Network Science **9** (3): 312–327, 2021
doi:10.1017/nws.2021.10

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Efficient Laplacian spectral density computations for networks with arbitrary degree distributions

Grover E. C. Guzman[1], Peter F. Stadler[2] and André Fujita[1,*]

[1]Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010 - São Paulo - SP, 05508-090 Brazil (e-mail: grover@ime.usp.br) and [2]Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, Universität Leipzig, Härtelstraße 16-18, Leipzig, D-04107, Germany (e-mail: studla@bioinf.uni-leipzig.de)
*Corresponding author. Email: andrefujita@usp.br

Action Editor: Ulrik Brandes

Abstract

The network Laplacian spectral density calculation is critical in many fields, including physics, chemistry, statistics, and mathematics. It is highly computationally intensive, limiting the analysis to small networks. Therefore, we present two efficient alternatives: one based on the network's edges and another on the degrees. The former gives the exact spectral density of locally tree-like networks but requires iterative edge-based message-passing equations. In contrast, the latter obtains an approximation of the spectral density using only the degree distribution. The computational complexities are $\mathcal{O}(|E|\log(n))$ and $\mathcal{O}(n)$, respectively, in contrast to $\mathcal{O}(n^3)$ of the diagonalization method, where n is the number of vertices and $|E|$ is the number of edges.

Keywords: spectral density; Laplacian matrix; normalized Laplacian matrix; configuration model tree-like network

1. Introduction

Calculating the Laplacian spectrum of networks is critical in many fields, including physics, chemistry, statistics, and mathematics (Trinajstic *et al.*, 1994; Almendral & Díaz-Guilera, 2007; Nadler *et al.*, 2009; Hakimi-Nezhaad & Ashrafi, 2014). We know that the spectrum encodes structural characteristics of the network. For example, by analyzing the Laplacian spectrum, we can obtain its diameter (Chung, 1989), the number of spanning trees (Bollobás, 1998), resistance distance (Klein & Randić, 1993; Chen & Zhang, 2007), Kirchoff index (Wu, 2004), vertex covers (Chen & Jost, 2012), Kemeny's constant (Pan *et al.*, 2018), and chromatic number (Sun & Das, 2020). Furthermore, they are also associated with dynamic processes, such as network synchronizability (Barahona & Pecora, 2002), quantum transition probabilities (Mülken *et al.*, 2005), and global mean first-passage time (Sun *et al.*, 2017).

Besides, the Laplacian spectral density has been proposed for network structure and dynamics analysis (Jamakovic & Van Mieghem, 2006; de Lange *et al.*, 2014; Liu *et al.*, 2015), classification (Wegner *et al.*, 2018), model selection (Granziol *et al.*, 2018), eigenvector localization (Hata & Nakao, 2017), detectability transitions (Peixoto, 2013), and the number of clusters estimation (Granziol *et al.*, 2018).

However, the Laplacian spectrum (or spectral density) is known only for a few graph classes, including trees (Chen & Jost, 2012; Zhang *et al.*, 2008; Sun *et al.*, 2016), quadrilateral (Li & Hou, 2017), k-triangle, k-quadrilateral (Huang & Li, 2018), hexagonal Möbius (Ma & Bian,

2019), Erdös–Rényi random graphs, fractal trees and dendrimers (Julaiti *et al.*, 2013), 3-prism (Ding *et al.*, 2014), Dyson hierarchical (Agliari & Tavani, 2017), Dice Lattice (Li *et al.*, 2016), and some small-world networks (Liu & Zhang, 2013; Liu *et al.*, 2015). In general, the spectrum needs to be computed directly with the help of numerical methods such as the QR decomposition (Gander, 1980), the power iteration method (Booth, 2006), or the Jacobi eigenvalue algorithm (Sleijpen & Van der Vorst, 2000). A severe practical limitation of all these approaches is that the computational complexity is $\mathcal{O}(n^3)$ (where n is the number of vertices of the network), which precludes their application to large networks. Thus, approximation methods have been proposed. Most of them focus on calculating the network spectral density (not the Laplacian spectral density) (Chung *et al.*, 2004; Semerjian & Cugliandolo, 2002; Rogers *et al.*, 2008; Metz *et al.*, 2011; Nadakuditi & Newman, 2013; Newman *et al.*, 2019; Newman, 2019; Cantwell & Newman, 2019). Recently, (Cantwell & Newman, 2019) introduced a message-passing approach for the (normalized) Laplacian spectral density using the local neighborhood of each vertex. Still, it requires computing matrix inversions and matrix-vector multiplications, which are computationally expensive (Cantwell & Newman, 2019).

Here, we present two algorithms to compute the (normalized) Laplacian spectral density of a locally tree-like network, namely edge-based message-passing and degree-based algorithms, in which computational complexities are $\mathcal{O}(|E| \log n)$ and $\mathcal{O}(n)$, respectively. Our edge-based algorithm calculates the exact Laplacian spectral density of weighted locally tree-like networks with self-loops. Alternatively, our degree-based algorithm approximates the Laplacian spectral density of unweighted locally tree-like networks without self-loops using only the degree distribution. In contrast to (Cantwell & Newman, 2019), neither the edge- nor the degree-based algorithm requires the computation of matrix inversions or matrix-vector multiplications. Also, different from the method proposed by (Newman *et al.*, 2019), our edge-based algorithm allows weights at the edges and self-loops.

2. Spectral density of a matrix

We start with the derivation of a system of message-passing equations for computing the spectra of weighted networks that are locally tree-like and, in addition, may contain self-loops.

The main characteristic of a locally tree-like network is that the local neighborhood of any node is a tree (Kowalska, 1987; Adcock *et al.*, 2013). In other words, consider a specific starting node and all nodes at a distance of maximum l from the starting node. Then, the sub-graph induced by this node-set will take the form of a tree, with probability tending to one as the graph size tends to infinity (Newman, 2018).

Let $G = (V, E)$ be an undirected weighted graph, where V and E are the sets of vertices and edges, respectively, $n = |V|$ be the number of vertices, and $w_{u,v} \in \mathbb{R}$ be the weight of edge $(u, v) \in E$. Let $\mathbf{M}$ be the weighted $n \times n$ adjacency matrix of G defined as follows:

$$\mathbf{M}_{u,v} = \begin{cases} w_{u,v}, & \text{if } (u, v) \in E, \text{ and} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Since the graph G is undirected, $\mathbf{M}$ is symmetric, and its n eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are real. Notice we do not restrict the values that $w_{u,v}$ can assume; we only require $\mathbf{M}$ to be symmetric. Writing $\delta(x)$ for Dirac's delta function, the spectral density of $\mathbf{M}$ is defined as (Newman *et al.*, 2019):

$$\rho(x) = \frac{1}{n} \sum_{i=1}^{n} \delta(x - \lambda_i) \tag{2}$$

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

Expressing Dirac's delta function as the limit of a Lorentzian (or Cauchy) distribution (Newman *et al.*, 2019), we can rewrite Equation (2) as:

$$\rho(x) = -\frac{1}{n\pi} \lim_{\epsilon \to 0^+} Im \sum_{i=1}^{n} \frac{1}{x - \lambda_i + i\epsilon} \tag{3}$$

Writing $\mathbf{I}$ for the identity matrix and $z = x + i\epsilon$ to emphasize that the imaginary part of z is close to zero (Newman *et al.*, 2019), we obtain

$$\rho(z) = -\frac{1}{n\pi} \sum_{i=1}^{n} \frac{1}{z - \lambda_i} = -\frac{1}{n\pi} Tr(z\mathbf{I} - \mathbf{M})^{-1} \tag{4}$$

By expanding the matrix inverse in Equation (4) as a geometric series, we obtain

$$(z\mathbf{I} - \mathbf{M})^{-1} = \frac{1}{z} \sum_{k=0}^{\infty} \left(\frac{\mathbf{M}}{z}\right)^k \tag{5}$$

This series converges only if $|1/z| \leq |1/\lambda_1|$ (Newman *et al.*, 2019).

We can express every matrix $\mathbf{M}$ as the sum of a diagonal matrix $\mathbf{H}$ and a matrix $\mathbf{W}$ with vanishing diagonal elements. In our setting, $\mathbf{H}$ then contains the weights of all self-loops of the associated graph, and $\mathbf{W}$ is the weighted adjacency matrix of the associated graph without self-loops. By introducing the diagonal matrix $\mathbf{H}(z) = z\mathbf{I} - \mathbf{H}$ and its inverse $\mathbf{H}^{-1}(z)$ yields

$$\begin{aligned} \rho(z) &= -\frac{1}{n\pi} Tr(z\mathbf{I} - \mathbf{H} - \mathbf{W})^{-1} \\ &= -\frac{1}{n\pi} Tr(\mathbf{H}(z) - \mathbf{W})^{-1} \\ &= -\frac{1}{n\pi} Tr\left\{\mathbf{H}^{-1}(z)(\mathbf{I} - \mathbf{H}^{-1}(z)\mathbf{W})^{-1}\right\} \end{aligned} \tag{6}$$

By expanding $(\mathbf{I} - \mathbf{H}^{-1}(z)\mathbf{W})^{-1}$ in Equation (6), we obtain

$$(\mathbf{I} - \mathbf{H}^{-1}(z)\mathbf{W})^{-1} = \sum_{k=0}^{\infty} \left(\mathbf{H}^{-1}(z)\mathbf{W}\right)^k \tag{7}$$

Finally, using Equation (7) in Equation (6), we obtain

$$\rho(z) = -\frac{1}{n\pi} \sum_{k=0}^{\infty} Tr\left\{\mathbf{H}^{-1}(z)(\mathbf{H}^{-1}(z)\mathbf{W})^k\right\} \tag{8}$$

The convergence condition of Equation (8) is the same as presented for Equation (5).

The problem is thus reduced to the efficient computation of $Tr\{\mathbf{H}^{-1}(z)(\mathbf{H}^{-1}(z)\mathbf{W})^k\}$. This will be achieved with a message-passing algorithm.

2.1 Edge-based message-passing method for weighted locally tree-like networks

In this section, we describe an extension of the message-passing algorithm developed by (Newman *et al.*, 2019). (Newman *et al.*, 2019) developed a message-passing algorithm for unweighted symmetric adjacency matrices. We extended it for weighted symmetric adjacency matrices, which may include self-loops.

A characteristic of locally tree-like networks is that any closed walk's length has an even number of steps (because any closed walk traverses the same edge twice) (Newman *et al.*, 2019).

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

Since $\mathbf{W}$ is the weighted adjacency matrix of a locally tree-like network, $\mathbf{H}^{-1}(z)\mathbf{W}$ also describes a locally tree-like network. Therefore, $Tr\{\mathbf{H}^{-1}(z)(\mathbf{H}^{-1}(z)\mathbf{W})^k\}$ is equal to 0 only when the exponent k is odd; that is, there are only closed walks of even length.

We can consider the matrix $\mathbf{H}^{-1}(z)\mathbf{W}$ as a complex weighted adjacency matrix, where the (i, j)-th position indicates the weight of a walk of length one from node i to node j. To obtain the (i, j)-th position of $(\mathbf{H}^{-1}(z)\mathbf{W})^2$, we sum the values of $(\mathbf{H}^{-1}(z)\mathbf{W})_{ir} \times (\mathbf{H}^{-1}(z)\mathbf{W})_{rj} \forall 1 \leq r \leq n$. Notice that we multiply the weights of the edges (i, r) and (r, j). If the value of $(\mathbf{H}^{-1}(z)\mathbf{W})_{ir} \times (\mathbf{H}^{-1}(z)\mathbf{W})_{r,j}$ is different from zero, it means that there is a walk of length two from i to j by passing through the edges (i, r) and (r, j). Thus, by defining the weight of a walk as the product of the edge weights it traverses, $(\mathbf{H}^{-1}(z)\mathbf{W})^2_{i,j}$ represents the sum of the walks' weights from node i to node j passing through two edges. Hence, $(\mathbf{H}^{-1}(z)\mathbf{W})^k_{ij}$ represents the sum of all the walks' weights from node i to node j passing through k edges. Let r be a positive integer and $n^{uv}_{2r}(z)$ be the sum of the (complex-valued) weights of all closed walks of length $2r$ that begins by traversing the edge from node u to node v and ends by traversing the same edge from v to u. Other edges may be traversed many times, but the edge between u and v is traversed only once each way. Let N_v be the set of neighbor nodes of v, $s_{u,v}(z) = (\mathbf{W}_{u,v})^2/\{(z - \mathbf{H}_{u,u})(z - \mathbf{H}_{v,v})\}$, and $\delta(i, j)$ be the Kronecker delta function. Then, we have

$$n^{uv}_{2r}(z) = s_{u,v}(z) \sum_{m=1}^{\infty} \left[\sum_{\substack{w_1 \in N_v, \\ w_1 \neq u}} \cdots \sum_{\substack{w_m \in N_v, \\ w_m \neq u}} \right] \left[\sum_{r_1=1}^{\infty} \cdots \sum_{r_m=1}^{\infty} \right] \prod_{i=1}^{m} n^{vw_i}_{2r_i}(z) \times \delta\left(r - 1, \sum_{i=1}^{m} r_i\right) \tag{9}$$

To obtain $n^{uv}_{2r}(z)$, we consider that before the walk ends after traversing edge (u, v), it makes m consecutive excursions from node v to each of its neighbors except by u. The total sum of such walks weight is the product of the weights of distinct excursions, with individual lengths $2r_1, \ldots, 2r_m$, i.e. $2r - 2$. To solve Equation (9), we define

$$h^{uv}(z) = \sum_{r=1}^{\infty} n^{uv}_{2r}(z) \tag{10}$$

By using Equation (9) into Equation (10), we obtain

$$\begin{aligned} h^{uv}(z) &= s_{u,v}(z) \sum_{m=1}^{\infty} \left[\sum_{\substack{w_1 \in N_v, \\ w_1 \neq u}} \cdots \sum_{\substack{w_m \in N_v, \\ w_m \neq u}} \right] \prod_{i=1}^{m} \sum_{r_i=1}^{\infty} n^{vw_i(z)}_{2r_i} \\ &= s_{u,v}(z) \sum_{m=1}^{\infty} \left[\sum_{\substack{w \in N_v, \\ w \neq u}} h^{vw}(z) \right]^m \\ &= \frac{s_{u,v}(z)}{1 - \sum_{\substack{w \in N_v, \\ w \neq u}} h^{vw}(z)} \end{aligned} \tag{11}$$

We can interpret $h^{uv}(z)$ as a message that node u received from v. We calculate it from the messages received by v from its other neighbors.

To solve our original problem (Equation (8)), we need to obtain the total sum of weighted closed walks that start and end at node u with any number of steps. Let N_u be the set of neighborsof

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

node u, then the total sum of weighted closed walks of length $2r$ that starts and ends at node u (denoted by $n^u_{2r}(z)$) can be obtained by:

$$n^u_{2r}(z) = \sum_{m=1}^{\infty} \left[\sum_{v_1 \in N_u} \cdots \sum_{v_m \in N_u} \right] \left[\sum_{r_1=1}^{\infty} \cdots \sum_{r_m=1}^{\infty} \right] \prod_{i=1}^{m} n^{uv_i(z)}_{2r_i} \delta \left(r, \sum_{i=1}^{m} r_i \right) \tag{12}$$

To calculate the total sum of the weighted closed walks of node u, we need to sum the values of $n^u_{2r}(z)/(z - \mathbf{H}_{u,u})$, for all $r \geq 1$, i.e.,

$$\begin{aligned} g^u(z) &= \sum_{r=1}^{\infty} \frac{n^u_{2r}(z)}{z - \mathbf{H}_{u,u}} \\ &= \frac{1}{z - \mathbf{H}_{u,u}} \sum_{m=1}^{\infty} \left[\sum_{v_1 \in N_u} \cdots \sum_{v_m \in N_u} \right] \prod_{i=1}^{m} \sum_{r_i=1}^{\infty} n^{2r_i}_{uv_i}(z) \\ &= \frac{1}{z - \mathbf{H}_{u,u}} \sum_{m=1}^{\infty} \left[\sum_{v \in N_u} h^{uv}(z) \right]^m \\ &= \frac{1}{z - \mathbf{H}_{u,u}} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \end{aligned} \tag{13}$$

Now we can express Equation (8) using Equation (13) as follows:

$$\rho(z) = -\frac{1}{n\pi} \sum_{u=1}^{n} \frac{1}{z - \mathbf{H}_{u,u}} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \tag{14}$$

Finally, by using both Equations (11) and 14, we calculate the spectra of $\mathbf{M}$. First, initialize the messages $h^{uv}(z)$ randomly. Then, iterate Equation (11) until convergence (Newman *et al.*, 2019; Rogers *et al.*, 2008). Finally, input the obtained $h^{uv}(z)$ into Equation (14). This method has a computational complexity of $\mathcal{O}(d|E|)$, where d and $|E|$ are the diameter and number of edges of the network (Koller & Friedman, 2009). Nonetheless, in locally tree-like networks, d is proportional to $\log(n)$ (Newman, 2018). Thus, our method has a final complexity of $\mathcal{O}(\log(n)|E|)$.

2.2 Degree-based methods

This section will use the message-passing equations developed in Section 2.1 to develop a degree-based approximation for the (normalized) Laplacian spectra. We consider that the network is unweighted and also that it was generated by the configuration model. The network must be unweighted because, in this case, the degree of a node represents the number of neighbors. Thus, we can group the vertices by their degree and use the "centroid" as the representative. However, for weighted graphs, the degree is the sum of the weights of the edges incident to the node. Such values do not represent the number of neighbors. Thus, we cannot use the same idea. The rationale to use the configuration model is that it generates locally tree-like networks when the number of vertices tends to infinity, based on a given degree sequence (Newman, 2018).

In the following subsections, we show how to obtain a degree-based approximation for the Laplacian and normalized Laplacian spectra.

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

2.2.1 Degree-based method for the Laplacian spectral density

Let $\deg(u)$ be the degree of node u, $\mathbf{D}$ be the degree matrix, where $\mathbf{D}_{i,j} = \deg(i)$ if $i = j$ and $\mathbf{D}_{i,j} = 0$ otherwise, and $\mathbf{A}$ be the adjacency matrix of the network. The Laplacian matrix $\mathbf{L}$ of a network is defined as:

$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

To obtain the spectral density of $\mathbf{L}$, first, replace $\mathbf{H}$ and $\mathbf{W}$ by $\mathbf{D}$ and $-\mathbf{A}$, respectively. Then, set $\mathbf{D}(z)$ as a diagonal matrix with $\mathbf{D}_{i,j}(z) = z - \deg(i)$ if $i = j$, and $\mathbf{D}_{i,j}(z) = 0$, otherwise. Next, we use them in Equation (8) and obtain

$$\rho(z) = -\frac{1}{n\pi} \sum_{k=0}^{\infty} Tr\left\{\mathbf{D}^{-1}(z)(\mathbf{D}^{-1}(z)\mathbf{A})^k\right\} \tag{15}$$

Since the adjacency matrix $\mathbf{A}$ is composed of only 0s and 1s, that is, $\mathbf{W}_{u,v} = \mathbf{A}_{u,v} = 1$, then $s_{uv}(z) = 1/((z - \deg(u))(z - \deg(v)))$. By using it in Equation (11), we obtain

$$h^{uv}(z) = \frac{1}{(z - \deg(u))(z - \deg(v))} \frac{1}{1 - \sum_{\substack{w \in N_v, \\ w \neq u}} h^{vw}(z)} \tag{16}$$

Replacing $\mathbf{H}_{u,u}$ by $\mathbf{D}_{u,u} = \deg(u)$ in Equation (13), we obtain

$$\begin{aligned} g^u(z) &= \frac{1}{z - \mathbf{D}_{u,u}} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \\ &= \frac{1}{z - \deg(u)} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \end{aligned} \tag{17}$$

Since we can express Equation (8) using Equation (17), we have that

$$\rho(z) = -\frac{1}{n\pi} \sum_{u=1}^{n} \frac{1}{z - \deg(u)} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \tag{18}$$

Now we will show the intuition to construct an efficient degree-based Laplacian spectral density estimator. For this, we use the message-passing algorithm (Equation (14)) in a network generated by a configuration model with 10, 000 nodes where half of the nodes have degree 5 and the other half degree 10. Figure 1 shows a scatter plot of $(z - \deg(u))h^{uv}(z)$, where $z = x + i\epsilon$. We observe two distinct non-overlapping groups corresponding to the degrees of the node v: the red dots are nodes of degree 5, and the blue dots represent the nodes of degree 10. Therefore, similarly to (Newman *et al.*, 2019), we could approximate each message by the mean value (centroid) of the group to which it belongs.

Consider a configuration model where the fraction of nodes with degree k is p_k. Our approximation consists of replacing $(z - \deg(u))h^{uv}(z)$ by the mean value (centroid) $h_k(z)$ of messages sent from nodes with the same degree k of node v. By rearranging the message-passing equation (Equation (16)), we have that

$$(z - \deg(v))(z - \deg(u))h^{uv}(z) = 1 + (z - \deg(v))(z - \deg(u))h^{uv}(z) \sum_{\substack{w \in N_v, \\ w \neq u}} h^{vw}(z) \tag{19}$$

By replacing $(z - \deg(u))h^{uv}(z)$ by $h_k(z)$ and calculating the average over all edges (u, v), where v has degree k, we obtain

$$(z - \deg(v))h_k(z) = 1 + h_k(z) \frac{1}{nkp_k} \sum_{v:k_v=k}^{u \in N_v} \sum_{\substack{w \in N_v, \\ w \neq u}} (z - \deg(v))h^{vw}(z) \tag{20}$$

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

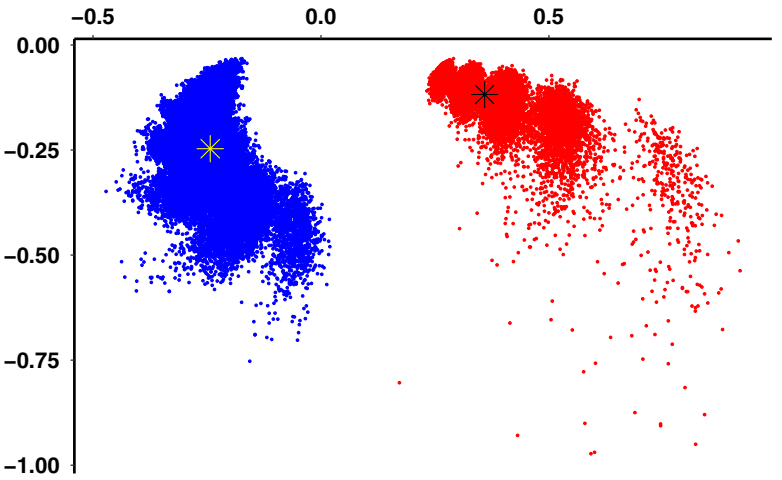

Figure 1. The values of $(z - \deg(u))h^{uv}(z)$ at $z = 7 + \epsilon i$ with $\epsilon = 0.01$ are plotted in the complex plane for a configuration model network composed of 10, 000 nodes. Half of the nodes have a degree of 5, and the other half have 10. The colored clouds correspond to nodes of degree 5 (red) and 10 (blue). The asterisks indicate the means of each cloud, which we used to develop the degree-based approximation.

The average of $(z - \deg(v))h^{vw}(z)$ over different w is independent of k, because in a configuration model, adjacent nodes v and w are uncorrelated. Then, for large networks, the average of $(z - \deg(v))h^{vw}(z)$ will be equal to the average of the whole network, denoted by $h(z)$:

$$\frac{1}{nkp_k} \sum_{v:k_v=k}^{u\in N_v} \sum_{\substack{w\in N_v,\\ w\neq u}} (z - \deg(v))h^{vw}(z) \to h(z) \tag{21}$$

Thus, we have $(z - k)h_k(z) = 1 + (k - 1)h_k(z)h(z)$ or

$$h_k(z) = \frac{1}{(z - k) - (k - 1)h(z)} \tag{22}$$

In a configuration model, given a node, its neighbor node's degree does not follow the overall network degree distribution p_k. In contrast, it follows a modified distribution named *excess degree distribution* (Newman, 2018). Let $c = \sum_k kp_k$; then, the excess degree distribution is expressed as:

$$q_k = \frac{(k + 1)p_{k+1}}{c} \tag{23}$$

Then, by using this quantity, the whole network average message is given by:

$$\begin{aligned} h(z) &= \sum_{k=1}^{\infty} \frac{kp_k}{c} h_k(z) \\ &= \frac{1}{c} \sum_{k=1}^{\infty} \frac{kp_k}{(z - k) - (k - 1)h(z)} \\ &= \sum_{k=0}^{\infty} \frac{q_k}{(z - k - 1) - kh(z)} \end{aligned} \tag{24}$$

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

We calculate the spectra by making a similar degree-based approximation to the quantities $g^u(z)$ of Equation (17). We assume that $g^u(z)$ is well approximated by the mean $g_k(z)$ (centroid) of $g^u(z)$ over all nodes u with degree k. In other words, this strategy does not work for graphs that $g_k(z)$ does not represent well the cluster, for example, when $g_k(z)$ is outside the cluster. Then, we obtain

$$g_k(z) = \frac{1}{(z-k) - kh(z)} \tag{25}$$

By summing up all $g^u(z)$, we then obtain

$$\sum_{u=1}^{n} g^u(z) = \sum_{k=0}^{\infty} np_k g_k(z) = n \sum_{k=0}^{\infty} \frac{p_k}{(z-k) - kh(z)} \tag{26}$$

Finally, by using Equation (26) into Equation (18), we calculate the Laplacian spectral density as follows:

$$\rho(z) = -\frac{1}{\pi} \sum_{k=0}^{\infty} \frac{p_k}{(z-k) - kh(z)} \tag{27}$$

Notice that we can calculate the Laplacian spectral density relying only on the degree distribution (Equations (24) and 27). First, initialize $h(z)$ (e.g., $h(z) = 0$), and iterate Equation (24) until convergence. Finally, use Equation (27) to obtain $\rho(z)$. The computational complexity of this method is dominated by Equation (24), which time complexity is $\mathcal{O}(n)$ per iteration. Usually, the number of iterations is much less than n, thus the computational complexity of our proposal is $\mathcal{O}(n)$.

2.2.2 Degree-based method for the normalized Laplacian spectral density

To develop the degree-based approximation for the normalized Laplacian spectral density, first we define the normalized Laplacian $\mathscr{L}$ of a network as:

$$\mathscr{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} \tag{28}$$

Thus, we have $\mathbf{H} = \mathbf{I}$ and $\mathbf{W} = -\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$, respectively. Then, we use $\mathscr{L}$ in Equation (8), and obtain

$$\rho(z) = -\frac{1}{n\pi} \sum_{k=0}^{\infty} Tr\left\{\mathbf{D}^{-1}(z)\left(\mathbf{D}^{-1}(z)\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}\right)^k\right\} \tag{29}$$

where $\mathbf{D}(z)$ is a diagonal matrix with $\mathbf{D}_{i,j}(z) = z - 1$ if $i = j$, and 0 otherwise. Since $\{\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}\}_{i,j} = 1/\sqrt{\deg(i)\deg(j)}$ if $\mathbf{A}_{i,j} \neq 0$, and 0 otherwise; we have that, $s_{uv}(z) = 1/(\deg(u)\deg(v)(z-1)^2)$. Then, we can rewrite Equation (11) as follows:

$$h^{uv}(z) = \frac{1}{\deg(u)\deg(v)(z-1)^2} \frac{1}{1 - \sum_{\substack{w \in N_v, \\ w \neq u}} h^{vw}(z)} \tag{30}$$

Since $\mathbf{D}_{u,u} = 1$, we rewrite Equation (13) as follows:

$$\begin{aligned} g^u(z) &= \frac{1}{z - \mathbf{D}_{u,u}} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \\ &= \frac{1}{(z-1)} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \end{aligned} \tag{31}$$

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

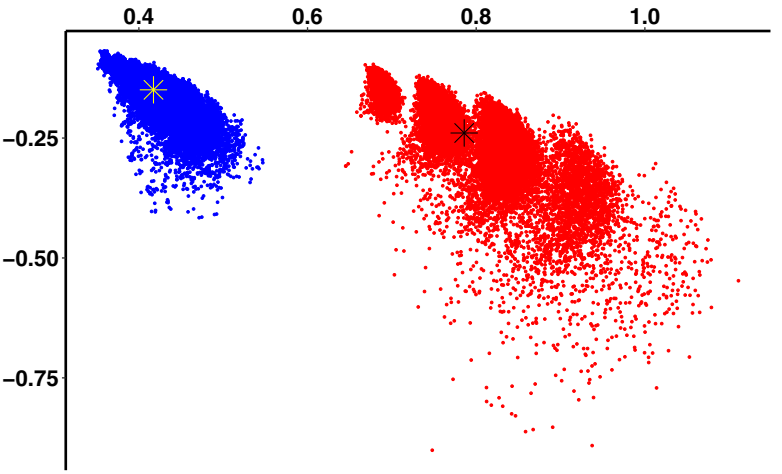

Figure 2. The values of $\deg(u)h^{uv}(z)$ at $z = 1 + \epsilon i$ with $\epsilon = 0.01$ are plotted in the complex plane for a configuration model network composed of 10, 000 nodes. Half of the nodes have a degree of 5, and the other half have 10. The colored clouds correspond to nodes of degree 5 (red) and 10 (blue). The asterisks indicate the means of each cloud, which we used to develop the degree-based approximation.

Then, we express Equation (8) by using Equation (31), and obtain

$$\rho(z) = -\frac{1}{n\pi(z-1)} \sum_{u=1}^{n} \frac{1}{1 - \sum_{v \in N_u} h^{uv}(z)} \tag{32}$$

Similarly to the Laplacian approximation, we consider again the configuration network of size $n = 10,000$, where half of the vertices have degree 5 and the other half have degree 10. Then, we use the message-passing equation (Equation (14)) to obtain all $h^{uv}(z)$. Figure 2 shows a scatter plot of $\deg(u)h^{uv}(z)$, where $z = x + i\epsilon$. We observe two distinct non-overlapping groups corresponding to the degrees of the node v: the red and blue dots correspond to nodes of degree 5 and 10, respectively. We could then approximate each message with the mean value of the group to which it belongs, similarly as done for the Laplacian approximation.

We consider again the configuration model where the fraction of nodes with degree k is p_k. Our approximation consists of replacing $\deg(u)h^{uv}(z)$ by the mean value (centroid) $h_k(z)$ of messages sent from nodes with the same degree k of node v. By rearranging the message-passing Equation (30), we obtain

$$(z-1)^2 \deg(v) \deg(u) h^{uv}(z) = 1 + (z-1)^2 \deg(v) \deg(u) h^{uv}(z) \sum_{w \in N_v, w \neq u} h^{vw}(z) \tag{33}$$

By replacing $\deg(u)h^{uv}(z)$ by $h_k(z)$, and averaging Equation (33) over all edges (u, v), where v has degree k, we obtain

$$(z-1)^2 \deg(v) h_k(z) = 1 + (z-1)^2 h_k(z) \frac{(k-1)}{nkp_k} \sum_{v: k_v = k}^{u \in N_v} \sum_{w \in N_v, w \neq u} \deg(v) h^{vw}(z) \tag{34}$$

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

Since adjacent nodes v and w are uncorrelated, the average of $\deg(v)h^{vw}(z)$ over different w is independent of k. Then, it will be equal to the average of the whole network ($h(z)$) as $n \to \infty$:

$$\frac{1}{nkp_k} \sum_{v:k_v=k}^{u \in N_v} \sum_{\substack{w \in N_v, \\ w \neq u}} \deg(v) h^{vw}(z) \to h(z) \tag{35}$$

Thus, we have $(z-1)^2 k h_k(z) = 1 + (z-1)^2(k-1)h_k(z)h(z)$ or

$$h_k(z) = \frac{1}{(z-1)^2} \frac{1}{k-(k-1)h(z)} \tag{36}$$

Here, we also consider the degree excess distribution. Then, the whole network average message is given by:

$$\begin{aligned} h(z) &= \sum_{k=1}^{\infty} \frac{kp_k}{c} h_k(z) \\ &= \frac{1}{(z-1)^2 c} \sum_{k=1}^{\infty} \frac{kp_k}{k-(k-1)h(z)} \\ &= \frac{1}{(z-1)^2} \sum_{k=0}^{\infty} \frac{q_k}{(k+1)-kh(z)} \end{aligned} \tag{37}$$

where we made a change of variable $k \to k+1$. Similarly to what we did in Section 2.2.1, we assume that $g^u(z)$ (Equation (31)) can be approximated by $g_k(z)$ (centroid), that is, the mean of $g^u(z)$ over all nodes u with degree k. Then, we obtain

$$g_k(z) = \frac{1}{(z-1)} \frac{1}{1-h(z)} \tag{38}$$

.

By summing up all nodes, we obtain

$$\sum_{u=1}^{n} g^u(z) = \sum_{k=0}^{\infty} np_k g_k(z) = \frac{n}{z-1} \sum_{k=0}^{\infty} \frac{p_k}{1-h(z)} = \frac{1}{z-1} \frac{n}{1-h(z)} \tag{39}$$

Finally, by inputing Equation (39) into Equation (32), the normalized Laplacian spectral density is given by:

$$\rho(z) = -\frac{1}{\pi(z-1)} \frac{1}{1-h(z)} \tag{40}$$

Then, by using Equations 37 and 40, we obtain the normalized Laplacian spectral density based only on the degree distribution. First, initialize $h(z)$ with any initial value (e.g., $h(z) = 0$), and iterate Equation (37) until convergence. Then, use the calculated value of $h(z)$ into Equation (40) to obtain $\rho(z)$. The computational complexity of this method is dominated by Equation (37), in which time complexity is $\mathcal{O}(n)$ per iteration. Since the number of iterations is usually much less than n, the overall time complexity of our proposal is $\mathcal{O}(n)$.

3. Experiments and results

To evaluate the advantages and limitations of our proposed methods, we applied them to graphs generated by three random graph models, namely the Poisson, configuration, and power-law

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

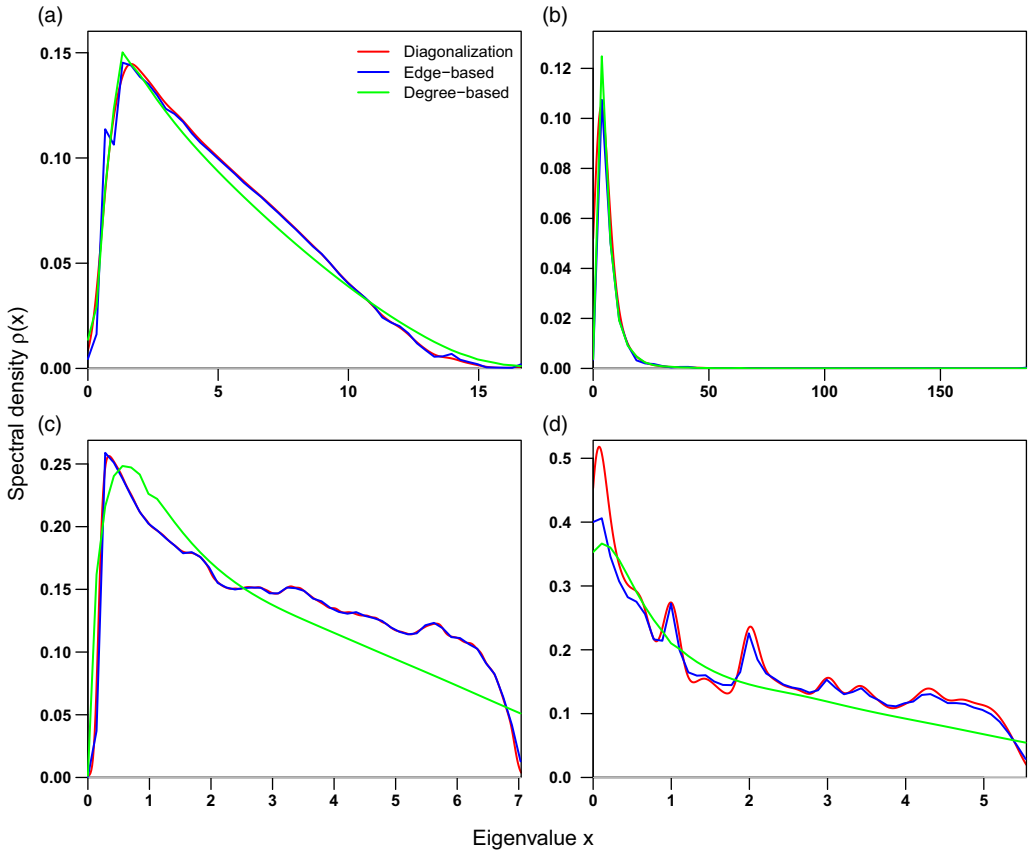

Figure 3. Laplacian spectral densities were obtained using the diagonalization (red), edge-based message-passing (blue), and degree-based approaches (green). (a) Poisson graph ($\lambda = 5$); (b) power-law graph generated by using the Barabási–Albert model ($\gamma = 1$ and $m = 3$); (c) Configuration model with degrees 2, 3, or 4 of equal probabilities; (d) configuration model with degrees 1, 2, or 3 of equal probabilities. We used the empirical degree distribution obtained from the network as an input of the degree-based approach. All graphs are composed of $n = 10,000$ nodes.

models (Newman, 2018). All three produce locally tree-like networks (Newman *et al.*, 2019). The Poisson model generates random graphs whose degree distribution follows a Poisson distribution with a specified mean λ. The configuration model creates random graphs with a given degree sequence. And the power-law model generates graphs whose degree distribution follows a power-law distribution (Adamic *et al.*, 2001; Adamic & Huberman, 2000). To generate power-law graphs, we used the Barabási–Albert model (Albert & Barabási, 2002).

Figures 3 and 4 show the spectra of the Laplacian and normalized Laplacian matrices of graphs generated with different random graph models. We produced four graphs with $n = 10,000$ and the following setup (similar to the setup presented in (Newman *et al.*, 2019)):

- Poisson graph with mean equals to five (Figures 3(a) and 4(a)).
- Power-law graph generated by using the Barabási–Albert model with parameters $\gamma = 1$ and $m = 3$ (Figures 3(b) and 4(b)).
- Configuration graph whose nodes have degrees 2, 3, or 4 with equal probabilities (Figures 3(c) and 4(c)).
- Configuration graph whose nodes have degrees 1, 2, or 3 with equal probabilities (Figures 3(d) and 4(d)).

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

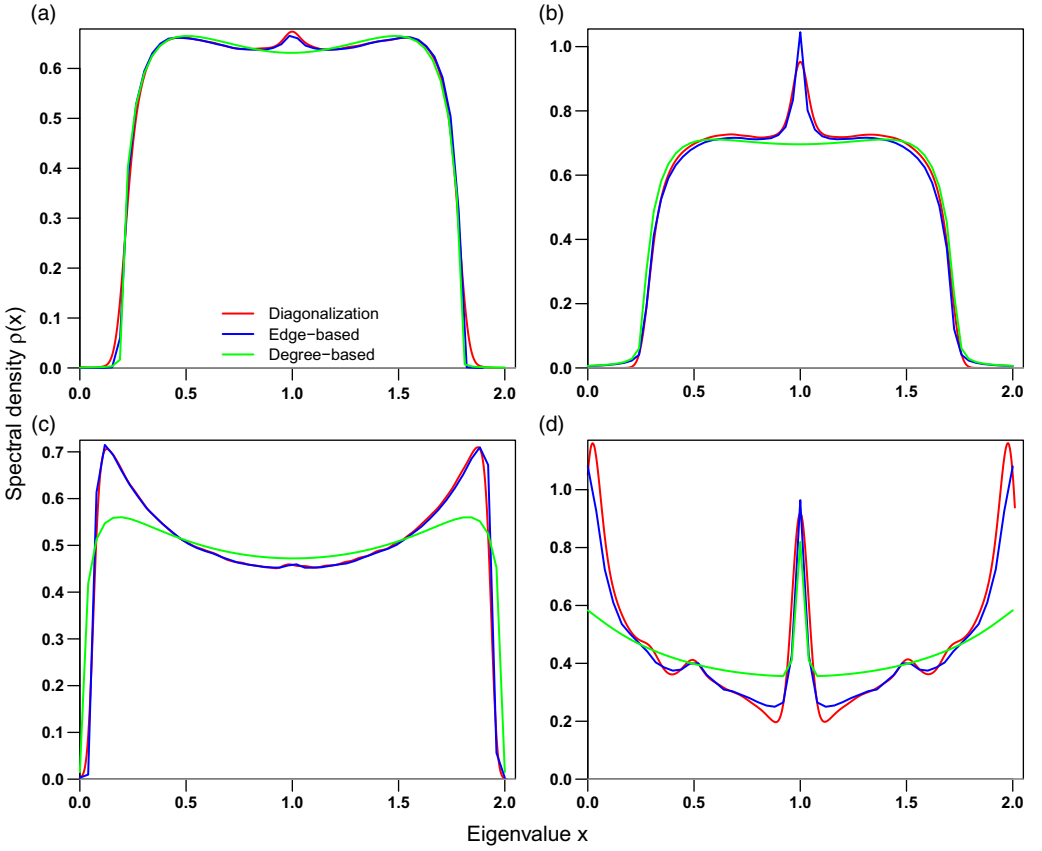

Figure 4. Normalized Laplacian spectral densities were obtained by using the diagonalization (red), edge-based message–passing (blue), and degree-based approaches (green). (a) Poisson graph (mean five); (b) power-law graph generated by using the Barabási–Albert model ($\gamma = 1$ and $m = 3$); (c) configuration model with degrees 2, 3, or 4 of equal probabilities; (d) configuration model with degrees 1, 2, or 3 of equal probabilities. We used the empirical degree distribution obtained from the network as an input of the degree-based approach. All graphs are composed of $n = 10,000$ nodes.

Notice that the edge-based method obtains good results for all models. The slight differences between the diagonalization and edge-based approaches are due to the smoothing parameters. The degree-based approximation presents a reduced performance. It happens because we approximate $g^u(z)$ (Equations (17) and 31) by $g_k(z)$ (Equations (25) and 38) (see Sections 2.2.1 and 2.2.2).

In Figure 4 panel (a), we would expect a semicircular spectral density (Jiang, 2012) because the Poisson graph is equal to the Erdös–Rényi graph where each edge is present with probability $p = \lambda/n$. However, we observe a non-semicircular shape and a peak at $x = 1$. It happens because the network is sparse (Newman *et al.*, 2019).

Figure 5 shows the spectra of the Laplacian matrix of graphs with different sizes ($n = 5,000$, $10,000$, $20,000$, and $40,000$) obtained by the diagonalization, edge-based, and degree-based approaches. We generated the graphs using the configuration model. Half of the nodes have degree 5, and the other half have degree 10. Notice that the estimations are very similar. The slight differences between the diagonalization and edge-based methods are due to the smoothing parameters used to estimate the spectral densities.

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

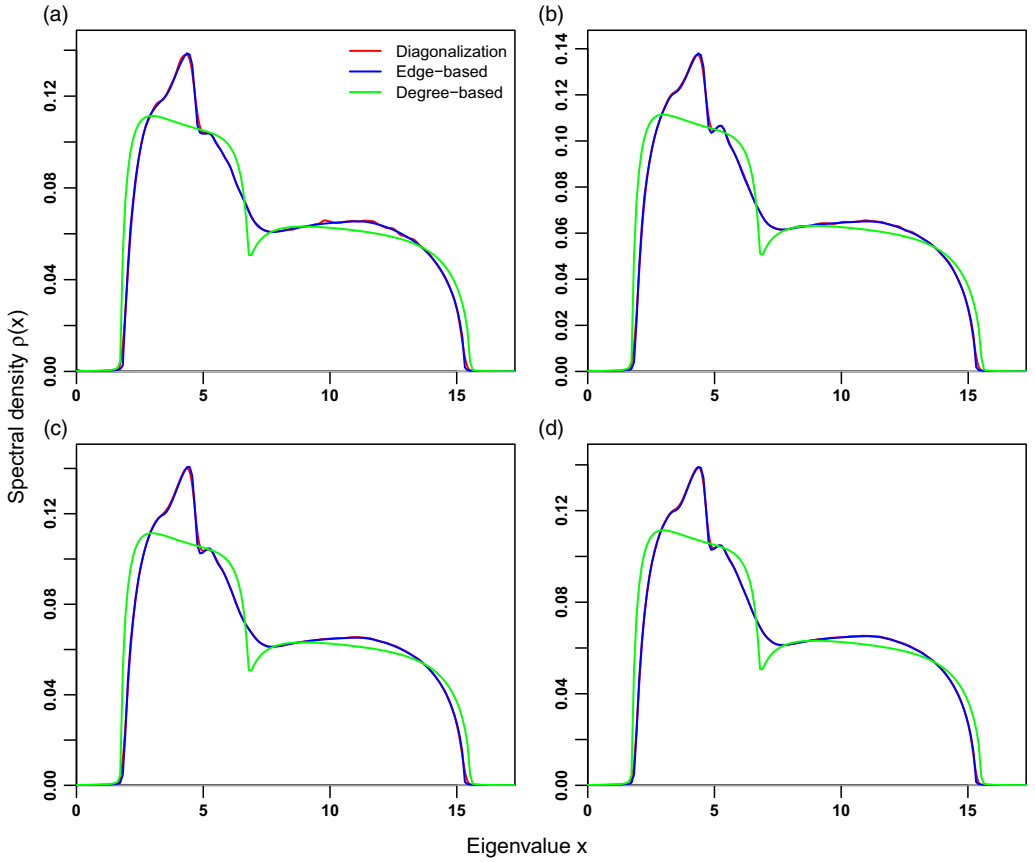

Figure 5. Laplacian spectral densities of graphs were obtained using diagonalization (red), edge-based message-passing (blue), and degree-based approaches. We generated the graphs using the configuration model. Half of the nodes have degree 5, and the other half have degree 10. (a) Laplacian spectral density for $n = 5,000$, (b) Laplacian spectral density for $n = 10,000$, (c) Laplacian spectral density for $n = 20,000$, and (d) Laplacian spectral density for $n = 40,000$. The slight difference between the diagonalization and edge-based methods is due to the smoothness parameter. The reduced performance of the degree-based approximation is because we approximate $g^u(z)$ by $g_k(z)$.

Figure 6 shows the Laplacian spectral density of weighted graphs. We sampled the edge weights from a Gaussian distribution with parameters mean zero and unit variance (Figure 6(a)), an uniform distribution (Figure 6(b)), an exponential distribution with mean one (Figure 6(c)), and Poisson distribution with parameter $\lambda = 2$ (Figure 6(d)). In both cases, when the edge weights are homogeneous or concentrated, the edge-based method presented good spectral densities estimations. It happens because we do not restrict the values that $w_{u,v}$ can take. We only require $\mathbf{M}$ to be symmetric. We did not apply the degree-based method here because this approach only works for unweighted locally tree-like graphs, as mentioned in Section 2.2.

Here, we presented two efficient methods to estimate the (normalized) Laplacian spectral density of big tree-like networks: an edge-based message-passing and a degree-based approach. As the number of big networks increases (e.g., social networks), we believe that more and more efficient (almost linear) methods become a must for network data analysis. As future work, we will modify our edge-based message-passing equations to contemplate weighted asymmetric matrices.

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

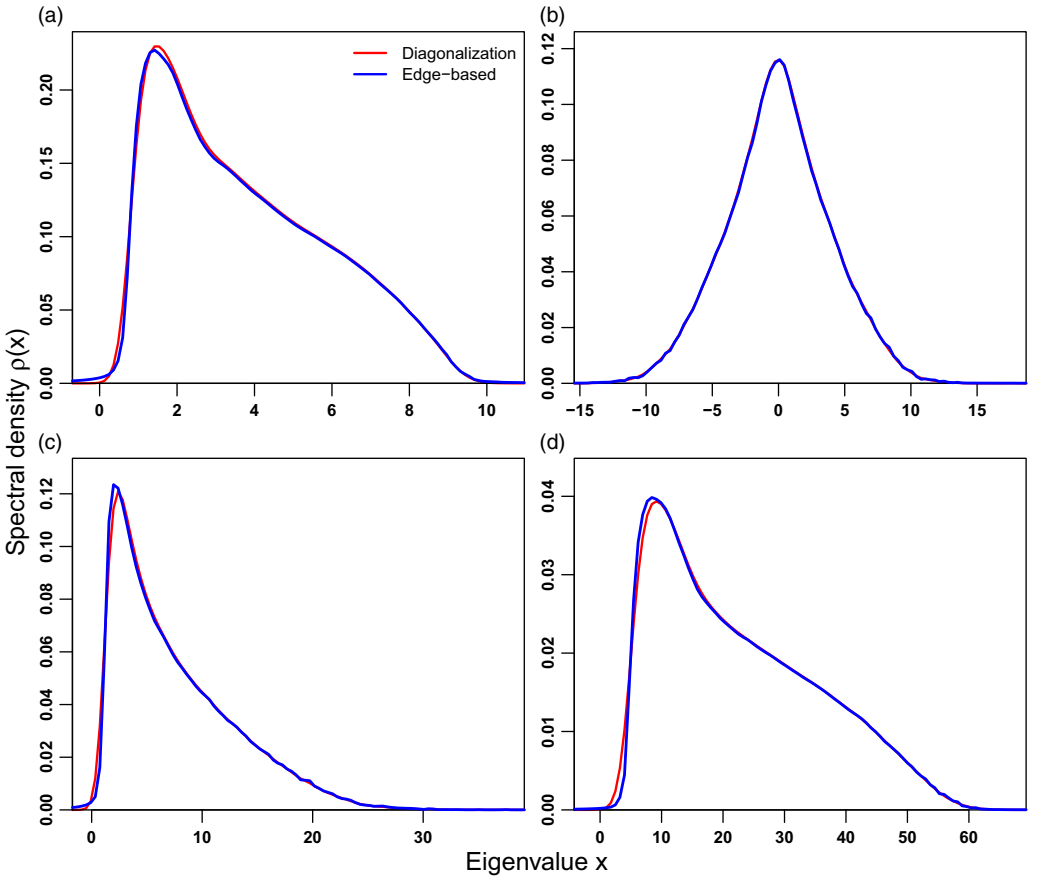

Figure 6. Laplacian spectral densities of weighted graphs were obtained using the diagonalization (red) and the edge-based methods (blue). We did not apply the degree-based method because this approach requires unweighted locally tree-like graphs. We generated the graphs using the configuration model with degrees 5 and 10 of equal probabilities. We sampled the edge weights from (a) a Gaussian distribution with mean 0 and unit variance, (b) an uniform distribution, (c) an exponential distribution with mean 1, and (d) a Poisson distribution with mean 2. All graphs are composed of $n = 10,000$ nodes.

Funding. G.E.C.G. was partially supported by CAPES (Finance code 001). P.F.S. was partially supported by Alexander von Humboldt Foundation. A.F. was partially supported by FAPESP (2018/21934-5, 2019-22845-9, 2020-08343-8), CNPq (303855/2019-3), CAPES, Alexander von Humboldt Foundation, The Academy of Medical Sciences - Newton Fund.

Conflicts of interest. None.

Supplementary materials. For supplementary material for this article, please visit http://dx.doi.org/10.1017/nws.2021.10

References

Adamic, L. A., & Huberman, B. A. (2000). Power-law distribution of the world wide web. *Science*, *287*(5461), 2115–2115.

Adamic, Lada A, Lukose, Rajan M, Puniyani, Amit R, & Huberman, Bernardo A. (2001). Search in power-law networks. *Physical Review E*, *64*(4), 046135.

Adcock, A. B., Sullivan, B. D., & Mahoney, M. W. (2013). Tree-like structure in large social and information networks. In*2013 IEEE 13th international conference on data mining* (pp. 1–10). IEEE.

Agliari, E., & Tavani, F.. (2017). The exact Laplacian spectrum for the Dyson hierarchical network. *Scientific Reports*, *7*(1), 1–21.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47.

Almendral, J. A., & Díaz-Guilera, A. (2007). Dynamical and spectral properties of complex networks. *New Journal of Physics*, *9*(6), 187.
Barahona, M., & Pecora, L. M. (2002). Synchronization in small-world systems. *Physical Review Letters*, *89*(5), 054101.
Bollobás, B. (1998). Random graphs. In*Modern graph theory* (pp. 215–252). Springer.
Booth, T. E. (2006). Power iteration method for the several largest eigenvalues and eigenfunctions. *Nuclear Science and Engineering*, *154*(1), 48–62.
Cantwell, G. T., & Newman, M. E. J. (2019). Message passing on networks with loops. *Proceedings of the National Academy of Sciences*, *116*(47), 23398–23403.
Chen, H., & Zhang, F. (2007). Resistance distance and the normalized Laplacian spectrum. *Discrete Applied Mathematics*, *155*(5), 654–661.
Chen, H., & Jost, J. (2012). Minimum vertex covers and the spectrum of the normalized Laplacian on trees. *Linear Algebra and Its Applications*, *437*(4), 1089–1101.
Chung, F., Lu, L., & Vu, V. (2004). The spectra of random graphs with given expected degrees. *Internet Mathematics*, *1*(3), 257–275.
Chung, F. R. K. (1989). Diameters and eigenvalues. *Journal of the American Mathematical Society*, *2*(2), 187–196.
de Lange, S., de Reus, M., & Van Den Heuvel, M. (2014). The Laplacian spectrum of neural networks. *Frontiers in Computational Neuroscience*, *7*, 189.
Ding, Q., Sun, W., & Chen, F. (2014). Applications of Laplacian spectra on a 3-prism graph. *Modern Physics Letters B*, *28*(02), 1450009.
Gander, W. (1980). Algorithms for the QR decomposition. *Research Report*, *80*(02), 1251–1268.
Granziol, D., Ru, B., Zohren, S., Dong, X., Osborne, M., & Roberts, S. (2018). Entropic spectral learning for large-scale graphs. *arxiv preprint arxiv:1804.06802*.
Hakimi-Nezhaad, M., & Ashrafi, A. R. (2014). A note on normalized Laplacian energy of graphs. *Journal of Contemporary Mathematical Analysis*, *49*(5), 207–211.
Hata, S., & Nakao, H. (2017). Localization of Laplacian eigenvectors on random networks. *Scientific Reports*, *7*(1), 1–11.
Huang, J., & Li, S. (2018). The normalized Laplacians on both k-triangle graph and k-quadrilateral graph with their applications. *Applied Mathematics and Computation*, *320*, 213–225.
Jamakovic, A., & Van Mieghem, P. (2006). The Laplacian spectrum of complex networks. *Proceedings of the european conference on complex systems, Oxford, September 25-29, 2006*.
Jiang, T. (2012). Empirical distributions of Laplacian matrices of large dilute random graphs. *Random Matrices: Theory and Applications*, *1*(03), 1250004.
Julaiti, A., Wu, B., & Zhang, Z. (2013). Eigenvalues of normalized Laplacian matrices of fractal trees and dendrimers: Analytical results and applications. *The Journal of Chemical Physics*, *138*(20), 204116.
Klein, D. J., & Randić, M. (1993). Resistance distance. *Journal of Mathematical Chemistry*, *12*(1), 81–95.
Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.
Kowalska, E. (1987). On locally tree-like graphs. *Applicationes Mathematicae*, *19*(3–4), 497–503.
Li, D., & Hou, Y. (2017). The normalized Laplacian spectrum of quadrilateral graphs and its applications. *Applied Mathematics and Computation*, *297*, 180–188.
Li, S., Yan, W., & Tian, T. (2016). The spectrum and Laplacian spectrum of the dice lattice. *Journal of Statistical Physics*, *164*(2), 449–462.
Liu, H., & Zhang, Z. (2013). Laplacian spectra of recursive treelike small-world polymer networks: Analytical solutions and applications. *The Journal of Chemical Physics*, *138*(11), 114904.
Liu, H., Dolgushev, M., Qi, Y., & Zhang, Z. (2015). Laplacian spectra of a class of small-world networks and their applications. *Scientific Reports*, *5*, 9024.
Ma, X., & Bian, H.. (2019). The normalized Laplacians, degree-kirchhoff index and the spanning trees of hexagonal möbius graphs. *Applied Mathematics and Computation*, *355*, 33–46.
Metz, F. L., Neri, I., & Bollé, D. (2011). Spectra of sparse regular graphs with loops. *Physical Review E*, *84*(5), 055101.
Mülken, O., Volta, A., & Blumen, A. (2005). Asymmetries in symmetric quantum walks on two-dimensional networks. *Physical Review A*, *72*(4), 042334.
Nadakuditi, R. R., & Newman, M. E. J. (2013). Spectra of random graphs with arbitrary expected degrees. *Physical Review E*, *87*(1), 012803.
Nadler, B., Srebro, N., & Zhou, X. (2009). Statistical analysis of semi-supervised learning: The limit of infinite unlabelled data. *Advances in neural information processing systems* (pp. 1330–1338).
Newman, M. (2018). *Networks*. Oxford university press.
Newman, M. E. J. (2019). Spectra of networks containing short loops. *Physical Review E*, *100*(1), 012314.
Newman, M. E. J., Zhang, X., & Nadakuditi, R. R. (2019). Spectra of random networks with arbitrary degrees. *Physical Review E*, *99*(4), 042309.
Pan, Y., Li, J., Li, S., & Luo, W. (2018). On the normalized Laplacians with some classical parameters involving graph transformations. *Linear and Multilinear Algebra*, 1–23.

https://doi.org/10.1017/nws.2021.10 Published online by Cambridge University Press

Peixoto, T. P. (2013). Eigenvalue spectra of modular networks. *Physical Review Letters*, *111*(9), 098701.
Rogers, T., Castillo, I. P., Kühn, R., & Takeda, K. (2008). Cavity approach to the spectral density of sparse symmetric random matrices. *Physical Review E*, *78*(3), 031116.
Semerjian, G., & Cugliandolo, L. F. (2002). Sparse random matrices: The eigenvalue spectrum revisited. *Journal of Physics A: Mathematical and General*, *35*(23), 4837.
Sleijpen, G. L. G., & Van der Vorst, H. A. (2000). A Jacobi–Davidson iteration method for linear eigenvalue problems. *SIAM Review*, *42*(2), 267–293.
Sun, S., & Das, K. Ch. (2020). Normalized Laplacian eigenvalues with chromatic number and independence number of graphs. *Linear and Multilinear Algebra*, *68*(1), 63–80.
Sun, W., Xuan, T., & Qin, S. (2016). Laplacian spectrum of a family of recursive trees and its applications in network coherence. *Journal of Statistical Mechanics: Theory and Experiment*, *2016*(6), 063205.
Sun, Y., Dai, M., Shao, S., & Su, W. (2017). The entire mean weighted first-passage time on infinite families of weighted tree networks. *Modern Physics Letters B*, *31*(07), 1750049.
Trinajstic, N., Babic, D., Nikolic, S., Plavsic, D., Amic, D., & Mihalic, Z. (1994). The Laplacian matrix in chemistry. *Journal of Chemical Information and Computer Sciences*, *34*(2), 368–376.
Wegner, A. E., Ospina-Forero, L., Gaunt, R. E., Deane, C. M., & Reinert, G. (2018). Identifying networks with common organizational principles. *Journal of Complex Networks*, *6*(6), 887–913.
Wu, F.-Y. (2004). Theory of resistor networks: The two-point resistance. *Journal of Physics A: Mathematical and General*, *37*(26), 6653.
Zhang, Z., Zhou, S., Qi, Y., & Guan, J. (2008). Topologies and Laplacian spectra of a deterministic uniform recursive tree. *The European Physical Journal B*, *63*(4), 507–513.

Cite this article: Guzman G. E. C., Stadler P. F. and Fujita A. (2021). Efficient Laplacian spectral density computations for networks with arbitrary degree distributions. *Network Science* **9**, 312–327. https://doi.org/10.1017/nws.2021.10