mathbf{V}^T) = Tr(e^{(-\Lambda t)}(\mathbf{V}^T\mathbf{V})) = Tr(e^{(-\Lambda t)}) \tag{12}$$

Thus, the trace function of the heat kernel is given by

$$Z(t) = Tr(H_t) = Tr(e^{(-\Lambda t)}) = \sum_{i=1}^{|V|} e^{-\lambda_i t}, \tag{13}$$

where λ_i is the eigenvalue of the normalised Laplacian matrix Xiao et al. (2009). For a connected graph ($\lambda_1 = 0$), Equation 13 can be written as

$$Z(t) = 1 + e^{-\lambda_2 t} + e^{-\lambda_3 t} + \cdots + e^{-\lambda_N t} \tag{14}$$

4

Xiao Xiao et al. (2009) developed a potential application of the trace formula as a basis for distinguishing different graphs 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 shows the plot heat kernel trace against time for the three graphs.
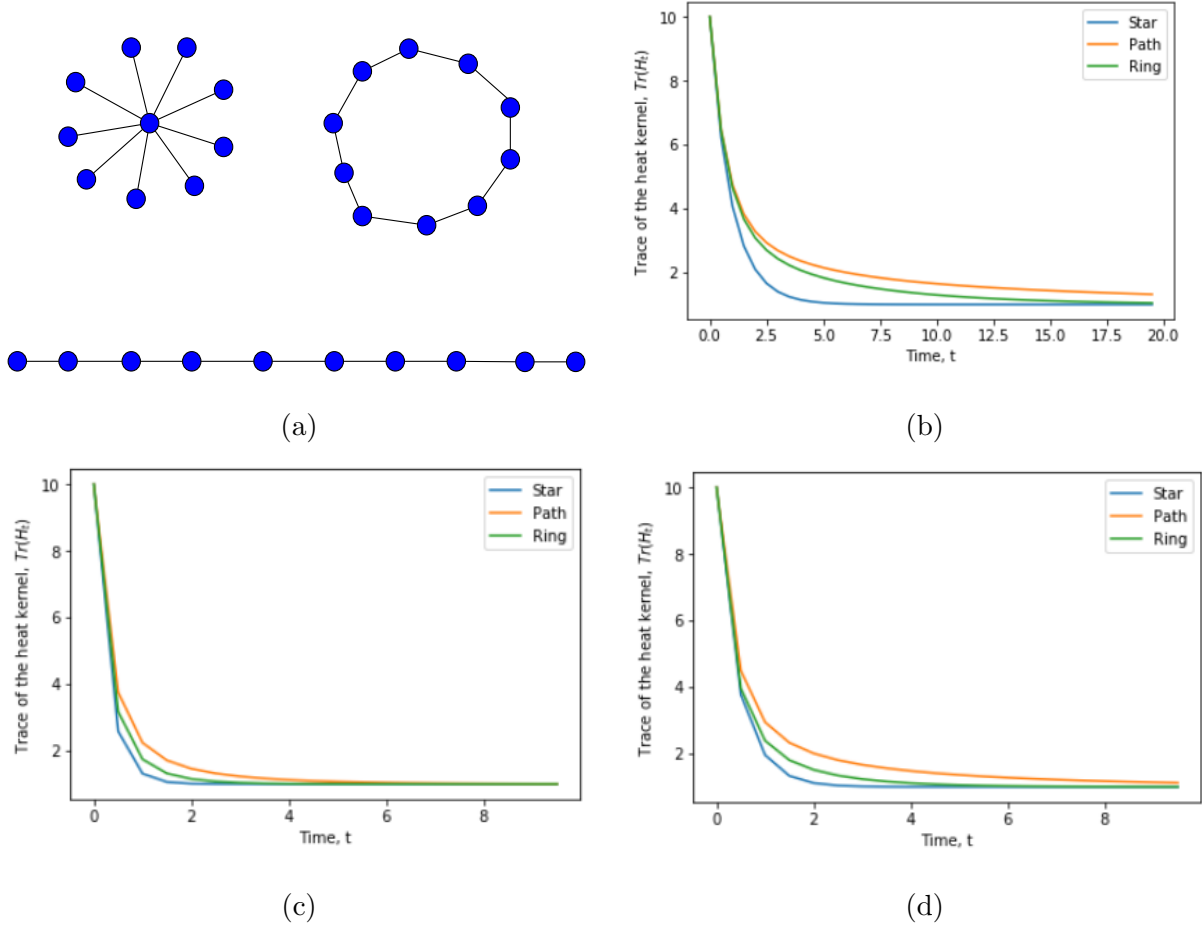

(a)

(b)

(c)

(d)

Figure 1: (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. Figures (c) and (d) are the plots of trace function against time for the 3 graphs based on the Mellin transforms of the k-Laplacians for $s = 2$ and $s = 3$ respectively.

From (b), we observe that since the 3 graphs are different in terms of structure, the curves too take on different shapes. It is evident that since path and 2-regular graphs have similar topologies, their corresponding graphs are close to each other while for star graph, its curve has a 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).

3   The Generalised Laplacian Matrix

In (Estrada, 2012), Estrada put forward the concept of the k-path Laplacian matrices which is based on the idea that one consider traversing a network by taking hops of length of $1, 2, \cdots, d_{max}$ where d_{max} is the diameter of the network. The k-path Laplacian matrices are therefore 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 by considering the analogy of polarisation process on a network as explained in (Estrada, 2012).

Following (Estrada et al., 2017b), we recall that 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:

$$L_k(ij) = \begin{cases} -1 & \text{if } d_{i,j} = k, \\ \delta_k(i) & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases} \qquad (15)$$

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.

The k-path Laplacian matrices are real, symmetric, positive semi-definite matrices. In addition, 0 is an eigenvalue of $\mathbf{L}_k$ with a $\mathbf{1}$ eigenvector that is, $\mathbf{L}_k \mathbf{1}^T = \mathbf{0}$ (Estrada, 2012).

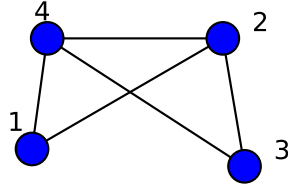For example, let us compute the k-path Laplacians for a simple graph G in Fig.2.

Figure 2: A network of size 4.

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

Using the k-path Laplacian, we can obtain an expression for the generalised Laplacian Matrix as a linear combination of the k-path Laplacian matrices given by

$$L_G = \left(\sum_{k=1}^{d_{max}} c_k \mathbf{L}_k \right) = c_1 L_1 + c_2 L_2 + \cdots + c_{d_{max}} L_{d_{max}}, \tag{16}$$

where d_{max} is the diameter and c_k are the co-efficients.

4   Choice of co-efficients, c_k

Considering the generalised Laplacian Matrix in Eqn.16, the co-efficients c_k are chosen in such a way that the effect of the long-range interactions weakens as the separation between nodes increases. The choice of the co-efficients c_k is evident in many man-made and naturally evolving systems where communication among the agents of the system follows a spatial decay. For instance in sensor systems, 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.

In this paper, we consider two forms that the spatial decay can take on namely:

7

1. Exponential form

   For the exponential form, Eqn.16 attains a generalised form based on the Laplace-transformed k-Laplacian:

   $$\tilde{\mathbf{L}}_{L,\lambda} = \mathbf{L} + \sum_{k=2}^{\Delta} e^{-\lambda k}\mathbf{L}_k, \tag{17}$$

   where $\lambda > 0$ is a parameter that depends on the specific situation to be modelled. Thus, the coefficients of Eqn.16 are $c_1 = 1$ and $c_{k \geq 2} = e^{-\lambda k}$.

2. Power-law

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

   $$\tilde{\mathbf{L}}_{M,s} = \sum_{k=1}^{\Delta} k^{-s}\mathbf{L}_k, \tag{18}$$

   where $s > 0$ is a parameter that depends on the specific situation to be modelled and the coefficients of Eqn.16 are $c_k = k^{-s}$.

Following the two transforms, we can the define the Laplacian matrix by **?**

$$\mathbf{L}_G = \begin{cases} \sum_{d=1}^{d_{max}} k^{-s} L_d & \text{for } \tau = Mell, s > 0 \\ L + \sum_{d=2}^{d_{max}} e^{-\lambda d} L_d & \text{for } \tau = Lapl, \lambda > 0. \end{cases} \tag{19}$$

4.1 Properties of the Generalised Laplacian matrix, L_G

Proposition 1. *Real and Symmetric*

For $k = 1, 2, \cdots, d_{max}$, let $c_1 L_1, c_2 L_2, \cdots, c_{d_{max}} L_{d_{max}}$ be symmetric matrices of the same size. Consider $L_G = c_1 L_1 + c_2 L_2 + \cdots + c_{d_{max}} c_{d_{max}} L_{d_{max}}$. Following the property of transposes; $(c_1 L_1 + c_2 L_2 + \cdots + L_{d_{max}})^T = c_1 L_1^T + c_2 L_2^T + \cdots + c_{d_{max}} L_{d_{max}}^T$. Since L_k are symmetric matrices, we know that for a given k, $L_k^T = L_k$. Thus, $(c_1 L_1 + c_2 L_2 + \cdots + c_{d_{max}} L_{d_{max}})^T = c_1 L_1 + c_2 L_2 + \cdots + c_{d_{max}} L_{d_{max}}$ which implies $L_G^T = L_G$.

Proposition 2. *Positive-Semi definite*

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

$$\mathbf{y}^T L_G \mathbf{y} = \mathbf{y}^T (c_1 L_1 + c_2 L_2 + \cdots + c_\delta L_\delta)\mathbf{y} = \mathbf{y}^T c_1 L_1 \mathbf{y} + \mathbf{y}^T c_2 L_2 \mathbf{y} + \cdots + \mathbf{y}^T c_\delta L_\delta \mathbf{y} \tag{20}$$

Since $\mathbf{y}^T c_k L_k \mathbf{y} \geq 0$ for $c_k > 0$ and $1 \leq k \leq d_{max}$, then

$$\mathbf{y}^T L_G \mathbf{y} \geq 0 \tag{21}$$

$\square$

Proposition 3. *Zero is an eigenvalue of L_G with 1 as an eigenvector. Let $\mathbf{v} = \mathbf{0}$, then $L_G\mathbf{v} = (c_1 L_1 + c_2 L_2 + \cdots + c_{d_{max}} L_{d_{max}})\mathbf{v}$ which gives*

$$L_G\mathbf{v} = c_1 L_1\mathbf{v} + c_2 L_2\mathbf{v} + \cdots + c_{d_{max}} L_{d_{max}}\mathbf{v}. \tag{22}$$

since $c_k L_k \mathbf{v} = 0$, Then $L_G\mathbf{v} = 0$ which implies that atleast 0 is an eigenvalue of L_G with 1 as an eigenvector.

5   Generalised Heat Diffusion Model

In the previous section, we have discussed the diffusion process that occurs through direct interactions between nearest neighbours in the network. In this section however, we consider the fact that in many real-world processes, it is observed that interactions with in networks do not only occur among nearest neighbours but also among non-nearest which we refer to as the long range interactions as illustrated in Fig. 3

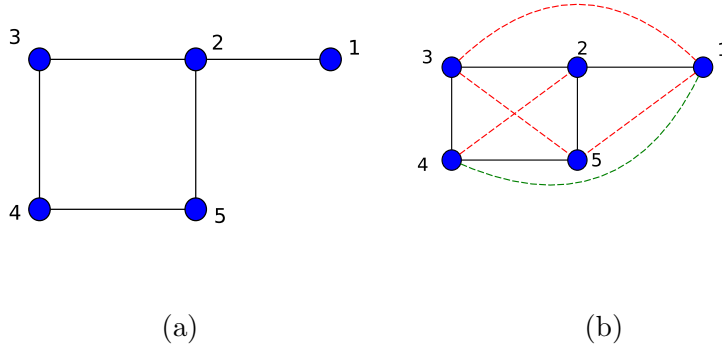(a)   (b)

Figure 3: (a) is a simple graph indicating the direct interactions followed in the normal diffusion model. (b) is a simple graph indicating both direct interactions (black solid lines), long-range interactions at hops of length 2 (red broken lines) and 3 (green-broken lines) that depict the generalised diffusion model.

5.1   The Generalised Heat Kernel

The diffusion process accounting for the long-range interactions is defined by

$$\frac{d\boldsymbol{\phi}}{dt} = -\varepsilon L_G \boldsymbol{\phi}, \quad \boldsymbol{\phi}(0) = \boldsymbol{\phi}_0, \tag{23}$$

where L_G is the generalised Laplacian matrix.

The heat kernel H_{Gt} of the generalised diffusion process which we refer to as the generalised heat kernel, the fundamental solution to eqn.23 is given by

$$H_{Gt} = e^{-L_G t},\tag{24}$$

which on expansion can be written as

$$H_{Gt} = e^{-t(c_1 L_1 + c_2 L_2 + \cdots + c_\Delta L_\Delta)},\tag{25}$$

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.6.

5.2 Trace of the Generalised Heat Kernel

The trace of the generalised heat kernel is therefore given by

$$Z_G(t) = Tr(H_{Gt}) = \sum_{i=1}^{|V|} e^{-\lambda_{iG}t}.\tag{26}$$
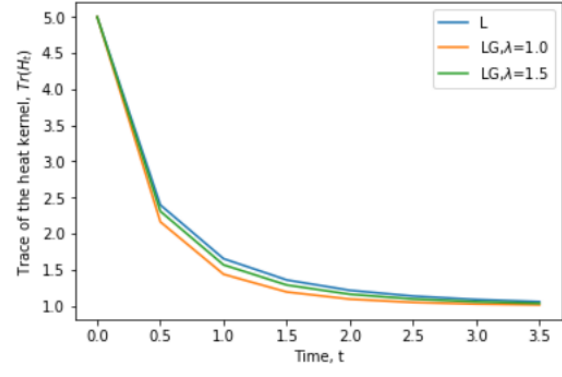
Alternatively, Eqn. **??** can be written as

$$Z_G(t) = 1 + e^{-\lambda_{2G}t} + e^{-\lambda_{3G}} + \cdots + e^{-\lambda_{NG}t}\tag{27}$$

6 Experiments and Results

Let us consider a simple graph in Fig. 3a. We then perform simulation for normal diffusion as well as generalised diffusion on the graph. We compare results for generalised diffusion based on the Mellin and Laplace transforms for different values of s and λ respectively.
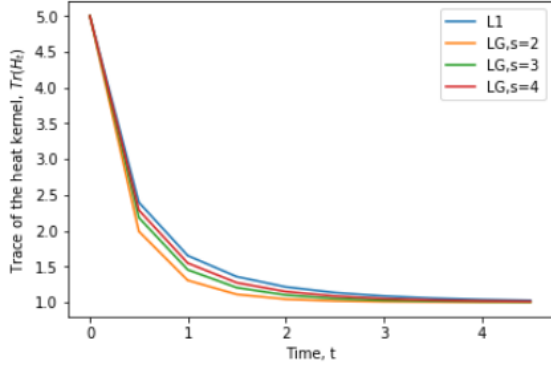
Figure 4: (a) is a plot of the heat kernel function against time for mellin-transform diffusion on graph in Fig. 3a. (b) is a plot of the heat kernel function against time for laplace -transform diffusion on graph in Fig. 3a

From Fig. 4, we that for both cases, as the value of the parameters s and λ increases, the curve of the heat kernel function approaches that corresponding heat kernel function with no long-range interactions.

6.1 Comparison of Barabasi and E-R networks for Mellin and Laplace

References

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

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

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

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

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, F. Kalala-Mutombo, and A. Valverde-Colmeiro. Epidemic spreading in networks with nonrandom 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.

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

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

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. 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.

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.

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

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