Spectral Analysis of the Adjacency Matrix of Random Geometric Graphs

Mounia Hamidouche[*], Laura Cottatellucci[†], Konstantin Avrachenkov[◇]

[*] Departement of Communication Systems, EURECOM, Campus SophiaTech, 06410 Biot, France
[†] Department of Electrical, Electronics, and Communication Engineering, FAU, 51098 Erlangen, Germany
[◇] Inria, 2004 Route des Lucioles, 06902 Valbonne, France

mounia.hamidouche@eurecom.fr, laura.cottatellucci@fau.de, k.avratchenkov@inria.fr.

Abstract—**In this article, we analyze the limiting eigenvalue distribution (LED) of random geometric graphs (RGGs). The RGG is constructed by uniformly distributing n nodes on the d-dimensional torus $\mathbb{T}^d \equiv [0,1]^d$ and connecting two nodes if their ℓ_p-distance, $p \in [1, \infty]$ is at most r_n. In particular, we study the LED of the adjacency matrix of RGGs in the connectivity regime, in which the average vertex degree scales as $\log(n)$ or faster, i.e., $\Omega(\log(n))$. In the connectivity regime and under some conditions on the radius r_n, we show that the LED of the adjacency matrix of RGGs converges to the LED of the adjacency matrix of a deterministic geometric graph (DGG) with nodes in a grid as n goes to infinity. Then, for n finite, we use the structure of the DGG to approximate the eigenvalues of the adjacency matrix of the RGG and provide an upper bound for the approximation error.**

Index Terms—**Random geometric graphs, adjacency matrix, limiting eigenvalue distribution, Levy distance.**

I. Introduction

In recent years, random graph theory has been applied to model many complex real-world phenomena. A basic random graph used to model complex networks is the Erdös-Rényi (ER) graph [1], where edges between the nodes appear with equal probabilities. In [2], the author introduces another random graph called random geometric graph (RGG) where nodes have some random position in a metric space and the edges are determined by the position of these nodes. Since then, RGG properties have been widely studied [3].

RGGs are very useful to model problems in which the geographical distance is a critical factor. For example, RGGs have been applied to wireless communication network [4], sensor network [5] and to study the dynamics of a viral spreading in a specific network of interactions [6], [7]. Another motivation for RGGs in arbitrary dimensions is

multivariate statistics of high-dimensional data. In this case, the coordinates of the nodes can represent the attributes of the data. Then, the metric imposed by the RGG depicts the similarity between the data.

In this work, the RGG is constructed by considering a finite set $\mathcal{X}_n$ of n nodes, $x_1, ..., x_n$, distributed uniformly and independently on the d-dimensional torus $\mathbb{T}^d \equiv [0,1]^d$. We choose a torus instead of a cube in order to avoid boundary effects. Given a geographical distance, $r_n > 0$, we form a graph by connecting two nodes $x_i, x_j \in \mathcal{X}_n$ if their ℓ_p-distance, $p \in [1, \infty]$ is at most r_n, i.e., $\|x_i - x_j\|_p \leq r_n$, where $\|.\|_p$ is the ℓ_p-metric defined as

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

The RGG is denoted by $G(\mathcal{X}_n, r_n)$. Note that for the case $p = 2$ we obtain the Euclidean metric on $\mathbb{R}^d$. Typically, the function r_n is chosen such that $r_n \to 0$ when $n \to \infty$.

The degree of a vertex in $G(\mathcal{X}_n, r_n)$ is the number of edges connected to it. The average vertex degree in $G(\mathcal{X}_n, r_n)$ is given by [3]

$$a_n = \theta^{(d)} n r_n^d,$$

where $\theta^{(d)} = \pi^{d/2}/\Gamma(d/2 + 1)$ denotes the volume of the d-dimensional unit hypersphere in $\mathbb{T}^d$ and $\Gamma(.)$ is the Gamma function.

Different values of r_n, or equivalently a_n, lead to different geometric structures in RGGs. In [3], different interesting regimes are introduced: the *connectivity regime* in which a_n scales as $\log(n)$ or faster, i.e., $\Omega(\log(n))$[1], the *thermodynamic regime* in which $a_n \equiv \gamma$, for $\gamma > 0$ and the *dense regime*, i.e., $a_n \equiv \Theta(n)$.

RGGs can be described by a variety of random matrices such as adjacency matrices, transition probability matrices and normalized Laplacian. The spectral properties of those

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

random matrices are powerful tools to predict and analyze complex networks behavior. In this work, we give a special attention to the limiting eigenvalue distribution (LED) of the adjacency matrix of RGGs in the connectivity regime.

Some works analyzed the spectral properties of RGGs in different regimes. In particular, in the thermodynamic regime, the authors in [8], [9] show that the spectral measure of the adjacency matrix of RGGs has a limit as $n \to \infty$. However, due to the difficulty to compute exactly this spectral measure, Bordenave in [8] proposes an approximation for it as $\gamma \to \infty$.

In the connectivity regime, the work in [6] provides a closed form expression for the asymptotic spectral moments of the adjacency matrix of $G(\mathcal{X}_n, r_n)$. Additionnaly, Bordenave in [8] characterizes the spectral measure of the adjacency matrix normalized by n in the dense regime. However, in the connectivity regime and as $n \to \infty$, the normalization factor n puts to zero all the eigenvalues of the adjacency matrix that are finite and only the infinite eigenvalues in the adjacency matrix are nonzero in the normalized adjacency matrix. Motivated by this results, in this work we analyze the behavior of the eigenvalues of the adjacency matrix without normalization in the connectivity regime and in a wider range of the connectivity regime.

First, we propose an approximation for the actual LED of the RGG. Then, we provide a bound on the Levy distance between this approximation and the actual distribution. More precisely, for $\epsilon > 0$ we show that the LEDs of the adjacency matrices of the RGG and the deterministic geometric graph (DGG) with nodes in a grid converge to the same limit when a_n scales as $\Omega(\log^\epsilon(n)\sqrt{n})$ for $d = 1$ and as $\Omega(\log^2(n))$ for $d \geq 2$. Then, under the ℓ_∞-metric we provide an analytical approximation for the eigenvalues of the adjacency matrix of RGGs by taking the d-dimensional discrete Fourier transform (DFT) of an $n = \mathrm{N}^d$ tensor of rank d obtained from the first block row of the adjacency matrix of the DGG.

The rest of this paper is organized as follows. In Section II we describe the model, then we present our main results on the concentration of the LED of large RGGs in connectivity regime. Numerical results are given in Section III to validate the theoretical results. Finally, conclusions are given in Section IV.

II. **Spectral Analysis of RGGs**

To study the spectrum of $G(\mathcal{X}_n, r_n)$ we introduce an auxiliary graph called the DGG. The DGG denoted by $G(\mathcal{D}_n, r_n)$ is formed by letting $\mathcal{D}_n$ be the set of n grid points that are at the intersections of axes parallel hyperplanes with separation $n^{-1/d}$, and connecting two points x_i', $x_j' \in \mathcal{D}_n$ if $\|x_i' - x_j'\|_p \leq r_n$ with $p \in [1, \infty]$. Given two nodes, we assume that there is always at most one edge between them. There is no edge from a vertex to itself. Moreover, we assume that the edges are not directed.

Let $\mathbf{A}(\mathcal{X}_n)$ be the adjacency matrix of $G(\mathcal{X}_n, r_n)$, with entries

$$\mathbf{A}(\mathcal{X}_n)_{ij} = \chi[x_i \sim x_j],$$

where the term $\chi[x_i \sim x_j]$ takes the value 1 when there is a connection between nodes x_i and x_j in $G(\mathcal{X}_n, r_n)$ and zero otherwise, represented as

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

A similar definition holds for $\mathbf{A}(\mathcal{D}_n)$ defined over $G(\mathcal{D}_n, r_n)$. The matrices $\mathbf{A}(\mathcal{X}_n)$ and $\mathbf{A}(\mathcal{D}_n)$ are symmetric and their spectrum consists of real eigenvalues. We denote by $\{\lambda_i, i = 1, .., n\}$ and $\{\mu_i, i = 1, .., n\}$ the sets of all real eigenvalues of the real symmetric square matrices $\mathbf{A}(\mathcal{D}_n)$ and $\mathbf{A}(\mathcal{X}_n)$ of order n, respectively. The empirical spectral distribution functions $v_n(x)$ and $v_n'(x)$ of the adjacency matrices of an RGG and a DGG, respectively are defined as

$$v_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{\mu_i \leq x\} \quad \text{and} \quad v_n'(x) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{\lambda_i \leq x\},$$

where $\mathrm{I}\{\mathrm{B}\}$ denotes the indicator of an event B.

Let a_n' be the degree of the nodes in $G(\mathcal{D}_n, r_n)$. In the following Lemma 1 we provide an upper bound for a_n' under any ℓ_p-metric.

Lemma 1. *For any chosen ℓ_p-metric with $p \in [1, \infty]$ and $d \geq 1$, we have*

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

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

To prove our result on the concentration of the LED of RGGs and investigate its relationship with the LED of DGGs under any ℓ_p-metric, we use the Levy distance between two distribution functions defined as follows.

Definition 1. *([10], page 257) Let v_n^A and v_n^B be two distribution functions on $\mathbb{R}$. The Levy distance $L(v_n^A, v_n^B)$ between them is the infimum of all positive ϵ such that, for all $x \in \mathbb{R}$*

$$v_n^A(x - \epsilon) - \epsilon \leq v_n^B(x) \leq v_n^A(x + \epsilon) + \epsilon.$$

Lemma 2. *([11], page 614) Let A and B be two $n \times n$ Hermitian matrices with eigenvalues $\lambda_1, ..., \lambda_n$ and $\mu_1, ..., \mu_n$, respectively. Then*

$$L^3(v_n^A, v_n^B) \leqslant \frac{1}{n} tr(A - B)^2,$$

where $L(v_n^A, v_n^B)$ denotes the Levy distance between the empirical distribution functions v_n^A and v_n^B of the eigenvalues of A and B, respectively.

Let M_n be the minimum bottleneck matching distance corresponding to the minimum length such that there exists a perfect matching of the random nodes to the grid points for which the distance between every pair of matched points is at most M_n.

Sharp bounds for M_n are given in [12][13][14]. We repeat them in the following lemma for convenience.

Lemma 3. *Under any ℓ_p-norm, the bottleneck matching is*

- $M_n = O\left(\left(\dfrac{\log n}{n}\right)^{1/d}\right)$, *when $d \geq 3$ [12].*

- $M_n = O\left(\left(\dfrac{\log^{3/2} n}{n}\right)^{1/2}\right)$, *when $d = 2$ [13].*

- $M_n = O\left(\sqrt{\dfrac{\log \epsilon^{-1}}{n}}\right)$, *with prob. $\geq 1 - \epsilon$, $d = 1$ [14].*

Under the condition $M_n = o(r_n)$, we provide an upper bound for the Levy distance between v_n and v'_n in the following lemma.

Lemma 4. *For $d \geq 1$, $p \in [1, \infty]$ and $M_n = o(r_n)$, the Levy distance between v_n and v'_n is upper bounded as*

$$L^3\left(v_n, v'_n\right) \leq d^{\frac{1}{p}} 2^{d+1} \left| \frac{1}{n} \sum_i \mathbf{N}(x_i) - a_n \right|$$
$$+ d^{\frac{1}{p}} 2^{d+1} \left| a_n - \frac{2}{n} \sum_i L_i \right| + a'_n, \qquad (1)$$

where, $\mathbf{N}(x_i)$ denotes the degree of x_i in $G(\mathcal{X}_n, r_n)$ and $L_i \sim \text{Bin}\left(n, \theta^{(d)}\left(r_n - 2M_n\right)\right)$.

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

The condition enforced on r_n, i.e., $M_n = o(r_n)$ implies that for $\epsilon > 0$, (1) holds when a_n scales as $\Omega(\log^\epsilon(n)\sqrt{n})$ for $d = 1$, as $\Omega(\log^{\frac{3}{2}+\epsilon}(n))$ for $d = 2$ and as $\Omega(\log^{1+\epsilon}(n))$ for $d \geq 3$.

In what follows, we show that the LED of the adjacency matrix of $G(\mathcal{X}_n, r_n)$ concentrate around the LED of the adjacency matrix of $G(\mathcal{D}_n, r_n)$ in the connectivity regime in the sense of convergence in probability.

Notice that the term $\sum_i \mathbf{N}(x_i)/2$ in Lemma 4 counts the number of edges in $G(\mathcal{X}_n, r_n)$. For convenience, we denote $\sum_i \mathbf{N}(x_i)/2$ as ξ_n. To show our main result we apply the Chebyshev inequality given in Lemma 5 on the random variable ξ_n. For that, we need to determine $\text{Var}(\xi_n)$

Lemma 5. *(Chebyshev Inequality) Let X be a random variable with an expected value $\mathbb{E}X$ and a variance $\text{Var}(X)$. Then, for any $t > 0$*

$$\mathbb{P}\{|X - \mathbb{E}X| \geq t\} \leq \frac{\text{Var}(X)}{t^2}.$$

Lemma 6. *When $x_1, ..., x_n$ are i.i.d. uniformly distributed in the d-dimensional unit torus $\mathbb{T}^d = [0, 1]$*

$$\text{Var}\left(\xi_n\right) \leq [\theta^{(d)} + 2\theta^{(d)}a_n].$$

Proof. The proof follows along the same lines of Proposition A.1 in [15] when extended to a unit torus and applied to i.i.d. and uniformly distributed nodes. $\qquad\square$

We can now state the main theorem on the concentration of the adjacency matrix of $G(\mathcal{X}_n, r_n)$.

Theorem 1. *For $d \geq 1$, $p \in [1, \infty]$, $a \geq 1$, $M_n = o(r_n)$ and $t > 0$, we have*

$$\mathbb{P}\{L^3\left(v_n, v'_n\right) > t\} \leq 2n \exp\left(\frac{-a_n \varepsilon^2}{3}\left(1 - \frac{2M_n}{r_n}\right)\right)$$
$$+ \frac{n\left[\theta^{(d)}(r_n - 2M_n)(a-1) + 1\right]^n}{a^{\left(\frac{t}{d^{\frac{1}{p}} 2^{d+3}} + \frac{a_n(2-c)}{4}\right)}}$$
$$+ \frac{d^{\frac{2}{p}} 2^{2d+6} \left[\theta^{(d)} + 2\theta^{(d)}a_n\right]}{n^2 t^2}, \qquad (2)$$

where $\varepsilon = \left(\dfrac{t}{d^{\frac{1}{p}} 2^{d+2} a_n} + \dfrac{(2-c)}{4} - \dfrac{2M_n}{r_n}\right)$ *and* $c = \left(1 + \dfrac{1}{2a_n^{1/d}}\right)^d$.

In particular, for every $t > 0$, $a \geq 2$, $\epsilon > 0$ and a_n that scales as $\Omega(\log^\epsilon(n)\sqrt{n})$ when $d = 1$, as $\Omega(\log^2(n))$ when $d \geq 2$, we have

$$\lim_{n \to \infty} \mathbb{P}\left\{L^3\left(v_n, v'_n\right) > t\right\} = 0.$$

Proof. See Appendix C. $\qquad\square$

This result is shown in the sense of convergence in probability by a straightforward application of Lemma 8 and 9 on the random variable L_i, then by applying Lemma 5 and 6 to ξ_n.

In what follows, we provide the eigenvalues of $\mathbf{A}(\mathcal{D}_n)$ which approximates the eigenvalues of $\mathbf{A}(\mathcal{X}_n)$ for n sufficiently large.

Lemma 7. *For $d \geq 1$ and using the ℓ_∞-metric, the eigenvalues of $\mathbf{A}(\mathcal{D}_n)$ are given by*

$$\lambda_{m_1,...,m_d} = \prod_{s=1}^d \frac{\sin(\frac{m_s \pi}{N}(a'_n + 1)^{1/d})}{\sin(\frac{m_s \pi}{N})} - 1, \qquad (3)$$

where, $m_1, ..., m_d \in \{0, ...N-1\}$, $a'_n = (2k_n+1)^d - 1$, $k_n = \lfloor Nr_n \rfloor$ and $n = N^d$. The term $\lfloor x \rfloor$ is the integer part, i.e., the greatest integer less than or equal to x.

Proof. See Appendix D. $\qquad\square$

The proof utilizes the result in [16] which shows that the eigenvalues of the adjacency matrix of a DGG 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 $\mathbf{A}(\mathcal{D}_n)$.

For $\epsilon > 0$, Theorem 1 shows that when a_n scales as $\Omega(\log^\epsilon(n)\sqrt{n})$ for $d = 1$ and as $\Omega(\log^2(n))$ when $d \geq 2$, the LED of the adjacency matrix of an RGG concentrate around the LED of the adjacency matrix of a DGG as $n \to \infty$.
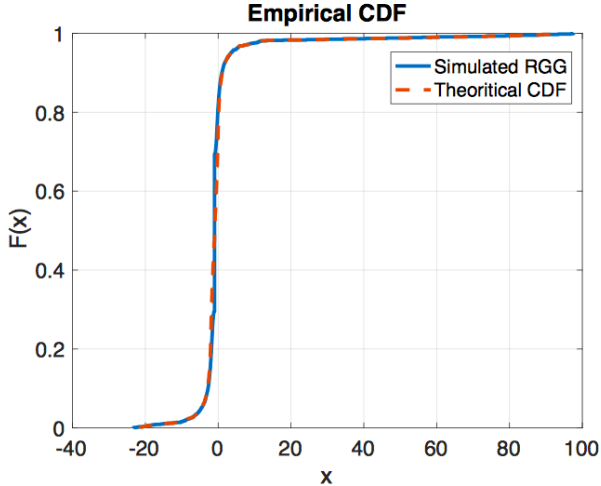
(a) Connectivity regime, $r_n = \frac{\log(n)}{\sqrt{n}}$, $n = 2000$.

Fig. 1. An illustration of the cumulative distribution function of the eigenvalues of an RGG.

Therefore, for n sufficiently large, the eigenvalues of the DGG given in (3) approximate very well the eigenvalues of the DGG.

III. **Numerical Results**

We present simulations to validate the results obtained in Section II. More specifically, we corroborate our results on the spectrum of the adjacency matrix of RGGs in the connectivity regime by comparing the simulated and the analytical results.

Fig. 1(a) shows the cumulative distribution functions of the eigenvalues of the adjacency matrix of an RGG realization and the analytical spectral distribution in the connectivity regime. We notice that for the chosen average vertex degree $a_n = \log(n)\sqrt{n}$ and $d = 1$, the curves corresponding to the eigenvalues of the RGG and the DGG fit very well for a large value of n.

IV. **Conclusion**

In this work, we study the spectrum of the adjacency matrix of RGGs in the connectivity regime. Under some conditions on the average vertex degree a_n, we show that the LEDs of the adjacency matrices of an RGG and a DGG converge to the same limit as $n \to \infty$. Then, based on the regular structure of the DGG, we approximate the eigenvalues of $\mathbf{A}(\mathcal{X}_n)$ by the eigenvalues of $\mathbf{A}(\mathcal{D}_n)$ by taking the d-dimensional DFT of an N^d tensor of rank d obtained from the first block row of $\mathbf{A}(\mathcal{D}_n)$.

V. **Acknowledgement**

REFERENCES

[1] P. Erdos, "On random graphs," *Publicationes Mathematicae*, vol. 6, pp. 290–297, 1959.

[2] E. N. Gilbert, "Random plane networks," *Journal of the Society for Industrial and Applied Mathematics*, vol. 9, no. 4, pp. 533–543, 1961.

[3] M. Penrose, *Random geometric graphs*. Oxford University Press, 2003.

[4] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proceedings of the 3rd ACM International Symposium on Mobile ad hoc Networking & computing*, 2002, pp. 80–91.

[5] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[6] V. M. Preciado and A. Jadbabaie, "Spectral analysis of virus spreading in random geometric networks," *IEEE Conference on Decision and Control*, 2009.

[7] A. Ganesh, L. Massoulié, and D. Towsley, "The effect of network topology on the spread of epidemics," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, 2005.

[8] C. Bordenave, "Eigenvalues of Euclidean random matrices," *Random Structures & Algorithms*, vol. 33, no. 4, pp. 515–532, 2008.

[9] P. Blackwell, M. Edmondson-Jones, and J. Jordan, *Spectra of adjacency matrices of random geometric graphs*. Unpublished, 2007.

[10] J. C. Taylor, *An introduction to measure and probability*. Springer Science & Business Media, 2012.

[11] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, vol. 9, pp. 611–677, 1999.

[12] P. W. Shor and J. E. Yukich, "Minimax grid matching and empirical measures," *The Annals of Probability*, pp. 1338–1348, 1991.

[13] F. T. Leighton and P. Shor, "Tight bounds for minimax grid matching, with applications to the average case analysis of algorithms," in *Proc. 18th Annual ACM Symposium on Theory of Computing*, 1986, pp. 91–103.

[14] A. Goel, S. Rai, and B. Krishnamachari, "Sharp thresholds for monotone properties in random geometric graphs," in *Proc. 36th Annual ACM Symposium on Theory of Somputing*, 2004, pp. 580–586.

[15] T. Müller, "Two-point concentration in random geometric graphs," *Combinatorica*, vol. 28, no. 5, p. 529, 2008.

[16] A. Nyberg, "The Laplacian spectra of random geometric graphs," Ph.D. dissertation, 2014.

[17] S. Rai, "The spectrum of a random geometric graph is concentrated," *Journal of Theoretical Probability*, vol. 20, no. 2, pp. 119–132, 2007.

[18] S. Janson, T. Luczak, and A. Rucinski, *Random graphs*. John Wiley & Sons, 2000.

[19] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends ® in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.

APPENDIX A
PROOF OF LEMMA 1

In this Appendix, we upper bound the vertex degree a'_n under any ℓ_p-metric, $p \in [1, \infty]$.

Assume that $G(\mathcal{X}_n, r_n)$ and $G(\mathcal{D}_n, r_n)$ are formed using the ℓ_∞-metric and let a_n and a_n' be their average vertex degree and vertex degree, respectively.

In this case, for a d-dimensional DGG with $n = \mathrm{N}^d$ nodes, the vertex degree a_n' is given by [16]

$$a_n' = (2k_n + 1)^d - 1, \quad \text{with} \quad k_n = \lfloor \mathrm{N}r_n \rfloor \quad \text{and} \quad n = \mathrm{N}^d.$$

Therefore, for $\theta^{(d)} \geq 2$ and $d \geq 1$, we have

$$a_n' = (2k_n + 1)^d - 1 \leq \frac{2^d a_n}{\theta^{(d)}} \left(1 + \frac{1}{2n^{1/d}r_n} \right)^d$$
$$\leq 2^d a_n \left(1 + \frac{1}{2a_n^{1/d}} \right)^d.$$

Now, let b_n' and b_n be the vertex degree and the average vertex degree in $G(\mathcal{D}_n, r_n)$ and $G(\mathcal{X}_n, r_n)$, respectively when using any ℓ_p-metric, $p \in [1, \infty]$. Notice that for any $p \in [1, \infty]$, we have

$$\|x\|_\infty \leq \|x\|_p.$$

Then, the number of nodes a_n' that falls in the ball of radius r_n is greater or equal than b_n', i.e., $a_n' \geq b_n'$. Hence,

$$b_n' \leq a_n' \leq 2^d a_n \left(1 + \frac{1}{2n^{1/d}r_n} \right)^d$$
$$= \frac{d^{1/p} 2^d a_n}{d^{1/p}} \left(1 + \frac{1}{2n^{1/d}r_n} \right)^d.$$

It remains to show the relation between b_n' and b_n.

Assume that the RGG is formed by connecting each two nodes when $d^{1/p}\|x_i - x_j\|_\infty \leq r_n$. This simply means that the graph is obtained using the ℓ_∞-metric with a radius equal to $\frac{r_n}{d^{1/p}}$. Then, the average vertex degree of this graph is $\frac{a_n}{d^{1/p}}$. In addition, we have

$$\|x\|_p \leq d^{\frac{1}{p}} \|x\|_\infty.$$

Therefore,

$$b_n' \leq d^{1/p} 2^d b_n \left(1 + \frac{1}{2n^{1/d}r_n} \right)^d.$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

APPENDIX B
PROOF OF LEMMA 4

In this Appendix, we upper bound the Levy distance between the distribution functions v_n and v_n'.

By a straightforward application of Lemma 2, we have

$$L^3 (v_n, v_n') \leq \frac{1}{n} \mathrm{Trace} \left[(\mathbf{A}(\mathcal{X}_n) - \mathbf{A}(\mathcal{D}_n)]^2 \right.$$
$$= \frac{1}{n} \sum_i \sum_j \left[\chi[x_i \sim x_j] - \chi[x_i' \sim x_j'] \right]^2$$
$$\stackrel{(a)}{=} \frac{1}{n} \sum_i \mathbf{N}(x_i) + a_n' - \frac{2}{n} \sum_i \mathbf{N}(x_i, x_i')$$
$$\stackrel{(b)}{\leq} \frac{1}{n} \sum_i \mathbf{N}(x_i) + a_n' - \frac{2}{n} \sum_i \mathrm{L}_i$$
$$\leq \left| \frac{1}{n} \sum_i \mathbf{N}(x_i) - \frac{2}{n} \sum_i \mathrm{L}_i \right| + a_n'$$
$$\leq d^{\frac{1}{p}} 2^{d+1} \left| \frac{1}{n} \sum_i \mathbf{N}(x_i) - \frac{2}{n} \sum_i \mathrm{L}_i \right| + a_n'$$
$$\leq d^{\frac{1}{p}} 2^{d+1} \left| \frac{1}{n} \sum_i \mathbf{N}(x_i) - a_n \right|$$
$$+ d^{\frac{1}{p}} 2^{d+1} \left| a_n - \frac{2}{n} \sum_i \mathrm{L}_i \right| + a_n'.$$

Step (a) follows from $\mathbf{N}(x_i) = \sum_j \chi[x_i \sim x_j]$ and $a_n' = \sum_j \chi[x_i' \sim x_j']$, and by defining $\mathbf{N}(x_i, x_i') = \sum_j \chi[x_i \sim x_j]\chi[x_i' \sim x_j']$. Step (b) follows from noticing that when $\|x_i - x_j\|_p \leq r_n - 2\mathrm{M}_n$, then $\|x_i' - x_j'\|_p \leq r_n$. So, all points within a radius of $r_n - 2\mathrm{M}_n$ of x_i map to the neighbors of x_i' [17]. Thus, $\mathbf{N}(x_i, x_i')$ is stochastically greater than the random variable $\mathrm{L}_i \sim \mathrm{Bin}(n, \theta^{(d)}(r_n - 2\mathrm{M}_n))$. $\qquad \square$

APPENDIX C
PROOF OF THEOREM 1

We provide an upper bound on the probability that the Levy distance between the distribution functions v_n and v_n' is higher than $t > 0$. The following lemmas are useful for the following studies.

Lemma 8. *(Chernoff Bound) Let* X *be a random variable. Then, for any* $t > 0$

$$\mathbb{P}\{\mathrm{X} \geq t\} \leq \frac{F(a)}{a^t},$$

where $F(a)$ *is the probability generating function and* $a \geq 1$.

Lemma 9. *([18], Corollary 2.3, page 27) If* $\mathrm{X} \in \mathrm{Bin}(n, p)$, $\mathbb{E}\mathrm{X} = np$ *and* $0 < \varepsilon \leq \frac{3}{2}$, *we have*

$$\mathbb{P}\{|\mathrm{X} - \mathbb{E}\mathrm{X}| \geq \varepsilon\mathbb{E}\mathrm{X}\} \leq 2 \exp(-\varepsilon^2 \mathbb{E}\mathrm{X}/3).$$

$$\mathbb{P}\left\{L^3\left(v_n,v_n'\right)>t\right\}\le\mathbb{P}\left\{d^{\frac{1}{p}}2^{d+1}\left|\frac{\sum\limits_i\mathbf{N}(x_i)}{n}-a_n\right|\right.$$

$$\left.+d^{\frac{1}{p}}2^{d+1}\left|a_n-\frac{2\sum\limits_i\mathrm{L}_i}{n}\right|+a_n'>t\right\}$$

$$\le\mathbb{P}\left\{\left|\sum_i\mathbf{N}(x_i)-na_n\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+2}}\right\}$$

$$+\mathbb{P}\left\{\left|na_n-2\sum_i\mathrm{L}_i\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+2}}-\frac{na_n'}{d^{\frac{1}{p}}2^{d+1}}\right\}.$$

Let

$$A=\mathbb{P}\left\{\left|\sum_i\mathbf{N}(x_i)-na_n\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+2}}\right\},$$

and

$$B=\mathbb{P}\left\{\left|na_n-2\sum_i\mathrm{L}_i\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+2}}-\frac{na_n'}{d^{\frac{1}{p}}2^{d+1}}\right\}.$$

We first upper bound the term A using Lemma 5 and 6.

$$A=\mathbb{P}\left\{\left|\sum_i\mathbf{N}(x_i)-na_n\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+2}}\right\}$$

$$=\mathbb{P}\left\{\left|\xi_n-\mathbb{E}\xi_n\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+3}}\right\}$$

$$\le\frac{d^{\frac{2}{p}}2^{2d+6}\mathrm{Var}(\xi_n)}{n^2t^2}\le\frac{d^{\frac{2}{p}}2^{2d+6}\left[\theta^{(d)}+2\theta^{(d)}a_n\right]}{n^2t^2}.$$

Next, we upper bound the term B.

$$B=\mathbb{P}\left\{\left|na_n-2\sum_i\mathrm{L}_i\right|>\frac{nt}{d^{\frac{1}{p}}2^{d+2}}-\frac{na_n'}{d^{\frac{1}{p}}2^{d+1}}\right\}$$

$$\le n\mathbb{P}\left\{a_n-2\mathrm{L}_i>\frac{t}{d^{\frac{1}{p}}2^{d+2}}-\frac{a_n'}{d^{\frac{1}{p}}2^{d+1}}\right\}$$

$$+n\mathbb{P}\left\{2\mathrm{L}_i-a_n>\frac{t}{d^{\frac{1}{p}}2^{d+2}}-\frac{a_n'}{d^{\frac{1}{p}}2^{d+1}}\right\}$$

$$\le n\mathbb{P}\left\{|a_n-\mathrm{L}_i|>\frac{t}{d^{\frac{1}{p}}2^{d+3}}-\frac{a_n'}{d^{\frac{1}{p}}2^{d+2}}+\frac{a_n}{2}\right\}$$

$$+n\mathbb{P}\left\{\mathrm{L}_i>\frac{t}{d^{\frac{1}{p}}2^{d+3}}-\frac{a_n'}{d^{\frac{1}{p}}2^{d+2}}+\frac{a_n}{2}\right\}$$

$$\overset{(a)}{\le}n\mathbb{P}\left\{|a_n-\mathrm{L}_i|>\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}\right\}$$

$$+n\mathbb{P}\left\{\mathrm{L}_i>\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}\right\}.$$

Step (a) follows by applying Lemma 1 and $c=\left(1+\frac{1}{2a_n^{1/d}}\right)^d$. Then,

$$B\le$$

$$n\mathbb{P}\left\{|\mathbb{E}\mathrm{L}_i-\mathrm{L}_i|>\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}-2\theta^{(d)}n\mathrm{M}_n\right\}$$

$$+n\mathbb{P}\left\{\mathrm{L}_i>\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}\right\}$$

$$\le n\mathbb{P}\left\{|\mathbb{E}\mathrm{L}_i-\mathrm{L}_i|>a_n\varepsilon\right\}$$

$$+n\mathbb{P}\left\{\mathrm{L}_i>\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}\right\},$$

where

$$\varepsilon=\left(\frac{t}{d^{\frac{1}{p}}2^{d+2}a_n}+\frac{(2-c)}{4}-\frac{2\mathrm{M}_n}{r_n}\right).$$

We continue by letting

$$B_1=\mathbb{P}\left\{|\mathbb{E}\mathrm{L}_i-\mathrm{L}_i|>a_n\varepsilon\right\}.$$

$$B_2=\mathbb{P}\left\{\mathrm{L}_i>\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}\right\}.$$

For n sufficiently large and consequently a_n sufficiently large, we have $1\le c<2$ and $0<\varepsilon\le\frac{3}{2}$. Therefore, by applying Lemma 9, we upper bound B_1 as

$$\mathbb{P}\left\{|\mathbb{E}\mathrm{L}_i-\mathrm{L}_i|>a_n\varepsilon\right\}$$

$$\le\mathbb{P}\left\{|\mathbb{E}\mathrm{L}_i-\mathrm{L}_i|>(a_n-2n\theta^{(d)}\mathrm{M}_n)\varepsilon\right\}$$

$$\le2\exp\left(\frac{-\varepsilon^2}{3}\left(a_n-2n\theta^{(d)}\mathrm{M}_n\right)\right).$$

The last term B_1 is upper bounded by using the Chernoff bound in Lemma 8.

The probability generating function of the binomial random variable L_i is given by

$$\left[a\theta^{(d)}(r_n-2\mathrm{M}_n)+1-\theta^{(d)}(r_n-2\mathrm{M}_n)\right]^n.$$

Therefore, for n sufficiently large, $1\le c<2$ and $a\ge1$, we have

$$B_2\le\frac{\left[\theta^{(d)}(r_n-2\mathrm{M}_n)(a-1)+1\right]^n}{a^{\left(\frac{t}{d^{\frac{1}{p}}2^{d+3}}+\frac{a_n(2-c)}{4}\right)}}.$$

Finally, taking the upper bounds of A and B obtained from the upper bounds of B_1 and B_2 all together, Theorem 1 follows. $\square$

In this appendix, we provide the eigenvalues of the adjacency matrix of the DGG using the ℓ_∞-metric.

When $d = 1$, the adjacency matrix $\mathbf{A}(\mathcal{D}_n)$ of a DGG in $\mathbb{T}^1$ with n nodes is a circulant matrix. A well known result appearing in [19], states that the eigenvalues of a circulant matrix are given by the discrete Fourier transform (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 $\mathrm{N}^{d-1} \times \mathrm{N}^{d-1}$ circulant blocks, each of size $\mathrm{N} \times \mathrm{N}$. The author in [16], pages 85-87, utilizes the result in [19], 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 the matrix $\mathbf{A}(\mathcal{D}_n)$

$$\lambda_{m_1,\ldots,m_d} = \sum_{h_1,\ldots,h_d=0}^{\mathrm{N}-1} c_{h_1,\ldots,h_d} \exp\left(-\frac{2\pi i}{\mathrm{N}}\mathbf{m}.\mathbf{h}\right), \quad (4)$$

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

$$c_{h_1,\ldots,h_d} = \begin{cases} 0, & \text{for } k_n < h_1,\ldots,h_d \leq \mathrm{N}-k_n-1 \\ & \text{or } h_1,\ldots h_d = 0, \\ 1, & \text{otherwise.} \end{cases}$$
$$(5)$$

The eigenvalues of the block circulant matrix $\mathbf{A}(\mathcal{X}_n)$ 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 $\mathbf{A}(\mathcal{X}_n)$, 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

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

When $d > 1$, the block circulant matrix $\mathbf{A}$ is diagonalized by the d-dimensional DFT matrix $\mathbf{F} = \mathbf{F}_{\mathrm{N}_1} \otimes \ldots \otimes \mathbf{F}_{\mathrm{N}_d}$, i.e., tensor product, where $\mathbf{F}_{\mathrm{N}_d}$ is the N_d-point DFT matrix.

Using (4) and (5), the eigenvalues of $\mathbf{A}(\mathcal{D}_n)$ in $\mathbb{T}^d$ are given by

$$\lambda_{m_1,\ldots,m_d} = \left[\sum_{h_1,\ldots,h_d=0}^{\mathrm{N}-1} \exp\left(-\frac{2\pi i m \mathbf{h}}{\mathrm{N}}\right) \right.$$
$$\left. - \sum_{h_1,\ldots,h_d=k_n+1}^{\mathrm{N}-k_n-1} \exp\left(-\frac{2\pi i m \mathbf{h}}{\mathrm{N}}\right) \right] - 1$$

$$= \sum_{h_1,\ldots,h_d=k_n+1}^{\mathrm{N}-k_n-1} \exp\left(-\frac{2\pi i m \mathbf{h}}{\mathrm{N}}\right) - 1$$

$$= \prod_{s=1}^{d} \frac{\left(e^{\frac{-2im_s\pi}{\mathrm{N}}k_n} - e^{\frac{2im_s\pi}{\mathrm{N}}(1+k_n)}\right)}{\left(-1 + e^{\frac{2im_s\pi}{\mathrm{N}}}\right)} - 1$$

$$= \prod_{s=1}^{d} \frac{\left(e^{\frac{2im_s\pi}{\mathrm{N}}(1+k_n)} - e^{\frac{-2im_s\pi}{\mathrm{N}}k_n}\right)}{\left(-1 + e^{\frac{2im_s\pi}{\mathrm{N}}}\right)} - 1$$

$$= \prod_{s=1}^{d} \frac{\sin(\frac{m_s\pi}{\mathrm{N}}(2k_n+1))}{\sin(\frac{m_s\pi}{\mathrm{N}})} - 1$$

$$= \prod_{s=1}^{d} \frac{\sin(\frac{m_s\pi}{\mathrm{N}}(a_n'+1)^{1/d})}{\sin(\frac{m_s\pi}{\mathrm{N}})} - 1.$$

$\square$