



**HAL**  
open science

# On the Normalized Laplacian Spectra of Random Geometric Graphs

Mounia Hamidouche, Laura Cottatellucci, Konstantin E Avrachenkov

► **To cite this version:**

Mounia Hamidouche, Laura Cottatellucci, Konstantin E Avrachenkov. On the Normalized Laplacian Spectra of Random Geometric Graphs. *Journal of Theoretical Probability*, 2022, 10.1007/s10959-022-01158-0 . hal-03582734v1

**HAL Id: hal-03582734**

**<https://inria.hal.science/hal-03582734v1>**

Submitted on 21 Feb 2022 (v1), last revised 4 Jan 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# On the Normalized Laplacian Spectra of Random Geometric Graphs

Mounia Hamidouche<sup>1</sup> · Laura Cottatellucci<sup>2</sup>  · Konstantin Avrachenkov<sup>3</sup>

Received: 15 June 2019 / Revised: 26 December 2021 / Accepted: 5 January 2022

© The Author(s) 2022

## Abstract

In this work, we study the spectrum of the normalized Laplacian and its regularized version for random geometric graphs (RGGs) in various scaling regimes. Two scaling regimes are of special interest, the connectivity and the thermodynamic regime. In the connectivity regime, the average vertex degree grows logarithmically in the graph size or faster. In the thermodynamic regime, the average vertex degree is a constant. We introduce a deterministic geometric graph (DGG) with nodes in a grid and provide an upper bound to the probability that the Hilbert–Schmidt norm of the difference between the normalized Laplacian matrices of the RGG and DGG is greater than a certain threshold in both the connectivity and thermodynamic regime. Using this result, we show that the RGG and DGG normalized Laplacian matrices are asymptotically equivalent with high probability (w.h.p.) in the full range of the connectivity regime. The equivalence is even stronger and holds almost surely when the average vertex degree  $a_n$  satisfies the inequality  $a_n > 24 \log(n)$ . Therefore, we use the regular structure of the DGG to show that the limiting eigenvalue distribution of the RGG normalized Laplacian matrix converges to a distribution with a Dirac atomic measure at zero. In the thermodynamic regime, we approximate the eigenvalues of the regularized normalized Laplacian matrix of the RGG by the eigenvalues of the DGG regularized normalized Laplacian and we provide an error bound which is valid w.h.p. and depends upon the average vertex degree.

---

✉ Laura Cottatellucci  
laura.cottatellucci@fau.de

Mounia Hamidouche  
mounia.hamidouche@eurecom.fr

Konstantin Avrachenkov  
k.avrachenkov@inria.fr

<sup>1</sup> Departement of communication systems, EURECOM, Campus SophiaTech, 06410 Biot, France

<sup>2</sup> Department of Electrical, Electronics, and Communication Engineering, FAU, Cauerstraße 7, 91058 Erlangen, Germany

<sup>3</sup> Inria, 2004 Route des Lucioles, 06902 Biot, France

**Keywords** Random geometric graph · Normalized Laplacian · Limiting eigenvalue distribution · Connectivity regime · Thermodynamic regime

**Mathematics Subject Classification 2020** 60B20 random matrices · 05C80 random graphs

## 1 Introduction

The spectra of random matrices and random graphs have been extensively studied in the literature [1–5]. Spectral graph methods have become a fundamental tool in the analysis of large complex networks, and related disciplines, with a broad range of applications in telecommunication, machine learning, data mining, and web search. A first natural random graph model of complex networks is the Erdős-Rényi (ER) random graph [6] where edges between nodes appear with equal probabilities. This model has many appealing analytical properties but does not model important features of many real complex networks. In particular, the ER graph fails in describing clustering properties of graphs in which the geographical distance is a critical factor, as for example wireless ad-hoc network [7], sensor network [8], and the study of the dynamics of a viral spreading in a specific network of interactions [9], [10]. To properly model such networks, we consider a special class of graphs known as random geometric graphs (RGGs) [11]. Another very important motivation for the study of RGGs is their applications to statistics and learning. Many clustering techniques such as the nearest-neighbor technique in statistics and machine learning are based on the spatial structure of RGGs [12].

Let us analytically define the RGG studied in this work. We consider a finite set  $\mathcal{X}_n$  of  $n$  nodes,<sup>1</sup>  $x_1, \dots, x_n$ , distributed uniformly and independently on the  $d$ -dimensional torus  $\mathbb{T}^d \equiv [0, 1]^d$ . Taking a torus  $\mathbb{T}^d$  instead of a cube exempts us from considering boundary effects. Given a threshold  $r_n > 0$ , we form a graph by connecting two nodes  $x_i, x_j \in \mathcal{X}_n$  if the  $\ell_p$ -distance between them is at most  $r_n$ , i.e.,  $\|x_i - x_j\|_p \leq r_n$  with  $p \in [1, \infty]$  (that is, either  $p \in [1, \infty)$  or  $p = \infty$ , see Figure 1(a). Here  $\|\cdot\|_p$  is the  $\ell_p$ -distance on  $\mathbb{R}^d$  defined as

$$\|x_i - x_j\|_p = \begin{cases} \left( \sum_{k=1}^d |x_i^{(k)} - x_j^{(k)}|^p \right)^{1/p} & \text{for } p \in [1, \infty), \\ \max\{|x_i^{(k)} - x_j^{(k)}|, 1 \leq k \leq d\} & \text{for } p = \infty. \end{cases}$$

For  $p = 2$ , the  $\ell_2$ -distance defines the standard Euclidean distance between two points in  $\mathbb{R}^d$ . When  $p = \infty$ , we obtain the Chebyshev distance, i.e., the maximum of the differences between the coordinates in any dimension of the two points. Such graphs, denoted by  $G(\mathcal{X}_n, r_n)$ , are called RGGs and are extensively discussed in [11]. Typically, the function  $r_n$  is chosen such that  $r_n \rightarrow 0$  when  $n \rightarrow \infty$ . Unlike ER graphs, the RGG is an inherently harder model to work with since the nature of the graph induces dependencies between edges.

<sup>1</sup> Throughout this work, the nodes of a graph are denoted by their spatial coordinates.

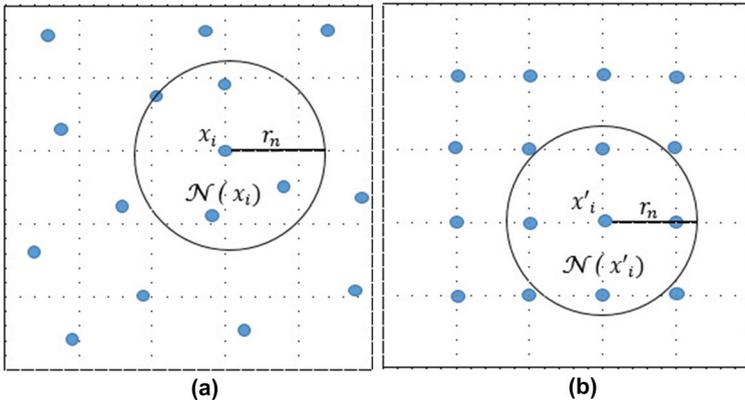


Fig. 1 Illustration of an RGG (a) and a (DGG) (b) for  $n = 16$

The degree of a vertex in  $G(\mathcal{X}_n, r_n)$  is the number of edges incident to it. Let  $\theta^{(d)}$  denote the volume of the  $d$ -dimensional unit hypersphere in  $\mathbb{T}^d$ . Then, the average vertex degree in  $G(\mathcal{X}_n, r_n)$  is equal to  $a_n = \theta^{(d)} n r_n^d$  [11]. The properties of RGGs, including spectral properties, often depend on the average vertex degree  $a_n$ . Two different scaling regimes for  $a_n$  are of interest in this work. The first one is the connectivity regime in which the average vertex degree  $a_n$  grows logarithmically in  $n$  or faster, i.e.,  $a_n = \Omega(\log(n))^2$  [11]. The second scaling regime of interest is the thermodynamic regime, in which the average vertex degree is a constant  $\gamma$ , i.e.,  $a_n \equiv \gamma$  [11]. This regime models real-world complex networks that consist of disconnected components, e.g., social interaction networks [13] and web graphs [14].

This work focuses on the analysis of the spectra of the normalized Laplacian matrix of  $G(\mathcal{X}_n, r_n)$  in the connectivity and thermodynamic regimes under any  $\ell_p$ -distance as  $d$  remains fixed and  $n \rightarrow \infty$ . To this end, we introduce the set  $\mathcal{D}_n$  of the  $n$  nodes that are at the intersections of all parallel hyperplanes with separation  $n^{-1/d}$ , see Fig. 1(b), and define a deterministic graph  $G(\mathcal{D}_n, r_n)$  by connecting any two nodes  $x'_i, x'_j \in \mathcal{D}_n$  if  $\|x'_i - x'_j\|_p \leq r_n$  for  $p \in [1, \infty]$ . For both  $G(\mathcal{X}_n, r_n)$  and  $G(\mathcal{D}_n, r_n)$ , we assume that there is at most one edge between any pair of nodes. There is no edge from a vertex to itself. Moreover, we assume that the edges are not directed.

In the following, we define the normalized Laplacian matrix for  $G(\mathcal{X}_n, r_n)$  and  $G(\mathcal{D}_n, r_n)$ .

Let  $\mathcal{N}(x_i)$  be the set of neighbors of vertex  $x_i$  in  $G(\mathcal{X}_n, r_n)$  and  $\mathcal{N}(x'_i)$  be the set of neighbors of vertex  $x'_i$  in  $G(\mathcal{D}_n, r_n)$ . The normalized Laplacian matrices  $\mathcal{L}(\mathcal{X}_n)$  and  $\mathcal{L}(\mathcal{D}_n)$  for  $G(\mathcal{X}_n, r_n)$  and  $G(\mathcal{D}_n, r_n)$ , respectively, have entries defined by

$$\mathcal{L}(\mathcal{X}_n)_{ij} = \delta_{ij} - \frac{\chi[x_i \sim x_j]}{\sqrt{\mathbf{N}(x_i)\mathbf{N}(x_j)}}, \quad \mathcal{L}(\mathcal{D}_n)_{ij} = \delta_{ij} - \frac{\chi[x'_i \sim x'_j]}{\sqrt{\mathbf{N}(x'_i)\mathbf{N}(x'_j)}}, \quad (1)$$

<sup>2</sup> The notation  $f(n) = \Omega(g(n))$  indicates that  $f(n)$  is bounded below by  $g(n)$  asymptotically, i.e.,  $\exists K > 0$  and  $n_0 \in \mathbb{N}$  such that  $\forall n > n_0 \ f(n) \geq K g(n)$ .

where  $\mathbf{N}(x_i)$  and  $\mathbf{N}(x'_i)$  are the sizes of the two sets  $\mathcal{N}(x_i)$  and  $\mathcal{N}(x'_i)$ , respectively, and  $\delta_{ij}$  is the Kronecker delta function. The term  $\chi[x_i \sim x_j]$  takes unit value when there is an edge between nodes  $x_i$  and  $x_j$  in  $G(\mathcal{X}_n, r_n)$  and zero otherwise, i.e.,

$$\chi[x_i \sim x_j] = \begin{cases} 1, & \|x_i - x_j\|_p \leq r_n, \quad i \neq j \\ 0, & \text{otherwise.} \end{cases}$$

A similar definition holds for  $\chi[x'_i \sim x'_j]$  defined over the nodes in  $G(\mathcal{D}_n, r_n)$ . Recall that  $a_n$  denotes the average vertex degree in  $G(\mathcal{X}_n, r_n)$  and, is constant and independent of  $n$  in the thermodynamic regime, i.e.,  $a_n \equiv \gamma > 0$ . In  $G(\mathcal{D}_n, r_n)$ , the number of neighbors of each vertex is the same and we denote it by  $a'_n = \mathbf{N}(x'_i)$ . In particular, in the thermodynamic regime  $\mathbf{N}(x'_i) = \gamma'$ . Additionally,  $\mathbf{N}(x_i, x'_i) = \sum_j \chi[x_i \sim x_j]\chi[x'_i \sim x'_j] \leq a'_n \forall i, j$ .

Note that the above formal definition of the normalized Laplacian in (1) requires  $G(\mathcal{X}_n, r_n)$  and  $G(\mathcal{D}_n, r_n)$  not to have isolated vertices. To overcome the problem of singularities due to isolated vertices in the thermodynamic regime, we follow the scheme proposed in [15]. It corresponds to the normalized Laplacian matrix on a modified graph constructed by adding auxiliary edges among all the nodes with weight  $\frac{\alpha}{n} > 0$ . Specifically, the entries of the normalized Laplacian matrices are modified as

$$\hat{\mathcal{L}}(\mathcal{X}_n)_{ij} = \delta_{ij} - \frac{\chi[x_i \sim x_j] + \frac{\alpha}{n}}{\sqrt{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)}}, \quad \hat{\mathcal{L}}(\mathcal{D}_n)_{ij} = \delta_{ij} - \frac{\chi[x'_i \sim x'_j] + \frac{\alpha}{n}}{\sqrt{(a'_n + \alpha)(a'_n + \alpha)}}. \quad (2)$$

The corresponding matrices are referred to as regularized normalized Laplacian matrices [15]. Observe that for  $\alpha = 0$ , (2) reduces to (1). The matrices  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$  are symmetric, and consequently, their spectra consist of real eigenvalues. We denote by  $\{\hat{\mu}_i, i = 1, \dots, n\}$  and  $\{\hat{\lambda}_i, i = 1, \dots, n\}$  the sets of all real eigenvalues of the real symmetric square matrices  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$  of order  $n$ , respectively. Then, the empirical spectral distribution functions of  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$  are defined as

$$F^{\hat{\mathcal{L}}(\mathcal{X}_n)}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{\hat{\mu}_i \leq x\}, \quad \text{and} \quad F^{\hat{\mathcal{L}}(\mathcal{D}_n)}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{\hat{\lambda}_i \leq x\},$$

where  $\mathbb{I}\{B\}$  denotes the indicator of an event  $B$ . The goal of this work is to show that  $F^{\hat{\mathcal{L}}(\mathcal{D}_n)}$  is a good approximation for  $F^{\hat{\mathcal{L}}(\mathcal{X}_n)}$  when  $n$  is large in both the connectivity and thermodynamic regimes. For that, we use a methodology that relates the asymptotic behavior of two sequences of matrices.

The rest of this paper is organized as follows. In Sect. 2, we provide an overview of related work. In Sect. 3, we present the main results of this work. In Sect. 4, we show how the results presented in Sect. 3 on the relation between the spectrum of RGGs and DGGs are obtained, and how to use the DGG structure to provide analytical approximations of the RGG eigenvalues. Numerical results are given in Sect. 5 to validate the theoretical results by comparing the limiting eigenvalue dis-

tribution (LED) obtained analytically and by simulation. Finally, conclusions and implications are drawn in Sect. 6.

## 2 Related Work

The spectral properties of random matrices are fundamental tools to predict and analyze complex network behavior. The work in [16] investigates the combinatorial Laplacian spectra of RGGs and shows that the spectrum consists of both discrete and continuous parts. The discrete part of the spectrum is a collection of Dirac atomic measures at integer values. The work in [17] shows that the Dirac atomic measure appear mainly due to the existence of symmetric motifs<sup>3</sup> that occur abundantly in RGGs compared to ER graphs. Other works also analyzed the symmetric motifs in RGGs and studied the probabilities of their appearance (see e.g., [18]).

Several works analyzed the spectra of Euclidean random matrices given by  $H_n = f(\|x_i - x_j\|_2)$  when  $n$  and  $d$  grows large and function  $f(\cdot)$  belongs to a given class [19], [20]. However, the obtained results cannot be applied to our problem as we assume that the dimension  $d$  stays fixed. The work in [20] also studies the LED of Euclidean random matrices  $H_n$  when the dimension  $d$  remains fixed. It shows that under some conditions the LED of  $H_n$  converges almost surely to a Dirac atomic measure at zero. However, the results in [20] require the continuity of function  $f$  and they cannot be applied to the step function considered in this work.

Regarding spectral properties of the adjacency matrix of (RGGs), in [21] and [22], the authors show that the spectral distribution of the adjacency matrix has a limit in the thermodynamic regime as  $n \rightarrow \infty$ . Due to the difficulty to compute exactly this spectral measure, in [21], the author proposes an approximation for it as  $\gamma \rightarrow \infty$ . In [9], a closed form expression for the asymptotic spectral moments of the adjacency matrix of  $G(\mathcal{X}_n, r_n)$  is derived in the connectivity regime. Then, an analytical upper bound for the spectral radius is derived in order to study the behavior of the viral infection in an RGG. Furthermore, in [23], the author shows that in the connectivity regime, the spectral measures of the transition probability matrix of the random walk in an RGG and in a DGG with nodes in a grid converge to the same limit as  $n \rightarrow \infty$ . However, the author in [23] does not study the full range of the connectivity regime and in the proof of this result, a condition is enforced on the average vertex degree  $a_n$  using the concept of the minimum bottleneck matching distance. The condition enforced on  $a_n$  in [23] implies that for  $\epsilon > 0$ , the result holds only for RGGs with an average vertex degree  $a_n$  that scales as  $\Omega(\log^\epsilon(n)\sqrt{n})$ ,  $\Omega(\log^{\frac{3}{2}+\epsilon}(n))$ ,  $\Omega(\log^{1+\epsilon}(n))$  for  $d = 1$ ,  $d = 2$  and  $d \geq 3$ , respectively.

To the best of our knowledge, explicit expressions for the LED or the eigenvalues of the combinatorial Laplacian and normalized Laplacian for RGGs are still not known in the full range of the scaling laws for the average vertex degree  $a_n$  in the connectivity regime, nor in the thermodynamic regime. Compared to [23], in this work we study the spectrum of  $\mathcal{L}(\mathcal{X}_n)$  in the connectivity regime for a wider range of scaling laws of the average vertex degree  $a_n$ , or equivalently a wider range of scaling laws of the

<sup>3</sup> Symmetric motifs are subsets of nodes which have the same adjacencies.

radius  $r_n^d$  and under any  $\ell_p$ -distance,  $p \in [1, \infty]$ . More specifically, for  $d \geq 1$ , we show that the LEDs of the normalized Laplacian  $\mathcal{L}(\mathcal{X}_n)$  for RGGs and for DGGs converge to the same limit when  $a_n = \Omega(\log(n))$ . Moreover, we study the spectrum of  $\hat{\mathcal{L}}(\mathcal{X}_n)$  in the thermodynamic regime. To overcome the problem of singularities due to isolated nodes in the thermodynamic regime, we investigate the normalized Laplacian on a graph modified by adding auxiliary edges among all the nodes with a certain weight. The corresponding normalized Laplacian is known as the regularized normalized Laplacian [15].

### 3 Main Results

In this section we present the main results of our work for both the connectivity and thermodynamic regimes. To this end, we introduce the Hilbert–Schmidt norm, also known as weak norm, and defined as

$$\|\mathbf{A}\|_{\text{HS}} = \frac{1}{n} \text{trace}(\mathbf{A}^\top \mathbf{A}).$$

The following Theorem 1 summarizes our main result in the connectivity regime. It shows that the spectrum of the RGG normalized Laplacian matrix converges in Hilbert–Schmidt norm almost surely (a.s.) to the spectrum of the DGG normalized Laplacian matrix in the connectivity regime as  $a_n > 24 \log(n)$ , i.e.,

$$\|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}} \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{a.s. for } a_n > 24 \log(n), \quad (3)$$

whereas the convergence holds w.h.p. for  $a_n = c \log(n)$  and  $c \leq 24$ .

**Theorem 1** *In the connectivity regime, i.e., for  $a_n = \Omega(\log(n))$ , the spectrum of the RGG normalized Laplacian matrix  $\mathcal{L}(\mathcal{X}_n)$  converges to the spectrum of the DGG normalized Laplacian matrix  $\mathcal{L}(\mathcal{D}_n)$  with convergence rate depending on  $a_n$  as follows:<sup>4</sup>*

$$\mathbb{P} \left\{ \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} = \begin{cases} \mathcal{O}(1/n), & a_n = c \log(n) \text{ and } c \leq 24, \\ \mathcal{O}(1/n^{c/12-1}), & a_n = c \log(n) \text{ and } c > 24, \\ \mathcal{O}\left(1/n^{\left(\frac{a_n}{12 \log(n)}\right)^{-1}}\right), & a_n \geq \log^{1+\epsilon}(n) \text{ and} \\ & a_n = o(n) \text{ with } \epsilon > 0, \\ \mathcal{O}(ne^{-n/12}), & a_n = \Omega(n). \end{cases}$$

**Proof** See Appendix D. □

Theorem 1 shows that when  $a_n \geq \log^{1+\epsilon}(n)$  for  $\epsilon > 0$ , the rate of convergence of  $\mathbb{P} \left\{ \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\}$  is  $\mathcal{O}(1/n^{(a_n/12 \log(n))^{-1}})$ . In particular, it shows that, when  $a_n = c \log(n)$ , for  $c > 24$ , the convergence rate is  $\mathcal{O}(1/n^{c/12-1})$ , and when

<sup>4</sup> The notation  $f(x) = \mathcal{O}(g(x))$  as  $x \rightarrow \infty$  indicates that there exists a positive real number  $K$  and a real number  $x_0$  such that  $|f(x)| \leq Kg(x)$  for all  $x \geq x_0$ . Additionally,  $f(x) = o(g(x))$  as  $x \rightarrow \infty$  indicates that for every positive constant  $\epsilon$  there exists a constant  $x_0$  such that  $|f(x)| \leq \epsilon g(x)$  for all  $x \geq x_0$ .

$c \leq 24$ , a slower convergence rate holds which scales as  $\mathcal{O}(1/n)$ . When the graph is dense, i.e.,  $a_n = \Omega(n)$ , the convergence rate is  $\mathcal{O}(ne^{-n/12})$ . Notice that the almost sure convergence in (3) follows from Theorem 1 and the Borel-Cantelli Lemma, e.g., [24], for  $a_n > 24 \log(n)$ .

The following Theorem 2 presents our result for the thermodynamic regime. The spectrum of the regularized normalized Laplacian of an RGG can be asymptotically approximated in Hilbert–Schmidt norm by the spectrum of the regularized normalized Laplacian of a DGG. In particular, we show that the Hilbert–Schmidt norm of the error is bounded with high probability by an explicit positive constant that depends on the average vertex degree  $a_n = \gamma$ , i.e.,

$$\mathbb{P} \left\{ \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} \xrightarrow{n \rightarrow \infty} 0,$$

where  $t > 0$  is an explicit positive constant.

**Theorem 2** *In the thermodynamic regime, with  $\gamma \geq \frac{2d^{1+1/p}}{2d-1}$ ,  $d \geq 1$  and  $t > \frac{8(n+2\alpha)\gamma+4\alpha^2}{n(\gamma+\alpha)^2}$ , the following bound holds*

$$\mathbb{P} \left\{ \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} \leq \frac{320(n-1)\vartheta}{tn^2(\gamma+\alpha)^2},$$

where  $\vartheta = \theta^{(d)} \left[ 1 + 2\gamma - \frac{4\gamma}{n} \right]$ .

**Proof** See Appendix C. □

This result implies that for  $n \rightarrow \infty$  and for every  $t > \frac{8\gamma}{(\gamma+\alpha)^2}$ , the following limit holds

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} = 0.$$

Informally, this entails that for any chosen positive  $t$  satisfying the above inequality, as  $n \rightarrow \infty$ , the Hilbert–Schmidt norm  $\|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}$  is not greater than  $t$  with high probability. Therefore, in the thermodynamic regime, we can approximate  $F^{\hat{\mathcal{L}}(\mathcal{X}_n)}$  by  $F^{\hat{\mathcal{L}}(\mathcal{D}_n)}$  with an error bound of  $\frac{8}{\gamma}$  that holds with high probability when  $n \rightarrow \infty$  and  $\alpha \rightarrow 0$ . It is worth noting that this error bound, valid with high probability, is small when  $a_n = \gamma$  is large.

Based on Theorem 1, we can use the LED of  $\mathcal{L}(\mathcal{D}_n)$ , the normalized Laplacian matrix for a DGG defined by the Chebyshev distance to determine the LED of the normalized Laplacian in the connectivity regime as  $n \rightarrow \infty$ . Corollary 1 provides the expression of the eigenvalues of  $\mathcal{L}(\mathcal{D}_n)$ . Similarly, Theorem 2 indicates that the LED of  $\hat{\mathcal{L}}(\mathcal{D}_n)$  provides an analytical approximation of the LED of  $\hat{\mathcal{L}}(\mathcal{X}_n)$  in the thermodynamic regime as  $n$  goes to infinity. The expression of the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{D}_n)$  are derived in Corollary 2.

**Corollary 1** *In the connectivity regime, i.e.,  $a_n = \Omega(\log(n))$ , for RGGs defined by the Chebyshev distance,  $a_n \geq \frac{2d}{2d-1}$ ,  $d \geq 1$ , and letting  $\alpha \rightarrow 0$ , the eigenvalues of  $\mathcal{L}(\mathcal{D}_n)$  are given by*

$$\lambda_{m_1, \dots, m_d} = 1 - \frac{1}{a'_n} \prod_{s=1}^d \frac{\sin(\frac{m_s \pi}{N} (a'_n + 1)^{1/d})}{\sin(\frac{m_s \pi}{N})} + \frac{1}{a'_n}, \quad (4)$$

with  $m_1, \dots, m_d \in \{0, \dots, N-1\}$ . In (4),  $n = N^d$ ,  $a'_n = (2k_n + 1)^d - 1$  and  $k_n = \lfloor Nr_n \rfloor$ . Then, in particular, as  $n \rightarrow \infty$ , the LED of  $\mathcal{L}(\mathcal{D}_n)$  converges almost surely to the Dirac measure at one.

**Proof** See Appendix E. □

**Corollary 2** *In the thermodynamic regime, for RGGs defined by the Chebyshev distance,  $\gamma \geq \frac{2d}{2d-1}$ ,  $d \geq 1$ , as  $n \rightarrow \infty$ , the eigenvalues of  $\mathcal{L}(\mathcal{D}_n)$  are approximated by*

$$\hat{\lambda}_{f_1, \dots, f_d} = 1 - \frac{1}{(\gamma' + \alpha)} \prod_{s=1}^d \frac{\sin(\pi f_s (\gamma' + 1)^{1/d})}{\sin(\pi f_s)} + \frac{1 - \alpha \delta_{f_1, \dots, f_d}}{(\gamma' + \alpha)}, \quad (5)$$

where  $s \in \{1, \dots, d\}$ ,  $m_s \in \{0, \dots, N\}$  and  $f_s = \frac{m_s}{N}$  in  $\mathbb{Q} \cap [0, 1]$  with  $\mathbb{Q}$  denotes the set of rational numbers. Also,  $\gamma' = (2 \lfloor \gamma^{1/d} \rfloor + 1)^d - 1$  and  $\delta_{f_1, \dots, f_d} = 1$  when  $f_1, \dots, f_d = 0$  otherwise  $\delta_{f_1, \dots, f_d} = 0$ .

**Proof** See Appendix E. □

## 4 On the Relation Between the Spectra of Laplacians for RGGs and DGGs

In this section, we analyze the asymptotic behavior of the RGG normalized Laplacian matrix in the connectivity and thermodynamic regimes. In particular, we provide all intermediate results that led to our main results presented in Sect. 3. To this end, first we present a well-known methodology to reduce the asymptotic spectral analysis of RGG Laplacians to the asymptotic spectral analysis of DGG Laplacians by introducing the concept of *asymptotic equivalence of sequences of matrices*. Then, we derive a fundamental bound on  $\|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2$  to establish the asymptotic equivalence between the two matrix sequences  $\{\hat{\mathcal{L}}(\mathcal{X}_n)\}$  and  $\{\hat{\mathcal{L}}(\mathcal{D}_n)\}$ .

In the following we introduce the definition of asymptotic equivalent sequences of matrices and its implications on their eigenvalue spectra.

**Definition 1** [25] Two sequences of  $n \times n$  matrices  $\{\mathbf{A}_n\}$  and  $\{\mathbf{B}_n\}$  are said to be asymptotically equivalent if the following conditions are satisfied.

1.  $\mathbf{A}_n$  and  $\mathbf{B}_n$  are uniformly bounded in strong (and hence in weak) norm:

$$\|\mathbf{A}_n\|, \|\mathbf{B}_n\| \leq M < \infty, \quad n = 1, 2, \dots$$

2.  $\mathbf{A}_n - \mathbf{B}_n = \mathbf{D}_n$  goes to zero in weak norm as  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \|\mathbf{A}_n - \mathbf{B}_n\|_{\text{HS}} = \lim_{n \rightarrow \infty} \|\mathbf{D}_n\|_{\text{HS}} = 0.$$

The asymptotic equivalence between two sequences of matrices implies also the convergence to zero of both the Levy distance between the distributions of the eigenvalues of the two sequences of matrices and the mean squared Euclidean distance or error between pairs of ordered eigenvalues of the two matrices as follows from the next lemma.

**Lemma 1** *Let  $\mathbf{A}$  and  $\mathbf{B}$  be two  $n \times n$  symmetric matrices with ordered eigenvalues  $\lambda_1, \dots, \lambda_n$  and  $\mu_1, \dots, \mu_n$ , respectively. Then, the following inequalities hold:*

1. *Wielandt–Hoffman Inequality [26]*

$$\frac{1}{n} \sum_{k=1}^n |\lambda_k - \mu_k|^2 \leq \|\mathbf{A} - \mathbf{B}\|_{\text{HS}}^2,$$

2. *Inequality on the Levy distance (Theorem A.38, [1])*

$$L^3(F_n^{\mathbf{A}}, F_n^{\mathbf{B}}) \leq \|\mathbf{A} - \mathbf{B}\|_{\text{HS}}^2,$$

where  $L(F_n^{\mathbf{A}}, F_n^{\mathbf{B}})$  denotes the Levy distance between the empirical distribution functions  $F_n^{\mathbf{A}}$  and  $F_n^{\mathbf{B}}$  of the eigenvalues of  $\mathbf{A}$  and  $\mathbf{B}$ , respectively.

We observe that both  $\mathcal{L}(\mathcal{X}_n)$  and  $\mathcal{L}(\mathcal{D}_n)$  are uniformly bounded in strong norm since the spectral radius of any normalized (and also regularized) Laplacian matrix is bounded by 2. Therefore, to prove the asymptotic equivalence of the two sequences  $\{\mathcal{L}(\mathcal{X}_n)\}$  and  $\{\mathcal{L}(\mathcal{D}_n)\}$ , in the following Lemma 2, we provide an upper bound on  $\|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2$  which holds with high probability for any  $a_n$  in the connectivity and thermodynamic regimes.

**Lemma 2** *For  $d \geq 1$ , the Hilbert–Schmidt norm of the difference between  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$  is upper bounded as follows:*

$$\begin{aligned} \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 &\leq \left| \frac{1}{n} \sum_i \sum_j \frac{\left(\chi[x_i \sim x_j] + \frac{\alpha}{n}\right)^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \frac{b}{n(a'_n + \alpha)^2} \right| \\ &+ \left| \frac{2b}{n(a'_n + \alpha)^2} - \frac{2\left(\sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2\right)}{n(a'_n + \alpha) \left(\sum_i \sqrt{\mathbf{N}(x_i) + \alpha}\right)^2} \right|, \end{aligned} \tag{6}$$

where  $b = na'_n + \alpha^2 + 2\alpha a'_n$ .

**Proof** See Appendix A.  $\square$

In the thermodynamic regime, for  $d \geq 1$ , an upper bound for  $\|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2$  is obtained by setting  $a'_n \equiv \gamma'$ , where  $\gamma'$  is a positive constant. When the graph is connected as in the connectivity regime, it is not necessary to utilize the regularized normalized Laplacian and it is sufficient to consider the normalized Laplacian which is obtained from the regularized normalized Laplacian by setting  $\alpha = 0$  in (1). Then, in the connectivity regime, the upper bound in (6) reduces to the following:

$$\begin{aligned} \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 &\leq \left| \frac{1}{n} \sum_i \sum_j \frac{\chi[x_i \sim x_j]^2}{\mathbf{N}(x_i)\mathbf{N}(x_j)} - \frac{1}{a'_n} \right| \\ &\quad + \left| \frac{2}{a'_n} - \frac{2 \sum_i \mathbf{N}(x_i, x'_i)}{na'_n \left( \sum_i \sqrt{\mathbf{N}(x_i)} \right)^2} \right|. \end{aligned}$$

We provide an upper bound for the probability that the Hilbert–Schmidt norm of the difference between the DGG and RGG regularized normalized Laplacian matrices is greater than a certain value  $t$ , for any  $a_n$ .

**Proposition 1** For  $d \geq 1$ ,  $a_n \geq \frac{2d^{1+1/p}}{2d-1}$  and  $t > \frac{8(n+2\alpha)a_n+4\alpha^2}{n(a_n+\alpha)^2}$ , the following inequality holds:

$$\begin{aligned} \mathbb{P} \left\{ \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} &\leq \mathbb{P} \left\{ \sum_i |a_n - \mathbf{N}(x_i)|^2 > \frac{tn(a_n + \alpha)^2}{8} \right\} \\ &\quad + 2\mathbb{P} \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{tn(a_n + \alpha)^2 - 4\alpha^2}{8 \left(1 + \frac{2\alpha}{n}\right)} - na_n \right\} \\ &\quad + 2\mathbb{P} \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{tn(a_n + \alpha)}{16} \right\}. \end{aligned}$$

**Proof** See Appendix B.  $\square$

In the following, we show that the RGG and DGG normalized Laplacian matrices are asymptotically equivalent with high probability in the connectivity regime, i.e.,  $a_n = \Omega(\log(n))$ , as  $n \rightarrow \infty$ . For this purpose, it is sufficient to prove that  $\|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 \rightarrow 0$  as  $n \rightarrow +\infty$  with high probability, since the normalized Laplacian matrices  $\mathcal{L}(\mathcal{X}_n)$  and  $\mathcal{L}(\mathcal{D}_n)$  are uniformly bounded in strong norm.

**Proposition 2** *In the connectivity regime, i.e.,  $a_n = \Omega(\log(n))$ , for  $d \geq 1$ ,  $a_n \geq \frac{2d^{1+1/p}}{2d-1}$ ,  $p \in [1, \infty]$  and  $t > \frac{8}{a_n}$ , we have*

$$\mathbb{P} \left\{ \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} \leq \min \left[ \frac{320(n-1)\vartheta}{tn^2\gamma^2}, 6n \exp \left( -\frac{a_n(1 - \frac{1}{n\theta^{(d)}})}{12} \right) \right], \tag{7}$$

where  $\vartheta = \theta^{(d)} \left[ 1 + 2a_n - \frac{4a_n}{n} \right]$ .

**Proof** See Appendix C. □

Observe that the right hand side (r.h.s.) in (7) tends to zero as  $n \rightarrow +\infty$  for any  $a_n$  in the connectivity regime. Then, for any fixed  $t$ , it is possible to choose  $n > n_0$  to obtain a probability

$$\mathbb{P} \left\{ \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\}$$

arbitrarily small. Consequently, as  $n \rightarrow \infty$ , the value  $\|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2$  is not greater than a fixed  $t > 0$  with high probability. Hence, the difference of the RGG and DGG normalized Laplacian matrices goes to zero in weak norm with high probability and the two matrices are asymptotically equivalent. As a result, using Proposition 2, we obtain our main result in Theorem 1.

For the thermodynamic regime, we obtain the main result in Theorem 2 by applying directly Proposition 1 and setting  $a_n = \gamma$  and  $\alpha > 0$ .

Since the LEDs of the normalized Laplacian for  $G(\mathcal{X}_n, r_n)$  and  $G(\mathcal{D}_n, r_n)$  converge to the same limit as  $n \rightarrow \infty$  for any chosen  $\ell_p$ -distance and the convergence holds in the full range of the connectivity regime, i.e.,  $a_n = \Omega(\log(n))$ , then, we use the structure of the DGG to approximate the eigenvalues of the regularized normalized Laplacian matrix for  $G(\mathcal{X}_n, r_n)$  in the connectivity regimes. Theorem 2 enables a similar but looser approximation in the thermodynamic regime.

In the following, we derive the eigenvalues of  $G(\mathcal{D}_n, r_n)$ , when the geometric graphs are defined based on the Chebyshev distance. Let us consider a  $d$ -dimensional DGG with  $n = N^d$  nodes defined by the Chebyshev distance. Then, the degree of a vertex in  $G(\mathcal{D}_n, r_n)$  is given as [16]

$$a'_n = (2k_n + 1)^d - 1, \quad \text{with } k_n = \lfloor Nr_n \rfloor,$$

where  $\lfloor x \rfloor$  is the integer part, i.e., the greatest integer lower than or equal to  $x$ . Note that when  $d = 1$ , the Chebyshev distance and the Euclidean distance coincide.

In the following Lemma 3, by using the expression of the eigenvalues of the adjacency matrix for a DGG under the Chebyshev distance [16], we approximate the eigenvalues of the regularized normalized Laplacian for  $G(\mathcal{X}_n, r_n)$  when the number of nodes  $n$  is fixed and for any  $a_n$ .

**Lemma 3** When using the Chebyshev distance and  $d \geq 1$ , the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{D}_n)$  are given by

$$\hat{\lambda}_{m_1, \dots, m_d} = 1 - \frac{1}{(a'_n + \alpha)} \prod_{s=1}^d \frac{\sin(\frac{m_s \pi}{N} (a'_n + 1)^{1/d})}{\sin(\frac{m_s \pi}{N})} + \frac{1 - \alpha \delta_{m_1, \dots, m_d}}{(a'_n + \alpha)}, \quad (8)$$

with  $m_1, \dots, m_d \in \{0, \dots, N-1\}$  and  $\delta_{m_1, \dots, m_d} = 1$  when  $m_1, \dots, m_d = 0$  otherwise  $\delta_{m_1, \dots, m_d} = 0$ . In (8),  $n = N^d$ ,  $a'_n = (2k_n + 1)^d - 1$  and  $k_n = \lfloor Nr_n \rfloor$ .

**Proof** See Appendix E. □

Lemma 3 is utilized along with the almost sure asymptotic equivalence of the two matrix sequences  $\{\mathcal{L}(\mathcal{X}_n)\}$  and  $\{\mathcal{L}(\mathcal{D}_n)\}$  to obtain the LED of  $\{\mathcal{L}(\mathcal{X}_n)\}$  in the connectivity regime in Corollary 1. Additionally, it is exploited to determine an approximation of the LED of  $\{\hat{\mathcal{L}}(\mathcal{X}_n)\}$  with controlled quadratic error in the thermodynamic regime.

## 5 Numerical Results

In this section, we validate our analytical results obtained in Sect. 3 by numerical computations. More specifically, we compare the simulated and the analytical spectra of the regularized normalized Laplacian matrix of RGGs in the connectivity and thermodynamic regimes.

Figure 2(a) illustrates the empirical spectral distribution in the thermodynamic regime of a realization for an RGG with  $n = 30000$  vertices,  $\alpha = 0.001$  and the corresponding DGG. The theoretical distribution is obtained from Corollary 2. We notice that the gap that appears between the eigenvalue distributions of the RGG and the DGG is upper bounded as in Theorem 2.

Here, we provide an additional example in the thermodynamic regime to quantify the error between  $F^{\hat{\mathcal{L}}(\mathcal{X}_n)}$  and  $F^{\hat{\mathcal{L}}(\mathcal{D}_n)}$  for different values of  $\gamma$  when the random graphs are defined by the Chebyshev distance.

(1) When  $\gamma = 100$  and  $\alpha = 10^{-3}$ ,  $d = 1$  then as  $n \rightarrow \infty$

$$\mathbb{P} \left\{ \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 > \mathbf{0.019} \right\} \rightarrow 0.$$

(2) When  $\gamma = 120$  and  $\alpha = 10^{-3}$ ,  $d = 1$  then as  $n \rightarrow \infty$

$$\mathbb{P} \left\{ \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 > \mathbf{0.015} \right\} \rightarrow 0.$$

From these examples, we notice that for  $\gamma = 100$ , the LED of the regularized normalized Laplacian in the RGG can be approximated by the LED of a DGG with an error bound of 0.019 when  $\alpha = 10^{-3}$ . Then, as we increase the average vertex degree  $\gamma$  from  $\gamma = 100$  to  $\gamma = 120$ , the error bound decreases. Therefore, the larger the average vertex degree  $\gamma$  is, the tighter the approximation becomes.

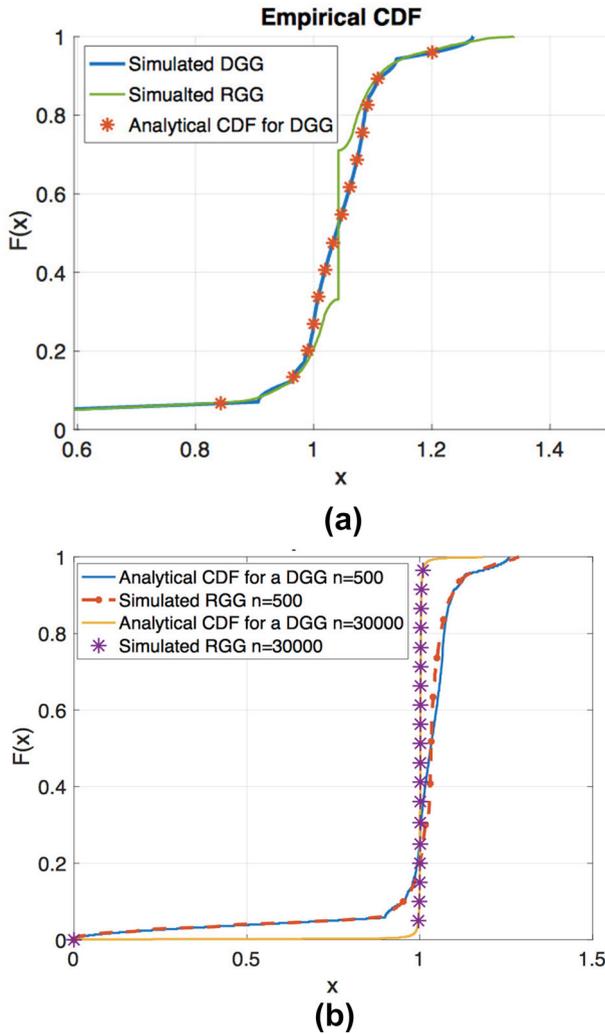


Fig. 2 Comparison between the simulated and the analytical spectral distributions of an RGG for  $d = 1$

In Fig. 2(b) we compare the spectral distribution of a DGG (continuous lines) with the one for an RGG for an increasing number of nodes  $n$  (dashed line for  $n = 500$  and star markers for  $n = 30000$ ) in the connectivity regime. We notice that the curves corresponding to the RGG and the DGG match very well when  $n$  is large confirming the asymptotic equivalence provided in Proposition 2. Additionally, it becomes apparent that by increasing  $n$ , the eigenvalue distribution converges almost surely to the Dirac measure at one. This behaviour confirms the result obtained in Corollary 1.

## 6 Conclusion

In this work, we studied the spectrum of RGGs in both the connectivity and thermodynamic regime. We first study the asymptotic equivalence of the sequences of the regularized normalized Laplacian matrix of a DGG and RGG and obtained an upper bound on the Hilbert–Schmidt norm of the difference between the regularized normalized Laplacian matrix of RGG and DGG. In the connectivity regime, as  $n \rightarrow \infty$ , we showed that the DGG and RGG regularized normalized Laplacian matrices are asymptotically equivalent with high probability. Then, using the structure of the DGG, we found that the LED of the RGG regularized normalized Laplacian matrix converges to the Dirac measure at one in the full range of the connectivity regime. In the thermodynamic regime, we showed that the LED of the RGG regularized normalized Laplacian matrix can be approximated in Hilbert–Schmidt norm by the LED of the DGG regularized normalized Laplacian matrix with an error bound which is valid with high probability. Then, with the regular structure of the DGG, we provided an analytical approximation for the eigenvalues of the RGG regularized normalized Laplacian matrix.

**Acknowledgements** This research was funded by the French Government through the Investments for the Future Program with Reference: Labex UCN@Sophia-UDCBWN.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data Availability** The authors declare that the data supporting the findings of this study are available within the article. Specifically, data for figures are provided with the paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendices

### A Proof of Lemma 2

To upper bound the Hilbert–Schmidt norm of the difference between the regularized normalized Laplacian matrices  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$ , we first give the following lemma.

**Lemma 4** *If  $a_i \geq 0$  and  $b_i > 0$  for all  $i$ , and there exists an  $a_i > 0$ , then*

$$\sum_{i=1}^n \frac{a_i}{b_i} > \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}.$$

By using Lemma 4, we upper bound the Hilbert–Schmidt norm of the difference between  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$  as

$$\begin{aligned} \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 &= \frac{1}{n} \text{Trace} \left[ \hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n) \right]^2 \\ &= \frac{1}{n} \sum_i \sum_j \left[ \frac{\chi[x_i \sim x_j] + \frac{\alpha}{n}}{\sqrt{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)}} - \frac{\chi[x'_i \sim x'_j] + \frac{\alpha}{n}}{\sqrt{(a'_n + \alpha)(a'_n + \alpha)}} \right]^2 \\ &= \frac{1}{n} \sum_i \sum_j \left[ \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} + \frac{(\chi[x'_i \sim x'_j] + \frac{\alpha}{n})^2}{(a'_n + \alpha)^2} \right] \\ &\quad - \frac{2}{n} \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})(\chi[x'_i \sim x'_j] + \frac{\alpha}{n})}{(a'_n + \alpha)\sqrt{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)}}. \end{aligned}$$

We notice from the last equation that  $\sum_j \chi[x_i \sim x_j] = \mathbf{N}(x_i)$  corresponds to the number of neighbors of vertex  $x_i$  in the RGG, and similarly  $\sum_j \chi[x'_i \sim x'_j] = a'_n$  corresponds to the number of neighbors of vertex  $x'_i$  in the DGG. Let  $b = na'_n + 2\alpha a'_n + \alpha^2$  and  $\mathbf{N}(x_i, x'_i) = \sum_j \chi[x_i \sim x_j]\chi[x'_i \sim x'_j]$ . Then, by applying Lemma 4 we obtain the following upper bound.

$$\begin{aligned} \|\hat{\mathcal{L}}(\mathcal{X}_n) - \hat{\mathcal{L}}(\mathcal{D}_n)\|_{\text{HS}}^2 &\leq \frac{b}{n(a'_n + \alpha)^2} - \frac{2 \left( \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right)}{n(a'_n + \alpha) \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \\ &\quad + \frac{1}{n} \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)}. \end{aligned}$$

Using the triangle inequality, the result follows. □

### B Proof of Proposition 1

In this appendix, we provide an upper bound for the probability that the Hilbert–Schmidt norm of the difference between the matrices  $\hat{\mathcal{L}}(\mathcal{X}_n)$  and  $\hat{\mathcal{L}}(\mathcal{D}_n)$  is greater than  $t > \frac{8(n+2\alpha)a_n+4\alpha^2}{n(a_n+\alpha)^2}$ . To prove this result, we first present the following Lemma 5 and Lemma 6.

**Lemma 5** For any chosen  $\ell_p$ -distance with  $p \in [1, \infty]$ ,  $d \geq 1$  and  $a_n \geq \frac{2d^{1+1/p}}{2d-1}$ , the following inequality holds

$$a'_n \geq \frac{a_n}{2d^{1+1/p}}, \quad (9)$$

where  $a_n$  is the average vertex degree of  $G(\mathcal{X}_n, r_n)$  and  $a'_n$  is the degree of each vertex in  $G(\mathcal{D}_n, r_n)$ .

**Proof** First, we show that (9) holds under the Chebyshev distance. Let an RGG and a DGG be obtained by connecting two nodes if the Chebyshev distance between them is at most  $r_n > 0$ . Recall that the Chebyshev distance corresponds to the metric given by the  $\ell_\infty$ -norm. Then, the degree of a  $d$ -dimensional DGG with  $n$  nodes formed by using the Chebyshev distance is given by [16]

$$\begin{aligned} a'_n &= \left(2\lfloor n^{1/d} r_n \rfloor + 1\right)^d - 1 \\ &= \left(2\lfloor a_n^{1/d} \rfloor + 1\right)^d - 1, \end{aligned}$$

where  $\lfloor x \rfloor$  is the integer part, i.e., the greatest integer less than or equal to  $x$ .

For  $p = \infty$ , we have

$$\begin{aligned} \frac{a_n}{2d} \leq a'_n &\iff a_n \leq 2da'_n \iff a_n \leq 2d \left(2\lfloor a_n^{1/d} \rfloor + 1\right)^d - 2d \\ &\iff (a_n + 2d) \leq 2d \left(2\lfloor a_n^{1/d} \rfloor + 1\right)^d. \end{aligned}$$

Notice that  $\lfloor a_n^{1/d} \rfloor \geq (a_n^{1/d} - 1)$ , then it is sufficient to show that

$$\begin{aligned} (a_n + 2d) \leq 2d \left(2(a_n^{1/d} - 1) + 1\right)^d &\iff (a_n + 2d) \leq 2d \left(2a_n^{1/d} - 1\right)^d \\ &\iff \left(\frac{1}{2d} + \frac{1}{a_n}\right) \leq \left(2 - \frac{1}{a_n^{1/d}}\right)^d. \end{aligned}$$

By taking the log in both sides of the last inequality, yields

$$\ln \left(\frac{1}{2d} + \frac{1}{a_n}\right) \leq d \ln \left(2 - \frac{1}{a_n^{1/d}}\right).$$

Consequently, under the Chebyshev distance, (9) holds for  $a_n \geq \frac{2d}{2d-1}$ .

Next, we show that (9) holds under any  $\ell_p$ -distance,  $p \in [1, \infty]$ . Let  $b_n$  and  $b'_n$  be the degrees of an RGG and a DGG formed by connecting each two nodes when  $d^{1/p} \|x_i - x_j\|_\infty \leq r_n$ . This simply means that the graphs are obtained using the Chebyshev distance with a radius equal to  $\frac{r_n}{d^{1/p}}$ . Then, the degree of the DGG can be

written as

$$b'_n = \left(2 \lfloor b_n^{1/d} \rfloor + 1\right)^d - 1.$$

When  $p = \infty$ , we have that (9) holds. Therefore, we deduce that for  $b_n \geq \frac{2d}{2d-1}$ , we get

$$\frac{b_n}{2d} \leq b'_n.$$

Note that for any  $\ell_p$ -distance with  $p \in [1, \infty)$  in  $\mathbb{R}^d$ , we have

$$\|x_i - x_j\|_p \leq d^{1/p} \|x_i - x_j\|_\infty.$$

Then the number of nodes  $a'_n$  in the DGG that falls in the ball of radius  $r_n$  is greater or equal than  $b'_n$ , i.e.,  $b'_n \leq a'_n$ . Therefore,

$$\frac{b_n}{2d} = \frac{a_n}{2d^{1+1/p}} \leq b'_n \leq a'_n.$$

Hence, for  $a_n \geq \frac{2d^{1+1/p}}{2d-1}$ , we get

$$\frac{a_n}{2d^{1+1/p}} \leq a'_n. \quad \square$$

**Lemma 6** For any chosen  $\ell_p$ -distance with  $p \in [1, \infty)$ ,  $d \geq 1$ ,  $a_n \geq \frac{2d^{1+1/p}}{2d-1}$ ,  $a'_n > \frac{a_n}{1d^{1+\frac{1}{p}}}$ , we have

$$t > \max \left[ \frac{4(n+2\alpha)a'_n + 4\alpha^2}{n(a'_n + \alpha)^2}, \frac{8(n+2\alpha)a_n + 4\alpha^2}{n(a_n + \alpha)^2} \right] \simeq t > \frac{8(n+2\alpha)a_n + 4\alpha^2}{n(a_n + \alpha)^2}.$$

**Proof** For  $\alpha$  sufficiently small, we let  $c = \frac{\alpha^2}{a_n} \simeq \frac{\alpha^2}{a'_n}$ , then we have,

$$\begin{aligned} t &> \max \left[ \frac{4(n+2\alpha)a'_n + 4\alpha^2}{n(a'_n + \alpha)^2}, \frac{8(n+2\alpha)a_n + 4\alpha^2}{n(a_n + \alpha)^2} \right] \\ &\simeq t > \max \left[ \frac{4[(n+2\alpha) + c]}{a'_n n(1+c)^2}, \frac{8[(n+2\alpha) + \frac{c}{2}]}{a_n n(1+c)^2} \right] \end{aligned}$$

Therefore, if  $\frac{a_n}{2d^{1+\frac{1}{p}}} < a_n < a'_n$ , then

$$a_n < a'_n \Leftrightarrow \frac{4}{a'_n} < \frac{4}{a_n} < \frac{8}{a_n}.$$

If  $a_n > a'_n$ , and  $a'_n > \frac{a_n}{2d^{1+\frac{1}{p}}}$ , then

$$\frac{4}{a_n} < \frac{4}{a'_n} < \frac{8d^{1+\frac{1}{p}}}{a_n}.$$

□

Now, we prove the main result presented in the proposition.

$$\begin{aligned} & \mathbb{P} \left\{ L^3 \left( F^{\hat{\mathcal{L}}(\mathcal{X}_n)}, F^{\hat{\mathcal{L}}(\mathcal{D}_n)} \right) > t \right\} \\ & \leq \mathbb{P} \left\{ \left| \frac{1}{n} \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \frac{b}{n(a'_n + \alpha)^2} \right| \right. \\ & \quad \left. + \left| \frac{2b}{n(a'_n + \alpha)^2} - \frac{2 \left( \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right)}{n(a'_n + \alpha) \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \right| > t \right\} \\ & \leq \mathbb{P} \left\{ \left| \frac{1}{n} \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \frac{b}{n(a'_n + \alpha)^2} \right| > \frac{t}{2} \right\} \\ & \quad + \mathbb{P} \left\{ \left| \frac{2b}{n(a'_n + \alpha)^2} - \frac{2 \left( \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right)}{n(a'_n + \alpha) \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \right| > \frac{t}{2} \right\}. \end{aligned}$$

Define,

$$A = \left| \frac{2b}{n(a'_n + \alpha)^2} - \frac{2 \left( \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right)}{n(a'_n + \alpha) \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \right|,$$

and

$$B = \left| \frac{1}{n} \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \frac{b}{n(a'_n + \alpha)^2} \right|.$$

In the following, we upper bound  $\mathbb{P} \left\{ A > \frac{t}{2} \right\}$  and  $\mathbb{P} \left\{ B > \frac{t}{2} \right\}$ .

$$\begin{aligned}
 &\text{First, we write } P \left\{ A > \frac{t}{2} \right\} = \\
 &P \left\{ \left| \frac{2b}{n(a'_n + \alpha)^2} - \frac{2 \left( \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right)}{n(a'_n + \alpha) \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \right| > \frac{t}{2} \right\} \\
 &= P \left\{ \left| 1 - \frac{(a'_n + \alpha) \left[ \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right]}{b \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \right| > \frac{tn(a'_n + \alpha)^2}{4b} \right\} \\
 &\stackrel{(a)}{=} P \left\{ b \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2 - (a'_n + \alpha) \left[ \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) \right. \right. \\
 &\quad \left. \left. + \alpha a'_n + \alpha^2 \right] > \frac{tn(a'_n + \alpha)^2}{4} \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2 \right\} \\
 &= P \left\{ \left( b - \frac{tn(a'_n + \alpha)^2}{4} \right) \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2 \right. \\
 &\quad \left. > (a'_n + \alpha) \left[ \sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2 \right] \right\}.
 \end{aligned}$$

Note that  $\sum_i \mathbf{N}(x_i, x'_i) \leq na'_n$  and  $\mathbf{N}(x_i) \leq n$ . Then, in step (a) for  $n$  sufficiently large, we remove the absolute value because

$$\begin{aligned}
 \frac{\sum_i \mathbf{N}(x_i, x'_i) + \frac{\alpha}{n} \sum_i \mathbf{N}(x_i) + \alpha a'_n + \alpha^2}{b \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} &\leq \frac{na'_n + \alpha n + \alpha a'_n + \alpha^2}{b \left( \sum_i \sqrt{\mathbf{N}(x_i) + \alpha} \right)^2} \\
 &\leq \frac{\left( 1 + \frac{\alpha}{a'_n} + \frac{\alpha}{n} + \frac{\alpha^2}{na'_n} \right)}{\left( 1 + \frac{2\alpha}{n} + \frac{\alpha^2}{na'_n} \right) n^2 \alpha} \leq 1.
 \end{aligned}$$

Notice from the last equality that,  $\frac{tn(a'_n + \alpha)^2}{4} > b \Leftrightarrow t > \frac{4na'_n + 4\alpha^2 + 8\alpha a'_n}{n(a'_n + \alpha)^2}$ .

Then,

$$P \left\{ A > \frac{t}{2} \right\} = 0 \text{ for } t > \frac{4na'_n + 4\alpha^2 + 8\alpha a'_n}{n(a'_n + \alpha)^2}. \tag{10}$$

We continue further by bounding  $P\left\{B > \frac{t}{2}\right\} =$

$$\begin{aligned} & P\left\{\left|\frac{1}{n}\sum_i\sum_j\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(\mathbf{N}(x_i)+\alpha)(\mathbf{N}(x_j)+\alpha)}-\frac{b}{n(a'_n+\alpha)^2}\right|>\frac{t}{2}\right\} \\ &= P\left\{\left|\frac{1}{n}\sum_i\sum_j\left(\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(\mathbf{N}(x_i)+\alpha)(\mathbf{N}(x_j)+\alpha)}-\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}\right.\right.\right. \\ &\quad \left.\left.\left.+\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}\right)-\frac{b}{n(a'_n+\alpha)^2}\right|>\frac{t}{2}\right\} \\ &\leq P\left\{\left|\frac{1}{n}\sum_i\sum_j\left(\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(\mathbf{N}(x_i)+\alpha)(\mathbf{N}(x_j)+\alpha)}-\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}\right)\right|>\frac{t}{4}\right\} \\ &\quad + P\left\{\left|\frac{1}{n}\sum_i\sum_j\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}-\frac{b}{n(a'_n+\alpha)^2}\right|>\frac{t}{4}\right\}. \end{aligned}$$

Let

$$B_1 = \left|\frac{1}{n}\sum_i\sum_j\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}-\frac{b}{n(a'_n+\alpha)^2}\right|,$$

and

$$B_2 = \left|\frac{1}{n}\sum_i\sum_j\left(\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(\mathbf{N}(x_i)+\alpha)(\mathbf{N}(x_j)+\alpha)}-\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}\right)\right|.$$

Now, we upper bound the two probabilities  $P\left\{B_1 > \frac{t}{4}\right\}$  and  $P\left\{B_2 > \frac{t}{4}\right\}$ .

First, for  $t > \frac{4na'_n+4\alpha^2+8\alpha a'_n}{n(a'_n+\alpha)^2}$ , we write  $P\left\{B_1 > \frac{t}{4}\right\} =$

$$\begin{aligned} &= P\left\{\left|\frac{1}{n}\sum_i\sum_j\frac{(\chi[x_i\sim x_j]+\frac{\alpha}{n})^2}{(a_n+\alpha)^2}-\frac{b}{n(a'_n+\alpha)^2}\right|>\frac{t}{4}\right\} \\ &= P\left\{\left|\sum_i\sum_j\left(\chi[x_i\sim x_j]+\frac{\alpha}{n}\right)^2-\frac{b(a_n+\alpha)^2}{(a'_n+\alpha)^2}\right|>\frac{nt(a_n+\alpha)^2}{4}\right\} \\ &\stackrel{(a)}{\leq} P\left\{\left|\sum_i\mathbf{N}(x_i)-na_n\right|>\frac{nt(a_n+\alpha)^2-4\alpha^2}{4\left(1+\frac{2\alpha}{n}\right)}-na_n\right\}. \end{aligned} \tag{11}$$

Step (a) follows from  $\sum_i \sum_j (\chi[x_i \sim x_j] + \frac{\alpha}{n})^2 = (1 + \frac{2\alpha}{n}) \sum_i \mathbf{N}(x_i) + \alpha^2$  and substituting the value of  $b$ .

We continue further by upper bounding the term  $\mathbb{P} \left\{ B_2 > \frac{t}{4} \right\} =$

$$= \mathbb{P} \left\{ \left| \sum_i \sum_j \left( \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} \right) \right| > \frac{tn}{4} \right\}$$

$$= \mathbb{P} \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} > \frac{tn}{4} \right\}$$

$$+ \mathbb{P} \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} > \frac{tn}{4} \right\}.$$

Define

$$C_1 = \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)},$$

$$C_2 = \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2}.$$

In the following, we upper bound  $\mathbb{P} \left\{ C_1 > \frac{tn}{4} \right\}$  and  $\mathbb{P} \left\{ C_2 > \frac{tn}{4} \right\}$ .

We start first with  $\mathbb{P} \left\{ C_1 > \frac{tn}{4} \right\} =$

$$= \mathbb{P} \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} > \frac{tn}{4} \right\}$$

$$\stackrel{(a)}{\leq} \mathbb{P} \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} - \frac{(1 + \frac{2\alpha}{n}) \sum_i \mathbf{N}(x_i) + \alpha^2}{\left( \sum_i \mathbf{N}(x_i) + n\alpha \right) \sum_j (\mathbf{N}(x_j) + \alpha)} > \frac{tn}{4} \right\}$$

$$= \mathbb{P} \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} - \frac{\alpha \left[ \left( \frac{1}{\alpha} + \frac{2}{n} \right) \sum_i \mathbf{N}(x_i) + \alpha \right]}{n \left( \frac{1}{n} \sum_i \mathbf{N}(x_i) + \alpha \right) \sum_j (\mathbf{N}(x_j) + \alpha)} > \frac{tn}{4} \right\}$$

$$\stackrel{(b)}{\leq} \mathbb{P} \left\{ \sum_{i,j} \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} - \frac{\alpha \left[ \left( \frac{1}{\alpha} + \frac{2}{n} \right) \sum_i \mathbf{N}(x_i) + \alpha \right]}{n \left[ \left( \frac{1}{\alpha} + \frac{2}{n} \right) \sum_i \mathbf{N}(x_i) + \alpha \right] \sum_j (\mathbf{N}(x_j) + \alpha)} > \frac{tn}{4} \right\}.$$

Step (a) follows from Lemma 4 and step (b) from  $(\frac{1}{\alpha} + \frac{2}{n}) \sum_i \mathbf{N}(x_i) + \alpha > \frac{1}{n} \sum_i \mathbf{N}(x_i) + \alpha$ .

$$\begin{aligned} \mathbb{P} \left\{ C_1 > \frac{tn}{4} \right\} &= \mathbb{P} \left\{ (n + 2\alpha) \sum_i \mathbf{N}(x_i) - \frac{\alpha(a_n + \alpha)^2}{\sum_j \mathbf{N}(x_j) + n\alpha} > \frac{tn^2(a_n + \alpha)^2}{4} - n\alpha^2 \right\} \\ &\leq \mathbb{P} \left\{ (n + 2\alpha) \left| \sum_i \mathbf{N}(x_i) - na_n \right| + \left| (n + 2\alpha) na_n - \frac{\alpha(a_n + \alpha)^2}{\sum_j \mathbf{N}(x_j) + n\alpha} \right| > \frac{tn^2(a_n + \alpha)^2}{4} - n\alpha^2 \right\} \\ &\leq \mathbb{P} \left\{ \left| (n + 2\alpha) na_n - \frac{\alpha(a_n + \alpha)^2}{\sum_j \mathbf{N}(x_j) + n\alpha} \right| > \frac{tn^2(a_n + \alpha)^2 - 4n\alpha^2}{8} \right\} \\ &\quad + \mathbb{P} \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{tn(a_n + \alpha)^2 - 4\alpha^2}{8(1 + \frac{2\alpha}{n})} \right\}. \end{aligned}$$

Let

$$C_3 = \mathbb{P} \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{tn(a_n + \alpha)^2 - 4\alpha^2}{8(1 + \frac{2\alpha}{n})} \right\}, \tag{12}$$

and

$$\begin{aligned} D &= \left| (n + 2\alpha) na_n - \frac{\alpha(a_n + \alpha)^2}{\sum_j \mathbf{N}(x_j) + n\alpha} \right| \\ \mathbb{P} \left\{ D > \frac{tn^2(a_n + \alpha)^2 - 4n\alpha^2}{8} \right\} \\ &= \mathbb{P} \left\{ \left| (n + 2\alpha) na_n - \frac{\alpha(a_n + \alpha)^2}{\sum_j \mathbf{N}(x_j) + n\alpha} \right| > \frac{tn^2(a_n + \alpha)^2 - 4n\alpha^2}{8} \right\} \\ &\stackrel{(a)}{\leq} \mathbb{P} \left\{ \left[ (n + 2\alpha) na_n + \frac{4n\alpha^2 - tn^2(a_n + \alpha)^2}{8} \right] \times \left( \sum_j \mathbf{N}(x_j) + n\alpha \right) > \alpha(a_n + \alpha)^2 \right\}. \end{aligned}$$

Step (a) follows from the inequality  $(n + 2\alpha) na_n > \frac{\alpha(a_n + \alpha)^2}{\sum_j \mathbf{N}(x_j) + n\alpha}$ . Notice that

$$8(n + 2\alpha) na_n + 4n\alpha^2 < tn^2(a_n + \alpha)^2 \Leftrightarrow t > \frac{8(n + 2\alpha) a_n + 4\alpha^2}{n(a_n + \alpha)^2}. \text{ Then,}$$

$$P \left\{ D > \frac{tn^2(a_n + \alpha)^2 - 4n\alpha^2}{8} \right\} = 0 \text{ for } t > \frac{8(n + 2\alpha) a_n + 4\alpha^2}{n(a_n + \alpha)^2}. \tag{13}$$

Finally, we upper bound the remaining probability

$$\begin{aligned} P \left\{ C_2 > \frac{tn}{4} \right\} &= P \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} > \frac{tn}{4} \right\} \\ &= P \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})}{(\mathbf{N}(x_i) + \alpha)} \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})}{(\mathbf{N}(x_j) + \alpha)} \right. \\ &\quad \left. - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} > \frac{tn}{4} \right\} \\ &\stackrel{(a)}{\leq} P \left\{ \left( \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)^2} \right)^{\frac{1}{2}} \left( \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_j) + \alpha)^2} \right)^{\frac{1}{2}} \right. \\ &\quad \left. - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} > \frac{tn}{4} \right\} \\ &= P \left\{ \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(\mathbf{N}(x_i) + \alpha)^2} - \sum_i \sum_j \frac{(\chi[x_i \sim x_j] + \frac{\alpha}{n})^2}{(a_n + \alpha)^2} > \frac{tn}{4} \right\}. \tag{14} \end{aligned}$$

Step (a) follows from applying Cauchy-Schwarz inequality. Then,  $P \left\{ C_2 > \frac{tn}{4} \right\} \leq$

$$\begin{aligned} &P \left\{ \sum_i \sum_j \left( \chi[x_i \sim x_j] + \frac{\alpha}{n} \right)^2 \left| \frac{1}{(\mathbf{N}(x_i) + \alpha)^2} - \frac{1}{(a_n + \alpha)^2} \right| > \frac{tn}{4} \right\} \\ &\leq P \left\{ \sum_i \left| (a_n + \alpha)^2 - (\mathbf{N}(x_i) + \alpha)^2 \right| > \frac{tn(a_n + \alpha)^2}{4} \right\} \\ &\leq P \left\{ \sum_i |a_n - \mathbf{N}(x_i)|^2 + 2(a_n + \alpha) \sum_i |a_n - \mathbf{N}(x_i)| > \frac{tn(a_n + \alpha)^2}{4} \right\} \\ &\leq P \left\{ \sum_i |a_n - \mathbf{N}(x_i)|^2 > \frac{tn(a_n + \alpha)^2}{8} \right\} + \\ &2P \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{tn(a_n + \alpha)}{16} \right\}. \tag{15} \end{aligned}$$

Finally,  $P\{C_1 > \frac{tn}{4}\}$  is upper bounded by the sum of (12) and (13). We Use this result combined with the upper bound of  $P\{C_2 > \frac{tn}{4}\}$  given in (15) to upper bound the term  $P\{B_2 > \frac{t}{4}\}$ . Then, apply the new upper bound with (10) and (11) to upper bound (1) and with Lemma 6, we get Proposition 1 follows.  $\square$

### C Proof of Theorem 2 and Proposition 2

In this Appendix, we show that the LED of the regularized normalized Laplacian for a DGG is a good approximation for the LED of the regularized normalized Laplacian for an RGG in both the connectivity and thermodynamic regimes.

To upper bound the terms obtained in Proposition 1, we use the Chebyshev inequality. Notice that  $\sum_i N(x_i)/2$  that appears in Proposition 1 counts the number of edges in  $G(\mathcal{X}_n, r_n)$ . For convenience, we denote  $\sum_i N(x_i)/2$  as  $\xi_n$ . In order to apply the Chebyshev inequality, we determine the variance of the number of edges, i.e.,  $\text{Var}(\xi_n)$  in the following lemma.

**Lemma 7** (Variance of  $\xi_n$ ) *When  $x_1, \dots, x_n$  are i.i.d. uniformly distributed in the  $d$ -dimensional unit torus  $\mathbb{T}^d = [0, 1]$*

$$\begin{aligned} \text{Var}(\xi_n) &\leq \frac{(n-1)}{n} \left[ \theta^{(d)} + 2(n-2)(\theta^{(d)})^2 r_n^d \right]. \\ &= \frac{\theta^{(d)}(n-1)}{n} \left[ 1 + 2a_n - \frac{4a_n}{n} \right]. \end{aligned}$$

**Proof** The proof follows along the same lines of Proposition A.1 in [27] when extended to a unit torus and applied to i.i.d. and uniformly distributed nodes.  $\square$

Let  $\vartheta = \theta^{(d)} \left[ 1 + 2a_n - \frac{4a_n}{n} \right]$ . In the following, we upper bound the probabilities given in Proposition 1 using Lemma 7 and the Chebyshev inequality. We start by upper bounding the first term as follows:

$$\begin{aligned} &P \left\{ \left| \sum_i N(x_i) - na_n \right| > \frac{tn(a_n + \alpha)^2 - 4\alpha^2}{4 \left(1 + \frac{2\alpha}{n}\right)} - na_n \right\} \\ &= P \left\{ |\xi_n - \mathbb{E}\xi_n| > \frac{tn(a_n + \alpha)^2 - 4\alpha^2}{8 \left(1 + \frac{2\alpha}{n}\right)} - na_n \right\} \\ &\leq \frac{8^2 \left(1 + \frac{2\alpha}{n}\right)^2 \text{Var}(\xi_n)}{\left[tn(a_n + \alpha)^2 - 4\alpha^2 - 8\left(1 + \frac{2\alpha}{n}\right)na_n\right]^2} \\ &= \frac{8^2 \left(1 + \frac{2\alpha}{n}\right)^2 (n-1)\vartheta}{n \left[tn(a_n + \alpha)^2 - 4\alpha^2 - 8\left(1 + \frac{2\alpha}{n}\right)na_n\right]^2}, \end{aligned} \tag{16}$$

and

$$\begin{aligned} \mathbb{P} \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{tn(a_n + \alpha)}{16} \right\} &\leq \frac{32^2 \text{Var}(\xi_n)}{t^2 n^2 (a_n + \alpha)^2} \\ &= \frac{32^2 (n - 1) \vartheta}{t^2 n^3 (a_n + \alpha)^2}. \end{aligned} \tag{17}$$

Finally, we upper bound the last term as

$$\begin{aligned} \mathbb{P} \left\{ \sum_i |\mathbf{N}(x_i) - a_n|^2 > \frac{nt(a_n + \alpha)^2}{8} \right\} &\leq \mathbb{P} \left\{ \left( \sum_i |\mathbf{N}(x_i) - a_n| \right)^2 > \frac{nt(a_n + \alpha)^2}{8} \right\} \\ &= \mathbb{P} \left\{ \sum_i |\mathbf{N}(x_i) - a_n| > \frac{(a_n + \alpha)\sqrt{nt}}{2\sqrt{2}} \right\}. \end{aligned}$$

Then,

$$\begin{aligned} \mathbb{P} \left\{ \sum_i |\mathbf{N}(x_i) - a_n|^2 > \frac{nt(a_n + \alpha)^2}{8} \right\} &\leq 2\mathbb{P} \left\{ \left| \sum_i \mathbf{N}(x_i) - na_n \right| > \frac{(a_n + \alpha)\sqrt{nt}}{2\sqrt{2}} \right\} \\ &= 2\mathbb{P} \left\{ |\xi_n - \mathbb{E}\xi_n| > \frac{(a_n + \alpha)\sqrt{nt}}{4\sqrt{2}} \right\} \\ &\leq \frac{64(n - 1)\vartheta}{tn^2(a_n + \alpha)^2}. \end{aligned} \tag{18}$$

Corollary 2 for the thermodynamic regime is obtained from upper bounding the terms in the r.h.s of (1) in Proposition 1 obtained in (16), (17) and (18). In addition, by letting  $\alpha \rightarrow 0$ , the provided upper bounds hold in the connectivity case.

In the following, we propose a tighter upper bound for the Levy distance between  $F^{\mathcal{L}(\mathcal{D}_n)}$  and  $F^{\mathcal{L}(\mathcal{X}_n)}$  in the connectivity regime. More precisely, the following upper bounds are tighter when the average vertex degree scales as  $\Omega(\log^{1+\epsilon}(n))$  or for  $a_n = c \log(n)$  when  $c > 24$ .

First, observe that the number of nodes that fall in an arbitrary interval of radius  $r_n$  follows a binomial distribution. Then in order to derive the distribution of  $\mathbf{N}(x_i)$ , we need to derive the distribution of the nodes that fall in a ball centered in  $x_i$ . To derive the distribution of  $\mathbf{N}(x_i)$  in the connectivity regime, we throw at random a node which will be in a random position and we are left with  $n - 1$  nodes. Then, we take a ball of size  $2r_n$ , centered around the thrown node. If we throw randomly the remaining  $n - 1$  nodes, then,  $\mathbf{N}(x_i)$  will be a random variable binomially distributed with parameters  $(n - 1, \theta^{(d)}r_n)$ , i.e.,

$$\mathbb{P}(\mathbf{N}(x_i) = k) = \binom{n - 1}{k} (\theta^{(d)}r_n)^k (1 - \theta^{(d)}r_n)^{n - k - 1}, \text{ for } k = 0, \dots, n - 1.$$

To upper bound the terms in the r.h.s of (1) given in Proposition 1, we introduce upper bounds for a binomially distributed random variable  $X$  appropriate for large deviations. These results play a key role to establish the relation between  $F^{\mathcal{L}(\mathcal{X}_n)}$  and  $F^{\mathcal{L}(\mathcal{D}_n)}$ .

**Lemma 8** ([28], Theorem 2.1, page 26) *Let  $X \sim \text{Bin}(n, p)$ ,  $\mathbb{E}X = np$ , then*

$$\mathbb{P}\{X \geq \mathbb{E}X + t\} \leq \exp\left(-\frac{t^2}{2(\mathbb{E}X + t/3)}\right), \quad t \geq 0.$$

**Lemma 9** ([28], corollary 2.3, page 27) *Let  $X \sim \text{Bin}(n, p)$ ,  $\mathbb{E}X = np$  and  $0 < t \leq 3/2$ , then*

$$\mathbb{P}\{|X - \mathbb{E}X| \geq t\mathbb{E}X\} \leq 2 \exp\left(-\frac{t^2}{3}\mathbb{E}X\right).$$

We upper bound the first term in the r.h.s of (1) in Proposition 1 by using Lemmas 8 and 9 as follows. For  $t > \frac{8}{a_n}$ , we have

$$\begin{aligned} \mathbb{P}\left\{\left|\sum_i \mathbf{N}(x_i) - na_n\right| > \frac{tna_n^2}{4} - na_n\right\} &\stackrel{(a)}{\leq} \mathbb{P}\left\{\sum_i |\mathbf{N}(x_i) - a_n| > \frac{nta_n^2}{4} - na_n\right\} \\ &\stackrel{(b)}{\leq} n\mathbb{P}\left\{|\mathbf{N}(x_i) - a_n| > \frac{ta_n^2}{4} - a_n\right\} \\ &\leq n\mathbb{P}\left\{\mathbf{N}(x_i) > \frac{ta_n^2}{4}\right\} \\ &\quad + n\mathbb{P}\left\{\mathbf{N}(x_i) < 2a_n - \frac{ta_n^2}{4}\right\} \\ &\leq n \exp\left(-\frac{(ta_n - 4)^2}{16\left(\frac{t}{6} - \frac{2}{3a_n} + \frac{2}{a_n}\right)}\right). \end{aligned} \tag{19}$$

Step (a) follows from  $\left|\sum_i z_i\right| \leq \sum_i |z_i|$  and step (b) from  $\sum_i |\mathbf{N}(x_i) - a_n| \leq n|\mathbf{N}(x_i) - a_n|$ .

Now, instead of upper bounding the two last probabilities given in Proposition 1, we go back to Appendix B and upper bound (14) by letting  $\alpha \rightarrow 0$ .

$$\begin{aligned} \mathbb{P}\left\{C_2 > \frac{tn}{4}\right\} &\leq \mathbb{P}\left\{\sum_i \sum_j \frac{\chi[x_i \sim x_j]^2}{\mathbf{N}(x_i)^2} - \sum_i \sum_j \frac{\chi[x_i \sim x_j]^2}{a_n^2} > \frac{tn}{4}\right\} \\ &\leq n\mathbb{P}\left\{\frac{1}{\mathbf{N}(x_i)} - \frac{\mathbf{N}(x_i)}{a_n^2} > \frac{t}{4}\right\} \end{aligned}$$

$$\begin{aligned}
 &= n\mathbb{P} \left\{ a_n^2 - \mathbf{N}(x_i)^2 > \frac{ta_n^2\mathbf{N}(x_i)}{4} \right\} \\
 \mathbb{P} \left\{ C_2 > \frac{tn}{4} \right\} &= n\mathbb{P} \left\{ \left[ a_n^2 - \mathbf{N}(x_i)^2 \right] + \frac{ta_n^2}{4} [a_n - \mathbf{N}(x_i)] > \frac{ta_n^3}{4} \right\} \\
 &\leq n\mathbb{P} \left\{ \left| a_n^2 - \mathbf{N}(x_i)^2 \right| > \frac{ta_n^3}{8} \right\} + n\mathbb{P} \left\{ |a_n - \mathbf{N}(x_i)| > \frac{a_n}{2} \right\} \\
 &\leq n\mathbb{P} \left\{ -\mathbf{N}(x_i)^2 > \frac{ta_n^3}{8} - a_n^2 \right\} + n\mathbb{P} \left\{ \left| a_n - r_n^d - \mathbf{N}(x_i) \right| > \frac{a_n - r_n^d}{2} \right\} \\
 &\quad + n\mathbb{P} \left\{ \mathbf{N}(x_i) > a_n - r_n^d + \sqrt{\frac{ta_n^3}{8} + a_n^2 - a_n + r_n^d} \right\} \\
 &= n\mathbb{P} \left\{ -\mathbf{N}(x_i)^2 > \frac{ta_n^3}{8} - a_n^2 \right\} + n\mathbb{P} \left\{ |\mathbf{N}(x_i) - \mathbb{E}\mathbf{N}(x_i)| > \frac{a_n - r_n^d}{2} \right\} \\
 &\quad + n\mathbb{P} \left\{ \mathbf{N}(x_i) > \mathbb{E}\mathbf{N}(x_i) + \sqrt{\frac{ta_n^3}{8} + a_n^2 - a_n + r_n^d} \right\}.
 \end{aligned}$$

Then, applying Lemma 8 and 9, we have for  $t > \frac{8}{a_n}$ ,  $\mathbb{P} \left\{ C_2 > \frac{tn}{4} \right\} \leq$

$$2n \exp \left( -\frac{(a_n - r_n^d)}{12} \right) + n \exp \left( \frac{-3a_n \left[ \sqrt{\frac{t}{8}a_n + 1} - 1 + \frac{r_n^d}{a_n} \right]^2}{2 \left[ 2 - \frac{2r_n^d}{a_n} + \sqrt{\frac{t}{8}a_n + 1} \right]} \right). \tag{20}$$

Finally, taking the upper bounds found by using the Chebyshev inequality in (16), (17) and (18) combined with the upper bounds found by using Lemmas 8 and 9, i.e., (19), (20) all together, then by applying Lemma 5 and letting  $\alpha \rightarrow 0$ , Proposition 2 follows.  $\square$

### D Proof of Theorem 1

Since Proposition 2 holds under the conditions that  $\vartheta < \theta^{(d)}a_n \left( \frac{1}{a_n} + 2 \right)$  and  $t > \frac{8d^{1+1/p}}{a_n}$ , we apply these conditions on (7). Then,

$$A = \frac{320(n-1)\vartheta}{\text{tn}^2a_n^2} < \frac{320\vartheta}{\text{tn}a_n^2} < \frac{40\vartheta}{\text{na}_nd^{1+1/p}} = \frac{40\theta^{(d)} \left( \frac{1}{a_n} + 2 \right)}{\text{nd}^{1+1/p}},$$

and

$$B = \frac{6n}{\exp\left(\frac{a_n}{12} \left[1 - \frac{1}{n\theta^{(d)}}\right]\right)} = \frac{6}{n \frac{a_n}{12 \log(n)} \left[1 - \frac{1}{n\theta^{(d)}}\right]^{-1}}.$$

Consequently,

$$\mathbb{P} \left\{ \|\mathcal{L}(\mathcal{X}_n) - \mathcal{L}(\mathcal{D}_n)\|_{\text{HS}}^2 > t \right\} < \min \left[ \frac{40\theta^{(d)} \left(\frac{1}{a_n} + 2\right)}{nd^{1+1/p}}, \frac{6}{n \frac{a_n}{12 \log(n)} \left[1 - \frac{1}{n\theta^{(d)}}\right]^{-1}} \right].$$

When  $a_n = c \log(n)$  and  $c > 24$ , then  $B < A$  and the rate of convergence is  $\mathcal{O}(1/n^{c/12-1})$ , whereas, when  $c \leq 24$ , the rate of convergence is  $\mathcal{O}(1/n)$ . For  $\epsilon > 0$  and  $a_n \geq \log^{1+\epsilon}(n)$ , the rate of convergence is  $\mathcal{O}(1/n^{(a_n/12 \log(n))^{-1}})$ . When the graph is dense, i.e.,  $a_n$  scales as  $\Omega(n)$ , the LED of the normalized Laplacian matrix of the RGG converges to the Dirac measure at one with rate of convergence  $\mathcal{O}(ne^{-n/12})$ . Hence, from the result in Proposition 2 follows that the LEDs of the normalized Laplacian matrix for  $G(\mathcal{X}_n, r_n)$  and  $G(\mathcal{D}_n, r_n)$  converge to the same limit as  $n \rightarrow \infty$  under any chosen  $\ell_p$ -distance in the definition of the graphs, and the convergence holds in the full range of the connectivity regime, i.e.,  $a_n = \Omega(\log(n))$ .  $\square$

### E Proof of Lemma 3

In this appendix, we provide the eigenvalues of the regularized normalized Laplacian matrix for a DGG using the Chebyshev distance. Then, the degree of a vertex in  $G(\mathcal{D}_n, r_n)$  is given as [16]

$$a'_n = (2k_n + 1)^d - 1, \quad \text{with } k_n = \lfloor Nr_n \rfloor,$$

where  $\lfloor x \rfloor$  is the integer part, i.e., the greatest integer less than or equal to  $x$ . The regularized normalized Laplacian can be written as

$$\hat{\mathcal{L}}(\mathcal{D}_n) = \mathbf{I} - \frac{1}{(a'_n + \alpha)} \mathbf{A} - \frac{\alpha}{n(a'_n + \alpha)} \mathbf{1}\mathbf{1}^\top, \quad (\text{A})$$

where  $\mathbf{I}$  is the identity matrix,  $\mathbf{1}^\top = [1, \dots, 1]^\top$  is the vector of all ones and  $\mathbf{A}$  is the adjacency matrix defined as

$$\mathbf{A}_{ij} = \begin{cases} 1, & \|x_i - x_j\|_p \leq r_n, \quad i \neq j, \quad \text{and } p \in [1, \infty), \\ 0, & \text{otherwise.} \end{cases}$$

When  $d = 1$ , the adjacency matrix  $\mathbf{A}$  of a DGG in  $\mathbb{T}^1$  with  $n$  nodes is a circulant matrix. A well known result appearing in [25], states that the eigenvalues of a circulant

matrix are given by the (DFT) of the first row of the matrix. When  $d > 1$ , the adjacency matrix of a DGG is no longer circulant but it is block circulant with  $N^{d-1} \times N^{d-1}$  circulant blocks, each of size  $N \times N$ . The author in [16], pages 85–87, utilizes the result in [25], and shows that the eigenvalues of the adjacency matrix in  $\mathbb{T}^d$  are found by taking the  $d$ -dimensional DFT of an  $N^d$  tensor of rank  $d$  obtained from the first block row of (A)

$$\lambda_{m_1, \dots, m_d} = \sum_{h_1, \dots, h_d=0}^{N-1} c_{h_1, \dots, h_d} \exp\left(-\frac{2\pi i}{N} \mathbf{m} \cdot \mathbf{h}\right), \tag{21}$$

where  $\mathbf{m}$  and  $\mathbf{h}$  are vectors of elements  $m_i$  and  $h_i$ , respectively, with  $m_1, \dots, m_d \in \{0, 1, \dots, N - 1\}$  and  $c_{h_1, \dots, h_d}$  defined as [16]

$$c_{h_1, \dots, h_d} = \begin{cases} 0, & \text{for } k_n < h_1, \dots, h_d \leq N - k_n - 1 \text{ or } h_1, \dots, h_d = 0, \\ 1, & \text{otherwise.} \end{cases} \tag{22}$$

The eigenvalues of the block circulant matrix  $\mathbf{A}$  follow the spectral decomposition [16], page 86,

$$\mathbf{A} = \mathbf{F}^H \mathbf{\Lambda} \mathbf{F},$$

where  $\mathbf{\Lambda}$  is a diagonal matrix whose entries are the eigenvalues of  $\mathbf{A}$ , and  $\mathbf{F}$  is the  $d$ -dimensional DFT matrix. It is well known that when  $d = 1$ , the DFT of an  $n \times n$  matrix is the matrix of the same size with entries

$$\mathbf{F}_{m,k} = \frac{1}{\sqrt{n}} \exp(-2\pi i m k / n) \quad \text{for } m, k = \{0, 1, \dots, n - 1\}.$$

When  $d > 1$ , the block circulant matrix  $\mathbf{A}$  is diagonalized by the  $d$ -dimensional DFT matrix  $\mathbf{F} = \mathbf{F}_{N_1} \otimes \dots \otimes \mathbf{F}_{N_d}$ , i.e., tensor product, where  $\mathbf{F}_{N_d}$  is the  $N_d$ -point DFT matrix.

Notice that all the matrices in (A) have a common eigenvector that is  $\left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$  and this eigenvector coincides with the first row and column of  $\mathbf{F}$ . Then,  $\left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$  is also an eigenvector of  $\hat{\mathcal{L}}(\mathcal{D}_n)$ .

The regularized normalized Laplacian can be expressed as

$$\begin{aligned} \hat{\mathcal{L}}(\mathcal{D}_n) &= \mathbf{I} - \frac{1}{(a'_n + \alpha)} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} - \frac{\alpha}{n(a'_n + \alpha)} \mathbf{1}\mathbf{1}^\top \\ &= \mathbf{F}^H \left( \mathbf{I} - \frac{1}{(a'_n + \alpha)} \mathbf{\Lambda} - \frac{\alpha}{n(a'_n + \alpha)} \mathbf{F}\mathbf{1}\mathbf{1}^\top \mathbf{F}^H \right) \mathbf{F} \\ &= \mathbf{F}^H \left( \mathbf{I} - \frac{1}{(a'_n + \alpha)} \mathbf{\Lambda} - \frac{n\alpha}{n(a'_n + \alpha)} \mathbf{e}_1^\top \mathbf{e}_1 \right) \mathbf{F} \\ &= \mathbf{F}^H \mathbf{\Lambda}_1 \mathbf{F}, \end{aligned} \tag{23}$$

where  $\mathbf{e}_1 = (1, 0 \dots 0)$  and  $\mathbf{\Lambda}_1 = \left( \mathbf{I} - \frac{1}{(a'_n + \alpha)} \mathbf{\Lambda} - \frac{n\alpha}{n(a'_n + \alpha)} \mathbf{e}_1 \mathbf{e}_1^\top \right)$  is a diagonal matrix whose diagonal elements are the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{D}_n)$ . Then, from (23), the derivation of the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{D}_n)$  reduces to the derivation of the eigenvalues of the normalized adjacency matrix  $\mathbf{A}'$ .

The normalized adjacency matrix  $\mathbf{A}'$  is defined as

$$\mathbf{A}' = \frac{1}{(a'_n + \alpha)} \mathbf{A}.$$

By using (21) and (22), the eigenvalues of  $\mathbf{A}'$  for a DGG in  $\mathbb{T}^d$  are given as

$$\begin{aligned} \lambda'_{m_1, \dots, m_d} &= \frac{1}{(a'_n + \alpha)} \left[ \sum_{h_1, \dots, h_d=0}^{N-1} \exp\left(-\frac{2\pi i \mathbf{m} \mathbf{h}}{N}\right) \right. \\ &\quad \left. - \sum_{h_1, \dots, h_d=k_n+1}^{N-k_n-1} \exp\left(-\frac{2\pi i \mathbf{m} \mathbf{h}}{N}\right) \right] - \frac{1}{(a'_n + \alpha)} \\ &= -\frac{1}{(a'_n + \alpha)} \sum_{h_1, \dots, h_d=k_n+1}^{N-k_n-1} \exp\left(-\frac{2\pi i \mathbf{m} \mathbf{h}}{N}\right) - \frac{1}{(a'_n + \alpha)} \\ \lambda'_{m_1, \dots, m_d} &= -\frac{1}{(a'_n + \alpha)} \prod_{s=1}^d \frac{\left( e^{-\frac{2im_s\pi}{N} k_n} - e^{\frac{2im_s\pi}{N} (1+k_n)} \right)}{\left( -1 + e^{\frac{2im_s\pi}{N}} \right)} - \frac{1}{(a'_n + \alpha)} \\ &= \frac{1}{(a'_n + \alpha)} \prod_{s=1}^d \frac{\left( e^{\frac{2im_s\pi}{N} (1+k_n)} - e^{-\frac{2im_s\pi}{N} k_n} \right)}{\left( -1 + e^{\frac{2im_s\pi}{N}} \right)} - \frac{1}{(a'_n + \alpha)} \\ &= \frac{1}{(a'_n + \alpha)} \prod_{s=1}^d \frac{\sin\left(\frac{m_s\pi}{N} (2k_n + 1)\right)}{\sin\left(\frac{m_s\pi}{N}\right)} - \frac{1}{(a'_n + \alpha)}. \end{aligned}$$

Then, we conclude that the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{D}_n)$  for  $n$  finite are given by

$$\begin{aligned} \hat{\lambda}_{m_1, \dots, m_d} &= 1 - \frac{1}{(a'_n + \alpha)} \prod_{s=1}^d \frac{\sin\left(\frac{m_s\pi}{N} (2k_n + 1)\right)}{\sin\left(\frac{m_s\pi}{N}\right)} + \frac{1 - \alpha \delta_{m_1, \dots, m_d}}{(a'_n + \alpha)} \\ \hat{\lambda}_{m_1, \dots, m_d} &= 1 - \frac{1}{(a'_n + \alpha)} \prod_{s=1}^d \frac{\sin\left(\frac{m_s\pi}{N} (a'_n + 1)^{1/d}\right)}{\sin\left(\frac{m_s\pi}{N}\right)} + \frac{1 - \alpha \delta_{m_1, \dots, m_d}}{(a'_n + \alpha)}, \end{aligned}$$

with  $m_1, \dots, m_d \in \{0, \dots, N - 1\}$  and  $\delta_{m_1, \dots, m_d} = 1$  when  $m_1, \dots, m_d = 0$  otherwise  $\delta_{m_1, \dots, m_d} = 0$ .

In particular, as  $\alpha \rightarrow 0$ , the eigenvalues of  $\mathcal{L}(\mathcal{D}_n)$  in the connectivity regime are given by

$$\lambda_{m_1, \dots, m_d} = 1 - \frac{1}{a'_n} \prod_{s=1}^d \frac{\sin(\frac{m_s \pi}{N} (a'_n + 1)^{1/d})}{\sin(\frac{m_s \pi}{N})} + \frac{1}{a'_n}, \quad (24)$$

with  $m_1, \dots, m_d \in \{0, \dots, N-1\}$ .

In the thermodynamic regime, for  $s \in \{0, \dots, d\}$  we let  $f_s = \frac{m_s}{N}$  then as  $n \rightarrow \infty$ ,  $f_s \in \mathbb{Q} \cap [0, 1]$  where  $\mathbb{Q}$  denotes the set of rational numbers. Therefore, for  $\gamma \geq 1$ , the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{X}_n)$  can be approximated by the eigenvalues of  $\hat{\mathcal{L}}(\mathcal{D}_n)$  given as

$$\hat{\lambda}_{f_1, \dots, f_d} = 1 - \frac{1}{(\gamma' + \alpha)} \prod_{s=1}^d \frac{\sin(\pi f_s (\gamma' + 1)^{1/d})}{\sin(\pi f_s)} + \frac{1 - \alpha \delta_{f_1, \dots, f_d}}{(\gamma' + \alpha)}, \quad (25)$$

where  $\gamma' = (2 \lfloor \gamma^{1/d} \rfloor + 1)^d - 1$  and  $\delta_{f_1, \dots, f_d} = 1$  when  $f_1, \dots, f_d = 0$ , otherwise  $\delta_{f_1, \dots, f_d} = 0$ .

## References

- Bai, Z., Silverstein, J.W.: Spectral Analysis of Large Dimensional Random Matrices. Springer, New York (2010)
- Bai, Z.D.: Methodologies in spectral analysis of large dimensional random matrices, a review. Stat. Sinica **9**, 611–677 (1999)
- Farkas, I.J., Derényi, I., Barabási, A.-L., Vicsek, T.: Spectra of “real-world” graphs: beyond the semi-circle law. Phys. Rev. E **64**(2), 026704 (2001)
- Van Mieghem, P.: Graph Spectra for Complex Networks. Cambridge University Press, Cambridge (2010)
- Couillet, R., Debbah, M.: Random Matrix Methods for Wireless Communications. Cambridge University Press, Cambridge (2011)
- Erdos, P.: On random graphs. Publicationes mathematicae **6**, 290–297 (1959)
- Haas, Z. J., Deng, J., Liang, B., Papadimitratos, P., Sajama, S.: “Wireless ad hoc networks,” *Encyclopedia of Telecommunications*, (2002)
- Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. Comput. Netw. **52**(12), 2292–2330 (2008)
- Preciado, V. M., Jadbabaie, A.: “Spectral analysis of virus spreading in random geometric networks,” *IEEE Conference on Decision and Control*, (2009)
- Ganesh, A., Massoulié, L., Towsley, D.: “The effect of network topology on the spread of epidemics,” in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, (2005)
- Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
- Maier, M., Hein, M., von Luxburg, U.: Optimal construction of k-nearest-neighbor graphs for identifying noisy clusters. Theor. Comput. Sci. **410**(19), 1749–1764 (2009)
- Sadri, A. M., Hasan, S., Ukkusuri, S. V., Lopez, J. E. S.: “Analyzing social interaction networks from twitter for planned special events,” *arXiv preprint arXiv:1704.02489*, (2017)
- Page, L., Brin, S., Motwani, R., Winograd, T.: “The pagerank citation ranking: Bringing order to the web.” Stanford InfoLab, Tech. Rep., (1999)
- Avrachenkov, K., Ribeiro, B., Towsley, D.: “Improving random walk estimation accuracy with uniform restarts,” in *International Workshop on Algorithms and Models for the Web-Graph*. Springer, pp. 98–109 (2010)
- Nyberg, A.: “The Laplacian spectra of random geometric graphs,” Ph.D. dissertation, University of Houston, (2014)

17. Nyberg, A., Gross, T., Bassler, K.E.: Mesoscopic structures and the Laplacian spectra of random geometric graphs. *J. Complex Netw.* **3**(4), 543–551 (2015)
18. Dettmann, C.P., Knight, G.: Symmetric motifs in random geometric graphs. *J. Complex Netw.* **6**(1), 95–105 (2017)
19. El Karoui, N.: The spectrum of kernel random matrices. *Ann. Stat.* **38**(1), 1–50 (2010)
20. Jiang, T.: Distributions of eigenvalues of large Euclidean matrices generated from lp balls and spheres. *Linear Algebra Appl.* **473**, 14–36 (2015)
21. Bordenave, C.: Eigenvalues of Euclidean random matrices. *Random Struct. Algorithms* **33**(4), 515–532 (2008)
22. Blackwell, P., Edmondson-Jones, M., Jordan, J.: *Spectra of adjacency matrices of random geometric graphs*. Unpublished, (2007)
23. Rai, S.: The spectrum of a random geometric graph is concentrated. *J. Theor. Probab.* **20**(2), 119–132 (2007)
24. Billingsley, P.: *Probability and Measure*. John Wiley & Sons, Hoboken (2008)
25. Gray, R.M.: “Toeplitz and circulant matrices: a review,” *Foundations and Trends®. Commun. Inf. Theory* **2**(3), 155–239 (2006)
26. Hoffman, A.J., Wielandt, H.: The variation of the spectrum of a normal matrix. *Duke Math. J.* **20**, 37–39 (1953)
27. Müller, T.: Two-point concentration in random geometric graphs. *Combinatorica* **28**(5), 529 (2008)
28. Janson, S., Luczak, T., Rucinski, A.: *Random Graphs*. John Wiley & Sons, Hoboken (2011)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.