GLOBAL INFORMATION LOSS AND CRITICALITY IN RESISTANCE MATRICES

by

Gergely Ódor

Supervisor:

Bálint Virág

THESIS

Presented to the Department of Mathematics,

Central European University

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Mathematics with Its Applications

CENTRAL EUROPEAN UNIVERSITY

May 26, 2017

Acknowledgements

I would like to express my gratitude to my advisor, Prof. Bálint Virág of the Alfréd Rényi Institute of Mathematics, for introducing me into research in mathematics. At first I had trouble asking good mathematical questions as all of my previous experience came from computer science, applied mathematics and neurobiology. Bálint showed me how to formalise my intuition, he pointed me in promising directions when I was stuck, and it is due to Bálint's vision that I can see how the topic of my thesis fits in a bigger picture. I would like to thank him for all the inspiration, which helped me find my place in a new environment after moving back to my home country.

I wish to thank the Mathematics Department of the Central European University for helping me expand my mathematical foundations and for providing financial support for this past year. I would especially like to thank Prof. Pál Hegedűs, Director of the MS program in Mathematics, for his constant attention to every aspect of my progress.

Finally, I would like to thank the support of my friends and family, and above all, my loving girlfriend, who stood by me and kept me motivated this entire time, despite the difficulties of writing this thesis and other concurrent issues.

Abstract

Resistance matrices of graphs exhibit non-trivial global information about the graph such as community structure. This property seams to make them good candidates for developing fast machine learning algorithms for very large graphs. However, it has been recently reported that for large graphs the resistance matrices are dominated by something as local as the degrees of the vertices [13]. This phenomenon is sometimes called *Global Information Loss* (GIL), and it suggests that using resistance matrices in data science should be discouraged. However, the significance of GIL is not yet well understood; there have been arguments that with simple operations the global information can still be accessed [13, 16]. Conversely it has also been reported that GIL does not have significant effect on the numerical results in applications [10].

In this thesis we aim to build the theoretical background to better understand the GIL phenomenon. To give a rigorous definition for GIL, first we build a new notion of graph convergence based on resistance matrices, and we say that GIL occurs if the limit object is trivial (low rank in some sense). The properties of the new graph convergence definition is also of independent interest. Our notion of limit object is similar to the graphon, although due to the GIL, as a simple corollary of previous results, the limit object of dense graphs is always trivial. We also prove that 2D grid graphs and binary trees have trivial limit; these are examples of sparse, bounded degree sequences that have the GIL property.

However, we also exhibit examples of sequences where the limit objects are non-trivial: path graphs and conditioned critical Galton-Watson trees. Since we suspect that more examples of such sequences with non-trivial limits will come from random graph models at the critical point (e.g. critical percolation) we call these sequences *critical*.

The main results of this thesis are the proofs of criticality and triviality of several examples of graph models. In the end we also show numerical simulations to strengthen the robustness of our arguments, and outline a direction for practical applications.

Table of Contents

Chapter 1

Introduction

Effective resistances have played an important role for a long time in physics and engineering and more recently in probability theory and computer science. The notion originates from the theory of electric networks. Given a resistor network, the electric resistance between two points i and j measures their potential difference if we push a unit current flow across the entire network from i to j. In other words, it is the resistance of the entire network if the input/output wires are attached to i and j. The same notion can be defined for any weighted graph $G = (V, E, w)$ where the edge weights set the conductances of the links of the corresponding electric network. The resistance matrix of a weighted graph is simply the matrix R where the (i, j) entry is the magnitude of effective resistance between i and j.

Resistance matrices connect many seemingly unrelated notions. On one hand they are tightly linked with the Laplacian inverse of G: $R_{i,j} = (\mathbf{e}_i - \mathbf{e}_j)^T L^\dagger (\mathbf{e}_i - \mathbf{e}_j)$. Laplacian matrices are important in the computation of global properties of graphs, such as the maximum-flow or the number of spanning trees. Their applications range over machine learning, partial differential equations, spectral graph theory, etc [18]. Moreover, taking advantage of the special structure of Laplacian matrices, the effective resistances can be computed in $\widetilde{O}(|E|)$ time with the Spielman-Teng solver [17]. On the other hand, entries of $R_{i,j}$ are just the normalised expected commute times between i and j for a simple random walk on G. For this reason instead effective resistance, often the term resistance distance is used. Again, this result implies that resistance matrices tell us something about the global structure of the graph.

The advantages of resistance matrices is thus two-fold; they are fast to (approximately)

compute and they carry non-trivial, global information. Therefore, they seam to be highly applicable for very large graphs, where more accurate approaches would be intractable. However, recently it was shown that for large and dense graphs the effective resistances loose their global meaning; they only contain something very local such as the degree of each point [13]. Intuitively, in a large and dense graph the random walk mixes very quickly, we reach the neighbourhood of every point. The bottleneck for the commute time is hitting the points exactly, which of course is dominated by the local information. This phenomenon is referred to as Global Information Loss by [16]. There have been attempts to correct for GIL, either by just subtracting the known local information the [13] or modifying slightly modifying the definition of effective resistances [16]. On the other hand, there have been reports where resistance matrices and the Spielman-Teng solver have been successfully applied and produced state of the art results for large-scale clustering [10]. While the authors mention the issue, they do not report any effect of GIL on their results.

The questions in this thesis are motivated by understanding the practical questions above, however, we take a more theoretical path. Since the GIL appears only for large graphs we attempt to study it in the limits of graph sequences. To do this, we define a new notion of graph convergence based on the resistance matrices, where the limit object is in some sense an generalised infinite resistance matrix. It is of independent theoretical interest how this new notion of graph convergence behaves with respect to existing notions. Due to the GIL phenomenon, for many dense graphs the limit will be trivial, however, we are also able to exhibit graph sequences for which the limit object has interesting structure. Since these interesting graph sequences seem to come from random graph models near the critical point (critical Galton-Watson tree, critical percolation), we call such sequences critical. Critical graphs are therefore graphs where the GIL phenomenon does not occur. This thesis mostly focuses on developing the theoretical background of this new notion of graph convergence and proving the criticality of the following graph sequences: path graphs, 2D grids, binary trees, critical Galton-Watson trees.

In section 2 we give a more rigorous treatment of the theory of electric resistances and

develop tools needed in the subsequent proofs. In section 3 we define the new notion of resistance based graph convergence and we present fundamental examples and results. In section 4 we define critical graphs and prove our three main results: path graphs and critical Galton-Watson trees are critical whereas 2D grid graphs are not. Finally, in section 5 we strengthen our theoretical arguments with numerical experiments.

Chapter 2

Effective resistance

In this section we provide the important rigorous definitions and theorems about effective resistances needed in later chapters. Complete proofs will be omitted; besides the exact statements only intuition and citations are given. Some of our notation and many of the definitions follow Section 2 of [15].

While effective resistances have already been (intuitively) defined in the Introduction, we prefer this more concise theoretical definition.

Definition 1. The *resistance matrix* of a weighted graph $G = (V, E, w)$ with edge is

$$R_{i,j} = L^{\dagger}_{i,i} + L^{\dagger}_{j,j} - 2L^{\dagger}_{i,j} = (\mathbf{e}_i - \mathbf{e}_j)^T L^{\dagger}(\mathbf{e}_i - \mathbf{e}_j) \tag{2.1}$$

where $L^{\dagger}$ is the Moore-Penrose inverse of the Laplacian matrix $L = D - A$ of the graph. Throughout this thesis, we only consider connected graphs with non-negative edge weights. This condition implies that apart from the lowest, all eigenvalues are strictly positive, and the nullspace of L is spanned by $\mathbf{1}$.

This definition agrees with the physics interpretation of the effective resistance. Let's consider an electrical network induced by the graph G, where the edge weights become the conductances of the connections in the network. If $\mathbf{i}_{ij} = \mathbf{e}_i - \mathbf{e}_j \in \mathbb{R}^n$ is a vector that represents and external battery connected to nodes i and j that maintains a unit current flow, then by Kirchoff's laws $\mathbf{v} = L^{\dagger}\mathbf{i}_{ij}$ is exactly the vector of electrical potentials in the network with this current flow (up to a shift with $\mathbf{1}$). Then the magnitude of the effective resistance between i and j is $|\mathbf{v}_i - \mathbf{v}_j|$ which agrees with our definition (see [15] for a detailed proof).

2.1 Probabilistic interpretations

As outlined in the Introduction, effective resistances have important probabilistic interpretations. The reason behind this is that electric current itself can be modelled as a random walk of tiny electrons. Let's we consider a resistor network and apply potential 0 to vertex a and 1 to set Z. Then by Proposition 2.2 in [15] the resulting current flow $\mathbf{i}$ can be viewed as $E[S_{xy} - S_{yx}]$, where S_{xy} is the number of transitions from x to y of the simple random walk started at a and absorbed by Z.

The following lemma makes the connection with effective resistances more direct. This statement is the way [15] defines effective resistance (page 25).

Lemma 1. Let $R(a, Z)$ be the effective resistance between a and z, where z is the supernode that collects all vertices of Z. Let $P(a \to Z)$ be the probability that a simple random walk started at a hits Z before returning to a. Then

$$R(a, Z) = \frac{1}{d_a P(a \to Z)} \tag{2.2}$$

where d_a is the degree of a.

A similar connection can be made for two-point resistances and commute times. This lemma will not be used in later proofs, but we feel that it cannot be omitted from such a review.

Lemma 2 (Corollary 2.21 in [15])**.** Let the random variable τ_{ij} be the number of steps it takes for a random walk on graph G to reach node j starting from node i. Then, for the *commute time* $\mathcal{C}_{ij} = \mathbf{E}[\tau_{ij}] + \mathbf{E}[\tau_{ji}]$ we have

$$R_{i,j} = \frac{1}{vol(G)} \mathcal{C}_{ij} \tag{2.3}$$

, where $vol(G) = \frac{1}{2} \sum_{i \in V} d_i$ is the sum of the edge weights.

2.2 Energy of current flow

We already saw that the effective resistance between nodes i and j is tightly linked with the unit flow from i to j. We will make this connection even tighter.

Definition 2. Let's direct the edges E in some arbitrary way. A *flow* between source $s \in V$ and sink $t \in V$ is a function $F : E \to \mathbb{R}$, which satisfies the conservation law $\sum_{e \in E_+(v)} F(e) - \sum_{e \in E_-(v)} F(e) = 0$ for every node $v \in V \setminus \{s, t\}$, where $E_+(v)$ and $E_-(v)$ are the edges entering and leaving v respectively. The *energy* of a flow is $\mathcal{E}(F) = \sum_{e \in E} \frac{1}{w(e)} F(e)^2$, where $w(e)$ is the weight of edge e in G.

The energy defined in the previous definition again follows the intuition from physics. The if the flow is the current flow, then the above formula gives the $\sum_{e \in E} R_e I_e^2$ formula from high-school physics. It turns out that indeed, the current flow has special meaning in our mathematical model as well. It is the flow with minimal energy.

Theorem 1 (Thompson's Principle, page 35 in [15])**.** *Let F be any unit flow with source i and sink j of unit magnitude (meaning $\sum_{e \in E_+(i)} F(e) - \sum_{e \in E_-(i)} F(e) = 1$), and let $\mathbf{i}_{i,j}$ be the unit current flow from i to j. Then if $F \neq \mathbf{i}_{i,j}$ we have*

$$R_{i,j} = \mathcal{E}(\mathbf{i}_{i,j}) < \mathcal{E}(F) \tag{2.4}$$

This theorem will be used to give upper bounds on $R_{i,j}$ by constructing unit flows in the proof of the triviality of 2D grids. For the proof of path and tree graphs the following simple corollary will be useful.

Corollary 1. The resistance distance on a tree G is the same as the distance on G' with reciprocal edge weights.

Proof. Since there is only one path between i and j, the only possible unit flow gives value 1 to each edge on the path and zero everywhere else. □

2.3 Global information loss

In this subsection we quickly review the results of [14] on GIL. The authors focus on two geometric graph models important in machine learning ($\epsilon-$graphs and knn-graphs) and they propose two techniques (flow based and spectral) to prove that the effective resistances converge to $\frac{1}{d_i}+\frac{1}{d_j}$, where $d_i = \sum_j w(i,j)$ denotes the degrees of the nodes and $w(i,j)$ are the edge weights of G. In this paper we are mostly interested in the spectral arguments since they generalise to more general graph models. The main result depends on the following proposition

Theorem 2 (Proposition 5 in [14])**.** *For any graph G*

$$\left| R_{ij} - \left(\frac{1}{d_i} + \frac{1}{d_j} \right) \right| \leq 2 \left(\frac{1}{1-\lambda_2} + 2 \right) \frac{w_{max}}{d_{min}^2} \tag{2.5}$$

where λ_2 is the second largest eigenvalue of the transition matrix of the random walk of G ($P = D^{-1}W$).

Luxburg et al. use this theorem to show that graphs with high minimum degree and not too low spectral gap, the GIL phenomenon holds. They are able to prove that a rather large class of graphs satisfy this property:

Corollary 2 (Corollary 13 in [14])**.** For random graphs on n nodes with expected minimum degree $\overline{d_{min}} = \omega(log(n))$ and $i \neq j$ we have

$$\left| R_{ij} - \left(\frac{1}{d_i} + \frac{1}{d_j} \right) \right| = O\left(\frac{1}{log(2n)} \right) \to 0 \tag{2.6}$$

In particular, Corollary 2 covers for Erdős Rényi graphs with $p = \omega\left(\frac{log(n)}{n}\right)$. The proof of Corollary 2 relies on advanced techniques by [6] to bound λ_2 by d_{min} in a random graph.

Luxburg et al. point out in their discussion that these methods don't generalise to graphs with bounded degree, e.g. grid graphs, and they are not aware of any previous work from which the convergence of the resistances of such graphs would follow easily. In chapter 4 we will prove a result for grid graphs similar to Corollary 2, although the limit of R_{ij} will be $\frac{1}{\pi}$ instead of $\left(\frac{1}{d_i} + \frac{1}{d_j}\right)$.

Chapter 3

Graph Convergence

3.1 Graph convergence

Convergence of graphs is a relatively young field of study. The first definitions for the convergence of bounded degree graphs was introduced in 2001 [4], and for dense graphs the definition of the limit first appeared in 2006 in [5], and later that year the corresponding limit appeared in [12]. The dense graph convergence boils down to the convergence of subgraph statistics:

Definition 3. A dense sequence of graphs G_n is convergent if

$$t(F, G_n) = \frac{N(F, G_n)}{n^{|V(F)|}}$$

converge for every connected graph F, where $N(F, G)$ is the number of subgraphs in G isomorphic of F.

Definition 4. The limit object of dense graph convergence is the so called graphon, a two-measurable function $W : [0, 1]^2 \to [0, 1]$, which satisfies

$$t(F, W) = \int_{[0,1]^k} \prod_{ij \in E(F)} W(x_i, x_j) dx_1 \dots dx_k \tag{3.1}$$

It is a non-trivial theorem in that every convergent sequence in the sense of Definition 3 has a limit graphon. We will revisit this idea in Chapter 3 in our definition of graph convergence.

For the sparse regime, the definition relies on a more involved probabilistic approach:

Definition 5. A bounded degree sequence of graphs G_n is (Benjamini-Schramm) convergent if the sequence $(0, G_n)$ of random rooted graphs (deterministic G_n with uniform random root 0) weekly converges to $(0, G_\infty)$, where G_∞ is a random graph on countably infinitely many vertices. In this case the limit object is a distribution over rooted countable graphs.

We need different definitions for the sparse and dense case because both definitions are meaningless (or trivial) in the other regime and there has not yet been developed any general theory that would apply to both cases. In fact the two regimes capture different properties, they are two different worlds. For a thorough treatment of graph convergence we refer to [11].

3.2 Resistance based graph convergence

The discrepancy in the different graph convergence notions suggests that there is room for new definitions. Our goal in this chapter is to start developing a new theory of graph convergence based on the effective resistance. Our hope is that this definition will also help understanding the GIL phenomenon.

Definition 6 (Resistance based graph convergence)**.** Let $(G_n = (V_n, E_n, w_n))_{n\in\mathbb{N}}$ be a sequence of weighted graphs and for a given k, let $R_{n,k}$ be the resistance matrix of k points uniformly sampled from V_n. We say that (G_n) is *convergent* if $(R_{n,k})$ is weekly convergent to some R_k for every k as $n \to \infty$.

Definition 7. The limit object of a sequence (G_n) is a measurable function $f_R : [0,1]^4 \to \mathbb{R}^+$ such $f_R(a,b,c,d) = f_R(a,c,b,d)$, and if $\alpha, (\xi_i)_{1\le i\le k}$ and $(\lambda_{\{i,j\}})_{1\le i\le j\le k}$ are uniform independent random variables on $[0,1]$, then the matrix with (i,j) entry $f_R(\alpha, \xi_i, \xi_j, \lambda_{\{i,j\}})$ has the same distribution as R_k. The notation is $(G_n) \xrightarrow{r} f_R$, where the r stands for resistance based convergence.

The limit object of the resistance based convergence is very similar to the graphon, although this definition seams slightly more complicated at first. We address these differences and similarities and prove the existence of such limit objects later; let's see a few examples first to understand the definitions more clearly.

For any sequence (G_n) we always scale the edge weights such that the maximum entry of $(R_{n,k})$ is $O(1)$ but larger than 0 with high probability, so that the R_ks and limit objects are meaningful. After such rescaling, our definition gives results for the convergence of both the dense and sparse graphs. We give and example to both cases.

Example 1 (Complete graph)**.** Let G_n be the sequence of complete graphs of size n with edge weights $w_n(i, i+1) = \frac{n}{2}$. Then $G_n \xrightarrow{r} f_R(\cdot,\cdot,\cdot,\cdot) = 1$, the constant 1 function.

Proof. It is not difficult to check that $(nI_n - \mathbf{1}\mathbf{1}^T)^\dagger = \frac{1}{n}I_n - \frac{1}{n^2}\mathbf{1}\mathbf{1}^T$, which implies that the effective resistance between any two points of the complete graph is $2(\frac{1}{n} - \frac{1}{n^2}) - 2(-\frac{1}{n^2}) = \frac{2}{n}$. After rescaling with conductances $\frac{2}{n}$, we get the the resistance matrix of G_n is $\mathbf{1}\mathbf{1}^T - I_n$, which is what we needed. □

Example 2 (Path graph)**.** Let G_n be the sequence of paths of length n with edge weights $w_n(i, i+1) = n$. Then $G_n \xrightarrow{r} |\xi_i - \xi_j|$.

Proof. By Corollary 1 and the convenient choice of edge weights, $R_{n,k} = Abs(Z_{n,k}\mathbf{1}^T - \mathbf{1}Z_{n,k}^T)$, where $Z_{n,k} \in \mathbb{R}^k$ contains the k randomly selected points from the set $\{\frac{i}{n} \mid 1 \geq i \geq n\}$ and Abs is the absolute value function taken entrywise. This expression weekly converges to $R_k = Abs(Z_k\mathbf{1}^T - \mathbf{1}Z_k^T)$ where $Z_k \in \mathbb{R}^k$ is now sampled uniformly randomly from $[0, 1]$. Hence, G_n is convergent to the given object. □

In these examples, after establishing that the sequences are convergent, we proved the existence of the limit object by exhibiting it. However, it is not yet clear that every convergent sequence has a limit object. Proving this statement if in fact the main result of this chapter.

Theorem 3. *Every convergent sequence of graphs (G_n) in the sense of Definition 6 has a limit object as defined in Definition 7.*

Proof. The proof given here will be similar to one of the proofs for the existence of Lovász-Szegedy limit objects discussed in section 11.3.3. in [11]. Instead of considering $R_{n,k}$, consider $R_n \in \mathbb{R}^{\infty\times\infty}$ which is constructed by sampling countably infinity points from G_n instead of k. If (G_n) is convergent, the sequence (R_n) also weakly convergent to some infinite matrix R_∞. Moreover, R_∞ have the *symmetrially exachangability* property: the submatrix $(R_\infty)_{ij}$ for $1 \leq i, j \leq k$ has the same joint distribution as the submatrix $(R_\infty)_{\sigma(i)\sigma(j)}$ for for $1 \leq i, j \leq k$, since we the order of the graph vertices does not matter.

Such objects as R_∞ are called symmetrically or weakly exchangeable arrays, and there are many known results about them. Theorem 5.1 in [3] shows the existence of a measurable function that satisfies exactly the same conditions as our measurable function f_R in Definition 7. □

As stated several times, our resistance based graph convergence and especially our limit object bears resemblance to the dense graph limit and the graphon. This might not be obvious at first sight, since the graphon if a deterministic object, and all of our definitions made use of randomness so far. However, if we slightly rephrase the definition of the limit object of dense convergence, it could be a measurable function $f : [0, 1]^4 \to \{0, 1\}$, such that similarly to our definition, the submatrix generated by $f(\alpha, \xi_i, \xi_j, \lambda_{\{i,j\}})$ for $1 \leq i, j \leq k$ converges in distribution to the submatrix we get by sampling k vertices from the adjacency matrix. We even know what that f is

$$f = \begin{cases} 1 \text{ if } \lambda_{\{i,j\}} < W(\xi_i, \xi_j) \\ 0 \text{ otherwise} \end{cases} \tag{3.2}$$

Since the graphon W encodes exactly the subgraph densities the adjacency matrix of subgraph F will be generated exactly with the desired probability.

With this new point of view, it is clear the the graphon and the resistance based limit are tightly linked, however there are still notable differences:

- The obvious difference is that while the graphon ensures convergence of adjacency matrices whereas resistance based limit ensures convergence of resistance matrices.
- We already saw that as opposed to the dense graph limit, that the resistance based limit can gives non-zero limit object even if the graphs are sparse. In fact, we will see soon that the resistance based limit is most interesting in the sparse regime.
- For the graphon, it has been prooved by [8] that f is independent of α (the limit object is ergodic). We will see later that this is not the case for the resistance based limit object. On the other hand, we are yet to encounter a sequence of graphs where the resistance based limit object depends on $\lambda_{\{i,j\}}$.

3.3 Large degree graphs

While the resistance based graph convergence non-zero limit objects to both the dense and sparse regime, unfortunately, due to the GIL phenomenon, the definition is oblivious to global structure in the dense case. As shown in Corollary 2 many dense graphs, the effective resistance between two points (in the limit) is simply $\frac{1}{d_i} + \frac{1}{d_j}$. Thus, in some sense we can say that these graphs converge to the augmented star graph $R(\alpha, \xi_i, \xi_j, \lambda_{\{i,j\}}) = \frac{1}{d(\xi_i)} + \frac{1}{d(\xi_j)}$ with an auxiliary node in the middle that is connected to v_i with weight d_i.

One way to give more strength to our definition would be to subtract the limiting low-rank matrix and rescale the edge weights until we again get a $O(1)$, non-zero limit. This approach is similar to the idea of second order convergence in a Taylor series and has been already used in applications [13].

In this thesis, we do not explore this idea any further, instead we focus on characterising the graphs for which the limit object is meaningful without any modifications.

Chapter 4

Criticaility

It will be more convenient to give a definition for the graphs where the GIL phenomenon does occur, where the limit objects become trivial.

Definition 8. The the limit object of (G_n) be $R(\cdot,\cdot,\cdot,\cdot)$ and let's sample $\alpha, (\xi_i)_{1\leq i\leq 2k}$ and $(\lambda_{\{i,j\}})_{1\leq i\leq j\leq 2k}$. Let the (i,j) entry of the $k\times k$ the random matrix $\mathcal{R}_k$ be the value $R(\alpha,\xi_i,\xi_{j+k},\lambda_{\{i,j+k\}})$. We say that G is *trivial* if for every $\epsilon>0$ we have $P\left(\frac{rank(\mathcal{R})}{k}>\epsilon\right)\to 0$ as $k\to\infty$. We call a sequence of graphs trivial if they converge to a trivial object. A sequence of graphs is *critical* if it is not trivial.

The intuition is that row-rank $\mathcal{R}_k$ corresponds to trivial. Notice that since the (i,j) entry of $\mathcal{R}_k$ depends on ξs with indices i and $j+k$ we essentially measure the effective resistance between two disjoint set of points. In particular, the diagonal elements behave the same way as the off-diagonals unlike in R_k which allows $\mathcal{R}_k$ to be low-rank in many cases.

By Corollary 2, for large degree graphs the $\mathcal{R}_k$ matrices are all deterministic with rank 1 or 2, which immediately gives us our first result.

Theorem 4. *Limits of sequences of graphs with $\overline{d_{min}}=\omega(log(n))$ are trivial.*

Not only does this Definition 8 rule out the previously described $\frac{1}{d_i}+\frac{1}{d_j}$ type augmented star matrices and the limit objects of the grid graphs (see the proof later in the chapter), but also any combinations of them. For example one could consider a sequence of graphs G_n that consists of a union of complete graphs and grid graphs on $\frac{n}{2}$ nodes respectively, with appropriate edge weights (and connected with a single edge). Such objects should

still be characterised to be trivial, however, less robust definitions (e.g. trivial objects are defined to have local information based resistance: $R_{ij} = f(v_i) + f(v_j)$ for some f) would have missed these examples.

We showed several examples of trivial graphs sequences but we are yet to show critical ones ones. As one of the main results of this thesis, the next two sections will be about showing such examples.

4.1 Grid graphs

4.1.1 1D grid

Theorem 5. *The sequence of path graphs is critical.*

Proof. In the proof of Example 2 we showed that $R_k = Abs(Z_k \mathbf{1}^T - \mathbf{1} Z_k^T)$, where where the entries of $Z_k \in \mathbb{R}^k$ are sampled uniformly randomly from $[0, 1]$. We will drop the k subscript for better readability.

A similar argument shows that $\mathcal{R}_k = Abs(Z^{(1)}\mathbf{1}^T - \mathbf{1} Z^{(2)T})$, where $Z^{(1)}, Z^{(2)} \in \mathbb{R}^k$ are $2k$ independent uniform samples from $[0, 1]$. Let's assume $Z^{(1)}, Z^{(2)}$ are both in ascending order and let's also define a third vector $Z^{(3)} \in \mathbb{R}^{2k}$ that contains both the the values of $Z^{(1)}$ and $Z^{(2)}$, again in ascending order. Let $Z^{(4)} \in \{0, 1\}^{2k}$ have 1 at it's i^{th} entry if $Z_i^{(3)}$ is an element of $Z^{(1)}$, and 0 if $Z_i^{(3)}$ is an element of $Z^{(2)}$. A helpful schematic the illustrates these definitions on an example is included in Figure 4.1 (a).

Now consider the auxiliary matrix $C_k = \mathcal{R}_k + Z^{(1)}\mathbf{1}^T - \mathbf{1} Z^{(2)T}$. Since we changed $\mathcal{R}_k$ only by a rank 2 matrix, it will be enough to show $rank(C_k) = \Theta(n)$ to get the desired result. The block structure of of $sign(C_k)$ is shown on Figure 4.1 (b). Notice that the thick red line of $2k$ segments that separates non-zeros from the 0-block is exactly encoded in $Z^{(4)}$, as a 0 in $Z^{(4)}$ corresponds to a vertical segment and a 1 to a horizontal segment. This is because if for example we start with a zero, than $Z_1^{(2)}$ is smaller than all elements of $Z^{(1)}$, so the entire first row of $Z^{(1)}\mathbf{1}^T - \mathbf{1} Z^{(2)T}$ is positive which implies that the first row

of C_k is positive as well. This argument can be generalised to derive the correspondence between the red line and $Z^{(4)}$.

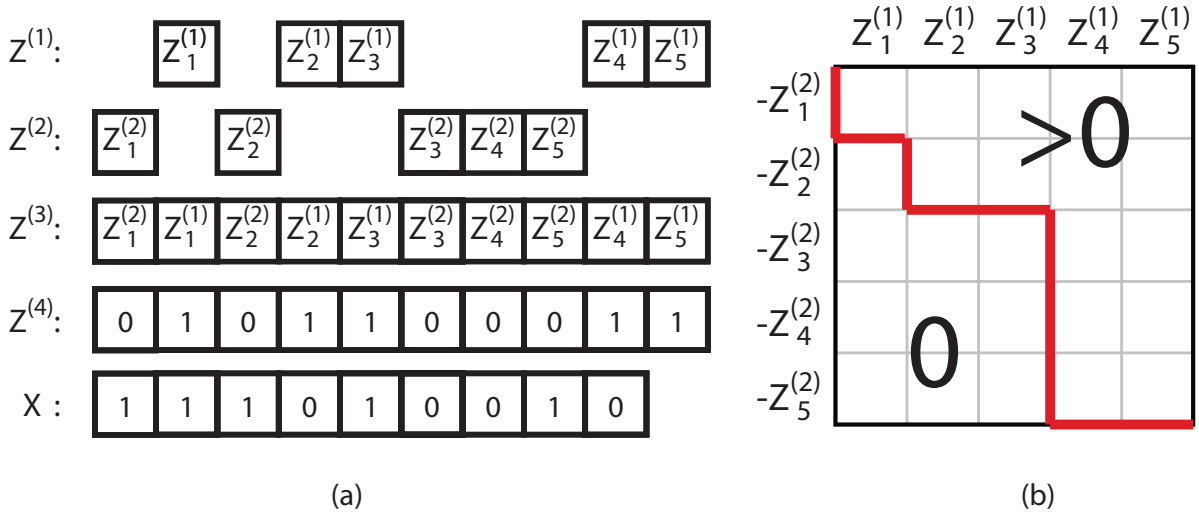

Figure 4.1:
(a): A helpful schematic for definitions the of $Z^{(1)}$, $Z^{(2)}$, $Z^{(3)}$, $Z^{(4)}$ and $\mathcal{X}$. The order of the interlacing of $Z^{(1)}$ and $Z^{(2)}$ is picked arbitrarily to show a particular example.
(b): The block structure of of $sign(C_k)$ corresponding to (a)

The rank of C_k can be lower bounded by the number of rows that start with different number of leading zeros (e.g. rows 1,2 and 5 on Figure 4.1 (b)). The number of such rows is exactly the same as the number of times we see a 01 subsequence in $Z^{(4)}$. Let $\mathcal{X}_i \in \mathbb{R}^{2k-1}$ be indicator vector of a 01 or 10 sequence starting at index i of $Z^{(4)}$. Since the number of 01 and 10 subsequences are the same it will be enough to show that $\sum_{i=i}^{2k-1} \mathcal{X}_i = \Theta(k)$ w.h.p. Unfortunately, the $\mathcal{X}_i$s are not independent since for example knowing the first $2k-2$ values completely determines $\mathcal{X}_{2k-1}$. Still, we can show that if we only know the first few values the dependence of the next $\mathcal{X}_i$ will not be too great. Let's make this more rigourous.

We can model $Z_i^{(4)}$ as a random permutation of k zeros and k ones. Given all previous $Z_j^{(4)}$s, $Z_i^{(4)}$ has the same distribution as $Ber\left(\frac{1}{2k-i}\left(k - \sum_{j=1}^{i-1} Z_j^{(4)}\right)\right)$, where $Ber\,(p)$ is a

Bernoulli random variable. This in particular implies that

$$P\left(\mathcal{X}_i = 1 \mid \mathcal{X}_j \text{ for all } j < i\right) > P\left(Ber\left(\frac{1}{4}\right) = 1\right) \qquad \text{if } i < \frac{k}{4} \tag{4.1}$$

since in that case $\frac{3}{4} > k - \sum_{j=1}^{i-1} Z_j^{(4)} > \frac{1}{4}$, so knowing the first $\frac{k}{4}$ still gives us enough freedom to choose the next element somewhat independently.

Then for any sufficiently small $\epsilon > 0$

$$P\left(\frac{1}{k}\sum_{i=i}^{k} \mathcal{X}_i > \epsilon\right) \geq P\left(\frac{1}{k}\sum_{i=i}^{\frac{k}{4}} \mathcal{X}_i > \epsilon\right) \geq P\left(\frac{4}{k}\sum_{i=1}^{\frac{k}{4}} Ber\left(\frac{1}{4}\right) > 4\epsilon\right) \to 1 \tag{4.2}$$

by the law of large numbers. Equation (4.2) implies that $P\left(\frac{rank(\mathcal{R})}{k} > \epsilon\right) \to 1$, so sequence of path graphs is critical.

□

4.1.2 2D grid

Our physical intuition tells us that the 2D grid should be trivial. Electric resistance on a metal plate does not depend on the two points chosen. The physics theory even provides a formula to compute the effective resistance on an Ohmic material: $R = \frac{\rho l}{A}$, where ρ is the resistivity, l is the length of the conductor and A is the cross section. We could apply this formula to compute the resistance between a circle of radius $\frac{1}{n}$ and a circle of radius 1

$$\int_{\frac{1}{n}}^{1} \frac{\rho}{2r\pi} dr = \frac{\rho}{2\pi} log(n) + o(\log(n)) \tag{4.3}$$

Intuitively, ρ corresponds to the inverse of the edge weights $\frac{1}{w_{ij}}$, and we should choose $\rho = \frac{1}{\log(n)}$ for proper scaling. Then, as we take $n \to \infty$ the expression above converges to around $\frac{1}{2\pi}$. The two-point resistance in the limit should be two times the expression above (because there are two points), so it should be around $\frac{1}{\pi}$. In the rest of this section we provide a rigorous statement and proof.

Theorem 6. *If (G_n) is the sequence of $n \times n$ 2D grid graphs with edge weights $w_n = \log(n)$, (G_n) is convergent and trivial. The matrix $\mathcal{R}_k = \frac{1}{\pi}\mathbf{1}\mathbf{1}^T$ is deterministic, rank 1 for every k.*

Proof. It is a known result by [7], that for an infinite grid graph with unit edge weights, the asymptotic form of the resistance between points $v_1 = (0, 0)$ and $v_2 = (v_{2x}, v_{2y})$ is

$$R_\infty(v_1, v_2) = \frac{1}{\pi} log\left(\sqrt{v_{2x}^2 + v_{2y}^2}\right) + o(1), \tag{4.4}$$

This formula already gives a lower bound on the entries of $\mathcal{R}_k$. By Rayleigh's Monotonicity Principle (page 35 of [15]), adding edges only decreases the effective resistances

$$R_n(v_1, v_2) \geq \frac{1}{\log(n)} R_\infty(v_1, v_2) = \frac{log\left(\sqrt{(v_{1x} - v_{2x})^2 + (v_{1y} - v_{2y})^2}\right)}{\pi \log(n)} + O(1) \to \frac{1}{\pi} \tag{4.5}$$

with high probability as $n \to \infty$, because if we sample k points from an $n \times n$ grid, their distance will be $\Theta(n)$ with high probability.

For the upper bound we need to work slightly harder. We will construct a flow on the $n \times n$ grid (with unit edge weights) that has energy $\frac{1}{\pi} log(n) + o(log(n))$, which by Thompson's principle will give us the desired upper bound.

The flow will be build up of three parts; a flow from v_1 a to circle of radius r around v_1 (denoted by $C(v_1, r)$), a flow between $C(v_1, r)$ and $C(v_2, r)$ and finally a flow from $B(v_2, r)$ to v_2. The circles should be understood as a discretised versions of the Euclidian circles; the boundary points of the set of grid points less than or exactly r, but more than $r - 1$ distance away from v_1 using the Euclidian distance.

Since the each of k points will also be $\Theta(n)$ distance away from the boundary of G_n we suppose that $r = \Theta(n)$ and the circles are still fully contained in the grid, non-intersecting. The intuition is that the flow between the two circles will have $o(log(n))$ energy because the flow is spread out in a very large space, compared to the flows concentrated around v_1 and v_2. This resembles the main idea behind GIL: the effective resistance only depends on

what happens in a neighbourhood of the points, although in our case we have to consider a larger neighbourhood then just the degree.

First let's construct the flow from v_1 to $C(v_1, r)$. The construction is not easy to be made explicit. Let's consider the electric flow that we get if we apply potential 0 to v_1 and as much potential to each node of $C(v_1, r)$ as needed for the flow to have magnitude 1. We have two claims about this flow: 1) That it brings approximately equal $\Theta(\frac{1}{r})$ current to each node of $C(v_1, r)$, and 2) the energy of this flow is $\frac{1}{2\pi}log(n) + o(log(n))$.

The first claim follows from symmetry; both the potential differences and the effective resistances (as seem in 4.4) are constant across $C(v_1, r)$, so the incoming current flow must be as well. An alternative argument would rely on the random walk interpretation of the current flow and the fact that random walk on a 2D grid converges to a normal distribution which is spherically symmetric.

For the second claim, we use the interpretation outlined in Lemma 1 and the result 4.4. The energy of the flow can be interpreted as $R_\infty(v_1, C(v_1, r))$, since the flow we are considering is oblivious to what is happening outside the circle.

Let l be the (discretised) perpendicular bisector of v_1 and v_2. We may assume that such a bisector exists; if v_1 and v_2 are odd distance away we may move v_2 one edge further without changing any of the resistances by $\omega(log(n))$. Since, the event of a random walk on the infinite grid hitting l is always preceded by the random walk hitting $C(v_1, r)$:

$$R_\infty(v_1, C(v_1, r)) = \frac{1}{4P(v_1 \to C(v_1, r))} < \frac{1}{4P(v_1 \to l)} = R_\infty(v_1, l)$$

Now consider the electrical flow on the infinite grid between v_1 and v_2. The line l splits this flow into two, and if we flip edges on the v_2 side of the plane we get two unit flows from a point to l. At least one of these flows must have energy $\frac{1}{2}R_\infty(v_1, v_2)$. Hence,

$$R_\infty(v_1, l) < \frac{1}{2}R_\infty(v_1, v_2) = \frac{1}{2\pi}log(n) + o(log(n))$$

Now we just need to show the construction of the flow between $B(v_1, r)$ and $B(v_2, r)$. We can be somewhat generous in our construction and our analysis since we only need

an asymptotic upper bound. For simplicity let's assume that v_1 and v_2 has the same y-coordinate.

The outline of the construction is on Figure 4.2. We first split the circles into four equal parts: (a), (b), (c) and (d) as shown. Each grid point $c_{1,a} \in C(v_1, r)$ in part (a) is is connected with a straight horizontal line to its pair in $c_{2,a} \in C(v_2, r)$, and a constant flow is sent across this path, which has value $\Theta(\frac{1}{r})$, the same as the incoming current to c_1 from our point-to-circle flow construction. Note that multiple points $c_{1,a}$ may have the same y-coordinate, in which case these flows collide and add up. Still, since we are in region (a), this can happen with only a constant number of grid points, so the value of the flow on each green edge is at most $\Theta(\frac{1}{r})$. Similar arguments can be said about regions (b), (c) and (d) although in these cases the flows start out in different direction to assure that there are only a constant number of collisions.

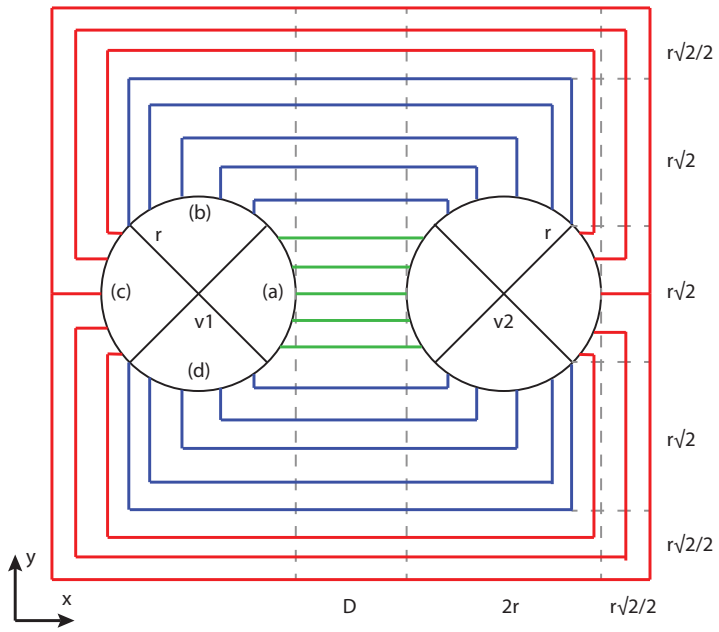

Figure 4.2: Flow construction with energy $\Theta(1)$ between two circles of radius r. Red, blue and yellow lines correspond to the flow lines in different sections. The dashed grey lines help count the number of flow lines in each section.

Now we have to upper bound the energy of the flow. The number of edges in this construction is upper bounded by $(D+4r+r\sqrt{2})(4r\sqrt{2}) = \Theta(n^2)$, since $D = dist(v_1, v_2) - 2r = \Theta(n)$ and $r = \Theta(n)$ by assumption. On the other RI^2 term for each edge is asymptotically $\Theta(\frac{1}{n^2})$ (since we set $R = 1$), hence the energy of the flow is $\Theta(1)$. Also since the number of rows are columns used are both $\Theta(r)$, the flow will fit inside the $n \times n$ grid.

Therefore, once we join the flow tree constructed flows, we get that the overall energy is $\frac{1}{\pi} log(n) + o(log(n))$ which is what we needed for the upper bound.

□

This is an important result, since this is the first example when the limiting effective resistance is trivial but does not depend on $\frac{1}{d_i}$. To be fair, the $\frac{1}{\pi}$ can still be thought of the the local neighbourhood of a point in a plane.

4.1.3 3D grid

We can essentially use the same technique to show that a 3D grid is trivial, only the flow construction is slightly more involved. For a 3D grid, the limiting resistance is much more complicated (see [7], Eq. (37))

4.2 Tree graphs

4.2.1 Binary tree

The binary tree is an interesting example of trivial graphs. Here, the triviality does not come from many disjoint paths which allow the random walker to get lost in space. Instead, all nodes behave very similarly, hence the trivial limit.

Theorem 7. *A sequence (G_n) of binary trees of depth n are convergent with edge weights $w_n = n$. The limiting object is trivial.*

Proof. By Corollary 1, the effective resistance between two points only depends on their distance. For a binary tree the distance is easily calculated as

$$(R_{n,k})_{ij} = l(v_i) + l(v_j) - 2l(p(v_i, v_j)) \tag{4.6}$$

where $l(v) : V_n \to \mathbb{N}$ is the level of a vertex on the tree and $p(v_1, v_2)$ is the first common parent of v_1, v_2.

Since the tree is growing exponentially, most of the mass is concentrated on the leaves. If we sample v_i and v_j uniformly randomly, they will most likely be on one of the lowest levels. For any $0 < C < 1$

$$P\left(l(v_i) < Cn\right) = \frac{2^{Cn} - 1}{2^n - 1} \to 0 \qquad \text{as } n \to \infty \tag{4.7}$$

We can also show that first their common parent will be around the root. This is because to have $p(v_i, v_j) \geq k$, even if we know that $l(v_i) \geq k$, we need that v_j to be chosen to be in the same subtree starting from the k^{th} level as v_i. There are 2^{k-1} disjoint subtrees starting from level k, so the probability of this event is low for large 0.

$$P\left(p(v_i, v_j) > Cn\right) \leq P\left(v_j \text{ and } v_j \text{ in same Cn-subtree} \mid l(v_i) \geq Cn\right) \leq \frac{1}{2^{Cn-1}} \to 0 \tag{4.8}$$

as $n \to \infty$. We have this shown that almost surely $\frac{l(v_i)}{n} = 1$ and $\frac{p(v_i,v_j)}{n} = 0$. This, together with Equation (4.6), shows that after rescaling, $(R_{n,k})_{i,j} = 2$ if $i \neq j$ almost surely, which proves the theorem. □

4.2.2 Non-critical Galton-Watson tree

We are interested in the limit of a sequence of graphs (G_n), that are are generated by a Galton-Watson process. We would like the size of G_ns to increase as n increases, we need need to condition on the size of the random tree. The first natural idea is to condition on $|G_n| = n$, however, as shown in Lemma 1.2.5. in [1], such trees actually behave like

subcritical trees. Instead, we condition on the event of non-extinction before level n and we set G_n to be the first n levels (and ignore the rest). We strongly suspect that a similar proof as for the case of the deterministic binary tree would work here since on these conditioned supercritical trees we know that the lowest levels grow exponentially w.h.p. and on any level there is a node which has at least two child branches that never become extinct with non-zero probability. However, no rigorous proof is shown here.

4.2.3 Critical Galton-Watson tree

Again, we need to condition on something to assure that G_n is increasing in size. For critical Galton-Watson tree it is standard to condition on $|G_n| = n$ (following the notation of [2], we call these CRT(n)). In fact, there are many known results known about the convergence of such graphs [1]. Convergence in a *local* sense looks at the trees at a finite size, it is oblivious to events that happen infinitely far from the root. The limit object of CRT(n) in the local sense is a so called Kesten tree, which is essentially an infinite path graph with long, but finite branches. This result already hints that similarly to the path graph, the critical Galton-Watson tree should also be critical.

For the proof, we need the other convergence notion, *scaling* convergence. In the scaling limit, the edge lengths are rescaled such that all vertices remain finite distance away from the root. Since distances are effective resistances are identical in a tree graph, scaling convergence agrees with our definition.

Theorem 8. *A sequence of graphs (G_n), where $G_n = CRT(n)$, with edge weights $w_n = \frac{1}{\sqrt{n}}$. The limiting object is critical.*

Proof. It is a previous result by [2] that CRT(n)s converge with $\frac{1}{\sqrt{n}}$ edge weights to the Continuum Random Tree (CRT). Since we only consider tree graphs their definition of convergence agrees with resistance based convergence. The limit object is $f_R = dist_\alpha(\xi_i, \xi_j)$, where $dist_\alpha(\cdot, \cdot)$ is the distance function on the CRT with parameter by α (this is our example where the first argument of f_R plays an important role).

We also know from [2] that if we sample k points from a CRT, the edge lengths of the resulting *k-subtree* is given by the density function

$$f = s \cdot exp(\frac{-s^2}{2}) dl_1 \dots dl_{2k-3} \qquad \text{where } s = \sum_i l_i \tag{4.9}$$

and topology of the tree is uniformly random across all binary trees. Another way to describe the random k-subtree of the CRT is by the line-breaking process: consider a non-homogenous Poisson process on $[0, \infty)$ with rate $\lambda(x) = x$, separate out the segments between the points, start with the first segment, and in the i^{th} step join one end of i^{th} segment to a uniform random point of the subtree built by the previous segments. After k steps this tree has the same distribution as the k-subtree. To show (G_n) is critical we need to show that if we sample $2k$ points - half of them colored red and half colored blue randomly - then the distance matrix between red and blue points is high rank, more precisely for some ϵ small enough $P\left(\frac{rank(\mathcal{R})}{k} > \epsilon\right) \to 1$.

As in the proof of the criticality of the path graph we will make a few rank-one modifications to $\mathcal{R}$ to get a high-rank matrix. We will call these rank one modifications *pruning*. The modification will always involve deleting all leaves of the tree and moving the red and blue points to the parent of the leaf. Let's label red points by r_i and blue points by b_i so we can keep track of them, and as always let $\mathcal{R}_{i,j}$ contains the distance between r_i and b_i. First, we will prune leaves that are colored red. Notice, that the distance from r_i to its parent (denoted by $d_p(r_i)$) is added to each entry of row i of $\mathcal{R}$. Therefore we can subtract the rank one matrix $d_p(r)\mathbf{1}^T$ from $\mathcal{R}$, and now $\mathcal{R}^{(1)}_{i,j}$ contains distances between $r^{(1)}_i$ and b_i, where $r^{(1)}_i$ is now the parent of r_i. Visually we can imagine this as moving the red dots up by one and deleting the edges which became irrelevant. After pruning red leaves, we can do the same with the blue ones. This time the matrix we subtract is $\mathbf{1}d_p(b)^T$, since we are now subtracting constant vectors from the columns. Again we can do this as in the line-breaking process it was assured that all colored nodes are leaves, and the leave edge s contain low-rank information. The pruning of red and blue edges together are shown on an example on Figure 4.3 (a) $\to$ (b).

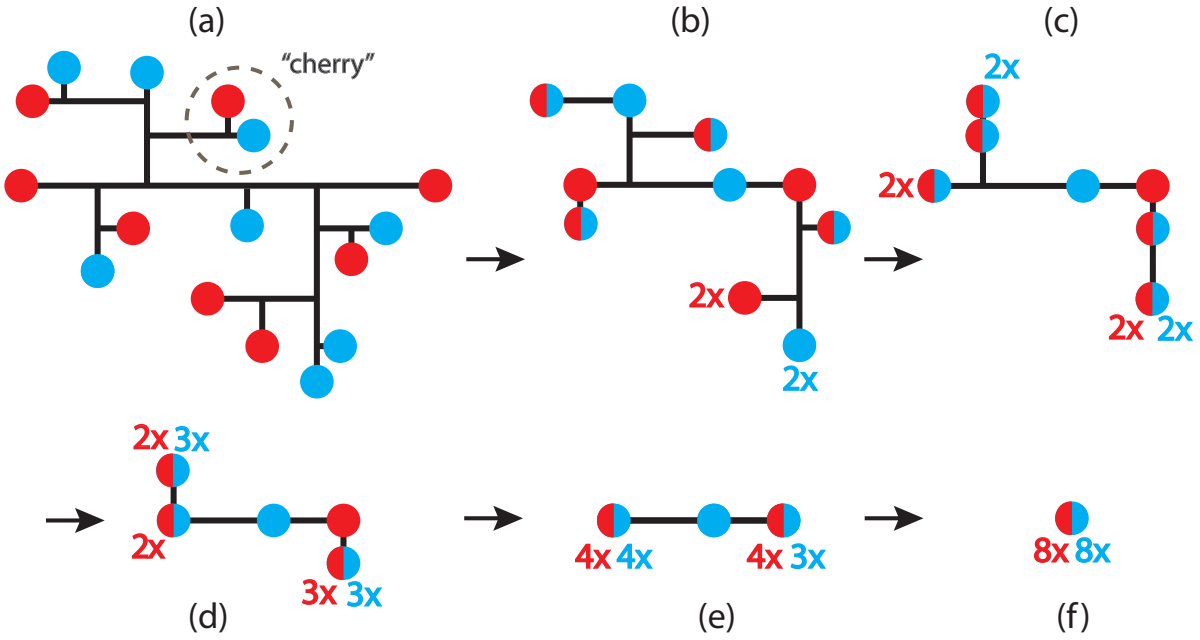

Figure 4.3: The pruning of a 16-subtree in 5 steps starting from (a), finishing at (f). The colored numbers next to the balls indicate how many of that color has been merged to that particular ball. For example the lone ball in (f) colored both red and blue indicates that all 8 red and all 8 blue balls has been merged into a single point.

Once we are done with our first round of pruning, it is possible that two blue, two red or a blue and a red point at the same node, they collide. This last case, the collision of differently colored balls is the key to this proof. It prevents us from doing an other round of pruning of either colors, since moving up a for instance a blue ball to its parent would decrease its distance from most of the red balls, but would increase it from the red ball it collided with previously. The indicator vector that would tells us which distances increase and which decrease differs for every b_i, which might turn out complicated.

The key idea is to continue with the pruning regardless and correct for the problem caused by the collisions. Let $CB_j^{(i)} = \{k \mid r_k^{(i)} \text{ collided with } b_j^{(i)}\}$ and $CR_j^{(i)}$ and the same for the collisions with $r_j^{(i)}$. Clearly these sets (or equivalence classes) are disjoint and monotone increasing in i for every j; balls only collide, don't separate. Initially these blocks

are small. We already saw that $|CB_j^{(0)}| = |CR_j^{(0)}| = 0$ and $|CB_j^{(1)}|, |CR_j^{(1)}| \leq 1$. By the end the whole tree will be pruned off, and we get $|CB_j^{(t)}| = |CB_j^{(t)}| = k$, where t is the number of rounds of pruning we need to take (see Figure 4.3 (f)). Finally, let $cb_j^{(i)}$ be a canonical (e.g. lowest) element of $CB_j^{(i)}$. Now we have enough notation to start the pruning.

We will prune the two colors at the same time, thus we will make rank-two subtractions of $\mathbf{1}d_p(b^{(i)})^T + d_p(r^{(i)})\mathbf{1}^T$ in the i^{th} step, and we set $r^{(i+1)}$ and $b^{(i+1)}$ be the parents of $r^{(i)}$ and $b^{(i)}$. The term we need to correct by to make the subtractions legitimate is a matrix supported only on the blocks defined by the sets of indices $(CB_j^{(i)}, CB^{(i)}_{cb_j^{(i)}})$. This is because these are precisely the terms which are already 0 in $\mathcal{R}^{(i)}$, and while all other terms should be decrease with the weight of the pruned edge, these terms should still stay 0. Denote this block matrix that we need to add back to the sum by $D^{(i)}$. Note that these $D^{(i)}$ are positive semidefinite and their rank is exactly the number of district sets defined by $CB_j^{(i)}$ (or in other words the number of balls where there is a collision of different colors).

By the end of the pruning process we are left with a single point and $\mathcal{R}^{(t)} = 0$. This means that we proved the equation

$$\mathcal{R} = \sum_{i=1}^{t} \left(\mathbf{1}d_p(b^{(i)})^T + d_p(r^{(i)})\mathbf{1}^T + D^{(i)}\right) \tag{4.10}$$

Since the diameter of the tree decreases by 2 at each step, we can bound t by the depth of the tree. It has been shown in [9], that the expected value of the depth of a uniformly random binary tree is $\Theta(\sqrt{n})$ and by Markov's inequality for any $c > 0$

$$P\left(\text{depth of unif. rand. bin. tree} > cn\right) \leq \frac{\Theta(\sqrt{n})}{cn} \to 0 \qquad \text{as } n \to \infty \tag{4.11}$$

Therefore with high probability, $t < cn$ so to show $\mathcal{R}$ is high rank it is enough to show that $\sum D^{(i)}$ is high rank. As noted before each $D^{(i)}$ is positive semidefinite, so for us it will be enough if $D^{(1)}$ is high rank, the addition of the other $D^{(i)}$s can only increase this value.

The rank $D^{(1)}$ is exactly the number of nodes with a red-blue collision after the first step. The preimage of such a collision is a so called *"cherry"*, a red and a blue ball with

common parent (see Figure 4.3 (a)).

Lemma 3. The number of cherries in a 2k-subtree with randomly chosen k red and k blue leaves is $\Theta(k)$.

Proof. It will be enough to show that the number of colorless cherries (leaves with common parent) is $\Theta(k)$, because the coloring is independent of the construction of the tree, and the probability for a colorless cherry to become differently colored is $\frac{1}{2}$.

We will show that when we add the next branch in the i^{th} step of the line-breaking process, we either create a new cherry with probability $\frac{1}{4}$. By equation (4.9), if we condition on the length of the whole tree, the edge lengths all have the same distribution. Since there i leaf edges and $i-3$ internal edges, this implies that we have more than $\frac{1}{2}$ probability of adding the next branch to a leaf edge. When we add the next branch to an internal edge, the number of cherries stay the same. When we add the next branch to a leaf edge, we always create a new cherry, however, we might also destroy an existing one. The chance of us creating a new cherry is thus

$$\frac{1}{2}\left(1-\frac{2\cdot \#\text{ of cherries}}{\#\text{ of leaves}}\right) \tag{4.12}$$

Solving for the stable point of the equation $\frac{1}{2}(1-2p)=p$ we get that the probability of creating a cherry is $\frac{1}{4}$.

□

With the proof of the Lemma 3, the proof of the theorem will be complete. □

Chapter 5

Numerical experiments

5.1 Measuring criticality

In this section we propose a method to measure how close the graphs are to criticality. Criticality depends on the rank of $\mathcal{R}_k$. Of course we cannot expect to have the exact ranks of finite size trivial graphs to be low, our results only hold in the limit. Instead, measurements we measure how close $\mathcal{R}_{n,k}$ is to a low-rank matrix. We compute all singular values of $R_{n,k}$, and measure how fast they decay.

Definition 9. Let the *approximate rank* of the matrix $R_{n,k}$ be given by:

$$Arank(\mathcal{R}_k) = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{k} \sigma_i} \tag{5.1}$$

The *Arank* function is in the range $[0, 1]$, and the closer the value is to 1 the closer the graph is to being critical. We computed the the *Arank* function for several graph models and the results are shown on Figure 5.1. The results agree with the theoretical arguments; for most graph models *Arank* is monotone increasing except for the path graph and the critical Galton-Watson tree where the function stays constant.

We do not have theoretical arguments for percolation, but Figure 5.1 suggests that the situation there is similar to the case of Galton-Watson trees; only the critical percolation seems to be critical.

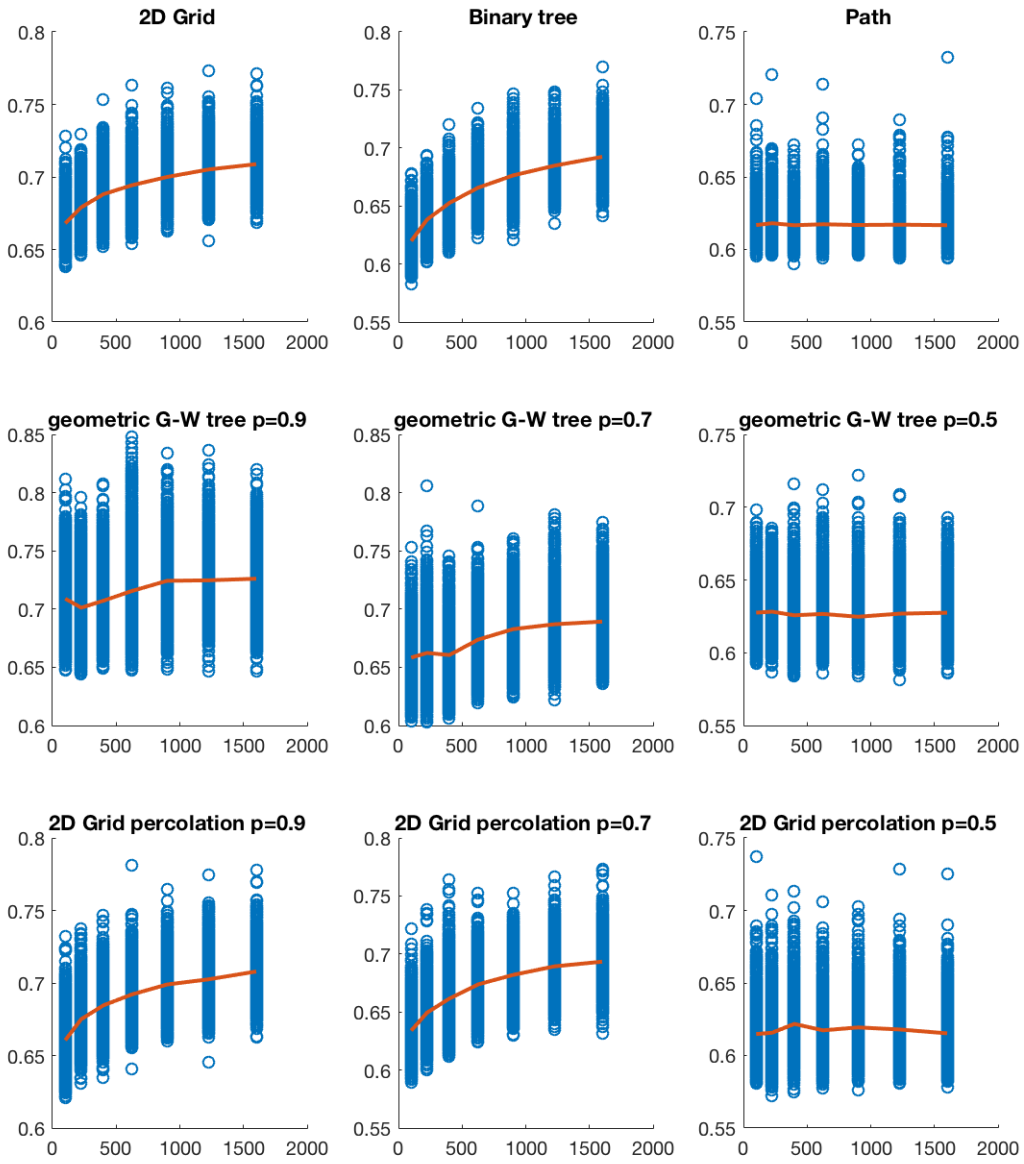

Figure 5.1: The graphs show that value of the $Arank(\mathcal{R}_k)$ function against the size graph sizes for different graph models (2D Grid graphs, Path graphs, Binary trees, Galton-Watson trees with geometric offspring distribution and Percolation on the 2D grid). For all graphs k=30. $\mathcal{R}_k$ was sampled 1000 times except for the G-W trees and the Percolation models where $\mathcal{R}_k$ was sampled 100 times for each of the 20 random graph samples. The blue dots show the $Arank$ values for each sample and the red lines show the average value.

5.2 Future applications

One of the main messages of this thesis is that while the GIL phenomenon seams to hold for most graph models, there are examples where the information carried by effective resistances is non-trivial. Moreover, for the path graph and the critical Galton-Watson tree we saw that effective resistances preserve the geometry of the path even in the limit. This doesn't necessary imply that for all critical graphs resistant distances will be meaningful, but our results serve as motivation continue working on this question. First, it would be interesting to expand the set of known critical graphs. Random graphs models around the critical point, e.g. critical percolation and real-world data around the percolation threshold are good candidates.

As described in the Introduction, the effect of GIL is not completely clear on real world applications as there are conflicting reports [13, 16, 10]. After a through review on this matter, our work in this thesis could help identify the conditions under which resistance based methods can still be useful.

References

[1] Romain Abraham and Jean-François Delmas. An introduction to galton-watson trees and their local limits. *arXiv preprint arXiv:1506.05571*, 2015.

[2] David Aldous. The continuum random tree ii: an overview. 1991.

[3] David J Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.

[4] Itai Benjamini, Oded Schramm, et al. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6, 2001.

[5] Christian Borgs, Jennifer Chayes, László Lovász, Vera T Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph limits and parameter testing. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 261–270. ACM, 2006.

[6] Fan Chung and Mary Radcliffe. On the spectra of general random graphs. *the electronic journal of combinatorics*, 18(1):P215, 2011.

[7] József Cserti. Application of the lattice green?s function for calculating the resistance of an infinite network of resistors. *American Journal of Physics*, 68(10):896–906, 2000.

[8] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *arXiv preprint arXiv:0712.2749*, 2007.

[9] Philippe Flajolet and Andrew Odlyzko. The average height of binary trees and other simple trees. *Journal of Computer and System Sciences*, 25(2):171–213, 1982.

[10] Nguyen Lu Dang Khoa and Sanjay Chawla. Large Scale Spectral Clustering Using Resistance Distance and Spielman-Teng Solvers. In Jean-Gabriel Ganascia, Philippe

Lenca, and Jean-Marc Petit, editors, *Discovery Science*, Lecture Notes in Computer Science, pages 7–21. Springer Berlin Heidelberg, October 2012. DOI: 10.1007/978-3-642-33492-4_4.

[11] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Society Providence, 2012.

[12] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.

[13] Ulrike von Luxburg, Agnes Radl, and Matthias Hein. Getting lost in space: Large sample analysis of the commute distance. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, pages 2622–2630. Curran Associates Inc., 2010.

[14] Ulrike von Luxburg, Agnes Radl, and Matthias Hein. Hitting and commute times in large random neighborhood graphs. *Journal of Machine Learning Research*, 15(1):1751–1798, 2014.

[15] Russell Lyons and Yuval Peres. *Probability on Trees and Networks*. Cambridge University Press, New York, 2016. Available at `http://pages.iu.edu/~rdlyons/`.

[16] Canh Hao Nguyen and Hiroshi Mamitsuka. New Resistance Distances with Global Information on Large Graphs. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 639–647, 2016.

[17] D. Spielman and N. Srivastava. Graph Sparsification by Effective Resistances. *SIAM J. Comput.*, 40(6):1913–1926, January 2011.

[18] Daniel A Spielman. Algorithms, graph theory, and linear equations in laplacian matrices.