Spectral bounds of the regularized normalized Laplacian for random geometric graphs

Mounia Hamidouche*, Laura Cottatellucci†, Konstantin Avrachenkov◇

* Departement of communication systems, EURECOM, France

† Department of Electrical, Electronics, and Communication Engineering, FAU, Germany

◇ INRIA Sophia Antipolis, France.

Abstract—In this work, we study the spectrum of the regularized normalized Laplacian for random geometric graphs (RGGs) in both the connectivity and thermodynamic regimes. We prove that the limiting eigenvalue distribution (LED) of the normalized Laplacian matrix for an RGG converges to the Dirac measure in one in the full range of the connectivity regime. In the thermodynamic regime, we propose an approximation for the LED and we provide a bound on the Levy distance between the approximation and the actual distribution. In particular, we show that the LED of the regularized normalized Laplacian matrix for an RGG can be approximated by the LED of the regularized normalized Laplacian for a deterministic geometric graph with nodes in a grid (DGG). Thereby, we obtain an explicit approximation of the eigenvalues in the thermodynamic regime.

I. Introduction

THE spectrum of the normalized Laplacian associated to random geometric graphs (RGGs) provides very useful insights on the structure and the dynamics of large complex networks in which the geographical distance is a critical factor.

An RGG is a graph with a finite set $\mathcal{X}_n$ of n points $x_1, ..., x_n$, distributed uniformly and independently on a d-dimensional torus $\mathbb{T}^d \equiv [0,1]^d$. Given a geographical distance $r_n > 0$, we form a graph by connecting two points $x_i, x_j \in \mathcal{X}_n$ if the distance between them is at most r_n, i.e., $\mathrm{d}\,(x_i, x_j) \le r_n$. The maximum distance r_n is a function of n chosen such that $r_n \to 0$ when $n \to \infty$. The average vertex degree in $G(\mathcal{X}_n, r_n)$ is $a_n = \theta^{(d)} n r_n^d$, where $\theta^{(d)}$ denotes the volume of a d-dimensional unit hypersphere in $\mathbb{T}^d$ [1]. An RGG presents different properties depending on the average vertex degree a_n. Two different scaling regimes for a_n are of interest. The first one is the connectivity regime, in which a_n scales as $\Omega(\log(n))$[1] [1]. The second scaling regime of interest is the thermodynamic regime in which $a_n \equiv \gamma$, for $\gamma > 0$.

In [2], the author shows that the limiting eigenvalue distributions (LEDs) of the transition probability matrix of random walks in an RGG and in a deterministic geometric graph with nodes in a grid (DGG) converge to the same limit in the connectivity regime. However, for $\epsilon > 0$ the result in [2] holds only when a_n scales as $\Omega\left(\sqrt{n}\log^\epsilon(n)\right)$ for $d = 1$, as $\Omega\left(\log^{\frac{3}{2}+\epsilon}(n)\right)$ for $d = 2$, and as $\Omega\left(\log^{1+\epsilon}(n)\right)$ for $d \ge 3$. Motivated by the result in [2], here we study the LED of

[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 $\forall n > n_0$ $f(n) \ge Kg(n)$.

the regularized normalized Laplacian for an RGG in the full range of the connectivity regime, i.e, $a_n = \Omega(\log(n))$, and in the thermodynamic regime. To the best of our knowledge, an explicit expression for the LED of the normalized Laplacian for RGGs is 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.

Using the above defined RGG, we show that the LEDs of the normalized Laplacian for an RGG and for a DGG converge to same limit for any dimension $d \ge 1$ as $n \to \infty$ in the full range of the connectivity regime. Then, we show that their corresponding LEDs converge to the Dirac measure in one as $n \to \infty$ and $d = 1$. In the thermodynamic regime, we provide an analytical approximation for the eigenvalues of the regularized normalized Laplacian for the RGG when $d = 1$. Then, we numerically validate our theoretical results by comparing the simulated and the analytical spectum.

II. Spectral Analysis

In order to study the LED of RGGs, we consider a finite set $\mathcal{D}_n$ of n grid points that are at the intersections of all parallel hyperplanes with separation $n^{-1/d}$. Then, we define a deterministic geometric graph in a grid $G(\mathcal{D}_n, r_n)$ by connecting two points in $\mathcal{D}_n$ when the distance between them is at most r_n. Given two nodes in $G(\mathcal{X}_n, r_n)$ or in $G(\mathcal{D}_n, r_n)$, we assume that there is at most one edges between them and there is no edge from a vertex to itself. We denote the degree of a vertex in $G(\mathcal{D}_n, r_n)$ by a'_n, and in particular $a'_n \equiv \gamma'$ in the thermodynamic regime. Let $\mathbf{N}(x_i)$ be the number of neighbors of a vertex x_i in $G(\mathcal{X}_n, r_n)$. Define $\hat{\mathcal{L}}(\mathcal{X}_n)$ as the regularized normalized Laplacian matrix for $G(\mathcal{X}_n, r_n)$ with entries,

$$\hat{\mathcal{L}}(\mathcal{X}_n)_{ij} = \delta_{ij} - \frac{\chi[x_i \sim x_j] + \frac{\alpha}{n}}{\sqrt{(\mathbf{N}(x_i) + \alpha)(\mathbf{N}(x_j) + \alpha)}}. \quad (1)$$

The term $\chi[x_i \sim x_j]$ takes unit value when there is an edge between nodes i and j in $G(\mathcal{X}_n, r_n)$ and zero otherwise. δ_{ij} is the Kronecker delta function and $\frac{\alpha}{n} \ge 0$. A similar definition holds for $\hat{\mathcal{L}}(\mathcal{D}_n)_{ij}$, the regularized normalized Laplacian defined over the nodes in $G(\mathcal{D}_n, r_n)$. The regularized version of the normalized Laplacian matrix is used to overcome the problem of singularities due to isolated vertices in the thermodynamic regime [3].

In the following Theorem 1 and 2, we show that the empirical spectral distribution function $F^{\hat{\mathcal{L}}(\mathcal{D}_n)}$ is a good approximation of the spectral distribution $F^{\hat{\mathcal{L}}(\mathcal{X}_n)}$ when n

is large in both the connectivity and thermodynamic regimes using the Levy distance L^3 [4].

Theorem 1. *In the connectivity regime, i.e., as $a_n = \Omega(\log(n))$, when $d \geq 1$, $a_n \geq 2d$, $\alpha \to 0$ and $n \to \infty$ then, for any $t > 0$,*

$$\lim_{n \to \infty} P\left\{L^3\left(F^{\hat{\mathcal{L}}(\mathcal{X}_n)}, F^{\hat{\mathcal{L}}(\mathcal{D}_n)}\right) > t\right\} = 0.$$

In particular, when $a_n = c\log(n)$, $c > 24$ then, the LED of the normalized Laplacian for an RGG converges to the LED of the normalized Laplacian for a DGG with rate of convergence $\mathcal{O}\left(1/n^{c/12-1}\right)$, and 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 $\mathcal{O}\left(1/n^{(a_n/12\log(n))-1}\right)$. An exponential rate of convergence $\mathcal{O}\left(ne^{-n/12}\right)$ holds when the graph is dense, i.e., a_n scales as $\Omega(n)$. Hence, the result given in Theorem 1 generalizes the work in [2] and shows that the LEDs of the normalized Laplacian of $G(\mathcal{X}_n, r_n)$ and $G(\mathcal{D}_n, r_n)$ converge to the same limit as $n \to \infty$ in the full range of the connectivity regime and for any chosen distance function.

Theorem 2. *In the thermodynamic regime, i.e., as $a_n \equiv \gamma$, for $d \geq 1$, $\gamma \geq 2d$ and for every $t > \dfrac{8\gamma}{(\gamma' + \alpha)^2}$*

$$\lim_{n \to \infty} P\left\{L^3\left(F^{\hat{\mathcal{L}}(\mathcal{X}_n)}, F^{\hat{\mathcal{L}}(\mathcal{D}_n)}\right) > t\right\} = 0.$$

From Theorem 2 it is apparent that, as $n \to \infty$ and $\alpha \to 0$, $F^{\hat{\mathcal{L}}(\mathcal{D}_n)}$ approximates $F^{\hat{\mathcal{L}}(\mathcal{X}_n)}$ with an error bound of $8\gamma/\gamma'^2$. Now, we provide the LED of the normalized Laplacian in the connectivity regime as $n \to \infty$. By utilizing Theorem 1 and the expression for the eigenvalues of the adjacency matrix of a DGG in [5], we can state the following Lemma 1.

Lemma 1. *In the connectivity regime, for $d = 1$, $\alpha \to 0$ and using the Euclidean distance, the eigenvalues λ_m of $\hat{\mathcal{L}}(\mathcal{D}_n)$ are obtained from*

$$\lambda_m = 1 - \frac{1}{a'_n}\left(\frac{\sin(\frac{m\pi}{n}(a'_n + 1))}{\sin(\frac{m\pi}{n})} - 1\right), \quad (2)$$

where $m \in \{0, ...n-1\}$, $a'_n = 2\lfloor nr_n \rfloor$ and $\lfloor x \rfloor$ is the integer part, i.e., the greatest integer less than or equal to x. Then, as $n \to \infty$, the LED of $\hat{\mathcal{L}}(\mathcal{D}_n)$ converges to the Dirac measure in one.

Now, we give an approximation of the eigenvalues for the regularized normalized Laplacian of an RGG in the thermodynamic regime as $n \to \infty$.
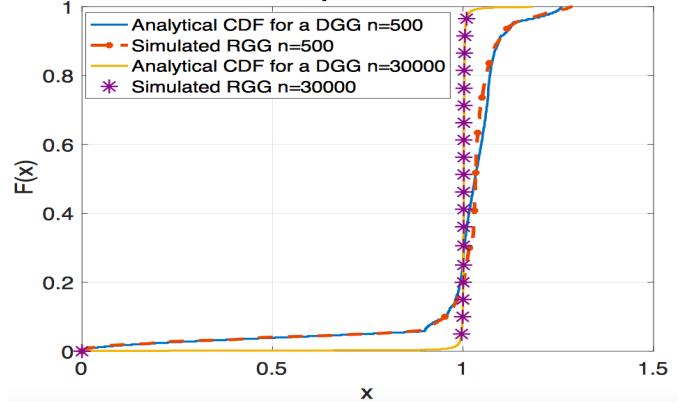
Lemma 2. *In the thermodynamic regime, for $d = 1$ and using the Euclidean distance, the asymptotic eigenvalues of $\hat{\mathcal{L}}(\mathcal{D}_n)$, as $n \to \infty$ are given by*

$$\lambda_w = 1 - \frac{1}{(\gamma' + \alpha)}\frac{\sin(w(\gamma' + 1))}{\sin(w)} + \frac{1 - \alpha\delta_w}{(\gamma' + \alpha)}, \quad (3)$$
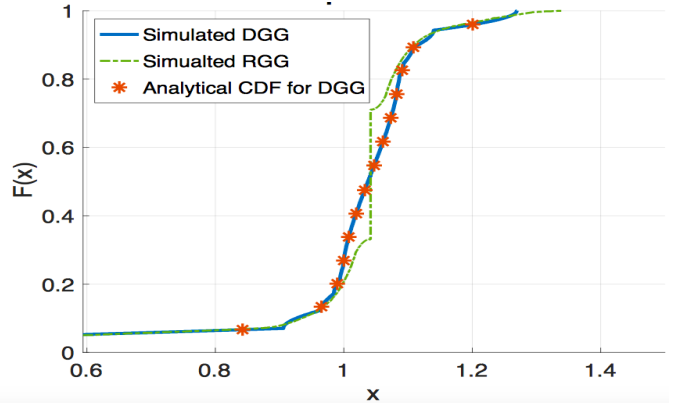
where $w \in [0, \pi]$, $\gamma' = 2\lfloor \gamma \rfloor$ for $\gamma \geq 2$ and $\delta_w = 1$ when $w = 0$ otherwise $\delta_w = 0$.

In particular, when $\alpha \to 0$ then, (3) reduces to

$$\lambda_w = 1 - \frac{1}{\gamma'}\left(\frac{\sin(w(\gamma' + 1))}{\sin(w)} - 1\right). \quad (4)$$

(a) Connectivity regime for different n and $r_n = \log^{\frac{3}{2}}(n)/n$.

(b) Thermodynamic regime for $\gamma = 12$ and $r_n = 12/n$.

Figure 1. Spectral distribution of an RGG and a DGG.

III. NUMERICAL EVALUATION

We validate our analytical results obtained in Section II on the spectrum of the regularized normalized Laplacian of RGGs by numerical computations. In Fig. 1.a we compare the spectral distribution of a DGG (continuous line) with the one of an RGG with increasing the 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. This confirms the result given in Theorem 1. It appears that by increasing n, the eigenvalue distribution converges to the Dirac measure in one as shown in Lemma 1. Fig. 1.b illustrates the empirical spectral distribution of a simulated regularized normalized Laplacian in the thermodynamic regime for a realization of an RGG and DGG with $n = 30000$ nodes and $\alpha = 0.001$. The analytical distribution is obtained using Lemma 2. The gap that appears between the eigenvalue distributions of the RGG and the DGG is quantified by the error bound in Theorem 2.

REFERENCES

[1] M. Penrose, *Random geometric graphs*. Oxford University Press, 2003.
[2] S. Rai, "The spectrum of a random geometric graph is concentrated," *Journal of Theoretical Probability*, vol. 20, no. 2, pp. 119–132, 2007.
[3] K. Avrachenkov, B. Ribeiro, and D. Towsley, "Improving random walk estimation accuracy with uniform restarts," in *International Workshop on Algorithms and Models for the Web-Graph*. Springer, 2010, pp. 98–109.
[4] J. C. Taylor, *An introduction to measure and probability*. Springer Science & Business Media, 2012.
[5] A. Nyberg, "The Laplacian spectra of random geometric graphs," Ph.D. dissertation, 2014.