

# IMPERIAL

IMPERIAL COLLEGE LONDON

DEPARTMENT OF MATHEMATICS

MSci RESEARCH PROJECT

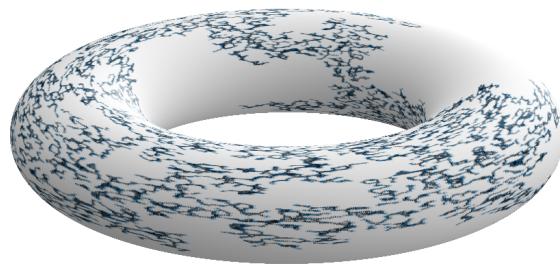
---

## **$Tr[L^+]$ as a Percolation Signature for Random Geometric Graphs**

---

*Author:*  
Ross Ah-Weng

*Supervisor(s):*  
Dr. Gunnar Pruessner, Emir Sezik



Submitted in partial fulfillment of the requirements for the MSci Mathematics at Imperial College London

June 9, 2025

## Abstract

Random Geometric Graphs (RGGs) model spatially embedded networks, where edges form based on proximity of nodes. In the thermodynamic limit, as the connection radius increases, RGGs exhibit a percolation transition from disconnected clusters to a giant connected component. Although this phenomenon has been studied in probability theory, no existing work has identified a signature of the percolation transition in the spectral statistics of the graph Laplacian, which belongs to an ensemble of Euclidean Random Matrices. This thesis addresses this gap by introducing a principled extension of the Kirchhoff index, a graph invariant based on effective resistance, to disconnected graphs. We demonstrate that this observable diverges at the percolation transition, and extract critical parameters from its finite-size scaling, matching known values. We then use tools from Random Matrix Theory to derive an analytical expression of its ensemble average, formulate a motif expansion for the spectral density of sparse RGGs, and investigate crossover behaviour in the soft kernel case.

## **Plagiarism statement**

The work contained in this thesis is my own work unless otherwise stated.

*Signature:* Ross Ah-Weng

*Date:* June 9, 2025

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Background</b>	<b>5</b>
2.1	Percolation Theory and Critical Phenomena	5
2.2	The Random Geometric Graph	6
2.3	The Kirchhoff Index	8
2.4	Random Matrix Theoretic Ingredients	10
<b>3</b>	<b>Methods</b>	<b>11</b>
3.1	Random Geometric Graph Generation	11
3.2	Finite-Size Scaling	14
<b>4</b>	<b>Results</b>	<b>16</b>
4.1	Conventional Statistics of the Laplacian	16
4.1.1	Trace	16
4.1.2	Dimension of the Null Space	17
4.1.3	Smallest Non-Zero Eigenvalue	18
4.2	$\text{Tr}[L^+]$ as a signature of the percolation transition	19
4.2.1	The Hard RGG	19
4.2.2	The Soft RGG	21
4.3	Connection to Euclidean Random Matrices	23
4.3.1	An analytical expression for $\langle \text{Tr}(L^+) \rangle$	23
4.3.2	Motif Expansion of the Hard RGG spectrum	24
4.3.3	Soft RGGs and Gaussian Euclidean Random Matrices	25
<b>5</b>	<b>Conclusion</b>	<b>26</b>
<b>A</b>	<b>Appendix</b>	<b>27</b>
A.1	Proof of Equation 4.6	27
A.2	Proof of Lemma 4.3.2	28
A.3	GitHub Repository	28

# Chapter 1

## Introduction

Consider a network of wireless communication stations scattered across the country, and suppose that each station can only communicate with others within a fixed physical distance. Can a message travel from one end of the country to another by hopping between stations? Or consider the spread of an infection in a forest, where trees can infect others within some infection distance - will the disease remain localised, or spread through the entire area? More abstractly, picture a cloud of data points, all related to each other via some similarity metric, and we want to study how tuning its threshold can partition the network into meaningful classes.

These are classical examples that motivate the study of Random Geometric Graphs (RGGs), first introduced by Gilbert as the Gilbert Disk Model [1]. By connecting points that lie within a fixed radius of each other, the RGG arises as a natural model of spatially embedded networks. RGGs and its variants have been studied under various names across many fields; namely the Poisson blob model [2], Continuum Percolation [3], Single-Linkage Clustering [4], Vietoris-Rips Complexes [5] and Gabriel graphs [6].

Percolation theory studies the emergence of large-scale connectivity in random systems. Percolation on the RGG has been studied extensively in probability theory, notably in the book Continuum Percolation by Meester and Roy [3], and extensively by Penrose, summarised in his monograph [7]. Numerical estimates of critical densities and exponents have been available for some time [8], and there is conjectural evidence that some variants share the same universality class as lattice percolation models [9].

A related object of study are Euclidean Random Matrices (ERMs), a matrix ensemble that encodes the spatial distances between randomly distributed points. First introduced by Mezard et al. [10], ERMs find application in modelling phenomena in disordered systems, such as vibrations in glasses [11] and random lasers [12]. Despite this, they are significantly less understood than classical Gaussian and Wishart ensembles.

In particular, the adjacency and Laplacian matrix of a RGG are special cases of ERMs, and their spectral properties encode graph structure and connectivity. While some existing work connects these two fields [13–15], no existing work identifies a matrix signature of the percolation transition in the thermodynamic limit. This thesis introduces a principled extension of the Kirchhoff Index to disconnected systems, which diverges at criticality in RGGs on both sides of the transition. We numerically investigate its scaling properties using finite-size scaling and data collapse, extracting the critical radius and exponents, that align with the literature. We finally derive an analytical expression for its ensemble average in random matrix theoretical language, positing a geometrical approximation for the spectrum of the sparse Hard RGG, and numerically explore the emergence of a Boson-like peak as  $r \rightarrow 0$  in the Soft RGG.

# Chapter 2

## Background

### 2.1 Percolation Theory and Critical Phenomena

In this section, consider Bernoulli Bond percolation on the square lattice  $\mathbb{Z}^2$ , where each edge is independently occupied with probability  $p$ , which acts as a control parameter. As  $p$  increases, the system undergoes a phase transition at the *Critical Probability*  $p_c$ , from a fragmented state consisting of many disjoint clusters, to a connected state with a giant component.

Formally, denote  $P_\infty(p)$  as the probability that the cluster containing the origin is infinite; then we define  $p_c = \inf_{p \in [0,1]} P_\infty(p) > 0$ . Construct the *Increasing Coupling* by assigning i.i.d. weights  $w_{ij} \sim U(0, 1)$  to every edge in the lattice, and defining the configuration  $\eta_p$  by letting each edge be occupied if  $w_{ij} \leq p$  [2]. Then  $\eta_p$  is equal in law to Bernoulli Bond percolation, and the family of configurations  $\{\eta_p\}_{p \in [0,1]}$  are coupled such that  $\forall p < q$ ,  $\eta_p$  is a subgraph of  $\eta_q$ . It follows that  $P_\infty(p)$  is non-decreasing.

It can be shown using Peierls-type arguments that for  $p < p_c$ ,  $P_\infty(p) = 0$  [2]. Then the derivative of  $P_\infty(p)$  w.r.t  $p$  is discontinuous at  $p_c$  - this is characteristic of a continuous phase transition. Near the critical point in the supercritical regime,  $p > p_c$ , we have the scaling relation  $P_\infty \sim (p - p_c)^\beta$ , where  $\beta = \frac{5}{36}$  [16]. More broadly, the system displays critical behaviour at  $p_c$ , where many macroscopic observables are governed by power laws, each with associated critical exponents.

Define the two-point correlation function  $G(x)$  as the probability that there exists a path of occupied bonds from  $\mathbf{0}$  to  $x \in \mathbb{Z}^2$ . In the subcritical regime,  $p < p_c$ , we  $G(x)$  decays exponentially with distance,  $G(x) = e^{-\frac{\|x\|_2}{\xi}}$ , where we define the *Correlation Length*  $\xi$ , which measures the characteristic length scale of the system. Near  $p_c$ , the correlation length diverges as a power law  $\xi \sim |p - p_c|^{-\nu}$ , where  $\nu = \frac{4}{3}$ . Further, the number of clusters per site, near  $p_c$  satisfies  $n_c \sim |p - p_c|^{2-\alpha}$ , where  $\alpha = \frac{2}{3}$ .

While many other order parameters and associated critical exponents exist, we note in particular the conductivity  $\Sigma$ . In a finite system of size  $L$ , resistors of unit conductance are placed at each occupied edge, then the *Effective Conductivity*  $\Sigma_L$  is defined as the current flow induced by a unit potential difference between opposite boundaries of the system. Note that a current only flows given the existence of a spanning cluster connecting the two boundaries. Taking the thermodynamic limit as  $L \rightarrow \infty$ , this converges to the bulk conductivity  $\Sigma_L \rightarrow \Sigma$ . Which for  $p > p_c$ , where an infinite cluster exists, we have the scaling relation  $\Sigma \sim |p - p_c|^t$  near  $p_c$  [16], where  $t$  is the *Conductivity Exponent*.

At the critical point, the percolating cluster becomes scale-invariant, meaning it lacks a characteristic lengthscale and has self-similar structure. For a wide class of lattices, the macroscopic

behaviour of the system becomes independent of its microscopic properties, and shares the same critical exponents and scaling relations. This holds true even for certain different percolation mechanisms; we classify these systems under the (*Standard*) *Percolation Universality Class*.

Classical 2D percolation lattice models study planar graphs, which admit a dual graph representation. Much analysis stems from this approach, for example the relation for certain lattices  $p_c + p_c^* = 1$ , where  $p_c^*$  is the critical probability on the dual graph [17]. Beyond lattice models, percolation in more general scenarios have been studied; Continuum Percolation studies off-lattice systems, and infinite-range models study systems where microscopic interactions can span arbitrary long distances.

## 2.2 The Random Geometric Graph

Fix  $N \in \mathbb{N}$ , and a density  $\rho \in \mathbb{R}^+$ , and let  $L := \sqrt{\frac{N}{\rho}}$ . Let  $\square_L := [-L/2, L/2]^2$  denote the two-dimensional square of side length  $L$ , centred at the origin.

**Definition 1.** We define a Point Configuration  $X_\rho^N$ :

$$\begin{aligned} X_\rho^N &:= \{x_i\}_{i=1}^N \subset \square_L, \\ x_i &\sim \text{Unif}(\square_L) \quad i.i.d \end{aligned} \tag{2.1}$$

This defines a Binomial point process of  $N$  points at uniform density on  $\square_L$ .

**Definition 2.** Define the Toroidal Distance  $d_T : \square_L \times \square_L \rightarrow \mathbb{R}^+$

$$d_T(x, y) = \left( \sum_{j=1}^2 \min \{ |x_j - y_j|, L - |x_j - y_j| \}^2 \right)^{1/2}, \quad x, y \in \square_L \tag{2.2}$$

Then we may view the metric space  $\square_L$  equipped with the toroidal distance  $d_T(x, y)$  as the 2D flat torus  $\mathbb{R}^2 / L\mathbb{Z}^2$ . By imposing periodic boundary conditions on  $\square_L$ , we may work with translation invariance, and avoid boundary effects in finite systems.

Let  $k_r : \square_L \times \square_L \rightarrow [0, \infty)$  be a family of symmetric kernels that depend on the toroidal distance between its arguments and some  $r \in \mathbb{R}^+$ . That is, there exists a family of functions  $f_r : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  such that  $k_r(x, y) = f_r(d_T(x, y))$ . In this thesis we consider the respective Top-Hat and Gaussian kernels:

**Definition 3.** Let  $r \in \mathbb{R}^+$ . Define the Hard Kernel  $\bar{k}_r$ :

$$\bar{k}_r(x, y) = \mathbb{1}_{d_T(x, y) \leq r} = \begin{cases} 1, & \text{if } d_T(x, y) \leq r \\ 0, & \text{otherwise} \end{cases} \tag{2.3}$$

Define the Soft Kernel  $\tilde{k}_r$ :

$$\tilde{k}_r(x, y) = e^{-\frac{(d_T(x, y))^2}{2r^2}} \tag{2.4}$$

**Definition 4.** Given a point configuration  $X_\rho^N$  and a symmetric kernel  $k_r$ ,  $r \in \mathbb{R}^+$ , define the Random Geometric Graph (RGG)  $G_\rho^N(X_\rho^N, k_r)$  as the weighted graph induced by the weighted adjacency matrix:

$$A \in \mathbb{R}^{N,N}, \quad A_{ij} = k_r(x_i, x_j) \tag{2.5}$$

Further, define the Hard RGG  $\bar{G}_\rho^N(X_\rho^N, r) := G_\rho^N(X_\rho^N, \bar{k}_r)$ , and the Soft RGG  $\tilde{G}_\rho^N(X_\rho^N, r) := G_\rho^N(X_\rho^N, \tilde{k}_r)$ .

Further, we may take the *Thermodynamic Limit* as  $N, L \rightarrow \infty$  with respect to a fixed density  $\rho < \infty$ , such that  $\frac{N}{L^2} = \rho$ , yielding an infinite graph  $G_\rho^\infty(X_\rho^\infty, k_r)$ . In this case,  $X_\rho^\infty$  is a Poisson point process with rate  $\rho$ .

Define the Degree Matrix  $D_{ij} = \delta_{ij} \sum_{k=1}^N A_{ik}$ , and the *Graph Laplacian*  $L = D - A$ . There exists a permutation matrix  $P$  such that  $PAP^{-1}$  and hence  $PLP^{-1}$  are block diagonal, each block corresponding to one of the  $k$  connected components, which we call *Clusters*. That is,  $PLP^{-1} = \bigoplus_{i=1}^k L_i$ , where  $\forall i \in 1, \dots, k$ ,  $L_i$  has eigenvalue 0 with algebraic multiplicity 1. Since this corresponds to relabelling i.i.d. points of  $X_\rho^N$ , we assume WLOG  $A, L$  are block diagonal.

Note that while Soft RGGs are fully connected  $\forall r > 0$ , Hard RGGs are disconnected when  $r < \max_{x,y \in X_\rho^N} d_T(x, y)$ . Further, for  $r' < r$ ,  $\bar{G}_\rho^N(X, r') \subseteq \bar{G}_\rho^N(X, r)$ ; where  $\bar{G}_\rho^N(X, 0)$  is a graph of  $N$  disconnected vertices, and  $\bar{G}_\rho^N(X, \max_{x,y \in X_\rho^N} d_T(x, y))$  is an  $N$ -clique. This is reminiscent of the increasing coupling in Bernouilli Bond percolation, where in this case  $\{\bar{G}_\rho^N(X, r)\}_{r \in \mathbb{R}^+}$  are a coupled monotone family of RGGs w.r.t a fixed point configuration. Thus,  $r$  naturally acts as a control parameter for percolation:

**Definition 5.**  $G_\rho^\infty(X_\rho^\infty, k_r)$  *Percolates* if the cluster of  $G_\rho^\infty(X_\rho^\infty \cap \{0\}, k_r)$  containing the origin is unbounded, which we then call a *Percolating Cluster*.

Further, let  $p_\infty(r)$  be the probability that  $G_\rho^\infty(X_\rho^\infty, k_r)$  percolates, taken over the ensemble of realisations of the Poisson point process  $X_\rho^\infty$ .

We often need to average quantities over the ensemble of all point configurations:

**Definition 6.** Define the (Microcanonical) Ensemble Average of an observable  $f : X_\rho^N \rightarrow \mathbb{R}$ :

$$\langle f \rangle = \frac{1}{L^{2N}} \int_{x_1 \in \square_L} \dots \int_{x_N \in \square_L} f(\mathbf{x}) dx_1 \dots dx_N \quad (2.6)$$

We also extend the definition of the ensemble average to the thermodynamic limit if it exists.

**Theorem 2.2.1.** Define the Critical Radius  $r_c$ :

$$r_c = \inf(r > 0 : p_\infty(r) > 0) \quad (2.7)$$

Then  $r_c \in (0, \infty)$ . [3]

In most of the literature, the radius is fixed at unity, and  $\lambda$  is varied as a control parameter. However,  $G_\rho^\infty(X_\rho^\infty \cap \{0\}, k_1)$  and  $G_\rho^\infty(X_\rho^\infty \cap \{0\}, k_r)$  are isomorphic under scaling the metric space by  $\frac{1}{r}$ , and thus we can discuss these cases interchangably. By fixing density and varying  $r$ , we align with statistical mechanics intuition, and connect naturally with increasing coupling and single-linkage clustering models.

It is important to distinguish between the percolation point, and the following [7]:

**Theorem 2.2.2.** Define  $r_{conn}$  as the radius at which  $\bar{G}$  is asymptotically almost surely connected. Then:

$$r_{conn} = \sqrt{\frac{\log(n) + C}{\pi n}}, \quad C \in \mathbb{R} \quad (2.8)$$

Note that the above analogues of percolation do not apply to the Soft RGG. We will explore further how we may be able to expand these definitions in the literature to this case.

## 2.3 The Kirchhoff Index

We now motivate the definition of the Kirchhoff Index using the notion of *Effective Resistance*, adapted from the texts [18–20].

**Definition 7.** Let  $G = (V, E)$  be a graph and  $P$  a stochastic matrix. Then  $h : V \rightarrow \mathbb{R}$  is a harmonic function w.r.t.  $P$  at vertex  $i \in V$  iff:

$$h(i) = \sum_{j \in V} P(i, j)h(j) \quad (2.9)$$

Further, let  $B \subseteq V$ ,  $|B| \geq 1$  and  $f : B \rightarrow \mathbb{R}$ . Then there exists a unique harmonic extension  $h$  that is harmonic on  $V \setminus B$  and  $\forall v \in B : h(v) = f(v)$ .

Let  $C$  be the weighted adjacency matrix of a connected graph  $G = (V, E)$ , where we view the weights as the conductance between nodes. Let  $P$  be the associated stochastic matrix by normalising rows of  $C$ . Define the *Voltage*  $W$  as the harmonic extension on  $V \setminus \{a, z\}$  w.r.t  $P$ , determined by the values  $W(a), W(z)$ .

We define the *Current Flow* via Ohm's law as  $I(i, j) = C(i, j)(W(i) - W(j))$ . Kirchhoff's current law gives us that  $\sum_{j \in V} I(i, j) = \|I\|(\delta_a(i) - \delta_z(i))$ , where we define  $\|I\|$  as the current strength. We thus define the *Effective Resistance* between  $a, z$  as  $R_{eff}(a, z) = \frac{W(a) - W(z)}{\|I\|}$ .

**Definition 8.** Define the *Kirchhoff Index*:

$$R_G = \sum_{i, j \in V, i < j} R_{eff}(i, j) \quad (2.10)$$

One may also establish a correspondence between random walks on Graphs and electrical resistance:

**Lemma 2.3.1.** Consider a random walk defined by a Markov chain on a graph  $G$  with transition matrix  $P$ , formed from normalising the rows of the weighted adjacency matrix  $A$ . Define the Commute Time,  $C(i, j)$  as the expected time for the random walk starting at  $i$  to visit  $j$  and return back to  $i$ . Then:

$$C(i, j) = R_{eff}(i, j) \cdot \sum_{i, j \in V} A(i, j) \quad (2.11)$$

We now derive an alternative formulation of the Kirchhoff Index in terms of the graph Laplacian.

**Definition 9.** Let  $A \in \mathbb{R}^{m,n}$ . Define the *Pseudoinverse* or *Moore-Penrose inverse* as the unique matrix  $A^+ \in \mathbb{R}^{m,n}$  satisfying the following conditions:

1.  $AA^+A = A$
2.  $A^+AA^+ = A^+$
3.  $(AA^+)^T = AA^+$
4.  $(A^+A)^T = A^+A$

In particular,  $A^+$  acts as the left inverse of  $A$  on the row space of  $A$ , and the zero map on the null space of  $A$ .

**Theorem 2.3.2.** *Let  $L$  be the Laplacian matrix of a connected graph  $G$ , with eigenvalues  $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ :*

$$R_G = N\text{Tr}(L^+) = N \sum_{i \in \{2, \dots, N\}} \frac{1}{\lambda_i} \quad (2.12)$$

*Proof.* WLOG consider a unit current injection  $\|I\| = 1$ , and let  $b$  be the current vector with entries  $b(i) = \delta_a(i) - \delta_z(i)$ . Kirchhoff's current law yields:

$$b(i) = \sum_j I(i, j) = \sum_j C(i, j)(W(i) - W(j)) = \sum_j L(i, j)W(j) \iff b = LW \quad (2.13)$$

Since  $L$  is singular, we seek solutions  $W$  in the row space of  $L$ , that is, perpendicular to  $\text{Span}(\mathbf{1})$ ;  $W = L^+b + c\mathbf{1}$ ,  $c \in \mathbb{R}$ . Thus:

$$\begin{aligned} R_{eff}(i, j) &= [L^+b](i) - [L^+b](j) = [L^+(e_i - e_j)](i) - [L^+(e_i - e_j)](j) \\ &= L^+(i, i) + L^+(j, j) - 2L^+(i, j) \end{aligned} \quad (2.14)$$

Using symmetry of  $L^+$ . Since  $\mathbf{1}$  is an eigenvector of  $L^+$ :

$$R_G = \sum_{i, j \in V, i < j} L^+(i, i) + L^+(j, j) - 2L^+(i, j) = N\text{Tr}(L^+) \quad (2.15)$$

Since the spectrum of  $L^+$  consists of the reciprocals of the eigenvalues of  $A$  associated with the row space of  $A$ , and 0 for each eigenvalue associated with the null space of  $A$ , the second equality follows.

□

Thus the leading-order term of  $\text{Tr}[L^+]$  is the reciprocal of the *Smallest Non-Zero Eigenvalue* of  $L$ ,  $\frac{1}{\lambda_{SNZE}}$ . Label the eigenvalues of  $L$  as  $0 = \lambda_1 \leq \dots \leq \lambda_N$ . It is well known that  $G$  is connected iff  $\lambda_2 > 0$ ;  $\lambda_2$  is also known as the *Algebraic Connectivity* of a graph, and satisfies Cheeger's Inequality [21].

**Lemma 2.3.3** (Cheeger's Inequality). *Let  $G = (V, E)$  be a connected graph, with graph Laplacian  $L$  and  $\lambda_2$  its second-smallest eigenvalue Then:*

$$\frac{h(G)^2}{2\Delta} \leq \lambda_2 \leq 2h(G) \quad (2.16)$$

where, for  $S \subseteq V$ ,  $\partial S$  the set of edges with one end in  $S$  and the other in  $V \setminus S$ :

$$h(G) = \min_{S \subset V, 0 < |S| \leq \frac{|V|}{2}} \frac{|\partial S|}{|S|} \quad (2.17)$$

Intuitively,  $\lambda_2$  quantifies the weakest point of connectivity of a graph - if there is a narrow "bottleneck" that sparsely connects two large regions in the graph then  $\lambda_2$  is small.

## 2.4 Random Matrix Theoretic Ingredients

A Random Matrix is a matrix whose entries are random variables. The entries of a Euclidean Random Matrix (ERM) are deterministic functions of the pairwise distances between  $N$  points that are randomly distributed in a finite region  $V \subset \mathbb{R}^d$ .

We study the ensemble of ERMs induced by point configurations and  $f : \square_L \times \square_L \rightarrow \mathbb{R}$ , with entries:

$$A_{ij} = f(x_i, x_j) - \sum_{k=1}^N f(x_i, x_k) \quad (2.18)$$

Let  $A$  be a matrix from the above ensemble. Define the empirical spectral density of  $A$  as the random measure:

$$\rho_A(\lambda) = \frac{1}{N} \sum_{i=1}^N \delta(\lambda - \lambda_i) \quad (2.19)$$

where  $\{\lambda_i\}$  are the eigenvalues of  $A$ ; these are random variables, with joint p.d.f.  $\rho(\lambda_1, \dots, \lambda_N)$ . Taking an average over the matrix ensemble yields the following [22]:

**Definition 10.** Define the Spectral Density of the ensemble:

$$\rho(\lambda) = \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{R}^N} \delta(\lambda - \lambda_i) \rho(\lambda_1, \dots, \lambda_N) d\lambda_1 \dots d\lambda_N \quad (2.20)$$

which, in our case is equal to the ensemble average over point configurations  $\rho(\lambda) = \langle \rho_A(\lambda) \rangle$ . We also define a useful quantity:

**Definition 11.** Define the Resolvent of  $A$  as:

$$R : \mathbb{C} \setminus \{\lambda_i\} \rightarrow \mathbb{R}, \quad R(z) = \frac{1}{N} \langle \text{Tr}[(A - zI)^{-1}] \rangle \quad (2.21)$$

which, in the limit as  $N \rightarrow \infty$ :

$$\lim_{N \rightarrow \infty} R(z) = \int \frac{\rho(\lambda)}{z - \lambda} d\lambda \quad (2.22)$$

We are also able to obtain the spectral density from the resolvent [22]:

$$\rho(\lambda) = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \text{Im} R(\lambda - i\epsilon) \quad (2.23)$$

# Chapter 3

## Methods

### 3.1 Random Geometric Graph Generation

To perform finite-size scaling analysis on the model, we must calculate statistics of the RGG Laplacian matrix over numerous point configurations over a fine parameter sweep of  $r$ , including for large system size  $N$ . This requires a computationally efficient method to generate RGGs. We treat the hard and soft cases differently, as each case benefits from different methods.

Firstly, we consider a numerical trick in which we may efficiently compute the toroidal distance between points using the Python mod operation:

**Lemma 3.1.1.**

$$\forall x, y \in \square_L : d_{\mathcal{T}}(x, y) \equiv \|((x - y + \frac{L}{2}) \bmod L) - \frac{L}{2}\|_2 \quad (3.1)$$

where we define:

$$\begin{aligned} x_i \in \mathbb{R} : x_i \bmod L &= \max_{m: x_i - mL \geq 0} x_i - mL \\ x \in \square_L : x \bmod L &:= (x_1 \bmod L, x_2 \bmod L)^T \end{aligned} \quad (3.2)$$

*Proof.* Let  $x, y \in \square_L$  and define  $d = x - y$ . Then for each coordinate  $j \in \{1, 2\}$ , observe:

$$[((d_j + \frac{L}{2}) \bmod L) - \frac{L}{2}]^2 = \begin{cases} d_j^2, & \text{if } |d_j| \leq \frac{L}{2} \\ (d_j - \text{sign}(d_j)L)^2, & \text{otherwise} \end{cases} = \min\{|d_j|, L - |d_j|\}^2 \quad (3.3)$$

$$\therefore \left\| ((x - y + \frac{L}{2}) \bmod L) - \frac{L}{2} \right\|_2^2 = \sum_{j=1}^2 \min\{|x_j - y_j|, L - |x_j - y_j|\}^2 = d_{\mathcal{T}}(x, y)^2. \quad (3.4)$$

Taking square roots on both sides completes the proof.  $\square$

Note that mod, exp, as well as basic arithmetic operations are implemented in Python and are defined to apply element-wise, in the Hadamard sense, to tensors. This naturally allows us to implement a vectorised algorithm that, given a list of co-ordinates from a point configuration, calculates the pairwise toroidal distance matrix, from which we easily construct the Laplacian matrix of the associated Soft RGG:

---

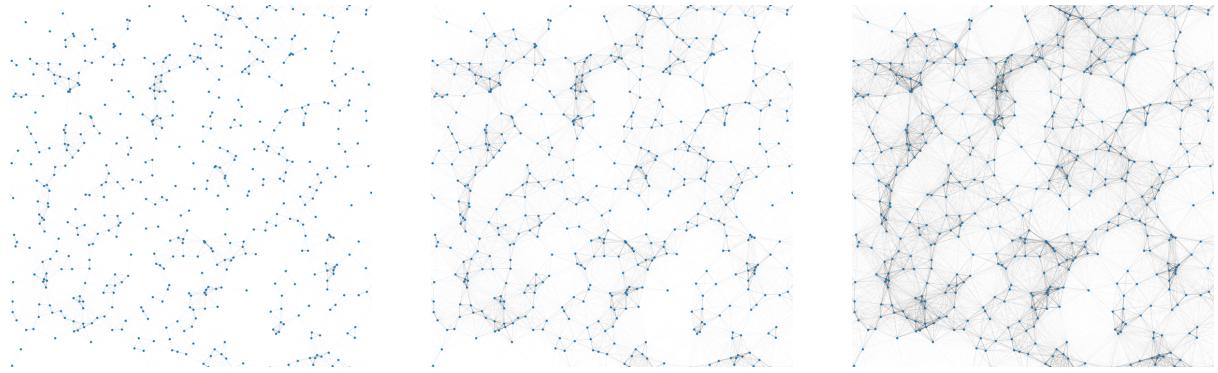
**Algorithm 1:** Vectorised Soft RGG Generation

---

**Input:**  $N$  (Number of nodes),  $\rho$  (Density),  $r$  (Radius)

**Output:** Soft-RGG Laplacian matrix  $\mathbf{L}$

- 1  $L \leftarrow \sqrt{N/\rho}$
  - 2 Sample  $N$  points  $p_i \sim \text{Uniform}([-L/2, L/2]^2)$ ;
  - 3 Construct tensor  $D \in \mathbb{R}^{N,N,2}$  with  $D[i, j] = p_i - p_j \in \mathbb{R}^2$  ;
  - 4 Apply toroidal wrapping:;
  - 5  $D_p \leftarrow (D + L/2) \bmod L - L/2$ ;
  - 6 Compute Euclidean distances:;
  - 7  $D_f[i, j] \leftarrow \|D_p[i, j]\|_2$ ;
  - 8 Set  $\mathbf{L}_{ij} = -\exp(-D_f[i, j]/(2r^2))$ ;
  - 9 Set diagonal:  $\mathbf{L}_{ii} \leftarrow -\sum_{j \neq i} M_{ij}$ ;
- 



**Figure 3.1** Soft RGGs generated by a point configuration with  $N = 500, r = 0.1, 0.125, 0.15$ .

While we may apply the same implementation to the case of the hard kernel, there are further optimisations to be made in virtue of the fact that each point only interacts with points in a radius of  $r$  w.r.t  $d_T$ . This is remedied by dividing  $\square_L$  into a square grid of  $n = \lfloor \frac{L}{r} \rfloor$  square cells along each side of length  $\frac{L}{n}$ . Then every point is located within one of the cells; WLOG pick an arbitrary neighbouring cell if a point lies on the boundary. Then it is easy to see that  $\forall y : d_T(x, y) < r$ ,  $y$  is contained within the  $3 \times 3$  grid of cells centered in the cell occupied by  $x$ , with periodic wraparound at the boundaries of  $\square_L$ .

Additionally, since the Hard RGG is generally disconnected, we require the Laplacian to be in block-diagonal form. We thus construct a Breadth-First Search algorithm, which constructs the Laplacian matrix for each connected cluster containing each node in the point configuration, if it has not yet been visited. Since each node is checked, we guarantee that all clusters of the system are accounted for.

We are thus able to generate numerous configurations of the RGG for different system sizes  $N$  and radii  $r$ ; this allows us to obtain a Monte Carlo estimate for  $\langle \text{Tr}[L^+] \rangle$  by taking an empirical mean of  $\text{Tr}[L^+]$  over the sample. This is parallelised via use of the High Performance Computing (HPC) cluster, provided by the Research Computing Services of Imperial College London.

The spectrum of each RGG is found using the `np.linalg.eigvalsh` function. To calculate the Kirchhoff index, we compute  $\tau \approx \sum_{\lambda_i > \epsilon} \frac{1}{\lambda_i}$ , where  $\epsilon = 10^{-12}$ .

---

**Algorithm 2:** Hard RGG Generation using Breadth-First Search

---

**Input:**  $N$  (Number of nodes),  $\rho$  (Density),  $r$  (Radius)

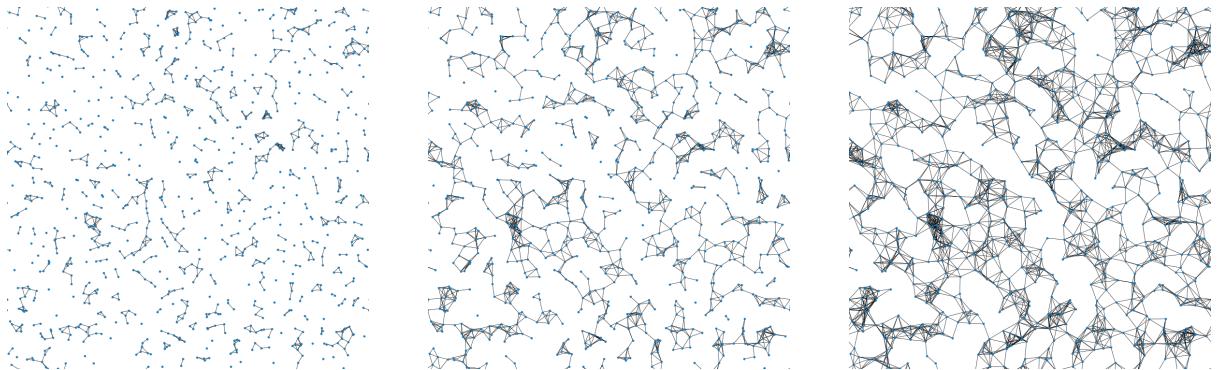
**Output:** Laplacian  $\mathbf{L}$  of Hard RGG

```

1  $L \leftarrow \sqrt{N/\rho};$ 
2 Sample  $N$  points  $p_i \sim \text{Uniform}([-L/2, L/2]^2);$ 
3  $n \leftarrow \lfloor L/r \rfloor, l \leftarrow L/n;$ 
4 Divide  $\square_L$  into grid of  $n^2$  cells of length  $l$ ;
5 Define grid cell indices  $(i, j)$  and dictionary of toroidal neighbours;
6 Assign each  $p_i$  to cell  $(i, j)$  via spatial hash;
7 Build dictionary from cell  $\mapsto$  point indices in cell;
8 Mark all nodes unvisited;
9 Initialise list of Laplacians of clusters  $\mathcal{L}$ ;
10 for  $i = 1$  to  $N$  do
11   if node  $i$  not visited then
12     Mark  $i$  as visited, initialise queue with  $(i, 0)$ ;
13     while queue not empty do
14       Pop  $(v, d)$  from queue;
15       Find candidate nodes in  $3 \times 3$  neighbour cells;
16       Filter candidates by torus distance  $< r$ ;
17       Add unvisited nodes to queue and node list, add edges to edge list;
18     end
19     Construct sparse Laplacian  $L_c$  from node/edge list, calculate and store spectrum
       and size of  $L_c$ ;
20     Append  $L_c$  to  $\mathcal{L}$  ;
21   end
22 end
23 Return  $\mathbf{L} = \bigoplus_{\mathcal{L}_j \in \mathcal{L}} \mathcal{L}_j;$ 

```

---



**Figure 3.2** Hard RGGs generated by a point configuration with  $N = 1000, r = 0.05, 0.075, 0.1$ .

**Lemma 3.1.2.** *The approximation, for small  $\epsilon > 0$ :*

$$\text{Tr}[L^+] \approx \sum_{\lambda_i > \epsilon} \frac{1}{\lambda_i} \quad (3.5)$$

*is exact in the case of the Hard RGG for  $N < \frac{\pi}{\arccos(1 - \frac{\epsilon}{2})}$*

*Proof.* The approximation for  $\text{Tr}[L^+]$ , by truncating eigenvalues smaller than  $\epsilon$ , is exact if the smallest non-zero eigenvalue  $\lambda_{SNZE} > \epsilon$ . It is known that for all connected graphs with unit weights of size  $N$ , the Path graph minimises  $\lambda_2$  [23], with value  $2(1 - \cos(\frac{\pi}{N}))$ . Since this is decreasing for  $N \geq 1$ , take  $N$  to be the system size of the RGG. Thus, a sufficient condition is:

$$2(1 - \cos(\frac{\pi}{N})) > \epsilon \iff N < \frac{\pi}{\arccos(1 - \frac{\epsilon}{2})} \quad (3.6)$$

□

For  $\epsilon = 10^{-12}$ , we have that the above approximation is exact for  $N < 3141454$ , which exceeds the maximum system size simulated. However, for soft RGGs, as  $N, L$  increase in the thermodynamic limit, the weights between the furthest nodes exponentially decay to zero, while weights between the closest nodes stay constant. Intuitively by Cheeger's inequality, this leads to a vanishing  $\lambda_2$  that quickly falls below machine epsilon, `numpy.finfo().eps`  $\approx \epsilon \sim 10^{-16}$ .

## 3.2 Finite-Size Scaling

Let  $\tau = \frac{\text{Tr}[L^+]}{N}$ . Assume that  $\xi, \tau$  scales as the power laws:

$$\xi \sim |r - r_c|^{-\nu}, \quad \tau \sim |r - r_c|^{\varsigma} \quad (3.7)$$

We thus assume the standard ansatz for the scaling form of  $\tau$  near criticality [24]:

$$\tau \sim \begin{cases} \xi^{-\frac{\varsigma}{\nu}}, & L \gg \xi \\ L^{-\frac{\varsigma}{\nu}}, & L \ll \xi \end{cases} = L^{-\frac{\varsigma}{\nu}} f\left(\frac{L}{\xi}\right) \quad (3.8)$$

for some universal scaling function  $f(u)$ , satisfying:

$$f(u) = \begin{cases} c \in \mathbb{R}, & u \ll 1 \\ u^{\frac{\varsigma}{\nu}}, & u \gg 1 \end{cases} \quad (3.9)$$

Then substituting  $\tilde{f}(u) := f(u^{\frac{1}{\nu}})$ :

$$\tau = L^{-\frac{\varsigma}{\nu}} \tilde{f}(L^{\frac{1}{\nu}} |r - r_c|) \iff L^{\frac{\varsigma}{\nu}} \tau = \tilde{f}(L^{\frac{1}{\nu}} |r - r_c|) \quad (3.10)$$

Thus plotting  $L^{\frac{1}{\nu}} |r - r_c|$  against  $L^{\frac{\varsigma}{\nu}} \tau$  for all  $N$  should yield the universal curve  $\tilde{f}(x)$ ; this is called the data collapse method [25].

In practice, we aim to minimise some distance function between the empirical data and the curve  $\tilde{f}(x)$  with respect to the parameters  $r_c, \nu, \varsigma$ . Suppose we compute  $m$  values of  $\tau$  over  $m$  point configurations for each of the  $n$  values of system size  $N$ , and each of the  $k$  values of  $r$ . Let

$x_{i,j} = L_i^{\frac{1}{\nu}} |r - r_c|$ ,  $y_{i,j} = L_i^{\frac{\varsigma}{\nu}} \langle \tau_j \rangle$ ,  $dy_{i,j} = L_i^{\frac{\varsigma}{\nu}} d\tau_j$ , where  $\langle \tau_j \rangle, d\tau_j$  is the mean and standard error on  $\tau_j$  over  $m$  realisations respectively.

We use the quality function from Houdayer and Hartmann [26], which resembles a reduced  $\chi^2$  statistic:

$$S = \frac{1}{N} \sum_{i,j} \frac{(y_{ij} - Y_{i,j})^2}{dy_{i,j}^2 + dY_{i,j}^2} \quad (3.11)$$

where each  $Y_{i,j}$  is the locally estimated value of  $\tilde{f}(x_{i,j})$ . This is calculated by selecting all datapoints from other system sizes  $x_{k,l}$ ,  $k \neq i$  such that  $x_{k,l} \leq x_{i,j} \leq x_{k,(l+1)}$ , and using weighted least squares, with respect to  $w = \frac{1}{dy_{k,l}}$ , to obtain an interpolated value at  $x_{i,j}$  accounting for the standard errors  $dy_{k,l}$ :

$$Y_{i,j} = \frac{KK_{xy} - K_x K_y}{\Delta} x_{ij} + \frac{K_{xx} K_y - K_x K_{xy}}{\Delta} \quad (3.12)$$

where  $K = \sum_{k,l} w_{k,l}$ ,  $K_x = \sum_{k,l} w_{k,l} x_{k,l}$ ,  $K_y = \sum_{k,l} w_{k,l} y_{k,l}$ ,  $K_{xx} = \sum_{k,l} w_{k,l} x_{k,l}^2$ ,  $K_{yy} = \sum_{k,l} w_{k,l} y_{k,l}^2$  and  $\Delta = KK_{xx} - K_x^2$ .

We use the implementation provided by the *pyfssa* Python package [27, 28] to perform the aforementioned minimisation of  $S$  with respect to parameters  $r_c, \nu, \varsigma$  using the SciPy implementation of the Nelder-Mead algorithm.

Alternatively, we may exploit the finite-size scaling behaviour of the peaks of  $\tau$  at different system sizes, which at size  $N$  achieves its maximum value  $\tau_{max}^N$  at  $\hat{r}_c^N$ . By respectively substituting  $r = r_c$  into 3.10, and using the fact that  $\tilde{f}$  is maximised at some  $u \in \mathbb{R}$ , we derive the respective scaling relations:

$$\tau_{max}^N \sim L^{-\frac{\varsigma}{\nu}}, \quad \hat{r}_c \sim L^{-\frac{1}{\nu}} + r_c \quad (3.13)$$

For each system size  $N$ , we estimate  $\hat{r}_c^N, \tau_{max}^N$  by locally fitting a quadratic polynomial,  $ax^2 + bx + c$ , on the empirical curve of  $\tau$  in a small window centred around the empirical maximum value of  $\tau$ . Then  $\hat{r}_c^N \approx -\frac{b}{2a}$ . We then repeat this process centred at the nearest datapoint  $r$  to  $\hat{r}_c^N$ , to ensure our window is as symmetric about  $\hat{r}_c$  as possible.

Then, given the correct value of  $r_c$ , the negative reciprocal of the gradient recovered from the straight line  $\log L$  against  $\log |\hat{r}_c^N - r|$  yields  $\nu$ . We thus perform linear regression, and optimise the  $R^2$  value with respect to  $r_c$  using `scipy.optimize.minimize`. Performing a further linear regression on  $\log L$  against  $\log \tau_{max}^N$ , we obtain the gradient  $-\frac{\varsigma}{\nu}$ , which in turn yields  $\varsigma$ .

# Chapter 4

## Results

### 4.1 Conventional Statistics of the Laplacian

#### 4.1.1 Trace

We recall that the trace of the Laplacian measures the sum of the degrees of the nodes in the network. An analytical expression for both the Hard and Soft RGG is easily calculated:

**Lemma 4.1.1.** *In the case of the Hard RGG, for  $r \leq \frac{L}{2}$ :*

$$\langle \text{Tr}[\bar{L}(r)] \rangle = (N - 1)\pi r^2 \rho \quad (4.1)$$

*Proof.* Consider  $x_i \in X_\rho^N$ , then for  $r \leq \frac{L}{2}$ :

$$P(d_T(x_i, x_j) < r) = \frac{\pi r^2}{L^2} \quad (4.2)$$

Then the number of points that lie within a toroidal distance of  $r$  of  $x_i$  is distributed Binomially as  $B(N - 1, \frac{\pi r^2}{L^2})$ . Thus the expected degree of node  $i$  is:

$$\langle k_i \rangle = (N - 1) \frac{\pi r^2}{L^2} = \pi r^2 \rho \left(1 - \frac{1}{N}\right) \quad (4.3)$$

Multiplying by  $N$  yields  $\langle \text{Tr}[\bar{L}(r)] \rangle$ . □

**Lemma 4.1.2.** *In the case of the Soft RGG:*

$$\langle \text{Tr}[\tilde{L}(r)] \rangle = (N - 1)2\pi r^2 \rho \cdot \text{erf}\left(\frac{L}{2\sqrt{2r}}\right)^2 \quad (4.4)$$

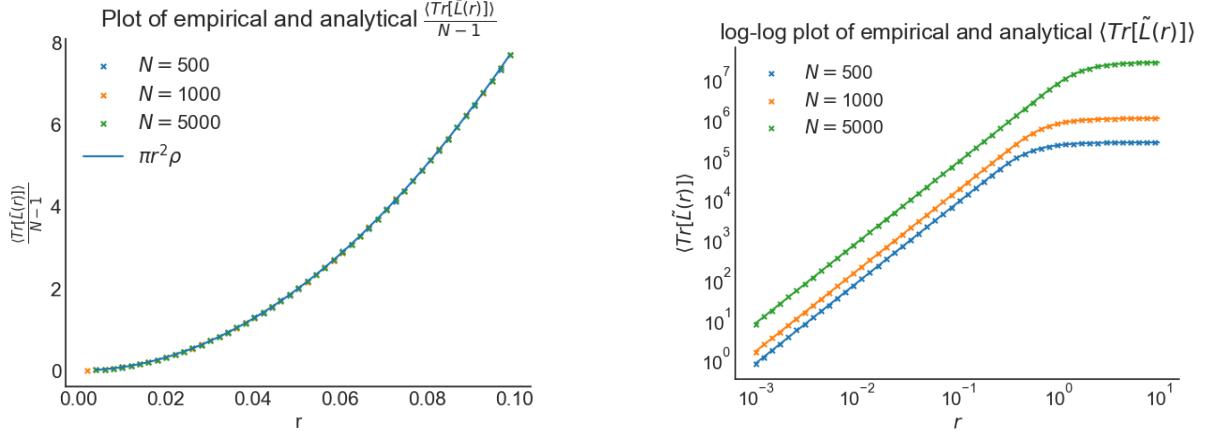
where  $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$  is the error function.

*Proof.* Consider  $x_i \in X_\rho^N$ . We may assume WLOG this point lies at the origin due to the translational invariance of  $d_T$  on  $\square_L$ . Then the expected (weighted) degree is:

$$\begin{aligned} \langle k_i \rangle &= \left\langle \sum_{j=1}^{N-1} e^{-\frac{\|x_j\|_2^2}{2r^2}} \right\rangle = (N - 1) \langle e^{-\frac{\|x\|_2^2}{2r^2}} \rangle \\ &= (N - 1) \frac{1}{L^2} \int_{\square_L} e^{-\frac{\|x\|_2^2}{2r^2}} dx \\ &= \frac{N - 1}{L^2} 2\pi r^2 \text{erf}\left(\frac{L}{2\sqrt{2r}}\right)^2 \end{aligned} \quad (4.5)$$

Multiplying by  $N$  and using  $L = \sqrt{\frac{N}{\rho}}$  yields  $\langle \text{Tr}[\tilde{L}(r)] \rangle$  as required.  $\square$

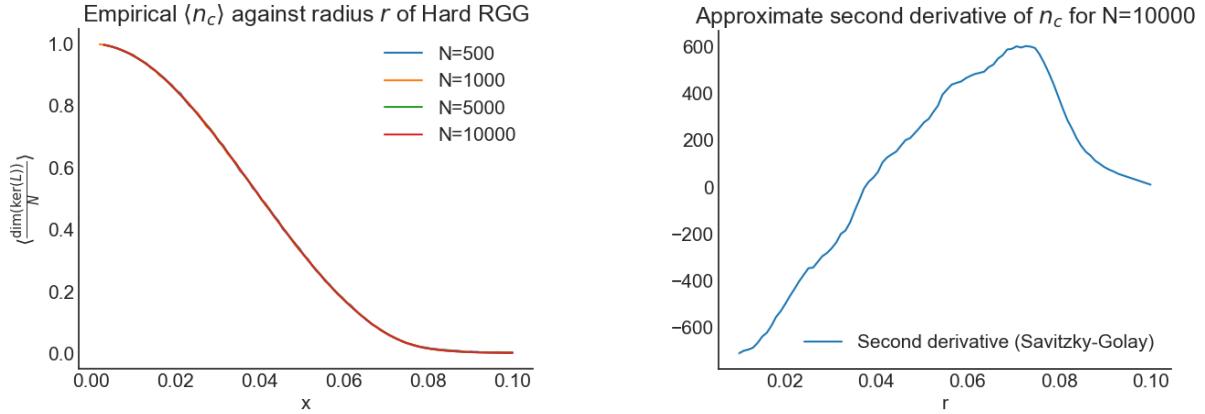
We empirically verify this relation against a Monte Carlo estimate of  $\text{Tr}[L]$  in Figure 4.1, and observe that our numerical results align exactly.



**Figure 4.1** Analytical and Monte-Carlo estimates of ensemble-averaged Trace of the RGG Laplacian against radius  $r$  in the Hard (left) and Soft (right) cases for system sizes  $N = 500, 1000, 5000$ .

#### 4.1.2 Dimension of the Null Space

The dimension of the null space of the Laplacian corresponds to the number of clusters in the graph; normalising by the system size  $N$  yields the number of clusters per site. We examine the Hard RGG case, as the soft RGG trivially has a null space dimension of one for all  $r$ .



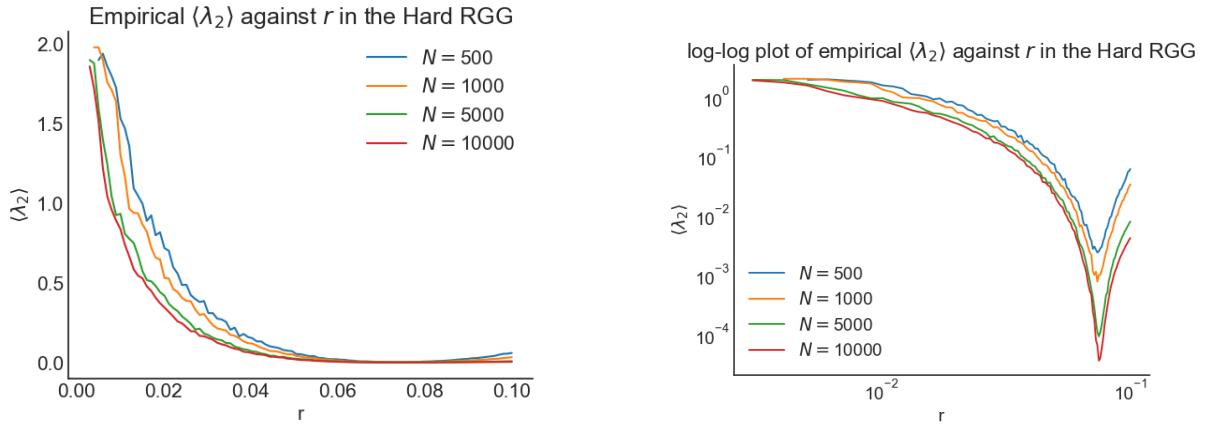
**Figure 4.2** Monte-Carlo estimate of ensemble-averaged number of clusters per site against radius  $r$  for system sizes  $N = 500, 1000, 5000, 10000$  (left), and approximation of its second derivative using the Savitzky-Golay filter with window size 15 and polynomial order 3, for  $N = 10000$  (right).

At first glance, the curve of the number of clusters per site,  $n_c$ , appears smooth with respect to  $r$ . However, it is instructive to examine the behaviour of higher-order derivatives - recall that in bond percolation on lattices, the third derivative of  $n_c$  with respect to the percolation probability diverges at the critical point. This motivates examining the second derivative of  $n_c$  with respect to  $r$ .

Due to the sensitivity of numerical differentiation to noise, we employ the Savitzky-Golay filter [29] to approximate the second derivative. This method performs a local polynomial fit in sliding windows over the data, thus providing a smoothed estimate of the derivative. The right panel of 4.2 shows a clear peak in the second derivative near  $r \approx 0.07$ , consistent with the emergence of a giant component. However, the location and sharpness of the peak are sensitive to window size, polynomial order and resolution of the original data. While this plot provides heuristic evidence for criticality, the presence of any non-analyticities remains inconclusive, and it cannot yield a precise estimate for  $r_c$ .

#### 4.1.3 Smallest Non-Zero Eigenvalue

We also plot the Smallest Non-Zero Eigenvalue,  $\lambda_{SNZE}$  for the Hard RGG Laplacian.



**Figure 4.3** Monte-Carlo estimate of ensemble-averaged smallest non-zero eigenvalue of the Hard RGG Laplacian against radius  $r$  for system sizes  $N = 500, 1000, 5000, 10000$ .

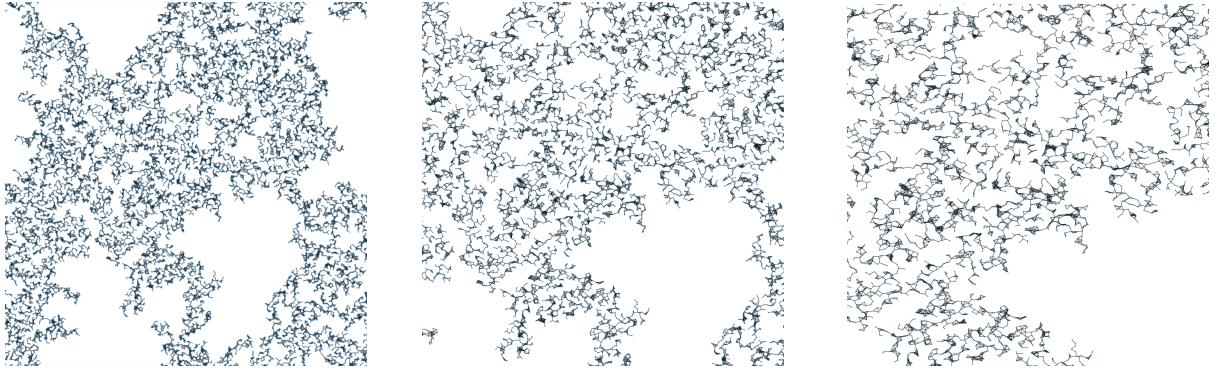
$\lambda_{SNZE}$  appears to reach a minimum at a critical value  $r_c$ , becoming sharper for larger system size. This aligns with our hypothesis that  $\text{Tr}[L^+]$  diverges at the critical point, as  $\frac{1}{\lambda_{SNZE}}$  is the leading order term. Heuristically, one may argue why this occurs using its interpretation as it is the algebraic connectivity  $\lambda_2$  for a cluster - in fact  $\lambda_{SNZE}$  is the minimum algebraic connectivity of all clusters in the Hard RGG.

In the subcritical regime, for small  $r$ , the graph is fragmented into many small, dense components, often resembling near-cliques. Thus  $\lambda_{SNZE}$  tends to be relatively large on average due to the high internal connectivity of these clusters, leading to a low trace of the Laplacian pseudoinverse. As the radius increases, the clusters expand and merge, forming bottlenecks between them, decreasing  $\lambda_{SNZE}$ .

Near the critical point, clusters merge to form a macroscopic cluster (infinite in the thermodynamic limit). The resulting structure exhibits an "infinite-barbell"-like topology, characterised by many relatively dense modules joined by narrow, path-like graphs. The algebraic connectivity of this topology vanishes in the thermodynamic limit - this can loosely be reasoned with using Cheeger's inequality; we can find an arbitrarily sparse cut between clusters due to the scale-invariant property of the percolating cluster.

Beyond the critical point, in the supercritical regime, the infinite component gains more internal connections as the radius increases. Thus the prevalence of bottlenecks reduces, leading to an increase in  $\lambda_{SNZE}$ .

While  $\lambda_2$  is a common measure of connectivity, one must be cautious when applying it in the setting of percolation. Crucially, it remains a local measure in two ways - it is sensitive to bot-



**Figure 4.4** Topology of the percolating cluster in the Hard RGG at different scales.

tlenecks within clusters, and crucially only captures the connectivity of the weakest connected cluster in the Hard RGG. Namely, for any fixed radius  $r > 0$ , one may argue that we can find an arbitrarily weakly connected cluster - for example two connected clusters separated by an arbitrarily long path. This would result in a low  $\lambda_2$ , even if the rest of the clusters in the graph were well connected.

For reasons as discussed previously, the Smallest Non-Zero Eigenvalue is not a reliable empirical measure of connectivity in the Soft RGG, as  $\lambda_2$  vanishes exponentially quickly in the thermodynamic regime for all values of  $r$ .

## 4.2 $Tr[L^+]$ as a signature of the percolation transition

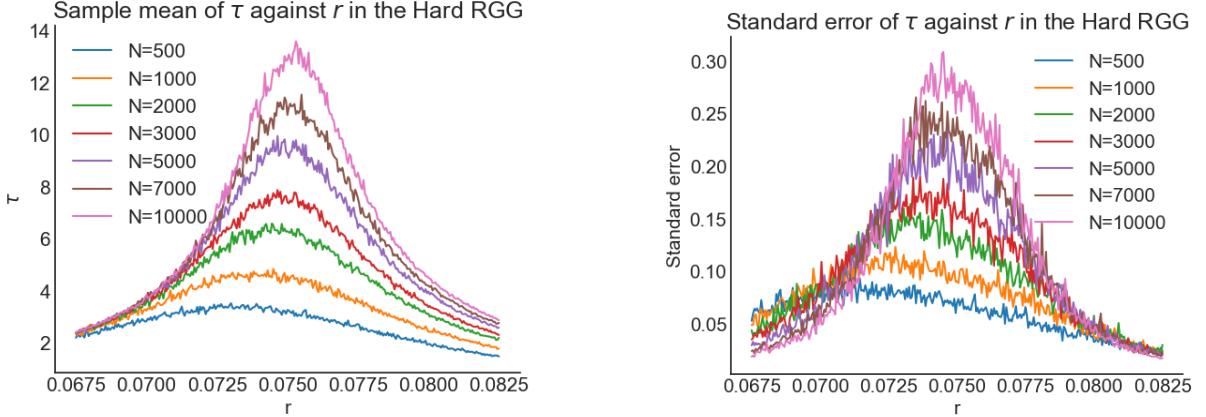
### 4.2.1 The Hard RGG

We now examine the quantity  $Tr[L^+]$  of the RGG, which corresponds to  $\frac{R_G}{N}$  for connected graphs. Classically, for disconnected graphs, the resistance distance is ill-defined as the effective resistance between nodes in disjoint clusters is infinite and so  $R_G = \infty$ . Typical conductance exponents are thus derived for the giant component in the supercritical regime, deriving scaling laws from the percolating cluster, namely the dimension of its conducting backbone.

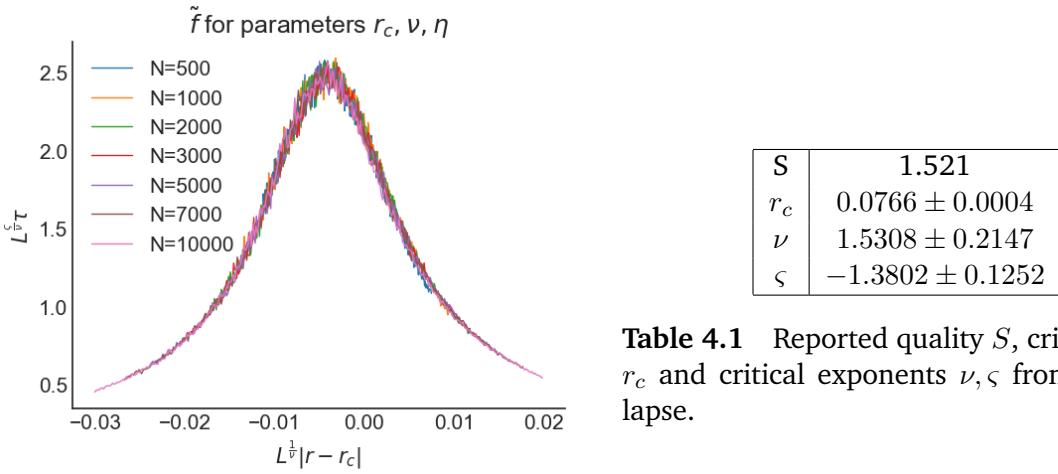
The key development in this work is a principled extension of the Kirchhoff Index to the entire graph, regardless of its connectivity. Specifically, we generalise this notion for disconnected graphs by exploiting the block-diagonal structure of the Laplacian as  $Tr[L^+] = \sum \frac{R_{G_i}}{|G_i|}$ , denoting each connected component as  $G_i$ . Crucially, this now corresponds to the sum of the internal effective resistances for each cluster, or alternatively the sum of within-cluster commute times. This generalisation provides a framework to examine conductivity properties of the system from both below and above criticality, as well as globally over all clusters.

In contrast to examining  $\lambda_2$ ,  $Tr[L^+]$  aggregates information from the entire Laplacian spectrum, most notably the smallest non-zero eigenvalues for every cluster in the graph. While it is dominated by  $\lambda_2$ , and hence also diverges at criticality by the same argument, it is more robust to local fluctuations, and is representative of the connectivity of the entire system.

As before, let  $\tau = \frac{Tr[L^+]}{N}$ ; recall this represents the average effective resistance between two nodes within the same cluster in an RGG. We first examine  $\langle \tau \rangle$  for the Hard RGG of size  $N$ . WLOG fix  $\rho = 250$ . We perform a Monte Carlo estimate of this quantity over 200 configurations, for 7 system sizes  $N \in \{500, 1000, 2000, 3000, 5000, 7000, 10000\}$ , and 301 equispaced values in a window around the critical point  $r \in [0.0675, 0.0825]$ .



**Figure 4.5** Monte-Carlo estimate (left) and Standard Error (right) of  $\tau$  against radius  $r$  around  $r_c$  for different system sizes.

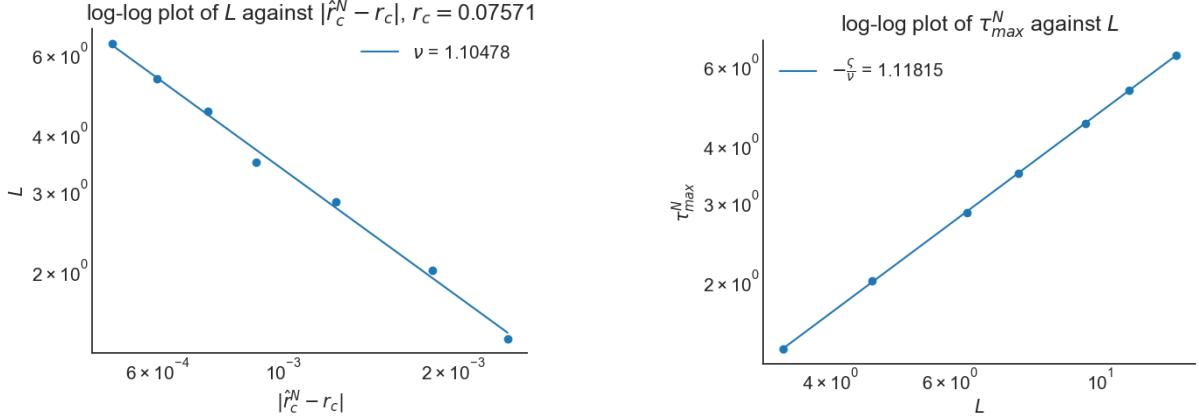


**Figure 4.6** Data collapse to  $\tilde{f}$ ;  $L^{\frac{1}{\nu}}|r - r_c|$  against  $L^{\frac{\zeta}{\nu}}\tau$  for multiple system sizes.

As expected, we observe a peak in the empirical average and standard error of  $\tau$  that sharpens and shifts for increasing  $N$  due to finite-size effects. Applying the data collapse method yields the master curve in Figure 4.6.

The curve collapse yields a visually satisfactory fit, supported by the quality metric  $S$  close to the optimal value of 1. Notably, we observe that the correlation length exponent for the 2D standard percolation universality class,  $\nu = \frac{4}{3}$  lies within the returned error bars, thus agreeing with the literature. However, the value of  $r_c$  slightly differs from the agreed empirical value in the literature,  $r_c \approx 2\sqrt{\frac{1.12808737}{250\pi}} \approx 0.075798$  [30, 31], indicating the presence of systematic error. Note that this fit takes the implicit assumption that the critical exponents are symmetric on both sides of the critical point; while this is true for  $\nu$  and other typical critical exponents, this may not be the case for  $\zeta$ , due to the difference in conductivity between subcritical and supercritical regimes. In addition, optimising with respect to three variables increases the risk of overfitting, or convergence to a local minima in  $S$ .

We compare these results to the finite-size scaling derived from the second method. We recover a more accurate value of  $r_c \approx 0.07571$ , however the resulting critical exponents differ from those obtained via the curve collapse method and the literature:  $\nu \approx 1.10478$ ,  $\zeta \approx 1.23532$ . While we have a strong linear fit, with respective  $R^2$  scores 0.9979 and 0.99998, we must be



**Figure 4.7** Derived exponents from finite-size scaling of pseudocritical points  $\hat{r}_c^N$ ,  $\tau_{max}^N$ .

cautious in interpreting these results. The method introduces systematic error by locally fitting for  $\tau_{max}$  using a quadratic, which is unrepresentative of the scaling function, and its potential asymmetry. This local polynomial fit may also introduce power-law artefacts into the fitting process.

#### 4.2.2 The Soft RGG

We now examine  $Tr[\tilde{L}^+]$  in the case of the Soft RGG. We present an important difference between these kernels:

**Lemma 4.2.1.** *For all Soft RGGs for fixed  $N$ ,  $Tr[\tilde{L}^+]$  is monotonically decreasing with respect to  $r$ .*

*Proof.* Consider a Soft RGG with Laplacian matrix  $\tilde{L}$ . We have by Appendix A.1:

$$\frac{\partial Tr[\tilde{L}^+]}{\partial w_{ij}} = -||\tilde{L}_i^+ - \tilde{L}_j^+||_2^2 \quad (4.6)$$

Then:

$$\frac{\partial Tr[\tilde{L}^+]}{\partial r} = \sum_{i,j} \frac{\partial Tr[\tilde{L}^+]}{\partial w_{ij}} \cdot \frac{\partial w_{ij}}{\partial r} = - \sum_{i,j} ||\tilde{L}_i^+ - \tilde{L}_j^+||_2^2 \cdot \frac{\partial f(x_i, x_j)}{\partial r} \quad (4.7)$$

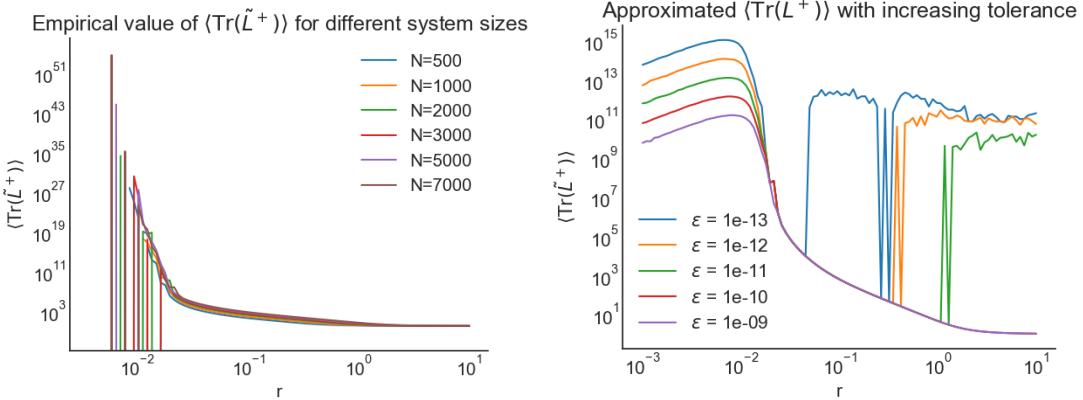
where:

$$\frac{\partial f(x_i, x_j)}{\partial r} = \frac{2}{r^3} e^{-\frac{d\tau(x_i, x_j)}{r^2}} > 0 \quad (4.8)$$

Since  $||\tilde{L}_i^+ - \tilde{L}_j^+||_2^2 > 0$  then  $\frac{\partial Tr[\tilde{L}^+]}{\partial r} < 0$  as required.  $\square$

Thus we cannot observe any divergence of  $Tr[\tilde{L}^+]$  to  $+\infty$  for  $r > 0$ . In fact:

**Lemma 4.2.2.** *For all Soft RGGs for fixed  $N$ ,  $\lim_{r \rightarrow 0^+} Tr[\tilde{L}^+] = +\infty$ .*



**Figure 4.8** log-log plots of a Monte-Carlo estimate (left) and approximation (right) of  $\text{Tr}[\tilde{L}^+]$  against  $r$ , exhibiting numerical error for small  $r$ .

*Proof.* Consider a Soft RGG induced by some point configuration  $X_\rho^N = \{x_i\}_{i=1}^N$ . Since the graph is connected, and the eigenvalues of  $\tilde{L}$  are positive, we use Cauchy-Schwarz:

$$\text{Tr}[\tilde{L}] \cdot \text{Tr}[\tilde{L}^+] = \sum_{i=2}^N \sqrt{\lambda_i^2} \cdot \sum_{i=2}^N \sqrt{\frac{1}{\lambda_i}} \geq (\sum_{i=2}^N 1)^2 = (N-1)^2 \quad (4.9)$$

$$\lim_{r \rightarrow 0^+} \text{Tr}[\tilde{L}] = \lim_{r \rightarrow 0^+} \sum_{i \neq j} w_{ij} = \lim_{r \rightarrow 0^+} \sum_{i \neq j} e^{-\frac{(d_T(x_i, x_j))^2}{2r^2}} = 0 \quad (4.10)$$

Thus:

$$\lim_{r \rightarrow 0^+} \text{Tr}[\tilde{L}^+] \geq \lim_{r \rightarrow 0^+} \frac{(N-1)^2}{\text{Tr}[\tilde{L}]} = +\infty \quad (4.11)$$

as  $\text{Tr}[\tilde{L}] \geq 0$ .

□

Since this holds true for all point configurations, we must have  $\lim_{r \rightarrow 0^+} \langle \text{Tr}[\tilde{L}^+] \rangle = +\infty$ .

To numerically explore the behaviour of  $\text{Tr}[\tilde{L}^+]$ , we first calculate the quantity using the full numerical spectrum of  $L$  returned by `np.linalg.eigvalsh`. For small values of  $r$ , we obtain numerically unstable results over all system sizes  $N$ . Nonetheless, we observe for larger values of  $r$  a monotonically decreasing curve that becomes shallower with increasing system size  $N$ .

To combat numerical issues, we use the approximation 3.1.2, using eigenvalues that exceed a tolerance  $\epsilon$ . However, this approach introduces its own complications - we first observe that below  $r \approx 10^{-2}$ ,  $\tilde{L}^+$  appears to be increasing, which conflicts with analytical results. In addition, more of the initial divergence is captured as tolerance decreases, reflected in the rising peak of the curve. However, decreasing tolerance past  $\epsilon \sim 10^{-10}$  introduces further instabilities for larger values of  $r$ , due to finite precision errors affecting the smallest eigenvalues. These observations should be taken as a caution as to the difficulty of numerically calculating the spectrum of Soft RGGs.

## 4.3 Connection to Euclidean Random Matrices

### 4.3.1 An analytical expression for $\langle \text{Tr}(L^+) \rangle$

By definition, the graph Laplacian of the Hard and Soft RGGs are ERMs. We now state an analytical expression for the ensemble-averaged  $\text{Tr}(L^+)$  using tools from Random Matrix Theory:

**Theorem 4.3.1.** *Let  $\rho_N(\lambda)$  be the spectral density of the Laplacian of the RGG of size  $N$  and  $R_L(z)$  is the resolvent of a given Laplacian  $L$  of the RGG. Suppose that  $\lambda^{-1}\rho_N(\lambda)$  is integrable near zero. Then:*

$$\langle \text{Tr}[L^+] \rangle = N \lim_{\epsilon \rightarrow 0^+} \int_\epsilon^\infty \frac{\rho_N(\lambda)}{\lambda} d\lambda = \left\langle \lim_{\epsilon \rightarrow 0^+} \left[ \frac{a(L)}{i\epsilon} - NR_L(i\epsilon) \right] \right\rangle \quad (4.12)$$

where:

$$a(L) = \begin{cases} 1, & \text{in the case of the Soft RGG} \\ n_c[L], & \text{in the case of the Hard RGG} \end{cases} \quad (4.13)$$

Further, the above holds in the thermodynamic limit if there exists a spectral density  $\rho_\infty$  such that  $\rho_N(\lambda) \rightarrow \rho_\infty(\lambda)$  and  $\lambda^{-1}\rho_\infty(\lambda)$  is integrable near zero.

*Proof.* Consider the Laplacian matrix,  $L$ , of a point configuration of the RGG of size  $N$ , with eigenvalues  $\lambda_1 \leq \dots \leq \lambda_N$ . Let  $\rho_L$  be the empirical spectral density of  $L$ .

$$\text{Tr}[L^+] = \sum_{i=a+1}^N \frac{1}{\lambda_i} = \lim_{\epsilon \rightarrow 0^+} \sum_{i=1}^N \frac{1}{\lambda_i} \mathbf{1}_{\{\lambda_i > \epsilon\}} = N \lim_{\epsilon \rightarrow 0^+} \int_\epsilon^\infty \frac{\rho_L(\lambda)}{\lambda} d\lambda \quad (4.14)$$

First, note that the ensemble average of the empirical spectral density of  $L$ ,  $\langle \rho_L(\lambda) \rangle = \frac{1}{N} \sum_{i=1}^N \langle \delta(\lambda - \lambda_i) \rangle = \rho_N(\lambda)$ . Secondly, note that the random variables  $X_\epsilon := \int_\epsilon^\infty \frac{\langle \rho_L(\lambda) \rangle}{\lambda} d\lambda$  are non-negative and monotonically increasing in  $\epsilon$  to  $\text{Tr}[L^+]$ . By the Monotone Convergence theorem, we have that  $\lim_{\epsilon \rightarrow 0^+} \langle X_\epsilon \rangle = \langle \lim_{\epsilon \rightarrow 0^+} X_\epsilon \rangle$ . Thus:

$$\langle \text{Tr}[L^+] \rangle = N \lim_{\epsilon \rightarrow 0^+} \int_\epsilon^\infty \frac{\langle \rho_L(\lambda) \rangle}{\lambda} d\lambda = N \lim_{\epsilon \rightarrow 0^+} \int_\epsilon^\infty \frac{\rho_N(\lambda)}{\lambda} d\lambda \quad (4.15)$$

As before, let  $L$  be the Laplacian of a point configuration of the RGG of size  $N$ . Then  $L$  has eigenvalue zero with algebraic multiplicity  $a$ . From the definition of its resolvent  $R_L(z)$ , and approaching zero from the upper half-plane, i.e.  $z = i\epsilon, \epsilon \in \mathbb{R}^+$ :

$$\lim_{\epsilon \rightarrow 0^+} \left[ \frac{a}{i\epsilon} - NR_L(i\epsilon) \right] = \lim_{\epsilon \rightarrow 0^+} \left[ \frac{a}{i\epsilon} - \text{Tr}[(i\epsilon - L)^{-1}] \right] = \text{Tr}[L^+] \quad (4.16)$$

as  $\text{Tr}[(z - L)^{-1}] = \sum_{i=1}^N \frac{1}{z - \lambda_i} = \frac{a}{z} + \sum_{i=a+1}^N \frac{1}{z - \lambda_i}$ , so the first term removes the contribution of the pole at zero in the resolvent. Taking an ensemble average on both sides yields the expression as required.

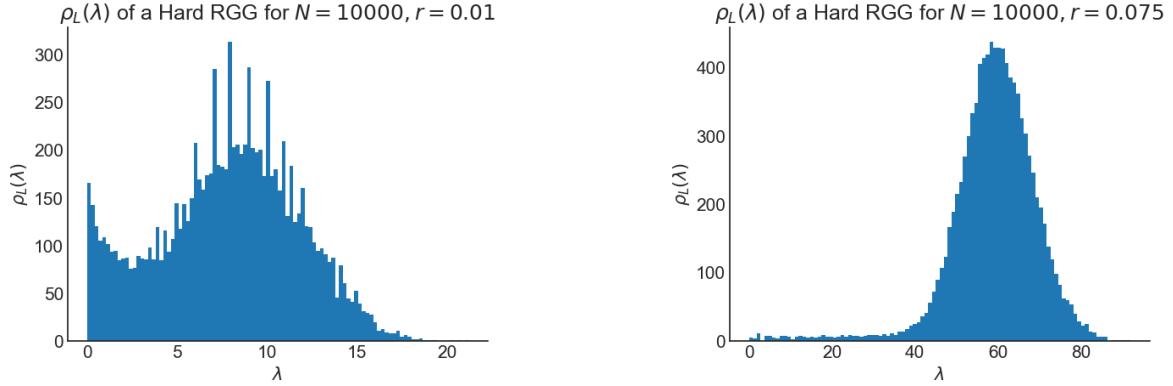
Note that the Monotone Convergence Theorem still holds if  $\text{Tr}[L^+]$  is infinite. Given the stated conditions, we thus have  $N \lim_{\epsilon \rightarrow 0^+} \int_\epsilon^\infty \frac{\rho_N(\lambda)}{\lambda} d\lambda \rightarrow \langle \text{Tr}[L_\infty^+] \rangle$  in the thermodynamic limit, and the second equality follows as the limit is pointwise.

□

We note that this result is easily extended to other graphs, defining  $a$  as the number of connected components.

### 4.3.2 Motif Expansion of the Hard RGG spectrum

While the above is an important analytical expression for  $\tau$ , it is generally the case that no closed-form expression for the spectral density of graphs induced by percolation processes. In particular, the spectrum of the Hard RGG is highly non-trivial, exhibiting delta spikes for low values of  $r$ :



**Figure 4.9** Empirical Spectral Densities of the Hard RGG for  $r = 0.01$ ,  $r = 0.075$ .

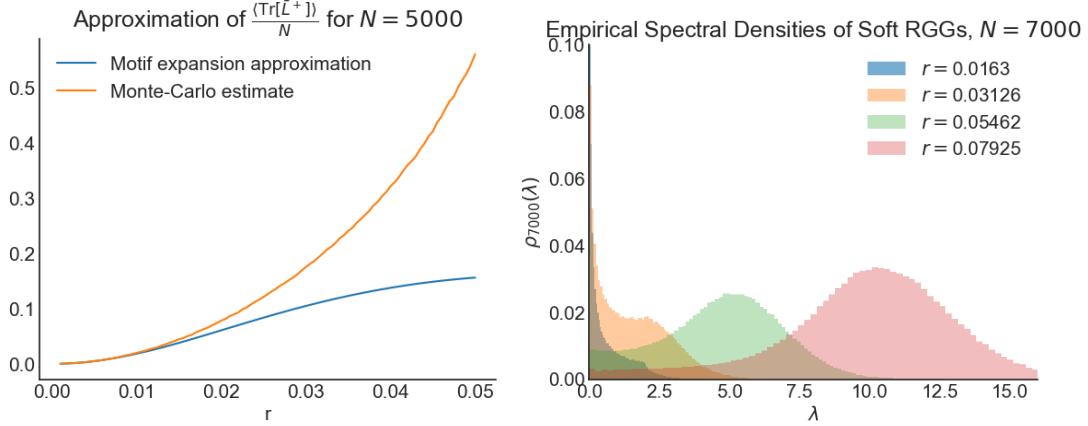
The most prominent spikes occur at integer values. This is explained by the prevalence of isolated cliques; recall that the spectrum of an  $N$ -clique consists of 0 with algebraic multiplicity 1, and  $N$  with algebraic multiplicity  $N - 1$ . At low radius  $r$ , small cliques tend to form in isolation, contributing integer eigenvalues to the spectrum. Other subgraph topologies, or motifs, also form and contribute to the spectral density. To approximate the spectral density of the Hard RGG  $\rho_N$  in the thermodynamic limit, we propose an approximate motif expansion in the sparse subcritical regime  $r \ll r_c$ ,  $N \gg 1$ :

$$\rho_N(z) \approx \frac{1}{Z} \sum_{n=1}^{n_{max}} \sum_{M \in \mathcal{M}_n} \rho_M \binom{N}{n} P(M) \quad (4.17)$$

where we sum up the spectral densities of each motif  $M \in \mathcal{M}_n$  of size  $n$ , multiplied by the (Binomially) expected number of  $M$ , where  $P(M)$  is the probability of a motif occurring in isolation, up to some maximum motif size  $n_{max}$ , controlling the order of truncation, and  $Z$  is a normalising constant.

This is analogous to cluster expansions in percolation theory, where one calculates the expected frequency of isolated subgraphs of a given size, especially in the subcritical regime [32, 33]; this develops on the existing framework in the 1D RGG [34]. This now allows us to view the spectral density in a geometric light, although calculating  $P(M)$  for large motifs is likely intractable. This expansion makes the strong assumption that motif occurrences are independent, which breaks down for small  $N$ , due to the constraints of system size and lack of self-averaging effects, and for large  $r$ , as motif isolation is less likely due to the percolating component. We also assume that motifs do not overlap, as the Binomial model breaks down for large  $P(M)$  or large motif size.

We now check the accuracy of the expansion for  $n_{max} = 2$ . The probability of an isolated node,  $M_1$ , occurring is clearly  $P(M_1) = (1 - \frac{\pi r^2}{L^2})^{N-1}$ . Denote  $M_2$  as the motif defined by two connected nodes; this is the only possible motif of size 2. We derive the following expression in Appendix 4.3.2:



**Figure 4.10** Comparison of Motif expansion approximation,  $n_{\max} = 2$  of  $\frac{\langle \text{Tr}[L^+] \rangle}{N}$  and Monte-Carlo estimate (left), Crossover of the Soft RGG spectral density (right)

**Lemma 4.3.2.** For  $r \leq \frac{L}{2}$ :

$$P(M_2) = \int_0^r (1 - \frac{A(x, r)}{L^2})^{N-2} \frac{2\pi x}{L^2} dx \quad (4.18)$$

where:

$$A(x, r) = 2\pi r^2 - (2r^2 \cos^{-1}(\frac{x}{2r}) - \frac{x}{2} \sqrt{4r^2 - x^2}) \quad (4.19)$$

Since the only non-zero eigenvalue is contributed by  $M_2$ , which has spectrum  $\{0, 2\}$ , then  $\langle \text{Tr}[L^+] \rangle \approx N \frac{N_2}{2(N_1 + N_2)}$ , where  $N_1 = NP(M_1)$ ,  $N_2 = \frac{N(N-1)}{2} P(M_2)$ . We test this approximation by comparing to a Monte-Carlo estimate of  $\frac{\langle \text{Tr}[\tilde{L}^+] \rangle}{N}$  as before, with  $N = 5000$  and 50 repetitions. We see that even for a large truncation of the expansion, we obtain an accurate approximation for small  $r \lesssim 0.01$ .

### 4.3.3 Soft RGGs and Gaussian Euclidean Random Matrices

In the case of the soft kernel, the graph Laplacian is a Euclidean random matrix with Gaussian kernel. While there is no closed form for the spectrum or resolvent, there exists high and low density approximations [10, 35] for the adjacency matrix.

As  $r \rightarrow 0$ , by 4.2.2 we have that  $\langle \text{Tr}(\tilde{L}) \rangle \rightarrow 0$ . Since  $\text{Tr}(\tilde{L}) = \sum_{i=1}^N \lambda_i$ , and  $\lambda_i \geq 0$ , we see that all eigenvalues must uniformly converge to zero, and hence  $\lim_{r \rightarrow 0} \rho(\lambda) = \delta(\lambda)$ . On the other hand, for a finite Soft RGG of size  $N$ ,  $\lim_{r \rightarrow \infty} \rho(\lambda) = \{0, N, \dots, N\}$ , as  $\lim_{r \rightarrow \infty} w_{ij} = \lim_{r \rightarrow \infty} e^{-\frac{d_{\tau}(x_i, x_k)}{r^2}} = 1$ . So, as  $r \rightarrow \infty$ ,  $\tilde{L}$  converges to the Laplacian of an (unweighted)  $N$ -clique. It is known that the limiting spectral density in the thermodynamic limit exists given some density  $\rho$  [36], and results exist in the limit as the dimension  $d \rightarrow \infty$  [37].

Plotting the evolution of the empirical spectral density for increasing  $r$  in Figure 4.10, we observe a smooth transition from a Boson-like regime, characterised by a high density of low-frequency modes, to a Gaussian-like profile that increases in mean. This transition appears to correspond to the radius in which the numerical estimate becomes accurate in Figure 4.8. We recall that  $\langle \text{Tr}[L^+] \rangle$  can be expressed as the limiting  $-1^{\text{th}}$  moment of the spectral density - thus it is an appropriate object to study the formation of a Boson-like peak for low  $r$  and the crossover to a connected regime.

# Chapter 5

## Conclusion

An article by Radl et al. argues (and is titled) "The Resistance Distance is Meaningless for Large Random Geometric Graphs" [38]; yet this thesis has shown quite the opposite. The crucial distinction is the limiting regime - in Radl's case, as  $N \rightarrow \infty$ ,  $\rho \rightarrow \infty$  with points sampled from density  $p$  in a compact space; then the resistance distance between points  $a$  and  $b$  is simply  $\frac{1}{p(a)} + \frac{1}{p(b)}$ . However, in the thermodynamic limit, we have observed that effective resistance intuitively diverges at criticality in the case of the Hard RGG, and captures the formation of a Boson-like peak as  $r \rightarrow 0$  in the case of the Soft RGG.

Concretely, we have unified concepts from Statistical Mechanics and Spectral Graph Theory to extend the notion of the Kirchhoff Index to disconnected graphs, and hence construct a percolation observable from the statistics of the graph Laplacian. Numerical experiments have demonstrated textbook finite-size scaling behaviour, and the critical radius and correlation length exponent have been extracted from this quantity, with values agreeing with the literature within an acceptable tolerance. An analytical expression for the ensemble average of  $\text{Tr}[L^+]$  was derived, connecting to standard Random Matrix Theoretical quantities such as the spectral density and the resolvent. We then proposed an approximate motif expansion of the sparse Hard RGG spectral density, and identified a crossover in the case of the Soft RGG, exhibiting a Boson-like peak as  $r \rightarrow 0$ .

Classical approaches to random resistor models rely on renormalisation group techniques, for example the Node-Link-Blob picture [39]. However, recent work by Armstrong et al. extends these ideas using tools from homogenisation theory, rigorously describing the behaviour of harmonic functions on the percolating cluster in the supercritical regime of the Hard RGG, above a certain lengthscale [40, 41]. Since  $\text{Tr}[L^+]$  can be expressed as the sum of harmonic Green's functions  $\sum_x G(x, x)$ , future work should leverage these results to formalise the divergence of  $\text{Tr}[L^+]$  as  $r \rightarrow r_c^+$ .

Crucially, the quantity  $\text{Tr}[L^+]$  is not a specific observable to the RGG, but is well-defined for any disconnected graph or percolation model. If, in fact, the Hard RGG in the thermodynamic limit belongs to the standard percolation universality class, then the same methodology presented in this thesis applies naturally to all percolation models in this class. Further work is imperative to confirm any universal behaviour of  $\text{Tr}[L^+]$  over different percolation models, estimate critical exponents, and relate them to the existing conductance coefficients.

# Appendix A

## Appendix

### A.1 Proof of Equation 4.6

Let  $L$  be the Laplacian matrix of an undirected connected graph. From the expression in Theorem 4.3 in [42], and using the fact that  $L$  is symmetric implies  $L^+$  is symmetric:

$$\frac{\partial L^+}{\partial w_{ij}} = -L^+ \frac{\partial L}{\partial w_{ij}} L^+ + (L^+)^2 \frac{\partial L^T}{\partial w_{ij}} (I - LL^+) + (I - LL^+) \frac{\partial L^T}{\partial w_{ij}} (L^+)^2 \quad (\text{A.1})$$

It is known that  $LL^+$  and  $L^+L$  are projection operators onto the image of  $L$ ,  $\text{Span}(\mathbf{1})^\perp$ ; equivalently,  $LL^+ = L^+L = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$  [43]. Since  $L = D - A$ , then:

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial D}{\partial w_{ij}} - \frac{\partial A}{\partial w_{ij}} = e_i e_i^T + e_j e_j^T - e_i e_j^T - e_j^T e_i \quad (\text{A.2})$$

which has  $\mathbf{1}$  as a left and right eigenvector. Then:

$$(L^+)^2 \frac{\partial L^T}{\partial w_{ij}} (I - LL^+) = (L^+)^2 (e_i e_i^T + e_j e_j^T - e_i e_j^T - e_j^T e_i) \left( \frac{1}{n} \mathbf{1} \mathbf{1}^T \right) = 0 \quad (\text{A.3})$$

and by the same logic,  $(I - LL^+) \frac{\partial L^T}{\partial w_{ij}} (L^+)^2 = 0$ . Using the above, and the linearity and cyclic property of the trace:

$$\frac{\partial \text{Tr}[L^+]}{\partial w_{ij}} = \text{Tr}\left[\frac{\partial L^+}{\partial w_{ij}}\right] = -\text{Tr}\left[-L^+ \frac{\partial L}{\partial w_{ij}} L^+\right] = -\text{Tr}\left[(L^+)^2 \frac{\partial L}{\partial w_{ij}}\right] \quad (\text{A.4})$$

Substituting A.2, and using  $\text{Tr}[Axy^T] = y^T Ax$ :

$$\frac{\partial \text{Tr}[L^+]}{\partial w_{ij}} = -[(L^+)^2]_{ii} - [(L^+)^2]_{jj} + -[(L^+)^2]_{ij} - [(L^+)^2]_{ji} \quad (\text{A.5})$$

Since  $L^+$  is symmetric, then  $(L^+)^2$  is symmetric. Then:

$$\frac{\partial \text{Tr}[L^+]}{\partial w_{ij}} = -[(L^+)^2]_{ii} - [(L^+)^2]_{jj} + 2[(L^+)^2]_{ij} = -(e_i - e_j)^T (L^+)^2 (e_i - e_j) \quad (\text{A.6})$$

Rewriting as an  $L^2$  norm:

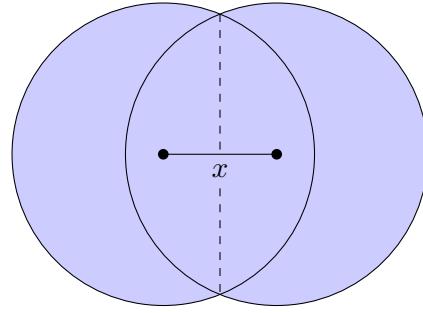
$$\frac{\partial \text{Tr}[L^+]}{\partial w_{ij}} = -\|L^+(e_i - e_j)\|_2^2 = -\|L_i^+ - L_j^+\|_2^2 \quad (\text{A.7})$$

as required.

## A.2 Proof of Lemma 4.3.2

$P(M_2)$  is equal to the probability that two points are within a distance of  $r$ , and none of the  $N - 2$  remaining points fall within the union of the two circles of radius  $r$  around each point.

We first consider the area of this region, given two points are a toroidal distance of  $x < r$ ,  $A(x, r)$ , pictured as the shaded region below:



The area of a circular segment of height  $h$  and radius  $r$  is:

$$A_{seg} = 2 \int_{r-h}^r \sqrt{r^2 - y^2} dy = r^2 \cos^{-1}(1 - \frac{h}{r}) - (r - h) \sqrt{r^2 - (r - h)^2} \quad (\text{A.8})$$

Substituting height  $h = r - \frac{x}{2}$  and radius  $r$ , we thus find  $A(x, r)$ :

$$A(x, r) = 2\pi r^2 - 2A_{seg} = 2\pi r^2 - (2r^2 \cos^{-1}(\frac{x}{2r}) - \frac{x}{2} \sqrt{4r^2 - x^2}) \quad (\text{A.9})$$

Then the probability that no other point falls within  $A(x, r)$  is simply  $(1 - \frac{A(x, r)}{L^2})^{N-2}$ .

Due to the translational invariance of  $\square_L$  under  $d_\tau$ , we may translate the space so that a given point lies at the origin; then the largest circle inscribed in  $\square_L$  has radius  $\frac{L}{2}$ . Within this radius, the set of points at distance  $x$  from the origin forms an annulus of area  $2\pi x dx$ ; thus the p.d.f. of  $x$  is  $\frac{2\pi x}{L^2}$  for  $x \leq \frac{L}{2}$ . Integrating over  $0 \leq x \leq r$ :

$$P(M_2) = \int_0^r (1 - \frac{A(x, r)}{L^2})^{N-2} \frac{2\pi x}{L^2} dx \quad (\text{A.10})$$

## A.3 GitHub Repository

<https://github.com/raahweng/Random-Geometric-Graphs>

# Bibliography

- [1] E. N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.
- [2] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2nd edition, 1999. See Section 2.2 for the monotone coupling construction.
- [3] R. Meester and R. Roy. *Continuum Percolation*. Cambridge Tracts in Mathematics. Cambridge University Press, 1996.
- [4] J. P. J. S. H. Z. S. Florek, K. Sur la liaison et la division des points d'un ensemble fini. *Colloquium Mathematicum*, 2(3-4):282–285, 1951.
- [5] L. Vietoris. Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen. *Mathematische Annalen*, 97:454–472, 1927.
- [6] K. R. Gabriel and R. R. Sokal. A new statistical approach to geographic variation analysis. *Systematic Biology*, 18(3):259–278, 09 1969.
- [7] M. Penrose. *Random Geometric Graphs*. Oxford University Press, 05 2003.
- [8] G. E. Pike and C. H. Seager. Percolation and conductivity: A computer study. i. *Phys. Rev. B*, 10:1421–1434, Aug 1974.
- [9] E. T. Gawlinski and H. E. Stanley. Continuum percolation in two dimensions: Monte Carlo tests of scaling and universality for non-interacting discs. *Journal of Physics A: Mathematical and General*, 14(8):L291, aug 1981.
- [10] M. Mézard, G. Parisi, and A. Zee. Spectra of Euclidean random matrices. *Nuclear Physics B*, 559(3):689–701, 1999.
- [11] C. Ganter and W. Schirmacher. Euclidean random matrix theory: Low-frequency non-analyticities and Rayleigh scattering. *Philosophical Magazine*, 91(13-15):1894–1909, 2011.
- [12] A. Goetschy. *Light in disordered atomic systems: Euclidean matrix theory of random lasing*. Theses, Université de Grenoble, Nov. 2011.
- [13] M. Hamidouche. *Spectral analysis of random geometric graphs*. PhD thesis, 05 2020.
- [14] C. P. Dettmann, O. Georgiou, and G. Knight. Spectral statistics of random geometric graphs. *EPL (Europhysics Letters)*, 118(1):18003, Apr. 2017.
- [15] K. Adhikari, R. J. Adler, O. Bobrowski, and R. Rosenthal. On the spectrum of dense random geometric graphs, 2020.

- [16] D. Stauffer and A. Aharony. *Introduction to Percolation Theory: Second Edition*. Taylor & Francis, London, 2nd edition, 1992.
- [17] H. Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Communications in Mathematical Physics*, 74:41–59, 1980.
- [18] G. Grimmett. *Probability on Graphs: Random Processes on Graphs and Lattices*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2 edition, 2018.
- [19] D. J. Klein and M. Randic. Resistance distance. *Journal of Mathematical Chemistry*, 12: 81–95, 1993.
- [20] A. Ghosh, S. Boyd, and A. Saberi. Minimizing effective resistance of a graph. *SIAM Review*, 50(1):37–66, 2008.
- [21] B. Mohar. Isoperimetric numbers of graphs. *Journal of Combinatorial Theory, Series B*, 47 (3):274–291, 1989.
- [22] G. Livan, M. Novaes, and P. Vivo. *Introduction to Random Matrices*. Springer International Publishing, 2018.
- [23] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2): 298–305, 1973.
- [24] D. P. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, 4 edition, 2014.
- [25] A. Malthe-Sørenssen. *Percolation Theory Using Python*. 01 2024.
- [26] J. Houdayer and A. K. Hartmann. Low-temperature behavior of two-dimensional gaussianising spin glasses. *Phys. Rev. B*, 70:014418, Jul 2004.
- [27] A. Sorge. pyfssa 0.7.6. <https://github.com/andsor/pyfssa/tree/master>, 2015. Accessed: 2025-05-30.
- [28] O. Melchert. autoscale.py - a program for automatic finite-size scaling analyses: A user’s guide, 2009.
- [29] A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964.
- [30] P. Balister, B. Bollobás, and M. Walters. Continuum percolation with steps in the square or the disc. *Random Structures & Algorithms*, 26(4):392–403, 2005.
- [31] S. Mertens and C. Moore. Continuum percolation thresholds in two dimensions. *Physical Review E*, 86(6), Dec. 2012.
- [32] M. F. Sykes and M. Glen. Percolation processes in two dimensions. i. low-density series expansions. *Journal of Physics A: Mathematical and General*, 9(1):87, jan 1976.
- [33] P. Mann, V. A. Smith, J. B. O. Mitchell, C. Jefferson, and S. Dobson. Exact formula for bond percolation on cliques. *Phys. Rev. E*, 104:024304, Aug 2021.
- [34] A. Nyberg, T. Gross, and K. E. Bassler. Mesoscopic structures and the laplacian spectra of random geometric graphs, 2014.
- [35] G. Parisi. Euclidean random matrices: Solved and open problems. In É. Brézin, V. Kazakov, D. Serban, P. Wiegmann, and A. Zabrodin, editors, *Applications of Random Matrices in Physics*, pages 219–260, Dordrecht, 2006. Springer Netherlands.

- [36] C. Bordenave. Eigenvalues of euclidean random matrices. *Random Structures & Algorithms*, 33(4):515–532, 2008.
- [37] X. Zeng. Spectrum of large euclidean random matrices generated from  $l^p$  ellipsoids. *Linear and Multilinear Algebra*, 70(8):1403–1422, 2022.
- [38] A. Radl, U. Luxburg, and M. Hein. The resistance distance is meaningless for large random geometric graphs. 01 2009.
- [39] A. Coniglio. Cluster structure near the percolation threshold. *Journal of Physics A: Mathematical and General*, 15(12):3829, dec 1982.
- [40] S. Armstrong and P. Dario. Elliptic regularity and quantitative homogenization on percolation clusters, 2017.
- [41] S. Armstrong and R. Venkatraman. Quantitative homogenization and large-scale regularity of poisson point clouds, 2023.
- [42] G. H. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2):413–432, 1973.
- [43] I. Gutman and W. Xiao. Generalized inverse of the laplacian matrix and some applications. *Bulletin (Académie serbe des sciences et des arts. Classe des sciences mathématiques et naturelles. Sciences mathématiques)*, (29):15–23, 2004.