The Laplacian of the Erdős-Rényi graph

Rivier, Renaud

This publication URL: https://archive-ouverte.unige.ch/unige:170196
Publication DOI: 10.13097/archive-ouverte/unige:170196

UNIVERSITÉ DE GENÈVE
Section de mathématiques

FACULTÉ DES SCIENCES
Dr. Antti KNOWLES

The Laplacian of the Erdős-Rényi graph

THÈSE

Présentée à la Faculté des Sciences de l'Université de Genève pour obtenir le grade de Docteur ès Sciences, mention Mathématiques

par

Renaud RIVIER

de
Lausanne (VD)

Thèse N° - 5736 -

GENÈVE, printemps 2023

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION MATHÉMATIQUES

Thèse de Monsieur Renaud RIVIER

intitulée :

«The Laplacian of the Erdős-Rényi Graph»

La Faculté des sciences, sur le préavis de Monsieur A. KNOWLES, professeur associé et directeur de thèse (Section de mathématiques), Monsieur Y. VELENIK, professeur ordinaire (Section de mathématiques), Monsieur K. TORBEN, professeur (chair of stochastics at FAU Erlangen-Nürnberg, Germany), et Monsieur A. JOYE, professeur (Université Grenoble Alpes, Saint-Martin d'Hères, France), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 20 juin 2023

Thèse - 5736 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

À ma mère et à mes soeurs

Renaud

Résumé

Cette thèse a pour sujet les propriétés spectrales de la matrice Laplacienne de graphes aléatoires. Notre travail se concentre plus particulièrement sur le graphe Erdős-Rényi, qui est le modèle le plus simple d'un graphe aléatoire. Le graphe Erdős-Rényi, noté G, est caractérisé par son nombre de sommets $N \in \mathbb{N}^*$ ainsi que son régime $0 < d < N$ qui controle le degré moyen de chaque sommet: chaque arête du graphe complet sur N sommets est ouverte indépendemment avec probabilité d/N. Les propriétés du graphe G varient en fonction de la valeur de d. Ces différents comportements probabilistes ont été étudiés en profondeur depuis le papier originel de P. Erdős et A. Rényi. La matrice d'adjacence A et la matrice laplacienne L caractérisent complètement le graphe G et leur spectre peut être utilisé pour obtenir des informations importantes sur le graphe. Ce sont donc des objets mathématiques naturels dont l'étude revêt un intérêt à la fois pratique et théorique.

La matrice laplacienne a une place centrale en théorie spectrale des graphes. L'énergie d'un vecteur selon l'opérateur L est communément appelée énergie de Dirichlet et sert à obtenir des bornes de concentration de mesure. Le trou spectral de L est directement relié au phénomène de *mixing* de la marche aléatoire sur le graphe G et à son temps de relaxation. De manière plus générale, L est l'hamiltonien quantique d'une particule qui se déplace aléatoirement sur G et la géométrie de ses vecteurs propres possède une interprétation physique concrète reliée aux notions d'isolant et de conducteur éléctrique. D'un point de vue de la recherche en matrices aléatoires, L est un exemple simple de matrice de Wigner déformée creuse. Une question centrale pour de tels modèles et de savoir jusqu'à quel degré de densité le spectre des matrices gardent les mêmes comportements asymptotiques que les matrices de Wigner déformées, et à quel moment le spectre s'altère de manière notoire (apparition de valeurs propres extrêmes, changements des statistiques des valeurs propres ou de la dimension des vecteurs propres).

Dans une première partie du travail, une loi locale sur la matrice de Green du laplacien est prouvée jusqu'au régime sous-critique $d \geq C\sqrt{\log N}$, $C > 1$, repoussant de manière significative les limites des résultats existants dans la littérature qui n'étaient valables que jusqu'à $d \geq N^{\varepsilon}$, $\varepsilon > 0$. La loi locale permet entre autre d'assurer la délocalisation totale des vecteurs propres dans le centre du spectre. Nous expliquons pourquoi ces résultats sont optimaux et ne peuvent pas être étendus hors du centre du spectre et à des régimes de d inférieurs. Les techniques et les conclusions présentées dans ce chapitre s'inscrivent dans le prolongements des lois locales, un sujet qui a reçu un attention soutenue ces quinze dernières années.

Dans un seconde partie du travail, nous étudions les statistiques des valeurs propres extrêmes et la forme des vecteurs propres associés. Nous montrons que, pour $1 \ll d \ll N^{1/3}$, le processus ponctuel générés par les plus grandes valeurs propres de L est asymptotiquement proche d'un processus de Poisson dont nous donnons la fonction de densité. En particulier, pour $\frac{1}{2}\log N \ll d \ll N^{1/3}$, nous pouvons décrire la distribution du trou spectral de L comme une fonction du voisinages des sommets de petits degrés. Finalement nous montrons que pour certains régimes sous-critiques le trou spectral de L est donné par une fonction implicite calculée sur des arbres de tailles finies.

Dans notre travail, nous montrons la co-existence à l'intérieur du spectre de L d'états délocalisés et localisés, en prolongement de la vaste littérature sur ce sujet et la fameuse conjecture autour du modèle d'Anderson. Nous pensons que cette thèse fournit une solide base théorique à la compréhension des algorithmes spectraux et à la théorie spectrale des graphes. Nous proposons également des possibles extensions et des nouveaux chemins de recherche.

Abstract

This thesis focuses on the spectral properties of the Laplacian matrix of random graphs. The specific emphasis is on the Erdős-Rényi graph, which is one of the simplest model of a random graph. The Erdős-Rényi graph, denoted by G, is characterized by the number of vertices $N \in \mathbb{N}^*$ and the density parameter $0 < d < N$, which controls the average degree of each vertex: each edge of the complete graph on N vertices is independently open with probability d/N. The probabilistic behaviors of G vary depending on the value of the parameter d and have been extensively studied since the seminal paper by P. Erdős and A. Rényi. The spectral properties of the two canonical matrices associated with G, namely the adjacency matrix A and the Laplacian matrix L, completely characterize G and their spectra can provide important information about the graph. Hence, they are natural mathematical objects with both practical and theoretical significance.

The Laplacian matrix holds a central position in the spectral theory of graphs. The quadratic form defined by L is commonly referred to as the Dirichlet energy and is used to obtain concentration of measure bounds. The spectral gap of L is directly related to the mixing phenomenon of the random walk on the graph G and to its relaxation time. More generally, L is the quantum Hamiltonian of a particle that moves randomly on G, and the geometric properties of its eigenvectors have a concrete physical interpretation related to the notions of insulator and electrical conductor. From the perspective of random matrices, L serves as a natural example of a sparse deformed Wigner matrix. A central question for such ensembles is to determine to what extent the spectrum of matrices sampled for these models maintains the same asymptotic behaviors as other deformed Wigner matrices. In particular, the spectrum might undergo notable changes (such as the appearance of extreme eigenvalues, changes in statistics, and eigenvector dimensions) and the mechanism underlying such transitions are of particular interest.

In the first part of this work, a local law for the Green matrix of the Laplacian is proved up to the subcritical regime $d \geq C\sqrt{\log N}$, where $C > 1$. This significantly extends results previously available in the literature, which were only valid up to $d \geq N^{\varepsilon}$, for $\varepsilon > 0$. The local law ensures, among other things, the complete delocalization of the eigenvectors in the center of the spectrum. We explain why these results are optimal and cannot be extended beyond the center of the spectrum nor to lower regimes of d. The techniques and conclusions presented in this chapter are in line with the extension of local laws for matrix ensembles, a topic that has received significant attention in the past fifteen years.

In the second part of this work, we study the statistics of extreme eigenvalues and the shape of the corresponding eigenvectors. We show that, for $1 \ll d \ll N^{1/3}$, the point process generated by the largest eigenvalues of L is asymptotically close to a Poisson process for which we provide the explicit intensity measure. Particularly, for $\frac{1}{2}\log N \ll d \ll N^{1/3}$, we can describe the distribution of the spectral gap of L as a function of the neighborhoods of low-degree vertices. Finally, we demonstrate that for certain subcritical regimes, the spectral gap of L is determined by an implicit function calculated on trees of finite sizes.

In our work, we demonstrate the coexistence of delocalized and localized states in the spectrum of L, extending the vast literature on this subject and the famous conjecture concerning the Anderson model. We believe that this thesis provides a solid theoretical foundation for understanding spectral algorithms and spectral graph theory. We also propose possible extensions and new research directions.

Remerciements

Les mathématiques sont souvent perçues comme une matière hermétique pour laquelle les prédispositions font tout. En ce qui me concerne, ce sont les professeurs qui ont été déterminants. À ce titre, je veux remercier ceux qui ont eu ce rôle pour moi. Augustin mon oncle qui m'a le premier introduit à la rigueur. Katheryn Hess Bellwald qui a créé le Cours Euler et Jérôme Scherer qui l'a mené durant des années. Je le remercie pour sa patience. Ma soeur Camille et Ancelin qui m'ont beaucoup appris dans ma jeunesse. Je tiens aussi à remercier Tom Mountford, avec qui j'ai découvert la théorie des probabilités et grâce à qui j'en ai fait ma spécialité. Alain-Sol Sznitman pour m'avoir supervisé durant mon travail de master. Johannes Alt et Raphael Ducatez, qui en plus d'avoir été des collègues et amis, m'ont énormément appris sur les mathématiques et les matrices. Finalement, je remercie Antti Knowles. D'abord pour m'avoir fait confiance il y a cinq ans pour commencer un doctorat sous sa direction. Ensuite pour m'avoir donné un bon sujet de recherche, des conditions de travail idéales ainsi que son soutien. Finalement, c'est sous sa supervision, au fil des nombreuses discussions, élaborations et petites disputes, que j'ai acquis une maturité, une confiance et une identité scientifique. La simplicité, la curiosité et l'endurance. Merci.

La thèse est l'achèvement d'un long voyage. Et les compagnons donnent tout leur sel aux périples. A ce titre je veux ici remercier les amis qui m'ont accompagné durant mes études. Louis, Gabriel, Nicolas, Moritz et Igor à l'EPFL. Alain, François et Alain-Sam à Zurich avec qui j'ai appris qu'être solide ne signifiait pas nécessairement se prendre la tête. Enfin à Genève, les personnes sont nombreuses. D'abord Grégoire et Thibaut avec qui j'ai partagé un bureau et mille conversations durant notre séjour aux Acacias. Adrien, caporal Eiichi, Marco, Tommaso, François, Giovanni, Yaroslav, Marco, Léo, Philippe, pour les parties de foot, la caféine partagée et les éclats de rire. Les adultes ne sont pas en reste: un grand merci à Joselle Besson, Rinat Kashaev, Max et Van Hongler, Sylvain Sardy, Dirk Van der Marel, Yvan Velenik.

Autour de la thèse, il y a la vie. Je remercie d'abord mes amis, d'enfance et d'ailleurs. On ne comprend jamais aussi bien l'importance de l'amitié qu'après une longue course et une belle discussion autour d'une bière. Ou lorsqu'un message d'encouragement tombe à point. Lara, Stan, Yannick, Gauvain, Simon un grand merci. Ma famille et mes proches ensuite, qui, évidemment, sont là depuis bien avant le début de ma thèse: Audrey, François, Georg, Jeeun et Sejong, Adrien, Florence et Kambiz, Marie-Christine, JC et le Valais, Cristina, Anouk, Camille et la famille, Solange et Léo. Et Roxane, évidemment.

Renaud

Contents

List of Figures

Chapter 1

Introduction

1.1 Random matrices and disordered systems

The study of complex systems has been an inspiring source of mathematical activity for more than a hundred years. As the scale to which physical science studied the world became smaller, understanding the properties of nature required physicists to handle large systems with many particles and interactions. The idea to use randomness to analyze simplified models turned out to be a very powerful technic. Perhaps the most popular model is the Ising model used to describe ferromagnetism in a piece of metal (see [18] for a historical note on the matter). As anybody who followed an introductory lecture to quantum mechanics knows, matrices are the mathematical tool used to describe matter on very small scales. Indeed, the Hamiltonian of a quantum system is a matrix. While the first systems taught in quantum mechanics class are simple (two-level systems, harmonic oscillators), some systems can become very large and too difficult to study using standard linear algebra technics. Computing the spectral decomposition of a two-by-two matrix is easy, doing it for a ten-by-ten matrix is already much more arduous. In parallel to statistical mechanics, the study of large quantum systems borrowed the formalism of probability theorem to mathematics. The combination of matrices (quantum Hamiltonians) and randomness (probability theory) resulted in random matrices introduced by Wigner [50] in his seminal work. It is also worth noting that, long before Wigner's work on the semi-circle law, random matrices were studied as a tool for statistics most notably by Fisher, Pearson and Wishart (see [41], [51], [26] and [52]).

The second half of the twentieth century witnessed the remarkable emergence of computer science, from physic-based hardware innovations to the establishment of the development and analysis of algorithms as a standalone mathematical field. The notion of a graph became extremely popular as a way to describe many real-life problems. For instance, the problem of finding the shortest itinerary between two cities, which has countless applications nowadays, was elegantly solved by E. W. Dijkstra in 1956 ([21]). Understanding large graphs became a practical problem that impacted many aspects of modern life. This mathematical object received even greater academic attention with the advent of the world wide web. Indeed the formalism of graphs was a natural way to encode the relationship between different internet agents. The Page-Rank algorithm, which was at the heart of Google's search engine for years ([40]), is a famous example of a real-life application of probability theory (Markov chains) and linear algebra (Perron-Frobenius theorem) combined to extract meaningful information from large random graphs.

In this thesis, we consider the simplest model for a random graph and investigate the spectral properties of its Laplacian matrix. The Erdős-Rényi graph on $n \in \mathbb{N}^*$ vertices is the graph obtained by keeping every edge of the complete graph with some probability $p \in [0, 1]$. In their seminal paper [22] P. Erdős and A. Rényi showed that this graph undergoes a strong phase transition around the value $p = 1/n$, where connected components of macroscopic size begin to appear. Following that paper, the asymptotic behavior of the Erdős-Rényi graph when $n \to +\infty$ was extensively studied (see [16] and [31] for a summary) for various regimes of $p = p(n)$ and a host of different techniques were developed (for instance see [36] for an elegant proof using the Breadth-First Search algorithm). Together with the adjacency matrix, the Laplacian is a canonical way to encode the structure of a graph. From a physical point of view, it is the Hamiltonian of a quantum particle living on the graph. From a mathematical point of view, it is an important object to study and understand the behavior of random walks on graphs. Our work focuses on the spectrum of the Laplacian: we analyze the eigenvalues and eigenvectors of the Laplacian matrix and their asymptotic behavior, as the size of the graph goes to infinity ($n \to +\infty$) and under various regimes ($d := n \cdot p$). Understanding the properties of the spectrum is essential when using algorithms relying on spectral analysis of the graph. In the same way that doing data analysis without knowing about the central limit theorem leads to flawed conclusions, carelessly using spectral algorithms without understanding the asymptotic behavior of large random matrices will inevitably lead to erroneous results. For instance, the Page Rank algorithm does not directly analyze the adjacency matrix of the internet graph but rather a mixture of the internet and a mean-field component called the random surfer model. Although the reason for their original choice is not rigorously justified in their paper, the authors of [40] obtain more meaningful results with that modification. The present thesis gives an explication of that phenomenon, in the case of the Laplacian matrix, by showing that only a limited amount of information can be extracted from the maximal eigenvector of the Laplacian. The current work also answers important graph theoretic questions, in particular about the spectral gap of random graphs. In [10], the authors showed that a positive fraction of random regular graphs is Ramanujan graphs. This result followed as an easy corollary of a more profound analysis of the law of the extreme eigenvalue of some class of random matrices (Tracy-Widom distribution). In this thesis, we provide an estimate for the spectral gap of Erdős-Rényi graphs in a wide range of regimes by using the same philosophy: first performing a complete analysis of the edge of the spectrum and then converting this information into an understanding of the spectral gap. In particular, we exhibit regimes of d for which the spectral gap is determined by functions of small finite trees. Finally, by far the most technical result in the present work is a local law for the Green function of the Laplacian, extending to critical and subcritical regimes previously known results; from an information point of view however, this local law is a negative result since it tells us that no meaningful information can be extracted from the bulk of the spectrum of the Laplacian. Indeed local laws are a reflection of the mean-field nature of the model, with their simplest being the complete eigenvector delocalization. In terms of information, we believe our results serve as a valuable, and mathematically rigorous, reminder of the inherent limitations of spectral algorithms. At the end of the introduction, we suggest lines of research that could open new possibilities to extract information from the spectrum of the Laplacian.

1.2 The Erdős-Rényi graph Laplacian

The Erdős-Rényi is the simplest model of a random graph. Let $N \in \mathbb{N}^*$ and K_N be the complete graph on N vertices.

Notation. For $n \in \mathbb{N}$ a natural number we write

$$[n] := \{1, \dots, n\}. \tag{1.1}$$

Let $d \in [0, N]$. For every edge of $E(K_N)$ we toss a biased coin with head probability d/N and tail probability $1 - d/N$. We denote by X_e the random variable that realizes this experiment. A realization of the Erdős-Rényi graph is given by keeping all edges for which the result was head,

$$\mathbb{G}(\omega) = \big(V, E(\omega)\big), \qquad E(\omega) := \{e \in E(K_N) : X_e \text{ is head}\}.$$

Notation. For a graph G, we denote by $V = V(G)$ the set of vertices and by $E = E(G)$ the set of edges. The graph distance is the length of the shortest path on G and is denoted by $d = d_G$.
For $x, y \in V(G)$ we write $x \sim y$ to mean that x and y are connected by an edge.
We use the capital letter N to denote the size of the system. In the case of the Erdős-Rényi graph, N denotes the number of vertices of $\mathbb{G}$.

This construction is best summarized as

The Erdős-Rényi graph is the subgraph of the complete graph obtained by removing every edge independently at random.

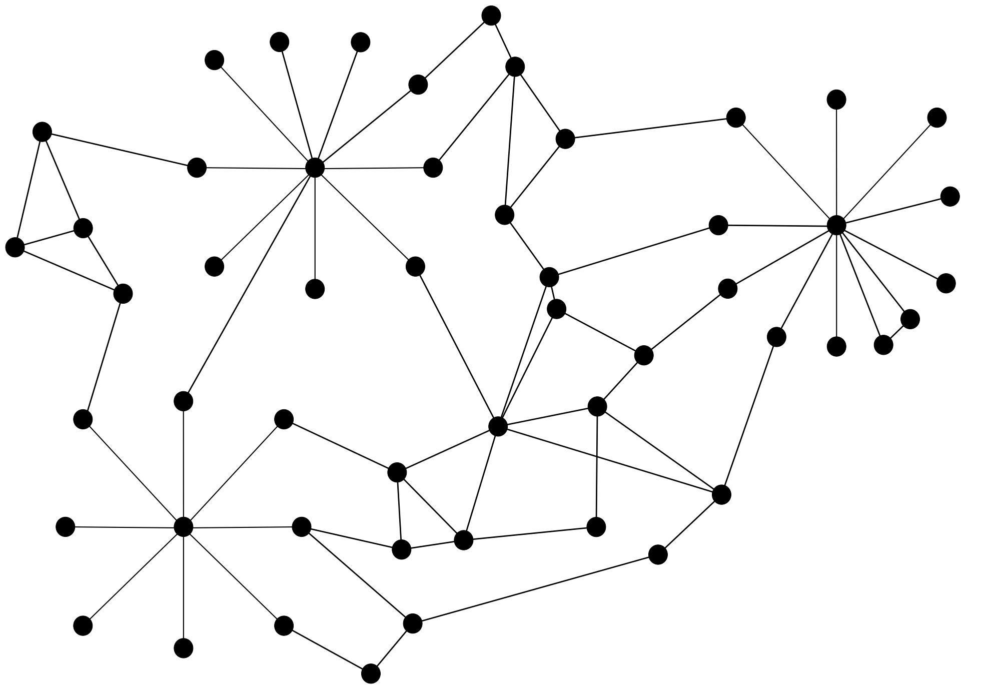


Figure 1.1: A realisation of the Erdős-Rényi graph for $N = 53$ vertices.

Formal construction and behavior of the degree sequence

Let $N \in \mathbb{N}^*$ and $0 < d < N$. We denote by $\mathrm{Bern}(d/N)$ a Bernoulli distribution of parameter d/N. The Erdős-Rényi graph with parameters N and d is a random graph with vertex set $[N]$ and whose adjacency matrix A is defined via the law of its entries, $(A_{xy} : x < y)$, which are independent Bernoulli random variables with parameter d/N. The graph has no loop and so $A_{xx} = 0$ for every $x \in [N]$.
Let $D = \mathrm{Diag}(D_1, \dots, D_N)$ the diagonal matrix with the degrees defined as

$$D_x = \sum_y A_{xy}, \quad x \in [N].$$

The Laplacian matrix is defined as

$$L = D - A. \tag{1.2}$$

Definition. Let $n \in \mathbb{N}^*$ H be a matrix. A number $\lambda \in \mathbb{C}$ is an eigenvalue of H if it satisfies the equality

$$H\mathbf{v} = \lambda \mathbf{v},$$

for some vector $\mathbf{v} \in \mathbb{R}^n$, called the eigenvector.
The (geometric) multiplicity of λ is the dimension of the vector space spanned by all the eigenvectors of λ.

Lemma. *The matrix L is a symmetric, positive definite, real N-by-N matrix. The Laplacian has an eigenvalue at 0 whose multiplicity is equal to the number of connected components of $\mathbb{G}$.*
Moreover, for every connected component $C \subseteq V$, the vector $\mathbf{q}_C := \frac{1}{\sqrt{|C|}}\mathbf{1}_C$ is a normalized eigenvector of L with corresponding eigenvalue 0.

Proof. The proof of that lemma is standard and we only give a short explanation. The fact that L is symmetric follows from the construction of the adjacency matrix A and the fact that the edges of $\mathbb{G}$ have

no orientation. If C is a connected component of $\mathbb{G}$, using the fact that $\sum_{y\neq x} A_{xy} = -D_x$, $x \in [N]$, it is straightforward to see that $\mathbf{q}_C$ is an eigenvalue of L.
If $C_1, \dots, C_n$, $n \in \mathbb{N}$, are all the connected components of $\mathbb{G}$, it is clear that $(\mathbf{q}_{C_i} : i \in [n])$ form an system of orthonormal eigenvectors. This shows that the multiplicity of 0 is greater than the number of connected components. If $\mathbf{v} \in \mathbb{R}^{|V|}$ is an eigenvector of L with eigenvalue 0, then

$$0 = \langle \mathbf{v}, L\mathbf{v} \rangle = \sum_{C \text{ c.c. of } \mathbb{G}} \sum_{x,y \in C} \big(\mathbf{v}_x - \mathbf{v}_y\big)^2,$$

and thus $\mathbf{v}$ must be constant on every connected component of $\mathbb{G}$. This shows that $\mathbf{v}$ is a linear combination of $\mathbf{q}_{C_i}$, $i \in [n]$ and that the multiplicity of 0 is smaller than the number of connected components. This concludes the proof. □

Remark. The Laplacian of the Erdős-Rényi can be studied as a toy model for quantum mechanics. Indeed L is the quantum Hamiltonian of a free particle on $\mathbb{G}$.

As we mentioned at the beginning of the chapter, this thesis is focused on the analysis of the spectrum of the matrix L. As elementary as the notion of eigenvalues and eigenvectors may sound, it is not at all obvious how one should study those objects for large, possibly random, matrices. The only facts we know *a priori* about the spectrum of L, are that the eigenvectors of L form an orthonormal basis of $\mathbb{R}^N$ and the eigenvalues are real, since the matrix is symmetric (c.f. the spectral theorem [46, Theorem 1.3.1]). However, how is one supposed to go about computing the (non-trivial) eigenvalues of L? While elegant and simple to memorize, the method of the characteristic polynomial to find the eigenvalues of L followed by Gauss elimination to solve the equation $(L - \lambda)\mathbf{v} = 0$ seems too cumbersome for dimensions greater than 3.
The Courant-Fisher characterization of the eigenvalues (also known as the min-max principle) is a set of powerful equalities that expresses the eigenvalues as the solutions of a variational problem ([29, Theorem 4.2.6]). If $\lambda_{\max}$ denotes the largest eigenvalue of L (equivalently the spectral radius of L since L is positive definite), the Courant-Fisher inequalities simplify to the Rayleigh quotient and yield

$$\lambda_{\max} = \max_{\mathbf{v}\in\mathbb{R}^N,\, \|\mathbf{v}\|=1} \langle \mathbf{v}, L\mathbf{v} \rangle$$

In particular, consider $x \in [N]$ the vertex with maximal degree $D_x = \max_{y\in[N]} D_y$ and consider the test vector $\mathbf{v} = \mathbf{1}_x$. Then

$$\lambda_{\max} \geq \langle \mathbf{1}_x , L\mathbf{1}_x \rangle = \Big\langle \mathbf{1}_x , D_x\mathbf{1}_x - \sum_{y\sim x} \mathbf{1}_y \Big\rangle = \max_{y\in[N]} D_y. \tag{1.3}$$

While the degrees can take any value from 0 to N, there are some indications that the matrix A should be much smaller than D. First of all the entries of A are bounded (by 1) whereas the degrees can take any value between 0 and N (but typically the largest degree is of size $O(d + \log N)$). Moreover, some heuristics indicate that the graph $\mathbb{G}$ contains very few cycles and locally resembles a tree. In this case, basic spectral theory tells us that $\|A\| \leq 2\sqrt{\Delta}$, where Δ is the maximum degree of the graph. In [11], the authors showed that the matrix $\underline{A} := \frac{1}{\sqrt{d}}\big(A - \mathbb{E}A\big)$ was shown to have extreme eigenvalues located at

$$\lambda_{(N-k)}(\underline{A}) = \sqrt{D_{(k)}/d} + O\Big(1 + \Big(\frac{\log N}{d\log(\log N/d)}\Big)^{1/2}\Big),$$

where $D_{(k)}$ is the k-th largest degree.
Considering regimes for which $d \gg \log N$ and glossing over some technicalities, we can use this information and standard perturbation theory to deduce that

$$\lambda_{\max} \leq \max_{x\in[N]} D_x + \sqrt{d}\|\underline{A}\| \leq \max_{x\in[N]} D_x + \sqrt{D_x} + o(1) \leq \max_{x\in[N]} D_x\big(1 + \sqrt{1/D_x}\big) + o(1). \tag{1.4}$$

Putting (1.3) and (1.4) together, we find $\lambda_{\max} = \max_{x\in D_x} D_x(1 + o(1))$.

This rough analysis suggests that understanding the statistics of the largest degree vertices is a good starting point to understand the extreme eigenvalues of L. As it turns out, the degree sequence of the Erdős-Rényi has been well-understood (see [16]).

Lemma. *The degrees* $\big(D_x : x \in [N]\big)$ *follow a binomial distribution of parameter* $(N-1, d/N)$*. Moreover, the random variables are weakly correlated, for* $x, y \in [N]$*,* $x \neq y$*,,*

$$\mathbb{E}(D_x) = d\Big(1 - \frac{1}{N}\Big), \quad \mathrm{Var}(D_x) = \sqrt{d}\Big(1 - \frac{1}{N}\Big), \quad \mathrm{Cov}(D_x, D_y) = \frac{d}{N}\Big(1 - \frac{d}{N}\Big),$$

Proof. The first statement follows from the definition of D_x and the independence of the random variables A_{xy}. Computations of the mean and variance are standard. For $x \neq y$, we have

$$\begin{aligned}\mathrm{Cov}(D_x, D_y) &= \mathbb{E}\big[D_x D_y\big] - \mathbb{E}[D_x]\mathbb{E}[D_y] = \mathbb{E}\Big(\sum_{u\neq x, v\neq y} A_{ux}A_{yv}\Big) - d^2\Big(1 - \frac{1}{N}\Big)^2 \\ &= \mathbb{E}\big[A_{xy}^2\big] + N^2\Big(1 - \frac{2}{N}\Big)\frac{d^2}{N^2} - d^2\Big(1 - \frac{1}{N}\Big)^2 = \frac{d}{N}\Big(1 - \frac{d}{N}\Big).\end{aligned}$$

□

Remark. Some authors prefer to consider the Erdős-Rényi on $N+1$ vertices and to have $\mathbb{E}D_x = d$. However since we usually consider parameters for which $d \ll N$ the factor $1 - \frac{d}{N}$ has no impact on the computations.

The best way to analyze the distribution of the degree is to use the Poisson approximation of the binomial law. Indeed, as is well-known from elementary probability, if $X_n \sim \mathcal{B}_{n,p}$, $n \in \mathbb{N}$ with the parameter p depending on n and such that $\lim p \cdot n = d \in \mathbb{R}$, then the binomial random variables converge in distribution towards a Poisson distribution, $X_n \Rightarrow Y$, with $Y \sim \mathcal{P}(d)$. This result is formalized in Lemma B.5. As the central limit theorem states (Proposition B.1), after an appropriate scaling the binomial distribution converges toward the normal law, $\frac{1}{\sqrt{d}}(X_n - d) \Rightarrow Z$, $Z \sim \mathcal{N}(0,1)$. A perhaps less famous asymptotic property of Poisson variables is that for $d \gg \log N$, the tails of the distribution of X_n are well-described by a normal law $Z \sim \mathcal{N}(0,1)$ (see Remark B.6) but as $d \lesssim \log N$, this is no longer the case (see Section B.2). These approximations give us an important heuristic that we will use throughout this work

$$D_x \overset{(\mathrm{d})}{=} \mathcal{B}_{N,d,} \sim \mathcal{P}(d) \sim \begin{cases} \mathcal{N}(d, \sqrt{d}), & \text{everywhere for } d \gg \log N \\ \mathcal{N}(d, \sqrt{d}), & \text{in the neighborhood of } d, \text{ for all regimes.} \end{cases} \tag{1.5}$$

The regime $d \asymp \log N$ is thus a limit at which the statistics of $\max_{x\in[N]} D_x$ undergo a strong transition. This corresponds to the apparition of *inhomogeneity* in the graph (see [6, Section 2]).

The Erdős-Rényi graph undergoes two phase transitions as a function of d. As d becomes larger than 1, there is a unique large connected component. As d crosses the value $\log N$, the graph becomes connected with high probability ([16, Theorem 7.3]). We usually distinguish three regimes

1. The supercritical regime: $d \gg \log N$, the graph is connected (with high probability and homogeneous);
2. The critical regime: $d \asymp \log N$, the graph becomes inhomogeneous;
3. The subcritical regime: $d \ll \log N$, the graph is inhomogeneous and disconnected.

An intuitive way to understand the second phase transition is to consider the expected number of isolated vertices in $\mathbb{G}$ as a function of d. Let $x \in [N]$. Then by definition of the model the degree D_x follows a binomial distribution of parameter N and d/N. By elementary analysis, the binomial law is well approximated by a Poisson d distribution. In this sense we have

$$\mathbb{E}\big[\big|\{x \in [N] : D_x = 0\}\big|\big] \sim N\mathbb{P}\big(\mathcal{P}_d = 0\big) = N\mathrm{e}^{-d} \begin{cases} \ll 1 & \text{if } d \gg \log N, \\ \equiv 1 & \text{if } d \asymp \log N, \\ \gg 1 & \text{if } d \ll \log N. \end{cases} \tag{1.6}$$

Now isolated vertices are the simplest disconnected components and we can expect that they can be used as canaries for the emergence of disconnected components. This observation has an immediate consequence on the spectrum of L as it increases the multiplicity of the trivial eigenvalue.

Bulk versus Edge of L

We explained above how the degrees of $\mathbb{G}$ can be thought of as weakly correlated Poisson variables of mean d. A naive guess would be to locate the eigenvalues of L roughly at the same place as the eigenvalues of D, i.e. at the degrees of $\mathbb{G}$. An important heuristic in random matrix theory is to divide the spectrum of random matrix ensembles into two parts, the *bulk* of the spectrum and the *edge*. The bulk is the region where the expected density of states is positive (i.e. in our case the regions where there are $O(N)$ eigenvalues). The edge is typically defined as the region where the maximal (or minimal) eigenvalue resides.
After the brief intuition we developed about the relationship between degrees of $\mathbb{G}$ and eigenvalues of L, it is natural to use the diagonal matrix D to define the bulk, the edge and the mid-region. We define informally the bulk as all energies that lie within $O(1)$ standard deviations of the expected degree and the edge as the region on which $\max_{x\in[N]} D_x$ is supported. We informally define those areas as

$$\text{Bulk} := [d - C\sqrt{d}, d + C\sqrt{d}], \qquad \text{Edge} := [\Delta - \omega, \Delta + \omega], \qquad \Delta := \max_{x\in[N]} D_x,$$

for some $C > 0$ and $\omega \gg 1$.

It is more convenient to study the *rescaled* version of the Laplacian defined by

$$\underline{L} := \frac{L - d}{\sqrt{d}}. \tag{1.7}$$

With this rescaling, the diagonal entries of our matrix are the *rescaled degrees*, denoted by $v_x := \frac{D_x - d}{\sqrt{d}}$, and the bulk-edge dichotomy becomes

$$\text{Bulk} := [-C, C], \qquad \text{Edge} := [\Delta - \omega, \Delta + \omega], \quad \Delta = \begin{cases} \sqrt{2\log N} + \log\sqrt{2\pi}\log N, & d \gg \log N, \\ \frac{\log N}{\sqrt{d}\log(\log N/d)}, & d \ll \log N. \end{cases}$$

1.3 Simulations and overview of results

The analysis of large random graphs is the subject of abundant literature (see for instance [16]). The most famous and impactful example of the efficiency of spectral analysis of random graphs is probably the PageRank algorithm that was introduced in [40]. A central question throughout this work is the following

Can meaningful information be extracted from the eigenvalues and eigenvectors of $\underline{L}$?

For a Hermitian matrix H, we usually denote by $\operatorname{Spec} H$ the set of its eigenvalues and for $\lambda \in \operatorname{Spec} H$, we write $\mathbf{w}_\lambda$ its eigenvector. Note that eigenvectors are defined up to a phase, which we assume to be 1 for simplicity.
An important question in random matrix theory is the distribution of eigenvalues. An even more important question, and also a more difficult one, is the behavior of eigenvectors. Eigenvectors are typically normalized in such a way that $\|\mathbf{w}_\lambda\|^2 = 1$, $\lambda \in \operatorname{Spec}(L)$. For a given eigenvector $\mathbf{w}_\lambda$, $\lambda \in \operatorname{Spec} L$, the square of its coefficients naturally defines a probability distribution on the set of vertices of $\mathbb{G}$. Indeed the function $f : V(\mathbb{G}) \to [0,1]$, $f(x) = |\mathbf{w}_\lambda(x)|^2$, is a density function whose total mass is 1. What does the graph of f look like?

Remark. The question of localization and delocalization of eigenvector can be understood by looking at a simple two-level system. Let us consider

$$H = \begin{bmatrix} v_1 & \varepsilon \\ \varepsilon & v_2 \end{bmatrix}, \qquad v_1, v_2, \varepsilon \in \mathbb{R}, \quad v_1 \neq v_2.$$

The spectrum of H is

$$\lambda_\pm = \frac{1}{2}\big(v_1 + v_2 \pm \sqrt{(v_1 - v_2)^2 + 4\varepsilon^2}\big) = \frac{1}{2}\big(v_1 + v_2 \pm (v_1 - v_2)\sqrt{1 - 4\Delta^2}\big), \quad \Delta := \frac{\varepsilon}{v_1 - v_2}.$$

Using the first order approximation $\sqrt{1-t} = 1 - \frac{1}{2}t + O(t^2)$, we see that

$$\mathbf{w}_{\lambda_+} \sim \frac{1}{\sqrt{1+\Delta^2}} \begin{pmatrix} 1 \\ \Delta \end{pmatrix}, \quad \mathbf{w}_{\lambda_-} \sim \frac{1}{\sqrt{1+\Delta^2}} \begin{pmatrix} -\Delta \\ 1 \end{pmatrix}.$$

We see that $\mathbf{w}_{\lambda_+}$ goes from being completely localised at $\varepsilon = 0$ to being completely delocalised for $\varepsilon = v_1 - v_2$. The mechanism is driven by the *ratio* between the spectral gap $v_1 - v_2$ and the interaction strength ε.

Suppose $x \in [N]$ is an isolated vertex. In that case it is supports an eigenvector $\mathbf{w} = \mathbf{1}_x$ with eigenvalue 0. The function f corresponding to $\mathbf{w}$ is thus completely *localized* on one point of the graph. On the other suppose the graph $\mathbb{G}$ is connected. In that case, the eigenvalue of 0 has multiplicity one and corresponding eigenvector $:= \frac{1}{\sqrt{N}}\mathbf{1}_{[N]}$. In that case the function f is constant and *delocalized*, since $f(x) = 1/N$.

Those two scenarios represent the extreme behaviors of eigenvectors. However, the shape of the eigenvector is not constrained to those two extreme postures and can vary between being completely flat and being completely localized. In fact, in the spectrum of the Laplacian matrix, the whole range of behavior seems to coexist simultaneously (i.e. for one realization of $\mathbb{G}$) depending on the region of the spectrum. In the following simulation, we consider the rescaled Laplacian matrix (defined (1.7)).

Remark. (Physical and informational interpretation of $|\mathbf{w}(x)|^2$.) The shape of eigenvectors has some important interpretations in quantum mechanics. A delocalized eigenvector is a synonym for a state in which the electron can travel through the medium. On the other hand, a localized eigenvector corresponds to a state which is trapped in some region of the medium, meaning that the electron cannot travel and the medium is an isolator.
In the field of computer sciences, the Perron-Frobenius eigenvector of the *random surfer* matrix, used in the PageRank algorithm ([40]) is the mathematical object used to rank the webpages according to their importance in the web. Let's say that $x, y \in V(G)$ are two vertices (webpages) of the internet graph, then the page x is recommended to you before the page y if and only if $|\mathbf{w}(x)|^2 \geq |\mathbf{w}(x)|^2$, where $\mathbf{w}$ is the Perron-Frobenius eigenvector.

Note that similar pictures would appear for different values of d. For $d \gg \log N$, the picture is more symmetric around 0 while for $d \ll \log N$, more points accumulate at the left of the spectrum and the multiplicity of $-\sqrt{d}$ increases.

As a point of comparison, we plot the same picture but for a matrix sampled from the Gaussian Orthogonal Ensemble (GOE). Let $H \in \mathbb{R}^{N\times N}$ be a Hermitian matrix distributed as

$$H_{xy} = H_{yx} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1), \quad x, y \in [N].$$

It is well-known in the literature that the eigenvectors of H are completely delocalized and that the density of state converges to the semi-circle law. The argument relies on the symmetry under orthogonal transformations of the normal law. The interested reader can also refer to [5, Figure A.1] to see how the same plot looks when we consider only the adjacency matrix.

In the present thesis, we investigate various regions of the spectrum. We structure our research around four results, Theorems A, B, C and D.

The three regions of the spectrum depicted in Figure 1.3 all bear a special importance in spectral analysis. The analysis of the edge of the spectrum has a long-standing tradition both in mathematics and physics, either when studying the spectral gap and the ground state energy of a quantum system or when analyzing the law of the largest eigenvalues and the spectral radius of some matrix ensemble. The bulk of the spectrum on the other hand has been a mainstay in random matrix theory ever since the seminal paper of [37].

Since the Erdős-Rényi graph is a random object, all of our results must incorporate this notion of randomness.

Notation. An event Ω (that may depend on N) holds with high probability if

$$\mathbb{P}(\Omega) \geq 1 - o(1).$$

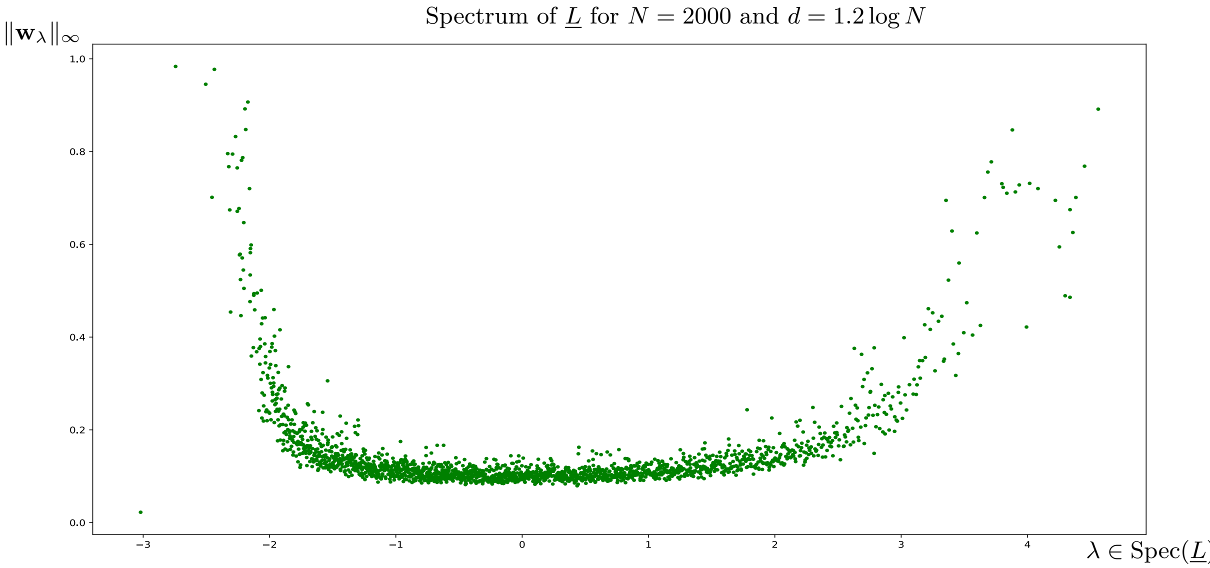


Figure 1.2: The spectrum of the Laplacian matrix is presented here as a scatter plot. Simulation for $N = 2000$ and $d = C\log N$, $C = 1.2$. We plotted the points $(\lambda, \|\mathbf{w}_\lambda\|_\infty)$ for $\lambda \in \mathrm{Spec}\,\underline{L}$. The spectrum is not uniformly distributed as the region around 0 seems to be much more populated than the region at the right and left extremities. The simulation shows four distinct behaviors. In the middle, the eigenvectors seem to be flat. At the extremities of the spectrum, the eigenvectors are localized. One exception to that observation is the eigenvector that corresponds to the smallest eigenvalue: however, this is easily explained once we recognize that this is the trivial eigenvalue of the Laplacian matrix whose eigenvector is given by $\frac{1}{\sqrt{N}}\mathbf{1}_{[N]}$. Finally we see that there is a smooth transition between the middle and the two extremities of the spectrum.

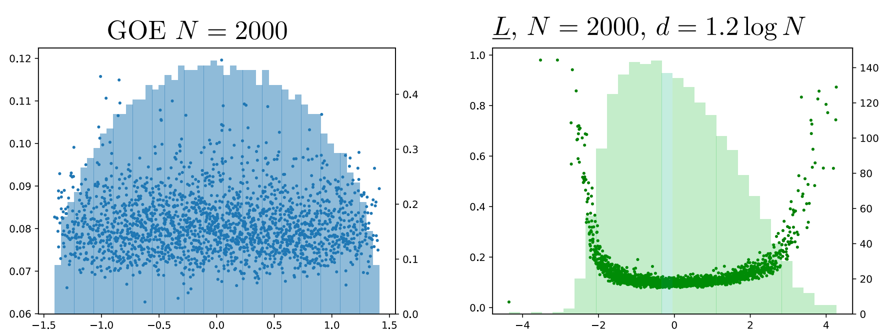


Figure 1.3: On the left-hand side, the simulation for a GOE. The numerical results agree with the results proved in the literature. Namely that the density of states converges to a semi-circle law (clear from the picture) and that the eigenvectors of a GOE are all "flat". On the right-hand side, the same simulation as above but this time with the density of states plotted in the background as a histogram.

The thesis is organized as follows.

- In Chapter 2 we investigate the bulk of the spectrum and prove Theorem A. The results and techniques fall into the literature on local laws.
- In Chapter 3 we investigate the right and left edge of the spectrum and prove Theorems B and C. The general philosophy of the proof is to build a bijection between high-degree (respectively low-degree) vertices and the maximal (respectively minimal) eigenvalues of $\underline{L}$. The techniques rely on perturbation

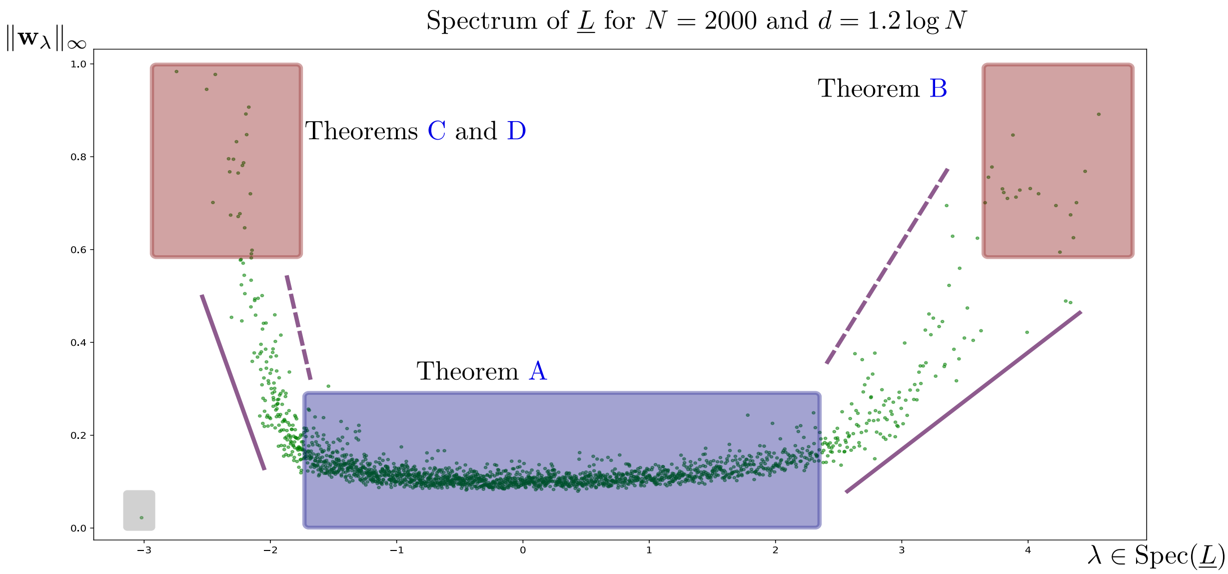


Figure 1.4: The above simulation is the same as the one in Figure 1.2. Theorem A provides a rigorous understanding of the bulk of the spectrum of $\underline{L}$ (blue region), while Theorems B, C and D describe the edge of the spectrum (red regions). The transition between bulk and edge is explained by a partial localization argument (continuous purple lines). The partial delocalization counterpart is unknown (dashed purple lines).

theory with a spectral gap.

- In Chapter 4 we investigate the left edge of the spectrum. We prove that for regimes of d smaller than $\frac{1}{2} \log N$, there is a natural matching between the smallest eigenvalues of $\underline{L}$ and trees embedded in $\mathbb{G}$. Our argument relies on rank-one perturbation theory.
- In Chapter 5 we collect quantitative estimates and geometric results on the Erdős-Rényi . Many arguments in Chapter 3 require those informations.
- In Appendix A we collect result of linear algebra. Appendix B contains various probabilistic results. Appendix C summarizes all notations used in the thesis.

The results in this thesis have been proved under different assumptions. For instance, the extreme eigenvalue statistics of Laplacian matrices has been studied with different assumptions and results (see [33] or [20]). The bulk of the spectrum of $\underline{L}$ was first studied in [32] and [19] and in [30] the authors managed to prove a local law for polynomial values of d. In figure 1.5, we provide a word map of the various results already known on the subject of Laplacian matrices and where our results fall.

1.4 Local laws for sparse deformed Wigner matrices

Local laws provide control over the limiting behavior of the Green function entries. Such local laws have been a central part of random matrices theory ever since the seminal works of [23] and [24]. More precisely for $(H_N, N \geq 1)$ a sequence of random matrices, local laws aim to understand the entries of the matrices

$$G_N(z) := \frac{1}{H_N - z}, \quad z \in \mathbb{H}.$$

Remark. The following section explains the methods used to prove the local law. We only describe the methods that are relatively original and often refer to "typical strategies" or "most common techniques".

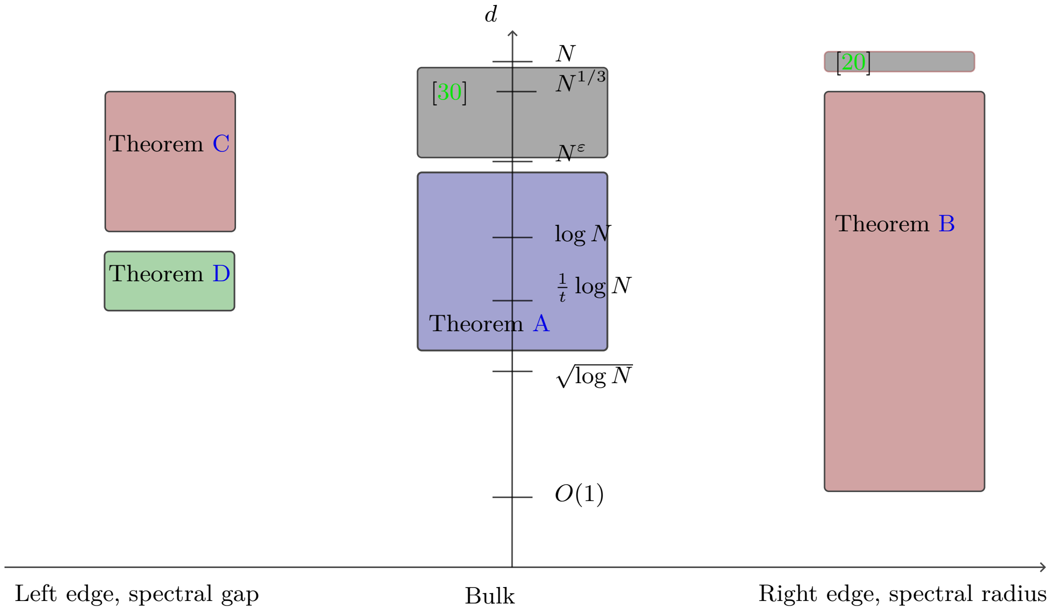


Figure 1.5: The different regimes covered by the thesis. In grey, results previously known. In blue, the results found in Chapter 2, where we improved [30] arguments to push down the result to the regimes $\sqrt{\log N}$. In red, the results found in Chapter 3 where we show how spectral statistics are closely related to extreme degrees and how the corresponding eigenvectors have clear geometric interpretations. In green, the results found in Chapter 4, where we show how trees embedded in the macroscopic connected component generated the spectral gap of $\underline{L}$.

This chapter is designed for a public familiar with local laws and is by no means a crash course on that subject (it is a difficult subject). However, the interested reader can find good lecture notes on the matter in [14] (shorter) and [25] (longer).

The most common ensemble of random matrices is the *Wigner* ensemble which satisfies the following properties

1. the matrix H is symmetric.
2. the entries are independent (up to the symmetry constraint), i.e. the collection $(H_{xy} : 1 \leq x \leq y \leq N)$ form an independent family of random variables.
3. the entries are centered, with variance $\mathbb{E}H_{xy}^2 = \frac{1}{N}$ and the random variables $\sqrt{N}H_{xy}$ are bounded in any L^p space, uniformly in N, i, j.

For a good introduction to local laws for Wigner matrices see for instance [14].

The starting point of any local law is the Schur complement formula

$$G_{xx} = \frac{1}{H_{xx} - z - \sum_{y,z}^{(x)} H_{xy} G_{yz}^{(z)} H_{zx}}. \tag{1.8}$$

All notations including $\sum^{()}$ and $G^{()}$ will be introduced in Chapter 2. The general strategy in the proof of local laws is to use large deviation estimates to show that

$$\sum_{y,z}^{(x)} H_{xy} G_{yz}^{(z)} H_{zx} = g + \sum_{y}^{(x)} \Big(H_{xy}^2 - \frac{1}{N} \Big) G_{yy} + \mathcal{E}_x = g + o(1) \quad g(z) := \frac{1}{N} \operatorname{Tr} G_{zz}(z). \tag{1.9}$$

Plugging this back into (1.8) yields a self-consistent equation for the normalized trace g which, up to some stability estimates, defines the asymptotic limit of G_{xx}.
For instance, in the case of Wigner matrices, the diagonal term is $H_{xx} = o(1)$ and thus the limit of G_{xx} does not depend on x, namely G_{xx} is asymptotically constant in x. The following local and global laws are well-known (see [14, Theorem 2.4 and 2.6]).

Proposition (Local law for Wigner matrices)**.** *Under the above assumption, we have*

$$G_{xy}(z) - m_{\mathrm{sc}}(z)\delta_{xy} = o(1).$$

uniformly for $x, y = 1, \ldots, N$ *and* $z = E + \mathrm{i}\eta$, $\eta \gg N^{-1}, E \in [-2, 2]$ *In particular,* $g \to m_{\mathrm{sc}}$ *and thus the sample density of states* $\rho := \frac{1}{N}\sum_{\lambda \in \mathrm{Spec}(H)} \delta_\lambda$ *converges weakly to the semi-circle law.*

The Laplacian matrix cannot be normalized in such a way that it becomes a Wigner matrix. Since the off-diagonal matrix A and the diagonal entries D fluctuate on different scales, it is not possible to normalize L so that all rescaled entries simultaneously fluctuate on the same $1/\sqrt{N}$.
This is a typical situation in the study of so-called *deformed Wigner matrices*. Such matrices have been notably studied in [37].
In our case, we will analyze the matrix

$$M := V - H + R, \qquad G := (M - z)^{-1}, \qquad z \in \mathbb{H},$$

where H is a *sparse* Wigner matrix (see below), V is a diagonal matrix whose entries can be morally thought of as $\mathcal{N}(0, 1)$ variables and R is a rank-one perturbation (see Section 2.1). Our first result, Theorem 2.4 shows that

$$G_{xx} \sim \frac{1}{v_x - z - \overline{m}(z)}, \quad z \in \mathrm{Bulk},$$

where $\overline{m}$ is a random meromorphic function on the complex upper-half plane that depends only on the variables $(v_x : x \in [N])$. In that sense, the asymptotic behavior of G does not depend on the geometry of the graph (mean-field behavior).

We prove a local law down to optimal scale and state the result in the shortest form possible here.

Theorem A. *For any constant* $\tau > 0$ *there exists a constant* $C > 0$ *such that if* $d \geq C\sqrt{\log N}$*, then with high probability* $|G_{xy}(z)| \leq C$*, for every* $x, y \in [N]$ *and* $z = E + \mathrm{i}\eta$*,* $|E| \leq \tau^{-1}$ *and* $\eta \geq N^{-1+\tau}$*.*

An immediate consequence of our result is that the eigenvectors of M corresponding to the eigenvalues in the bulk are delocalized. Since by construction the matrices M and L are related by a linear transformation, $M = \frac{1}{\sqrt{d}}(L - d)$, we deduce that with high probability, for any $\tau, R > 0$,

$$\|\mathbf{w}_\lambda\|_\infty = O\big(N^{-1+\tau}\big), \quad \forall \lambda \in \mathrm{Spec}(L) \cap [-R, R]. \tag{1.10}$$

Going back to Figure 1.3, we see that (1.10) explains why in the blue region the eigenvectors' infinity norm is bounded.

The problem of sparsity

In the case of the Laplacian matrix, a few of the usual conditions are not satisfied. It is clear that we cannot simultaneously scale both the diagonal part (degrees) and the off-diagonal part (adjacency) in such a way that the entries of both blocks have variance $1/N$. Let us focus on the adjacency matrix. First of all the entries are not centered random variable of variance $1/N$. This issue can be easily solved by centering and rescaling the off-diagonal entries

$$H_{xy} := \frac{1}{\sqrt{\gamma}}\Big(A_{xy} - \frac{d}{N}\Big), \qquad \gamma := d\Big(1 - \frac{d}{N}\Big).$$

However the assumption 3 for Wigner matrices is not true. Indeed we have

$$\mathbb{E}|H_{xy}|^p = \frac{1}{Nd^{\frac{p}{2}-1}}\Big(1+O\Big(\frac{d}{N}\Big)\Big) \ll \frac{1}{N^{p/2}},$$

as soon as $d \ll N$. But we consider *sparse* Erdős-Rényi graphs and actually in this thesis we consider only values $d \leq N^{1/3}$. This means that in our model the of diagonal entries of H have heavy tails. Such random matrix ensembles are notoriously harder to analyze since the large deviation estimates that can be derived are much weaker than for Wigner matrices. In the context of the Erdős-Rényi graph, [28] circumvented this difficulty by introducing a new class of large deviation for sparse random vectors. Their method is relevant in our setting but with some modifications as explained below. The proof relies on estimating

$$\mathbb{E}\Big|\sum_{i=1}^{n} a_i X_i\Big|^p, \qquad p = O(\log N) \tag{1.11}$$

where $a_i \in \mathbb{C}$ are constant coefficients and X_i are centered, independent random variables with variance $1/N$ and $\mathbb{E}|X_i|^p \leq N^{-1}d^{-\frac{p}{2}-1}$.
Typically the variables are set to $X_i := H_{\ell i}$ for some fix $\ell \in [N]$ and the coefficients are the entries of the matrix $G^{(\ell)} := \big(M^{(\ell)} - z\big)^{-1}$ where $M^{(\ell)}$ is the sub-matrix of M obtained by removing the ℓ-th. Usual estimates (see for instance [14, Lemma 3.6]) are thus harder to get with the weaker control on the moments of X.

The problem of correlations

The matrix L has a non-trivial correlation structure. Since the diagonal entries are the sum of the off-diagonal entries, it is clear that the variables $(L_{xy} : 1 \leq x \leq y \leq N)$ are not independent. Thus the assumption 2 for Wigner matrices is not satisfied for the Laplacian matrix.
These correlations make it impossible to derive large deviation estimates in the usual way (for a definition of usual see [14, Lemma 3.6]). This is an important obstacle to the proof of any local law. Indeed
To illustrate this fact, let us fix $\ell, k \in [N]$ and recall the definition of $G^{(\ell)}$. Let $\tilde{M}^{(\ell)}$ be the analogue of M but for the graph $\mathbb{G}^{(\ell)}$ obtained by removing the vertex ℓ. Let a_{ij} be the coefficients of the matrix $G^{(\ell)}$ and b_{ij} the coefficients of the matrix $\widetilde{G}^{(\ell)} := \big(\tilde{M}^{(\ell)} - z\big)^{-1}$. It is clear that the variables $(b_{ij} : i, j \neq \ell)$ are measurable with respect to $\big(H_{ij} : i, j \in [N] \setminus \{\ell\}\big)$ and in particular independent of X_i. This is not the case for the coefficients $(a_{ij} : i, j \neq \ell)$. Bounding (1.11) is already a difficult task, but it is even harder if the variables $(X_i : i \in [N])$ are not independent from the coefficients.

In [30], the authors address this issue by using the resolvent formalism to replace the coefficients a_{ij} by the coefficients b_{ij}. The second resolvent identity (see Lemma A.3) states that $G_A - G_B = G_A(A-B)G_B$, for $A, B \in \mathbb{C}^{N\times N}$. By setting $A = M^{(\ell)}$ and $B^{(\ell)}$, it is possible to *filter out* the correlations between $\big(H_{i\ell} : i \neq \ell\big)$ at the cost of adding an extra sum. To illustrate their method, we apply the resolvent identity two times

$$\sum_i a_i X_i = \sum_i b_i X_i + \sum_{i,j} b_i X_i a_{ij} X_j = \sum_i b_i X_i + \sum_{i,j} b_i X_i b_{ij} X_j + \sum_{i,j,k} b_i X_i b_{ij} X_j a_{jk} X_k.$$

Now the first two sums on the right hand can be estimated by first conditioning $H^{(\ell)}$ and then bounding moments of random vectors with deterministic coefficients. Applying this trick many times allows us to push further the correlated term a_{ij} and eventually to control it.
In [30], the authors use the resolvent identity $O(1)$ times and thus have to prove large deviation estimates on random vectors of the form

$$\sum_{i_1,\dots,i_l} b_{i_1} X_{i_1} b_{i_1 i_2} \cdots b_{i_{l-1} i_l} X_{i_l}, \qquad l = O(1).$$

One of the key technical achievements in Chapter 2 is to derive large deviation estimates that are efficient for $l = O(\log N)$. The proof relies on two mechanism

1. An analysis of large computation graphs.

2. The fact that off-diagonal entries of the Green function are with high probability of order $o(1)$.

The first mechanism is a general technic that can be found in many proofs of local laws (see for instance [45] and [23]). The second mechanism is somewhat surprisingly not often used in these proofs but it is an important ingredient to overcome the difficulties arising from sparsity and heavy tails.

The problem of inhomogeneity of $\mathbb{G}$

Once the large deviation estimates have been derived and upgraded, a local law up to $d \gtrsim \log N$ can be quickly obtained by following the pipeline laid out in [30]. A central difficulty to go lower than $\log N$ is that the graph becomes *inhomogeneous.* First of all some disconnected components appear as we go below $\log N$ but even before that the behavior of the degree sequences changes (see the next section for more on the matter).

A delicate consequence of this fact is that the *local average* of the Green function is not anymore necessarily well-approximated by the *global average.* The estimate

$$\sum_{y,z}^{(x)}\Big(H_{xy}^2-\frac{1}{N}\Big)G_{yy}=o(1),\qquad \forall x\in[N],$$

is not true anymore. Rather we can say that

$$\exists T\subset[N],\quad |[N]\setminus T|\ll N,\quad \sum_{y,z}^{(x)}\Big(H_{xy}^2-\frac{1}{N}\Big)G_{yy}=o(1),\qquad \forall x\in T.$$

This set T is called the set of typical vertices while the complement is the set of atypical vertices.

Separating typical and atypical vertices is an idea that was developed in [5]. In that article, the authors considered the adjacency matrix of the Erdős-Rényi in the same regimes of d. The phenomenon of inhomogeneity was thus the same. However, the (small but very real) correlations between entries of the Laplacian matrix make the analysis much more complicated.

1.5 Largest eigenvalues, right of the spectrum

The analysis of the extreme eigenvalues of random matrix ensembles has been long-standing (see [13] and [9] for instance). In the case of the Erdős-Rényi graph, as long as $d \ll \log N$, the top eigenvalues were located close to the interval $[-2,2]$. In [6] and [48], the authors showed that the appearance of eigenvalues outside of $[-2,2]$ was shown to be related to the emergence of very large vertices in $\mathbb{G}$. Our main theorem regarding the right of the spectrum states that the eigenvalue process is asymptotically close to a sequence of Poisson Point Processes (PPP). As a by-product of the proof, we also show that the eigenvectors corresponding to the maximal eigenvalues are localized in the graph on balls around large degree vertices.

Our result covers a region of the spectrum in which there are at most $\mathcal{K}=O(\log\log N)$ eigenvalues.

Theorem B. *Let $\varepsilon>0$. There exists a constant $K\geq 0$ such that if $K\leq d\leq N^{\frac{1}{3}-\varepsilon}$ then the following holds with high probability. The point process of the $\mathcal{K}$ largest eigenvalues is asymptotically close to a sequence of PPP . Furthermore, if $(\log\log N)^{1/4}\leq d$, the eigenvectors corresponding to the $\sqrt{\mathcal{K}}$ largest eigenvalues are localized on the balls that surround the vertices with the largest degrees and largest sphere of radius two.*

The law of the largest eigenvalue $\lambda_1(\underline{L})$ is explicitly described. Theorem B explains why in Figure 1.3, the eigenvectors located at the right edge of the spectrum have a high infinity norm, $\|\mathbf{w}_\lambda\|_\infty=1-o(1)$. Indeed we will see in Chapter 3 that the top eigenvectors of $\underline{L}$ have an *exponentially decaying* mass away from their centers of localization.

Neighborhood of large degree vertices

To study the eigenvalues of L it is convenient to study a linear shift of the matrix L, namely

$$\underline{L} := \frac{L-d}{\sqrt{d}}.$$

In that way the diagonal entries are (almost) centered and have variance 1. We define the *rescaled degree* as

$$v_x := \frac{D_x - d}{\sqrt{d}}.$$

Since the largest entries of $\underline{L}$ are given by the largest normalized degrees, understanding the extreme statistics of the variables $(v_x : x \in [N])$ seems to be a natural place to start our investigation of the extreme eigenvalue statistics. Note that the variables v_x can take negative values, but not smaller than $-\sqrt{d}$. Recalling (1.5), we see that the rescaled degrees typically behave as $\mathcal{N}(0,1)$ variables and we expect most of them to take values in a compact interval around 0.

For clarity let us start with the easiest case when $\mathbb{G}$ is in the supercritical regime, meaning that $d \gg \log N$. The matrix $\underline{L}$ can be viewed as a perturbation of the matrix $\underline{D} := \frac{D-d}{\sqrt{d}}$. In [11] and [12] the authors showed that [6], $\frac{1}{d}\|A - \mathbb{E}A| = O(1)$. In addition, as we explained in (1.5), the statistics of v_x are close to the statistics of N weakly correlated normal variables. It is well known from extreme value theory (see for instance [17, Section 14]) that if $X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, the maximum is located around $\sqrt{2\log N}$ and fluctuates on the scale $(\log N)^{-1/2}$.

$$\frac{1}{\tau}\big(\max_i X_i - \sigma\big) \Rightarrow \mathrm{e}^{\mathrm{e}^{-x}}\mathrm{d}x, \tag{1.12}$$

where $\tau = \sqrt{2\log N}$ and $\sigma = \tau(1 + o(1))$.
Let us consider $x_{\max} \in [N]$ such that $v_{x_{\max}} = \max_{y\in[N]} v_y$. The strategy is to study the neighborhood of the vertex x.

Notation. We equip the graph $\mathbb{G}$ with the graph distance $d(\cdot,\cdot)$. For $x \in [N]$ and $i \in \mathbb{N}$, we define the sphere and the ball of radius i around x as

$$S_i(x) := \{y \in [N] : d(x,y) = i\}, \qquad B_i(x) := \{y \in [N] : d(x,y) \le i\}.$$

For $T \subset [N]$, we denote by $\mathbb{G}|_T$ the graph restricted to the vertex set T,

$$\mathbb{G}|_T := \big(T, \{(x,y) \in E(\mathbb{G}) : x, y \in T\}\big)$$

For $M \in \mathbb{R}^{N\times N}$ and $T \subset [N]$, we denote by $M|_T$ the $|T|$-by-$|T|$ sub-matrix

$$\big(M|_T\big)_{xy} = M_{xy}, \quad x, y \in T.$$

For $d \le N^{1/3-\varepsilon}$, $\varepsilon > 0$, it is possible by cutting a small number of edges (typically no more than $O(1)$), to remove cycles and large vertices (i.e. $y \in [N]$ such that $v_y \ge \sqrt{1.5\log N}$) from $B_1(x_{\max})$. After this *pruning* procedure, $\mathbb{G}|_{B_1(x_{\max})}$ is a tree with one root vertex $v_x \asymp \sqrt{2\log N}$ and $\max_{y\in B_1(x)} |v_y| = O(\sqrt{\log N})$. Actually, we will show that we can do this procedure *simultaneously* for all vertices $x \in \mathcal{V}$, where $\mathcal{V} := \{x \in [N] : v_x \ge \max_{y\in[N]} v_y - \sqrt{\tau \log N}\}$, for some small enough constant $\tau > 0$ (see Figure).

The strategy is then broken into three parts

1. Section 3.5 : Compute the largest eigenvalue $\lambda(x)$ of $\underline{L}|_{B_1(x)}$, for $x \in \mathcal{V}$. Show that $\lambda(x) = v_x + \frac{1}{v_x} + \varepsilon_x$, $\varepsilon_x = O\big((\log N)^{-1}\big)$.
2. Section 3.2 : Use a block diagonal approximation of $\underline{L}$ to show that the point processes

$$\Phi := \sum_{\lambda\in\operatorname{Spec}\underline{L}} \delta_\lambda, \qquad \text{and} \qquad \Psi := \sum_{x : v_x \sqrt{(2-\tau)\log N}} \delta_{v(x)+\frac{1}{\sigma}+\varepsilon_x}, \tag{1.13}$$

agree on the interval $\big[\sigma - \sqrt{\tau\log N}, +\infty\big)$, for some error terms $\varepsilon_x = o\big((\log N)^{-1/2}\big)$.

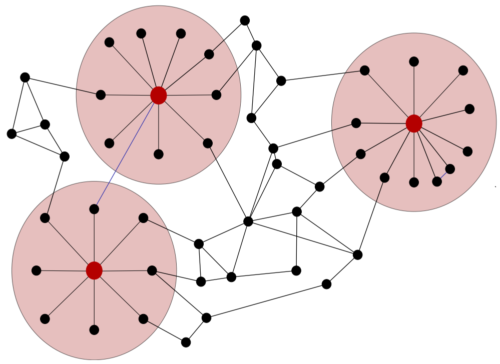


Figure 1.6: Illustration of the pruning procedure on the graph from Figure 1.1. The vertices drawn in red represent the large degree vertices and the colored edges are the ones removed during the pruning procedure.

3. Section 3.7 : Show that the law of the variables $\left(v_x + \frac{1}{v_x} : x \in \mathcal{V}\right)$ factorize asymptotically and that the process Ψ converges to a Poisson Point Process with intensity measure $\mathrm{e}^{-x}\mathrm{d}x$.

Perturbation theory

Let V be a Hermitian matrix and H a "small" Hermitian matrix. If everything is known about the spectrum of V, what can we say about the spectrum of $V+V$? This is a central question in random matrix theory but more generally it is a very fruitful source of exercises for quantum mechanics classes (see [44, Chapter 5]). While the tools used are elementary (see for instance [46, Chapter 1]), it is rarely taught in mathematics class. There are many ways to go about perturbation theory, perhaps the most historically relevant are the so-called Rayleigh-Schrödinger coefficients.
In our setup, we need to understand the top eigenvalue of the Laplacian matrix of a rooted tree, whose root vertex has a very large rescaled degree. Let us denote by x_* the root vertex and $B_r(x)$, $r \in \mathbb{N}$ the tree Setting V to be the diagonal matrix of the rescaled degrees and H the adjacency matrix, we can treat that question as a perturbation problem. An adaption of an argument found in [35, Chapter 4] allows to give an approximate value for the maximal eigenvalue of $V+H$ by computing all possible cycles in the graph $\mathbb{G}$ that starts at x_*. This is a self-contained argument that is elaborated in Proposition **??**. The argument relies only on the existence of a spectral gap (i.e that $\lambda_2(V) \leq \lambda_1(V) - 2\|H\|$) and on Cauchy's integral theorem. We believe it can be used in other contexts. An illustration is proposed in Figure 1.7.

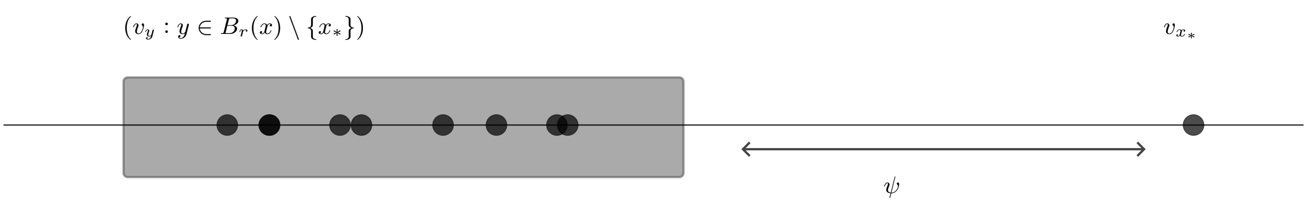


Figure 1.7: Computing the maximal eigenvalue of $\underline{L}|_{B_r(x_*)}$ with perturbation theory. Here x_* is supposed to be a large degree vertex surrounded by vertices with *average* degree, i.e. $v_y = O(1) \ll v_{x_*}$.

Remark. This perturbation argument works best when we can consider cycles of length greater than 4. This is however not always possible if we consider $d \geq N^{1/6}$. To fill the gap between $N^{1/6}$ and $N^{1/3}$ we use an alternative technique based on the Courant-Fisher principle (min-max characterization of the eigenvalues).

An adaptation of an argument from [35] allows us to elegantly express the eigenvalue of L

Rigidity of Poisson statistics

A good analysis of the eigenvalue process should allow us to distinguish the top eigenvalues one from the other. As explained previously, for regime $d \gg \log N$, we can do this by matching the maximal eigenvalues with a function of the maximal degree vertices. As d becomes of order $\log N$ however, it is possible to see more than one vertex of maximal degree. Let us consider independent Poisson variables of parameter d as a toy model for the eigenvalue process at the edge.
Let $(Y_i)_{i=1}^N$ be i.i.d. random variables with distribution $\mathcal{P}_d$, $i = 1, \dots, N$. Then if $d \gg \log N$ the maximum of the Y_i is with high probability unique and at the edge the rescaled Y_is form a
PPP. If $d \asymp \log N$ the distribution of $\max_i Y_i$ has bounded support and with positive probability there are many $j \in [N]$ such that $Y_j = \max_i Y_i$. Finally if $d \ll \log N$, the distribution of $\max_i Y_i$ is concentrated on 1 or 2 points and almost surely there are many $j \in [N]$ such that $Y_j = \max_i Y_i$. This emerging *rigidity* in the distribution of the extremes of Y_i is an adversarial mechanism when we want to distinguish top eigenvalues in critical regimes. See also [11, Remark 4.14].
To circumvent this obstacle, we will do a finer perturbation analysis on $\underline{L}|_{B_5(x)}$, where x is again a high-degree vertex. We then show that on the right half-infinite interval the point processes

$$\Phi := \sum_{\lambda \in \operatorname{Spec} \underline{L}} \delta_\lambda, \qquad \text{and} \qquad \Psi := \sum_{x \text{ large degree vertex}} \delta_{\Lambda(\alpha_x, \beta_x) + \varepsilon_x}, \tag{1.14}$$

agree. Here $\Lambda(\cdot,\cdot)$ is a function of $\alpha_x := \frac{D_x}{d}$ and $\beta := \frac{|S_2(x)|}{|S_1(x)|d} - 1$. The difference between the (1.13) and (1.14) is illustrated in Figure 1.8

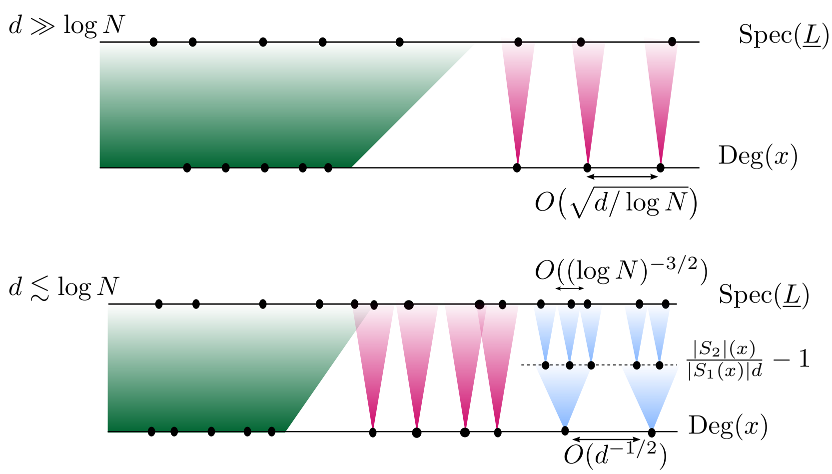


Figure 1.8: Illustration of the correspondence between high-degree vertices and maximal eigenvalues. In the supercritical regime, the maximal degrees are with high probability distinct and fluctuate, like a sample of N independent normal variables, on a scale $(\log N)^{-1/2}$. In the critical, respectively in the subcritical regimes, the maximal degrees can accumulate, respectively accumulate with high probability, on a few values. Then an extra layer of information is needed to distinguish the eigenvalues generated by $\underline{L}|_{B_r(x)}$, for $x \in \mathcal{W}^+$, the set of high-degree vertices.

1.6 Smallest eigenvalues and spectral gap

The left edge of the spectrum (c.f. Figure 1.2) looks similar to the right edge. This impression is vindicated by our first theorem concerning the smallest eigenvalues of $\underline{L}$ which is the almost symmetric counterpart to Theorem B.

Theorem C. *Let $\varepsilon > 0$. If $\frac{1}{2}\log N + (\log N)^{\varepsilon} \le d \le N^{\frac{1}{3}-\varepsilon}$, the point process of the $\mathcal{K}$ (non-trivial) eigenvalues is asymptotically close to a sequence of PPP . Furthermore, the top eigenvectors are localized around*

1. *minimal degree vertices when $d \ge \log N - \log\log N$;*
2. *leaves with minimal sphere of radius three when $d \le \log N - \log\log N$.*

Too many leaves

Let us denote by $\mathcal{L} := \{x \in [N] : D_x = 1\}$, the set of degree-one vertices Just like the statement of Theorem C is similar to the one of Theorem B, the ideas and techniques used in the proof are similar. There is however one key difference when $d = (1-c)\log N$, for $c \in (0, 1/2)$. In that case the number of leaves becomes polynomial: indeed adapting (1.6) we see that

$$\mathbb{E}\big[|\mathcal{L}|\big] = N\mathrm{e}^{-d}d = N^c(1+o(1)), \qquad \mathcal{L} := \{x \in [N] : D_x = 1\}$$

Because of this polynomial accumulation of vertices, the extreme value statistics of the variables $\big(\beta_x : x \in \mathcal{L}\big)$ undergo the same transition as the extreme value statistics of the maximum of n i.i.d. $\mathcal{P}_d$ variables. We need to proceed carefully and expand to approximate the smallest eigenvalues of $\underline{L}|_{B_10(x)}$, $x \in \mathcal{L}$, to the third order as

$$\lambda_1(\underline{L}|_{B_{10}(x)}) = \Lambda^{\mathcal{L}}\big(D_x, |S_2(x)|, |S_3(x)|\big) + \varepsilon_x = \Lambda^{\mathcal{L}}\big(1, |S_2(x)|, |S_3(x)|\big) + \varepsilon_x,$$

for ε_x sufficiently small.

An additional difficulty arises when $d = \frac{1}{2}(1+\varepsilon)\log N$ and $\varepsilon = o(1)$. In that case it is no longer possible to guarantee that the neighborhood of leaves is free of other small degree vertices and the perturbation argument (as illustrated in Figure 1.7) becomes increasingly difficult. We illustrate the difficulty by pushing our result to regimes $d \ge \frac{1}{2}\log N + (\log N)^{\varepsilon}$, for $\varepsilon > 0$ constant. We believe the optimal regime should be $d \ge \frac{1}{2}\log N + C\log\log N$, for some large enough constant $C \ge 0$.

What is smaller than a leaf? Maximal trees and spectral gap

A key assumption in any perturbation analysis argument is to have a *spectral gap* (c.f. Figure 1.7). However for $d \le \frac{1}{2}\log N$, there is with high probability a positive number of regions in $\mathbb{G}$ where small degree vertices are neighbors one of the other. In that case, it is impossible to perform perturbation analysis of the diagonal matrix $\underline{D}$. On the other hand, first-order perturbation analysis suggests that the eigenvalue generated by such a configuration should be of order $-\sqrt{d}+\frac{1}{2\sqrt{d}}(1+o(1))$, which is much different than the $-\sqrt{d}+\frac{1}{\sqrt{d}}(1+o(1))$ obtained from perturbation analysis of a rooted tree with a rooted vertex of degree 1. A similar analysis for a group of $t \in \mathbb{N}^*$, $t \ge 2$, vertices of degree $O(1)$ attached together suggests that the eigenvalue should be of the order $-\sqrt{d} + \frac{1}{t\sqrt{d}}(1+o(1))$. This observation suggests that the correct geometric shape to analyze is not any more isolated vertices of small-degree, but rather trees. Let $\mathbb{T}_t$, $t \in \mathbb{N}^*$, be the set of finite trees with vertices labeled by $[t]$. By Cayley's theorem, this set has cardinality t^{t-2}. In Chapter 4, we construct a function $\gamma_* : \mathbb{N}^* \to \mathbb{R}_{>0}$ such that $\gamma_*(t)$ is an implicit function of $\mathbb{T}_t$ and relate it to the smallest eigenvalues of $\underline{L}$.

Theorem D. *For $\frac{1}{t+1}\log N \ll d \ll \frac{1}{t}\log N$, $t \in \mathbb{N}^*$, with high probability, we have*

$$\lambda_2(\underline{L}) = -\sqrt{d} + \frac{\gamma_*(t)}{\sqrt{d}} + O\big(d^{-1}\big).$$

Although we do not prove eigenvector localization and convergence towards a P.P.P. in the present work, those results are believed to hold with a very high degree of confidence (and also with high probability). Indeed the error bounds obtained in the proof of Theorem D can be actually quite easily refined and the

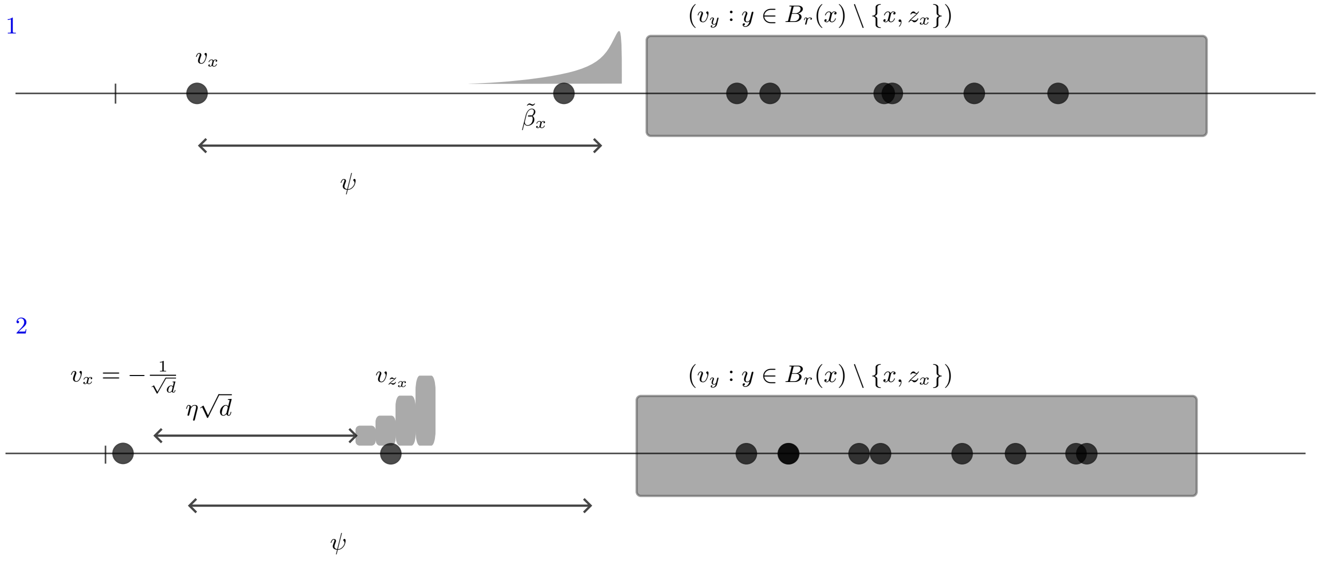


Figure 1.9: Illustration of the perturbation argument for points 1 (above) and 2 (below) of Theorem C. In the first case, the perturbation theory works well and the eigenvalues are well-described by a function of v_x and $\tilde{\beta}_x := \frac{|S_2(x)|-|S_1(x)|d}{\sqrt{d}}$ which has a continuous distribution. On the other hand, as $d = \eta \log N$, $\eta \in (1/2, 1)$, the smallest degree vertices are all leaves and the statistics of $\tilde{\beta} = v_{z_x} - d^{-1/2}$ become discrete (rigidity phenomenon of the maximum of n $\mathcal{P}_{\log n}$ variables). In that case the value of D_x (which is 1 for leaves) and of $|S_1| = D_{z_x}$ is not enough to distinguish the smallest eigenvalues. Finally, observe that the interval between v_x and v_{z_x} closes as $\eta \to 0$. This explains the limitation in the hypotheses of the theorem.

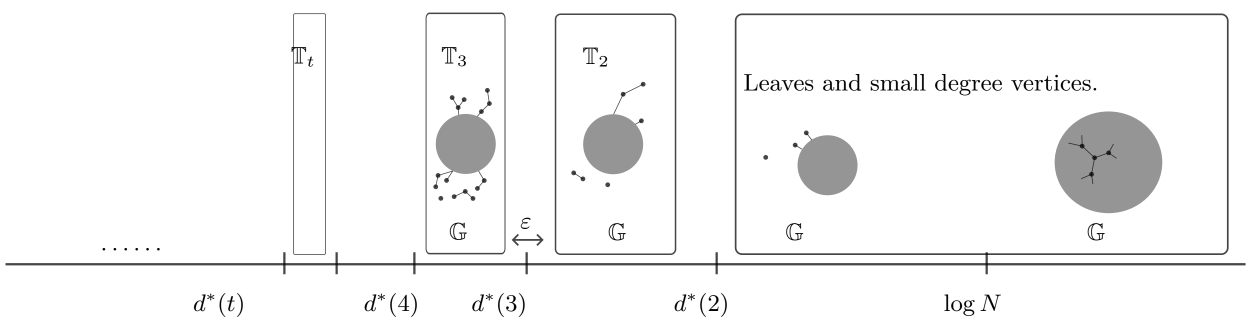


Figure 1.10: Illustration of Theorem D

pipeline for Theorem B can be adapted in a straightforward way to handle the point process of minimal eigenvalues in the regimes covered by Theorem D.

The mechanism underlying the proof of Theorems B and C rely on the analysis of the matrix $H = H_0 + H'$, where the largest eigenvalue of H_0 is separated by a spectral gap ψ from the rest of Spec H_0 and where H' is a small matrix, in the sense that $\|H'\| \ll \psi$. The mechanics at work in Theorem D are different: we consider the spectrum of rank one perturbation of a block diagonal matrix

$$H = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix} + \theta \mathbf{v}\mathbf{v}^*, \tag{1.15}$$

where $\theta \in \mathbb{R}$ and $\mathbf{v}$ is a real vector.

Now imagine the Erdős-Rényi is connected and admits a subset of vertices $T \subseteq [N]$ such that

$$|T| = t, \qquad \max_{x \in T} D_x \leq 2t, \qquad \min_{x \notin T} D_x \geq d - O(\sqrt{d}).$$

Suppose in addition that there is exactly one edge that links T and T^c, namely $e = (x, z)$, $x \in T$ and $z \in [N] \setminus T$. In that case, the matrix $\underline{L}$ can be written as the matrix H of (1.15) with H_1 is the (rescaled) Laplacian matrix of a tree $T \in \mathbb{T}$, H_2 is the rest of the graph, i.e. $\underline{L}|_{T^c}$, $\theta = \frac{1}{\sqrt{d}}$ and $\mathbf{v} = \mathbf{1}_x - \mathbf{1}_z$, for $x \in T$ and $z \in T^c$ is used to "link" those two disjoint connected components. H_1 and H_2 are both (rescaled) Laplacian matrices of graphs and therefore each generates a trivial eigenvalue $-\sqrt{d}$ (recall the rescaling of (1.7)).

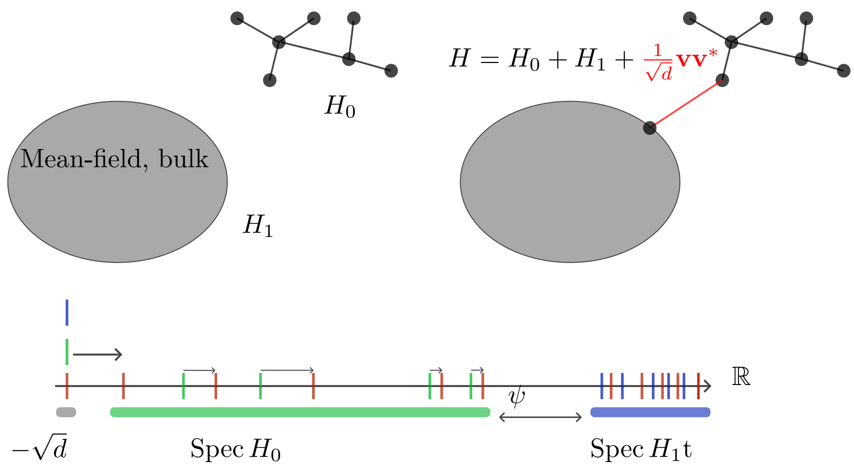


Figure 1.11: Illustration of rank-one perturbation theory. In blue, we denote the eigenvalues of H_1, in red the ones of H_0. In our setup we suppose that the spectrums of H_0 and H_1 are separated by a spectral gap ψ (this is similar to the arguments illustrated in Figure 1.7). However here many "extreme" eigenvalues are perturbed (if the tree is of size $t \in \mathbb{N}^*$, then there are t such extreme eigenvalues). We see that analyzing the spectral gap of H, i.e. the distance between $\lambda_1(H) = -\sqrt{d}$ and $\lambda_2(H)$ is equivalent to understanding how the trivial eigenvalue $\lambda_1(H_0) = -\sqrt{d}$ is perturbed when H_0 is "anchored" to H_1.

An interesting corollary of Theorems C and D is the following characterization of the *spectral gap* of $\underline{L}$ restricted to the giant connected component.

Notation. We denote by $\mathbb{G}_{cc}$ the connected component of $\mathbb{G}$ with the highest number of vertices.

Corollary E (Spectral gap of $\underline{L}|_{\mathbb{G}_{cc}}$)**.** *Let $\varepsilon > 0$ and $\Lambda := \lambda_2(\underline{L}|_{\mathbb{G}_{cc}}) + \sqrt{d}$ be the spectral gap of the rescaled Laplacian matrix. Then the following holds*

(i) For $d \geq \log N - \log\log N$, the spectral gap is given by

$$\frac{\min_{x \in \mathbb{G}_{cc}} D_x}{\sqrt{d}} + O(d^{-1}).$$

(ii) For $\frac{1}{2}\log N + (\log N)^{\varepsilon} \leq d \leq \log N - \log\log N$, the spectral gap is given by

$$\frac{1}{\sqrt{d}} - \frac{\min_{x \in \mathcal{L}} D_{z_x}}{d} + O(d^{-2}).$$

(iii) For $\frac{1+\varepsilon}{t+1}\log N \leq d \leq \frac{1-\varepsilon}{t}\log N$, $t \in \mathbb{N}^*$, the spectral gap is given by

$$\frac{\textit{Universal function of finite trees of size } t}{\sqrt{d}} + O(d^{-1})$$

1.7 Outlook

What did we learn?

There are many ways to interpret the results obtained in this work. The first way is to view them as a collection of mathematical results the relevance of which lies in the hypotheses under which the various theorems hold and the techniques we use to prove them. For instance, as pointed out in Figure 1.5, proving a local law for the Laplacian of the Erdős-Rényi graph is not new but the regimes of d under which such results are proved is sensibly better than previous work in this domain. Similarly, our local law falls in the study of deformed Wigner matrices and, while the methods we use a very different from the ones found in [37], [38] and [39], Theorem A can be seen as a companion result of the ones found in those papers.

Another way is to consider the physical meaning of our results. In [50], Wigner used random matrices to describe the energy levels of particles in heavy atoms. Wigner managed to describe the distribution of the energy levels of the atoms and show that it had the shape of a semi-circle, but, to our knowledge, he did not make mention of the eigenvectors. In [8], Anderson proposed a way to describe the movement of electrons on a lattice with random potential by considering an operator $\Delta + V$, where Δ is the hopping term and V the potential. In this case, the eigenvectors had an explicit interpretation as the location of the electron's wave function on the grid. In their famous paper [27], Fröhlich and Spencer proved rigorously that for a large disorder, the top eigenvectors of the model are localized (see [2] for an overview of results and techniques). This result has a direct physical interpretation as it tells that if the level of impurity is high enough, the electron is trapped in some region of the grid (insulator, localized eigenvector, pure point spectrum) and cannot travel in the medium (conductor, delocalized eigenvector, absolutely continuous spectrum). In this regard, Theorems B and C are an analog of their result for the Laplacian: as mentioned earlier, the Laplacian is the quantum Hamiltonian of a particle moving on the graph $\mathbb{G}$. Proving localization of eigenvectors around large-, respectively small-degree vertices (both analog to high disorder) means that for those level of energy, the particle is trapped in the regions surrounding those vertices. This was already proved for the adjacency matrix of $\mathbb{G}$ in [7]. While we do not prove eigenvector localization around trees, we strongly believe that Theorem D can be extended to eigenvector localization exactly as its two counterparts.

A third way to understand our results is through the lens of information. As we mentioned earlier, spectral analysis of graphs is a very popular tool in computer science, statistics and data analysis. Let us remind the reader of the question we asked at the beginning of the chapter.

Can meaningful information be extracted from the eigenvalues and eigenvectors of $\underline{L}$?

In his famous paper [34], Mark Kac asked whether it was possible to hear the shape of a drum. Or in other words, what does the spectrum tell us about our system? Understanding this question is important because we want to know what predictive powers our spectral algorithms can have. In that regard, the main conclusion of this thesis is the following: spectral analysis of the bulk and the edge eigenvectors of L yields no interesting information. Let us elaborate. In Theorem A, we claim that a local law holds down to the regime $\sqrt{\log N}$ in the bulk. This means that the eigenvectors are completely delocalized. Local laws are the typical result for *mean-field* models which by definition have no interesting structure. In Theorems B, C and D, we claim that the extreme eigenvalues of the matrix $\underline{L}$ are determined by *local configurations*, such as large degree vertices, small degree vertices and embedded trees. In particular, they do not reflect any general effect of the system. The only information that was given is the characterization of the spectral gap (Corollary E) of $\underline{L}$ down to regimes $\frac{1}{C}\log N$, for $C = O(1)$. In particular, our results do not give ways to identify anomalies or communities in graphs, with the exception of finding trees that are embedded in the graph.

Non-ergodic delocalization

An interesting direction for future analysis would be to understand the partially delocalized phases of the spectrum of $\underline{L}$. As is evident from Figure 1.2 (purple lines on both sides of the bulk), such phases exist in any regime of d and seem to be characteristic of regions of the spectrum that correspond to a polynomial density of states. A partial delocalization (also known as non-ergodic delocalization) mechanism seems to be at play in the spectral regions located slightly above $E = \sqrt{\log N}$, where

$$\big|\{x \in [N] : |v_x - E| \le 1\} = N^{-\frac{1}{2}+c}(1 + o(1)), \quad c \in (0, 1).$$

While we believe our explanation of the partial localization phenomenon (lower bound on $\|\mathbf{w}\|_\infty$), given in Chapter 2, could be made rigorous without much effort (as it was done for instance in [5, Theorem 1.2]), we do not know how to prove partial delocalization (upper bound on $\|\mathbf{w}\|_\infty$). There are only a few known techniques that allow for a rigorous derivation of partial delocalization estimates. Indeed, apart from the local law framework, the proof of delocalization results is extremely difficult and remains a big open question in the field (Anderson extended state conjecture). Some heuristics have been developed for the Erdős-Rényi graph [47] and general physical models and rigorous mathematical proofs have been shown for specific cases ([1], [15], [49]).

Another way to study non-ergodic delocalization would be to modify the model by adding a small mean-field component. This is in line with the idea of [40]. For instance does the localization result of Theorem B remain valid if we consider

$$\underline{L}_1(\alpha) := (1-\alpha)\underline{L} + \alpha\frac{1}{N}\mathbf{1}_{[N]}\mathbf{1}_{[N]}^*, \quad \underline{L}_2(\alpha) := (1-\alpha)\underline{L} + \alpha \cdot \mathrm{GOE}, \qquad \alpha \in [0, 1],$$

with α possibly dependent on N? Clearly as $\alpha \to 1$, the matrix $\underline{L}_1(\alpha)$ moves closer to a rank-one matrix while $\underline{L}_2(\alpha)$ becomes a GOE. As Google's algorithm shows, we could expect a partial delocalization on the most relevant vertices.

Very large trees

The behavior of the spectrum for regimes $d = o(1)\log N$ would be a natural way to expand Theorem D. Would the localization conjecture remain valid? Could trees of different sizes simultaneously contribute to the spectral gap?

As explained in Chapter 2, when d approaches $\sqrt{\log N}$, large stars (resonant and non-resonant) start to contribute significantly to the bulk of the spectrum. As is evident from the computations in Chapter **??**, as long as embedded trees are on finite size, their contribution to the spectral gap can be formally understood as

$$\lambda(\text{spectral gap}) = -\sqrt{d} + \frac{\text{Universal function of } L(T)}{\sqrt{d}} + \frac{\mathcal{N}(0,1)}{d^{3/2}}. \tag{1.16}$$

Because the spectrum of $L(T)$ lives at a scale of $O(1)$ and the fluctuation in the neighborhood in the graph of T fluctuate on the scale $d^{-3/2}$, as long as $|T| = O(1)$ the scales in (1.16) allow to distinguish between the contribution of different trees. It is not at all evident that, as much larger trees begin to appear, which will happen as soon as $d = o(\log N)$, this relation still holds. In particular, we do not know if the spectral gap of $L(T)$ is still generated by maximal trees. It might be that smaller trees with large neighborhood fluctuations yield a smaller spectral gap than larger tree with small neighborhood fluctuations. We expect the spectral gap to close and the rate at which it does would be an interesting question.

Very dense regimes

As explained above, the proof of Theorem B relies on a sparsity property of the graph that allows us to separate the neighborhood around high-degree vertices (see Figure 1.6). For $d = N^\varepsilon$, $\varepsilon > 0$, the radius of $\mathbb{G}$ (the maximal distance between two points in the graph) is, with high probability $1/\varepsilon$. Therefore, as ε becomes larger than 1/3, it is no longer possible to do so and our proof breaks down (see Figure 1.5). However, we believe that the law of the extreme eigenvalues remains the same, and our conjecture is supported by results about dense Laplacian matrices (see [20]). However, we believe new techniques are required since, for the adjacency matrix, the regime $N^{1/3}$ is also the boundary between two kinds of behaviors for the extreme eigenvalues.

Chapter 2

Delocalization in the bulk

In this chapter, we prove two local laws for the Laplacian matrix of the Erdős-Rényi graph. The result is valid for regimes as low as $C\sqrt{\log N}$. In particular, we obtain that the entries of the Green function $G(z)$ remain uniformly bounded down to the optimal scale $\eta \gg N^{-1}$ on any compact interval. We explain how these results allow us to prove eigenvector delocalization in the bulk and why the result is optimal.

2.1 Main results

In this chapter, we study the spectrum of the Laplacian matrix of the Erdős-Rényi graph. Recalling the adjacency matrix A, we define

$$H := \frac{1}{\sqrt{\gamma}}(A - \mathbb{E}[A]), \quad \gamma := d\Big(1 - \frac{d}{N}\Big). \tag{2.1}$$

The matrix H is Hermitian and the entries $(H_{xy} \colon 1 \le x < y \le N)$ form a family of independent centered random variables satisfying

$$\mathbb{E}|H_{xy}|^2 = \frac{1}{N}, \quad |H_{xy}| \le K d^{-1/2}, \tag{2.2}$$

for some $K \ge 0$, as soon as $d \le N/2$.
We define the diagonal matrix V and the rank one matrix R as

$$V := \operatorname{Diag}\big(v_1, \dots, v_N\big), \quad R = \frac{d}{\sqrt{\gamma}}\mathbf{e}\mathbf{e}^*, \tag{2.3}$$

where $\mathbf{e} := \frac{1}{\sqrt{N}}\mathbf{1}_{[N]}$ and

$$v_x := \sum_y H_{xy}. \tag{2.4}$$

We will prove a local law for the matrix

$$M := V - H - R. \tag{2.5}$$

The matrix M is a linear transform of L defined in (1.2) since

$$M = \frac{1}{\sqrt{\gamma}}\Big(L - d - \frac{d}{N}\Big). \tag{2.6}$$

We study M instead of L to comply with the conventions of the literature on local laws. In particular, matrices of the form (2.5) are instances of so-called *deformed Wigner matrices* which have been studied in the literature, for instance in [30], [5] and [37].

Convention 2.1. *For now on every quantity depends implicitly on N unless we explicitly define it as a constance.*

Let $\Xi := \Xi_{N,\nu}$ be a family of events parametrized by $N \in \mathbb{N}$ and $\nu > 0$. We say that Ξ holds with very high probability if for every $\nu > 0$ there exists a constant $\mathcal{C}_\nu$ such that

$$\mathbb{P}\big(\Xi_{N,\nu}\big) \geq 1 - \mathcal{C}_\nu N^{-\nu}$$

for all $N \in \mathbb{N}$.
In particular, the estimate $X \leq \mathcal{C}Y$ with very high probability means that, for each $\nu > 0$, there are constants $\mathcal{C}_\nu, c_\nu > 0$ depending on ν such that

$$\mathbb{P}\big(|X| \leq \mathcal{C}_\nu Y\big) \geq 1 - c_\nu N^{-\nu}.$$

We sometimes abbreviate that statement by $X = \mathcal{O}(Y)$.
Note that $X \leq Y$ with very high probability means

$$\mathbb{P}\big(|X| \leq Y\big) \geq 1 - c_\nu N^{-\nu}$$

for some $c_\nu \geq 0$.

For constants $\kappa \in (0,1)$ and $R \geq 1$ we define the spectral domain

$$\mathbb{S}_{\kappa,R} := \{z \in \mathbb{C}\colon\ |\operatorname{Re} z| \leq R,\ N^{-1+\kappa} \leq \operatorname{Im} z \leq 2\}. \tag{2.7}$$

We prove a local law on the entries of the Green function and on the Stieltjes transform g of the empirical spectral measure of M,

$$G(z) := (M - z)^{-1}, \quad g(z) := \frac{1}{N} \sum_{\lambda \in \operatorname{Spec} M} \frac{1}{\lambda - z} = \frac{1}{N} \operatorname{Tr} G(z). \tag{2.8}$$

We always assume that z lies in the complex upper-half plane $\mathbb{H}$. The limiting behavior of G and g is governed by the following deterministic quantities.

Definition 2.2 (Quadratic vector equation)**.** Let $v_x \in \mathbb{R}$, $x \in [N]$, $v = (v_x)_{x\in[N]} \in \mathbb{R}^N$. We define the vector $m := m^v = (m_x^v)_{x\in[N]} \in \mathbb{H}^N$ to be the solution of the system of equations

$$\frac{1}{m_x} = v_x - z - \frac{1}{N} \sum_{y \in [N]} m_y, \qquad x \in [N], \tag{2.9}$$

for $z \in \mathbb{H}$. We also introduce

$$\overline{m} := \frac{1}{N} \sum_{x \in [N]} m_x. \tag{2.10}$$

As was observed in [37], g is the additive-free convolution of the semi-circle law $\mu_{\mathbf{sc}}$, and μ_v, the empirical distribution of the entries of V.

Lemma 2.3 (Existence and uniqueness of (2.9))**.** *For each $z \in \mathbb{H}$ and $v \in \mathbb{R}^N$, the system of equations (2.9) has a unique solution m in $\mathbb{H}$.*

The proof of Lemma 2.3 is a classical result that can be found for instance in [4]. In Section 2.9, we study in more detail the properties m when the v_x are defined as in (2.4).

Theorem 2.4. *Let M be as in (2.5) and m be the solution to (2.9) with v as defined (2.4). Then with high probability, there exists $D \geq 0$ such that if $D\sqrt{\log N} \leq d \leq (\log N)^{3/2}$, then*

$$\max_{x,y} |G_{xy}(z)| = O(1), \quad \max_{x,y} \left| G_{xy}(z) - \frac{\delta_{xy}}{v_x - z - \frac{1}{d}\sum_y A_{xy} m_y} \right| = O\left(\frac{\log N}{d^2}\right)^{1/3}, \quad \forall z \in \mathbb{S}_{\kappa,R}, \tag{2.11}$$

holds with probability $1 - O\big(d^{-1}\big)$.

Definition 2.5 (Free convolution)**.** The Stieltjes transform $\mathrm{m_{fc}}$ is defined as the solution to the following functional equation

$$\mathrm{m_{fc}}(z) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{\mathrm{e}^{-x^2/2}}{x - z - \mathrm{m_{fc}}(z)} \mathrm{d}x, \quad z \in \mathbb{H}.$$

For $d \gg \log N$, the last term of the denominator in (2.11), can be shown to converge $\mathrm{m_{fc}}$. We have

$$\Big|\frac{1}{d}\sum_y A_{xy} m_y - \mathrm{m_{fc}}(z)\Big| \leq \Big|\frac{D_x}{d} - 1\Big| \max_y |m_y| + |\overline{m}(z) - \mathrm{m_{fc}}(z)|.$$

Using Bennett's inequality, Lemma B.3, we see that $\max_x \big|\frac{D_x}{d} - 1\big| = \mathcal{O}(\sqrt{\log N/d})$. Moreover, it can be proved that $\overline{m} = \mathrm{m_{fc}} + \mathcal{O}(\sqrt{\log N/d})$, see for instance Lemma 2.34. We can conclude that for $d \gg \log N$,

$$G_{xx}(z) = \frac{1}{v_x - z - \mathrm{m_{fc}}(z)} + o(1).$$

This was known for $d \geq N^{\varepsilon}$, $\varepsilon > 0$, from [30]. The next theorem closes the gap between their results and Theorem 2.4.

Theorem 2.6. *For* $(\log N)^{1+\kappa} \leq d \leq N^{\kappa/12}$,

$$\max_{x,y} \Big|G_{xy}(z) - \frac{\delta_{xy}}{v_x - z - \mathrm{m_{fc}}(z)}\Big| = \mathcal{O}\bigg(\sqrt{\frac{\log N}{d}}\bigg), \quad \forall z \in \mathbb{S}_{\kappa,R}. \tag{2.12}$$

Remark 2.7 (Extension to generic sparse Laplacian matrices)**.** It is interesting to consider matrices M as defined in (2.5) but with more general conditions on V, H and R. The matrix H is typically viewed as some Wigner matrix and the conditions given by (2.2) are already quite general. A possible extension could be to modify the law of V, making it independent of H for instance. or to replace the factor $d/\sqrt{\gamma}$ in the definition of R by some $0 \leq f \ll N$ is some rank one matrix (see for instance [5, (4.1)].
As will become apparent from its proof, Theorem 2.6 is easily extended to this setup. However, the techniques used in the proof of Theorem 2.4 require some control on the relationship between the quantities $\alpha_x := \sum_y H_{xy}^2$, and v_x. For instance, in our proof, we need to know that if $|\alpha_x - 1| > c > 0$ for some constant $c > 0$, then $|v_x| \gg 1$. This is not true in general, as the simple example where H_{xy} are i.i.d. with uniform probability distribution on $\Big[-\frac{\sqrt{12}}{N^{1/2}}, \frac{\sqrt{12}}{N^{1/2}}\Big]$ shows. However, we believe many natural models, for instance, weighted random graphs, should be amenable to similar proofs.

Consequences and limits of the local law

Once a local law on some random matrix ensemble has been proved, many useful consequences can be drawn. A very important is the delocalization of the eigenvectors associated with the eigenvalues present in the interval of the local law. In this section, we use the following convention. If $\lambda \in \mathbb{R}$ is defined as the eigenvalue of some Hermitian matrix H, then $\mathbf{w}_\lambda$ denotes the associated eigenvector.

Corollary 2.8 (Eigenvector delocalization in the bulk)**.** *Let* $\kappa > 0$ *and* $R \geq 1$. *Let us consider the Laplacian matrix of the Erdős-Rényi graph* L *and write* $\mathbf{w}_\mu$, $\mu \in \operatorname{Spec} L$, *the eigenvector associated to the eigenvalue* μ. *There exists* $D \geq 0$ *such that if* $D\sqrt{\log N} \leq d \leq N^{\kappa/12}$, *then*

$$\big\|\mathbf{w}_\mu\big\|_\infty = O\Big(N^{\kappa-1}\Big), \qquad \forall \mu \in \big[d - R\sqrt{d}, d + R\sqrt{d}\big]$$

Proof. Let $\mu \in \operatorname{Spec} L$ such that $|\mu - d| \leq R\sqrt{d}$. By (2.6), we know that $\mathbf{w}_\mu$ is an eigenvector of M for the eigenvalue $\lambda = \frac{1}{\gamma}(\mu - d - d/N)$. Applying Theorems 2.4 and 2.6 with R as $R+1$, and choosing the random spectral parameter $z := \lambda + \mathrm{i}\eta$ with $\eta := N^{-1+\kappa}$, we find a constant $C \geq 0$, such that

$$C \geq \operatorname{Im} G_{kk}(z) = \sum_{\lambda' \in \operatorname{Spec}(M)} \frac{\eta}{(\lambda' - \lambda)^2 + \eta^2} |\mathbf{w}_{\lambda'}(k)|^2 \geq \frac{1}{\eta}|\mathbf{w}_\lambda(k)|^2 = N^{1-\kappa}|\mathbf{w}_\lambda(k)|^2.$$

This shows the claim. □

Localization on tuning forks and trees

Follow the construction in [5] using tuning forks. The same argument works here. We can even use a (longer) perturbation argument to show localization happens on big starts.As d gets closer to $\sqrt{\log N}$, isolated trees of size $\sqrt{d}$ will appear outside of the macroscopic connected component. Such trees can have eigenvalue of size $\sqrt{d}$ (for instance the star on $d-1$ vertices). However, restricting the spectral analysis to the macroscopic component does not solve this issue, since tuning forks of size d also appear with positive probability, giving rise to eigenstates with energy $\sqrt{d}$ localized on $2d$ vertices.

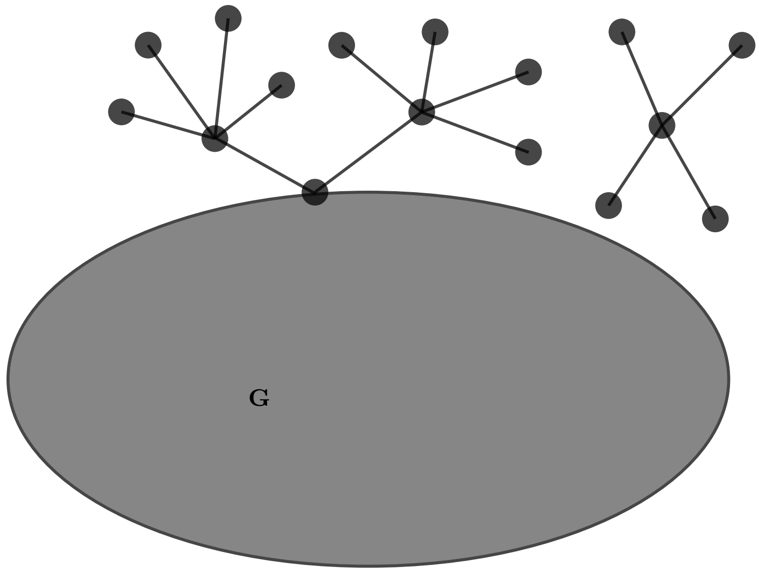


Figure 2.1: A tuning fork of size 2 times 5 and a star of size 5.

Partial localization outside of the bulk

In this subsection, we give a heuristic argument on why the conclusion of Corollary 2.8 cannot hold in regions of the spectrum where the density of states of the entries of V is $o(1)$. We know that this argument could be made rigorous by following the arguments of [5, Theorem 3.4] but for the sake of brevity we restrain from doing so.

The argument relies on the fact that M can be seen as a perturbation of the diagonal matrix V by the matrix H. The size of the perturbation is small compared to V. Indeed the entries of V can become very large (see Lemma 2.48) while we know (see Proposition 3.25) that with very high probability

$$\|H\| \leq 2 + \mathcal{C}\sqrt{\frac{\log N}{d}}. \tag{2.13}$$

Suppose $R \gg 1$ is an energy that depends on N such that the number of vertices with normalized degrees close to R is polynomially small, i.e.

$$\frac{1}{N}\big|\{x \in [N] : R - \phi \leq v_x \leq R + \phi\big|\} \asymp N^{-c/2}, \quad \sqrt{d}N^{-c/2} = o(1), \quad \big\|H\big\| = O(1), \tag{2.14}$$

for some constant $c > 0$ and some quantity ϕ, that may depend on N. Let us call $\mathcal{V} := \mathcal{V}(R, \phi)$ the subset of vertices that contribute to that set.

Remark 2.9 (Relevance of the assumptions made above)**.** Compare (2.14) with the analog condition in the bulk (2.109). Note that the entries of V can be morally thought of as identically distributed, weakly

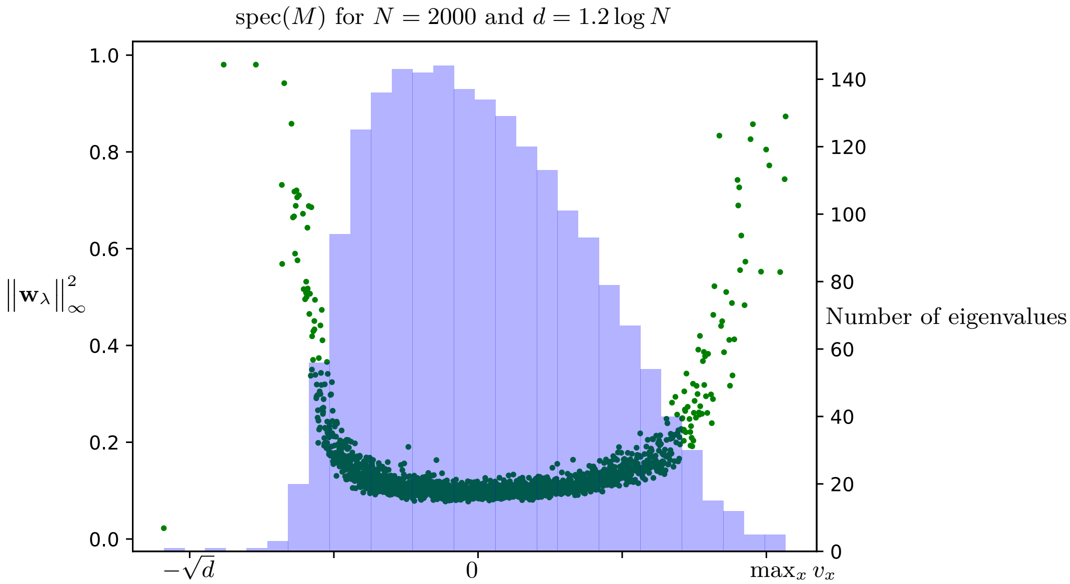


Figure 2.2: For the above simulation, we used a scatter plot $(\lambda, \|\mathbf{w}_\lambda\|_\infty^2)$, $\lambda \in \mathrm{Spec}(M)$ to illustrate the negative correlation between the density of states in some region of the spectrum and the delocalization of the eigenvectors in that region. The lower-left point corresponds to the trivial eigenvector $\frac{1}{\sqrt{N}}\mathbf{e}$. The bulk of the spectrum can be identified by the region where the green dots are the lowest and the density of states is high.

correlated, normal variable in the regime $d \gg \log N$. Therefore (2.14) is satisfied for $E = \sqrt{c \log N}$, $c \in (0, 2)$. The argument is the same for $R \ll -1$, as long as $R \gg -\sqrt{d}$ (remember that L is positive definite and thus $\mathrm{Spec}(\underline{L}) \subseteq [-\sqrt{d}, +\infty)$ by definition).
For $c > 0$ and $d = (\log N)^2$, all the conditions of (2.14) are satisfied.

Suppose $\lambda \in \mathrm{Spec}(\underline{L}) \cap \left[R - \frac{\phi}{2}, R + \frac{\phi}{2}\right]$ and $\mathbf{w}_\lambda$ is the corresponding eigenvector. We will show that

$$\left\|\mathbf{w}_\lambda|_{\mathcal{V}^c}\right\|_2^2 = o(1). \tag{2.15}$$

From (2.15), we will immediately be able to conclude that $\left\|\mathbf{w}_\lambda|_{\mathcal{V}}\right\|_2^2 > 1 - o(1)$ and thus, by Dirichlet principle,

$$\left\|\mathbf{w}_\lambda\right\|_\infty^2 \geq \frac{1}{2|\mathcal{V}|} \gtrsim \frac{1}{2} N^{-c/2} \gg N^{\kappa - 1},$$

for any $\kappa < 1 - c/2$.

Since the matrix M can be seen as a small perturbation of the diagonal matrix V, we could expect that if $\lambda \in \mathbb{R}$ were an eigenvalue of M close to R, then the eigenvector $\mathbf{w}_\lambda$ would be supported mainly on the set $\mathcal{V}$. Let us introduce the projection operators

$$\Pi := \sum_{x \in \mathcal{V}} \mathbf{1}_x \mathbf{1}_x^*, \qquad \overline{\Pi} = 1 - \Pi.$$

Projected on $\mathrm{Ran}(\Pi)$, the eigenvalue eigenvector equation for $(\lambda, \mathbf{w}_\lambda)$ becomes

$$0 = \overline{\Pi}(M - \lambda)\mathbf{w}_\lambda = \overline{\Pi}(M - \lambda)\big(\Pi + \overline{\Pi}\big)\mathbf{w}_\lambda = \big(\overline{\Pi} M \overline{\Pi} - \lambda\big)\overline{\Pi}\mathbf{w}_\lambda + \overline{\Pi} M \Pi \mathbf{w}_\lambda. \tag{2.16}$$

From (2.16), we deduce

$$\overline{\Pi}\mathbf{w}_\lambda = \frac{1}{\overline{\Pi}M\overline{\Pi} - \lambda}\overline{\Pi}M\Pi\mathbf{w}_\lambda. \tag{2.17}$$

By (2.3) and (2.14), we find that

$$\left\|\overline{\Pi}R\overline{\Pi}\right\| \le \frac{|\mathcal{V}|}{N}\sqrt{d} = o(1), \quad \left\|\overline{\Pi}H\overline{\Pi}\right\| \le C.$$

Since V is a diagonal matrix we see that

$$\operatorname{Spec}(V) = \{v_x : x \in [N]\}, \quad \operatorname{Spec}\big(\overline{\Pi}V\overline{\Pi}\big) = \{v_x : x \in \mathcal{V}\}.$$

We conclude that if $\phi \ge 2C$

$$\operatorname{Spec}\big(\overline{\Pi}M\overline{\Pi} - \lambda\big) \subseteq \mathbb{R} \setminus \big[R - \phi/3, R + \phi/3\big]. \tag{2.18}$$

Using the fact that $\mathbf{e}^*\mathbf{w}_\lambda = 0$ and that V is diagonal and R is a multiple of a projection, we find

$$\left\|\overline{\Pi}M\Pi\mathbf{w}_\lambda\right\| \le \left\|\overline{\Pi}H\Pi\mathbf{w}_\lambda\right\| \le C.$$

Thus, (2.16) yields

$$\left\|\overline{\Pi}_I\mathbf{w}_\lambda\right\| \le \frac{3C}{\phi} = o(1), \tag{2.19}$$

as soon as $\phi \gg 1$.

Outline of the proof of Theorem 2.4

The Laplacian matrix is an instance of so-called deformed matrix ensembles. Such ensembles consist generally of a mean-field matrix appropriately normalized to which a diagonal term, called the *random potential* is added. In some cases the potential is completely decoupled from the rest of the matrix: the distribution of the diagonal entries is arbitrary and the strength of the perturbation can be tuned by an external parameter see for instance [37]. In the case of the Laplacian matrix, the diagonal entries are obviously correlated with the off-diag entries. Moreover, they have an unbounded distribution as seen for the Central Limit Theorem approximation (Lemma 2.52) This model has been studied up to values $d \gg N^\alpha$, $\alpha > 0$, in [30].

Our proof differs from the preceding works in four ways. To begin with, we do not show convergence of g towards the deterministic function $\mathrm{m_{fc}}$ but only to the solution of (2.9).
The second difference is one of the key instruments to reach the scale $d \gg \sqrt{\log N}$ and was largely developed [5]. In that paper, the authors introduced the notion of typical and atypical vertices, characterizing those $x \in [N]$ for which the quantity

$$S_x^d := \sum_y^{(x)} \Big(|H_{xy}|^2 - \frac{1}{N}\Big)\widetilde{G}_{yy}^{(x)}$$

does not concentrate well. Although some non-trivial adaptations are required to make this line of argument applicable to our setup, the structure of the proof is similar.
The third difference is a new observation. In most local laws, the bootstrap assumption simply passes the information that G_{xy} is bounded uniformly in $x, y \in [N]$. In our proof, in we need more information, namely

(i) that the diagonal entries of G are comparable with v_x^{-1} for large degrees;

(ii) that the off-diagonal entries of G are smaller by a factor of $d^{-1/2}$ than any of the corresponding diagonal entry.

This information is encoded in the definition of the bootstrapping event θ defined in (2.24) and their derivation from large deviation techniques is fairly routine (see Proposition 2.32).
The final difference is technical and has to do with the ability to control multilinear large deviations of sparse random vectors. It expands on the techniques developed in both [30] and [28] and uses in addition the tweaked information found in θ.

2.2 Typical vertices in sparse regimes

Throughout this section $\sqrt{\log N} \leq d \leq (\log N)^{3/2}$. This section is devoted to an adaptation of the argument of [5]. The notion of typical and atypical vertices is introduced. Both the definition of typical vertices and the structure of the proof are identical, the techniques employed are however different.

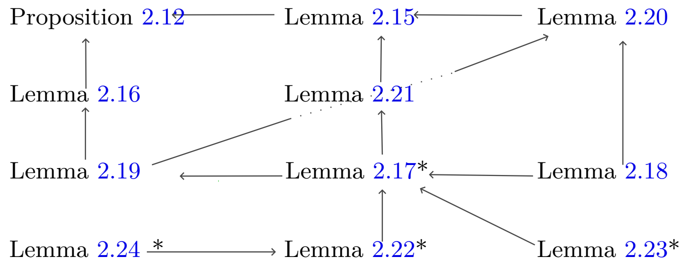


Figure 2.3: Dependencies in the proof of Proposition 2.12. The statements written with a star rely on estimates from Section 2.7. NotThe main erre that in Lemma 2.23, we use Proposition 2.32 in order to avoid rewriting the same argument.

Definition 2.10. For $x \in [N]$ we define $M^{(x)} := (M_{xy})_{x,y\in[N]\setminus T}$ to be the submatrix of M with the xth line and column removed. We generalize this to any $T \subset [N]$. We write

$$G^{(T)}(z) := \big(M^{(T)} - z\big)^{-1}. \tag{2.20}$$

We define $A^{(T)}$ and $H^{(T)}$ in the same way.

Let us introduce the main error parameter

$$\varphi_{\mathbf{a}} := \mathbf{a}\left(\frac{\log N}{d^2}\right)^{1/3}, \quad \mathbf{a} > 0. \tag{2.21}$$

The following definition is an analog of [5, Definition 4.6].

Definition 2.11 (Typical vertices)**.** Let $\mathbf{a} > 0$ be a constant, and define the set of *typical vertices* as

$$\mathcal{T}_{\mathbf{a}} := \{x \in [N] : |\Phi_x| \vee |\Psi_x| \leq \varphi_{\mathbf{a}}\}, \tag{2.22}$$

where

$$\Psi_x := \sum_y^{(x)} \Big(|H_{xy}|^2 - \frac{1}{N}\Big) G_{yy}^{(x)}, \quad \Phi_x := \sum_y^{(x)} \Big(|H_{xy}|^2 - \frac{1}{N}\Big), \tag{2.23}$$

Note that the matrix $G^{(T)}$ and $(H_{xy} \colon y \in [N])$ are correlated (see Definition 2.13). For this reason, the term Ψ_x is not amenable to large deviation estimates in the usual way. This is a major source of complications.

The bootstrapping events that control the entries of the Green function all depend on the parameter $\Gamma := \Gamma(\kappa) > 0$ which will be chosen large enough later. For $p > 0$ we define

$$\begin{gathered} \theta_d(\Gamma) := \mathbf{1}_{\max_{x,y}|G_{xy}|\leq\Gamma}, \quad \theta_o(p) := \mathbf{1}_{\max_{x\neq y[N]} \tilde{v}_{xy}|G_{xy}|\leq p}, \theta_i(\Gamma) := \mathbf{1}_{\min_{x\in[N]}|G_{xx}|(\tilde{v}_x\vee\Gamma)\geq 1/4}, \\ \theta(\Gamma, p) := \theta_d(\Gamma)\theta_o(p)\theta_i(\Gamma) \end{gathered} \tag{2.24}$$

where we introduced the quantities

$$\tilde{v}_x := |v_x - \operatorname{Re} z|, \quad \tilde{v}_{xy} := 1 \vee \tilde{v}_x \vee \tilde{v}_y, \quad x, y \in [N]. \tag{2.25}$$

The following proposition is the analog of [5, Proposition 4.8].

Proposition 2.12. *There exist $p, q \in (0,1)$, depending only on Γ, and $\mathbf{a} > 0$, depending on ν and q such that on the event $\{\theta(\Gamma, p) = 1\}$, the following statements hold with very high probability.*

(i) Most vertices are typical

$$|\mathcal{T}_{\mathbf{a}}^c| \le \exp(q\varphi_{\mathbf{a}}^2 d) + N\exp(-2q\varphi_{\mathbf{a}}^2 d). \tag{2.26}$$

(ii) Most neighbours of any vertex are typical

$$\sum_{y\in\mathcal{T}_{\mathbf{a}}^c}^{(x)} |H_{xy}|^2 \le 8\varphi_{\mathbf{a}}. \tag{2.27}$$

Note that for $\sqrt{\log N} \le d \le (\log N)$ we have $(\log N)^{1/6} \le \varphi_{\mathbf{a}}^2 d \le \sqrt{\log N}$, and therefore (2.26) implies

$$|\mathcal{T}_{\mathbf{a}}^c| \le \exp\big(\sqrt{\log N}\big) \vee N\exp\big(-(\log N)^{1/6}\big) \le N\exp\big(-(\log N)^{1/6}\big) \tag{2.28}$$

with very high probability.

The rest of this section is devoted to the proof of Proposition 2.12. We will need the following definitions.

Definition 2.13 (Decorellated submatrices)**.** For $T \subset [N]$ we define $\tilde{M}^{(T)}$ to be the $N - |T| \times N - |T|$ matrix as

$$\tilde{M}_{xy}^{(T)} = H_{xy} - \delta_{xy}\sum_{u\in T} H_{xu} - R_{xy}.$$

For $u \notin T$ we define $\tilde{M}^{(T,u)} = (\tilde{M}_{xy}^{(T,u)})_{x,y\in N\setminus T\cup\{u\}}$ to be the minor of $\tilde{M}^{(T)}$ obtained by removing the u-th row and column. We define

$$\widetilde{G}^{(T)}(z) := \big(\tilde{M}^{(T)} - z\big)^{-1}, \quad \widetilde{G}^{(T,u)}(z) := \big(\tilde{M}^{(T,u)} - z\big)^{-1}. \tag{2.29}$$

In Appendix A.1, we recall the standard identities that relate the entries of G and $G^{(x)}$, for $x \in [N]$. For $x \in [N]$ the entries of $\widetilde{G}^{(x)}$ and $G^{(x)}$ are related by the by the second resolvent identity (A.3)

$$\widetilde{G}_{ab}^{(x)} = G_{ab}^{(x)} - \sum_{c}^{(x)} \widetilde{G}_{ac}^{(x)} H_{xc} G_{cb}^{(x)}, \quad a, b \ne x.$$

The following definition has no direct analog in [5] but it is a generalization of [30, (3.1)].

Definition 2.14. For $x \in [N]$ and $T \subset [N]$ we define

$$\Psi_x^{(T)} := \sum_{y}^{(Tx)} \Big(|H_{xy}|^2 - \frac{1}{N}\Big)\widetilde{G}_{yy}^{(Tx)}, \quad \Phi_x^{(T)} := \sum_{y}^{(Tx)} \Big(|H_{xy}|^2 - \frac{1}{N}\Big),$$

and

$$\mathcal{T}_{\mathbf{a}}^{(T)} := \{x \in [N] : |\Phi_x^{(T)}| \vee |\Psi_x^{(T)}| \le \varphi_{\mathbf{a}}\}.$$

Note that $\Phi_x^{(\emptyset)} = \Phi_x$ but $\Psi_x^{(\emptyset)} \ne \Psi_x$ since $\widetilde{G}^{(x)} \ne G^{(x)}$. The following is the analog of [5, Lemma 4.15].

Lemma 2.15. *There are constants $0 < q \le 1$, depending on $\Gamma > 0$, and $\mathbf{a} > 0$, depending only on ν and q such that, for any deterministic set $X \subset [N]$, the following holds with very high probability on the event $\{\theta(q) = 1\}$.*

(i) $|X \cap \mathcal{T}_{\mathbf{a}/2}^c| \le \exp(q\varphi_{\mathbf{a}}^2 d) + |X|\exp(-2q\varphi_{\mathbf{a}}^2 d)$;

(ii) If $|X| \le \exp(2q\varphi_{\mathbf{a}}^2 d)$, *then* $|X \cap \mathcal{T}_{\mathbf{a}/2}^c| \le \varphi_{\mathbf{a}} d$.

For any deterministic $x \in [N]$ *the same estimates hold for* $(\mathcal{T}_{\mathbf{a}/2}^{(x)})^c$ *and a random set* $X \subset [N] \setminus \{x\}$ *that is independent of* $H^{(x)}$.

The proof of Lemma 2.15 is given after Lemma 2.21. The following is the analog of [5, Lemma 4.11].

Lemma 2.16. *There exists* $0 < q < 1$ *such that with very high probability, for any* $\mathbf{a} > 0$,

$$\theta|\Phi_y^{(x)} - \Phi_y| \le \varphi_{\mathbf{a}} \quad \theta|\Psi_y^{(x)} - \Psi_y| \le \varphi_{\mathbf{a}},$$

Proof of Proposition 2.12. For (i) we choose $X = [N]$ and use Lemma 2.15 (i) and the fact that $\mathcal{T}_{\mathbf{a}/2} \subset \mathcal{T}_{\mathbf{a}}$. By Lemma 2.16 we have $\mathcal{T}_{\mathbf{a}}^c \subset (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c$ with very high probability hence

$$\theta \sum_{y \in \mathcal{T}_{\mathbf{a}}^c}^{(x)} |H_{xy}|^2 \le \theta \sum_{y \in (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c}^{(x)} |H_{xy}|^2$$

with very high probability. Using the fact that $|H_{xy}| \le \frac{2A_{xy}}{\sqrt{d}} + \frac{2\sqrt{d}}{N}$, we find

$$\sum_{y \in (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c}^{(x)} |H_{xy}|^2 \le \frac{4}{d}\big|\{y \in S_1(x) \cap \big(\mathcal{T}_{\mathbf{a}/2}^{(x)}\big)^c\}\big| + \frac{4d}{N}.$$

Now observe that $S_1(x)$ is a measurable function of the family $(H_{xy})_{y\in[N]}$ and it is thus independent of $H^{(x)}$. Moreover by Lemma 2.48, we have $|S_1(x)| \le \log N \le q\varphi_{\mathbf{a}}^2 d$, for any fixed q and $\mathbf{a}$, and so $|S_1(x)| \le \exp(2q\varphi_{\mathbf{a}}^2 d)$. Applying Lemma 2.15 (ii) we find

$$\sum_{y \in (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c}^{(x)} |H_{xy}|^2 \le 4\varphi_{\mathbf{a}} + 4\frac{d}{N} \le 8\varphi_{\mathbf{a}}.$$

This concludes the proof. □

For $T \subset [N]$ we define $v_x^{(T)} := \sum_y^{(T)} H_{xy}$ and

$$\tilde{v}_x^{(T)} := \big|v_x^{(T)} - \operatorname{Re} z\big|, \quad \tilde{v}_{xy}^{(T)} := \tilde{v}_x^{(T)} \vee \tilde{v}_y^{(T)} \vee 2. \tag{2.30}$$

For $p > 0$, we define the following analog of (2.24),

$$\theta_d^{(T)}(\Gamma) := \mathbf{1}_{\max_{x,y\notin T}|\widetilde{G}_{xy}^{(T)}| \le 2\Gamma}, \quad \theta_o^{(T)}(p) := \mathbf{1}_{\max_{x\neq y, x,y\notin T}|\tilde{v}_{xy}^{(T)}\widetilde{G}_{xy}^{(T)}| \le p}, \theta_i^{(T)}(\Gamma) := \mathbf{1}_{\min_{x\notin T}|\widetilde{G}_{xx}^{(T)}|(\tilde{v}_x^{(T)}\vee\Gamma)\ge 1/16},$$
$$\theta^{(T)}(\Gamma, p) := \theta_d^{(T)}(\Gamma)\theta_o^{(T)}(p)\theta_i^{(T)}(\Gamma). \tag{2.31}$$

The following lemma is the analog of [5, Lemma 4.14]. However, its proof is much more involved in the case of M because in general, for $T \subset [N]$, the matrices $\widetilde{G}^{(T)}$ and $H^{(T)}$ are correlated.

Lemma 2.17. *There exists* $\mathbf{c} = \mathbf{c}_\nu$ *depending on* ν *and* κ, *and* $p > 0$, *depending on* Γ, *such that, for any* $\nu > 0$ *and any deterministic set* $T \subset [N]$ *satisfying* $|T| \le \mathbf{c}_\nu d/\Gamma^2$,

$$\theta(\Gamma, p) \max_{x,y\notin T} |\tilde{G}_{xy}^{(T)}| \le 4\Gamma, \tag{2.32}$$

and

$$\theta(\Gamma, p) \max_{x,y} |\widetilde{G}_{xy}^{(Tu)} - \widetilde{G}_{xy}^{(T)}| \le \frac{\mathcal{C}\Gamma}{d} + \mathcal{C}\Gamma \mathbf{1}_{x=y}\frac{A_{ux}}{\sqrt{d}}, \tag{2.33}$$

hold with probability $1 - O(N^{-\nu})$.

The proof of Lemma 2.17 is postponed to Section 2.3.

The following lemma can be compared with Lemma 2.48.

Lemma 2.18. *Let $T \subset [N]$ be deterministic. Then if $d|T| \leq \sqrt{N}$, we have*

$$\max_{x\in[N]} \big|S_1(x) \cap T\big| \leq \mathcal{C}, \tag{2.34}$$

with very high probability.

Proof. Let $x \in [N]$. Then $Z_x := \big|S_1(x) \cap T\big|$ is a binomial random variable with parameters $\mathcal{B}_{n,p}$, $n := |T|$, $p := \frac{d}{N}$. By Lemma B.3, we see that

$$\mathbb{P}\Big(Z_x - \mu \geq s\Big) \leq \exp\Big(-\mu h(a/\mu)\Big), \quad s \geq 0,$$

where $h(x) := x \log x - x + 1$ is defined in (B.1) and $\mu = \mathbb{E} Z_x = \frac{d|T|}{N}$. By assumption on d and $|T|$, $\mu \leq 1$ and we find

$$\mathbb{P}\Big(Z_x \geq C + 1\Big) \leq \exp\bigg(-C\Big[\log\Big(\frac{CN}{d|T|}\Big) - 1\Big]\bigg) \leq N^{-C/4}.$$

Choosing $C = \mathcal{C}_\nu$ large enough and applying a union bound, we get the desired result. ☐

The following is the analog of [5, Lemma 4.15].

Lemma 2.19. *There is a constant $0 < q \leq 1$, depending only on Γ, such that the following holds with very high probability.*
For any deterministic $T \subset [N]$,

$$\theta^{(T)}\mathbb{P}(|\Phi_x^{(T)}| \geq \varepsilon | H^{(T)}) \leq \mathrm{e}^{-32q\varepsilon^2 d}, \quad \theta^{(T)}\mathbb{P}(|\Psi_x^{(T)}| \geq \varepsilon | H^{(T)}) \leq \mathrm{e}^{-32q\varepsilon^2 d}. \tag{2.35}$$

Moreover for any $u \notin T$,

$$\Phi_x^{(Tu)} - \Phi_x^{(T)} = O(d^{-1}), \quad \theta^{(T)}\Big(\Psi_x^{(Tu)} - \Psi_x^{(T)}\Big) = \mathcal{O}\bigg(\big(1+\alpha_x\big)\bigg(\frac{A_{xu}}{\sqrt{d}} + \frac{1}{d}\bigg)\bigg). \tag{2.36}$$

and

$$\theta\big(\Psi_x^{(\emptyset)} - \Psi_x\big) = \mathcal{O}\left(\frac{1+\alpha_x}{\sqrt{d}}\right). \tag{2.37}$$

Note that $\theta^{(T)}$ is measurable with respect to $H^{(T)}$. This explains the position of $\theta^{(T)}$ outside of the conditional probability in (2.35). Due to the correlations inside of M, that lemma has weaker bounds than its analog in [5].

Proof of Lemma 2.16. Follows from (2.36) for $T = \emptyset$. ☐

Proof of Lemma 2.19. Since $\widetilde{G}^{(T)}$ and $\theta^{(T)}$ are measurable with respect to $H_{(T)}$, we can use (2.83b) in Proposition 2.38 with $\psi = \Gamma d^{-1/2}$ to get

$$\bigg\|\theta^{(T)} \sum_y^{(Tx)} \big(H_{xy}^2 - \frac{1}{N}\big)\widetilde{G}_{yy}^{(Tx)} \Big| H^{(T)}\bigg\|_r \leq 6\Gamma\sqrt{\frac{r}{d}}.$$

Applying Chebyshev's inequality with $r = 32q\varepsilon^2 d$ and $q = \frac{1}{(24\Gamma \mathrm{e})^2}$, we find

$$\theta^{(T)}\mathbb{P}(\Psi_x^{(T)} \geq \varepsilon) \leq \left(\frac{6\Gamma}{\varepsilon}\sqrt{\frac{r}{d}}\right)^r \leq \mathrm{e}^{-q\varepsilon^2 d}.$$

The large deviation bound on $\Phi_x^{(T)}$ is proved similarly using Proposition 2.38 with $a_i = 1$.
Equation (2.36) is a comparison argument and is derived similarly to [5, (4.32)] but using (2.33) instead of [5, (4.25)]. Note that the bounds we obtain are weaker than their counterparts.
On the event $\{\theta = 1\}$, we know that with very high probability $\max_{x,y,z}|\widetilde{G}_{xy}^{(z)}| \vee |G_{xy}^{(z)}| \leq 2\Gamma$, by Lemma 2.17. Therefore (2.37) follows from (A.8) and a double application of (2.97a),

$$\theta \sum_y^{(x)} (H_{xa}^2 - \frac{1}{N}) G_{yy}^{(x)} - \widetilde{G}_{yy}^{(x)} \leq \theta(\alpha_x + 1) \max_{y \neq x} |G_{yy}^{(x)} - \widetilde{G}_{yy}^{(x)}| \leq \frac{\mathcal{C}(\alpha_x + 1)}{\sqrt{d}}.$$

□

The following lemma is the analog of [5, Lemma 4.12]. Its proof is postponed to the end of this section.

Lemma 2.20. *There exist $p, q > 0$ depending only on Γ, such that, for any constants $\nu, \mathbf{a} > 0$, the following holds with very high probability. If $x \notin T \subset [N]$ are deterministic with $|T| \leq \varphi_{\mathbf{a}} d/\mathcal{C}$ then*

$$\mathbb{P}(T \subset \mathcal{T}_{\mathbf{a}/2}^c, \theta(p) = 1) \leq \mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 d|T|} + \mathcal{C}N^{-\nu},$$
$$\mathbb{P}\big(T \subset (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c, \theta(p) = 1\big) \leq \mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 d|T|} + \mathcal{C}N^{-\nu}.$$

Before proving Lemma 2.15, we need one last result, which is the analog of [5, Lemma 4.11].

Lemma 2.21. *There exists $p > 0$, depending only on Γ, such that for any deterministic $T \subset [N]$ satisfying $|T| \leq \frac{d}{\mathcal{C}\Gamma^2}$ we have $\theta(p) \leq \theta^{(T)}(p)$.*

Proof. For $p > 0$ small enough and $\mathcal{C} = \frac{2}{\mathbf{c}_\nu}$ the assumptions of Lemma 2.17 are satisfied. Using the bound (2.32), we conclude that $\theta = \theta\theta^{(T)}$ with very high probability. Since $\theta \leq 1$, the proof is complete. □

We are now ready to prove Lemma 2.15. The proof is essentially built around the same blueprint as the one in [5].

Proof of Lemma 2.15. Throughout the proof we abbreviate $\mathbb{P}_\theta(\Xi) := \mathbb{P}(\Xi \cap \{\theta = 1\})$. Let $\mathcal{C}$ be the constant from Lemma 2.20, and set

$$\mathbf{a} := \left(\frac{\mathcal{C}\nu}{4q}\right)^{1/3}. \tag{2.38}$$

For the proof of (ii), $k = \varphi_{\mathbf{a}} d/\mathcal{C}$ and use Lemma 2.20 to estimate

$$\begin{aligned}\mathbb{P}_\theta(|X \cap \mathcal{T}_{\mathbf{a}/2}^c| \geq k) &\leq \sum_{Y \subset X : |Y| \geq k} \mathbb{P}_\theta(Y \subset \mathcal{T}_{\mathbf{a}/2}^c) \leq \binom{|X|}{k} \big(\mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 dk} + \mathcal{C}N^{-\nu}\big) \\ &\leq \big(|X|\mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 d}\big)^k + \mathcal{C}|X|^k N^{-\nu} \leq \mathrm{e}^{-2q\varphi_{\mathbf{a}}^2 dk} + \mathcal{C}\mathrm{e}^{2q\varphi_{\mathbf{a}}^2 dk} N^{-\nu} \leq N^{-2q\mathbf{a}^3/\mathcal{C}} + \mathcal{C}N^{2q\mathbf{a}^3\mathcal{C}-\nu}.\end{aligned}$$

In the second step we used Lemma 2.19 and in the final step the assumption $d \leq (\log N)^{3/2}$.

To prove (i) we estimate for $t > 0$ and $l \in \mathbb{N}$,

$$\mathbb{P}_\theta(|X \cap \mathcal{T}_{\mathbf{a}/2}^c| \geq t) \leq \frac{1}{t^l} \sum_{x_1, \dots, x_l \in X} \mathbb{P}_\theta\big(x_i \in \mathcal{T}_{\mathbf{a}/2}^c,\, i \in [l]\big).$$

Choosing $l = \varphi_{\mathbf{a}} d/\mathcal{C}$, regrouping the summation according to the partition of coincidences, and using Lemma 2.19 yield

$$\begin{aligned}\mathbb{P}_\theta(|X \cap \mathcal{T}_{\mathbf{a}/2}^c| \geq t) &\leq \frac{1}{t^l} \sum_{\pi \in \mathcal{P}_l} |X|^{|\pi|} (\mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 d|\pi|} + \mathcal{C}N^{-\nu}) \\ &\leq \frac{1}{t^l} \sum_{k=0}^{l} \binom{l}{k} l^{l-k} |X|^k (\mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 dk} + \mathcal{C}N^{-\nu}) = \frac{(l + |X|\mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 d})^l + \mathcal{C}N^{-\nu}(l + |X|)^l}{t^l}.\end{aligned}$$

Here we denoted by $\mathcal{P}_l$ the set of partitions of $[l]$ and we denote by $k = |\pi|$ the number of blocks in the partition $\pi \in \mathcal{P}_l$. We also bounded the number of partitions of size k by $\binom{l}{k} l^{l-k}$. Using $l = \varphi_{\mathbf{a}} d/\mathcal{C}$ and choosing $t = \mathrm{e}^{q\varphi_{\mathbf{a}}^2 d} + |X|\mathrm{e}^{-2q\varphi_{\mathbf{a}}^2 d}$ as well as $\mathcal{C}$ and ν sufficiently large gives, using $d \geq \mathcal{C}\sqrt{\log N}$, $l \leq \mathrm{e}^{q\varphi_{\mathbf{a}} d} \leq N^{-\nu}$. Moreover since $d \leq (\log N)^{3/2}$ we get $\mathrm{e}^{-2q\varphi_{\mathbf{a}}^2 d} \leq N^{-\nu}$. Therefore

$$\mathbb{P}_\theta(|X \cap \mathcal{T}_{\mathbf{a}/2}^c| \geq k) \leq \mathcal{C}N^{-\nu}.$$

To obtain the same statements for $\big(\mathcal{T}_{\mathbf{a}/2}^{(x)}\big)^c$ we estimate

$$\mathbb{P}_\theta(|X \cap (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c| \geq k) \leq \mathbb{E}\big[\mathbb{P}_\theta\big(|X \cap (\mathcal{T}_{\mathbf{a}/2}^{(x)})^c| \geq k, \theta^{(x)}(p) = 1|X\big)\big] + \mathbb{P}(\theta^{(x)}(p) = 0, \theta(p) = 1).$$

Now, since the set $(\mathcal{T}_{\mathbf{a}/2}^{(x)})^c$ and the indicator function $\theta^{(x)}$ are independent of X we bound the conditional probability as before. Finally $\mathbb{P}(\theta^{(x)} = 0, \theta = 1) \leq N^{-\nu}$ is a consequence of Lemma 2.21. This concludes the proof of Lemma 2.15. □

Proof of Lemma 2.20. Throughout the proof we abbreviate $\mathbb{P}_\theta(\Xi) := \mathbb{P}(\Xi \cap \{\theta = 1\})$. Let us define the events

$$\Omega_x := \{|\Phi_x| \geq \varphi_{\mathbf{a}/2}\} \cup \{|\Psi_x| \geq \varphi_{\mathbf{a}/2}\}, \quad \Omega_x^{(T)} := \{|\Phi_x^{(T)}| > \varphi_{\mathbf{a}/4}\} \cup \{|\Psi_x^{(T)} > \varphi_{\mathbf{a}/4}|\}.$$

We have

$$\mathbb{P}\big(T \subset \mathcal{T}_{\mathbf{a}/2}^c, \theta = 1\big) = \mathbb{P}_\theta\Big(\bigcap_{x \in T} \Omega_x\Big),$$

and using a union bound we deduce

$$\begin{aligned}\mathbb{P}_\theta\Big(\bigcap_{x \in T} \Omega_x\Big) \leq \mathbb{P}_\theta\Big(\bigcap_{x \in T} \Omega_x^{(T)}\Big) &+ \sum_{x \in T} \mathbb{P}_\theta\big(|\Phi_x - \Phi_x^{(T)} > \varphi_{\mathbf{a}/4}\big) \\ &+ \sum_{x \in T} \mathbb{P}_\theta\big(|\Phi_x| \leq \varphi_{\mathbf{a}/2}, |\Psi_x - \Psi_x^{(T)}| > \varphi_{\mathbf{a}/4}\big).\end{aligned}$$

The first term is an intersection of events that are independent after conditioning on $H^{(T)}$. Using (2.35) we find

$$\mathbb{P}_\theta\Big(\bigcap_{x \in T} \Omega_x^{(T)}\Big) = \mathbb{E}\Big[\prod_{x \in T} \mathbb{P}\Big[\Omega_x^{(T)}\big|H^{(T)}\Big]\Big] \leq \mathrm{e}^{-4q\varphi_{\mathbf{a}}^2 d|T|}. \tag{2.39}$$

On the event $\big\{\theta = 1\big\} \cap \big\{|\Phi_x| \leq \varphi_{\mathbf{a}/2}\big\}$, using Lemma 2.18 and (2.36), we find that

$$|\Psi_x^{(\emptyset)} - \Psi_x^{(T)}| \leq \mathcal{C}\big(1 + \alpha_x\big)\bigg(\frac{|S_1(x) \cap T|}{\sqrt{d}} + \frac{|T|}{d}\bigg) \leq \mathcal{C}\bigg(\frac{1}{\sqrt{d}} + \frac{|T|}{d}\bigg) \leq \varphi_{\mathbf{a}/4},$$

holds with very high probability. Moreover observe that if $\Phi_x \leq \varphi_{\mathbf{a}}$ then $\alpha_x \leq 2$ and, using (2.37), we have

$$\big|\Psi_x - \Psi_x^{(\emptyset)}\big| \leq \frac{\mathcal{C}}{\sqrt{d}} \leq \varphi_{\mathbf{a}/8},$$

with very high probability. We conclude that $\Psi_x - \Psi_x^{(T)} \leq \varphi_{\mathbf{a}/4}$. Therefore, choosing $\mathbf{a}$ large enough, we see that the right-hand side of (2.39) is bounded by $\mathcal{C}N^{-\nu}$. This concludes the proof. □

2.3 Proof of Lemma 2.17

Throughout this section, we work on the very high probability event defined in Lemma 2.48, we fix $T \subset [N]$ satisfying the assumptions of Lemma 2.17. After a relabelling of the vertices, we can suppose that $T = \{1, \dots, |T|\}$. Let us introduce the following variables.

$$\Gamma_k := 1 \vee \max_x |\widetilde{G}_{xx}^{([k])}|, \quad P_k := \max_x \big|\big(V(k,x)\widetilde{G}_{xx}^{([k])}\big)^{-1}\big|, \quad Q_k := \max_{x \neq y} |V(k,x,y)\widetilde{G}_{xy}^{([k])}|.$$

where

$$V(k,x,y) := 2\max(\tilde{v}_x^{([k])}, \tilde{v}_y^{([k])}, \Gamma), \quad V(k,x) := V(k,x,x).$$

In particular not that $\max_{k,x,y}|V(k,x,y)| \le \log N$ with very high probability by Lemma 2.48.

We also introduce the following quantity that bounds the maximum of the Green function entries and all of its submatrices at stage k,

$$\tilde{\Gamma}(k) := \max_{x,y,u}\Big(\big|\widetilde{G}_{xy}^{([k])}\big|, \big|\widetilde{G}_{xy}^{([k]u)}\big|, \big|\widetilde{G}_{xy}^{([k],u)}\big|\Big),$$

where the maximum is taken over all $u \notin [k]$ and then over all $x, y \notin [k] \cup \{u\}$. Note that $\tilde{\Gamma}(0)$ is exactly $\tilde{\Gamma}$ defined in (2.54).

The key difficulty is to show that, as we increase k, the entries of the Green functions remain bounded. We focus on this issue in the proof of Lemma 2.17. However, we need *a priori* bounds on P_k and Q_k. We also need to be sure that Γ_{k+1} is not too large if Γ_k is bounded. These facts are the contents of the two following auxiliary lemmas, which are proved at the end of the section.

Let us define

$$\Gamma_k^{(+)} := 1 \vee \max_x |\widetilde{G}_{xx}^{([k],k+1)}|.$$

Lemma 2.22. *For any $\Gamma = O(1)$, there exists $p > 0$, depending only on Γ, such that on the event $\{Q_k \le 2p\} \cap \{P_k, \Gamma_k \le \Gamma\}$,*

$$\tilde{\Gamma}(k) \le 2\Gamma_k, \quad Q_{k+1} \le \frac{\mathcal{C}\Gamma_k}{\sqrt{d}}, \quad P_{k+1} \le 2P_k. \tag{2.40}$$

In particular $\Gamma_k^{(+)} \vee \Gamma_{k+1} \le 2\Gamma_k$ holds with very high probability.

Lemma 2.23. *For any $\Gamma = O(1)$,*

$$\frac{1}{2} \le V(x,k)\big|\widetilde{G}_{xx}^{([k])}\big| \le 10\Gamma^2, \quad x \notin [k], \tag{2.41}$$

holds with very high probability on the the event $\{\tilde{\Gamma}(k) \le \Gamma\}$. In particular, $P_k \le 2$ holds with very high probability on that event.

Proof of Lemma 2.17. For a deterministic set $A \subset [N]$, we introduce the random variable

$$\Delta(k,x) := |\{S_1(x) \cap [k]\}|, \quad x \in [N]$$

and $\Delta(k) := \max_{x\in[N]} \Delta(k,x)$. We will prove by induction on k that there is $\mathcal{C} > 0$ such that with very high probability

$$\max_x \Big(1 + \frac{64\mathcal{C}\Gamma^2}{\sqrt{d}}\Big)^{-\Delta(k,x)} \widetilde{G}_{xx}^{(k)} \le \Gamma_0\Big(1 + \frac{64\mathcal{C}\Gamma^2}{d}\Big)^k, \quad Q_k \le 2p, \quad P_k \le 2\Gamma \tag{2.42}$$

for all $k \in \mathbb{N}$ satisfying $k \leq \frac{d}{128\mathcal{C}\Gamma^2}$. In particular, by Lemma 2.18, $\max_{1\leq k\leq |T|}\Delta(k) \leq \mathcal{C}$ with very high probability. Therefore (2.42) implies

$$\Gamma_k \leq \Gamma_0\Big(1+\frac{64\mathcal{C}\Gamma^2}{\sqrt{d}}\Big)^{\Delta(k)}\Big(1+\frac{64\mathcal{C}\Gamma^2}{d}\Big)^k \leq 2\Gamma_0 \leq 4\Gamma.$$

It thus suffices to establish (2.42) to prove (2.32).
The initialization $k=0$ is trivial. Suppose the induction holds up to k and denote $z=k+1$. Then by our upper bound on k, we know that $\Gamma_k \leq 2\Gamma_0 \leq 4\Gamma$. Therefore by Lemma 2.22 (renaming Γ as 4Γ), for $p>0$ small enough depending only on Γ, we have with very high probability,

$$\Gamma_k^{(+)} \vee \Gamma_{k+1} \leq \tilde{\Gamma}(k) \leq 2\Gamma_k, \quad Q_{k+1} \leq \frac{\mathcal{C}\Gamma_k}{\sqrt{d}}, \quad P_{k+1} \leq 2P_k. \tag{2.43}$$

By induction hypothesis, $\Gamma_k \leq 2\Gamma_0 \leq 4\Gamma$, we conclude that for any constant $p>0$, $Q_{k+1} \leq \frac{\mathcal{C}\Gamma_k}{\sqrt{d}} \leq 2p$ holds with very high probability. Since $\tilde{\Gamma}(k) \leq 4\Gamma = O(1)$, we can apply Lemma 2.23 with Γ replaced by 8Γ and we find that $P_{k+1} \leq 2$ holds with very high probability. Thus we have proved the second and third inequalities of (2.42).
Of course (2.43) is not sufficient to establish the first inequality in (2.42) if we want $|T|$ to be comparable to d. We will now improve on those estimates using Proposition 2.49 and with the stronger control $Q_{k+1} \leq \frac{\mathcal{C}\Gamma_k}{\sqrt{d}}$.

The starting point is the equality

$$\widetilde{G}_{xy}^{([k+1])} = \widetilde{G}_{xy}^{([k])} + \Big(\widetilde{G}_{xy}^{([k],z)} - \widetilde{G}_{xy}^{([k])}\Big) + \Big(\widetilde{G}_{xy}^{([k+1])} - \widetilde{G}_{xy}^{([k],z)}\Big). \tag{2.44}$$

We expand the first term on the right-hand side of (2.44) using (A.8b) with $T=[k]$ and $u=z$. Using Lemma A.1 and $|H_{au}| \leq Kd^{-1/2}$, we find

$$\Bigg|\frac{f}{N}\widetilde{G}_{zz}^{([k])}\sum_a^{(Tu)}\widetilde{G}_{xa}^{([k],z)}\Bigg| \leq 4\Gamma_k f\sqrt{\frac{\Gamma_k^{(+)}}{N\eta}}, \qquad \frac{f^2}{N^2}\Bigg|\widetilde{G}_{zz}^{([k])}\sum_{a,b}^{([k])}\widetilde{G}_{ab}^{([k],z)}\Bigg| \leq 4\Gamma_k f^2\sqrt{\frac{\Gamma_k^{(+)}}{N\eta}}.$$

Using (2.97a) Proposition 2.49 we see that, on $\{\Gamma_k^{(+)}, \Gamma_{k+1} \leq 2\Gamma_k\}$, and using again the fact that $\Gamma_k = O(1)$ to check (2.96), we find

$$\Bigg|\sum_a^{([k+1])} H_{xa}\widetilde{G}_{ay}^{([k],z)}\Bigg| \leq \frac{\mathcal{C}\Gamma_k}{\sqrt{d}},$$

and

$$\widetilde{G}_{zz}^{([k])}\sum_a^{([k+1])} H_{ua}\widetilde{G}_{ay}^{([k],z)}\sum_a^{([k+1])} H_{ua}\widetilde{G}_{ax}^{([k],z)} \leq \frac{\mathcal{C}\Gamma_k^3}{d} \leq \frac{16\Gamma\mathcal{C}\Gamma_k}{d},$$

hold with very high probability. Using the fact that $d, f = O\big(N^{\tau/6}\big)$ and $\Gamma = O(1)$, we conclude that

$$\Big|\widetilde{G}_{xy}^{([k],z)} - \widetilde{G}_{xy}^{([k])}\Big| \leq \frac{64\mathcal{C}\Gamma_k}{d}, \tag{2.45}$$

holds with very high probability.
Using (2.43), (A.4) and (2.98b) of Proposition 2.49 with $T=[k]$, $u=z$, $\Gamma(T)=4\Gamma_k$ and $Q(T\cup\{u\}) \leq \mathcal{C}\Gamma_k d^{-1/2}$, we find that

$$\widetilde{G}_{xy}^{([k+1])} - \widetilde{G}_{xy}^{([k],z)} = \sum_a^{[k+1]}\widetilde{G}_{xa}^{([k],z)}H_{za}\widetilde{G}_{ay}^{([k+1])} \leq \mathbf{1}_{x=y}A_{xz}\frac{\mathcal{C}\Gamma_k}{\sqrt{d}} + \frac{\mathcal{C}\Gamma_k}{d}, \tag{2.46}$$

holds with very high probability for $x, y \notin [k+1]$. Plugging the previous estimates into (2.44), we find that

$$\max_{x \neq y} |\widetilde{G}_{xy}^{([k+1])}| \leq \left(1 + \frac{64\Gamma^2\mathcal{C}}{d}\right)\Gamma_k, \quad \max_{x} |\widetilde{G}_{xx}^{([k+1])}| \leq \left(1 + \frac{64A_{xk+1}\Gamma^2\mathcal{C}}{\sqrt{d}} + \frac{64\Gamma^2\mathcal{C}}{d}\right)|\widetilde{G}_{xx}^{([k])}|,$$

We deduce from the induction hypothesis that

$$|\widetilde{G}_{xx}^{([k+1])}| \leq \left(1 + \frac{64\Gamma^2\mathcal{C}}{\sqrt{d}}\right)^{A_{zx}+\Delta(k,x)} \left(1 + \frac{64\Gamma^2\mathcal{C}}{d}\right)^{k+1} |\widetilde{G}_{xx}^{([k])}|\Gamma_0, \quad x \notin [k+1],$$

holds with very high probability. Now using the fact that $A_{zx} + \Delta(k,x) = \Delta(k+1,x)$, we can conclude the proof of the first inequality in (2.42). This concludes the proof of (2.42) and the induction.

Equation (2.33) follows from (2.44), (2.45) and (2.46). This concludes the proof. □

The rest of the section is devoted to the proof of Lemmas 2.22 and 2.23. We first prove this simple consequence of Lemma 2.47. Let us denote by $\|X\|_{r|H^{(z)}} := \left(\mathbb{E}\left[|X|^r|H^{(z)}\right]\right)^{1/r}$, for X a random variable.

Lemma 2.24. *Let $\Gamma > 0$ and $T \subset [N]$. Suppose $\max_{x,y\notin T} \left|\widetilde{G}_{xy}^{(T)}\right| \leq \Gamma$. Then*

$$\sum_a |H_{az}||\widetilde{G}_{ax}^{(z)}| \leq \frac{\mathcal{C}\Gamma}{\sqrt{d}}, \quad x \in [N] \setminus T,$$

holds with very high probability

Proof. Let us fix $x \in [N]$. Let $Z_a := |H_{za}| - \mathbb{E}|H_{za}|$. Then

$$\mathbb{E}\big||H_{za}| - \mathbb{E}|H_{za}|\big| \leq \frac{4\sqrt{d}}{N} - \frac{d^{3/2}}{N^2} \leq \frac{4\sqrt{d}}{N}, \quad \mathbb{E}\big(|H_{za}| - \mathbb{E}|H_{za}|\big)^2 \leq \frac{2}{N}.$$

Using Lemma A.1 and the fact that $\widetilde{G}^{(z)}$ is measurable with respect to of $H^{(z)}$, we can apply Lemma 2.47 with $\gamma = N^{-\kappa/4}$ and Γ as Γ to get

$$\left\| \sum_a^{(z)} \Big(|H_{za}| - \mathbb{E}|H_{za}|\Big)|\widetilde{G}_{ax}^{(z)}| \right\|_{r|H^{(z)}} \leq \frac{1}{\sqrt{d}} \frac{64r\Gamma}{1 + \kappa \log N}, \quad r \in \mathbb{N}^*.$$

Setting $r = \nu \log N$, $\mathcal{C} \geq 64\mathrm{e}\nu\Gamma$ and using Chebyshev's inequality yields a bound in very high probability. We conclude that

$$\begin{aligned}
\sum_a^{(z)} |H_{za}||\widetilde{G}_{ax}^{(z)}| &= \sum_a^{(z)} \big(|H_{za}| - \mathbb{E}|H_{za}|\big)|\widetilde{G}_{ay}^{(z)}| + \mathbb{E}|H_{za}| \sum_a^{(z)} |\widetilde{G}_{ax}^{(z)}| \\
&\leq \frac{\mathcal{C}\Gamma}{\sqrt{d}} + \frac{\Gamma\sqrt{d}}{N^{\kappa/4}} \leq \frac{\mathcal{C}\Gamma}{\sqrt{d}},
\end{aligned}$$

holds with very high probability. In the second inequality we used Lemma A.1. Since x was arbitrary and all bounds hold with very high probability, we conclude. □

Proof of Lemma 2.22. Without loss of generality, we set $k = 0$, $\Gamma_0 = \Gamma_k$ and $k+1 = z$. Let $V(x) := V(0,x,x)$ and $V(x,y) := V(0,x,y)$ and

$$\Gamma^{(+)} := \max_{x,y} \left|G_{xy}^{(z)}\right|, \quad P^{(+)} := \max_x \left|\left(G_{xx}^{(z)}V(x)\right)^{-1}\right|, \quad Q^{(+)} := \max_{x\neq y} \left|V(x,y)G_{xy}^{(z)}\right|.$$

We will first prove that

$$\Gamma^{(+)} \leq \frac{4}{3}\Gamma_0, \quad Q^{(+)} \leq 4p, \quad P^{(+)} \leq \frac{4}{3}P_0. \tag{2.47}$$

We choose $p>0$ such that $p\le \frac{1}{4\Gamma}$. Applying (A.7), with $T=\emptyset$ and for $x,y\neq z$ we find

$$|G^{(z)}_{xy}|\le |G_{xy}|+\Big|\frac{G_{xz}G_{zy}}{G_{zz}}\Big|\le \Gamma_0+\frac{p}{V(x,z)}\frac{p}{V(x,y)}2V(z)\le \Gamma_0+\frac{8p^2}{\Gamma_0}\le \frac{4}{3}\Gamma_0,$$

where we used that $V(x),V(z)\le V(x,z)$ and $2\Gamma\le V(x,y)$ and $8p^2\le\Gamma_0^2$. We conclude that $\Gamma^{(+)}\le\frac{4}{3}\Gamma_0$.
If $x\neq y$, using $|G_{xz}|\le\frac{2p}{V(x,z)}$, we find

$$\big|V(x,y)G^{(z)}_{xy}\big|\le V(x,y)\Big[\frac{2p}{V(x,y)}+\frac{p}{V(x,z)}\frac{p}{V(x,y)}2V(z)\Big]\le 4p.$$

We conclude that $Q^{(+)}\le 4p$.
If $x=y$, using $|G_{xz}|\le\frac{2p}{V(x,z)}$ and $V(x),V(z)\le V(x,z)$, we find

$$\begin{aligned}\Big|V(x)G^{(z)}_{xx}\Big|^{-1}=&\Big|V(x)\Big(G_{xx}+\frac{G_{xz}G_{zx}}{G_{zz}}\Big)\Big|^{-1}=\frac{1}{V(x)G_{xx}}\Big|1-\frac{G_{xz}G_{zx}}{G_{zz}G_{xx}}\Big|^{-1}\\ \le&\frac{1}{V(x)G_{xx}}\big(1-4p^2\Gamma^2\big)^{-1}\le\frac{4}{3}P_0.\end{aligned}$$

Here we used the estimates $\frac{1}{G_{xx}}\le P_0V(x)$ as well as $p<(4\Gamma)^{-1}$. We deduce that $P^{(+)}\le\frac{4}{3}P_0$. This establishes (2.47).

We will now prove

$$\Gamma_1\le\frac{4}{3}\Gamma^{(+)},\quad Q_1\le\frac{\mathcal{C}\Gamma_0}{\sqrt{d}}Q^{(+)},\quad P_1\le\frac{4}{3}P^{(+)}. \tag{2.48}$$

The proof is then complete since combining (2.47) and (2.48) yields (2.40) with very high probability. Since $z=k+1$ was arbitrary, $\tilde\Gamma(k)\le 2\Gamma_0$ follows from (2.48) and a union bound.

Using (A.4) and Lemma 2.24 with $\Gamma=\Gamma_1$, we find

$$|\widetilde G^{(z)}_{xy}|=\Big|G^{(z)}_{xy}+\sum_a^{(z)}G^{(z)}_{xa}H_{za}\widetilde G^{(z)}_{ay}\Big|\le\Gamma^{(+)}+\Gamma^{(+)}\sum_a^{(z)}|H_{za}||\widetilde G^{(z)}_{ay}|\le\frac{3}{2}\Gamma_0+\frac{3\mathcal{C}\Gamma_0\Gamma_1}{2\sqrt{d}}, \tag{2.49}$$

holds with very high probability for any $x,y\neq z$. We find that $\big(1-\frac{\mathcal{C}\Gamma_0}{\sqrt{d}}\big)\Gamma_1\le\Gamma^{(+)}\le\frac{4}{3}\Gamma_0$ and therefore

$$\Gamma_1\le 2\Gamma_0,$$

holds with very high probability. If $x\neq y$, then we can multiply (2.49) on both sides by $V(x,y)$ by substituting $Q^{(+)}_k$ for $\Gamma^{(+)}$. Since $Q^{(+)}_k\le 4p$, we find $Q_1\le 8p$.
Observe that the bounds derived on Γ_1 are independent of the choice of z and since those bounds hold with very high probability, they can be derived simultaneously for every possible choice of z. We conclude that

$$\max_{a,b,x}|\widetilde G^{(x)}_{ab}|\vee|G^{(x)}_{ab}|\le 2\Gamma_0,\quad \max_{a,b,x}\mathbf{1}_{a\neq b}V(a,b)\Big(|\widetilde G^{(x)}_{ab}|\vee|G^{(x)}_{ab}|\Big)\le 8p, \tag{2.50}$$

holds with very high probability.

We conclude that the event $\tilde\Gamma(k)\le 2\Gamma_0$ holds with very high probability and since $\Gamma=O(1)$, we can apply Lemma 2.23. From (2.41) we see that there exists a constant $C\ge 0$ such that $\max_{x\in[N]}V(x)|G_{xx}|\le 40\Gamma_0^2\le C$ with very high probability. Let $x\neq y$ and suppose, without loss of generality, that $V(y)\le V(x)$. Using (A.6), we find the identity

$$V(x)G_{xy}=V(x)G_{xx}\sum_a^{(x)}H_{xa}G^{(x)}_{ay}+\frac{f}{N}V(x)G_{xx}\sum_a^{(x)}G_{ay}.$$

The second sum can be controlled by $Cf\sqrt{\Gamma_0 N^{-\kappa}} \le Cd^{-1/2}$ after using Lemma A.1. For the second term we use Proposition 2.49 (2.97a), with $\Gamma = 2\Gamma_0$, and find that

$$\Big|V(x)G_{xx}\Big|\Big|\sum_a^{(x)} H_{xa}G_{ay}^{(x)}\Big| \le C\frac{\mathcal{C}\Gamma_0}{\sqrt{d}} \le \frac{\mathcal{C}\Gamma_0}{\sqrt{d}}, \quad x \neq z,$$

holds with very high probability. If $V(y) \le V(x)$ we permute x and y in the above equation. We conclude that $Q_0 \le \frac{\mathcal{C}\Gamma_0}{\sqrt{d}}$. We can use this bound and (A.8) to find that

$$V(x,y)|G_{xy}^{(z)}| \le V(x,y)|G_{xy}| + V(x,y)|G_{xz}|\frac{|G_{zx}|}{|G_{zz}|} \le Q_0 + \frac{\Gamma V(x,y)V(z)}{dV(x,z)V(z,y)} \le 2Q_0 \le \frac{2\mathcal{C}\Gamma_0}{\sqrt{d}}, \quad x \neq y,$$

holds with very high probability. Since x and y are arbitrary we deduce $Q^{(+)} \le \frac{\mathcal{C}\Gamma_0}{\sqrt{d}}$.
We will now prove $Q_1 \le \frac{\mathcal{C}\Gamma_0}{\sqrt{d}}$ from which (2.48) will follow. Let $x \neq y$ and suppose, without loss of generality, that $V(y) \le V(x)$. Using (A.4) we find

$$\widetilde{G}_{xy}^{(z)} = G_{xy}^{(z)} - \sum_a^{(xz)} G_{xa}^{(z)}H_{za}\widetilde{G}_{ay}^{(z)} - G_{xx}^{(z)}H_{zy}\widetilde{G}_{yx}^{(z)}.$$

Multiplying on both sides by $V(x,y) = V(x)$ and using $Q^{(+)} \le \frac{\mathcal{C}\Gamma_0}{\sqrt{d}}$ we find

$$V(x)|\widetilde{G}_{xy}^{(z)}| \le Q^{(+)} + \sum_a^{(zy)} \big|\widetilde{G}_{xa}^{(z)}\big||H_{za}| \max_{a\neq z,x}\Big(V(x)\big|G_{xa}^{(z)}\big|\Big) + V(x)\big|\widetilde{G}_{yx}^{(z)}\big|\frac{2K\Gamma_0}{\sqrt{d}}.$$

Using Lemma 2.24 with Γ as $2\Gamma_0$, we deduce

$$V(x)|\widetilde{G}_{xy}^{(z)}| \le 2Q^{(+)}\Big(1 + \frac{\mathcal{C}\Gamma_1}{\sqrt{d}}\Big)$$

If $V(x) \le V(y)$, we exchange the x and y in the above equations and we can replace $V(x)$ by $V(x,y)$ in the last line. Since x and y were arbitrary, we deduce that $Q_1 \le 4Q^{(+)} \le \frac{\mathcal{C}\Gamma_0}{\sqrt{d}}$. This concludes the proof. ☐

There only remains to prove Lemma 2.23. Note that in its proof, we use a result from Proposition 2.32 (the very high probability bound (2.65c) to be precise). There is no logical loop, the only hypothesis of Proposition 2.32 is $\tilde{\Gamma}(0) = O(1)$ and we do not use Lemma 2.24 in its proof.

Proof of Lemma 2.23. Without loss of generality we let $k = 0$, $k+1 = z$, $\tilde{v}^{(T)}(x) = \tilde{v}(x)$. Let $x \in [N]$, $x \neq z$ and $V(x) = V(0,x)$. We want to show that

$$\frac{1}{2} \le V(x)|G_{xx}| \le 10\Gamma^2, \tag{2.51}$$

holds with very high probability. The starting point is the algebraic identity (A.2)

$$\frac{1}{G_{xx}} = v_x - z - \sum_a^{(xz)} H_{xa}^2 G_{yy}^{(x)} + Y_x = \tilde{v}(x) - \operatorname{Im} z - \sum_a^{(xz)} H_{xa}^2 G_{yy}^{(x)} + Y_x, \tag{2.52}$$

with Y_x defined as in (2.64). Using Proposition 2.32 below, we see that $Y_x \le \frac{\mathcal{C}\Gamma}{\sqrt{d}}$ with very high probability on the event $\{\tilde{\Gamma}(0) \le \Gamma\}$. Moreover, by definition we have

$$\sum_a^{(xz)} H_{xa}^2 G_{yy}^{(x)} = \frac{1}{d}\sum_{y\sim x} G_{yy}^{(x)} + O\Big(\frac{\Gamma d^2}{N}\Big),$$

and that $v_x = \sqrt{d}(\alpha_x - 1)$ and $\alpha_x = \frac{v_x}{\sqrt{d}} + 1$.

Suppose $|\tilde{v}_x| \le 2\Gamma$. In that case $V(x) = 2\Gamma$ and $\alpha_x \le \frac{3}{2}$ since $\frac{v_x}{\sqrt{d}} \le \frac{1}{2}$. We immediately have $|G_{xx}|V(x) \le 2\Gamma^2$ and

$$\left|\frac{1}{G_{xx}}\right| \le |\tilde{v}_x| + \alpha_x \Gamma + \frac{\mathcal{C}\Gamma}{\sqrt{d}} \le 4\Gamma.$$

Dividing the above inequation by $V(x)$ and using the fact that $\Gamma \ge 1$, we see that (2.51) holds in that case.

If $2\Gamma \le |\tilde{v}_x| \le \sqrt{d}$, then $V(x) = 2\tilde{v}_x$ and $\alpha_x \le \frac{3}{2}$ and we find

$$\frac{1}{2} \le \frac{2\tilde{v}_x}{\tilde{v}(x) + \frac{3}{2}\Gamma + \frac{\mathcal{C}\Gamma}{\sqrt{d}}} \le V(x)|G_{xx}| \le \frac{2\tilde{v}_x}{\tilde{v}(x) - \frac{3}{2}\Gamma - \frac{\mathcal{C}\Gamma}{\sqrt{d}}} \le 6 \le 10\Gamma^2.$$

Using the fact that $\frac{\alpha_x \Gamma}{\tilde{v}_x} = O\big(d^{-1/2}\big)$, we see that the right-hand side of the above equation satisfies the lower and upper bounds of (2.51).

Finally suppose $\sqrt{d} \le |\tilde{v}(x)|$. Then $\frac{\alpha_x}{\tilde{v}(x)} \le 2$ and from (2.52) we see that

$$\left|\frac{1}{G_{xx}}\right| = V(x)\big(1 + O(\Gamma/d)\big) + \mathcal{O}(\Gamma\mathcal{C}/\sqrt{d}).$$

where we used $\tilde{v}(x) = V(x)\big(1 + O\big(d^{-1/2}\big)\big)$. Multiplying the above by $V(x)$ yields (2.51). □

2.4 Proof of Theorem 2.4

In this section, we prove Theorem 2.4 and assume $d \le (\log N)^{3/2}$ throughout. Let us define the error parameters, for $T \subset [N]$,

$$\Lambda_d(T) := \max_{x \in T} |G_{xx} - m_x|, \qquad \Lambda_a(T) := \max_{x \in T} \left| G_{xx} - \frac{1}{v_x - z - \sum_y H_{xy}^2 m_y} \right|, \qquad \Lambda_o := \max_{x \ne y} |G_{xy}|. \tag{2.53}$$

The goal is to prove that $\Lambda_a([N])$ is small. As an intermediate step, we will show that $\Lambda_d(T)$ is small for $T = \mathcal{T}_{\mathbf{a}}$ the typical vertices. For $d \le \log N$, that $\Lambda_d([N])$ we cannot show that is small. Indeed for atypical vertices, m_x might not be a good approximation of G_{xx}. On the other hand, off-diagonal entries of G are always much smaller than 1, and this explains why Λ_o is introduced.

Let us also introduce the quantity $\tilde{\Gamma}$, that bounds the entries of the Green function and its associated modifications,

$$\tilde{\Gamma} = \max_{x,y,z} |G_{xy}| \vee |G_{xy}^{(z)}| \vee |\widetilde{G}_{xy}^{(z)}|, \tag{2.54}$$

where the maximum is taken over all $z \in [N]$ and $x, y \in [N] \setminus \{z\}$.

Let us introduce the following set of self-consistent equations.

Definition 2.25 (Restricted quadratic vector equation)**.** Let $v_x \in \mathbb{R}$, $x \in [N]$, $v = (v_x)_{x \in [N]} \in \mathbb{R}^N$. For $\mathcal{X} \subset [N]$. We define the vector $m := (m_x^{\mathcal{X},v})_{x \in \mathcal{X}}$ to be the unique solution of

$$\frac{1}{m_x^{\mathcal{X}}} = v_x - z - \frac{1}{|\mathcal{X}|} \sum_{y \in \mathcal{X}} m_y^{\mathcal{X}}, \qquad x \in \mathcal{X}, \quad z \in \mathbb{H}, \tag{2.55}$$

such that $m_x^{\mathcal{X}} \in \mathbb{H}$, for all $x \in \mathcal{X}$. We also introduce

$$\overline{m}^{\mathcal{X}} := \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} m_x^{\mathcal{X}}. \tag{2.56}$$

The vector m satisfies the following stability and uniqueness result which is the analog of [5, Lemma 4.16]. The condition $|\operatorname{Re} z| \leq 2-\tau$, which in that paper insures that the imaginary part of m is bounded away from zero, is replaced by (2.57) in our lemma.

Lemma 2.26. *Let $\mathcal{X} \subset [N]$, $(v_x)_{x\in\mathcal{X}} \in \mathbb{R}^{\mathcal{X}}$ and $m = m^{\mathcal{X},v}$ be the solution of (2.55). Let $\tau > 0$ be such that*

$$\operatorname{Im}\Big(\frac{1}{|\mathcal{X}|}\sum_{y\in\mathcal{X}} m_y\Big) \geq \tau. \tag{2.57}$$

Assume that for two vectors $(g_x)_{x\in\mathcal{X}}, (\varepsilon_x)_{x\in\mathcal{X}} \in \mathbb{C}^{\mathcal{X}}$, the identities

$$\frac{1}{g_x} = v_x - z - \frac{1}{|\mathcal{X}|}\sum_{y\in\mathcal{X}} g_y + \varepsilon_x, \tag{2.58}$$

hold for all $x \in \mathcal{X}$. Then there are constants $b, C \in (0,\infty)$, depending only on τ, such that if $\max_{x\in\mathcal{X}}|g_x - m_x| \leq b$ then

$$\max_{x\in\mathcal{X}}|g_x - m_x| \leq C \max_{x\in\mathcal{X}} \varepsilon_x. \tag{2.59}$$

The proof of Lemma 2.26 is deferred to Section 2.8.
The condition (2.57) is here to insure that we are in the bulk. Recalling (2.28), we anticipate that our result should hold even if we remove a small number of vertices. Let us define the event

$$\Xi_1(\tau) := \Big\{ \inf_{\mathcal{X}} \inf_{z\in\mathbb{S}_{\tau,R}} \operatorname{Im} \overline{m}^{\mathcal{X}}(z) \geq \tau \Big\}, \quad \tau > 0, \tag{2.60}$$

where the first infimum is taken over all sets $\mathcal{X} \subset [N]$ such that $|\mathcal{X}^c| \leq N\exp(-(\log N)^{1/6})$ (see (2.28)). The next lemma shows that with high probability if $z \in \mathbb{S}_{\tau,R}$, then we are in the bulk of the spectrum. It even states that we can remove $o(N)$ vertices and still be in the bulk.

Lemma 2.27. *Let v_x, $x \in [N]$, be defined as in (2.4). There is a constant $\tau = \tau(R) > 0$ such that*

$$\mathbb{P}(\Xi_1(\tau)) \geq 1 - O\big(d^{-1}\big).$$

Proof. We combine Lemma 2.51 together with Lemma 2.52 below. □

Lemma 2.28. *Let $\tau > 0$ be a constant. On the event $\Xi_1(\tau)$ we have*

$$\sup_{\mathcal{X}} \sup_{x\in\mathcal{X}} \big|m_x^{\mathcal{X}} - m_x^{[N]}\big| \leq \frac{|\mathcal{X}|}{N}, \quad z \in \mathbb{S}_{\tau,R}, \tag{2.61}$$

where the first infimum is taken over all sets $\mathcal{X} \subset [N]$ such that $|\mathcal{X}^c| \leq N\exp(-(\log N)^{1/6})$.

The proof of Lemma 2.28 is deferred to the end of Section 2.8.
Let us define the event

$$\Xi_2(\tau) := \Big\{ \inf_{x\in[N]:\, v_x \geq -R-2} \inf_{z\in\mathbb{S}_{\tau,R}} \operatorname{Im}\Big(\sum_{y\in[N]} H_{xy}^2 m_y^{[N]}(z)\Big) \geq \tau, \Big\}, \quad \tau > 0. \tag{2.62}$$

The next lemma is a *local* equivalent of Lemma 2.27 around any vertex.

Lemma 2.29. *Let v_x, $x \in [N]$, be defined as in (2.4). There is a constant $\tau = \tau(R) > 0$ such that*

$$\mathbb{P}(\Xi_2(\tau)) \geq 1 - O\big(d^{-1}\big).$$

Proof. We combine Lemma 2.51 together with Lemma 2.53 below. □

Lemma 2.30. *Let $\tau > 0$ be a constant. Then there exists $D > 0$ depending only on τ such that on the event $\Xi_1(\tau) \cap \Xi_2(\tau)$, we have*

$$\inf_{\mathcal{X}} \inf_{x:\, v_x \geq -R-C} \operatorname{Im}\Big(\sum_{y\in\mathcal{X}} H_{xy}^2 m_y^{[N]}(z)\Big) \geq \tau/2,$$

where the infimum is taken over all set $\mathcal{X} \subset [N]$ that satisfy

$$\sup_{x\in[N]} \frac{\big|S_1(x)\cap\mathcal{X}^c\big|}{d} \leq \frac{1}{D}. \tag{2.63}$$

Proof. On $\Xi_1(\tau)$, we have the bound $\max_x |m_x| \leq 1/\tau = O(1)$. Therefore

$$\operatorname{Im}\Big(\sum_{y\in\mathcal{X}} H_{xy}^2 m_y^{[N]}(z)\Big) \geq \operatorname{Im}\bigg(\frac{1}{d}\sum_{y\in\mathcal{X}\cap S_1(x)} m_y^{[N]}(z)\bigg) \geq \operatorname{Im}\bigg(\frac{1}{d}\sum_{y\in S_1(x)} m_y^{[N]}(z)\bigg) - \frac{1}{\tau D} \geq \frac{\tau}{2},$$

for $D \geq 2/\tau^2$ where we used (2.62), (2.63) and the fact that $\max_{x\in[N]} \big|m_x^{[N]}\big| \leq \tau^{-1}$ on $\Xi_1(\tau)$. ☐

The following lemma is the standard starting point to prove a local law. It is a straightforward application of Lemma A.2.

Lemma 2.31 (Schur complement formula)**.** *For any $x \in [N]$ and $z \in \mathbb{C}_+$, we have*

$$\frac{1}{G_{xx}} = v_x - z - \sum_{y}^{(x)} |H_{xy}|^2 G_{yy}^{(x)} + Y_x,$$

where

$$Y_x := -\frac{f}{N} + \sum_{a\neq b}^{(x)} H_{xa} G_{ab}^{(x)} H_{by} + \frac{f}{N}\Bigg[\sum_{a,b}^{(x)} G_{ab}^{(x)} H_{by} + \sum_{a,b}^{(x)} H_{xa} G_{ab}^{(x)}\Bigg] + \frac{f^2}{N^2}\sum_{a,b}^{(x)} G_{ab}^{(x)}. \tag{2.64}$$

where we defined $f = \left(\frac{d}{(1-d/N)}\right)^{1/2}$.

Let us recall (2.24) and (2.25). The next proposition establishes bounds that are typical in the proof of local laws.

Proposition 2.32 (Main error bounds)**.** *For $\Gamma = O(1)$ the following estimates hold with with very high probability*

$$\mathbf{1}_{\tilde\Gamma\leq\Gamma} \max_{x\neq y} |G_{xy}| \leq \Gamma \frac{\mathcal{C}}{\sqrt{d}}, \tag{2.65a}$$

$$\mathbf{1}_{\tilde\Gamma\leq\Gamma} \max_{x\neq u\neq y} |G_{xy} - G_{xy}^{(u)}| \leq \Gamma \frac{\mathcal{C}}{\sqrt{d}} \tag{2.65b}$$

$$\mathbf{1}_{\tilde\Gamma\leq\Gamma} \max_{x} |Y_x| \leq \Gamma \frac{\mathcal{C}}{\sqrt{d}}, \tag{2.65c}$$

Proof. Using (2.97a) for $T = \emptyset$ (in which case $\widetilde{G}^{(T,u)} = G^{(u)}$) and (A.6) and Lemma A.1, we find, for $x \neq y$,

$$|G_{xy}| \leq |G_{xx}| \Big|\sum_{a}^{(x)} H_{xa} G_{ay}^{(x)}\Big| + |G_{xx}| \frac{f}{N} \Big|\sum_{a}^{(x)} G_{ay}^{(x)}\Big| \leq \mathcal{C}\frac{\Gamma}{\sqrt{d}}.$$

Using (A.8a) with $T=\emptyset$,we find

$$G_{xy}-G_{xy}^{(u)}=-G_{xu}\Big[\sum_a^{(u)}H_{ua}G_{ay}^{(u)}+\frac{f}{N}\sum_a^{(u)}G_{ay}^{(y)}\Big].$$

The size of the terms on the right-hand side can be bounded using Proposition 2.49 and Lemma A.1 to obtain (2.65b).

To prove (2.65c), we first use Lemma A.1

$$\frac{f^2}{N^2}\Big|\sum_{a,b}^{(x)}G_{ab}^{(x)}\Big|\le\frac{f^2}{\sqrt{N\operatorname{Im}z}}\le N^{-\kappa/6}\le\mathcal{C}d^{-3/2},$$

and, introducing $S_a:=N^{-1}\sum_b^{(x)}G_{ab}^{(x)}$ we have

$$\frac{f}{N}\Big|\sum_{ab}^{(x)}G_{ab}^{(x)}H_{ax}\Big|\le\Big|f\sum_a^{(x)}S_aH_{ax}\Big|\le|S_a|f\sum_a^{(x)}|H_{ax}|\le\Gamma\frac{f}{\sqrt{N\eta}}\frac{\log N+d}{\sqrt{d}}\le\frac{\mathcal{C}}{d}$$

where used Ward's identity to bound $S_a\le\Gamma(N\eta)^{-1/2}$ and Lemma 2.48 to bound

$$\sum_a^{(x)}|H_{ax}|\le\sum_{a\colon A_{ax}=1}^{(x)}\frac{1}{\sqrt{d}}+\sqrt{d}\le\frac{\log N}{\sqrt{d}}+\sqrt{d}.$$

We use Proposition 2.50 to get

$$\sum_{x\ne y}^{(a)}H_{ax}G_{xy}^{(a)}H_{ya}=\frac{\mathcal{C}\Gamma}{d}.$$

□

We now turn to the bootstrapping argument which is the core of the proof of Theorem 2.4. Bootstrapping is a standard technique in the proof of local laws (see for instance [14] and references therein). The bootstrapping hypothesis usually takes the form of a uniform bound on the entries of the Green function. In our proof however, we need more information. The bootstrapping argument will need the following conditions.

1. The two events Ξ_1 and Ξ_2 analysed in Lemma 2.27 and 2.30 respectively should hold, i.e. $\mathbb{S}_{\kappa,R}$ lies in the bulk spectrum.
2. The entries of the Green function should be bounded, that is $\tilde\Gamma=O(1)$ for $\tilde\Gamma$ defined in (2.54);
3. The set of atypical vertices, introduced in Definition 2.11, should be amenable to Proposition 2.12, i.e. the condition $\theta(\Gamma,p)=1$ should hold;
4. The error parameters introduced in (2.53) need *a priori* control. We want to find a subset $U\subset[N]$ such that $\Lambda_d(U)$ and $\Lambda_a(U^c)$ are small and $|U^c|=o(N)$.

We call a set $U\subset[N]$, *full* if $|U^c|\le N\exp\big(-(\log N)^{1/6}\big)$ and (2.63) holds. Note that the notion of full set depends only on the constant τ. We introduce the indicator function

$$\phi:=\phi(\Gamma,p\lambda,\tau)=\mathbf{1}_{\{\tilde\Gamma\le\Gamma\}}\theta(\Gamma,p)\mathbf{1}_{\exists U\subset[N]\text{ full }\Lambda_d(U)\vee\Lambda_a(U^c)\le\lambda},\tag{2.66}$$

as well as the event

$$\Xi(\tau):=\Xi_1(\tau)\cap\Xi_2(\tau).$$

The reason we consider ϕ and Ξ separately comes from the fact that we only have $O(d^{-1})$ bounds on $\mathbb{P}(\Xi^c)$ while bounds on ϕ are with very high probability.

Proposition 2.33. *There exists $\tau > 0$ depending on κ and R, constants $\Gamma, \lambda, p > 0$, depending only on τ and constants $\mathbf{a}, \mathcal{D} \geq 0$ depending on τ and ν such that if*

$$\mathcal{D}\sqrt{\log N} \leq d \leq (\log N)^{3/2},$$

then for all $z \in \mathbb{S}_{\kappa,R}$, there exists $T \subset [N]$ which is full such that

$$\mathbf{1}_{\Xi(\tau)}\phi(2\Gamma, 2p, \lambda, \tau) \leq \theta(\Gamma, p), \quad \mathbf{1}_{\Xi(\tau)}\phi(2\Gamma, 2p, \lambda, \tau)\big(\Lambda_d(T) + \Lambda_o + \Lambda_a([N])\big) \leq C_*\varphi_{\mathbf{a}}, \tag{2.67}$$

holds with very high probability for some constant $C_* = C_*(\tau) > 0$.

Proof. Let $\tau > 0$ be defined so that $\Xi_1(\tau) \cap \Xi_2(\tau)$ both hold and choose $\Gamma = 2/\tau$. Let $p, q > 0$ be chosen from Γ and $\mathbf{a}$ be chosen from q and ν so that Proposition 2.12 holds for $2p, q$ and such that Lemma 2.22 holds for p and Γ.
Then we see that $\mathcal{T}_{\mathbf{a}}$ as defined in (2.22) is $q-$full since by (2.26) and (2.27) respectively

$$\frac{|\mathcal{T}_{\mathbf{a}}^c|}{N} \leq \frac{\mathrm{e}^{2q\varphi_{\mathbf{a}}^2 d}}{N} + \mathrm{e}^{-2q\varphi_{\mathbf{a}}^2 d} \leq \mathrm{e}^{-q\sqrt{\log N}}, \quad \sup_{x\in[N]} \frac{\big|S_1(x) \cap \mathcal{T}_{\mathbf{a}}^c\big|}{d} \leq 10\varphi_{\mathbf{a}} \leq 10\mathcal{D}^{-1/3}, \tag{2.68}$$

with very high probability, where we used the fact that $\varphi_{\mathbf{a}}^2 d \leq \frac{\mathbf{a}}{\mathcal{D}^{1/3}}\sqrt{\log N}$ and we chose $\mathcal{D} \geq \mathbf{a}^3$. Using Proposition 2.12 (ii) and (2.65c) we find that, with very high probability, there exist $\varepsilon_x \leq 2\varphi_{\mathbf{a}}$, $x \in \mathcal{T}_{\mathbf{a}}$, such that

$$\begin{aligned} \frac{1}{G_{xx}} &= v_x - z - \sum_y^{(x)} H_{xy}^2 G_{yy}^{(x)} + Y_x \\ &= v_x - z - \frac{1}{|\mathcal{T}_{\mathbf{a}}|}\sum_{y\in\mathcal{T}_{\mathbf{a}}} G_{yy} + \Psi_x + Y_x + O\bigg(\frac{|\mathcal{T}_{\mathbf{a}}^c|}{N|\mathcal{T}_{\mathbf{a}}|}\bigg) \\ &= v_x - z - \frac{1}{|\mathcal{T}_{\mathbf{a}}|}\sum_{y\in\mathcal{T}_{\mathbf{a}}} G_{yy} + \varepsilon_x. \end{aligned} \tag{2.69}$$

Here we used the fact that $\mathcal{C}d^{-1/2} \leq \varphi_{\mathbf{a}}$ since $d \leq (\log N)^{3/2}$.
Let U be a $q-$full set that satisfies $\Lambda_a(U^c), \Lambda_d(U) \leq \lambda$. Such a set exists by definition of ϕ, see (2.66). Let $T_1 := \mathcal{T}_{\mathbf{a}} \cap U$ and $T_2 := \mathcal{T}_{\mathbf{a}} \setminus U$ and $T := T_1 \cup T_2$. We will show that

$$\Lambda_d(T_1), \quad \Lambda_d(T_2), \quad \Lambda_a([N]) \leq C\varphi_{\mathbf{a}}, \qquad C = C(\tau) > 0. \tag{2.70}$$

Let us conclude the proof using (2.70). On the event $\theta(2\Gamma, 2p)$ we know by Lemma 2.22, that $\theta(2\Gamma, \frac{\mathcal{C}\Gamma}{\sqrt{d}})$ holds with very high probability. Using (2.65a) we see that $\Lambda_o \leq 2\Gamma\mathcal{C}d^{-1/2} \leq \varphi_{\mathbf{a}}$ holds with very high probability. Moreover since on $\Xi(\tau)$ we have $\operatorname{Im} \frac{1}{d}\sum_{y\in[N]} A_{xy} m_y \geq \tau$, we find that

$$|G_{xx}| \leq \frac{1}{\tau} + O\big(\varphi_{\mathbf{a}}\big) \leq \frac{2}{\tau} = \Gamma.$$

By Lemma 2.23 since $\tilde{\Gamma} \leq 2\Gamma = O(1)$, $\theta_i(\Gamma)$ holds. We conclude that $\theta(\Gamma, \frac{\mathcal{C}\Gamma}{\sqrt{d}}) \leq \theta(\Gamma, p)$ holds. Thus it suffices to establish (2.70) to conclude the proof.

We start with $\Lambda_d(T_1)$. By definition ϕ and (2.68) we see that $|T_1^c| \leq |T^c| + |U^c| \leq 2\mathrm{e}^{q\sqrt{\log N}}$ and $\frac{|T_1|}{|T|} = 1 - \frac{|T\setminus U|}{|T|} \geq 1 - O\big(N^{-1/2}\big)$. Plugging this into (2.69), we find

$$\frac{1}{G_{xx}} = v_x - z - \frac{1}{|T_1|}\sum_{y\in T_1} G_{yy} + \varepsilon_x, \quad x \in T_1. \tag{2.71}$$

By Corollary 2.28, since $|T_1^c| \leq \sqrt{N}$ we know that $\operatorname{Im} \overline{m}^{T_1} \geq \tau$. We want to apply Lemma 2.26 to $\mathcal{X} = T_1$ with $\max_x \varepsilon = O(\varphi_{\mathbf{a}})$. By definition of ϕ, we know that $\max_{x \in T_1} |G_{xx} - m_x| \leq \lambda$ and by Lemma 2.28 we know that

$$\max_{x \in T_1} |m_x^{[N]} - m_x^{T_1}| \leq \frac{|T_1^c|}{N} \leq \exp(-(\log N)^{1/6}) \leq \varphi_{\mathbf{a}}. \tag{2.72}$$

Therefore choosing λ as $b/2$ from Lemma 2.26, we have

$$\max_{x \in T_1} |G_{xx} - m_x^{T_1}| \leq \max_{x \in T_1} |m_x^{[N]} - m_x^{T_1}| + \max_{x \in T_1} |G_{xx} - m_x^{[N]}| \leq 2\lambda \leq b,$$

and we can conclude that $\max_{x \in T_1} \left| G_{xx} - m_x^{[T_1]} \right| \leq C\varphi_{\mathbf{a}}$ for some $C > 0$ depending on τ. Using one more time (2.72), we see that $\Lambda_d(T_1) \leq C\varphi_{\mathbf{a}}$. This concludes the estimate of $\Lambda_d(T_1)$ in (2.70).

We now estimate $\Lambda_d(T_2)$. Using the fact that $\max_{x \in [N]} |G_{xx}| \leq 2\Gamma$ we see that

$$\frac{1}{|T|} \sum_{x \in T} G_{xx} - \frac{1}{|T_1|} \sum_{x \in T_1} G_{xx} = O\left(\frac{\Gamma |T_2|}{|T||T_1|} \right) \leq N^{-1/2}. \tag{2.73}$$

Using (2.73) and $\Lambda_d(T_1) = O(\varphi_{\mathbf{a}})$, we see that

$$\operatorname{Im}\left(\frac{1}{|T|} \sum_{x \in T} G_{xx} \right) \geq \operatorname{Im} \overline{m}^{T_1} - O\left(\varphi_{\mathbf{a}} + \frac{\Gamma |T_2|}{|T||T_1|} + \varepsilon_x \right) \geq \frac{1}{2\tau}, \tag{2.74}$$

holds with very high probability. For $x \in T_2$, using (2.69), the fact that $\operatorname{Im} \overline{m}^{[N]} \geq \tau$ on Ξ, (2.73) and (2.74), we find that

$$\begin{aligned} \left| G_{xx} - m_x^{[N]} \right| &= \left| \frac{1}{v_x - z - \frac{1}{|T|} \sum_{y \in T} G_{yy} + \varepsilon_x} - \frac{1}{v_x - z - \overline{m}^{[N]}} \right| \\ &\leq \frac{2}{\tau^2} \left| \overline{m}^{[N]} - \frac{1}{|T|} \sum_{y \in T} G_{yy} + \varepsilon_x \right| \leq \frac{2}{\tau^2} \left| \overline{m}^{[N]} - \overline{m}^{T_1} + \Lambda_d(T_1) + 2\varepsilon_x \right| \leq C\varphi_{\mathbf{a}}, \end{aligned}$$

holds with very high probability with $C = C(\tau) > 0$. This shows that $\Lambda_d(T_2) = O(\varphi_{\mathbf{a}})$.

Note that using $\Lambda_d(T_1 \cup T_2) = O(\varphi_{\mathbf{a}})$ and (2.68), we have

$$\frac{1}{N} \sum_{x \in [N]} G_{xx} = \overline{m}^{[N]} + O(\varphi_{\mathbf{a}}), \tag{2.75}$$

and in particular $\operatorname{Im} \frac{1}{N} \sum_{x \in [N]} G_{xx} \geq \tau/2$ on $\Xi_1(\tau)$.

We now turn to atypical vertices and the estimate of $\Lambda_a(T^c)$. For $x \notin T$, we find from (A.2) that

$$\begin{aligned} \frac{1}{G_{xx}} &= v_x - z - \sum_{y}^{(x)} H_{xy}^2 G_{yy}^{(x)} + O(\varphi_{\mathbf{a}}) = v_x - z - \sum_{y \in \mathcal{T}_{\mathbf{a}}}^{(x)} H_{xy}^2 G_{yy}^{(x)} + O(\varphi_{\mathbf{a}}) \\ &= v_x - z - \sum_{y \in \mathcal{T}_{\mathbf{a}}}^{(x)} H_{xy}^2 m_y + O((1 + \alpha_x)\varphi_{\mathbf{a}}) = v_x - z - \sum_{y}^{(x)} H_{xy}^2 m_y + O((1 + \alpha_x)\varphi_{\mathbf{a}}). \end{aligned} \tag{2.76}$$

In the second and fourth equalities, we used the bound on $\max_x \Big(|m_x| \vee G_{xx} \Big) \leq \Gamma$. We now proceed as in the proof of Lemma 2.23 and distinguish between two cases.

In the first case, if $\alpha_x \geq 2$, we deduce that $v_x \geq \sqrt{d}$ and, using the identity $\alpha_x = \frac{v_x}{\sqrt{d}} + 1$,

$$\left| v_x - z - \sum_{y}^{(x)} H_{xy}^2 G_{yy}^{(x)} \right| \geq |v_x| \left(1 - \frac{2\Gamma}{\sqrt{d}} \right) - 1 - R \geq \frac{\sqrt{d}}{2}.$$

Using (2.75), we get

$$\left|G_{xx} - \frac{1}{v_x - z - \sum_y^{(x)} H_{xy}^2 m_y}\right| = O\left(\frac{(1+\alpha_x)\varphi_{\mathbf{a}}}{|v_x|\tau}\right) = O(\varphi_{\mathbf{a}}).$$

If $\alpha_x \le 2$, we use (2.68) and $\Xi_2(\tau)$ to see that

$$\sum_{y \in \mathcal{T}_{\mathbf{a}}} H_{xy}^2 m_y^{[N]} \ge \tau/2.$$

Moreover $\operatorname{Im} \overline{m}^{[N]} \ge \tau$ on $\Xi_1(\tau)$. Using those two lower bounds, we can invert (2.76) to find that for $x \notin T$ with $\alpha_x \le 2$, we have

$$\left|G_{xx} - \frac{1}{v_x - z - \sum_{y \in \mathcal{T}_{\mathbf{a}}} H_{xy}^2 m_y}\right| \le \frac{4\varphi_{\mathbf{a}}}{\tau^2}.$$

Now observe that $H_{xy}^2 = \frac{1}{d} A_{xy} + O(dN^{-2})$. We deduce that $\Lambda_a(T^c) = O(\varphi_{\mathbf{a}})$. Note that the same argument shows that $\lambda_a(T) = O(\varphi_{\mathbf{a}})$ from which we get (2.70). □

Proof of Theorem 2.4. Let us choose $\tau > 0$ as in Lemmas 2.27 and 2.29 such that $\Xi(\tau) := \Xi_1(\tau) \cap \Xi_2(\tau)$ holds with probability $1 - O(d^{-1})$.
Let $E \in [-R, R]$, $L \ge 1$ and $z_k := E + i\eta_k$, $\eta_k = L - kN^{-3}$ for $k = 0, \dots, k_*$ with $k_* := \inf\{k \in \mathbb{Z}, \eta_k \le N^{-1+\kappa}\}$. We introduce the events

$$\begin{aligned}
&\Omega_k(\Gamma, \lambda) := \big\{ \max\big(\Lambda_o(z_k), \Lambda_a([N])(z_k), \Lambda_d(T)(z_k)\big) \le \lambda, \text{ for some full set } T \subset [N]\big\},\\
&\Upsilon_k(\Gamma, p) := \{\theta(\Gamma, p) = 1\}, \qquad \Gamma, \Lambda, q, p > 0.
\end{aligned}$$

By Proposition 2.33, we know that there exists $C_* = C_*(\tau) > 0$ and constant $\Gamma, p, q, \mathbf{a}$, all depending, *in fine*, only on τ, such that, for any $\nu \ge 0$,

$$\mathbb{P}\big[\big(\Upsilon_k(\Gamma, p) \cap \Omega_k(\Gamma, C_*\varphi_{\mathbf{a}})\big)^c \cap \Omega_k(2\Gamma, 2p) \cap \Xi(\tau)\big] \le \mathcal{C}_\nu N^{-\nu},$$

for some $\mathcal{C}_\nu \ge 0$.
Moreover, by Lipschitz continuity of the entries of G and m (see the remark after Lemma 2.51), we know that

$$\Upsilon_k(\Gamma, p) \le \Upsilon_{k+1}(2\Gamma, 2p), \quad \Omega_k(\Gamma, \lambda) \le \Omega_{k+1}(2\Gamma, 2\lambda),$$

for any constants Γ, p and $\lambda \ge \frac{1}{N}$.

Finally observe that choosing $L \ge 0$ large enough in the definition of η_0 insures that $\Omega_0(2\Gamma, 2p, q)$ holds deterministically since then all quantities are smaller than L^{-1}. We can now conclude, by a standard induction argument that there exists constants $\Gamma, \mathbf{a}, q$ all depending on τ such that, for any $E \in [-R, R]$,

$$\prod_{k=1}^{k_*} \Omega_k(2\Gamma, 2C_*\varphi_{\mathbf{a}}) \mathbf{1}_{\Xi(\tau)} = \mathbf{1}_{\Xi(\tau)}$$

with very high probability.

Using the Lipschitz continuity and a grid of mesh size N^{-3}, we can extend this result to get

$$\begin{aligned}
&\mathbb{P}\Big[\Lambda_o(z) + \Lambda_a([N])(z) \le 4C_*\varphi_{\mathbf{a}},\, \tilde{\Gamma}(z) \le 4\Gamma,\, \forall z \in \mathbb{S}_{\kappa,R}\Big]\\
&\le \mathbb{P}\big[\Xi(\tau)^c\big] + 2RN^3\mathbb{P}\left[\left(\bigcap_{k=0}^{k_*} \Lambda_o(E + \mathrm{i}\eta_k) + \Lambda_a([N])(E + \mathrm{i}\eta_k) \le 4C_*\varphi_{\mathbf{a}},\, \tilde{\Gamma}(E + \mathrm{i}\eta_k) \le 4\Gamma, \Xi(\tau)\right)^c\right]\\
&\le Cd^{-1} + \mathcal{C}_\nu N^{-\nu+6} = O(d^{-1}).
\end{aligned}$$

Choosing $D = \mathcal{D}_{10}$ we conclude the proof. □

2.5 Proof of Theorem 2.6

In this section, we prove Theorem 2.6. We write

$$\tau := \tau(R,\kappa) = \inf_{z\in\mathbb{S}_{\kappa,R}} \operatorname{Im} \mathrm{m_{fc}}(z).$$

By Lemma 2.54, we know that $\tau \geq \frac{1}{100}\mathrm{e}^{-R^2/2}$. Recalling the definition of g from (2.8) we define the two quantities

$$\Lambda := \left| G_{xy} - \frac{\delta_{xy}}{v_x - z - \mathrm{m_{fc}}(z)} \right|, \quad \Theta := |g(z) - \mathrm{m_{fc}}(z)|.$$

For $v = (v_x)_{x\in[N]} \in \mathbb{R}^N$, we define

$$\hat{\mathrm{m}}_{\mathrm{fc}}(z,v) := \frac{1}{N}\sum_{i=1}^{N} \frac{1}{-v_i - z - \mathrm{m_{fc}}(z)}. \tag{2.77}$$

The quantity $\hat{\mathrm{m}}_{\mathrm{fc}}$ is a good approximation of $\mathrm{m_{fc}}$ as the following lemma suggests.

Lemma 2.34. *Let v_x be as in (2.4) and $\mathbb{S}_{\tau,R}$ as in (2.7). There exists $\mathcal{D} \geq 0$, depending on the notion of very high probability, such that for $\mathcal{D}\log N \leq d \leq \sqrt{N}$, then*

$$\left| \hat{\mathrm{m}}_{\mathrm{fc}}(z,v) - \mathrm{m_{fc}}(z) \right| \leq \mathcal{C}\sqrt{\frac{\log N}{d}}, \quad z \in \mathbb{S}_{\tau,R},$$

holds with very high probability.

The proof of Lemma 2.34 is a simple application of McDiarmid inequality and is given in Section 2.9. The next lemma is the analog of Lemma 2.26 and is stated in [30, Proposition 3.5].

Lemma 2.35. *There exists $c_* > 0$ depending only on R and κ such that if $w_1, \dots, w_N \in \mathbb{H}$ satisfy*

$$\left| \frac{1}{N}\sum_{i\in[N]} \frac{1}{w_i - z - \mathrm{m_{fc}}(z)} - \mathrm{m_{fc}}(z) \right| \leq c_*,$$

then

$$c_* \leq \left| 1 - \frac{1}{N}\sum_{i\in[N]} \frac{1}{(w_i - z - \mathrm{m_{fc}}(z))^2} \right| \leq 1 + c_*.$$

Proof. Since $R > 0$ is of order 1, by Lemma 2.54 there exists $\tau > 0$ such that $\operatorname{Im} \mathrm{m_{fc}}(z) \geq \tau$ for every $z \in \mathbb{S}_{\kappa,R}$. Following the proof in [30, Proposition 3.5] and writing $c_* = \min(c_1, \tau^2/16, 1/\tau)$, with c_1 is defined therein, we conclude. □

We also have the following analog to Proposition 2.32. Note that (2.78d) does not have an equivalent in Proposition 2.32. Instead, it is equivalent to the statement $\max_x \Psi_x = o(1)$. This is only true for $d \gg \log N$.

Proposition 2.36 (Error bounds)**.** *Let $\Gamma > 0$. There exists $\mathcal{C} \geq 0$ depending on ν such that for $d \geq \mathcal{C}\log N$, on the event $\tilde{\Gamma} \leq \Gamma$ we have*

$$\mathbf{1}_{\tilde{\Gamma}\leq\Gamma} \max_{x\neq y} |G_{xy}| \leq \Gamma\frac{\mathcal{C}}{\sqrt{d}}, \tag{2.78a}$$

$$\mathbf{1}_{\tilde{\Gamma}\leq\Gamma} \max_{x\neq a\neq y} |G_{xy} - G_{xy}^{(a)}| + |G_{xy}^{(a)} - \widetilde{G}_{xy}^{(a)}| \leq \Gamma\frac{\mathcal{C}}{\sqrt{d}} \tag{2.78b}$$

$$\mathbf{1}_{\tilde{\Gamma}\leq\Gamma} \max_{x} |Y_x| \leq \Gamma\frac{\mathcal{C}}{\sqrt{d}}, \tag{2.78c}$$

$$\mathbf{1}_{\tilde{\Gamma}\leq\Gamma} \sum_{y}^{(x)} \Big(H_{xy}^2 - \frac{1}{N}\Big) G_{yy}^{(x)} \leq \Gamma\mathcal{C}\sqrt{\frac{\log N}{d}} \tag{2.78d}$$

Proof. The proof of (2.78a)-(2.78c) is the same as in Proposition 2.32. To prove (2.78d), we expand $G_{yy}^{(x)}$ using (A.4) and we find

$$\sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})G_{yy}^{(x)} = \sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})\widetilde{G}_{yy}^{(x)} + \sum_{l=1}^{M}\sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})S_l(y,y)$$
$$+ \sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})\sum_a^{(x)} S_M(y,a)H_{ax}G_{ay}^{(x)},$$

Let us fix $M = 10$. Using Proposition 2.38 (2.83c), we see that

$$\Big|\sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})\widetilde{G}_{yy}^{(x)}\Big| \le \Gamma\mathcal{C}\sqrt{\frac{\log N}{d}},$$

with very high probability.
Using Lemma 2.48, we know that $\sum_y^{(x)}\big(H_{xy}^2 - \frac{1}{N}\big) \le \mathcal{C}$, with very high probability. Moreover by Proposition 2.39,

$$\max_{y\in[N]}|S_l(y,y)| \le \Big(\frac{\mathcal{C}\Gamma}{\sqrt{d}}\Big)^l, \qquad l \ge 1, \tag{2.79}$$

from which we conclude that

$$\sum_{l=1}^{M}\sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})S_l(y,y) \le \frac{\mathcal{C}\Gamma}{\sqrt{d}}.$$

Again using (2.79), we find that

$$\Big|\sum_y^{(x)}\Big(H_{xy}^2 - \frac{1}{N}\Big)\sum_a^{(x)} S_M(y,a)H_{ax}G_{ay}^{(x)}\Big| \le \frac{\Gamma}{d}\Big(\frac{\mathcal{C}\Gamma}{\sqrt{d}}\Big)^M \sum_a^{(x)}|H_{xa}| \le \frac{\Gamma(\log N + d)}{d^{3/2}}\Big(\frac{\mathcal{C}\Gamma}{\sqrt{d}}\Big)^M \sum_a^{(x)} \le \Gamma\frac{\mathcal{C}}{\sqrt{d}},$$

holds with very high probability. This concludes the proof of (2.78d). □

Proposition 2.37 (Bootstrap in dense regimes)**.** *There exists $\Gamma, \lambda > 0$, depending only on R and κ, such that for every $z \in \mathbb{S}_{\kappa,R}$, the following holds with very high probability*

$$\mathbf{1}_{\tilde\Gamma\le 2\Gamma,\Theta\le\lambda}\tilde\Gamma \le \Gamma, \quad \mathbf{1}_{\tilde\Gamma\le 2\Gamma,\Theta\le\lambda}(\Lambda\vee\Theta) \le \mathcal{C}\sqrt{\frac{\log N}{d}}.$$

Proof. Let us choose $\Gamma = 2/\tau$ and $\lambda \le c_*$ with c_* defined in Lemma 2.35 and write $\varphi := \sqrt{\frac{\log N}{d}}$. We start by improving the bound on Θ to $\Theta \le \varphi$. Recalling (2.64), we define

$$Z_x := Y_x + \sum_y^{(x)}(H_{xy}^2 - \frac{1}{N})G_{yy}^{(x)}.$$

Using Lemma 2.31 and (2.77), we find

$$g(z) - \mathrm{m}_{\mathrm{fc}}(z) = \frac{1}{N}\sum_x \frac{1}{v_x - z - g(z) + Z_x} - \frac{1}{v_x - z - \mathrm{m}_{\mathrm{fc}}(z)} + \hat{\mathrm{m}}_{\mathrm{fc}}(z) - \mathrm{m}_{\mathrm{fc}}(z). \tag{2.80}$$

Using $\Theta \le \lambda$ and Lemma 2.54, we find that

$$\mathrm{Im}(v_x - z - g(z) + Z_x) \ge \mathrm{Im}(\mathrm{m}_{\mathrm{fc}}) - \lambda - \varphi \ge \tau/2, \tag{2.81}$$

for $\tau = \mathrm{e}^{-R^2/2}/10$, $\lambda \le \tau/3$ and $d \ge 10 \log N$. Using Lemma 2.35 we see that for $\lambda > 0$ small enough, depending only on τ, we have

$$\left|1 - \frac{1}{N}\sum_x \frac{1}{(v_x - z - \mathrm{m_{fc}}(z))^2}\right| \ge c_*.$$

By Lemma 2.34, we know that $|\,\mathrm{m_{fc}}(z) - \hat{\mathrm{m}}_{\mathrm{fc}}(z)| \le \mathcal{C}\varphi$ holds with very high probability. We conclude that (2.80) implies that

$$\frac{c_*}{2}\Theta \le \left|1 - \frac{1}{N}\sum_x \frac{1}{(v_x - z - \mathrm{m_{fc}}(z))^2} - \frac{8\max_x|Z_x|}{\tau^3}\right|\Theta \le \frac{4\max_x|Z_x|}{\tau^2} + |\,\mathrm{m_{fc}}(z) - \hat{\mathrm{m}}_{\mathrm{fc}}(z)| \le \mathcal{C}\varphi,$$

holds with very high probability. Multiplying the above by $2/c_*$ we deduce that $\Theta \le \mathcal{C}\varphi$ holds with very high probability.

The bound on G_{xy} for $x \ne y$ follows from (2.78a). For the diagonal terms we have, using again Lemma 2.31, we prove that

$$\left|G_{xx} - \frac{1}{v_x - z - \mathrm{m_{fc}}(z)}\right| = \left|\frac{g(z) - \mathrm{m_{fc}}(z) + Z_x}{(v_x - z - g(z) + Z_x)(v_x - z - \mathrm{m_{fc}}(z))}\right| \le \Gamma\frac{\mathcal{C}}{\tau^2}\varphi,$$

holds with very high probability, where we used (2.81). Using (2.78b) and $\operatorname{Im}\mathrm{m_{fc}}(z) \ge \tau$, we see that $\tilde\Gamma \le \frac{1}{\tau} + \mathcal{C}\Gamma d^{-1/2} \le \Gamma$ with very high probability. This concludes the proof. □

Proof of Theorem 2.6. We proceed by induction on $z_k = E + i\eta_k$ for $|E| \le R$ and $\eta_k = L - \mathrm{i}kN^{-3}$ for $L \ge 1$ chosen large enough later and $1 \le k \le k^*$. Here $k^* := \sup\{k \in \mathbb{N},\ \eta_k \in S_{\tau,R}\}$. Throughout this proof we will use $\Gamma = \frac{2}{\tau}$ for some large enough constant $C > 0$. For z_0, we immediately have $\tilde\Gamma \le \Gamma$. Furthermore we have for $L \ge 2\lambda^{-1}$ we immediately have that $\Theta(z_0) \le |g(z_0)| + |\mathrm{m_{fc}}(z_0)| \le 2L^{-1} \le \lambda$. The induction hypothesis for z_0 is thus fulfilled. For the step $k \to k+1$, we use the Lipschitz continuity, the fact that $\varphi = o(1)$ and $|v_x - z - \mathrm{m_{fc}}(z)|^{-1} \le \tau^{-1}$, to show that

$$\mathbf{1}_{\{\Theta(z_k) \le \min(\lambda/2, \tau/2)\}} \ge \mathbf{1}_{\{\Theta(z_{k+1}) \le \lambda, \tilde\Gamma \le 2\Gamma\}},$$

which allows the induction to work.
We conclude the proof by noting that the intersection of $O(N^3)$ very high probability events is still a very high probability event. □

2.6 Large deviations estimates

In this section, we derive large deviation estimates for multilinear forms of sparse random vectors with independent components. The results here are independent from the rest of the paper.
In this section, we consider $N \in \mathbb{N}^*$ and $q = q(N) \ge 1$ and X_i, $i \in [N]$, to be independent, centered random variables satisfying

$$\mathbb{E}|X_i|^2 = \frac{1}{N}, \quad \mathbb{E}|X_i|^k \le \frac{1}{Nq^{k-2}}. \tag{2.82}$$

We also introduce $a_i, a_{ij} \in \mathbb{C}$, $i, j \in [N]$ to be deterministic complex numbers.

The following results were probed in [28, Section 3].

Proposition 2.38 (Proposition 3.1-3.2 of [28])**.** *Let $N, q \ge 0$, X_i and a_{ij}, a_i be defined as in (2.86). Suppose that*

$$\left(\frac{1}{N}\sum_i |a_i|^2\right)^{1/2} \le \gamma, \quad \frac{\max_i|a_i|}{q} \le \psi,$$

for some $\gamma, \psi \geq 0$. *Then*

$$\Big\|\sum_i a_i X_i\Big\|_r \leq \Big(\frac{2r}{1+2(\log(\psi/\gamma))_+}\Big)(\gamma \vee \psi), \tag{2.83a}$$

$$\Big\|\sum_i a_i(|X_i|^2 - \mathbb{E}|X_i^2|)\Big\|_r \leq 2\Big(1+\frac{2q^2}{N}\Big)\max_i|a_i|\Big(\frac{r}{q^2} \vee \sqrt{\frac{r}{q^2}}\Big), \tag{2.83b}$$

If γ *and* ψ *satisfy*

$$\Big(\max_i \frac{1}{N}\sum_j a_{ij}^2\Big)^{1/2} \vee \Big(\max_j \frac{1}{N}\sum_i a_{ij}^2\Big)^{1/2} \leq \gamma, \quad \max_{i\neq j}|a_{ij}| \leq \psi,$$

then

$$\Big\|\sum_{i\neq j} X_i a_{ij} X_j\Big\|_r \leq \frac{\Gamma}{d}\Big(\frac{r}{1+\log(\psi/\gamma)}\Big)^2. \tag{2.83c}$$

In [30, Section 3.3], the following quantities were introduced

$$S_l := \sum_{i_1,\dots,i_l} X_{i_1} a_{i_1 i_2} X_{i_2} \dots a_{i_{l-1} i_l} X_{i_l}, \tag{2.84a}$$

$$S_l(i) := \sum_{i_1,\dots,i_l} a_{i i_1} X_{i_1} a_{i_1 i_2} \dots a_{i_{l-1} i_l} X_{i_l}, \tag{2.84b}$$

$$S_l(i,j) := \sum_{i_1,\dots,i_l} a_{i i_1} X_{i_1} a_{i_1 i_2} \dots a_{i_{l-1} i_l} X_{i_l} a_{i_l j}. \tag{2.84c}$$

Note that (2.83a) is a bound on $\|S_1(j)\|_r$ if we set $a_{ij} \equiv a_i$ for some $j \in [N]$. The next result are an amelioration of [30, Proposition 3.1] that controls the L^r norm of S_l, $S_l(i)$ and $S_l(i,j)$.

Let us introduce the main error parameter

$$\mathcal{E} := \mathcal{E}(r,l,\phi,\gamma,q) = \left[\frac{1}{q^l}\max\left(\psi^{l-1}\left(\frac{rl}{1+\log(\psi/\gamma q^4)}\right)^l \vee \frac{r\psi}{1+\log(\psi/\gamma q^4)}\right)\right] \vee r^l\left(\frac{\gamma}{\psi}\right)^{l/4} \tag{2.85}$$

Proposition 2.39. *Let* X_i *be as in (2.82) and suppose* $(a_{ij})_{i,j}$ *are complex number that satisfy*

$$\left(\max_i \frac{1}{N}\sum_j a_{ij}^2\right)^{1/2} \vee \left(\max_j \frac{1}{N}\sum_i a_{ij}^2\right)^{1/2} \leq \gamma, \quad \max_{i\neq j}|a_{ij}| \leq \psi\Gamma, \quad \max_{i\in[N]}|a_{ii}| \leq \Gamma. \tag{2.86}$$

For $l \geq 2$ *we have*

$$\|S_l\|_r, \leq (8\Gamma)^l \mathcal{E} + \frac{(8\Gamma)^l r}{q^{l-1}}, \tag{2.87a}$$

$$\|S_l(i)\|_r, \|S_l(i,j)\|_r \leq (8\Gamma)^l \mathcal{E}. \tag{2.87b}$$

Remark 2.40. In the case where $\gamma \leq \psi \vee \frac{q^4}{\psi}$ and $l = O(1)$, $\mathcal{E}$ simplifies to

$$\mathcal{E} \leq \left(\frac{8\Gamma}{q}\left(\frac{rl}{1+(\log(\psi/q^4\gamma))_+} \vee 1\right)\right)^l \psi. \tag{2.88}$$

In some cases, we might need to consider cases where $l \gg 1$. In this case, we can use the factor ψ^{l-1} in (2.85) to offset the growth of l. A typical use case could be $l = \log N$ and $\psi = O\big((\log N)^{-1/2}\big)$.

Before we start the proof of Proposition 2.39, let us introduce the tools we will use. The study of large powers of polynomials arises often in random matrix theory. In our case we are interested in understanding the expression

$$\mathbb{E}(S_l)^r = \sum_{i_1,\dots i_{rl}\in[N]} \prod_{i=0}^{r-1}\prod_{j=1}^{l} a_{j+l(i-1)j+1+l(i-1)}\mathbb{E}\left[X_{i_1}\dots X_{i_{rl}}\right]. \tag{2.89}$$

An approach commonly used is to use so-called *computation graphs*.

Definition 2.41 (Computation graph)**.** Let $G=(V,E)$ be a finite graph. We define

$$\mathrm{Val}(G) := \frac{1}{N^{|V|}} \sum_{s\in[N]^V} \prod_{(i,j)\in E} |a_{s_i s_j}|. \tag{2.90}$$

Let Π be a partition of V. We define G_Π to be the graph whose vertex set are the block of Π and whose edges are $e_{\pi_x\pi_y}$. We denote the edge set of G_Π by $E(G,\Pi)$ Similarly to (2.90) we define

$$\mathrm{Val}(G,\Pi) := \frac{1}{N^{|\Pi|}} \sum_{s\in[N]^\Pi} \prod_{(\pi_1,\pi_2)\in E(G,\Pi)} |a_{s_{\pi_1} s_{\pi_2}}|. \tag{2.91}$$

Using this notion of computation graph and viewing $\{i_1,\dots,i_{rl}\}$ as V and using the independence of X_i and the bound $\mathbb{E}|X_i|^k \leq N^{-1}q^{2-k}$, we can transform (2.89) into

$$\mathbb{E}(S_l)^r = \sum_G \mathrm{Val}(G)\frac{1}{q^{|V(G)|}},$$

where the sum is taken among all graphs obtained by partitioning $[rl]$. As a first observation, we see, using the fact that the variables X_i are centered and independent if a block the partition of $[rl]$ has only one element, then the contribution of this graph is null.
Given the particular structure of S_l as defined in (2.84) the appropriate representation uses so-called *line graphs*.

Definition 2.42. (Line graphs and graphs induced by partitions) Let $l\in\mathbb{N}^*$. We define the *l-line graph* to be the graph $L_{[l]} := ([l],\{(i,i+1): 1\leq i<l\})$. The $l-1$ edges of $L_{[l]}$ are naturally indexed by $e_i=(i,i+1)$, $i\in[l-1]$.
Let $\alpha\subset[l-1]$. We define the (l,α)-line graph, denoted by $L_{[l]}^{(\alpha)}$, to be the graph obtained by partitioning a $L_{[l]}$ by the equivalence relation $x\sim_\alpha y$ if and only if $(xy)\in\alpha$.
For $r\in\mathbb{N}^*$, we define $\mathcal{P}_{\geq 2,\alpha}$ to be the set of partition of $[r]\times[l]$ satisfying

(i) Each block $\pi\in\Pi$ has size at least 2;

(ii) The partition must preserve the structure induced by α on every subset $\{i\}\times[l]$, $1\leq i\leq r$.

The following lemma relates the topology of a graph G and its value in the context of Definition 2.42. It is the key result of this section.

Lemma 2.43. *Let $r,l\geq 1$, $\alpha\subsetneq[l-1]$ and $G:=\sqcup_{i=1}^r L_{[l]}$ graph. Then, for $\Pi\in\mathcal{P}_{\geq 2,\alpha}$ we have*

$$\mathrm{Val}(G,\Pi) \leq \Gamma^{rl}\psi^{(r(l-|\alpha|-1))-|\Pi|/2}\gamma^{|\Pi|/2}. \tag{2.92}$$

Proof. Let $\Pi\in\mathcal{P}_{\geq 2,\alpha}([rl])$, G_Π as defined in Definition 2.41. Let T be a spanning forest of G_Π. We enumerate the vertices of T in such a way that π_1 is a leaf of T and π_2 is its neighbor in T. Edges in $E(G_\Pi)\setminus E(T)$ are either *loops*, meaning edges with of the form (x,x) for $x\in V(G_\Pi)$, or *extra edges*. We proceed as in [30, Lemma 3.13] to associate with every edge of T a factor γ using the Cauchy-Schwartz inequality.

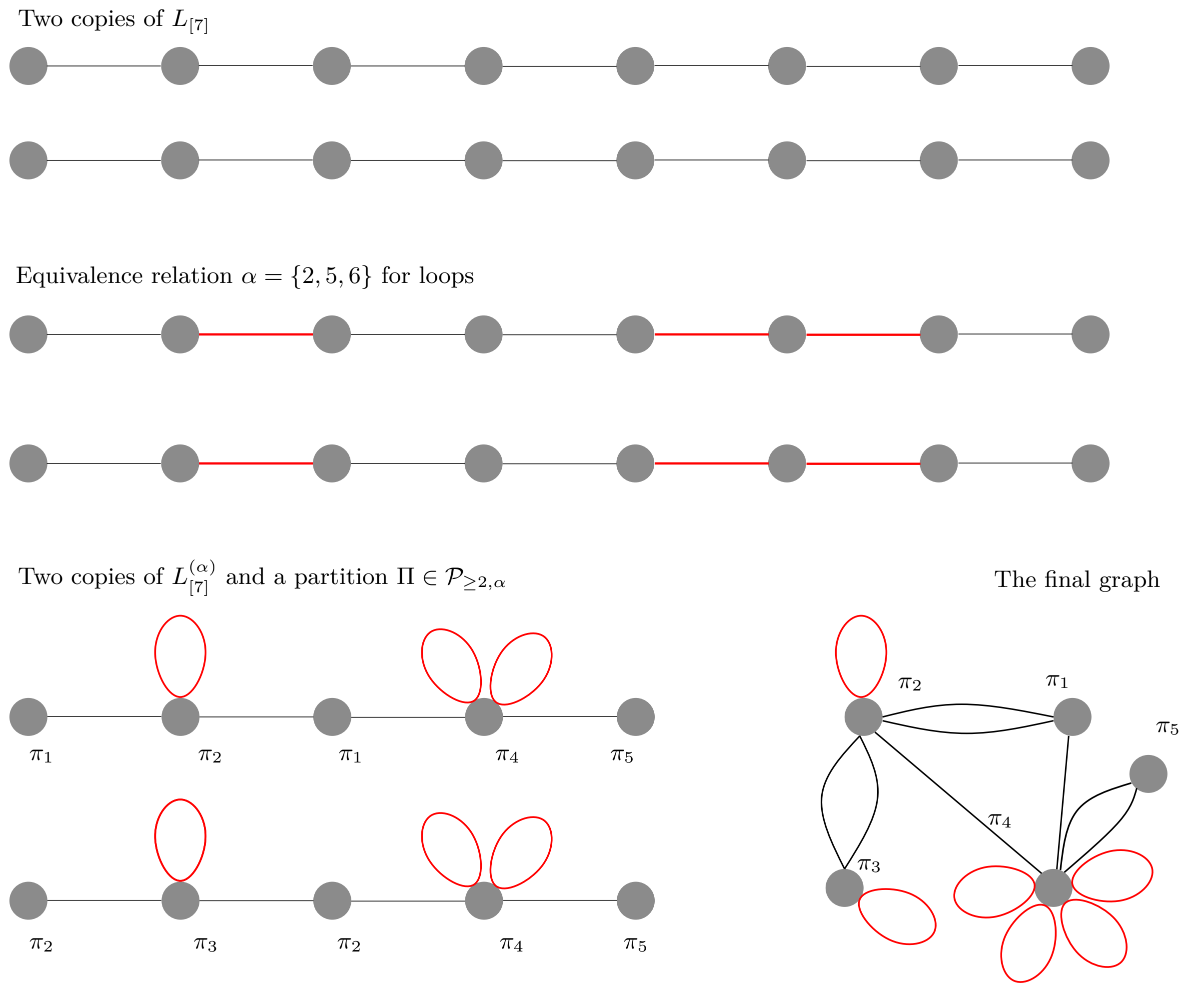


Figure 2.4: An illustration for Definition 2.42 in the case $r = 2$, $l = 7$, $\alpha = \{2, 5, 6\}$ and $\Pi \in \mathcal{P}_{\geq 2,\alpha}$ a partition on $[2] \times [7]$. Labeling the vertices from 1 to 14 from left to right and top to bottom, the blocks are $\pi_1 = \{1, 4\}$, $\pi_2 = \{2, 3, 8, 11\}$, $\pi_3 = \{9, 10\}$, $\pi_4 = \{5, 6, 11, 13\}$ and $\pi_5 = \{7, 14\}$. Note that between graph 3 and 4, no new loops are created.

As a toy example, consider the graph G' on two vertices x, y, z with edge set $\big\{(x, x), (x, y), (x, y), (y, z)\big\}$. Then $E(T) = \big\{(x, y), (y, z)\big\}$, we have one extra edge (x, y) and one loop (x, x). We find that

$$\begin{aligned}\mathrm{Val}(G') = \frac{1}{N^3} \sum_{i,j,k\in[N]} |a_{ii}||a_{ij}|^2|a_{jk}| \leq \max_i |a_{ii}| \max_{i\neq j} |a_{ij}| \left(\frac{1}{N}\sum_{i\in[N]} 1\right)\left(\frac{1}{N}\sum_{j,k\in[N]} |a_{ij}||a_{jk}|\right) \\ \leq \Gamma^2\psi \frac{1}{N}\sum_k |a_{jk}| \left(\sum_j |a_{ij}|\right) \leq \Gamma^2\psi\frac{1}{N}\sum_k |a_{jk}|\left(\frac{1}{N}\sum_j |a_{ij}|^2\right)^{1/2} \leq \Gamma^2\psi\gamma\left(\frac{1}{N}\sum_k |a_k|^2\right)^{1/2} \leq \Gamma^2\psi\gamma,\end{aligned}$$

where we used the Cauchy-Schwartz inequality in the two last steps. In this case, the enumeration of T was $\pi_1 = x$, $\pi_2 = y$ and $\pi_3 = z$.

In the case of G_Π we find

$$\begin{aligned}\mathrm{Val}(G) &\leq \Gamma^{rl}\frac{1}{N^{|V(G)|}}\sum_{s\in[N]^{\Pi}_*}\prod_{\substack{(\pi,\pi')\in E(G)\\ \pi\neq\pi'}} a_{s_\pi s_{\pi'}}\\ &\leq \Gamma^{rl}\psi^{r(l-|\alpha|-1)-|E(T)|}\,\mathrm{Val}(T) \leq \Gamma^{rl}\psi^{r(l-|\alpha|-1)-|E(T)|}(\sqrt{\Gamma}\gamma\wedge\psi)^{|E(T)|}.\end{aligned}$$

In the first inequality, we used the fact that, by definition, for $\Pi \in \mathcal{P}_{\geq 2,\alpha}([rl])$ there are exactly $r|\alpha|$ loops. Once all loops are removed, there remain $r(l-|\alpha|-1)$ edges G_Π.. In the second inequality, we estimated all extra edges by ψ. In the last equality, we used the Cauchy-Schwartz inequality repeatedly, as was illustrated on G'.

If c is the number of connected components of G and k is its number of vertices, we have $|E(T)| = k - c$. The key observation is now that, for $\alpha \neq [l-1]$ and $\Pi \in \mathcal{P}_{\geq 2,\alpha}([rl])$, every connected component has at least two vertices. Indeed every copy of $L^{(\alpha)}_{[l]}$ has at least two vertices and those vertices cannot belong to the same block of Π, by ((ii)) Definition 2.42. Therefore $k \geq 2c$ and since $\gamma \leq \psi$ we find that

$$\begin{aligned}\mathrm{Val}(G_{r,l,\Pi}) &\leq \Gamma^{rl}\psi^{r(l-|\alpha|-1)}\left(\frac{\sqrt{\Gamma}\gamma\wedge\psi}{\psi}\right)^{|E(T)|} \leq \Gamma^{rl}\psi^{r(l-|\alpha|-1)}\left(\frac{\sqrt{\Gamma}\gamma\wedge\psi}{\psi}\right)^{k/2}\\ &\leq \Gamma^{rl}\psi^{r(l-|\alpha|-1)-k/2}(\sqrt{\Gamma}\gamma)^{k/2} \leq \Gamma^{rl}\psi^{r(l-|\alpha|-1)-k/2}\gamma^{k/2}.\end{aligned}$$

□

Proof of Proposition 2.39. We start from (2.89). The family of graphs induced by all possibles partitions of $[rl]$ is difficult to handle and therefore we first decouple the problem, by considering separately all possible (l,α)-line graphs. This is done by using the identity

$$1 = \sum_{\alpha\subset[l]}\sum_{i_1,\dots,i_l}\prod_{(k,k+1)\in\alpha} 1_{i_k=i_{k+1}}\prod_{(k,k+1)\notin\alpha} 1_{i_k\neq 1_{k+1}},$$

and Minkowski's inequality to find that

$$\|S\|_r \leq \sum_{\alpha\subset[l-1]}\|S^\alpha\|_r$$

where

$$S_l^{(\alpha)} := \sum_{i_1,\dots,i_l} X_{i_1}a_{i_1i_2}X_{i_2}a_{i_2i_3}\dots a_{i_{l-1}i_l}X_{i_l}\prod_{(k,k+1)\in\alpha}\mathbf{1}_{i_k=i_{k+1}}\prod_{(k,k+1)\notin\alpha}\mathbf{1}_{i_k\neq i_{k+1}},\quad \alpha\subset[l-1].$$

The first product imposes that the vertices $i_k, i_{k+1} \in [N]$, for $(k,k+1)\in\alpha$ belong to the same block of the partition Π. The second product forbids vertices $i_k, i_{k+1} \in [N]$ from belonging to the same block, if $(k,k_1)\notin\alpha$.

Let us now fix $\alpha \subsetneq [l-1]$ and set $G := \sqcup_{i=1}^r L^{(l)}$. Using (2.82) and (2.90), we find

$$\begin{aligned}\mathbb{E}(S_l^{(\alpha)}(i))^r &\leq \sum_{\Pi\in\mathcal{P}_{\geq 2}([rl],\alpha)}\sum_{s\in[N]^{V(G)}}\frac{1}{N^{|\Pi|}q^{rl-2|\Pi|}}\prod_{(\pi,\pi')\in E(G_\Pi)} a_{s_\pi s_{\pi'}}\\ &\leq \sum_{\Pi\in\mathcal{P}_{\geq 2}([rl],\alpha)}\frac{1}{q^{rl-2|\Pi|}}\,\mathrm{Val}(G_\Pi).\end{aligned}$$

We can now control $\mathrm{Val}(G_\Pi)$ using Lemma 2.43. Note that the bound (2.92) only depends on $|V(G_\Pi)| = |\Pi|$. Using the Stirling numbers of the second kind, denoted Str, we see that

$$\left|\left\{\Pi\in\mathcal{P}_{\geq 2,rl}\colon\ |\Pi| = k\right\}\right| \leq \mathrm{Str}(rl,k) \leq \binom{rl}{k}k^{rl-k},\quad k\in\mathbb{N}^*.$$

Setting $b := |\alpha|$, we find

$$\begin{aligned}\mathbb{E}(S_l^{(\alpha)})^r &\leq \sum_{k=2}^{r(l-b)\wedge rl/2} \operatorname{Str}(r(l-b),k)\frac{1}{q^{rl-2k}} \operatorname{Val}(G_\Pi : |V(G_\Pi)| = k) \\ &\leq 2^r\Gamma^{rl} \max_{k,b} \frac{k^{r(l-b)-k}}{q^{rl-2k}} \psi^{r(l-b-1)-k/2}\gamma^{k/2},\end{aligned}$$

where the maximum is taken over $1 \leq k \leq r(l-b) \wedge rl/2$ and $0 \leq b \leq l-2$.

Lemma 2.44. *Let $r, l \geq 0$ and $0 \leq b \leq r(l-1)$. Then*

$$\max_{k,b} \frac{k^{r(l-b)-k}}{q^{rl-2k}} \psi^{r(l-b-1)-k/2}\gamma^{k/2} \leq \left(\frac{4\Gamma}{q}\frac{rl}{1+(\log(\tilde{\psi}/\gamma))_+}\right)^{rl}\left(\sqrt{\frac{\gamma}{\psi}} \vee \psi\right)^r,$$

where the maximum is taken over $1 \leq k \leq r(l-b) \wedge rl/2$ and $0 \leq b \leq l-2$.

Before proving Lemma 2.44 we use it to conclude the proof. We find

$$\max_{\alpha \subsetneq [l-1]} \left\| S_l^{(\alpha)} \right\|_r \leq (2\Gamma)^l \mathcal{E}.$$

There remains to control the case $\alpha = [l-1]$, i.e.the term $S_l^{[l-1]} = \sum_i a_{ii}^{l-1} X_i^l$. Writing $Y_i = X_i^l - EX_i^l$, $i \in [N]$ and using the fact that $\mathbb{E}Y_i = 0$ and $\mathbb{E}|Y_i|^k \leq \frac{2^k}{Nq^{kl-2}}$ for $k \geq 1$ as well as (2.83a) with $a_i := a_{ii}^{l-1}$, we find that

$$\left\|\sum_i a_{ii}^{l-1} Y_i\right\|_r \leq \frac{4^l r\Gamma^{l-1}}{q^{l-1}}, \qquad \left\|\sum_i E(X_i^l)\right\|_r \leq \frac{(2\Gamma)^l}{q^{l-2}}.$$

Therefore we get, as soon as $r \geq 2$,

$$\|S_l^{[l-1]}\|_r \leq \frac{4^l\Gamma^{l-1}r}{q^{l-1}} + \frac{2^l\Gamma^{l-1}}{q^{l-2}} \leq \frac{4^l\Gamma^{l-1}r}{q^{l-1}},$$

and we can conclude that

$$\|S_l\|_r \leq \|S_l^{([l-1])}\|_r + \sum_{\alpha \subsetneq [l-1]} \|S_l^{(\alpha)}\|_r \leq \frac{8^l\Gamma^{l-1}r}{q^{l-1}} + (4\Gamma)^l\mathcal{E}.$$

This concludes the proof of (2.87a).
For $S(i)$ and $S(i,j)$ the only difference is that the case $\alpha = [l-1]$ leads to a better error term. Indeed we have

$$\left\|S_l^{[l-1]}(i)\right\|_r \leq \left\|\sum_j a_{ij}a_{jj}^{l-1}Y_i\right\|_r + \left\|\sum_j a_{jj}^{l-1}E(X_i^l)\right\|_r.$$

The first sum can be controlled using (2.83a) to obtain

$$\left\|\sum_j a_{ij}a_{jj}^{l-1}Y_i\right\|_r \leq \Gamma^{l-1}\Big(\frac{2r}{1+2\log(\psi/\gamma)} \vee 2\Big)(\gamma \vee \psi),$$

The adaptations for $S(i,j)$ are straightforward and we skip the details. This concludes the proof. □

Remark 2.45. In the control of $S_l(i)$ and $S_l(i,j)$ we could consider the specificities of the computation graph generated by the extra term a_{ii_1} and $a_{i_l j}$. This is done for instance in [30, Definition 3.12] where they introduce so-called *white vertices*. In our case this only lead to minor improvements (the term $\psi/q^4\gamma$ in the denominator of (2.87a) becomes $\psi/q^2\gamma$) and we do not pursue this amelioration further.

Proof of Lemma 2.44. We first bound the contribution of Γ uniformly by Γ^{rl}. Then we introduce the function $f : [0, l-1] \times [1, rl] \to \mathbb{R}_+$ defined by

$$f(b,k) = \mathbf{1}_{2 \le k \le r(l-b) \wedge \frac{rl}{2}} \frac{k^{r(l-b)-k}}{q^{rl-2k}} \gamma^{k/2} \psi^{r(l-b-1)-\frac{k}{2}}.$$

Our goal is to bound the function f. We start with the case where $\gamma \le \psi/q^4$. In this case we consider, for fixed $\bar{b} \in [l-2]$, the function $\log f(k,\bar{b})$. A standard analysis of that function yields that

$$k_* = \frac{r(l-b)}{1 + \log k_* + \log(\psi/\gamma q^4)},$$

is a critical point and that the function itself is concave in $[1, rl]$. We conclude that

$$f(k,\bar{b}) \le f(1,\bar{b}) + f(k_*,\bar{b}) \le \left(\frac{\psi^{l-b-1}}{q^l}\right)^r + \left(\frac{r(l-b)}{1+\log(\psi/\gamma q^4)}\right)^{r(l-b)} \left(\frac{\psi^{l-b-1}}{q^l}\right)^r.$$

Observe that the right-hand side of the above equation is a function of b, which we call $g := g(b)$. A short analysis of $\log g$ shows that it is convex and thus maximized at the border of its domain of definition, that is at $b = 0$ and $b = l-2$. We find

$$\max f(k,b) \le \frac{1}{q^{rl}} \left(\psi^{l-1} \vee \psi\right)^r + \frac{1}{q^{rl}} \left(\psi^{l-1} \left(\frac{rl}{1+\log(\psi/\gamma q^4)}\right)^l \vee \frac{r\psi}{1+\log(\psi/\gamma q^4)}\right)^r \le \mathcal{E}^r. \tag{2.93}$$

The case $\gamma \ge \psi/q^4$ is handled in another way. Here we need to only optimize over b. We find

$$\max_{k,b} f(k,b) \le \max_b \frac{r^{r(l-b)}}{q^{rl}} \left(\frac{\gamma q^4}{\psi}\right)^{k_{\max}/2}, \quad k_{\max} := k_{\max}(b) = r(l-b) \wedge \frac{rl}{2}.$$

The logarithm of left hand side of the above inequality is a piecewise linear function in b with a non-positive derivative. It is therefore maximized at $b = 0$ in which case we find

$$\max_{k,b} f(k,b) \le \frac{r^{rl}}{q^{rl}} \left(\frac{\gamma q^4}{\psi}\right)^{rl/4} \le r^{rl} \left(\frac{\gamma}{\psi}\right)^{rl/4}.$$

Combining this bound with (2.93), we conclude. ☐

We will also need the following refinement of Proposition 2.49 which is the analog of (2.83c).

$$S_l^o := \sum_{i_1,\dots,i_l \ne i_1} X_{i_1} a_{i_1 i_2} X_{i_2} \dots a_{i_{l-1} i_l} X_{i_l}. \tag{2.94}$$

Proposition 2.46. *Suppose the assumptions of Proposition 2.49 are satisfied. Then*

$$\|S_l^o\|_r \le (4\Gamma)^l \mathcal{E}.$$

Proof. The proof is the same as for (2.87a) except that we do not have to consider the case where $\alpha = [l-1]$ since this is prohibited by the condition $i_1 \ne i_l$. ☐

Lemma 2.47. *Let Z_k be centered random variables with $\mathbb{E}|Z_k|^2 \leq \frac{Kq}{N}$ and $\|Z_k\|_\infty \leq \frac{K}{q}$, for some constant $K > 0$. Suppose $a_{ij} \in \mathbb{C}$ for $i, j \in [N]$ satisfy*

$$\Big(\frac{1}{N}\sum_k a_k^2\Big)^{1/2} \leq \gamma, \quad \max_k |a_k| \leq \Gamma.$$

Suppose $q^3 \geq \gamma$

$$\Big\|\sum_k a_k Z_k\Big\|_r \leq \Big(\frac{2K\Gamma r}{1+\log(q^3/\gamma)}\Big)\frac{1}{q} \tag{2.95}$$

Proof. We have

$$\mathbb{E}\left(\sum_k a_k Z_k\right)^r \leq \sum_{k=1}^{r/2} k^{r-k} K^r \Gamma^{r-k} \frac{(\gamma q)^k}{q^{r-2k}} \leq \left(\frac{2K\Gamma}{q}\right)^r \left(\frac{r}{1+\log(q^3/\gamma)}\right)^r$$

where we used the binomial theorem to estimate the sum by 2^r times the value at the maximum k_* with

$$k_* \leq \frac{r}{1+(\log(q^3/\gamma))_+}.$$

□

2.7 Consequences of large deviations

In this section, we prove estimates on multilinear sums of Green function entries. We need the following bound on the size of $\sum_y H_{xy}$ for $x \in [N]$.

Lemma 2.48. *For $1 \leq d \leq N/2$, for any $\nu > 0$ there exists $\mathcal{C}_\nu > 0$ such that*

$$\mathbb{P}\Big[\max_{x\in[N]} \sum_y |H_{xy}| \leq \mathcal{C}_\nu \frac{\log N + d}{\sqrt{d}}\Big] \leq C_\nu N^{-\nu}.$$

We denote $\Delta(\nu)$ as the indicator function of the event in the above probability.

Proof. Let $x \in [N]$ and $Y_y := |H_{xy}| - \mathbb{E}|H_{xy}|$, $y \in [N]$. By (2.1), Y_y is centered and satisfies $\mathbb{E}Y_y^2 \leq 1/N$ and $|Y_y| \leq \frac{K}{\sqrt{d}}$ for some constant $K \geq 0$. We can thus use Lemma 2.47 to bound with very high probability

$$\sum_y Y_y = \mathcal{O}\left(\frac{\log N}{\sqrt{d}}\right).$$

We conclude that

$$\sum_y |H_{xy}| = \sum_y Y_y + N\mathbb{E}|H_{xy}| \leq \mathcal{C}\frac{\log N}{\sqrt{d}} + 4\sqrt{d} \leq \mathcal{C}\frac{\log N + d}{\sqrt{d}}.$$

where $\mathcal{C} \geq 0$ depends on the notion of very high probability. □

Let us define

$$\Gamma(T) := \max_{x,y\notin T} \big|\widetilde{G}_{xy}^{(T)}\big|, \quad \Gamma(T;z) := \max_{x,y\notin T\cup\{z\}} \Big|\widetilde{G}_{xy}^{(T,z)}\Big|, \quad Q(T) := \max_{x\neq y,\, x,y\notin T} \big|\widetilde{G}_{xy}^{(T)}\big|.$$

Proposition 2.49. *Let $\Gamma > 0$, $T \subset [N]$ and $u \in [N] \setminus T$. Suppose there exists $\kappa > 0$ such that*

$$d^2 \leq N\eta N^{-\kappa/4}, \quad d^2|T|^2 \leq N^{\kappa/2}, \quad \Gamma = O(1) \tag{2.96}$$

Then for any $\nu > 0$, there exists $\mathcal{C} > 0$ such that on the event $\{\Gamma(Tu), \Gamma(T,u) \leq \Gamma\}$ we have

$$\sum_a^{(T\cup\{u\})} H_{ua}\widetilde{G}^{(T,u)}_{ax} \leq \frac{\mathcal{C}\Gamma}{\sqrt{d}}, \tag{2.97a}$$

$$\sum_a^{(T\cup\{u\})} \widetilde{G}^{(T,u)}_{xa} H_{ua}\widetilde{G}^{(T\cup\{u\})}_{ay} \leq \frac{\mathcal{C}\Gamma}{\sqrt{d}}, \tag{2.97b}$$

with very high probability.
Moreover, for any constant $C' > 0$, if $\{\Gamma(T), \Gamma(Tu) \leq \Gamma,\, Q(Tu) \leq \Gamma C' d^{-1/2}\}$ holds then we have

$$\sum_a^{(T\cup\{u\})} H_{ua}\widetilde{G}^{(T,u)}_{ax} \leq A_{ux}\frac{C'\mathcal{C}}{\sqrt{d}} + C'\frac{\mathcal{C}\Gamma}{d}, \tag{2.98a}$$

$$\sum_a^{(T\cup\{u\})} \widetilde{G}^{(T,u)}_{xa} H_{ua}\widetilde{G}^{(T\cup\{u\})}_{ay} \leq \mathbf{1}_{x=y}A_{xu}\frac{C'\mathcal{C}}{\sqrt{d}} + C'\frac{\mathcal{C}\Gamma}{d}, \tag{2.98b}$$

hold with very high probability.

Note that the condition on Γ in (2.96) could be restated as $\frac{\Gamma}{\sqrt{d}} \leq \frac{1}{C}$ for some $C > 0$ large enough depending on the notion of very high probability.

Proof. Without loss of generality, we suppose that $T = \emptyset$. The general case is obtained by writing $\tilde{H}^{(T)}$ and $\widetilde{G}^{(Tu)}_{xy}$ instead of H and $\widetilde{G}^{(u)}_{xy}$ respectively. We restrict ourselves to the event $\Delta(\nu)$ defined in Lemma 2.48. Since this event holds with very high probability, it suffices to prove our results on it.
Let us consider the sums expression defined in (2.84) with $a_{ij} := \widetilde{G}^{(u)}_{ij}$, $X_i := H_{ui}$, $\psi = 1$. Applying (A.4), we see that

$$\sum_a^{(u)} H_{ua}G^{(u)}_{ax} = \sum_a^{(u)} H_{ua}\widetilde{G}^{(u)}_{ax} + \sum_{a,i_1}^{(u)} H_{ua}\widetilde{G}^{(u)}_{ai_1}H_{i_1 b}G^{(u)}_{i_1 x} = S_1(x) + \sum_{i_1}^{(u)} S_1(i_1)H_{ui_1}G^{(u)}_{bx}.$$

We now iterate (A.4) M times, for some fixed $M \in \mathbb{N}_{>0}$ and find

$$\sum_a^{(u)} H_{ua}G^{(u)}_{ax} = \sum_{l=1}^{M} S_l(x) + \sum_b^{(u)} S_M(b)H_{ub}G^{(u)}_{bx}. \tag{2.99}$$

Now by Proposition 2.49 and (2.88), using that $\log(\psi/d\gamma) \geq \frac{\kappa}{4}\log N$, see that (2.87b) becomes

$$\big\|S_l(x)\big\|_{r|H^{(u)}} \leq \left(\frac{32\Gamma l r}{\kappa \log N}\right)^l \leq \left(\frac{32\Gamma M r}{\kappa \log N}\right)^l$$

for $1 \leq l \leq M$. Therefore, setting $r = \nu \log N$ and $M = 10$, $\mathcal{C} \geq \frac{80M\mathrm{e}}{\kappa}$ and $\varepsilon = \mathcal{C}\Gamma/\sqrt{d}$, we find, using Chebyshev's inequality,

$$\mathbf{1}_{\Gamma_0,\Gamma_1\leq\Gamma}\mathbb{P}\big[\max_{1\leq i\leq M} S_l(x) > \varepsilon^l \big| H^{(u)}\big] \leq \mathbf{1}_{\Gamma_0,\Gamma_1\leq\Gamma} M \max_{1\leq i\leq M}\mathbb{P}\big[S_l(x) > \varepsilon^l\big|H^{(u)}\big] \leq M \max_{1\leq l\leq M} \mathrm{e}^{-lr} \leq \mathcal{C}_\nu N^{-\nu}. \tag{2.100}$$

In the same way, we that, setting $\varepsilon = \mathcal{C}\Gamma/d^{-5}$ for $\mathcal{C} > 0$ large enough,

$$\mathbb{P}\big(\max_{b\in[N]} S_M(b) \geq \varepsilon\big) \leq N\varepsilon^{-r}\left(\frac{80\nu\Gamma}{\kappa\sqrt{d}}\right)^{10r} \leq \mathcal{C}_\nu N^{-\nu}.$$

Therefore, using Lemma 2.48, we have

$$\sum_a^{(u)} H_{ua}G_{ax}^{(u)} \leq \sum_{l=1}^{M}\Big(\frac{\mathcal{C}\Gamma}{\sqrt{d}}\Big)^l + \frac{\mathcal{C}\Gamma}{d^5}\max_{b\in[N]}|G_{bx}^{(u)}|\sum_b^{(u)}|H_{ub}| \leq 2\frac{\mathcal{C}\Gamma}{\sqrt{d}} + \frac{\mathcal{C}\Gamma}{d^{3/2}},$$

where we used the fact that $\frac{\mathcal{C}\Gamma}{\sqrt{d}} \leq \frac{1}{2}$ and the definition of $\Delta(\nu)$ to bound the contribution of the last sum. This proves (2.97a).

Equation (2.97b) is proved analogously using the bounds on $S_l(i,j)$ and the identity

$$\sum_a^{(u)} \widetilde{G}_{xa}^{(u)}H_{ua}\widetilde{G}_{ay}^{(u)} = \sum_{l=1}^{M} S_l(x,y) + \sum_b^{(u)} S_l(x,b)H_{ub}G_{by}^{(u)}, \quad M\in\mathbb{N}^*. \tag{2.101}$$

To prove (2.98), under the additional assumption that $\max_{a\neq b}\widetilde{G}_{ab}^{(u)} \leq C'd^{-1/2}$ we only need to improve the bounds on

$$\sum_a^{(u)} H_{ua}\widetilde{G}_{ax}^{(u)}, \quad \sum_a^{(u)} \widetilde{G}_{xa}^{(u)}H_{ua}\widetilde{G}_{ay}^{(u)}, \tag{2.102}$$

since $S_l(x), S_l(x,y) = \mathcal{O}(d^{-1})$ for $l \geq 2$. Indeed observe that in (2.85), we collect an extra ψ factor as soon as $l > 1$.

For the first sum in (2.102), if $A_{xa} = 0$, Here we use the fact that H_{xy} is A_{xy} measurable and that $\big(H_{xy}\colon x,y\in[N]\big)$ form a family of independent random variables. Therefore

$$\begin{aligned}\Big\|\mathbf{1}_{A_{ux}=0}\sum_a^{(u)} H_{ua}\widetilde{G}_{ax}^{(u)}\Big\|_r &\leq \bigg\|\Big\|\mathbf{1}_{A_{ux}=0}\sum_a^{(ux)} H_{ua}\widetilde{G}_{ax}^{(u)}\Big\|_{r|H^{(u)}} + \Big\|\frac{K\sqrt{d}|\widetilde{G}_{xx}^{(u)}|}{N}\Big\|_{r|H^{(u)}}\bigg\|_r \\ &= \bigg\|\big\|\mathbf{1}_{A_{ux}=0}\big\|_{r|H^{(u)}}\Big\|\sum_a^{(ux)} H_{ua}\widetilde{G}_{ax}^{(u)}\Big\|_{r|H^{(u)}}\bigg\|_r + \bigg\|\frac{K\sqrt{d}|\widetilde{G}_{xx}^{(u)}|}{N}\bigg\|_r \\ &\leq \bigg\|\Big\|\sum_a^{(ux)} H_{ua}\widetilde{G}_{ax}^{(u)}\Big\|_{r|H^{(u)}}\bigg\| + \frac{K\sqrt{d}\max_x|\widetilde{G}_{xx}^{(u)}|}{N}.\end{aligned}$$

In the first step, we used Minkowski's inequality on the conditional L^r norm and $\mathbf{1}_{A_{xu}=0}|H_{xu}| \leq \frac{K\sqrt{d}}{N}$. In the second step we used the independence between A_{xu} and $(H_{ua}\colon a\neq x)$ and the fact that $\widetilde{G}_{xx}^{(u)}$ is $H^{(u)}$-measurable. In the third inequality, we again used Minkowski's and the bound 1 on the indicator function. We can then apply (2.83a) with $\psi = \frac{C'\Gamma}{d}$ to conclude that

$$\bigg\|\mathbf{1}_{\tilde{\Gamma}\leq\Gamma}\Big\|\mathbf{1}_{A_{ux}=0}\sum_a^{(u)} H_{ua}\widetilde{G}_{ax}^{(u)}\Big\|_{r|H^{(u)}}\bigg\|_r \leq \frac{C\Gamma}{d}\left(\frac{r}{1+\log(\psi/\gamma)}\right).$$

This translates into a very high probability bound as before.

For the second sum in (2.102), we split the sum into three parts

$$\sum_a^{(u)} \widetilde{G}_{xa}^{(u)}H_{ua}\widetilde{G}_{ay}^{(u)} = \widetilde{G}_{xx}^{(u)}H_{ux}\widetilde{G}_{xy}^{(u)} + \widetilde{G}_{xy}^{(u)}H_{uy}\widetilde{G}_{yy}^{(u)} + \sum_a^{(uxy)} \widetilde{G}_{xa}^{(u)}H_{ua}\widetilde{G}_{ay}^{(u)}. \tag{2.103}$$

If $x = y$ then the bound is the same as in (2.97b) unless $A_{xu} = 0$ in which case we get

$$\widetilde{G}^{(u)}_{xx} H_{ux} \widetilde{G}^{(u)}_{xy} + \widetilde{G}^{(u)}_{xy} H_{uy} \widetilde{G}^{(u)}_{yy} \leq \frac{K\Gamma\sqrt{d}}{N} = O(1/d).$$

On the other hand if $x \neq y$, we can bound the sum using (2.87b) with $\psi = C' d^{-1/2}$ instead of $\psi = \Gamma$ and as for the two first terms we find

$$|\widetilde{G}^{(u)}_{xx} H_{ux} \widetilde{G}^{(u)}_{xy}| \leq \Gamma \frac{1}{\sqrt{d}} \frac{C'\Gamma}{\sqrt{d}} \leq \frac{\Gamma^2}{d}.$$

Finally, $\sum_a^{(uxy)} \widetilde{G}^{(u)}_{xa} H_{ua} \widetilde{G}^{(u)}_{ay}$ is bounded using (2.87b) to bound $S_1(x, y)$ with $\max_{a \neq b} \widetilde{G}^{(u)}_{ab} \leq \frac{C'\Gamma}{\sqrt{d}}$. This concludes the proof. □

Proposition 2.50. *We have on the event* $\max_{x,y,a} |G_{xy}| \vee |\widetilde{G}^{(a)}_{xy}| \leq \Gamma$, *we have with very high probability*

$$\sum_{x \neq y}^{(a)} H_{xa} G^{(a)}_{xy} H_{ay} \leq \frac{\Gamma \mathcal{C}}{d}. \tag{2.104}$$

Proof. Let us recall S_l^o defined in (2.94), (A.4) to expand the sum and find

$$\sum_{x \neq y}^{(a)} H_{xa} G^{(a)}_{xy} H_{ay} = \sum_{l \geq 2} S_l^o + \sum_b^{(a)} S_l(b) H_{by}.$$

The last sum is estimated using Lemma 2.48,

$$\sum_b^{(a)} S_l(b) H_{by} \leq S_1 + \max_l |S_l(b)| \sum_b^{(a)} |H_{by}| \leq \max_b |S_M(b)| \frac{\log N + d}{\sqrt{d}}$$

with very high probability.
By (2.83b) and a Chebyshev's argument, we see that $S_1 = \mathcal{O}\big(\frac{\Gamma}{d}\big)$. For $l = 2, \ldots, M$, we use Proposition 2.46 and find

$$\max_b |S_l^o| \leq \left(\frac{\mathcal{C}\Gamma}{\sqrt{d}}\right)^l$$

with very high probability. We conclude as before. □

2.8 Quadratic vector equations

The results In this section, are focused on the stability of the self-consistent equation (2.55).
For $N \in \mathbb{N}$, $\mathbf{a} := (a_x)_{x \in [N]} \in \mathbb{C}^N$, we define the matrix

$$B(\mathbf{a}) := 1 - \mathbf{1}\mathbf{a}^* = \begin{bmatrix} 1 - a_1 & -a_1 & -a_1 & \ldots & -a_1 \\ -a_2 & 1 - a_2 & -a_2 & \ldots & -a_2 \\ -a_3 & -a_3 & 1 - a_3 & \ldots & -a_3 \\ -a_4 & \ddots & \ddots & \ddots & -a_4 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ -a_N & -a_N & -a_N \ldots & \ldots & 1 - a_N \end{bmatrix}. \tag{2.105}$$

Writing the Von Neumann series we see that

$$B^{-1}(\mathbf{a}) = \sum_{n \geq 0} (\mathbf{1}a^*)^n = I + \sum_{n \geq 0} \langle \mathbf{a}, \mathbf{1} \rangle^n \mathbf{1}\mathbf{a}^* = I + \frac{1}{1 - \langle \mathbf{a}, \mathbf{1} \rangle} \mathbf{1}\mathbf{a}^* = I + \frac{1}{\det B(\mathbf{a})} \mathbf{1}\mathbf{a}^*. \tag{2.106}$$

Therefore for B^{-1} to be bounded, we need to ensure that $1 - \frac{1}{N} \sum_i a_i$ remains bounded away from zero.

Proof of Lemma 2.26. Let $\mathbf{e} := \frac{1}{\sqrt{|\mathcal{X}|}}\mathbf{1}_{\mathcal{X}}$, $\mathbf{m} = (m_1, \dots, m_{|\mathcal{X}|})$, $\mathbf{g} = (g_1, \dots, g_{|\mathcal{X}|})$ and $\boldsymbol{\varepsilon} := (\varepsilon_1, \dots, \varepsilon_{|\mathcal{X}|})$. Throughout the proof we multiply column vectors entrywise
Subtracting (2.58) and (2.55) yields

$$\frac{1}{g_x m_x}(g_x - m_x) = \big(\mathbf{e}\mathbf{e}^*(\mathbf{g}-\mathbf{m})\big)_x + \varepsilon_x, \quad x \in \mathcal{X}.$$

Multiplying this equation by on both sides $g_x m_x$ and using $g_x m_x = m_x(g_x - m_x) + m_x^2$, we get

$$g_x - m_x = m_x(g_x - m_x)\big(\mathbf{e}\mathbf{e}^*(\mathbf{g}-\mathbf{m})\big)_x + m_x^2\big(\mathbf{e}\mathbf{e}^*(\mathbf{g}-\mathbf{m})\big)_x + m_x(g_x - m_x)\varepsilon_x + m_x^2\varepsilon_x, \quad x \in \mathcal{X}.$$

Recalling (2.105), we define the matrix $B := B(\mathbf{m}^2) = 1 - (\mathbf{m}^2\mathbf{e})\mathbf{e}^*$, where $\mathbf{m}^2 = (m_1^2, \dots, m_{|\mathcal{X}|}^2)$. Subtracting the above equation by $m_x^2\big(\mathbf{e}\mathbf{e}^*(\mathbf{g}-\mathbf{m})\big)_x$, we find the vectorial equation

$$B(\mathbf{g}-\mathbf{m}) = \mathbf{m}(\mathbf{g}-\mathbf{m})(\mathbf{e}\mathbf{e}^*(\mathbf{g}-\mathbf{m})) - \mathbf{m}(\mathbf{g}-\mathbf{m})\boldsymbol{\varepsilon} + \mathbf{m}^2\boldsymbol{\varepsilon}. \tag{2.107}$$

For a matrix $R \in \mathbb{C}^{\mathcal{X}\times\mathcal{X}}$, we write $\|R\|_{\infty\to\infty}$ for the operator norm induced by the norm $\|\mathbf{r}\|_\infty = \max_{x\in\mathcal{X}}|r_x|$. In particular, we have $\|\mathbf{m}^2\|_\infty \le \tau^2$, since since $|m_x| \le \frac{1}{\operatorname{Im}(z+\overline{m})} \le \tau$.
In order to bound $\|\mathbf{g}-\mathbf{m}\|_\infty$, we want to find the inverse of B and then multiply (2.107) by B^{-1}. Recalling (2.106), we see that in order to control the $\|B\|_\infty$, we need to study the quantity $\det B = 1 - \frac{1}{\mathcal{X}}\sum_x m_x^2$. We have

$$\begin{aligned}
&\operatorname{Re}\Big[\frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}}\frac{1}{(v_x - z - \overline{m})^2}\Big] = \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}}\frac{\big(\operatorname{Re} v_x - z - \overline{m}\big)^2 - \big(\operatorname{Im} v_x - z - \overline{m}\big)^2}{|v_x - z - \overline{m}|^4}\\
&\le \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}}\frac{1}{|v_x - z - \overline{m}|^2} - \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}}\frac{\big(\operatorname{Im} v_x - z - \overline{m}\big)^2}{|v_x - z - \overline{m}|^4} \le \frac{\operatorname{Im}\overline{m}}{\operatorname{Im}(\overline{m}+z)} - \bigg(\frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}}\frac{\operatorname{Im}(z+\overline{m})}{|v_x - z - \overline{m}|^2}\bigg)^2\\
&\le 1 - (\operatorname{Im}\overline{m})^2,
\end{aligned}$$

where we used the Cauchy-Schwartz inequality in the second inequality. We conclude that $\det B$ is bounded away from zero by a constant as soon as (2.57) holds. In particular

$$\|B^{-1}\|_{\infty\to\infty} \le \|1\|_{\infty\to\infty} + \frac{\tau^2}{1-\tau^2} \le C,$$

for some $C > 0$ depending only on τ.
Combining the above estimate with (2.107), we find

$$\|\mathbf{g}-\mathbf{m}\|_\infty \le \|B^{-1}\|_{\infty\to\infty}\Big[\tau\|\mathbf{g}-\mathbf{m}\|_\infty^2 - \tau\|\mathbf{g}-\mathbf{m}\|_\infty\|\varepsilon\|_\infty + \tau^2\|\varepsilon\|_\infty\Big].$$

If we know *a priori* that $\|\mathbf{g}-\mathbf{m}\|_\infty \le \lambda$ and that $\lambda \le \frac{1}{2C}$, we can move the term quadratic in $\|\mathbf{g}-\mathbf{m}\|_\infty$ to the left-hand side and estimate $\|\mathbf{g}-\mathbf{m}\|_\infty = O(\|\boldsymbol{\varepsilon}\|_\infty)$. This concludes the proof. □

Proof of Corollary 2.28. This is done by induction. Fix $U \subset [N]$ and $E \in [-R, R]$. Consider $z_k := E + \mathrm{i}L(1 - kN^{-2})$, $k = 0, \dots, k_*$. Here we define $k_* := \sup\{k \in \mathbb{N} : z_k \in \mathbb{S}\}$ and $L \ge 2/\lambda$. Then (2.61) holds for $g_x = m_x^{[N]}$ with $\varepsilon_x = \frac{|U|}{N}$ and for z_0 we trivially have $\sup_{x\in\mathcal{X}}|g_x - m_x| \le \lambda$. The constant $b > 0$ from Lemma 2.26 depends only on R. We conclude that $\sup_{x\in U}\big|m_x^U - m_x^{[N]}\big| \le C\frac{|U|}{N} = o(1)$, for some $C > 0$ depending on R. We can then bootstrap this using the Lipschitz continuity of both $m^{[N]}$ and m^U, and using the fact that $C\frac{|U|}{N} \le \lambda$. This concludes the proof. □

2.9 Estimates on the imaginary parts of the normalized Stieltjes transform

The following result gives lower bound on $\overline{m}$ defined in (2.10) in terms of the density of the $(v_x)_{x\in[N]} \in \mathbb{R}^N$. The following lemma contains no randomness. It is stated for (2.9) but the result is also valid for (2.55) by replacing $[N]$ by $\mathcal{X} \subset [N]$ and N by $|\mathcal{X}|$ and m_x by $m_x^{\mathcal{X}}$.

Lemma 2.51 (Existence, uniqueness and characterisation in the bulk)**.** *There exists a unique solution to (2.9) in $\mathbb{H}^N$ and*

$$|\overline{m}(z)| \le 1, \quad z \in \mathbb{H}. \tag{2.108}$$

Recall the definition of $\mathbb{S}_{\kappa,R}$. If there exists $c := c(R) \in (0,1)$ depending on R such that for any $t \in [c^2/8, 2]$ (independent of N) we have

$$\frac{|\{x \in [N] : |v_x - E| \le t\}|}{N} \ge ct, \quad E \in (-R-1, R+1), \tag{2.109}$$

then

$$\operatorname{Im} \overline{m}(z) \ge \frac{c^2}{8}, \quad z \in \mathbb{S}_{\kappa,R}. \tag{2.110}$$

Suppose $v_x \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Then by a second-moment argument we can convince ourselves that (2.109) is satisfied with $\frac{\mathrm{e}^{-R^2/2}}{C}$ for some large enough $C > 0$ as soon as N is large enough with high probability.
Let us first make some comments on $\overline{m}$. A general result on Nevanlinna functions states that if $m : \mathbb{H} \to \mathbb{H}$ is analytic and satisfies the limit condition $\lim_{\eta\to\infty} \mathrm{i}\eta m(\mathrm{i}\eta) = -1$, then m is the Stieltjes transform of some probability measure. Indeed we have

$$|m(z) - m(w)| = \left| \int_{\mathbb{R}} \frac{1}{x-z} - \frac{1}{x-w} \mathrm{d}\mu(x) \right| \le |z-w| \int_{\mathbb{R}} \frac{\mathrm{d}\mu(x)}{|x-z||x-w|} \le |z-w| \frac{1}{\operatorname{Im}(z)\operatorname{Im}(w)},$$

where μ is the associated probability measure. Alternatively we can use [3, Theorem 2.1] with the kernel $S = \frac{1}{N}\mathbf{1}_{[N]}\mathbf{1}_{[N]}^*$ to get this result. Since (2.10) satisfies all of these conditions and so it is N^2-Lipschitz continuous on $\mathbb{S}$.

Proof of Lemma 2.51. The existence and uniqueness result can be found in a very general formulation for instance in [4, Proposition 2.1] or in the Appendix of [37].
For (2.108) we observe that

$$|\overline{m}(z)|^2 \le \frac{1}{N} \sum_{x\in[N]} \frac{1}{|v_x - z - \overline{m}|^2} = \frac{\operatorname{Im}(\overline{m})}{\operatorname{Im}(\overline{m} + z)} < 1.$$

This proves (2.108). To prove (2.110) we first observe that (2.109) implies

$$\frac{1}{N} \sum_x \frac{1}{(\frac{v_x - E}{t})^2 + 1} \ge \frac{1}{N} \sum_{x:|v_x - E| \le t} \frac{1}{1+1} \ge \frac{c}{2} t, \quad t \ge c^2/8. \tag{2.111}$$

Let $E_0 \in (-R, R)$ and $z_k := E_0 + \mathrm{i}\eta_k$, $\eta_k = 1 - kN^{-3}$ and $k = 0, \dots, k_*$ with $k_* := \inf\{k \in \mathbb{N}, z_k \notin \mathbb{S}_{\kappa,R}\} - 1$. We will proceed by induction on k making use of the N^2-Lipschitz continuity of $\overline{m}(z)$.
Denoting $\zeta(z) := z + \overline{m}(z)$ $E_k := \operatorname{Re}\zeta(z_k)$ and $I_k := \operatorname{Im}\zeta(z_k)$, we observe that since $|\overline{m}(z)| \le 1$ and $\operatorname{Im}\overline{m}(z) \ge 0$, we have $E_k \in (-R-1, R+1)$ and $\frac{1}{N} \le I_k \le 2$. For $k = 0$ we have

$$\operatorname{Im}(\overline{m}(z_0)) = \frac{1}{N} \sum_x \frac{I_0}{(v_x - E_0)^2 + I_0^2} = \frac{1}{I_0} \frac{1}{N} \sum_x \frac{1}{(\frac{v_x - E_0}{I_0})^2 + 1} \ge \frac{c}{2} \ge \frac{c^2}{8},$$

where we applied (2.111) with $t = I_0 \geq 1 \geq c^2/8$ in the third step.
For the induction step we suppose that $\mathrm{Im}(\overline{m}(z_k)) \geq \frac{c^2}{8}$. By Lipschitz-continuity and because $\eta_{k+1} \geq \frac{1}{N}$ we have $I_{k+1} \geq \mathrm{Im}(\overline{m}(z_k)) - \frac{1}{N} + \eta_{k+1} \geq \frac{c^2}{8} - \frac{1}{N} + \frac{1}{N} \geq \frac{c^2}{8}$. We now distinguish between two cases.
On the one hand if $I_{k+1} \geq c/2$, we have

$$\mathrm{Im}(\overline{m}(z_{k+1})) = \frac{1}{I_{k+1}} \frac{1}{N} \sum_x \frac{1}{(\frac{v_x - E_0}{I_{k+1}})^2 + 1} \geq \frac{1}{I_{k+1}} \frac{1}{N} \sum_x \frac{1}{(\frac{v_x - E_0}{c/2})^2 + 1} \geq \frac{1}{2} \frac{c}{2} \frac{c}{2} = \frac{c^2}{8}.$$

On the other hand if $I_{k+1} \leq c/2$, we have

$$\mathrm{Im}(\overline{m}(z_{k+1})) \geq \frac{1}{I_{k+1}} \frac{1}{N} \sum_x \frac{1}{(\frac{v_x - E_0}{c^2/8})^2 + 1} \geq \frac{1}{I_{k+1}} \frac{c}{2} \frac{c^2}{8} \geq \frac{c^2}{8},$$

where we used $\frac{1}{I_{k+1}} \geq 2/c$ in the last inequality. This concludes the induction and the proof. □

The condition (2.109) means that we are in the bulk of the spectrum. For a given $E \in \mathbb{R}$ and $t \geq 0$, and v_x, $x \in [N]$, as in (2.4), we introduce the function

$$f(E,t) := \mathbb{P}\big[v_x \in [E-t, E+t]\big]. \tag{2.112}$$

Note that f is independent of the choice of $x \in [N]$ as the v_x are identically distributed. If we fix a constant $R \geq 0$, then by Proposition B.1 we know that

$$f(E,t) \geq \frac{1}{2\sqrt{2\pi}} \int_{E-t}^{E+t} \mathrm{e}^{-x^2/2} \mathrm{d}x \geq \frac{\mathrm{e}^{-(R+2)^2/2}}{10}, \quad E \in [-R,R], \quad 0 \leq t \leq 2, \tag{2.113}$$

for N large enough, and that there exists a constant $C_L := C_L(E) > 0$ such that

$$|f(E,t) - f(E,s)| = \mathbb{P}\big(t < |v_x - E| \leq s\big) = \frac{1 + o(1)}{\sqrt{2\pi}} \int_{I(t) \setminus I(s)} \mathrm{e}^{-x^2/2} \mathrm{d}x \leq C_L |s-t|, \tag{2.114}$$

for all $E \in [-R,R]$ and $0 \leq s < t \leq R$. Note that under these constraints, C_L depends only on R and is uniform in E.

Lemma 2.52. *For $d \gg 1$ and $R = O(1)$ and v_x be as in (2.4). Then there exists $c_* = c_*(R) \in (0,1)$ such that for every $\frac{c_*^2}{8} \leq t \leq 2$,*

$$\inf_{\mathcal{X}:\ |\mathcal{X}^c| = o(N)} \inf_{E \in [-R,R]} \frac{|\{x \in \mathcal{X} : |v_x - E| \leq t\}|}{|\mathcal{X}|} \geq t c_*, \tag{2.115}$$

with high probability.

Proof. Let $R, C > 0$ be constants. Fix $t > 0$, $E \in [-R,R]$ and $I := [E-t, E+t]$ Let us first observe that since

$$\sup_{|\mathcal{X}^c| \leq \sqrt{N}} \left| \frac{|\{x \in \mathcal{X} : v_x \in I\}|}{|\mathcal{X}|} - \frac{|\{x \in [N] : v_x \in I\}|}{N} \right| \leq \sup_{|\mathcal{X}^c| \leq \sqrt{N}} \frac{2|\mathcal{X}^c||\mathcal{X}|}{N|\mathcal{X}|} = o(1),$$

it suffices to show that (2.115) holds of $\mathcal{X} = [N]$ up to taking a smaller constant.

Let us define the enlargement of I to be $I(s) := [E-t-s, E+t+s]$, for $s \geq 0$. We introduce $Z_x(I) := \mathbf{1}_{v_x \in I}$, $x \in [N]$. Then recalling (2.112), $\mathbb{E}[Z_x(I(s))] = f(E, t+s)$ we find

$$\mathbb{E}[Z_x(I) Z_y(I)] = \mathbb{E}\Big[\mathbb{E}\big[Z_x(I)|H_{xy}\big]\mathbb{E}\big[Z_y(I)|H_{xy}\big]\Big] = \mathbb{E}\big[f(E, s + H_{xy})^2\big],$$

where we used the fact that v_x and v_z are independent conditioned on H_{xz}. Using the fact that $H_{xz} \leq \frac{K}{\sqrt{d}}$ and the Lipschitz continuity of $f(E, \cdot)$ from (2.114), we see that

$$\left|\mathbb{E}[Z_x(I)Z_w(I)] - f(E,t)^2\right| \leq \frac{C_L^2 K^2}{d},$$

and

$$\operatorname{Var}\Big(\frac{1}{N}\sum_{x\in[N]} Z_x(I)\Big) = \frac{1}{N^2}\sum_{x,y\in[N]}\Big(\mathbb{E}[Z_x(I)Z_y(I)] - \mathbb{E}[Z_x(I)]\mathbb{E}[Z_y(I)]\Big) \leq \frac{C_L^2K^2}{d}.$$

Writing $Z := \frac{1}{N}\sum_x Z_x$, and using Chebyshev's inequality, we find

$$\mathbb{P}\Big(\frac{|\{x\in[N] : v_x \in I\}|}{N} \leq \frac{f(E,t)}{2}\Big) \leq \mathbb{P}\Big(\left|Z - \mathbb{E}Z\right| \geq \frac{f(E,t)}{2}\Big) \leq \frac{4C_L^2K^2}{df(E,t)^2} = O\big(d^{-1}\big), \tag{2.116}$$

where we used (2.113) in the last equality. Since the right-hand side of (2.116) is independent of E, we deduce that there exists $\tau = \tau(R)$, that can be chosen as the right-hand side of (2.113), such that for any constant $t > 0$ and $E \in [-R, R]$,

$$\frac{|\{x\in[N] : |v_x - E| \leq t\}|}{N} \geq t\tau,$$

holds with probability $1 - O\big(d^{-1}\big)$.
Let $u = \frac{\tau^2}{2^7}$ and $t = \frac{u}{4} = \frac{\tau^2}{2^9}$. Define $E_n := -R + nt$ for $n = 0, \dots, \lceil 2R/t\rceil$. Then the intervals $I_n := [E_n - t, E_n + t]$ are a covering of $[-R, R]$ and for any $E \in [-R, R]$, there is a $n \in \mathbb{N}^*$ such that $[E_n - t, E_n + t] \subset [E - u, E + u]$. Therefore

$$\begin{aligned}&\mathbb{P}\Big(\inf_{E\in[-R,R]} \frac{\left|\left\{x\in[N]\colon\ |v_x - E| \leq u\right\}\right|}{N} < u\tau/4\Big)\\ &\leq \mathbb{P}\Big(\exists n \in [\lceil R/t\rceil]\colon\ \frac{\left|\left\{x\in[N]\colon\ |v_x - E_n| \leq u/4\right\}\right|}{N} < u\tau/4\Big) \leq \frac{2RC}{td} = O\Big(\frac{1}{d}\Big).\end{aligned}$$

We conclude that (2.115) holds for $c_* = \frac{\tau}{4}$ and for fixed $t = \frac{\tau^2}{2^7}$. Let us call Ξ this event.
Let $c_* = \tau/8$. If $u \geq t$, then the interval $[E - u, E + u]$ contains $t\lfloor\frac{u}{t}\rfloor$ disjoint intervals of size t. Then on Ξ we have, for any $\mathcal{X}$ and E as in (2.115),

$$\frac{\left|\left\{x\in[N]\colon\ |v_x - E| \leq u\right\}\right|}{N} \geq t\Big\lfloor\frac{u}{t}\Big\rfloor\frac{\tau}{4} \geq \frac{u}{2}\frac{\tau}{4} = u\frac{\tau}{8} = uc_*,$$

holds for any $u \geq t = \frac{\tau^2}{2^9} = \frac{c_*^2}{8}$. Here we used that $t\lfloor\frac{u}{t}\rfloor \geq \frac{u}{2}$. We conclude that on Ξ, we have

$$\inf_{\mathcal{X}\colon\ |\mathcal{X}^c|\leq\sqrt{N}}\ \inf_{E\in[-R,R]} \frac{|\{x\in\mathcal{X} : |v_x - E| \leq t\}|}{|\mathcal{X}|} \geq tc_*, \qquad \frac{c_*}{2} \leq t \leq 2.$$

Since Ξ holds with high probability this concludes the proof. □

Lemma 2.53. *Let $m = m^{(v)}$ be as in (2.9). With high probability, there exists $\tau > 0$ depending only on R such that for every $x \in [N]$ with $|v_x| \leq R + 2$ we have*

$$\inf_{x\in[N]\colon\ |v_x|\leq R+2} \operatorname{Im}\Big(\frac{1}{d}\sum_{y\in[N]} A_{xy}m_y\Big) \geq \tau, \qquad z \in \mathbb{S}_{\kappa,R}. \tag{2.117}$$

Proof. Let us fix $x \in [N]$ such that $|v_x| \le R+2$ and write $S_1(x) := \big\{y \in [N]\colon\ A_{xy} = 1\big\}$. Observe that this means that $|S_1(x)| = d(1+O(d^{-1/2}))$.
Combining Lemma 2.52 for $\mathcal{X} = [N]$ with Lemma 2.51, we deduce that there exists a constant $c_* := c_*(R) > 0$ such that $c_* \le \operatorname{Im}(z + \overline{m}(z)) \le 2$. As in (2.111), we find

$$\operatorname{Im}\Big(\frac{1}{d}\sum_{y\in S_1(x)} m_y\Big) = \frac{1}{d\operatorname{Im}(z+\overline{m}(z))}\sum_{y\in S_1(x)}\frac{1}{\big(\frac{v_y-\operatorname{Re} z}{\operatorname{Im}(z+\overline{m}(z))^2}\big)^2+1} \ge \frac{1}{2d}\sum_{y\in S_1(x)}\frac{1}{\big(\frac{v_y-\operatorname{Re} z}{c_*}\big)^2+1}. \tag{2.118}$$

Let us define $Z_y := \mathbf{1}_{|v_y - \operatorname{Re} z|\le c_*}$, $y\in[N]$, and let $f := \mathbb{E}[Z_y]$. (Note that f is the analog of $f(c_*)$ introduced in the proof of Lemma 2.52) Then by Proposition B.1, $f > c > 0$ for some constant $c > 0$ depending on R. Conditioning on A_{xy} and using the Lipschitz continuity argument similar to the one used in the proof of Lemma 2.52 (2.115), we find that

$$\mathbb{E}[Z_y\mathbf{1}_{y\in S_1(x)}] = \mathbb{E}[\mathbb{E}[Z_y|A_{xy}]\mathbf{1}_{y\in S_1(x)}] = \frac{fd}{N}\left(1+O\left(\frac{K}{\sqrt{d}}\right)\right)$$

We also have that

$$\mathbb{E}[Z_yZ_w\mathbf{1}_{y,w\in S_1(x)}] = \mathbb{E}\Big[\mathbb{E}[Z_y|H_{xy},H_{yw}]\mathbb{E}[Z_w|H_{xw},H_{yw}]\mathbf{1}_{y,w\in S_1(x)}\Big] = f^2\frac{d^2}{N^2}(1+O(d^{-1/2}))^2$$

We therefore have

$$\operatorname{Var}\Big(\frac{1}{d}\sum_{y\in S_1(x)} Z_y\Big) = \frac{1}{d^2}\operatorname{Var}\Big(\frac{1}{N}\sum_{y\in[N]} Z_y\mathbf{1}_{y\in S_1(x)}\Big) = O\Big(\frac{1}{\sqrt{d}}\Big).$$

Using Chebyshev's inequality as in (2.116) we see that there exists $\tau = \tau(R) > 0$ such that we have, with probability $1 - O(d^{-1})$,

$$\frac{1}{d}\sum_{y\in S_1(x)} Z_y \ge \frac{1}{2}\mathbb{E}\Big[\frac{1}{d}\sum_{y\in S_1(x)} Z_y\Big] \ge \tau.$$

Plugging this into (2.118), $\operatorname{Im}\big(\frac{1}{d}\sum_{y\in S_1(x)} m_y\big) \ge \tau/4$.
We use this pointwise lower bound to deduce a uniform lower bound in $\operatorname{Re} z$ as in the proof of Lemma 2.52. This concludes the proof. □

Proof of Lemma 2.29. Using Lemma 2.52 and the fact that $|T^c| \le \sqrt{N}$, (2.109) remains valid for some $c = c(R) > 0$. This proves the statement about Ξ_1. To prove the statement about Ξ_2, we use (ii) Proposition 2.12, Lemma 2.53 and $\sup_x m_x = O(1)$. By (2.26), $N^{-1}|\mathcal{T}_{\mathbf{a}}^c| = O(N^{-1/2})$ and we have

$$\operatorname{Im}\Big(\sum_{y\in\mathcal{T}_{\mathbf{a}}} H_{xy}^2 m^{[N]}\Big) \ge \operatorname{Im}\Big(\frac{1}{d}\sum_{y\in S_1(x)\cap\mathcal{T}_{\mathbf{a}}} m^{[N]}\Big) \ge \tau$$

for some $\tau > 0$. □

Lemma 2.54 (Imaginary part of the free convolution)**.** *For $z \in \mathbb{H}$,*

$$\operatorname{Im} \mathrm{m}_{\mathrm{fc}}(z) \ge \frac{1}{10}\mathrm{e}^{-\operatorname{Re}(z)^2/2}.$$

Proof. Let $z\in\mathbb{H}$ and $I = \operatorname{Im}(z+\mathrm{m}_{\mathrm{fc}}(z))$ and $R = \operatorname{Re}(z+\mathrm{m}_{\mathrm{fc}}(z))$. We denote by $d\mu_G$ the standard normal distribution. Using (2.108) and the fact that $I \ge 0$, we find

$$\begin{aligned}\operatorname{Im}(\mathrm{m}_{\mathrm{fc}}(z)) &= \int\frac{I}{(x-R)^2+I^2}d\mu_G(x) \ge \int_{[R-1-I^{-\alpha},R-1]}\frac{I}{(x-R)^2+I^2}d\mu_G(x)\\ &\ge \frac{I^{1-\alpha}}{\sqrt{2\pi}}\min_{x\in[R-1-I^{-\alpha},R-1]}\frac{e^{-x^2/2}}{(x-R+1)^2+I^2} \ge \frac{I^{1-\alpha}}{\sqrt{2\pi}}\frac{e^{-(\operatorname{Re} z)^2/2}}{I^2} = \frac{e^{-(\operatorname{Re} z)^2/2}}{\sqrt{2\pi}I^{1+\alpha}}.\end{aligned}$$

Since $I \le 3$ we have for $\alpha = 0$, $\operatorname{Im}\mathrm{m}_{\mathrm{fc}}(z) \ge \frac{e^{-(\operatorname{Re} z)^2/2}}{3\sqrt{2\pi}}$. □

Proof of Lemma 2.34. Using McDiarmid's inequality and the independence of the entries H_{xy}, $x, y \in [N]$, we can show that

$$\mathbb{P}\Big[\Big|\,\hat{\mathrm{m}}_{\mathrm{fc}}(v,z) - \mathbb{E}\,\hat{\mathrm{m}}_{\mathrm{fc}}(v,z)\Big| \geq t\Big] \leq C\mathrm{e}^{-ct^2 d},$$

for some constant $C, c > 0$. Setting $t = \mathcal{C}\sqrt{\log N/d}$ yields a very high probability bound.
Then we can use a Lindeberg to compare the expectations of $\hat{\mathrm{m}}_{\mathrm{fc}}(v,z)$ and $\hat{\mathrm{m}}_{\mathrm{fc}}(u,z)$ with $u \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. We find that

$$\Big|\mathbb{E}\,\hat{\mathrm{m}}_{\mathrm{fc}}(v,z) - \mathbb{E}\,\hat{\mathrm{m}}_{\mathrm{fc}}(u,z)\Big| \leq C\frac{1}{\sqrt{N}},$$

where the constant depends on R and can be chosen as $\mathrm{e}^{3R^2/2}$. We conclude by observing that $\mathbb{E}\,\hat{\mathrm{m}}_{\mathrm{fc}}(u,z) = \mathrm{m}_{\mathrm{fc}}(z)$ by definition of m_{fc} .

□

Chapter 3

Localization at the edge

In this chapter, we study the law of the extreme eigenvalues of the Laplacian matrix. We show that after an appropriate rescaling, the eigenvalue point process of L is asymptotically close to a sequence of Poisson Point Processes. Throughout this chapter, we consider the Erdős-Rényi graph on N vertices with connection probability d/N.

3.1 Main results

In this chapter, we prove statements about the statistics of extreme eigenvalues and eigenvectors of the normalized Laplacian matrix $\underline{L}$ which is defined through

$$\underline{L} := \frac{D - A - d}{\sqrt{d}}, \tag{3.1}$$

where A is the adjacency matrix of $\mathbb{G}$ and $D = \mathrm{Diag}(D_1, \dots, D_N)$ the diagonal matrix of the degrees defined through $D_x := \sum_y A_{xy}$.

Remark 3.1 (Convention on constants)**.** By convention, all objects introduced may depend implicitly on N, the number of vertices in the graph $\mathbb{G}$, unless explicitly mentioned otherwise.

The central idea of this chapter is to build a bijection between top eigenvalues and vertices that have large degrees. Following this idea, three different regimes appear.

(i) For $d \gg \log N$, the point process of the maximal and minimal eigenvalues can be well approximated by the PPP corresponding to independent normal variables.

(ii) For $d \asymp \log N$, the maximal degree of the graph has a bounded distribution and is with high probability not unique. At a scale much smaller than $d^{-1/2}$, the eigenvalues form a PPP .

(iii) For $d \ll \log N$, the maximal degree of the graph is deterministic and a PPP appears at scale $\sqrt{d \log N}$.

For the left edge of the spectrum, the idea of mapping the smallest eigenvalues to vertices with the smallest degree remains valid to some extent.

(i) For $d \geq \log N$, the techniques employed for the left edge of the spectrum remain valid for the right edge.

(ii) For $\frac{1}{2} \log N \ll d \leq \log N$, the smallest degree vertices are with high probability leaves. We analyze the neighborhood of the leaves on the graph to understand the distribution of the smallest eigenvalues. However, the number of leaves can become polynomially large which requires extra care in the analysis.

(iii) For $d \leq \frac{1}{2}\log N$, leaves are no longer the structures that generate the smallest eigenvalues. Trees that *dangle* from the macroscopic connected component are the key objects. We study this idea further in Chapter 4.

Remark 3.2 (Rigidification of extreme values of Poisson variables)**.** Let $(Y_i)_{i=1}^N$ be i.i.d. random variables with distribution $\mathcal{P}_d$, $i = 1, \dots, n$. Let us use this family as a toy model for the largest eigenvalues. Then if $d \gg \log n$ the maximum of the Y_i is with high probability unique and at the edge the rescaled Y_is form a PPP . If $d \asymp \log n$ the distribution of $\max_i Y_i$ has bounded support and with positive probability there exists many $j \in [n]$ such that $Y_j = \max_i Y_i$. Finally if $d \ll \log n$, the distribution of $\max_i Y_i$ is concentrated on 1 or 2 points and with high probability there are many $j \in [n]$ such that $Y_j = \max_i Y_i$. This emerging *rigidity* in the distribution of the extremes of Y_i is an adversarial mechanism when we want to distinguish top eigenvalues in critical regimes. See also [11, Remark 4.14].

Remark 3.3 (Conventions)**.** We use bold symbols to represent vectors of dimension N as well as sequences of real numbers indexed by N. For instance $\mathfrak{u}_+$ defined in (3.3) is a sequence of real numbers indexed by N.

The following deterministic objects are used for centering and scaling. Define the function

$$f : (0,\infty) \to \mathbb{R}, \quad f(x) = x\log x - (x-1) + \frac{1}{2d}\log(2\pi d x), \tag{3.2}$$

and let $x_* \in \mathbb{R}$ be the point above which f is increasing, i.e. the largest solution of the equation $x \log x = -\frac{1}{2d}$. We introduce the restrictions $f_- := f|_{[\frac{1}{2d}, x_*]}$, $f_+ := f|_{[x_*,\infty)}$ and

$$\mathfrak{u}_\pm := f_\pm^{-1}\Big(\frac{\log N}{d}\Big) \vee \frac{1}{d}. \tag{3.3}$$

The random variables D_x are well approximated by the Poisson law of parameter d (see Lemma B.5). The quantity $d\mathfrak{u}_+$ plays therefore the role of a proxy for the quantity $\max_{x\in[N]} D_x$, the largest degree of the graph. Similarly, $d\mathfrak{u}_-$ is a proxy for the smallest degree of the Erdős-Rényi graph, zero excluded. Various properties of $\mathfrak{u}_\pm$ are given in Section B.2. We see that the maximal and minimal entries on $\operatorname{Diag}\underline{L}$, are approximately given by

$$\max_{x\in[N]} \underline{L}_{xx} \approx v(\mathfrak{u}_+) \approx \begin{cases} \sqrt{2\log N}, & d \gg \log N \\ \frac{\log N}{\sqrt{d}\log(\log N/d)}, & d \ll \log N, \end{cases} , \quad \min_{x\in[N]} \underline{L}_{xx} \approx v(\mathfrak{u}_-) \approx \begin{cases} -\sqrt{2\log N}, & d \gg \log N \\ -\sqrt{d}, & d \leq \log N, \end{cases}$$

where we introduced the function

$$v : [0,\infty) \to [0,\infty), \quad t \mapsto \sqrt{d}(t-1). \tag{3.4}$$

We are now ready to introduce the point processes we will compare. As a general rule, eigenvalue processes are denoted by the Greek letter Φ and abstract Poisson point processes are denoted by the Greek letter Ψ. Superscript $+$ refers to point processes at the right edge (top eigenvalues), and $-$ to processes at the left edge (bottom eigenvalues). The scaling and centering we use for supercritical regimes, $d \gg \log N$, would not make sense for subcritical regimes. We therefore need to introduce different eigenvalue processes, depending on the regime we want to consider. The subscript sup refers to supercritical regimes and the subscript crit refers to critical regimes. For the right edge of the spectrum, the scaling and centering for critical, $d \asymp \log N$, and subcritical, $d \ll \log N$, remain the same and so $\Phi^+_{\mathrm{sub}} = \Phi^+_{\mathrm{crit}}$. On the other hand, for the left edge of the spectrum, we consider different processes depending on whether $d \geq \log N$ or $d \leq \log N$. We introduce the process Φ^-_{sub} later.

Definition 3.4 (Eigenvalue process at the edge)**.** We define the rescaled eigenvalue point process

$$\Phi^+ := \sum_{\lambda \in \operatorname{Spec}\underline{L}} \delta_{\tau(\lambda-\sigma)} \tag{3.5}$$

where the rescaling parameters are defined as

$$\sigma := v(\mathfrak{u}_+) + \frac{\mathfrak{u}_+}{v(\mathfrak{u}_+)}\Big(1 + \frac{1}{v(\mathfrak{u}_+)^2}\Big), \qquad \tau := \frac{\sqrt{d}v(\mathfrak{u}_+)^2}{\sqrt{\mathfrak{u}_+}}. \tag{3.6}$$

Our first result Theorem 3.8 shows that Φ_+ is asymptotically close to a sequence of Poisson Point Processes Ψ.

Definition 3.5 (Poisson reference processes)**.** Given $m \subseteq \mathbb{R}$ (which might depend on N), we define $\Psi_{\mathrm{crit},m}$ to be the sequence of Poisson Point Processes with intensity measures

$$\rho(\mathrm{d}s) = \sum_{k\in\mathbb{Z}} \mathfrak{u}_+^k g\big(s + k\tau\theta\big)\mathrm{d}s, \quad g(t) := \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}, \quad \theta := \frac{1}{\sqrt{d}}\Big(1 - \frac{1}{v(\mathfrak{u}_+)^2}\Big)$$

The density ρ is plotted for three different values of d in Figure 3.1.

Our results hold in a region $[-\kappa, +\infty)$ containing an expected number $\mathcal{K}$ of rescaled eigenvalues. Let

$$\mathcal{K} := \log\log N, \tag{3.7}$$

we define

$$\kappa_+ := -\inf\big\{s \in \mathbb{R} : \rho([-s,\infty)) \le \mathcal{K}\big\}. \tag{3.8}$$

We define $\mathcal{K}$ in order to push Theorem 3.8 as far as possible, i.e.for regimes d as small as $O(1)$. We could fine-tune the value of $\mathcal{K}$ for the other regimes, but we do not pursue this here.

Remark 3.6 (Thumb rules)**.** At this point, it might be useful to consider the order of magnitude of the different parameters introduced. The two most important ones are $\mathfrak{u}_+$ and $v(\mathfrak{u}_+)$ which are related by the equalities

$$v(\mathfrak{u}_+) = \sqrt{\mathfrak{u}_+}(d-1), \quad \mathfrak{u}_+ \asymp \begin{cases} \frac{\log N}{d\log(\log N/d)}, & d \le \frac{1}{2}\log N, \\ 1, & d \ge \frac{1}{2}\log N. \end{cases} \quad , \quad v(\mathfrak{u}_+) \asymp \sqrt{\log N} \vee \frac{\log N}{\sqrt{d}\log(\log N/d)}.$$

If $Z_i \overset{\text{i.i.d.}}{\sim} \mathcal{P}_d$, $i = 1, \dots N$, then we expect the maximum to satisfy $\max_i Z_i = d\mathfrak{u}_+$ and the area around $\mathfrak{u}_+$ to be populated by the Z_is in the following way $|\{i \in [N]\colon\ Z_i = \mathfrak{u}_+ - \ell\}| \equiv (\mathfrak{u}_+)^\ell$, $\ell \in \mathbb{Z}$.

The asymptotic closeness of the rescaled eigenvalue and reference processes can be given a precise meaning by introducing the metric of convergence of point processes on compact sets.

Definition 3.7. For Φ and Ψ two point processes and $\kappa > 0$, we define

$$\mathcal{D}_\kappa(\Phi,\Psi) := \sum_{n\in\mathbb{N}} 2^{-n} \sup_{s_1,\dots,s_n \ge -\kappa}\ \sup_{k_1,\dots,k_n\in\mathbb{N}} \Big|\mathbb{P}\Big(\bigcap_{i\in[n]} (\Phi([s_i,\infty)) \ge k_i)\Big) - \mathbb{P}\Big(\bigcap_{i\in[n]} (\Psi([s_i,\infty)) \ge k_i)\Big)\Big|.$$

We say that two point processes Ψ and Φ are *asymptotically close* if $\mathcal{D}(\Phi,\Psi) \to 0$, as $N \to +\infty$.

Theorem 3.8 (Eigenvalue right edge)**.** *There is a constant $K > 0$ such that for any constant $\varepsilon > 0$, if*

$$K \le d \le N^{\frac{1}{3}-\varepsilon},$$

then

$$\mathcal{D}_{\kappa_+}\big(\Phi^+, \Psi^+\big) \longrightarrow 0.$$

Remark 3.9 (Distribution of the maximal eigenvalue of $\underline{L}$)**.** Let $\Lambda := \max_{\lambda\in\mathrm{Spec}\,\underline{L}}$ be the maximal eigenvalue of $\underline{L}$. We can read off from Theorem 3.8 the law of the rescaled version of Λ. In the supercritical regime $d \gg \log N$, we have $\mathbb{P}(\mathbf{b}(X_1 - \mathbf{a}) \le s) = \mathrm{e}^{\mathrm{e}^{-s}} + o(1)$ for $\mathbf{b} = \sqrt{2\log N}$ and $\mathbf{a} = \mathbf{b} + \frac{1}{2\log N} + \frac{1}{2}\log\frac{d}{2\log N}$. This behavior and rescaling are reminiscent of the one observed for the extreme eigenvalue statistics of N independent normal variables (but with an extra small shift in our case). This result is similar to the one found in [20, Theorem 1.2].

In the critical and subcritical regimes, the relevant scaling of Λ is $X_1 := \tau(\Lambda - \sigma)$. In the critical regime

$d \asymp \log N$, the distribution of Λ does not admit a limit. Instead, it is a mixture of two distributions on different scales. The maximal degree in the graph has a law given by

$$\mathbb{P}\Big(\max_{x\in[N]} D_x - \lfloor d\mathfrak{u}_+\rfloor = t\Big) = c^t$$

for some constant $c \in (0,1)$ (c.f. Lemma B.7). This randomness influences Λ on a scale of order $d^{-1/2}$. If $\Delta := \max_{x\in[N]} D_x - \lfloor d\mathfrak{u}_+\rfloor$, the *randomly centered* eigenvalue $\tilde{X}_1 := \tau(\Lambda - \sigma - \theta(\Delta - d\mathfrak{u}_+))$ follows the law of the maximum of $\mathfrak{u}_+^{\Delta}$ independent normal variables.
In the subcritical regime, one must distinguish between resonant and nonresonant regimes (see [6, Remark 1.5] for a more detailed discussion). In the non-resonant regime, $\max_x D_x = K$ with high probability, for some deterministic K (that depends on N and d) and after a suitable affine rescaling, X_1 has a Gumble distribution. In the resonant regime, asymptotically with probability $1-1/\mathrm{e}$, the variable X_1 has a standard normal, and with probability $1/\mathrm{e}$ it has a Gumble distribution centered around $-\frac{v(\mathfrak{u}_+)^2}{\sqrt{\mathfrak{u}_+}}(1 - O\big((\log N)^{-1}\big)$.

We now move on to the eigenvectors of $\underline{L}$.

Theorem 3.10 (Eigenvector localization). *Let $\varepsilon > 0$, $\mathcal{K} \le (\log\log N)^{1/2}$, and*

$$(\log\log N)^{\frac{1}{5}+\varepsilon} \le d \le N^{1/3-\varepsilon}.$$

The $\mathcal{K}$ eigenvectors corresponding to the $\mathcal{K}$ highest eigenvalues of $\underline{L}$ are localized around some vertex $x \in [N]$, in the sense that

$$\big\|\mathbf{w}_{\lambda(i)}|_{B_{10}(x(i))}\big\| = o(1), \quad i \le \mathcal{K}.$$

To study the neighborhood around the smallest degree vertices we need to introduce new objects. Indeed when the number of *minimal vertices* becomes exponentially large, the statistics of the sphere of radius two around leaves are no longer well approximated by a normal law, as explained in Remark 3.2.
Let

$$\mathfrak{u}_\gamma(k) := f^{-1}\left(\frac{(1-\gamma)\log N + k\log d}{d}\right) \vee \frac{1}{d}, \quad \gamma \in [0,1),\ k \in \mathbb{N}^*, \tag{3.9}$$

and

$$v_\gamma := \sqrt{d}(\mathfrak{u}_\gamma - 1). \tag{3.10}$$

We abbreviate $\mathfrak{u}_\gamma := \mathfrak{u}_\gamma(1)$. The parameter v_γ will be the approximate value of $\max_{x\in\mathcal{L}, z\sim x} v_z$, where $\mathcal{L}$ denotes the set of leaves in $\mathbb{G}$ and $x \sim z$ means that the two vertices are connected.

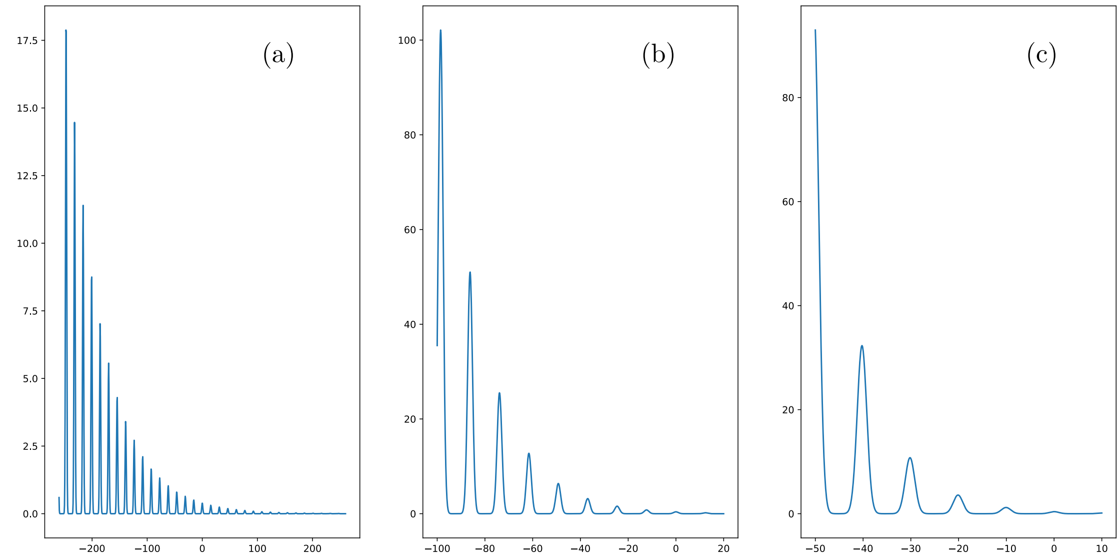
Definition 3.11 (Eigenvalue process at the left edge, leaves). Let $\gamma \ge 0$. We define the rescaled eigenvalue point process

$$\Phi^- := \sum_{\lambda\in\mathrm{Spec}(\underline{L})} \delta_{\tau_-(\sigma_- - \lambda)}, \qquad \Phi^\gamma := \sum_{\lambda\in\mathrm{Spec}(\underline{L})} \delta_{\tau_\gamma(\sigma_\gamma - \lambda)},$$

where the scaling parameters are defined as

$$\sigma_- := v(\mathfrak{u}_-) + \frac{\mathfrak{u}_-}{v(\mathfrak{u}_-)}\Big(1 + \frac{1}{v(\mathfrak{u}_-)^2}\Big), \qquad \tau_- := \frac{\sqrt{d}v(\mathfrak{u}_-)^2}{\sqrt{\mathfrak{u}_-}},$$

$$\sigma_\gamma := v_{(1)} + \frac{1}{d(v_{(1)} - v_\gamma)}\left[1 + \frac{\mathfrak{u}_\gamma}{v_{(1)}(v_{(1)} - v_\gamma)}\right], \qquad \tau_\gamma := \frac{d^2(v_{(1)} - v_\gamma)^2}{\sqrt{\mathfrak{u}_\gamma}}.$$



$d = 251$, $\mathfrak{u}_+ = 1.3$, $v(\mathfrak{u}_+) = 319$ $\quad$ $d = 18$, $\mathfrak{u}_+ = 2$, $v(\mathfrak{u}_+) = 37$ $\quad$ $d = 4, \mathfrak{u}_+ = 3$, $v(\mathfrak{u}_+) = 14$

Figure 3.1: An illustration of the density function $s \mapsto \rho(s)$, with $N = 10^4$. We plotted in (a) the supercritical regime, $d = N^{0.6}$, in (b) the critical regime, $d = 2\log(N)$, and in (c) the subcritical regime, $d = \frac{1}{2}\log N$. At first sight, the three plots look very similar. However, the scale of the y-axis is different in each case. In (a) the intensity of ρ remains smaller than 20 even for very small values of x and can be seen to be slightly positive even when $x > 50$. In addition, the intensity between different peaks does not seem to change a lot. This is because each term in the sum in the definition of ρ comes with a factor $\mathfrak{u}_+^k$ and $\mathfrak{u}_+ = 1 + o(1)$ in the supercritical regimes. In (b) and (c) the intensity of ρ becomes larger than 20 as soon as $x < 60$, respectively $x < 40$. This is caused by the fact that the coefficient $\mathfrak{u}_+$ increases as d decreases. Moreover, the height of the peaks varies markedly in (b) and even more vigorously in (c): indeed $\mathfrak{u}_+ > 1 + \varepsilon$, for $d \asymp \log N$ and $\mathfrak{u}_+ > 2$ for $d = o(\log N)$. Remember that we do not expect to see multiple vertices with the same degree (near $d\mathfrak{u}_+$) when $d \gg \log N$ but rather a cloud of points once we zoom out enough. On the other hand, we know (see Remark 3.2) that maximal degrees will accumulated on $O(1)$ possible values when $d \asymp \log N$ and then on one or two values when $d = o(\log N)$.
In addition, it is remarkable that the spacing between the bumps become wider as d becomes smaller. In (a), one can even imagine a continuous function by tracing a line between the peaks and the function e^{-x} seems to be a good approximation (see Remark 3.9). Comparing the above pictures with [7, Figure 1.1], we can better understand how the impact of the degrees on the limiting distribution of the maximal eigenvalue is much stronger in the case of the Laplacian than in the case of the adjacency matrix.

Let us define the intensity measures ρ_- and ρ_γ exactly as we defined ρ in Definition 3.5 but replacing τ and θ by τ_-, respectively τ_γ, and

$$\theta_- := \frac{1}{\sqrt{d}}\Big(1 - \frac{1}{v(\mathfrak{u}_-)^2}\Big),$$

respectively

$$\begin{aligned} \theta_\gamma := & \frac{1}{d(v_{(1)} - v_\gamma)} \cdot \\ & \left[1 + \frac{\mathfrak{u}_\gamma}{v_{(1)}(v_{(1)} - v_\gamma)} + \frac{1}{\sqrt{d}(v_{(1)} - v_\gamma)^3} + \frac{1}{\sqrt{d}(v_{(1)} - v_\gamma)^4} + \frac{1}{\sqrt{d}(v_{(1)} - v_\gamma)^2}\right]. \end{aligned} \tag{3.11}$$

We will see along the proof, that for values of d much larger than $\frac{1}{2}\log N$, the parameter θ_γ takes the much simpler form $\theta_\gamma := \frac{k}{d^{3/2}(v_{(1)}-v_\gamma)}$, since in that regime $v_{(1)} - v_\gamma \asymp \sqrt{d}$. However, in order to push our analysis as far as possible, we need the complicated formula from (3.11).
We also define κ_-, resp. κ_γ, and Ψ^-, resp. Ψ^γ, analogously to κ, defined in (3.8) and Definition 3.5.

Theorem 3.12 (Left edge of the spectrum)**.** *Let $\varepsilon > 0$.*

(i) If $\log N - (\log\log N)^2 \le d \le N^{-\varepsilon+1/3}$, then

$$\mathcal{D}_{\kappa_-}\left(\Psi^-, \Phi^-\right) \longrightarrow 0.$$

(ii) If $\frac{1}{2}\log N + (\log N)^\varepsilon \le d \le \log N - \log\log N$ and $\gamma := 1 - \frac{d}{\log N}$, then

$$\mathcal{D}_{\kappa_\gamma}\left(\Psi^\gamma, \Phi^\gamma\right) \longrightarrow 0.$$

For the next theorem, we denote by $\lambda(1) \ge \lambda(2) \ge \dots \lambda(N) = -\sqrt{d}$ the eigenvalues of $\underline{L}$ sorted decreasingly. We also denote by $x(i)$ the vertices of $\mathbb{G}$ sorted lexicographically first by $|S_1(x)|$ then by $|S_2(x)|$ and finally by $|S_3(x)|$.

Theorem 3.13 (Eigenvector localization)**.** *Let $\varepsilon > 0$, $\mathcal{K} \le (\log\log N)^{1/2}$, and*

$$\frac{1}{2}\log N + (\log N)^\varepsilon \le d \le N^{-1/3+\varepsilon}.$$

Then the following holds with high probability

1. *The $\mathcal{K}$ eigenvectors corresponding to the $\mathcal{K}$ smallest non-trivial eigenvalues of $\underline{L}$ are localized around some vertex $x \in [N]$, in the sense that*

$$\left\|\mathbf{w}_{\lambda(i)}|_{B_{10}(x(i))}\right\| = o(1), \quad i \le \mathcal{K}.$$

2. *In particular, for $d \le \log N - (\log\log N)$, the localization centers of those $\mathcal{K}$ eigenvectors are known to be leaves.*

Remark 3.14 (Lexicographic ordering of vertices)**.** Let us introduce the following three-level lexicographic ordering of the vertices $[N]$. For $x \in [N]$, the three levels are defined by $L_i(x) := |S_i(x)|$, $i = 1, 2, 3$, and $L_i(x) = +\infty$ if $S_i(x) = \emptyset$ (this case is relevant for isolated vertices). We can then compare any two vertices $x, y \in [N]$ by lexicographically ordering their $(L_1(x), L_2(x), L_3(x))$ and $(L_1(y), L_2(y), L_3(y))$.
As it turns out from our statical and geometric analysis of $\mathbb{G}$, this three-level ordering is enough to distinguish strictly, with high probability, the vertices from the sets

$$S_- := \{x \in [N] : D_x = \min_{y\in[N]} D_y + \omega\}, \quad \text{and} \quad S_+ := \{x \in [N] : D_x = \max_{x\in[N]} D_x - \omega\}, \quad \omega = \log\log N.$$

Theorem 3.10 could be stated as the $\mathcal{K}$ eigenvectors corresponding to the $\mathcal{K}$ largest eigenvalues are localized around the $\mathcal{K}$ largest vertices with respect to the three-level lexicographic ordering. An analogous reformulation in terms of the smallest vertices with respect to the three-level lexicographic ordering would also be valid for Theorem 3.13.

3.2 Block decomposition

In this section, we construct an approximate block diagonal decomposition of $\underline{L}$ in some bases of approximate eigenvectors. The first decomposition, Proposition 3.15 is relevant for small and large eigenvalues in the regime $\log N \ll d \ll N^{1/6}$. The second decomposition, Proposition 3.18 is relevant only for large eigenvalues, but in the regime $1 \ll d \ll (\log N)^2$. The third decomposition, Proposition 3.21 is relevant for small eigenvalues in the regime $\frac{1}{2}\log N \ll d \ll (\log N)^2$. In the regime $N^{1/6} \leq d \ll N^{1/3}$, we do not manage to prove a block diagonal decomposition but we can nevertheless describe the extreme eigenvalues (see Proposition 3.16). The proofs of all three propositions are deferred to Section **??**. Each proposition is followed by a corollary that makes explicit the matching between the eigenvalue process of $\underline{L}$ at the edge and the point process of approximate eigenvalue generated by top, respectively bottom, degree vertices.

For $\varepsilon > 0$ we introduce the vertex sets $\mathcal{U}^+(\varepsilon)$ and $\mathcal{U}^-(\varepsilon)$ which represent the vertices that have a large, respectively small, degree. As we will see, the set $\mathcal{U}^{\pm}(\varepsilon)$ are with high probability r-packings in $\mathbb{G}$, for any $r = O(1)$. This property is crucial for the construction of the block diagonal approximations. Let

$$\mathcal{U}^{\pm}(\varepsilon) := \big\{x \in [N] \colon |v_x| > \mathfrak{u}^{\pm}(\varepsilon), \operatorname{sign}(v_x) = \pm\big\}, \tag{3.12}$$

where

$$\mathfrak{u}^+(\varepsilon) := \max\Big\{\sqrt{(1+\varepsilon)\log N},\ \sqrt{d}\Big(\frac{1+\varepsilon}{2}\mathfrak{u}_+ - 1\Big)\Big\},$$

$$\mathfrak{u}^-(\varepsilon) := \begin{cases} \min\Big(\sqrt{(1+\varepsilon)\log N}, \varepsilon\sqrt{d}\Big) & \text{if } d \geq \log N - \log\log N \\ \sqrt{d} - \frac{1}{\sqrt{d}\varepsilon} & \text{if } d < \log N - \log\log N. \end{cases}$$

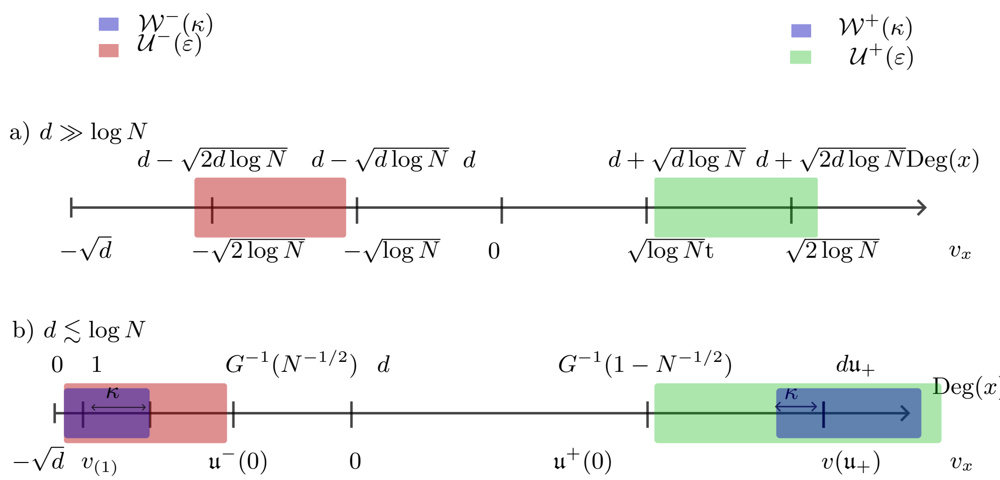


Figure 3.2: Illustration of the different vertex subgroups used in the block diagonal approximation of $\underline{L}$. As is apparent from the figure, the distribution of the degrees is symmetric around d in the dense regime. In the sparse regime, the distribution of the degrees is skewed towards small degrees. We denote by G the distribution function of the degrees. We choose $\mathfrak{u}^+(\varepsilon)$ and $\mathfrak{u}^-(\varepsilon)$ in such a way that vertices in $\mathcal{U}^{\pm}(\varepsilon)$ are never close one from the other.

In dense regimes, the distribution of the degrees of $\mathbb{G}$ is well approximated by a Gaussian random variable of mean d and variance d and is symmetric around d.

Proposition 3.15 (Block decomposition for $d \gg \log N$)**.** *Let $\varepsilon > 0$ and*

$$(\log N)^{1+\varepsilon} \leq d \leq N^{\frac{1}{6}-\varepsilon} \tag{3.13}$$

For $\tau \in (1-2\varepsilon, 1)$, the following holds with high probability. The exists a Hermitian matrix M and an orthogonal matrix U such that

$$\|\underline{L} - M\| = O\big(d^{-1/2}\big), \qquad UMU^* = \begin{bmatrix} \nu & 0 & E_d^* \\ 0 & \mathcal{D} & E^* \\ E_d & E & X \end{bmatrix}, \tag{3.14}$$

where

(i) $\mathcal{D} = \mathrm{Diag}\big(v_x + \frac{\alpha_x}{v_x} + \varepsilon_x \colon x \in \mathcal{U}^+(\tau) \cup \mathcal{U}^-(\tau)\big)$ *and* $\max_x |\varepsilon_x| = O\big(\frac{1}{\log N}\big)$.

(ii) $\|X\| = \sqrt{(1+\tau)\log N} + O(1)$ *and* $\|E\| = O\big((\log N)^{-1}\big)$;

(iii) $\nu + \sqrt{d} = O(d^{-1/2})$ *and* $\|E_d\| = O(d^{-1})$.

We can push further the analysis of the extreme eigenvalues of $\underline{L}$ to the regime $d \ll N^{1/3}$. However, we do not obtain a block diagonal decomposition of $\underline{L}$. The reason for that is the fact that the balls around large-degree vertices are too close to obtain good bounds on the norm of the off-diagonal block E.

Proposition 3.16 (Very dense regime)**.** *Let* $\varepsilon > 0$*,* $\tau \in (1-\varepsilon/2, 1)$ *and* $N^\varepsilon \leq d \leq N^{\frac{1}{3}-\varepsilon}$*. Let* $I^\pm$ *be the* $[N]$*-valued random variable defined by*

$$I^+ := \inf\big\{i \in [N] \colon v_{(i)} \leq \mathfrak{u}_+(1-\tau/2)\big\}, \quad I^- := \inf\big\{i \in [N] \colon v_{(N-i)} \geq \mathfrak{u}_-(1-\tau/2)\big\}.$$

Then there exist error terms $\varepsilon_i^\pm$*,* $i \in [I^\pm]$ *such that* $\max_i |\varepsilon_i^\pm| = O\big((\log N)^{-3/2} + d^{-1}\big)$ *and*

$$\lambda_{(i)} = v_{(i)} + \frac{1}{v_{(i)}} + \varepsilon_i^+, \quad \lambda_{(N-i+1)} = v_{(N-i)} + \frac{1}{v_{(N-i)}} + \varepsilon_i^-.$$

Recalling the scaling parameters τ_{sup} and σ_{sup} from Definition 3.11, let us define the intervals

$$\chi_{+,\text{dense}}(c_*) := \big[\sigma_{\text{sup}} - c_*\sqrt{\log N}, +\infty\big), \qquad \chi_{-,\text{dense}}(c_*) := \big(-\infty, -\sigma_{\text{sup}} + c_*\sqrt{\log N}\big].$$

Corollary 3.17. *Let* $\varepsilon > 0$ *and* $(\log N)^{1+\varepsilon} \leq d \leq N^{1/3-\varepsilon}$*. Then with high probability, there exists a constant* $c_* > 0$*, error terms* ε_x*,* $x \in [N]$*, such that* $\max_x |\varepsilon_x| = O\big((\log N)^{-\frac{1+\varepsilon}{2}}\big)$ *and the processes*

$$\sum_{\lambda \in \mathrm{Spec}\,\underline{L} \setminus \{-\sqrt{d}\}} \delta_\lambda, \quad \textit{and} \quad \sum_{x \in \mathcal{U}^+(\tau) \cup \mathcal{U}^-(\tau)} \delta_{v_x + \frac{\alpha_x}{v_x} + \varepsilon_x}$$

coincide on the spectral domain $\chi_{+,\text{dense}}(c_*) \cup \chi_{-,\text{dense}}(c_*)$*.*

Proof. We first prove the statement for d as in (3.13). By (3.14) and perturbation theory, Lemma A.7, we know that the eigenvalue processes of M and $\underline{L}$ coincide up to a small error

$$\sum_{\lambda \in \underline{L}} \delta_\lambda = \sum_{\mu \in M} \delta_{\mu + \varepsilon_\mu}, \quad \max_\mu |\varepsilon_\mu| = O\big(d^{-1/2}\big).$$

Moreover if $Y := \begin{pmatrix} \nu & 0 & 0 \\ 0 & \mathcal{D} & 0 \\ 0 & 0 & X \end{pmatrix}$, then again by perturbation theory, we see that the eigenvalue process $\sum_{\mu \in \mathrm{Spec}(M)} \delta_\mu$ coincides with the process $\sum_{\tilde{\mu} \in \mathrm{Spec}(Y)} \delta_{\tilde{\mu} + \varepsilon_{\tilde{\mu}}}$, Now observe that for $c_* = \tau/2$, by (ii), we have $\nu \notin \chi_{\pm,\text{dense}}(c_*)$ and, by (iii),

$$\mathrm{Spec}\, X \cap \big(\chi_{+,\text{dense}}(c_*) \cup \chi_{-,\text{dense}}(c_*)\big) = \emptyset.$$

Therefore, using (i), we see that the processes

$$\sum_{\tilde{\mu} \in \mathrm{Spec}(Y)} \delta_{\tilde{\mu}}, \quad \text{and} \quad \sum_{x \in \mathcal{U}^\pm(\tau)} \delta_{v_x + \frac{\alpha_x}{v_x} + \varepsilon_x}$$

where $|\varepsilon| = O\big((\log N)^{-1}\big)$, coincide on $\chi_{\pm,\text{dense}}(c_*)$.

We now prove the statement for $N^{1/6-\varepsilon} \leq d \leq N^{1/3-\varepsilon}$. This is immediately follows from Proposition 3.16, the definition of $I_\pm$, the choice $c_* = \tau/2$ and the fact that $|\alpha_x - 1| = O\Big(\sqrt{\frac{\log N}{d}}\Big) = O\Big(\big(\log N\big)^{-1}\Big)$. We skip the details. This concludes the proof. ∎

To obtain good approximate eigenvalues in sparse regimes, we will need stronger estimates than those provided by Proposition 3.15. Such estimates can only be achieved by selecting an even more restrictive subset of vertices. For the rest of this section, we introduce the parameter κ which we use to tune the size of this subset. Let

$$\kappa = (\log\log N)^2 \vee \frac{\log\log N}{\log(\mathfrak{u}_+)}. \tag{3.15}$$

and

$$\mathcal{W}^+(\kappa) := \{x \in [N] : D_x \geq d\mathfrak{u}_+ - \kappa\}, \qquad \mathcal{W}^-(\kappa) := \{x \in [N] : D_x \leq d\mathfrak{u}_- + \kappa\}. \tag{3.16}$$

Observe that for $d \ll (\log\log N)^2 \log N$, the first term in (3.15) dominates and $\kappa \leq \frac{\log N}{\sqrt{d}\log(\log N/d)}$. On the other hand $d \gg \log N$, we always have $\kappa \asymp \frac{\sqrt{d}\log\log N}{\sqrt{\log N}}$. We conclude that $\mathcal{W}^+ \subseteq \mathcal{U}^+(\varepsilon)$ for any constant $\varepsilon \in (0,1)$ in all regimes (see Lemma 3.50).
For $x \in [N]$, we define $S_i(x) := \{y \in [N]\colon\ \operatorname{dist}(x,y) = i\}$, $i \geq 0$, where dist is the usual graph distance $\mathbb{G}$, as well as

$$\alpha_x := \frac{D_x}{d}, \qquad \beta_x := \frac{|S_2(x)|}{|S_1(x)|d}. \tag{3.17}$$

Note that $v_x = \sqrt{d}(\alpha_x - 1)$. We can also define the functions

$$\alpha(t) = \frac{t}{d}, \quad v(t) := \sqrt{d}\big(\alpha(t) - 1\big), \quad \rho(t) := \frac{\alpha(t)}{v(t)}, \quad t \geq 0. \tag{3.18}$$

We define the *approximate eigenvalue around a vertex x* as the function

$$\Lambda_x := \Lambda_x(\alpha_x, \beta_x) = v_x + \frac{\alpha_x}{v_x}\left(1 + \frac{1}{v_x^2}\right) + \frac{\sqrt{d}\alpha_x}{v_x^2}(\beta_x - 1). \tag{3.19}$$

This quantity can be understood as the extreme eigenvalue of an (α_x, β_x)-rooted tree.
Let us introduce the generic error parameter

$$\omega(\alpha) := d^{2\alpha} \vee \log\log N, \quad \alpha \in (0,1). \tag{3.20}$$

Proposition 3.18 (Block decomposition for top eigenvalues)**.** *For $\alpha \in (0, 1/12)$ and κ as in (3.15), there exist $K := K(\alpha) \geq 1$ and $c_* := c_*(\alpha,\kappa) > 0$ such that if $\varepsilon > 0$ and*

$$K \leq d \leq (\log N)^2,$$

then the following holds with probability $1 - O(\mathrm{e}^{-c_*\omega(\alpha)^2})$. There exists an orthogonal matrix U such that

$$U^{-1}\underline{L}U = \begin{bmatrix} \mathcal{D}_{\mathcal{W}} & 0 & E_{\mathcal{W}}^* \\ 0 & \mathcal{D}_{\mathcal{U}} & E_{\mathcal{U}}^* \\ E_{\mathcal{W}} & E_{\mathcal{U}} & X \end{bmatrix}, \tag{3.21}$$

where

(i) $\mathcal{D}_{\mathcal{W}} = \operatorname{Diag}\big(\Lambda_x + \varepsilon_x\colon\ x \in \mathcal{W}^+(\kappa)\big)$, *and*

$$\max_{x\in\mathcal{W}^+} |\varepsilon_x| = O\left(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\right) \tag{3.22}$$

(ii) $\mathcal{D}_{\mathcal{U}} = \operatorname{Diag}\big(v_x + \frac{\alpha_x}{v_x} + \varepsilon_x\colon\ x \in \mathcal{U}^+(\varepsilon) \setminus \mathcal{W}(\kappa)\big)$. *and* $\max_{x\in\mathcal{U}^+} |\varepsilon_x| = O\left(\frac{\omega(\alpha)}{\sqrt{d\log N}}\right)$,

(iii) $\|E_\sharp\| = O\big(\frac{\omega}{\log N}\big)^4$ *for* $\sharp \in \{\mathcal{U}, \mathcal{W}\}$ *and*

$$\operatorname{Spec} X \subseteq \Big(-\infty, \mathfrak{u}^+(\varepsilon)(1+\varepsilon)\Big]. \tag{3.23}$$

Moreover $\sup_{x\in\mathcal{W}^+(\kappa)}|\Lambda_x - v_x + \frac{\alpha_x}{v_x}| = O(\frac{\omega}{\log N})$ *and* $\sup_{x\in\mathcal{W}^+(\kappa)}|\beta_x - 1| = O\left(\frac{\kappa \log\log N}{\log N}\right)$.

Let us define the intervals

$$\chi_+ := \big[v(\mathfrak{u}_+) - \frac{\kappa}{2\sqrt{d}}, +\infty\big]. \tag{3.24}$$

Corollary 3.19. *For* $\alpha \in (0, 1/12)$*, there exists* $K := K(\alpha) > 1$*, such if* $K \le d \le \log N$*, with high probability there exist error terms* ε_x*,* $x \in [N]$*, such that*

$$\max_x|\varepsilon_x| = O\left(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\right),$$

and the processes

$$\sum_{\lambda\in\operatorname{Spec}\underline{L}} \delta_\lambda \quad \textit{and} \quad \sum_{x\in\mathcal{W}^+(\kappa)} \delta_{\Lambda_x+\varepsilon_x}$$

coincide on the spectral domain $\chi_+(\kappa)$*.*

Proof. By definition, $x \in \mathcal{W}(\kappa)$ if and only if $v_x := \sqrt{d}(\mathfrak{u}_+ - 1) - \frac{\kappa}{\sqrt{d}}$. Moreover for any $x \in \mathcal{U}(\varepsilon) \setminus \mathcal{W}(\kappa)$, $\varepsilon > 0$, then, using (ii),

$$v_x + \frac{\alpha_x}{v_x} + \varepsilon_x \le v(\mathfrak{u}_+) - \frac{\kappa}{\sqrt{d}}\left(1 + O\left(\frac{1}{\sqrt{d\log N}}\right)\right) \le v(\mathfrak{u}_+) - \frac{\kappa}{2\sqrt{d}}.$$

We can conclude as in Corollary 3.17 using the estimates from Proposition 3.18. □

For regimes $\frac{1}{2}\log N \le d \le \log N$, we will show that bottom eigenvalues are in bijection to some subset of the set of leaves

$$\mathcal{L} := \big\{x \in [N] : D_x = 1\big\}. \tag{3.25}$$

For for $d = \gamma \log N$, with $\gamma \in (\frac{1}{2}, 1)$, the number of leaves becomes polynomial as $\mathbb{E}|\mathcal{L}| \asymp N^{1-\gamma}$. This has two consequences. First of all, we should expect the eigenvalue spacing to become polynomially small, which would render error estimates such as those found in Proposition 3.18 useless to distinguish eigenvalues. Secondly, the sphere of radius two around $x \in \mathcal{L}$ having a distribution close to a $\mathcal{P}_d$ variable, the observable β_x will become subject to the rigidity phenomenon as described in Remark 3.2.
Both these issues are tackled by looking at extreme value statistics of the family of random variables $\{\alpha_{z_x} : x \in \mathcal{L}\}$ where z_x denotes the unique neighbor of such x. Inside $\mathcal{W}^-$, we define the set

$$\mathcal{W}^\gamma(n) := \Big\{x \in \mathcal{W}^-(\kappa) \cap \mathcal{L} : \text{ and for } z \text{ the unique neighbor of } x, \, |D_z - d\mathfrak{u}_\gamma| \le n\Big\}, \quad n > 0. \tag{3.26}$$

The set $\mathcal{W}^\gamma$, $\gamma := \frac{d}{\log N}$, is only meaningful in regimes where $\gamma \le 1$ and we set the convention that $\mathcal{W}^\gamma = \emptyset$ if $\gamma \ge 1$.

Definition 3.20 (Approximate eigenvalues for leaves)**.** For $x \in \mathcal{L}_1$ and z_x denote its unique neighbor. We define

$$\begin{aligned}\Lambda_x^{\mathcal{L}} := \Lambda_x^{\mathcal{L}}(\alpha_{z_x}, \beta_{z_x}) &= v_{(1)} + \frac{1}{d(v_{(1)} - v_{z_x})}\left(1 - \frac{1}{d(v_{(1)} - v_{z_x})^2} + \frac{\alpha_{z_x}(1 + \frac{1}{\sqrt{d}v_{(1)}} + \frac{1}{v_{(1)}^2})}{v_{(1)}(v_{(1)} - v_{z_x})}\right) \\ &\quad + \frac{\alpha_{z_x}}{\sqrt{d}v_{(1)}(v_{(1)} - v_{z_x})^2}(\beta_{z_x} - 1), \\ \tilde{\Lambda}_x^{\mathcal{L}} := \tilde{\Lambda}_x^{\mathcal{L}}(\alpha_{z_x}) &= v_{(1)} + \frac{1}{d}\frac{1}{v_{(1)} - v_{z_x}} + \frac{1}{dv_{(1)}(v_{(1)} - v_{z_x})^2}\end{aligned} \tag{3.27}$$

where $v_{(1)}$ was introduced at the end of Definition 3.11.

Proposition 3.21 (Block decomposition for bottom eigenvalues)**.** *Let $K = O(1)$ and*

$$\frac{1}{2}\log N + (\log N)^{1/K} \leq d \leq (\log N)^2. \tag{3.28}$$

For $\varepsilon \in (0, \frac{1}{10})$, and $r \geq 10K$, $\gamma := \frac{d}{\log N}$ and v_γ as defined in Definition 3.11 and

$$\kappa := \begin{cases} \frac{1}{2\varepsilon} & \text{if, } d \leq \log N - \log\log N \\ \kappa & \text{else,} \end{cases} \tag{3.29}$$

there exists $c_* := c_*(K) > 0$ such that the following holds with probability $1 - O\big(\mathrm{e}^{-c_*(\log N)^{c_*}}\big)$. There exists an orthogonal matrix U such that

$$\underline{L} = U^{-1}\begin{bmatrix} \nu & 0 & 0 & E_\nu^* & 0 \\ 0 & D_{\mathcal{W}} & 0 & E_{\mathcal{W}}^* & 0 \\ 0 & 0 & D_{\mathcal{U}} & E_{\mathcal{U}}^* & 0 \\ E_\nu & E_{\mathcal{W}} & E_{\mathcal{U}} & X & 0 \\ 0 & 0 & 0 & 0 & Y \end{bmatrix} U, \quad D_{\mathcal{W}} = \begin{bmatrix} \mathcal{W}_1 & 0 & 0 \\ 0 & \mathcal{W}_2 & 0 \\ 0 & 0 & \mathcal{W}_3 \end{bmatrix}, \tag{3.30}$$

where

(i) $\mathcal{W}_1 = \mathrm{Diag}\Big(\Lambda_x^{\mathcal{L}} + \varepsilon_x\colon\ x \in \mathcal{W}^\gamma(\kappa)\Big)$,

$$\max_{x\in\mathcal{W}^\gamma(\kappa)}|\varepsilon_x| = O\bigg(\frac{\omega(\alpha)^3}{(\log N)^3(v_{(1)} - v_\gamma)^2}\bigg), \quad (\beta_{z_x} - 1) = O\bigg(\frac{\kappa\log d}{(\log N)^{3/4}}\bigg). \tag{3.31}$$

(ii) $\mathcal{W}_2 = \mathrm{Diag}\,\big(\tilde{\Lambda}_x^{\mathcal{L}} + \varepsilon_x\colon\ x \in \mathcal{L}\setminus\mathcal{W}^\gamma(\kappa)\big)$,

$$\max_x|\varepsilon_x| = O\left(\frac{\omega(2\alpha)}{(\log N)^{2-\frac{1}{6}}(v_{(1)} - v_\gamma)^2}\right).$$

(iii) $\mathcal{W}_3 = \mathrm{Diag}\,(\Lambda_x + \varepsilon_x\colon\ x \in \mathcal{W}^-(\kappa)\setminus\mathcal{L})$ *and*

$$\max_x|\varepsilon_x| = O\bigg(\frac{\omega}{d(\log N)^2}\mathbf{1}_{d\geq\log N - (\log\log N)^2} + \frac{1}{(\log N)^{\frac{1}{2}+2c_*}}\mathbf{1}_{d\leq\log N - \log\log N}\bigg).$$

(iv) $\mathcal{D}_{\mathcal{U}} = \mathrm{Diag}\,\Big(v_x + \frac{\alpha_x}{v_x} + \varepsilon_x\colon\ x \in \mathcal{U}^-(\varepsilon)\setminus\mathcal{W}^-(\kappa)\Big)$ *and* $\max_{x\in\mathcal{U}}|\varepsilon_x| = O\big((\log N)^{-\frac{1}{2}-2c_*}\big)$.

(v) The submatrix Y is of size $O\Big(Nd^4\mathrm{e}^{-d}\Big)$, $\|E_\sharp\| = O\Big(\big(\frac{\omega}{\log N}\big)^{r/2}\Big)$, for $\sharp \in \big\{\nu, \mathcal{W}, \mathcal{U}\big\}$, and

$$\nu = -\sqrt{d} + O\Big(N^{-\frac{1}{5}}\Big), \quad \mathrm{Spec}\,X \subseteq \bigg[-\sqrt{d}(\mathfrak{u}_- - 1) + \frac{3}{2}\kappa, +\infty\bigg), \quad \mathrm{Spec}\,Y = \Big\{-\sqrt{d}\Big\}. \tag{3.32}$$

(vi) For $d \geq \log N - (\log\log N)^2$,

$$\sup_{x\in\mathcal{W}^-}|\Lambda_x - v_x - \frac{\alpha_x}{v_x}| = O\Big(\frac{\omega}{\log N}\Big) \qquad \max_{x\in\mathcal{W}^-\cap\mathcal{L}}|\Lambda_x - \Lambda_x^{\mathcal{L}}| \vee |\Lambda_x^{\mathcal{L}} - \tilde{\Lambda}_x^{\mathcal{L}}| = O\bigg(\frac{\omega^3}{(\log N)^{5/2}}\bigg). \tag{3.33}$$

We could push this result further by only asking that $d \geq \frac{1}{2}\log N\big(1 + \frac{C}{\log d}\big)$ for some large enough constant $C \geq 0$. This leads to more complicated error terms and we do not push it further. We are content

with pushing the lower bound on d to $(\frac{1}{2}+o(1))\log N$.
Let us define the interval

$$\chi_- := \big(-\infty, \sqrt{d}(\mathfrak{u}_- - 1) + \frac{\kappa}{2\sqrt{d}}\big], \quad \chi_\gamma := \big(-\infty, \sqrt{d}(\mathfrak{u}_- - 1) - \chi_\gamma\big], \tag{3.34}$$

where

$$\chi_\gamma := \frac{1}{d(v_{(1)} - v_\gamma)}\left[1 + \frac{\kappa}{2\sqrt{d}}\frac{1}{(v_{(1)} - v_\gamma)}\right].$$

Note that $\sqrt{d}(\mathfrak{u}_- - 1) = -\sqrt{d} + \frac{1}{\sqrt{d}} = v_{(1)}$ and $|v_\gamma| = O\big(\omega^2\big)$ when $d \le \log N - \log\log N$.

Corollary 3.22. *Let $K > 0$. If $K = O(1)$ and $\frac{1}{2}\log N + (\log N)^{1/K} \le d \le (\log N)^2$ there exists $c_* := c_*(K) > 0$ such that the following holds with probability $1 - O\big(\mathrm{e}^{-c_*(\log N)^{c_*}}\big)$.*

(i) If $d \ge \log N - (\log\log N)^2$, the processes

$$\sum_{\lambda \in \operatorname{Spec} \underline{L}} \delta_\lambda \quad \text{and} \quad \delta_{-\sqrt{d}} + \sum_{x \in \mathcal{W}^-(\kappa)} \delta_{\Lambda_x + \varepsilon_x}$$

coincide on the spectral domain χ_- where ε_x, $x \in [N]$, are error terms satisfying $\max_{x\in\mathcal{W}^-}|\varepsilon_x| = O\big(\frac{\omega}{d(\log N)^2}\big)$.

(ii) If $d \le \log N - (\log\log N)$ and $\gamma = 1 - \frac{d}{\log N}$ there exist $C \ge 0$ and $M = O(N^{4\gamma/3})$ such that the processes

$$\sum_{\lambda \in \operatorname{Spec} \underline{L}} \delta_\lambda \quad \text{and} \quad M\delta_{-\sqrt{d}} + \sum_{x \in \mathcal{W}^\gamma(\kappa)} \delta_{\Lambda_x^{\mathcal{L}} + \varepsilon_x} \tag{3.35}$$

coincide on the spectral domain χ_e where ε_x, $x \in [N]$, are error terms satisfying $\max_{x\in\mathcal{W}^\gamma(\kappa)}|\varepsilon_x| = O\big(\frac{\omega}{(v_{(1)}-v_\gamma)^2(\log N)^3}\big)$.

Proof. Let $\varepsilon > 0$ be some small constant such that $\frac{1}{2\varepsilon} \ge \kappa$. The first point is proved the same way as Corollary 3.19. In this regime, we use(vi) to replace $\tilde{\Lambda}_x^{\mathcal{L}}$ and $\Lambda_x^{\mathcal{L}}$ for $x \in \mathcal{L}$ by Λ_x and we conclude by an argument analog to the one in the proof of Corollary 3.17.
For the second point first observe that, for any $c > 0$, $M = O\Big(N^{4\gamma/3}\Big)$, by ((v)). Removing the contribution of $\operatorname{Spec} Y$ correspond to the first sum of the second process. Let $x \in \mathcal{L} \setminus \mathcal{W}^\gamma(2\kappa)$ and $y \in \mathcal{W}(\kappa) \setminus \mathcal{L}$, and $z \in \mathcal{U} \setminus \mathcal{W}(\kappa)$. Then

$$\begin{aligned}\tilde{\Lambda}_x^{\mathcal{L}} &\ge v_{(1)} - \frac{1}{d(v_{(1)} - v_\gamma)} - \frac{\kappa}{d^{3/2}(v_{(1)} - v_\gamma)^2} + O\left(\frac{(\log\log N)^4 + \kappa^2}{(\log N)^{2-\frac{1}{6}}(v_{(1)} - v_\gamma)^2}\right)\\ &\ge v_{(1)} - \chi_\gamma,\end{aligned}$$

where we used $2 - \frac{1}{6} \ge \frac{3}{2}$ in the second inequality. Moreover

$$\Lambda_y \wedge v_z + \frac{\alpha_z}{v_z} \ge v_{(1)}.$$

These bounds remain correct even after we add the respective error terms ε_x. These bounds combined with those on $\|E_\sharp\|$ in (v), guarantee that only terms from $\mathcal{W}_1$ contribute in the interval χ_γ. We conclude with an argument analog to the one in the proof of Corollary 3.17. □

Remark 3.23. However for $d \le \frac{1}{2}\log N$, some fundamental differences arise. First of all, the number of disconnected components and their variety increases. But more importantly, even the left edge spectrum of $\underline{L}|_{\mathbb{G}_{cc}}$ exhibits different behavior. For $d \le \frac{1}{k}\log N$, $k \in \mathbb{N}^*$, some trees of size k start to *dangle* from $\mathbb{G}_{cc}$. They do not form disconnected components but only have a few connections to the rest of the graph. These configurations induce eigenvalues smaller than the eigenvalues induced by leaves. This phenomenon is explored further in Chapter 4.

Before turning to the proofs of the various propositions stated in this section, let us introduce some notations and *a priori* estimates.

We denote by $\underline{A} := \frac{1}{\sqrt{d}} A$ and by $\underline{D} := \frac{1}{\sqrt{d}}(D - d)$.

Definition 3.24 (Restriction of a matrix)**.** Let $N \in \mathbb{N}^*$, $M \in \mathbb{C}^{N\times N}$ and $T \subseteq [N]$. We define $M|_T$ to be the T-by-T matrix with

$$(M|_T)_{ij} = M_{ij}, \quad i, j \in T.$$

Any vector $\mathbf{w} = (w_x)_{x\in[T]}$ defined on some subset $T \subset [N]$ is naturally extended on $[N]$ by $\overline{\mathbf{w}} = (\overline{w}_x)_{x\in[N]}$, with $\overline{w}_x = w_x$ if $x \in T$ and zero otherwise. In general, we do not write the overline.

We also need the following bound on the operator norm of the adjacency matrix.

Proposition 3.25 (Bound on $\|\underline{A} - \mathbb{E}\underline{A}\|$)**.** *For $4 \le d \le \frac{1}{2}N$, we have, with very high probability*

$$\|\underline{A} - \mathbb{E}\underline{A}\| \le \begin{cases} 3 & \text{if } d \ge (\log N)^5 \\ 2 + \mathcal{C}\sqrt{\log N/d} & \text{if } \frac{1}{2}\log N \le d \le (\log N)^5 \\ 1 + \mathcal{C} + \mathcal{C}\frac{\log N}{d(\log\log N - \log d)} & \text{if } 4 \le d \le \frac{1}{2}\log N. \end{cases} \tag{3.36}$$

Proof. The first case is obtained by applying [14, Lemma G.2] with $\tilde{H} := \underline{A} - \mathbb{E}\underline{A}$. Since $\max_{i,j}\|\tilde{H}_{ij}\| = O(d^{-1/2}) = O((\log N)^{-5/2})$, the hypotheses are satisfied. The last two cases are [6, Corollary 6.2]. ☐

3.3 Dense regimes and min-max principle

In this section, we prove Proposition 3.16. While both Propositions 3.16 and 3.15 combine to give Corollary 3.17, the techniques used in their respective proofs are very different.

The proof of Proposition 3.16 relies on the min-max characterization of the eigenvalue problem and the graph constructed in Proposition 5.15. A crucial ingredient needed to derive the upper bound on $\lambda_{(1)}$ is the fact that, for $\tau > 0$ small enough, for any two large degrees $x, y \in \mathcal{U}(\tau)$, the sphere of radius one around y does not overlap with the sphere of radius 2 around x (see (i) of Proposition 5.16).

Proof of Proposition 3.16. Let $\mathbb{G}^\tau$ be the graph defined by Proposition 5.15 and M its rescaled Laplacian matrix. We work on the event defined by Proposition 3.25. Since the eigenvalues of $\underline{L}$ and M differ by at most $O\big(d^{-1/2}\big) = O\big((\log N)^{-1}\big)$, it suffices to prove the result for the eigenvalues and normalized degrees of M and $\mathbb{G}^\tau$ respectively. All quantities henceforth are related to those objects. We only prove our result for top eigenvalues as the proof for bottom eigenvalues is essentially the same.

Let $k \in [I]$, where we abbreviate $I := I^+$, $S_k := \operatorname{span}\big(\mathbf{w}_{x(i)} \colon\ i \in [k]\big)$, $k \in [I]$, where

$$\mathbf{w}_{x(i)} := \left(1 + \frac{1}{v_{x(i)}^2}\right)^{-1/2} \left[\mathbf{1}_{x(i)} - \frac{1}{\sqrt{d}v_{x(i)}}\mathbf{1}_{S_1(x(i))}\right], \qquad i \in [I]. \tag{3.37}$$

The rescaling in front is to insure that $\mathbf{w}$ are normalized. However, this does not impact the computations up to a factor $O\big((v_x)^{-2}\big) = O\big((\log N)^{-1}\big)$. For instance, writing $x = x_{(i)}$ and $D_x = |S_1(x)|$ for some $i \in [I]$, we find that

$$\begin{aligned} \langle \mathbf{w}_x , M\mathbf{w}_x\rangle &= \frac{1}{1+v_x^{-2}}\Big\langle \mathbf{1}_x - \frac{1}{\sqrt{d}v_x}\mathbf{1}_{S_1(x)} \,,\, (\underline{D}^\tau - \underline{A}^\tau)\Big(\mathbf{1}_x - \frac{1}{\sqrt{d}v_x}\mathbf{1}_{S_1(x)}\Big)\Big\rangle \\ &= \frac{1}{1+v_x^{-2}}\Big[v_x^\tau - \frac{1}{dv_x^2}\sum_{y\in S_1(x)} v_y^\tau - \Big\langle \mathbf{1}_x - \frac{\mathbf{1}_{S_1(x)}}{v_x\sqrt{d}} \,,\, \frac{\mathbf{1}_{S_1(x)}}{\sqrt{d}} + \frac{D_x}{dv_x}\mathbf{1}_x\Big\rangle\Big] \\ &= \frac{1}{1+v_x^{-2}}\Big[v_x + \frac{2}{v_x} + O\big(d^{-1/2} + (\log N)^{-1}\big)\Big] = v_x + \frac{1}{v_x} + \varepsilon_x, \end{aligned} \tag{3.38}$$

with $\varepsilon_x = O\big((\log N)^{-3/2} + d^{-1}\big)$. In the last equality we used the power series expansion of $(1+v_x^{-2})^{-1}$ to get the simplification.

We can derive a lower bound on $\lambda_{(k)}$ using the max-min principle

$$\lambda_{(k)} = \max_{\dim S = k} \min_{v\in S} \langle v\,, Mv\rangle \geq \min_{v\in S_k} \langle v\,, Mv\rangle.$$

By Proposition 5.15, the vertex sets $(B_1(x_{(i)}) : i \in [I])$ are disjoint and $S_1^\tau(x_{(i)}) \cap S_2^\tau(x_{(j)}) = \emptyset$, for $i,j\in[I]$, $i\neq j$. If $v = \sum_i a_i \mathbf{w}_{x(i)}$, with $\sum_i |a_i|^2 = 1$, $a_i\in\mathbb{R}$, we find

$$\langle v\,, Mv\rangle = \sum_{i=0}^{k} a_i^2 \left[v_{x(i)} + \frac{1}{v_{x(i)}} + \frac{1}{dv_{x(i)}^2} \sum_{y\in S_1(x(i))} v_y - 1 \right] \geq \min_{i\in[k]} \left(v_{x(i)} + \frac{1}{v_{x(i)}} + \varepsilon_{x(i)} \right).$$

Now since the function $f(x) := x + x^{-1}$ is strictly increasing for $x > 1$ we conclude that

$$\lambda_{(k)} \geq v_{x(k)} + \frac{1}{v_{x(k)}} + O\big((\log N)^{-3/2} + d^{-1}\big).$$

We now turn to the upper bound on $\lambda_{(k)}$. We will prove that

$$\lambda_{(1)} \leq \max_{i\in[I]} v_{x(i)} + \frac{1}{v_{x(i)}} + O\left(\frac{1}{\log N}\right). \tag{3.39}$$

The proof for $\lambda_{(k)}$, $2\leq k\leq I$ is similar, using the max-min principle

$$\lambda_{(k)} = \min_{S:\dim S = N-k+1} \max_{v\in S} \langle s\,, Ms\rangle \leq \max_{v\in U_k} \langle v\,, Mv\rangle \leq \max_{k\leq i\leq I} v_{x(i)} + \frac{1}{v_{x(i)}} + O\left(\frac{1}{\log N}\right),$$

with $U_k := S_k^\perp$. The details are left to the reader.
Let us consider the system of linear independent normalized vectors $W := \big(\mathbf{w}_{x(i)} : i\in[I]\big)$ (remember that in the graph $\mathbb{G}^\tau$ the balls $B_1(x_{(i)})$, $i\in[I]$ are disjoint). We complete W in a basis by first adding the vector

$$\mathbf{q} := \frac{1}{|A|^{1/2}} \mathbf{1}_A, \qquad A := [N] \setminus \bigcup_{i\in[I]} B_1(x_{(i)}),$$

and then by using Gram-Schmidt procedure to complete $W\cup\{\mathbf{q}\}$ with a collection of vectors that we call U. The vector $\mathbf{q}$ plays the role of the trivial eigenvector $\mathbf{e} := \frac{1}{\sqrt{N}}\mathbf{1}_{[N]}$ but its support excludes the support of the vectors of W (thus insuring orthogonality). However the difference is small since, using Lemma 5.2 we get

$$|A| \leq \sum_{i\in[I]} |B_1(x_{(i)})| \leq 2dN^{\tau/2} = O\big(N^{1/3}\big),$$

and thus

$$\langle \mathbf{e}, \mathbf{q}\rangle = \big(1 - O\big(N^{-2/3}\big)\big). \tag{3.40}$$

Note that since $W\cup\{\mathbf{q}\}\cup U$ is an orthonormal basis, we have

$$\sum_{u\in U} \big|u(x_{(i)})\big|^2 \leq \frac{1}{v_{x(i)}^2} \leq 1 - \big|\mathbf{w}_{x(i)}(x_{(i)})\big|^2 = \frac{1}{v_{x(i)}^2} \qquad i\in[I]. \tag{3.41}$$

Let $\mathbf{v}\in\mathbb{R}^N$ and

$$\mathbf{v} = \sum_{i\in[I]} a_i \mathbf{w}_{x(i)} + \sum_{\mathbf{u}\in U} b_u \mathbf{u} + c\mathbf{q}, \qquad a := \left(\sum_{i\in[I]} a_i^2\right)^{1/2}, \quad b := \left(\sum_{u\in U} b_u^2\right)^{1/2},$$

with $a_i, b_u, c \in \mathbb{R}$ and $a^2 + b^2 + c^2 = 1$.
Combining (3.38) with (i) of Proposition 5.16, we get

$$\langle \mathbf{w}_{x(i)}, M\mathbf{w}_{x(j)} \rangle = \delta_{ij}\Big(v_{x(i)} + \frac{1}{v_{x(i)}} + O\Big(\frac{1}{\log N}\Big)\Big), \quad i, j \in [I].$$

Since $\|\underline{A}^\tau - \underline{A}\| = O\big(d^{-1/2}\big)$, the vectors of U are orthogonal to $\mathbf{q}$ (and thus $\langle \mathbf{q}, \mathbf{u} \rangle = 0$, for any $\mathbf{u} \in U$) and $\mathbb{E}\underline{A}^\tau = \sqrt{d}\mathbf{e}\mathbf{e}^*$, we can use (3.40) and Proposition 3.25 to get

$$\begin{aligned} \big|\langle \mathbf{u}, \underline{A}^\tau \mathbf{u} \rangle\big| &= \big|\langle \mathbf{u}, (\underline{A}^\tau - \mathbb{E}\underline{A}^\tau)\mathbf{u} \rangle - \sqrt{d}\langle \mathbf{u}, \mathbf{e}\mathbf{e}^* \mathbf{u} \rangle\big| \\ &\leq \|\underline{A}^\tau - \mathbb{E}\underline{A}^\tau\| + \sqrt{d}|\langle \mathbf{u}, 1 - \mathbf{q} \rangle|^2 |\langle 1 - \mathbf{q}, \mathbf{e} \rangle|^2 \leq 3 + O\big(N^{-1/2}\big) \leq 4, \end{aligned}$$

where we used Cauchy-Schwartz inequality in the first inequality. holds with with very high probability and thus, using (3.41), we get

$$\langle \mathbf{u}, M\mathbf{u} \rangle = \langle \mathbf{u}, \underline{D}^\tau \mathbf{u} \rangle + 4 \leq \max_{y \notin \mathcal{U}(\tau)} v_y + \max_{i \in [I]} \frac{v_{x(i)}}{v_{x(i)}^2} + 4 \leq \sqrt{(2-\tau)\log N} + 5 \leq \sqrt{(2-\tau/2)\log N}.$$

Finally the cross-terms between U and W can be controlled by

$$\begin{aligned} \Big\langle \sum_i a_i \mathbf{w}_{x(i)}, M \sum_u b_u \mathbf{u} \Big\rangle &= \Big\langle \sum_i a_i \mathbf{w}_{x(i)}, \underline{D}^\tau \sum_u b_u \mathbf{u} \Big\rangle - \Big\langle \sum_i a_i \mathbf{w}_{x(i)}, \underline{A}^\tau \sum_u b_u \mathbf{u} \Big\rangle \\ &\leq \Big\langle \sum_i a_i \mathbf{1}_{x(i)} D, \sum_u b_u \mathbf{u}(x_{(i)}) \Big\rangle + \frac{1}{\sqrt{d}} \sum_i \frac{a_i}{v_{x(i)}} \sum_{y \in S_1(x(i))} v_y^\tau \sum_u b_u |\mathbf{u}(y)| + 3 \\ &\leq \Big(\sum_i a_i^2 v_{x(i)}^2\Big)^{1/2} \Big(\sum_u b_u^1 |\mathbf{u}(x_{(i)})|^2\Big)^{1/2} + Cba + 3 \\ &\leq \bigg(\frac{\max_i v_{x(i)}}{\min_i v_{x(i)}}\bigg)^{1/2} ab + Cba + 3 \leq 2Cba + 3. \end{aligned}$$

Here in the second inequality, we transferred the diagonal operator on the left side of the scalar product and used the precise structure of $\mathbf{w}$. We also used again the bound on the size of $\underline{A}^\tau - \mathbb{E}\underline{A}^\tau$. In the second inequality we used Cauchy-Schwartz as well as (iii) of Proposition 5.16 to control $\sum_{y \in S_1(x_{(i)})} v_y^\tau = Cd$ for $C \geq 1$ large enough. Finally, in the last step we used $\sqrt{(2-\tau)\log N} \leq \min_i v_{x(i)} \leq \max_i v_{x(i)} \leq \sqrt{2\log N}$.
Using (3.40), we see that $\langle \mathbf{q}, M\mathbf{q} \rangle = -\sqrt{d}\big(1 + O\big(N^{-1/3}\big)\big)$. The cross-terms with $\mathbf{q}$ are controlled by

$$\begin{aligned} \Big|\Big\langle \sum_i a_i \mathbf{w}_i, M\mathbf{q} \Big\rangle\Big| &= \frac{1}{d|A|^{1/2}} \sum_i \frac{a_i |S_2(x)|}{v_{x(i)}} = O\big(N^{-1/3}\big), \\ \Big|\Big\langle \sum_u b_u \mathbf{u}, M\mathbf{q} \Big\rangle\Big| &\leq \frac{1}{\sqrt{N}} \sum_u b_u \sum_{x \in [N]} v_x \mathbf{u}(x) + 3 \leq \sqrt{(2-\tau/2)\log N} + O(1), \end{aligned}$$

where for the second claim we proceed as before using (3.41) to bound the contribution of D applied to any vector of $\mathbf{u}$ and (3.40) to bound the contribution of $\underline{A}^\tau$. In the last inequality we used Cauchy-Schwartz to bound $N^{-1/2}\sum_u |b_u| \leq b$.
Writing $v_1 := \max_{i \in [I]} v_{x(i)}$, we find

$$\langle \mathbf{v}, M\mathbf{v} \rangle \leq a^2\Big(v_1 + \frac{1}{v_1} + \varepsilon\Big) + \big(b^2 + bc\big)\sqrt{(2-\tau)\log N} + 2Cba + 6 - c^2\frac{\sqrt{d}}{2} + O\big(N^{-1/3}\big),$$

for some $\varepsilon = O\big(d^{-1} + (\log N)^{-3/2}\big)$, $i \in [I]$. Now since $v_1 + \frac{1}{v_1} \geq \sqrt{2\log N} - \log\log N$ with high probability, we conclude that this inequality is maximal when a is maximal. This concludes the proof. □

3.4 Proof of block diagonal decompositions

In this section, we prove various block diagonal approximations of the matrix $\underline{L}$. While the results and computations might differ between the proofs of Propositions 3.15, 3.18 and 3.21, the pipeline is always the same. First, we analyse the top eigenvalue and eigenvector $(\lambda, \mathbf{w})$ of $\underline{L}$ restricted to the balls that surround our extremal vertices (meaning high- or low-degree vertices). Second, we obtain good bound on the radial decay of $\mathbf{w}$. Finally, we always have to account for the trivial eigenvector of the macroscopic connected component: indeed Proposition 3.25 does not give bounds on $\|\underline{A}\|$ but on $\|\underline{A} - \mathbb{E}\underline{A}\|$. All three elements are then combined to prove the block diagonal approximations.

Block diagonal decomposition in dense regimes

When $d \ll N^{1/6}$ we can use the graph constructed in Proposition 5.16 which has stronger separation properties (compare Proposition 5.16 (i) with Proposition 5.15 (i)). For $\tau \in (1-2\varepsilon, 1)$, let $\mathbb{G}^\tau$ be the graph constructed in Proposition 5.16. We define

$$M^\tau := \frac{1}{\sqrt{d}}\Big(D^\tau - A^\tau - d\Big), \tag{3.42}$$

where D^τ and A^τ are the degree matrix and the adjacency matrix of $\mathbb{G}^\tau$ respectively. In the rest of this section quantities related to $\mathbb{G}^\tau$ are indicated using a superscript τ.

Proposition 3.26. *Let d and $\tilde{\tau}$ be as in Proposition 3.15, $\tau \in (\tilde{\tau}, 1)$ and $M := M^{\tilde{\tau}}$ as defined in (3.42). There exists $\eta > 0$ such that the following holds with probability $1 - O\big(N^{-\eta}\big)$. For each $x \in \mathcal{U}(\tau)$ the matrix $M|_{B_2^{\tilde{\tau}}(x)}$ has a unique eigenvalue μ satisfying $|\mu| \geq \sqrt{(1+\tilde{\tau})\log N}$. The corresponding eigenvector is denoted $\mathbf{w}$ and satisfies*

$$\Big\|(M|_{B_2^{\tilde{\tau}}(x)} - (v_x^{\tilde{\tau}} + \frac{\alpha_x^{\tilde{\tau}}}{v_x^{\tilde{\tau}}}))\mathbf{w}\Big\| = O\Big(\frac{1}{\log N}\Big) \tag{3.43}$$

and

$$\|(M - \langle \mathbf{w}, M|_{B_2^{\tilde{\tau}}(x)}\mathbf{w}\rangle)\mathbf{w}\| = O\Big(\frac{1}{\log N}\Big). \tag{3.44}$$

Proof. Let us fix $\tilde{\tau} > 0$ and drop the superscript in this proof. We work on the event defined in Proposition 5.16 and Lemma 5.2 that hold with probability $1 - O\big(N^{-\eta}\big)$ for $\eta > 0$ small enough.
By Proposition 5.16 (i), there exists $c > 0$ such that $\min_{y \in B_2(x)} |v_x - v_y| > c\sqrt{\log N}$. Since $B_2(x)$ is a tree, by Lemma 5.10 and (5.2), we know that $\|A|_{B_2(x)}\| \leq 2\max_{y \in B_2(x)} \sqrt{D_y} \leq C\sqrt{d}$ for $C \geq 0$ large enough. We can therefore apply Proposition A.10 applied to the matrix $H = \frac{1}{\sqrt{d}} A^\tau|_{B_2^\tau(x)}$ and $V = \frac{1}{\sqrt{d}}\big(D^\tau|_{B_2^\tau(x)} - d\big)$ to conclude that $M|_{B_2(x)}$ has a unique eigenvalue larger than $\sqrt{(1+\tau)\log N} + C$. Moreover using (A.13) with $k = 2$, we find

$$\mu = v_x + \frac{1}{d}\sum_{y \in S_1(x)} \frac{1}{v_x - v_y} + O\left(\frac{1}{\min_{y \in B_2(x)} |v_x - v_y|^3}\right).$$

Proposition 5.16 (iii), we find

$$\frac{1}{d}\sum_{y \in S_1(x)} \frac{1}{v_x - v_y} = \frac{\alpha_x}{v_x} + C\frac{\alpha_x(\log N)^{1/4}}{v_x^2} + C\frac{\sqrt{\log N}}{dv_x^2} = \frac{\alpha_x}{v_x} + O\left(\frac{1}{(\log N)^{3/2}}\right).$$

Here, we introduced $C \geq 0$ as the constant coming from the first order Taylor expansion of the function $f(t) = \frac{1}{v_x - t}$. We deduce (3.43).

The second claim is proved analogously (3.53c), using a spectral gap argument. This concludes the proof. ∎

We denote $\mathbf{e} := N^{-1/2}\mathbf{1}_{[N]}$, $\mathcal{B}_\tau := \bigcup_{x\in\mathcal{U}(\tau)} B_2^\tau(x)$ and $\partial\mathcal{B}_\tau := \bigcup_{x\in\mathcal{U}(\tau)} S_2^\tau(x)$.

Proposition 3.27. *Under the assumptions of Proposition 3.15, there exists a normalized eigenvector* $\mathbf{q}$ *supported on the complement of* $\mathcal{B}_\tau$ *such that*

$$\|(M^\tau + \sqrt{d})\mathbf{q}\| = O(d^{-1/2}), \quad \|\mathbf{q} - \mathbf{e}\| = O(d^{-1}).$$

Proof. By Lemma 5.2 and Lemma 5.7, we know that

$$\mathbb{P}\left[\max_{x\in[N]} \frac{D_x}{d} \vee \frac{|B_2(x)|}{d^2} \geq C\right] = O\Big(N^{-1}\Big),$$

for $C \geq 0$ chose large enough. We conclude that

$$|\mathcal{B}_\tau| \leq CN^{1-\tau/2}d^2 \leq N^{1-\frac{1}{3}-\varepsilon-\frac{\tau}{2}} = O\Big(N^{-1/2}\Big),$$

holds with probability $1 - O\big(N^{-\eta}\big)$ for $\eta > 0$ small enough as soon as $\tau \geq 1/3$. Let $H := M^\tau|_{\mathcal{B}_\tau^c}$. By perturbation theory and Proposition 3.25, H has a unique eigenvalue outside of the interval $[-\sqrt{\tau \log N} - 2, \sqrt{\tau \log N} + 2]$. The vector $\mathbf{q} := |\mathcal{B}_\tau^c|^{-1/2}\mathbf{1}_{\mathcal{B}_\tau^c}$ satisfies $\|(H + \sqrt{d})\mathbf{q}\|^2 = O\big(N^{-1/4}\big)$, and so by Lemma A.6, we find

$$\|\mathbf{e} - \mathbf{q}\| = O\Big(\frac{1}{d}\Big).$$

This concludes the proof. □

Proof of Proposition 3.15. Let $\tau < \tilde{\tau} < 1$, $M := M^{\tilde{\tau}}$ as defined in (3.42) and $\mathcal{U}(\tau) := \mathcal{U}^+(\tau) \cup \mathcal{U}^-(\tau)$ Then by (iii) of Proposition 5.16

$$\|\underline{L} - M\| \leq \frac{2}{\sqrt{d}} \max_{x\in[N]} \big|D_x - D_x^\tau\big| = O(d^{-1/2}).$$

Let

$$\Pi := \sum_{x\in\mathcal{U}(\tau)} \mathbf{w}_x\mathbf{w}_x^* + \mathbf{q}\mathbf{q}^*, \quad \overline{\Pi} = 1 - \Pi$$

where $\mathbf{w}_x$ is the eigenvector of $M|_{B_2(x)}$ described in Proposition 3.26 and $\mathbf{q}$ is the approximate eigenvector constructed in Proposition 3.27. Suppose $\mathbf{u}$ is an eigenvector of $\overline{\Pi}M\overline{\Pi}$. Then $\max_{x\in\mathcal{U}(\tau)}|\mathbf{u}(x)| \leq \max_{y\notin\mathcal{U}(\tau)}|v_y|\frac{4}{v_x}$ since by orthonormality

$$|\mathbf{u}(x)|^2 \leq 1 - |\mathbf{w}(x)|^2 \leq 1 - |\mathbf{w}^x(x)|^2 + O\Big(\frac{1}{\log N}\Big) \leq 1 - \left(1 - \frac{1}{v_x}\right)^2 + O\Big(\frac{1}{\log N}\Big) \leq \frac{4}{v_x}.$$

Therefore, we find that with high probability

$$\|X\| \leq \max_{y\notin\mathcal{U}(\tau)}|v_y|\frac{2}{v_x} + \big\|\overline{\Pi}M|_{\mathbb{G}_\tau\setminus\mathcal{U}(\tau)}\overline{\Pi}\big\| \leq \max_{y\notin\mathcal{U}(\tau)}|v_y|\frac{2}{v_x} + \sqrt{\tau\log N} + 10.$$

Here we used the fact that $\frac{1}{\sqrt{d}}\|A^\tau - \mathbb{E}A^\tau\| \leq 3$ by Proposition 3.25. By (3.43), we find that $|\varepsilon_x| = O(1/\log N)$. Finally if $\mathbf{v} = \sum_{x\in\mathcal{U}(\tau)} a_x\mathbf{w}_x$ we find, using (3.44) and the fact that the balls $B_3(x)$, $x \in \mathcal{U}(\tau)$, are disjoint, we get

$$\big\|E_{\mathcal{U}(\tau)}\mathbf{v}\big\|^2 = \sum_{x\in\mathcal{U}(\tau)} a_x^2\|(M - \langle\mathbf{w}_x\,, M\mathbf{w}_x\rangle)\mathbf{w}_x\|.$$

This concludes the proof. □

Block diagonal decomposition for top eigenvalues

In this subsection, we prove Proposition 3.18. Let us recall the definition of κ and $\omega(\alpha)$, $\alpha > 0$, from (3.15) and (3.20) respectively.

Proposition 3.28 (Rigidity at $\mathcal{U}^+$)**.** *Let $\alpha \in (0, 1/12)$ and $r \geq 10$. There exists $K := K(\alpha, r) \geq 1$ and $c_* := c_*(\alpha, \kappa) > 0$ such that if $\varepsilon > 0$ and*

$$K \leq d \leq (\log N)^2,$$

then the following holds with probability $1 - O(\mathrm{e}^{-c_* \omega(\alpha)})$.

(i) For each $x \in \mathcal{W}^+(\kappa)$ the normalized eigenvector $\mathbf{w}$ of $\underline{L}|_{B_r(x)}$ corresponding to its largest eigenvalue satisfies

$$\big\|(\underline{L} - \Lambda_x)\mathbf{w}\big\| = O\bigg(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\bigg), \tag{3.45a}$$

where Λ_x is defined in (3.19).

(ii) For each $x \in \mathcal{U}^+(\varepsilon)$ the normalized eigenvector $\mathbf{w}$ of $\underline{L}|_{B_r(x)}$ corresponding to its largest eigenvalue satisfies

$$\big\|(\underline{L} - (v_x + \frac{\alpha_x}{v_x}))\mathbf{w}\big\| = O\Big(\frac{\log\log N}{\log N}\Big). \tag{3.45b}$$

(iii) For each $x \in \mathcal{U}^+(\varepsilon) \setminus \mathcal{W}^+(\kappa)$, if $1 \leq i \leq r$

$$\big\|\mathbf{w}|_{S_i(x)}\big\| = O\bigg(\Big(\frac{\log\log N}{(\log N \vee d)}\Big)^{\frac{i}{2}}\bigg), \quad \big\|(\underline{L} - \langle \mathbf{w}, \underline{L}\mathbf{w}\rangle)\mathbf{w}\big\| = O\bigg(\Big(\frac{\log\log N}{(\log N)}\Big)^{\frac{r}{2}-1}\bigg). \tag{3.45c}$$

Note that by Lemmas 5.2 and 5.18, $\max_{x \in \mathcal{W}^+(\kappa)} \big|\Lambda_x - v_x - \frac{\alpha_x}{v_x}\big| = O\big((\log N)^{-1}\big)$ on an event of good probability. Therefore (3.45a) implies (3.45b) for $x \in \mathcal{W}^+$.

Remark 3.29 (Non-isoradial nature of $\mathbf{w}$)**.** The eigenvector $\mathbf{w}$ constructed in (i) is not isoradial in the sense that exists $c > 0$ such that for each $x \in \mathcal{W}^+(\kappa)$,

$$\mathbb{P}\big[\min_{\gamma \in \mathbb{R}} \|\gamma \mathbf{s}_1 - \mathbf{w}(x)|_{S_1(x)}\| \geq c\big] \geq c. \tag{3.46}$$

The proof of Proposition 3.28 is deferred to the end of Section 3.5

Proposition 3.30 (Delocalisation estimate)**.** *Let $1 \leq d \leq \log N$, $\varepsilon > 0$ and $0 < \eta < \varepsilon/2$. If $\mathbf{u}$ is a vector orthogonal to $\operatorname{span}\big\{\mathbf{w}(x) : x \in \mathcal{U}^+(\varepsilon)\big\}$, where $\mathbf{w}_x$ is the eigenvector constructed in Proposition 3.28, then*

$$\frac{|\mathbf{u}(x)|}{\|\mathbf{u}\|} = O\Big(\frac{\omega}{\sqrt{\log N}}\Big), \quad x \in \mathcal{U}^+(\varepsilon), \tag{3.47}$$

holds with probability $1 - O(N^{-\eta})$.

Proof. Let us fix $\varepsilon > 0$, $x \in \mathcal{U}^+$ and $(\mu, \mathbf{w})$ the eigenvalue-eigenvector pair of $\underline{L}|_{B_r(x)}$ constructed in Proposition 3.28. Let $\mathbf{u} \in \operatorname{span}\{\mathbf{w}(x) : x \in \mathcal{U}^+\}^\perp$. Without loss of generality, we suppose that $\mathbf{u}$ is normalized. For $\eta > 0$, we work on the event defined in Proposition 5.19.
Since $B_{r+2}(x)$ is a tree for every $x \in \mathcal{U}^+$, Lemma 5.10 gives us that

$$\|\underline{A}|_{B_{r+1}(x)}\| \leq \frac{2}{\sqrt{d}}\Big(\max_{y \in B_r(x)} D_y\Big)^{1/2} \leq 2\sqrt{D_x/d} = 2\sqrt{\alpha_x}.$$

The orthogonality relation between $\mathbf{u}$ and $\mathbf{w}$ becomes

$$\begin{aligned} 0 =& \big|\langle \mathbf{u}, \mathbf{w}\rangle\big| = \Big|\Big\langle \mathbf{u}\,,\, \frac{\underline{L}|_{B_r(x)}}{\mu}\mathbf{w}\Big\rangle\Big| \geq \frac{1}{\mu}\Big[\Big|\langle \mathbf{u}, \underline{D}\mathbf{w}\rangle\Big| - \big|\langle \mathbf{u}, \underline{A}\mathbf{w}\rangle\big|\Big] \\ &\geq \frac{1}{\mu}\Big[|\langle \mathbf{u}, \underline{D}\mathbf{w}|_{B_0(x)}\rangle| - \langle \mathbf{u}, \underline{D}\mathbf{w}|_{B_r(x)\setminus\{x\}}\rangle| - \|\mathbf{u}\|\|\underline{A}|_{B_{r+1}(x)}\|\|\mathbf{w}\|\Big]. \end{aligned}$$

Using (3.45c) and the bound on $\|\underline{A}|_{B_{r+1}(x)}\|$, we find

$$\begin{aligned} 0 &\geq \frac{1}{\mu}\Big[|\mathbf{u}(x)v_x||\mathbf{w}(x)| - \|\mathbf{w}|_{B_r(x)\setminus\{x\}}\|\|\underline{D}|_{B_r(x)}\| - 2\sqrt{\alpha_x}\Big] \\ &\geq \frac{|\mathbf{u}(x)v_x|}{\mu}\Big[1 - C\Big(\frac{\omega(\alpha)}{\log N}\Big)^{1/2}\Big] - \frac{C}{\mu}\left(\frac{\omega(\alpha)}{\log N}\right)^{1/2} \max_{y\in B_r(x)\setminus\{x\}}|v_y| - \frac{2\sqrt{\alpha_x}}{\mu}. \end{aligned}$$

with C chosen as in (3.45c).
Using the fact that $\mu \geq v_x \geq \psi_+$ and $\max_{y\in B_r(x),\, y\neq x}|v_y| \leq \psi_+$, for ψ_+ defined in (5.22), we find

$$\frac{1}{2}|\mathbf{u}(x)| \leq \frac{1}{\mu}\left[\frac{\omega(\alpha)\psi_+}{\sqrt{\log N}} + 2\sqrt{\alpha_x}\right] \leq 2\frac{\omega(\alpha)}{\sqrt{\log N}}.$$

This concludes the proof. □

Proof of Proposition 3.18. Let us fix $\varepsilon, \alpha > 0$. We work on the event defined by Propositions 3.30, 5.19, for $r \geq 10$, and by Lemma 5.18.
Points (i) and (ii) follow from (i) and (ii) of Proposition 3.28 respectively. By (3.19), Lemma 5.18 and (i) from Proposition 5.21,

$$\sup_{x\in\mathcal{W}^+(\kappa)} \big|\Lambda_x - v_x + \frac{\alpha_x}{v_x}\big| = O\Big(\frac{\omega}{\log N}\Big),$$

holds with probability $1 - O\big(\mathrm{e}^{-c\omega(\alpha)}\big)$ for some constant $c > 0$ small enough.
Let $(\mu_x, \mathbf{w}_x)$, $x \in \mathcal{U}^+(\varepsilon)$, be the eigenvector-eigenvalue pair constructed for $\underline{L}|_{B_r(x)}$ in Proposition 3.28. Let $v := \sum_{x\in\mathcal{W}^+} c_x\mathbf{w}_x$, with $\sum_x c_x^2 = 1$. Since the balls $\big(B_{r+10}(x)\colon\ x \in \mathcal{U}^+(\varepsilon)\big)$ are disjoint, we have

$$\|E_{\mathcal{W}}v\|^2 \leq 2\sum_{x\in\mathcal{W}^+}|c_x|^2\|(\underline{L} - \langle \mathbf{w}_x\,, \underline{L}\mathbf{w}_x\rangle)\mathbf{w}_x\|^2 = O\big((\log N)^{-4}\big),$$

where we used (3.45c) and $r \geq 10$ in the last step. The bound over $\|E_{\mathcal{U}}\|$ is proved in the same way.
It now remains to show (3.23). Let us define the projections

$$\Pi := \sum_{x\in\mathcal{U}^+(\varepsilon)} \mathbf{w}_x\mathbf{w}_x^*, \quad \overline{\Pi} := 1 - \Pi. \tag{3.48}$$

By Proposition 3.30 and Lemma 5.2, if $\mathbf{v}$ is a vector of norm 1, we have that

$$\Big\langle \overline{\Pi}\mathbf{v}\,,\, D|_{\mathcal{U}^+(\varepsilon)}\overline{\Pi}\mathbf{v}\Big\rangle \leq \max_{x\in[N]}\big|\overline{\Pi}v(x)\big|\max_{y\in\mathcal{U}^+}|v_y| \leq C\frac{\omega(\alpha)^2\sqrt{d}(\mathfrak{u}_+ - 1)}{\log N},$$

holds with high probability for some constant $C \geq 0$ large enough. We find

$$\Big\langle \overline{\Pi}\mathbf{v}\,,\, \underline{L}\overline{\Pi}\mathbf{v}\Big\rangle \leq \Big\langle \overline{\Pi}\mathbf{v}\,,\, D|_{\mathcal{U}^+(\varepsilon)}\overline{\Pi}\mathbf{v}\Big\rangle + \Big\langle \overline{\Pi}\mathbf{v}\,,\, D|_{(\mathcal{U}^+(\varepsilon))^c}\overline{\Pi}\mathbf{v}\Big\rangle + \|\underline{A} - \mathbb{E}\underline{A}\| - \Big\langle \overline{\Pi}\mathbf{v}\,,\, \mathbb{E}\underline{A}\overline{\Pi}\mathbf{v}\Big\rangle.$$

Using Proposition 3.25 and the fact that $\mathbb{E}\underline{A} = d\mathbf{e}\mathbf{e}^* - \frac{d}{N}\,\mathrm{Id}_N$, we find

$$\begin{aligned} \big\langle \overline{\Pi}\mathbf{v}, \underline{L}\overline{\Pi}\mathbf{v}\big\rangle &\leq \max_{y\notin\mathcal{U}(\varepsilon)}|v_y| + C\frac{\omega(\alpha)^2\sqrt{d}(\mathfrak{u}_+ - 1)}{\log N} + C\mathfrak{u}_+ - d\sum_{x\in[N]}\big|\overline{\Pi}(v)\big|^2 + \frac{d}{N}\|\mathbf{v}\| \\ &\leq \mathfrak{u}^+(\varepsilon) + C\big(\omega(\alpha)^2 + 2\mathfrak{u}_+\big) \leq \mathfrak{u}^+(\varepsilon)(1+\varepsilon), \end{aligned}$$

holds with high probability. This proves (3.23) and concludes the proof. □

Block diagonal decomposition for bottom eigenvalues

Proposition 3.31 (Rigidity at $\mathcal{U}^-$)**.** *Let $K > 0$ and $\alpha, \varepsilon \in (0, 1/10)$ and κ and κ as in (3.29). If*

$$\frac{1}{2}\log N + (\log N)^{1/K} \leq d \leq (\log N)^2, \quad 0 < c_* < 2\alpha \wedge \frac{\log d}{K \log\log N},$$

then the following holds for $r \geq \max(10, 10/c_*)$, with probability $1 - O(\mathrm{e}^{-d^{c_*}})$.

(i) For each $x \in \mathcal{U}^-(\varepsilon)$, the normalized eigenvector $\mathbf{w}(x)$ of $\underline{L}|_{B_r(x)}$ corresponding to its smallest eigenvalue satisfies

$$\left\| \Big(\underline{L} - \Big(v_x + \frac{\alpha_x}{v_x}\Big)\Big)\mathbf{w}(x) \right\| = O\Big(\frac{1}{\sqrt{\log N}}\Big). \tag{3.49a}$$

(ii) For each $x \in \mathcal{W}^-(\kappa)$, the normalized eigenvector $\mathbf{w}(x)$ of $\underline{L}|_{B_r(x)}$ corresponding to its smallest eigenvalue satisfies

$$\|(\underline{L} - \Lambda)\mathbf{w}(x)\| = O\big(\varepsilon_x\big). \tag{3.49b}$$

where $\Lambda \in \big\{\Lambda_x, \Lambda_x^{\mathcal{L}}, \tilde{\Lambda}_x^{\mathcal{L}}\big\}$ according to the different cases listed Proposition 3.21 (i)-(iii) and ε_x defined accordingly therein.

(iii) For $d \geq \log N - (\log\log N)^2$ we have

$$\max_{x \in \mathcal{W}^-(\kappa)\cap\mathcal{L}} |\Lambda_x - \Lambda_x^{\mathcal{L}}| \vee |\Lambda_x - \tilde{\Lambda}_x^{\mathcal{L}}| = O\left(\frac{\omega(\alpha)}{(\log N)^2}\right). \tag{3.49c}$$

(iv) For each $x \in \mathcal{U}^-(\varepsilon)$,

$$\begin{aligned} \big\|\mathbf{w}(x)|_{S_i(x)}\big\| &= O\left(\frac{1}{(\log N)^{c_*}} \frac{\omega^i}{(\log N)^{\frac{i-1}{2}}}\right), \quad 1 \leq i \leq r, \\ \big\|(\underline{L} - \langle \mathbf{w}(x), \underline{L}\mathbf{w}(x)\rangle)\mathbf{w}(x)\big\| &= O\left(\frac{\omega^r}{(\log N)^{\frac{r-2}{2}}}\right). \end{aligned} \tag{3.49d}$$

The proof of Proposition 3.31 is deferred to the end of Section 3.6. Let us define

$$\mathcal{B}_r^-(\varepsilon) := \bigcup_{x \in \mathcal{U}^-(\varepsilon)} B_r(x), \quad r \in \mathbb{N}^*. \tag{3.50}$$

Lemma 3.32 (Connected components and trivial eigenvector)**.** *Let $K > 0$, $r \in \mathbb{N}^*$ and d be as in (3.28). If $r = O(1)$, there exists $c_* > 0$ such that with probability at least $1 - O(\mathrm{e}^{-d^{c_*}})$, the following hold.*

1. *the graph $\mathbb{G}$ consists of one connected component with more than $N/2$ vertices, denoted $\mathbb{G}_{cc}$ and at most $O\big(\mathrm{e}^{-d}N\big)$ isolated vertices.*
2. *There exists a vector $\mathbf{q}$ supported on $\mathbb{G}_{cc} \setminus \mathcal{B}_r^-(\varepsilon)$ such that*

$$\|(\underline{L} + \sqrt{d})\mathbf{q}\| = O\Big(N^{-\frac{1}{4}+\eta}\Big), \quad \eta > 0.$$

Proof. The first point follows from Lemmas 5.5. Recalling (3.12), Lemma B.3, we use Chebyshev's inequality to prove that $|\mathcal{U}^-(\varepsilon)| \leq N^{\frac{1}{2}+\eta/2}$ with probability $1 - O\big(N^{-\eta}\big)$, $\eta > 0$. In the rest of the proof, we work on the intersection of the event $\{|\mathcal{U}^-(\varepsilon)| \leq N^{\frac{1}{2}+\eta/2}\}$ and the one defined in Proposition 5.23, that we call Ξ. In particular, we know that on Ξ,

$$\big|\mathcal{B}_r^-(\varepsilon)\big| \leq CN^{\frac{1}{2}+\eta/2}d^{r+6} = O\big(N^{\frac{1}{2}+\eta}\big), \quad \big|\mathbb{G}_{cc} \setminus \mathcal{B}_r^-(\varepsilon)\big| \geq N^{1-\eta}.$$

The same estimates are true if we replace $\mathcal{B}_r^-$ by $\mathcal{B}_{r+3}^-$. Let

$$\mathbf{q} := |\mathcal{B}_r^c(\varepsilon)|^{-1/2}\mathbf{1}_{\mathcal{B}_r^c(\varepsilon)}.$$

We have

$$\left(\underline{L} + \sqrt{d}\right)\mathbf{q} = \frac{1}{\sqrt{|\mathcal{B}_{r+1}^c|}}\sum_{x\in\mathcal{B}_r^c}(v_x - \tilde{v}_x)\mathbf{1}_x + \frac{1}{\sqrt{|\mathcal{B}_{r+1}^c|}}\sum_{x\in\mathcal{B}_r^c\cap\mathcal{B}_{r+2}}\mathbf{1}_x,$$

where $\tilde{v}_x := \frac{1}{\sqrt{d}}\left(\sum_y (A|_{\mathcal{B}_r^c})_{xy} - d\right)$. Since on Ξ the balls $\left(B_{r+3}(x)\colon x\in\mathcal{U}^-(\varepsilon)\right)$ are disjoint trees, $0 \leq v_x - \tilde{v}_x \leq d^{-1/2}$ holds for $x\in\mathcal{B}_r^c$ and $v_x - \tilde{v}_x = 0$ for $[N]\setminus\mathcal{B}_{r+3}^c$. Therefore only vertices in $\mathcal{B}_{r+1}^c\cap\mathcal{B}_{r+2}$ contribute to the right-hand side of the above equation. Moreover, on Ξ,

$$\frac{|\mathcal{B}_{r+1}^c\cap\mathcal{B}_{r+2}|}{|\mathcal{B}_{r+2}^c|} \leq \frac{|\mathcal{U}^-(\varepsilon)|\max_{x\in\mathcal{U}^-(\varepsilon)}|B_r(x)|}{N^{1-\eta}} = \frac{N^{\frac{1}{2}+\eta}}{N^{1-\eta}} = O\left(N^{-\frac{1}{2}+2\eta}\right).$$

We concludes that $\|(\underline{L}+\sqrt{d})\mathbf{q}\|^2 = O\left(N^{-\frac{1}{2}-2\eta}\right)$ and the claim follows. □

Proposition 3.33 (Delocalisation estimate)**.** *Let $\varepsilon > 0$, $K > 0$,*

$$\frac{1}{2}\log N + (\log N)^{1/K} \leq d \leq \log N,$$

and $\mathbf{w}(x)$, $x\in\mathcal{U}^+(\varepsilon)$, be the vector constructed in Proposition 3.28. There exists $c_* > 0$ such that with probability at least $1 - O(\mathrm{e}^{-d^{c_*}})$, there is $C > 0$ such if $\mathbf{u}$ is a normalized vector orthogonal to $\operatorname{span}\{\mathbf{w}(x) : x\in\mathcal{U}^+(\varepsilon)\}$, then

$$|\mathbf{u}(x)| = O\left(\frac{\omega}{(\log N)^{c_*/2}}\right). \tag{3.51}$$

Proof. The proof is the same as the one of Proposition 3.30 using (3.49d) instead of (3.45c). □

Proof of Proposition 3.21. Almost all claims from (i)-(iv) and (vi) follow from Proposition 3.31 (i)-(iii). The only claim that is not proved is the bound on $(\beta_{z_x} - 1)$ in (3.31), which follows from Lemma 5.24. There only remains to prove (v). We begin by restricting ourselves to $\mathbb{G}_{cc}$ using Lemma 3.32. This proves the statement regarding $\operatorname{Spec} Y$. Let $H := \underline{L}|_{\mathbb{G}_{cc}}$. The statement about ν comes from Lemma 3.32. Estimates on $\|E_\sharp\|$, $\sharp\in\{\nu,\mathcal{U},\mathcal{W}\}$ are proved as in Proposition (iii) using (3.49d).
By Proposition 3.25, for $d \geq \frac{1}{2}\log N$, we have $\|\underline{A} - \mathbb{E}\underline{A}\| \leq 3$ with very high probability. Let us define

$$\Pi := \mathbf{q}\mathbf{q}^* + \sum_{x\in\mathcal{U}^-(\varepsilon)}\mathbf{w}(x)\mathbf{w}(x)^*, \quad \overline{\Pi} := 1 - \Pi,$$

where $\mathbf{w}(x)$ are the vector constructed in Proposition 3.31. Let $\mathbf{u}$ be a normalized vector supported on $\mathbb{G}_{cc}$. We find that

$$\begin{aligned}\left\langle\overline{\Pi}\mathbf{u}\,,\, H\overline{\Pi}\mathbf{u}\right\rangle &\geq \sum_{y\notin\mathcal{U}^-(\varepsilon)}\left|\overline{\Pi}\mathbf{u}(y)\right|^2 v_y + \sum_{y\in\mathcal{U}^-(\varepsilon)}\left|\overline{\Pi}\mathbf{u}(y)\right|^2 v_y - \|A - \mathbb{E}A\| - \langle 1-\mathbf{q}, \mathbb{E}A(1-\mathbf{q})\rangle \\ &\geq -\sqrt{d} + \mathfrak{u}^-(\varepsilon)\left[1 - \frac{C\omega}{(\log N)^{c_*}}\right] - 3 - \sqrt{d}\frac{|\mathcal{B}_r|}{N}, \geq -\sqrt{d}(\mathfrak{u}_- - 1) + \frac{1}{2}(\mathfrak{u}^-(\varepsilon) - \sqrt{d}\mathfrak{u}_-),\end{aligned}$$

where we used the fact that $\mathbb{E}\underline{A} = \sqrt{d}\mathbf{e}\mathbf{e}^*$ and the bound on $|\mathcal{B}_r|$ derived in the proof of Lemma 3.32. In the regime where $\mathfrak{u}^-(\varepsilon) \gg 1$ we can immediately conclude (3.32). If $\mathfrak{u}^-(\varepsilon) = \frac{1}{\varepsilon}$ we simply choose $\varepsilon > 0$ small enough so that $\frac{1}{\varepsilon} - 4 \geq \frac{3\kappa}{2}$, for instance $\varepsilon < 1/10$. This concludes the proof. □

3.5 Spectrum around large vertices

In this section, we prove Proposition 3.28. As a general rule, the error bounds relating to $\mathcal{W}^+(\kappa)$ are written using $\mathfrak{u}_+$-dependent parameters. This is done to obtain the degree of precision which we need in Section 3.7 to prove convergence to Poisson point processes. Bounds relating to $\mathcal{U}^+(\varepsilon)$ are not stated with this level of precision. Let us recall the definition of κ and $\omega(\alpha)$, $\alpha > 0$, from (3.15) and (3.20) respectively.

Let us introduce the spectral gap for this section

$$\psi_+ := \sqrt{\log N} \vee \frac{\log N}{\sqrt{d}\log(\log N/d)}. \tag{3.52}$$

Observe that there exists $C > 0$ such that $\frac{1}{C}v(\mathfrak{u}_+) \le \psi_+ \le Cv(\mathfrak{u}_+)$ by construction. However, these two parameters have different meanings and so we keep them distinct.

Proposition 3.34 (Spectrum of $\underline{L}|_{B_r(x)}$, $x \in \mathcal{U}^+$)**.** *Let $\alpha, \varepsilon \in (0, 1/12)$ and $r \ge 10$. There exists $K := K(\alpha, r) \ge 1$ and $c_* := c_*(\alpha, \kappa) > 0$ such that if*

$$K \le d \le (\log N)^2,$$

then the following holds with probability $1 - O(\mathrm{e}^{-c_*\omega(\alpha)})$.

1. *For each $x \in \mathcal{U}^+(\varepsilon)$, the largest eigenvalue of $\underline{L}|_{B_r(x)}$ written μ satisfy*

$$\mu = v_x + \frac{\alpha_x}{v_x} + O\biggl(\frac{\log\log N}{\log N}\biggr). \tag{3.53a}$$

2. *For each $x \in \mathcal{W}^+(\kappa)$ then*

$$\mu = \Lambda_x + O\biggl(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\biggr), \tag{3.53b}$$

where Λ_x is defined in (3.19).

Moreover if $\mathbf{w}$ is the eigenvector corresponding to μ we have

$$\|\mathbf{w}|_{B_r(x)\setminus B_i(x)}\| = O\biggl(\biggl(\frac{\log\log N}{\log N}\biggr)^{\frac{i}{2}}\biggr), \quad i \in [r]. \tag{3.53c}$$

Proof. Let us fix $r \ge 10$, $\varepsilon > 0$ and $\alpha \in (0, 1/12)$. We work on the event defined by Lemma 5.12 and Propositions 5.19 and 5.21. Then there exists $c_* > 0$, depending on κ and α and ε, such that the probability of this event is $1 - O\bigl(\mathrm{e}^{c_*\omega(\alpha)}\bigr)$ and

$$\max_{y \in B_r(x),\, y \ne x} |v_x - v_y| \ge c_*\psi_+, \quad x \in \mathcal{U}^+(\varepsilon), \tag{3.54}$$

where ψ_+ is defined in (3.52). Let us denote this event $\Xi := \Xi(c_*, \alpha)$.

We fix $x \in \mathcal{U}^+(\varepsilon)$ and write $V := \frac{D-d}{\sqrt{d}}|_{B_r(x)}$ and abbreviate $H := fA|_{B_r(x)}$ with $f = d^{-1/2}$. By (3.54) and Proposition 3.25, there exists $c > 0$ depending on c_* such that the hypotheses of Proposition A.10 are satisfied with $\psi = c_*\psi_+$ and $G = (B_r(x), A|_{B_r(x)}, x)$. Equation (A.13) with $k = 3$ becomes

$$\mu = v_x + E_2(0) + E_4(0) + E_4(1) + O\biggl(\frac{\|H\|^6(|v_x| + \psi)}{\psi^6}\biggr), \tag{3.55}$$

where $E_l(e)$, $l \ge 1$ and $e \in [l-1]$ is defined in (A.14).

Observe that, again by Lemma 5.10, since $B_r(x)$ is a tree $\|H\|^8 \le 2\alpha_x^4 \le 2\mathfrak{u}_+$. By (3.52) and the bound $v_x \le C\psi_+$ for $C \ge 0$ large enough, we see that

$$\frac{\|H\|^6(|v_x| + \psi_+)}{\psi_+^6} \le \frac{C\mathfrak{u}_+^3}{\psi_+^5} = O\biggl(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\biggr),$$

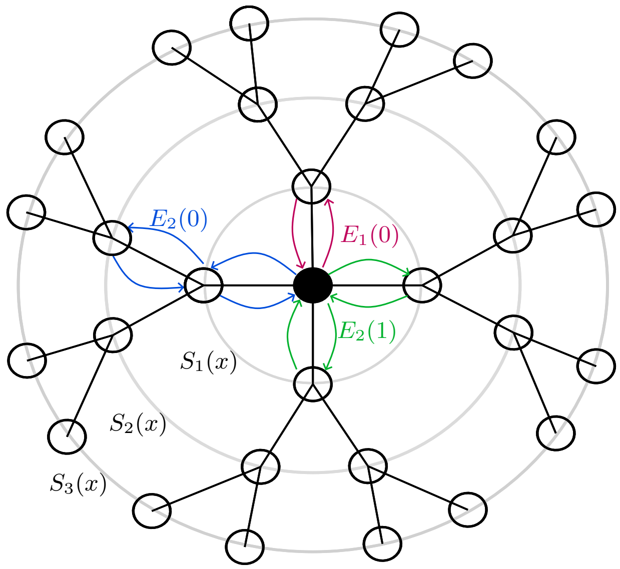


Figure 3.3: Illustration of the perturbation argument. The different terms that contribute to (3.55) appear as cycles that start and end at the root vertex (black vertex in the middle). The illustration makes clear what each index in $E_l(e)$ means: the subscript l stands for half the length of the cycle (in a tree there are no odd-length cycles) while de number e stands for the number of excess visits to the central vertex (i.e. the number of total visits minus 2).

where we used $\alpha < \frac{1}{12}$ in the last step.

The leading contribution to the correction away from v_x comes from the first term

$$E_2(0) := \frac{1}{d} \sum_{y \in S_1(x)} \frac{1}{v_x - v_y}.$$

Using the second order approximation $\frac{1}{v_x - t} = \frac{1}{v_x} + \frac{t}{v_x^2} + \frac{t^2}{v_x^3} + O\big(\frac{t^3}{v_x^4}\big)$. We find

$$\begin{aligned} E_2(0) &= \frac{1}{d} \sum_{y \in S_1(x)} \left[\frac{1}{v_x} + \frac{v_y}{v_x^2} + \frac{v_y^2}{v_x^3} + O\bigg(\frac{v_y^3}{|v_x - v_y|^4}\bigg)\right] \\ &= \frac{\alpha_x}{v_x} + \frac{\sqrt{d}\alpha_x(\beta_x - 1)}{v_x^2} + \frac{\alpha_x}{v_x^2\sqrt{d}} + \frac{1}{d} \sum_{y \in S_1(x)} \frac{v_y^2}{v_x^3} + O\bigg(\frac{v_y^3}{|v_x - v_y|^4}\bigg). \end{aligned} \tag{3.56}$$

Here the constant in the big-O is a universal one that depends only on the Taylor expansion of f. We also used the identity (valid since $B_r(x)$ is a tree)

$$\frac{1}{d} \sum_{y \in S_1(x)} v_y = \frac{|S_2(x)| + |S_1(x)| - d|S_1(x)|}{d^{3/2}} = \sqrt{d}\alpha_x(\beta_x - 1) + \frac{\alpha_x}{\sqrt{d}}. \tag{3.57}$$

Next we use the estimates (iv) of Proposition (iv), to see that on Ξ

$$\frac{|S_i(x)|}{d^i} = \mathfrak{u}_+\bigg(1 + O\Big(\frac{\log N}{dD_x}\Big)^{1/2}\bigg) = O\bigg(\mathfrak{u}_+\Big(1 \vee \frac{\log\log N}{d}\Big)\bigg), \quad x \in \mathcal{U}^+(\varepsilon),\ 1 \le i \le r. \tag{3.58}$$

By Proposition 5.21 we also have, for each $x \in \mathcal{W}^+(\kappa)$, on Ξ,

$$\begin{aligned} \frac{1}{d^i} \sum_{y \in S_i(x)} v_y^2 &= \alpha_x \left(1 + O\Big(\frac{\omega(\alpha)^2}{\sqrt{D_x}}\Big)\right) + (\log\log N)^6 \mathbf{1}_{d^{2\alpha} \le \log\log N}, \\ \frac{1}{d^i} \sum_{y \in S_i(x)} |v_y|^n &= \alpha_x d^{2\alpha} \left(1 + O\Big(\frac{\log N}{dD_x}\Big)^{1/2}\right) + C\kappa \frac{(\psi_+)^{n/2} \log\log N}{d^{i+2\alpha}}, \end{aligned} \tag{3.59}$$

for $1 \le i \le r$, $n \in \mathbb{N}^*$. By Proposition 5.19 we have, for each $x \in \mathcal{U}^+(\varepsilon)$, on Ξ, the bounds in the above equation become $\alpha_x (\log\log N)^2$ and $\alpha_x (\log N)^{\alpha n} + (\log N)^{1-2\alpha} d^{-i}$ respectively. We thus find, using $\kappa \log\log N \le d^{1+2\alpha}\sqrt{\mathfrak{u}_+}$, $(\log\log N)^6 \le \sqrt{v(\mathfrak{u}_+)}$ and $\alpha < 1/2$,

$$E_2(0) = \frac{\alpha_x}{v_x} + \frac{\alpha_x}{\sqrt{d} v_x^2} + \frac{\sqrt{d}\alpha_x}{v_x^2}(\beta_x - 1) + \begin{cases} \frac{\alpha_x}{v_x^3} + \varepsilon_x, \text{ with } \varepsilon_x \le \text{ RHS of (3.53b)}, & x \in \mathcal{W}^+(\kappa) \\ \varepsilon_x, \text{ with } \varepsilon_x \le \text{ RHS of (3.53a)} & x \in \mathcal{U}^+(\varepsilon). \end{cases} \tag{3.60}$$

The terms $E_4(e)$, $e = 0, 1$, correspond to cycles of length 4. $E_4(0)$ are those cycles that visit x only twice (at the beginning and the end), and thus they are in bijection with the points of $S_2(x)$. We have

$$E_4(0) := \frac{1}{d^2} \sum_{z \in S_2(x)} \frac{1}{(v_x - v_{(z-)})^2 (v_x - v_z)},$$

where we introduced the notation $(z-) := S_1(x) \cap S_1(z)$, for $z \in B_r(x) \setminus B_1(x)$ (see Picture).
The cycles that contribute to $E_4(1)$ visits x three times and are thus in bijection with $S_1(x) \times S_1(x)$. Counting the multiplicity 2 induced by reordering, we find

$$E_4(1) := \frac{-1}{2d^2} \sum_{(y,z) \in S_1(x) \times S_1(x)} \frac{1}{(v_x - v_y)(v_x - v_z)} \left[\frac{1}{v_x - v_y} + \frac{1}{v_x - v_z}\right] = \frac{-1}{d^2} \sum_{y,z \in S_1(x)} \frac{1}{(v_x - v_y)^2 (v_x - v_z)}$$

We can now conclude (3.53a) from (3.54) and (3.58), and the fact that for each $x \in \mathcal{U}^+(\varepsilon)$,

$$E_4(0) + E_4(1) \le \frac{C\mathfrak{u}_+}{v(\mathfrak{u}_+)^3} \left(1 + \frac{\log\log N}{d}\right) + \frac{\mathfrak{u}_+^2}{v(\mathfrak{u}_+)^3} = O\left(\frac{\log\log N}{\log N}\right).$$

In the rest of the argument, we consider only $x \in \mathcal{W}^+(\kappa)$. Proceeding as for $E_2(0)$ and using (3.59), we find

$$\begin{aligned} E_4(1) &= -\frac{1}{d^2} \sum_{y,z \in S_1(x)} \frac{1}{(v_x - v_y)^2 (v_x - v_z)} \\ &= -\frac{1}{d^2} \sum_{y,z \in S_1(x)} \left(\frac{1}{v_x^2} + \frac{2v_y}{v_x^3} + O\Big(\frac{|v_y|^2}{\psi_+^4}\Big)\right)\left(\frac{1}{v_x} + \frac{v_z}{v_x^2} + O\Big(\frac{|v_z|^2}{\psi_+^3}\Big)\right) \\ &= -\frac{\alpha_x^2}{v_x^3} - \frac{3\alpha_x}{v_x^4}\frac{1}{d} \sum_{y \in S_1(x)} v_y + O\left(\frac{\mathfrak{u}_+^2 d^{4\alpha}}{\psi_+^5}\right) = -\frac{\alpha_x^2}{v_x^3} + O\Big(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d} v(\mathfrak{u}_+)^2 \omega(\alpha)^2}\Big), \end{aligned} \tag{3.61}$$

where we used $\mathfrak{u}_+^2 \omega(\alpha)^2 \le v(\mathfrak{u}_+)^2$ in the last step.

The last term to control si $E_4(0)$. Using the linearization $\frac{1}{v_x - v_z} = \frac{1}{v_x} + O\Big(\frac{|v_z|}{\psi_+^2}\Big)$ and $\min_{y \in S_1(x)} |v_x - v_y| \ge c_* \psi_+$, we find

$$\begin{aligned} E_4(0) &= \frac{1}{d^2} \sum_{z \in S_2(x)} \frac{1}{(v_x - v_{(z-)})^2 (v_x - v_z)} = \frac{1}{d^2} \sum_{z \in S_2(x)} \frac{1}{v_x (v_x - v_{(z-)})^2} + O\Big(\frac{|v_z|}{\psi_+^4}\Big) \\ &= \frac{1}{d^2 v_x} \sum_{y \in S_1(x)} \frac{N_y}{(v_x - v_y)^2} + O\Big(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d} v(\mathfrak{u}_+)^2 \omega(\alpha)^2}\Big), \end{aligned}$$

where we introduced $N_y := |S_1(y) \cap S_2(x)| = (\sqrt{d}v_y + d - 1)$ and estimated the error term by considering the cases $d \geq (\log N)^{1/10}$ and $d \leq (\log N)^{1/10}$ separately.
Let $f(t) = \frac{\sqrt{d}t+d-1}{(v_x-t)^2}$, for $t \in \mathbb{R}$. Then a Taylor development to the second order gives

$$f(t) = \frac{d-1}{v_x^2} + t\Big[\frac{\sqrt{d}}{v_x^2} + \frac{2(d-1)}{v_x^3}\Big] + O\bigg(\frac{\sqrt{d}}{(v_x-\zeta)^3} + \frac{\sqrt{d}\zeta + d}{(v_x-\zeta)^4}\bigg), \quad 0 < \zeta < t.$$

Using (3.57), $v(\mathfrak{u}_+)^2 \geq \sqrt{\mathfrak{u}_+}\omega(\alpha)^2$ and Lemma 5.24, we deduce

$$\begin{aligned}\frac{1}{d^2 v_x}\sum_{y\in S_1(x)} \frac{N_y}{(v_x-v_y)^2} &= \frac{\alpha_x}{v_x^3} - \frac{\alpha_x}{dv_x^3} + \frac{1}{\sqrt{d}v_x^3}\frac{1}{d}\sum_{y\in S_1(x)} v_y + O\bigg(\frac{\alpha_x}{\sqrt{d}v_x\psi_+^3} + \frac{\alpha_x}{v_x\psi_+^4}\bigg)\\ &= \frac{\alpha_x}{v_x^3} + O\bigg(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\bigg).\end{aligned}$$

We conclude that

$$\begin{aligned}\mu &= v_x + \frac{\alpha_x}{v_x}\bigg[1 + \frac{1}{\sqrt{d}v_x} + \frac{2-\alpha_x}{v_x^2}\bigg] + \frac{\sqrt{d}\alpha_x}{v_x^2}(\beta_x - 1) + O\bigg(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\bigg)\\ &= v_x + \frac{\alpha_x}{v_x}\bigg[1 + \frac{1}{v_x^2}\bigg] + \frac{\sqrt{d}\alpha_x}{v_x^2}(\beta_x - 1) + O\bigg(\frac{\sqrt{\mathfrak{u}_+}}{\sqrt{d}v(\mathfrak{u}_+)^2\omega(\alpha)^2}\bigg)\end{aligned}$$

which proves (3.53b). (We used the identity $v_x = \sqrt{d}(\alpha_x - 1)$ in the last step.)

We now turn to the proof of (3.53c). Let us write $M := V - H$ and define $Q^{(k)} := \sum_{y\in B_r(x)\setminus B_k(x)} \mathbf{1}_y\mathbf{1}_y^*$, $k \in [r]$, to be the projection onto the coordinates $B_r(x) \setminus B_k(x)$. Then on Ξ there exists $c > 0$ such that $\mu - \|Q^{(k)}MQ^{(k)}\| \geq c\psi_+$, $k \in [r]$. The eigenvalue eigenvector equation $M\mathbf{w} = \mu\mathbf{w}$ becomes

$$(Q^{(k)}MQ^{(k)} - \mu)Q^{(k)}\mathbf{w} = -Q^{(k)}M(1-Q^{(k)})\mathbf{w} = -Q^{(k)}H(1-Q^{(k)})\mathbf{w}|_{S_{k-1}(x)}, \tag{3.62}$$

where in the last equality we used the fact that $B_r(x)$ is a tree and thus the only neighbors of $B_r(x) \setminus B_r(k)$ within $B_k(x)$ are precisely $S_k(x)$.
In other words, we have the following induction

$$\|Q^{(k)}\mathbf{w}\| \leq \frac{\|w|_{S_{k-1}(x)}\|\|Q^{(k)}H(1-Q^{(k)})\|}{c\psi_+}. \tag{3.63}$$

Using Lemma 5.10, we can bound $\|Q^{(k)}H(1-Q^{(k)})\| \leq 2\sqrt{\alpha_x}$. It follows that, there exists $C \geq 0$ large enough, such that

$$\|\mathbf{w}|_{B_r(x)\setminus B_k(x)}\| \leq C\frac{\sqrt{\alpha_x}\|\mathbf{w}|_{S_k}\|}{v(\mathfrak{u}_+)} = \|\mathbf{w}|_{S_k}\|O\left(\frac{\log\log N}{\sqrt{\log N}}\right), \tag{3.64}$$

where we used $\frac{\sqrt{\mathfrak{u}_+}}{v(\mathfrak{u}_+)} \leq \frac{1}{\sqrt{D_x}} \leq (\log N \vee d)$. Since $|\mathbf{w}(x)| \leq 1$, we conclude (3.53c). This concludes the proof. ☐

Proof of Proposition 3.28. We work on the event defined by Proposition 3.34 and on which $\big(B_{r+2}(x)\colon x \in \mathcal{U}^+(\varepsilon)\big)$ are disjoint trees.
Let us fix $x \in \mathcal{U}^+(\varepsilon)$. $B_{r+2}(x) \setminus B_r(x)$ is a forest and $\max_{y\in S_r(x)}|v_y| = O(\sqrt{\log N})$. The first assertion in (3.45c) follows immediately from (3.53c), as for the second claim we have, using again Lemma 5.10 to control the adjacency matrix,

$$\begin{aligned}\|(\underline{L} - \mu)\mathbf{w}\| &= \|(\underline{L} - \underline{L}|_{B_r(x)})\mathbf{w}\| \leq \|\underline{A}|_{S_{r+1}(x)}\|\|\mathbf{w}|_{S_r(x)}\|\\ &\leq 2\Big(\max_{z\in S_{r+1}(x)}\alpha_z\Big)^{1/2} O\Big(\Big(\frac{\omega}{\log N}\Big)^{r/2}\Big) = O\left(\frac{\omega^{r/2}}{(\log N)^{r/2-1}}\right).\end{aligned} \tag{3.65}$$

The two other points follow directly from the two claims of (3.53). ☐

3.6 Spectrum around small vertices

In this section, we study the eigenvalues generated by small degree vertices. We split our rigidity results into two different propositions. The first proposition includes regimes of d for which the number of leaves is very large and has weaker error bounds. The second one considers regimes of d where the number of small degree vertices remain small and has better error bounds.

Proposition 3.35 (Spectrum of $\underline{L}|_{B_r(x)}$, $x \in \mathcal{U}^-$ and d subcritical)**.** *Let $\alpha, \varepsilon \in (0, 1/3)$, $K > 0$. There exist $c_* > 0$ depending on K and α such that for any constant $r \geq \max(10, 10/c_*)$ and κ, κ as in (3.29), if*

$$\frac{1}{2}\log N + (\log N)^{1/K} \leq d \leq \log N - \log\log N,$$

then the following holds with probability $1 - O\big(\mathrm{e}^{-d^{c_*}}\big)$.

(i) For each $x \in \mathcal{U}^-(\varepsilon)$, the smallest eigenvalue of $\underline{L}|_{B_r(x)}$ written μ satisfy

$$\mu = v_x + \frac{\alpha_x}{v_x} + O\left(\frac{1}{\sqrt{d}(\log N)^{2c_*}}\right). \tag{3.66a}$$

(ii) If $x \in \mathcal{L} \setminus \mathcal{W}^\gamma(\kappa)$ then

$$\mu = \tilde{\Lambda}_x^{\mathcal{L}} + O\left(\frac{\omega(2\alpha)}{(\log N)^{2-\frac{1}{6}}(v_{(1)} - v_\gamma)^2}\right) \tag{3.66b}$$

with $\tilde{\Lambda}_x^{\mathcal{L}}$ defined in (3.27).

(iii) If $x \in \mathcal{W}^\gamma(\kappa)$ then

$$\mu = \Lambda_x^{\mathcal{L}} + O\bigg(\frac{(\log\log N)^4}{(\log N)^2(v_{(1)} - v_\gamma)^3}\bigg) \tag{3.66c}$$

with Λ_x defined in (3.27).

Moreover if $\mathbf{w}$ is the eigenvector corresponding to μ we have

$$\|\mathbf{w}|_{B_r(x)\setminus B_i(x)}\| = O\bigg(\frac{\omega(\alpha)}{\sqrt{\psi_-}}\bigg(\frac{\omega(\alpha)}{\sqrt{\log N}}\bigg)^{\frac{i}{2}}\bigg), \quad i \in [r] \tag{3.67}$$

where $\psi_- := (d - \frac{1}{2}\log N) \wedge \sqrt{\log N}$

Remark 3.36. If we suppose

$$\left(\frac{1}{2} + \varepsilon\right)\log N \leq d \leq (\log N)^2,$$

for some constant $\varepsilon > 0$, the proof becomes much shorted and the error bounds simpler. The underlying mechanism is that as d gets closer to $\frac{1}{2}\log N$, we cannot exclude the possibility that a ball $B_r(x)$ for some $x \in \mathcal{L}$ contains a vertex z such that $v_z - v_{(1)} \ll \sqrt{\log N}$. Therefore the perturbation analysis becomes more complicated. However, we can insure that there will be at most one such vertex, which makes the analysis feasible.

Proof. We work on the event defined in Propositions 5.23 and 5.26 and Lemmas 5.13, 5.22 and 5.24, which we denote by Ξ. Then there exists $c_* > 0$ such that $\mathbb{P}(\Xi) \geq 1 - O(\mathrm{e}^{-d^{c_*}})$. Note in particular that on Ξ the balls $\big(B_r(x) \colon x \in \mathcal{U}^-(\varepsilon)\big)$ are disjoint trees. Finally observe that for the values of d we consider, $x \in \mathcal{U}^-(\varepsilon)$ means that $D_x \leq \varepsilon^{-1} = O(1)$ (see (3.12)) and therefore on Ξ there exists $C > 0$ such that $\max_{i\in[r], x\in\mathcal{U}^-(\varepsilon)}|S_i(x)| \leq Cd^{i-1}$.

Let us fix $x \in \mathcal{U}^-(\varepsilon)$ and write $V := \frac{D-d}{\sqrt{d}}|_{B_r(x)}$ and abbreviate $H := fA|_{B_r(x)}$ with $f = d^{-1/2}$. By Lemma 5.10 and the fact that on Ξ we have

$$\min_{y \in B_r(x),\, y \neq x} |v_x - v_y| \geq c_* (\log N)^{c_*}, \qquad \max_{y \in B_r(x)} |v_y| = O\big(\sqrt{\log N}\big),$$

we see that $\|H\| \leq C$ for some constant $C \geq 0$ and that we are in the setup of Proposition A.10. Let $k := \lceil 5/c_* \rceil$ and $\psi = c_* (\log N)^{c_*}$, (A.13) becomes

$$\mu = v_x + \sum_{l=1}^{k} \frac{1}{d^l} \sum_{e=0}^{l-1} E_{2l}(e) + O\bigg(k \frac{\|H\|^{2k}\sqrt{d}}{\psi^{2k}}\bigg). \tag{3.68}$$

where $E_l(e)$ is defined in (A.14). Therefore we find that for our choice of k, the last term on the right-hand side of (3.68) is bounded by $O\big((\log N)^{-5}\big)$.

We define

$$y_* := \operatorname{argmin}\big\{v_y \colon\ y \in B_r(x)\, y \neq x\big\}, \quad r_* := \operatorname{dist}(x, y_*).$$

By Proposition 5.23 (ii), we have

$$\min_{y \in B_r(x) \setminus \{x, y_*\}} |v_x - v_y| \geq c_* \sqrt{\log N}, \quad |v_x - v_{y_*}| \geq c_* (\log N)^{c_*}. \tag{3.69}$$

To analyze (3.68), we will distinguish three cases.

The first case is when $r_* > 1$ and $D_x > 1$. The term $E_{2l}(e)$ has an expression in terms of cycles starting at x of length $2l$ (see (A.14)). As can be seen by a simple combinatorial argument, using the fact that on Ξ there exists $C > 0$ such that $\frac{|S_i(x)|}{|S_{i-1}(x)|} \leq Cd$, for $i = 2, \dots, r$, there are at most $D_x (Cd)^{l-1}$ such cycles. Moreover a cycle starting at x with length greater that 3 must visit at least twice $S_1(x)$, thus collecting at least to factors $\frac{1}{\min_{y \in S_1(x)} |v_x - v_y|}$. But since $r_* > 1$, such factors are of the order $(\log N)^{-c_*}$. We conclude that on Ξ,

$$|E_{2l}(e)| \leq \frac{D_x (Cd)^{l-1}}{d^l (\log N)} = O\Big(\frac{1}{d (\log N)^{2c_*}}\Big), \quad l \geq 2,\, e \geq 0. \tag{3.70}$$

If $x \in \mathcal{U}^-(\varepsilon)$ we get

$$\frac{1}{d} \sum_{y \in S_1(x)} \frac{1}{(v_x - v_y)} = \frac{\alpha_x}{v_x} + O\Big(\frac{D_x \max_{y \in S_1(x)} |v_y|}{\min_{y \in S_1(x)} |v_x - v_y|^2 d}\Big) = \frac{\alpha_x}{v_x} + O\Big(\frac{1}{\sqrt{d} (\log N)^{2c_*}}\Big), \tag{3.71}$$

which proves (i).

Next we consider the case when $D_x = 1$ and $x \in \mathcal{W}^\gamma(\kappa)$. Recalling Definition 3.20, we denote z_x the unique neighbour of x and write $v_x = v_{(1)}$. In particular $\max_{y \neq x, z_x} |v_y| \leq C (\log\log N)^2$ for some $C \geq 0$. If $x \in \mathcal{W}^\gamma(\kappa)$, this means that $r_* = 1$ and $v_{y_*} = v_\gamma + O\big(\frac{\kappa}{\sqrt{d}}\big)$. In this case, we consider the definition of $E_l(e)$ in (A.14) and see that at each step of the path either we are at z_x, in which case we gain a factor $(\log N)^{-c_*/2} d^{-1/2}$, or we are not at z_x, in which case we gain a factor $(\log N)^{-1/2}$ from (3.69). The last case is when we are at x and not at the end of the path which will create a term $(\log N)^{-c_*/2} d^{-1/2} \wedge (\log N)^{-1/2}$ by differentiation (see proof of Proposition A.10). This shows that

$$\max_{l \geq 8, e} |E_l(e)| = O\left(\frac{1}{(\log N)^3 (v_{(1)} - v_\gamma)^3}\right). \tag{3.72}$$

We compute explicitly

$$
\begin{aligned}
E_6(1) &= \frac{1}{d^3(v_{(1)}-v_{z_x})^3}\sum_{y\in S_2(x)}\frac{1}{v_{(1)}-v_y} \\
&= \frac{\alpha_{z_x}}{d^2(v_{(1)}-v_\gamma)^3 v_{(1)}} + \frac{1}{d^2(v_{(1)}-v_\gamma)^3 v_{(1)}^2}\frac{1}{d}\sum_{y\in S_2(x)} v_y + O\Big(\frac{1}{d^2(\log N)^3(v_{(1)}-v_\gamma)^3}\Big), \\
E_6(2) &= \frac{1}{d^3(v_{(1)}-v_{z_x})^3} = O\Big(\frac{(\log\log N)^4}{(\log N)^3(v_{(1)}-v_\gamma)^3}\Big),
\end{aligned}
\tag{3.73}
$$

and

$$
\begin{aligned}
E_2(0)+E_4(1) &= \frac{1}{d(v_{(1)}-v_{z_x})} + -\frac{1}{d^2(v_{(1)}-v_{z_x})^3} \\
&= \frac{1}{d(v_{(1)}-v_\gamma)} + \frac{v_\gamma - v_{z_x}}{d(v_{(1)}-v_\gamma)^2} + O\Big(\frac{\kappa^2}{(\log N)^2(v_{(1)}-v_\gamma)^3}\Big).
\end{aligned}
\tag{3.74}
$$

The only term remaining is $E_4(0)$. Proceeding as we did in (3.57) and using (3.57), we find

$$
\begin{aligned}
E_4(0) &= \frac{1}{d^2(v_{(1)}-v_{z_x})^2}\sum_{y\in S_2(x)}\frac{1}{v_{(1)}-v_y} \\
&= \Big[\frac{1}{d(v_{(1)}-v_\gamma)^2} + O\Big(\frac{\kappa}{d^{3/2}(v_{(1)}-v_\gamma)^3}\Big)\Big]\Big[\frac{\alpha_{z_x}}{v_{(1)}} + \frac{\sqrt{d}\alpha_{z_x}(\beta_{z_x}-1)}{v_{(1)}^2} - \frac{\alpha_{z_x}}{v_{(1)}^2\sqrt{d}} + O\Big(\frac{(\log\log N)^4}{(\log N)^{3/2}}\Big)\Big] \\
&= \frac{\alpha_{z_x}}{d(v_{(1)}-v_\gamma)^2 v_{(1)}} + \frac{\sqrt{d}\alpha_{z_x}(\beta_{z_x}-1)}{d(v_{(1)}-v_\gamma)^2 v_{(1)}^2} + O\Big(\frac{(\log\log N)^4}{(\log N)^{5/2}(v_{(1)}-v_\gamma)^2}\Big).
\end{aligned}
$$

Combining the above equation with (3.68), (3.72), (3.73) and (3.74), we conclude

$$
\mu = \Lambda_x^{\mathcal{L}} + O\Big(\frac{(\log\log N)^4}{(\log N)^2(v_{(1)}-v_\gamma)^3}\Big).
$$

where we used $\kappa = O(\log\log N)$ and $|v_{(1)}-v_\gamma| \le \sqrt{\log N}$.

Finally we consider the case where $D_x = 1$ but $x \notin \mathcal{W}^\gamma(\kappa)$. In this case we see that $E_6(1), E_6(2), E_4(1)$ can be bounded by the right-hand side of (3.66b). Recalling (3.27) we see that (3.68) becomes

$$
\mu = \tilde{\Lambda}_x^{\mathcal{L}} + \frac{v_z}{dv_{(1)}^2} + E_4(0) + O\left(\frac{(\log\log N)^4}{(\log N)^2(v_{(1)}-v_{z_x})^2}\right).
$$

Using Proposition 5.23 (iv) with $\alpha = 1/6$, we have

$$
\frac{1}{d}\sum_{y\in S_2(x)}\frac{1}{v_{(1)}-v_y} = \frac{\alpha_{z_x}}{v_{(1)}} + \frac{d^{1/6}}{\min_{y\in S_2(x)}|v_{(1)}-v_y|^2}.
$$

Either, $z_x = y_*$ and $v_{z_x} \le c_*\sqrt{d}$ and then we know, by Proposition 5.23 (iii) that $\min_{y\in S_2(x)}|v_{(1)}-v_y|^2 \ge c_*\log N$ and $v_{z_x} - v_{(1)} \ge c_*(\log N)^{c_*}$. Or $v_{z_x} \ge c_*\sqrt{d}$ in which case $\min_{y\in S_2(x)}|v_{(1)}-v_y|^2 \ge c_*^2(\log N)^{2c_*}$. In either case, we find,

$$
\begin{aligned}
E_4(0) &= \frac{\alpha_{z_x}}{dv_{(1)}(v_{(1)}-v_{z_x})^2} + O\Big(\frac{d^{1/6}}{d(v_{(1)}-v_{z_x})^2\min_{y\in S_2(x)}|v_{(1)}-v_y|^2}\Big) \\
&= \frac{\alpha_{z_x}}{dv_{(1)}(v_{(1)}-v_{z_x})^2} + O\left(\frac{(\log\log N)^4}{(\log N)^{2-\frac{1}{6}+2c_*}}\right).
\end{aligned}
$$

This proves (3.66b).
There only remains to prove (3.67). The proof goes exactly as the one for (3.53c). We use the fact that for all but at most one $i \in [r]$ the relation (3.63) holds with ψ_+ replaced by $c\sqrt{\log N}$ for $c > 0$ small enough. If there is $y \in S_i(x)$ such that $|v_x - v_y| \le c(\log N)^c$ for some $c > 0$, then we apply (3.63) with $c(\log N)^c$ instead of ψ_+. Since by Proposition 5.23 (iii) this situation happens for at most one $i \in [r]$ we conclude (3.49d). ☐

Proposition 3.37 (Spectrum of $\underline{L}|_{B_r(x)}$, $x \in \mathcal{U}^-$ and d critical). *Let $\alpha, \varepsilon > 0$ and $r \ge 10$. There exist $c_* > 0$ depending on ε and r such that if*

$$\log N - (\log\log N)^2 \le d \le (\log N)^2,$$

and κ, κ as in (3.29). Then the following holds with probability $1 - O\left(\mathrm{e}^{-d^{c_*}}\right)$.

(i) For each $x \in \mathcal{U}^-(\varepsilon)$, the largest eigenvalue of $\underline{L}|_{B_r(x)}$ written μ satisfy

$$\mu = v_x + \frac{\alpha_x}{v_x} + O\left(\frac{1}{\log N}\right). \tag{3.75a}$$

(ii) If $x \in \mathcal{W}^-(\kappa)$ then

$$\mu = \Lambda_x + O\left(\frac{\omega(\alpha)}{d(\log N)^2}\right) \tag{3.75b}$$

Moreover if $\mathbf{w}$ is the eigenvector corresponding to μ we have

$$\|\mathbf{w}|_{B_r(x)\setminus B_i(x)}\| = O\left(\left(\frac{1}{\log N}\right)^{\frac{i+1}{2}}\right), \quad i \in [r]. \tag{3.75c}$$

Proof. We work on the event defined by Propositions 5.23 and 5.25 and Lemma 5.24 that we denote Ξ. As in the proof of Proposition 3.35, we see that we are in the setup of Proposition A.10 with $\psi = c_*\sqrt{\log N}$, for some $c_* > 0$ small enough. For $k = 3$, (A.13) becomes

$$\mu = v_x + E_2(0) + E_4(1) + E_4(0) + O\left(\frac{\|H\|^6}{\psi_-^5}\right).$$

Since $B_r(x)$ is a tree, $\|H\| \le 2$ on Ξ.
We treat the three terms as we treated their analog in the proof of Proposition 3.34 (see (3.60) and (3.61)). We use $d \ge \frac{1}{2}\log N$ and Lemma 5.24 to remove the terms that are $O\left(\frac{d^{2\alpha}}{(\log N)^3}\right)$ In particular using Proposition 5.25 (ii) we find that

$$\frac{1}{dv_x^3}\frac{1}{d}\sum_y v_y^2 = \frac{\alpha_x}{v_x^3}\left[1 + O\left(\frac{d^\alpha}{(\log N)^{3/2}}\right)\right], \quad \frac{|S_1(x)|d}{d^2 v_x \psi_-^3} = O\left(\frac{\omega}{(\log N)^3}\right).$$

We skip the details as they are very similar to the proof of Proposition 3.34.
The proof of (3.75c) is the same as (3.53c) in Proposition 3.35. We skip the details and conclude the proof. ☐

Proof of Proposition 3.31. Exactly as the proof of Proposition 3.28, this time using a wider margin of safety. Setting $r = \frac{10}{c_*}$ we obtain exactly as in (3.65) that

$$\begin{aligned}\|(\underline{L} - \mu)\mathbf{w}\| =& \left\|(\underline{L} - \underline{L}|_{B_r(x)})\mathbf{w}\right\| \le \left\|A|_{S_{r+1}(x)}\right\|\left\|\mathbf{w}|_{S_r(x)}\right\| \\ \le& 2d^{-1/2}\left(\max_{z \in S_{r+1}(x)} D_z\right)^{1/2} O\left(\frac{\omega}{\sqrt{\psi_-}}\left(\frac{\omega}{\sqrt{\log N}}\right)^{10/c_* - 1}\right) = O\left(\frac{1}{(\log N)^{10}}\right).\end{aligned}$$

All of the equations (3.49) follow from their counterpart in Propositions 3.35 and 3.37.
In addition (3.49c) follows from the fact that for $d \geq \log N - (\log\log N)^2$ we have $v_{(1)} - v_\gamma \geq c\log N$ for some small enough $c > 0$. We have

$$\Lambda_x^{\mathcal{L}} - \tilde{\Lambda}_x^{\mathcal{L}} = \frac{\alpha_{z_x}}{\sqrt{d}v_{(1)}(v_{(1)} - v_{z_x})^2}\left[\frac{1}{\sqrt{d}} + (\beta_{z_x} - 1)\right] = O\left(\frac{1}{(\log N)^{5/2}}\right),$$

where we used Lemma 5.22 to bound $(\beta_{z_x} - 1) = O((\log N)^{-1/2})$.
We have, expanding the denominator $v_{(1)} - v_{z_x}$,

$$\begin{aligned}\tilde{\Lambda}_x^{\mathcal{L}} - \Lambda_x &= \frac{\alpha_1}{v_{(1)}} + \frac{\alpha_1}{v_{(1)}^3} + \frac{\alpha_1}{v_{(1)}^2}\sqrt{d}\left(\frac{|S_2(x)|}{d} - 1\right) - \frac{1}{d(v_{(1)} - v_{z_x})} - \frac{v_{z_x}}{dv_{(1)}^2} \\ &\leq \frac{1}{dv_{(1)}^3} - \frac{1}{d^{3/2}v_{(1)}^2} - \frac{v_{z_x}}{dv_{(1)}^3} + O\left(\frac{|v_{z_x}|^3}{d(v_{(1)} - v_{z_x})^4}\right) = O\left(\frac{1}{(\log N)^{5/2}}\right).\end{aligned}$$

This concludes the proof. □

3.7 Eigenvalue processes

In this section, we will prove that the point processes described in Corollaries 3.17, 3.19 and 3.22 are Poisson Point Processes.

Let us recall that we consider processes with $\mathcal{K} = \log\log N$ points (see (3.7)).

Let us introduce the quantity

$$Q(a,b) := \mathbb{P}(\mathcal{P}_{da} - da \geq b\sqrt{da}). \tag{3.76}$$

Lemma 3.38 (Linearization of the approximate eigenvalues)**.** *Let $K, \varepsilon > 0$ be constants and $\kappa > 0$ that may depend on N. Then*

(i) For $(\log N)^{1+\varepsilon} \leq d$ and $x \in \mathcal{W}^+(\kappa)$, we have

$$v_x + \frac{\alpha_x}{v_x} = v_x + \frac{1}{\sigma} + \frac{1}{\sqrt{d}} + O\Big(\frac{\kappa}{\log N}\Big) \tag{3.77}$$

(ii) For $1 \leq d \leq (\log N)^2$ and $x \in \mathcal{W}^+(\kappa)$, we have

$$\Lambda_x = \sigma + \sqrt{d}\theta(v_x - v(\mathfrak{u}_+)) + \frac{1}{\tau}\sqrt{\alpha_x}d(\beta_x - 1) + O\Big(\frac{\kappa^2}{\sqrt{d}v(\mathfrak{u}_+)^3}\Big) \tag{3.78}$$

(iii) For $\frac{1}{2}\log N \leq d \leq (\log N)^2$ and $x \in \mathcal{W}^-(\kappa)$, we have

$$\Lambda = \sigma(\mathfrak{u}_-) + \theta_-(\sqrt{d}(v_x - v(\mathfrak{u}_-))) + \frac{1}{\tau_{\text{crit},-}}\sqrt{\alpha_x}d(\beta_x - 1) + O\Big(\frac{\kappa^2}{\sqrt{d}v(\mathfrak{u}_+)^3}\Big). \tag{3.79}$$

for $\Lambda \in \{\Lambda_x, \Lambda_x^{\mathcal{L}}, \tilde{\Lambda}_x^{\mathcal{L}}\}$.

(iv) For $\frac{1}{2}\log N \leq d \leq \frac{\log N}{\log\log N}$ and $x \in \mathcal{W}^\gamma(\kappa)$, we have

$$\Lambda_x^{\mathcal{L}} = \sigma_\gamma + \theta_\gamma(v_{z_x} - v_\gamma) + \frac{1}{\tau_\gamma}(\beta_{z_x} - 1) + O\Big(\frac{\kappa^2\log(d)}{d^{9/4}(v_{(1)} - v_\gamma)^3} + \frac{\kappa^2}{d^{5/2}}\Big) \tag{3.80}$$

The proof of Lemma 3.38 relies on a Taylor expansion in $\Delta t \asymp O(\kappa/\sqrt{d})$. Note that if $t = v_x$ then $\frac{t}{\sqrt{d}} + 1 = \alpha_x$ and $\sqrt{d}(v_x - v(\mathfrak{u}_-)) = D_x - d\mathfrak{u}_+$. It is deferred to Section 3.9.

Asymptotic behavior of Φ^+ in critical and subcritical regime

In this section we prove Theorem 3.8 in the regime $K \leq d \leq (\log N)^2$. The starting point is Corollary 3.19 which tells us that to understand the eigenvalue point process in the window χ_+ (c.f. (3.24)) it suffices to look at the quantities Λ_x, $x \in \mathcal{W}(\kappa)$. The first step in the proof is to check that the window χ_+ (which morally lives on the space of degrees) is larger than the window defined by κ_+ (which lives on the space of point processes and depends on $\mathcal{K}$ and ρ). Indeed we have

$$v(\mathfrak{u}_+) - \frac{\kappa}{2\sqrt{d}} \leq \sigma - \frac{\kappa}{2\sqrt{d}} \ll \sigma - \frac{\kappa_+}{\tau}, \tag{3.81}$$

since $\tau \gg \sqrt{d}$ and $\kappa_+ \ll \kappa$ by definition.

Let us define the reference process

$$\Sigma := \sum_{x\in[N]} \delta_{Z_x}, \quad Z_x := \begin{cases} \tau\theta(D_x - d\mathfrak{u}_+) + d\sqrt{\alpha_x}(\beta_x - 1) & \text{if } D_x - d\mathfrak{u}_+ \geq -\sqrt{d}\kappa \\ -\infty & \text{otherwise.} \end{cases} \tag{3.82}$$

We define the error parameter

$$\eta := \frac{1}{(\log\log N)^2}, \tag{3.83}$$

chosen in order to have $\tau\varepsilon_x \ll \eta \ll \max_{s\geq\kappa} \frac{1}{\frac{\mathrm{d}}{\mathrm{d}s}\rho(E_s)}$ for ε_x defined in (3.22) and ρ defined below.

Lemma 3.39. *Under the assumptions of Corollary 3.19, we have for every $s \geq -\kappa_+$*

$$\Sigma\big(E_{s+\eta}\big) \leq \Phi^+\big(E_s\big) \leq \Sigma\big(E_{s-\eta}\big).$$

Proof. Let us introduce the intermediate process

$$\tilde{\Sigma} := \sum_{x\in[N]} \delta_{\tilde{Z}_x}, \quad Z_x := \begin{cases} \tau(\Lambda_x - \sigma(\mathfrak{u}_+)) & \text{if } v_x - v(\mathfrak{u}_+) \geq -\sqrt{d}\kappa \\ -\infty & \text{otherwise.} \end{cases}$$

By (3.81), only vertices belonging to $\mathcal{W}^+(\kappa)$ contribute to $[-\kappa_+, \infty)$. Moreover by Corollary 3.19 since $\eta \ll \tau\varepsilon_x$, for ε_x as defined therein, we have

$$\tilde{\Sigma}(E_{s+\eta/2}) \leq \Phi^+(E_s) \leq \tilde{\Sigma}(E_{s-\eta/2}).$$

Recalling Lemma 3.38 and in particular the error term on the right-hand side of (3.78), we see that

$$\tau\frac{\kappa^2}{\sqrt{d}v(\mathfrak{u}_+)^3} \ll \eta.$$

We conclude that $\Sigma(E_{s+\eta/2}) \leq \tilde{\Sigma}(E_s) \leq \Sigma(E_{s-\eta/2})$. □

We define the intensity measure

$$\tilde{\rho}(E_s) := \sum_{v\in\mathbb{N}} N\mathbb{P}(\mathcal{P}_d = v)Q(v, s - \tau\theta(v - d\mathfrak{u}_+))\mathbf{1}_{v\geq 5\vee\frac{\log\log N}{\log(\mathfrak{u}_+)}} \tag{3.84}$$

The next lemma states that $\tilde{\rho}$ is a good approximation of ρ.

Lemma 3.40. *For κ_+ as defined in (3.8) and for every $s \geq \kappa_+$*

$$\tilde{\rho}(E_s) = \rho(E_s)\bigg(1 + O\bigg(\frac{1}{(d\mathfrak{u}_+)^{1/5}}\bigg)\bigg) + O\big(\mathrm{e}^{-(\log N)^{1/5}}\big).$$

Moreover

$$\sup_{u\geq s}\left|\frac{\mathrm{d}}{\mathrm{d}s}\tilde{\rho}(E_{s-\eta})\right| \leq 3\mathcal{K}, \quad \tilde{\rho}(E_{s+-\eta}) \leq 3\mathcal{K}. \tag{3.85}$$

Proof. Let $K = (\log\log N)^2 \vee \frac{\log\log N}{\log(\mathfrak{u}_+)}$. By a Taylor expansion, we see that for $k \geq -K$,

$$N\mathbb{P}[\mathcal{P}_d = k + d\mathfrak{u}_+]\mathbf{1}_{k\geq \tilde{a}} = \mathfrak{u}_+^k\big(1 + O\big(\mathrm{e}^{K/d\mathfrak{u}_+}\big)\big)\Big(1 + O\Big(\frac{1}{d\mathfrak{u}_+}\Big)\Big) = \mathfrak{u}_+^k\Big(1 + O\Big(\frac{1}{d\mathfrak{u}_+}\Big)\Big).$$

Moreover, applying Lemma B.4 with $\mu = d^2\mathfrak{u}_+$ and $xi = \frac{1}{10}$ we find

$$Q(v, s - \tau\theta(v - d\mathfrak{u}_+)) = \begin{cases} G(s)(1 + (d\mathfrak{u}_+)^{-\frac{1}{5}}), & \text{if } |s - \tau\theta(v - d\mathfrak{u}_+)| \leq \mu^{\xi} \\ O\big(\mathrm{e}^{-(s^2\wedge\mu)}\big) & \text{else.} \end{cases}$$

We bound the contribution of the lower term by $O\big(\mathrm{e}^{-(\log N)^{1/5}}\big)$.

(3.85) is immediate from the definition of κ in (3.8) and the fact that

$$\sup_{s\geq\kappa-\eta} \rho(E_s) - \tilde{\rho}(E_s) \leq \rho(E_{\kappa-\eta})\frac{C\mathcal{K}}{(d\mathfrak{u}_+)^{c_*}} \leq 3\mathcal{K}.$$

This concludes the proof. □

The next step is to prove that the k-point correlation functions of the process Σ factorize asymptotically. Let us recall the definition of the $l-$point correlation measure q_Σ of a point process Σ, given in (3.91).

Lemma 3.41 (Inclusion exclusion for Σ)**.** *Let $\ell \leq c\frac{\log N}{\log \mathfrak{u}_+}$ for some small enough $c > 0$. For all $i \in [\ell]$, let $I_i = [a_i, b_i)$ with $-\kappa - \eta \leq a_i < b_i$. Then*

$$q_\Sigma(I_1 \times \cdots \times I_\ell) = \Big[1 - O\Big(\frac{l^2}{N}\Big)\Big] \prod_{i\in[\ell]} \tilde{\rho}(I_i) + O(N^{-1/5}).$$

Proof. Let $\ell \in \mathbb{N}$ and $a_1, \ldots, a_l \in \mathbb{R}$. A straightforward adaptation of Proposition 5.27 gives us

$$\begin{aligned} N^\ell\mathbb{P}\Big(\bigcap_{i\in[\ell]}\{Z_i \geq a_i\}\Big) &= N^\ell \sum_{v_1,\ldots,v_n\in\mathbb{N}} \Big(\prod_{i\in[\ell]} \mathbf{1}_{(v_i - d\mathfrak{u}_+)\tau\theta\geq\kappa}\Big)\mathbb{P}\Big(\bigcap_{i\in[\ell]}\{Z_i \geq a_i\}\Big) \\ &= \Big(\prod_{i\in[\ell]}\rho(E_i)\Big) + N^\ell O\Big(N^{-1/3}\mathbb{P}(\mathcal{P}_d \geq d\mathfrak{u}_+ - K)^\ell + \mathcal{K}^\ell N^{-\ell-1}\Big) \\ &= \Big(\prod_{i\in[\ell]}\rho(E_i)\Big) + O(N^{-1/4}), \end{aligned}$$

where we used (5.3) to estimate

$$\Big(N\mathbb{P}(\mathcal{P}_d \geq d\mathfrak{u}_+ - K)\Big) \leq \mathfrak{u}_+^{\ell K} \leq N^{1/5}$$

and the assumption on l for c small enough.

We have by Lemma 3.51,

$$\begin{aligned} N^\ell\mathbb{P}(Z_1 \in I_1, \ldots, Z_\ell \in I_\ell) &= \sum_{U\subset[\ell]} (-1)^{|U|}N^\ell\mathbb{P}\Big(\bigcap_{i\in U}\{Z_i \geq b_i\} \cap \bigcap_{i\in[\ell]\setminus U}\{Z_i \geq a_i\}\Big) \\ &= \sum_{U\subset[\ell]} (-1)^{|U|}N^\ell \prod_{\mathbb{P}}\Big(\prod_{i\in U}\rho(b_i) \prod_{i\in[\ell]\setminus U}\rho(b_i)\Big) + O\big(2^l N^{-1/4}\big) \\ &= \sum_{U\subset[\ell]} (-1)^{|U|}N^\ell \prod_{\mathbb{P}}\Big(\prod_{i\in U}\rho(b_i) \prod_{i\in[\ell]\setminus U}\rho(b_i)\Big) + O\big(N^{-1/5}\big), \end{aligned}$$

for $\ell \leq \frac{1}{20\log(2)}\log N$. We conclude by combining the two equations. □

Next we show that Σ behaves asymptotically as a PPP with density $\tilde{\rho}$.

Lemma 3.42 (Asymptotic Poisson Behavior of Σ)**.** *Suppose $n \in \mathbb{N}^*$ and $s \in \mathbb{R}$ satisfy*

$$n, \tilde{\rho}(E_s) \ll \frac{\log N}{\log \mathfrak{u}_+}.$$

Let $I_1, \dots, I_n \subseteq E_s$ be disjoint intervals of the form $I_i = [a_i, b_i)$ with $-\kappa - \eta \le a_i \le b_i$. Then for all $k_1, \dots, k_n \in \mathbb{N}$ we have

$$\mathbb{P}\Big(\bigcap_{i \in [n]} \big\{\Sigma(I_i) = k_i\big\}\Big) = \prod_i \frac{\tilde{\rho}(I_i) \mathrm{e}^{-\tilde{\rho}(I_i)}}{k_i!} + \mathcal{E}(k_1, \dots, k_n)$$

where the error term satisfies, for some small enough $c > 0$,

$$|\mathcal{E}(k_1, \dots, k_n)| \le \mathrm{e}^{-cm_*} + \mathrm{e}^{-\mathcal{K}}$$

Proof. Choose $m := c\frac{\log N}{\log \mathfrak{u}_+}$ for some constant $c > 0$ which will be chosen small enough in the following. We write $\rho = \tilde{\rho}$ in the rest of the proof. By Lemma 3.51, we find

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i \in [n]} \{\Sigma(I_i) = k_i\}\right) &= \frac{1}{k_1! \cdots k_n!} \sum_{l_1, \dots, l_n \in \mathbb{N}} \frac{(-1)^{\sum_i l_i}}{l_1! \cdots l_n!} q_\Sigma(I_1^{k_1 + l_1} \times \cdots \times I_n^{k_n + l_n}) \\
&= \frac{1}{k_1! \cdots k_n!} \sum_{l_1, \dots, l_n \in \mathbb{N}} \mathbf{1}_{\sum_i l_i \le m} \frac{(-1)^{\sum_i l_i}}{l_1! \cdots l_n!} q_\Sigma(I_1^{k_1 + l_1} \times \cdots \times I_n^{k_n + l_n}) + \mathcal{E}_0 \\
&= \frac{1}{k_1! \cdots k_n!} \sum_{l_1, \dots, l_n \in \mathbb{N}} \mathbf{1}_{\sum_i l_i \le m} \prod_{i \in [n]} \frac{(-1)^{l_i} \rho(I_i)^{k_i + l_i}}{l_i!} + \mathcal{E}_0 + \mathcal{E}_1 \\
&= \prod_{i \in [n]} \frac{\rho(I_i)^{k_i}}{k_i!} \mathrm{e}^{-\rho(I_i)} + \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2.
\end{aligned}$$

The first two equalities follow from Lemma 3.51 with $\ell = \sum_i l_i + k_i \le m$ (after choosing c small enough in the definition of m) and the error term is defined as

$$\mathcal{E}_0 := \frac{1}{k_1! \cdots k_n!} \sum_{l_1, \dots, l_n \in \mathbb{N}} \mathbf{1}_{\sum_i l_i = m+1} \frac{(-1)^{\sum_i l_i}}{l_1! \cdots l_n!} q_\Sigma(I_1^{k_1 + l_1} \times \cdots \times I_n^{k_n + l_n}).$$

The third equality follows after introducing

$$\begin{aligned}
\mathcal{E}_1 &:= \frac{1}{k_1! \cdots k_n!} \sum_{l_1, \dots, l_n \in N} \mathbf{1}_{\sum_i l_i \le m} \frac{1}{l_1! \cdots l_n!} \mathcal{E}_1(l_1, \dots, l_n), \\
\mathcal{E}_1(l_1, \dots, l_n) &:= q_\Sigma(I_1^{k_1 + l_1} \times \cdots \times I_n^{k_n + l_n}) - \prod_i \rho(I_i)^{k_i + l_i}.
\end{aligned}$$

The fourth equality is just the formula $\mathrm{e}^{-x} - \sum_{k \le m} \frac{x^k}{k!} \le \frac{1}{(k+1)!}$, $x > 0$, and

$$\mathcal{E}_2 := \prod \frac{\rho(I_i)^{k_i}}{k_i!} \sum_{l_1, \dots, l_n \in N} \mathbf{1}_{\sum_i l_i \le m+1} \prod_i \frac{(-\rho(I_i))^{l_i}}{l_i!}.$$

To control $\mathcal{E}_1$, we observe that $\sum_i k_i + l_i \leq 2m$ and so by Lemma 3.41

$$\begin{aligned}\sum_{l_1,\dots,l_n} \mathbf{1}_{\sum_i l_i = l} |\mathcal{E}_1(l_1,\dots,l_n)| \mathbf{1}_{\sum_i l_i \leq m} &\leq \frac{(\log N)^2}{N}\left(\prod_i \frac{\rho(I_i)^{k_i}}{k_i!}\right)\sum_{l_1,\dots,l_n} \mathbf{1}_{\sum_i l_i \geq m+1}\prod_i \frac{(-\rho(I_i))^{l_i}}{l_i!}\\ &\leq \frac{(\log N)^2}{N}\left(\prod_i \frac{\rho(I_i)^{k_i}}{k_i!}\right)\sum_{l\geq m+1}\frac{\left(3\sum_i \rho(I_i)\right)^l}{l!}\\ &= \frac{(\log N)^2}{N}\left(\prod_i \frac{\rho(I_i)^{k_i}}{k_i!}\right)\sum_{l\geq m+1}\frac{\left(3\rho(\bigcup_i I_i)\right)^l}{l!}.\end{aligned}$$

Similarly, we have

$$|\mathcal{E}_0| + |\mathcal{E}_2| = O\left(\prod_i \frac{\rho(I_i)^{k_i}}{k_i!}\right)\sum_{l\geq m+1}\frac{(3\rho(\bigcup_i I_i))}{l!}.$$

Using (3.85), it suffices to have $m \geq 3\mathrm{e}^2 n\mathcal{K}$ which is insured by the definition of m and the fact that $\log\log N \ll \frac{\log N}{\log \mathfrak{u}_+}$. Using the fact that the I_i are disjoint we see that $3\mathrm{e}^2\rho(\bigcup_i I_i) \leq m$ and so

$$|\mathcal{E}_0| + |\mathcal{E}_1| + |\mathcal{E}_2| \leq C(1 + \frac{(\log N)^2}{N})\left(\prod_i \frac{\rho(I_i)^{k_i}}{k_i!}\right)\mathrm{e}^{-2m} \leq \mathrm{e}^{-m}$$

Therefore

$$\sum_{k_1,\dots,k_n} \mathbf{1}_{\sum_i k_i \leq m}\mathrm{e}^{-m} \leq m^n \mathrm{e}^{-m} \leq \mathrm{e}^{-m/2}$$

for $n \leq m/\log m$. This concludes the proof for $\sum_i k_i \leq m$.
The case $\sum_i k_i \geq m$, we observe that

$$\begin{aligned}\sum_{k_1,\dots,k_n} \mathbf{1}_{\sum_i k_i > m}\prod_i \frac{\rho(I_i)^{k_i}\mathrm{e}^{-\rho(I+i)}}{k_i!} &= \sum_{l>m}\frac{(\sum_i \rho(I_i))^l \mathrm{e}^{-\sum_i \rho(I_i)}}{l!}\\ &\leq C\exp\left(-\frac{m}{\rho(\bigcup_i I_i)}\left(1+\rho\left(\bigcup_i I_i\right)\right)\log\left(\frac{m}{\rho(\bigcup_i I_i)}\right)\right)\\ &\leq C\mathrm{e}^{-\mathcal{K}}\end{aligned}$$

by (B.4). □

We are now ready to show the asymptotic closeness of Ψ^+ and Φ^+. Let $\tilde{\Psi}$ be the Poisson process with intensity measure $\tilde{\rho}$.

Lemma 3.43. *Fix $n \in \mathbb{N}^*$ and κ as in (3.8). Then*

$$\mathbb{P}\Big(\bigcap_{i\in[n]}\big\{\tilde{\Psi}(E_{t_i}) \leq k_i\big\}\Big) = \mathbb{P}\Big(\bigcap_{i\in[n]}\big\{\Psi(E_{s_i}) \leq k_i\big\}\Big) + o(1)$$

uniformly for $k_1,\dots,k_n \in \mathbb{N}$, $s_1,\dots s_n \geq -\kappa$ and $t_1,\dots,t_n$ satisfying $|t_i - s_i| \leq \eta$.

Proof. A simple exercise on Poisson processes shows that it suffices to establish $\tilde{\rho}(E_{t_i}) = \rho(E_{s_i}) + o(1)$. By (3.85) and Lemma 3.40, we have $\tilde{\rho}(E_{t_i}) = \rho(E_{t_i})$. Moreover an easy computation shows that $\frac{\mathrm{d}}{\mathrm{d}t}\rho(E_t) = O(\rho(E_t))$ and since $\eta = \frac{1}{(\log\log N)^2} = o\big(\frac{1}{\mathcal{K}}\big)$ and $\rho(E_t) = O(\mathcal{K})$, we conclude that $\rho(E_{t_i}) = \rho(E_{s_i}) + o(1)$. □

Proof of Theorem 3.8 for $d \le (\log N)^2$. By definition of κ_+ and κ in (3.24) and (3.15) respectively we see that

$$\kappa_+ \le 2\frac{\log(\mathcal{K})}{\log(\mathfrak{u}_+)} \ll \kappa,$$

since $\log \mathcal{K} \ll \log\log N$. By Lemmas 3.42 and 3.43 we can write for $t_1 \ge t_2 \ge \cdots \ge t_n \ge -\kappa - \eta$ and $r_1 \le \cdots \le r_n$ in $\mathbb{N}^*$,

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{i\in[n]}(\Phi(E_{t_i}) = r_i)\Big) &= \mathbb{P}\Big(\bigcap_{i\in[n]}(\Sigma(E_{t_i+\eta}) = r_i)\Big) = \mathbb{P}\Big(\bigcap_{i\in[n]}(\Sigma([t_i+\eta, t_{i-1}+\eta)) = r_i - r_{i-1})\Big)\\ &= \mathbb{P}\Big(\bigcap_{i\in[n]}(\tilde{\Psi}([t_i+\eta, t_{i-1}+\eta)) = r_i - r_{i-1})\Big) + \mathcal{E}(r_1,\dots,r_n)\\ &= \mathbb{P}\Big(\bigcap_{i\in[n]}(\tilde{\Psi}(E_{t_i+\eta}) = r_i)\Big) + o(1) = \mathbb{P}\Big(\bigcap_{i\in[n]}(\Psi(E_{t_i}) = r_i)\Big) + o(1).\end{aligned}$$

This concludes the proof. □

Asymptotic behavior of Φ^+ in supercritical regime

In this section we show that the eigenvalue process converges to Φ^+ in the supercritical regime $d \le (\log N)^{1+\varepsilon}$, $\varepsilon > 0$.

Definition 3.44. Let Φ be a random point process on $\mathbb{R}$. We can represent $\Phi = \sum_{x\in\mathcal{X}} \delta_{Z_x}$, where $\mathcal{X}$ is an index set and $(Z_x)_{x\in\mathcal{X}}$ is an exchangeable family of random variables. For $a, b \in \mathbb{R}$, $b \ne 0$, we denote by

$$\Phi^{(a,b)} := \sum_{x\in\mathcal{X}} \delta_{b(Z_x - a)},$$

The reference window is different, we introduce

$$\kappa_{\text{sup}} = (\log\log N)^2 \vee \frac{\log\log N}{\log \mathfrak{u}_+}$$

and the rescaling parameters

$$\alpha := \frac{1}{2}\log\frac{d}{2\log N}, \qquad \beta = \frac{1}{\sqrt{d}v(\mathfrak{u}_+)}.$$

In order to do that we introduce an intermediate point process Σ_{gumb}, which is the analog of (3.82),

$$\Sigma_{\text{gumb}} := \sum_{x\in[N]} \delta_{Z_x}, \quad Z_x := \begin{cases}\sqrt{2\log N}(v_x - v(\mathfrak{u}_+) - \alpha) & \text{if } v_x - v(\mathfrak{u}_+) \ge \kappa_{\text{sup}},\\ -\infty & \text{else.}\end{cases} \tag{3.86}$$

We will first show that Σ_{gumb} is asymptotically close to both Φ^+ and to Ψ_{sup}, a PPP with density $\rho_{\text{sup}} := \mathrm{e}^{-s}\mathrm{d}s$. Then we will show that Ψ_{sup} is asymptotically close to a rescaled version (zoomed-out version is more accurate) of ψ^+. This is illustrated by the chain of comparisons

$$\left(\Phi^+\right)^{(\alpha,\beta)} \overset{(1)}{\sim} \Sigma_{\text{gumb}} \overset{(2)}{\sim} \Psi_{\text{sup}} \overset{(3)}{\sim} \left(\Psi^+\right)^{(\alpha,\beta)}, \tag{3.87}$$

where $\sim$ denotes asymptotic closeness of point processes (with respect to the topology induced by the metric $\mathcal{D}_\kappa$), $\Psi^{(b,a)}$ denotes the process Ψ rescaled by $b \in R_+$ and shifted by $a \in \mathbb{R}$ (see Lemma 3.39 below). We will then be able to conclude that $\Phi^+ \sim \Psi^+$ by Lemma 3.46.

The step (1) in (3.87) is proved by adapting Lemma 3.39, using Corollary 3.17 instead of 3.19, the linearization (3.77) instead of (3.78) (in particular $\frac{1}{\sigma} - \frac{1}{v(\mathfrak{u}_+)} = O\big((\log N)^{-1}\big)$). This allows us to use Lemma 3.39 with the obvious modifications.

Step (2) of (3.87) is obtained by first proving asymptotic decoupling of the variables Z_x. This is an adaptation of Lemma 3.41 using Proposition 5.28 instead of 5.27 and the bound $\rho_{\sup}(\kappa) \le \mathcal{K}$.

Step (3) will follow from the next result.

Lemma 3.45. *The processes $\Psi_{\sup}$ and $\Psi^{(\alpha,\beta)}$ are asymptotically close in the sense that*

$$\mathcal{D}_{\kappa_{\sup}}(\Psi_{\sup}, \Psi^{(\alpha,\beta)}) \to 0.$$

Proof. Let $T \in [\kappa_{\sup}, +\infty)$. Let us write $\omega := \sqrt{\frac{2\log N}{d}}$ so that $\mathfrak{u}_+ = 1 + \omega$. Using $\frac{\theta\tau}{\sqrt{d}} = 2\log N(1 + o(1))$, we find

$$\begin{aligned}
\mathbb{E}\Big[\Psi\big([T, +\infty)\big)\Big] &= \int_T^\infty \sum_{k\in\mathbb{Z}} \mathfrak{u}_+^k g(s + k2\log N)\mathrm{d}s = \sum_{k\in\mathbb{Z}} \mathfrak{u}_+^k \int_T^\infty g(s + k2\log N)\mathrm{d}s \\
&= \sum_{k\in\mathbb{Z}} \mathfrak{u}_+^k \int_{T+k2\log N} g(s)\mathrm{d}s = \int_{\mathbb{R}} \mathrm{d}s \sum_{k\in\mathbb{Z}} \mathfrak{u}_+^k g(s)\mathbf{1}_{s\ge T+k2\log N} \\
&= \int_{\mathbb{R}} g(s) \sum_{k\le\frac{s-T}{2\log N}} \mathfrak{u}_+^k \mathrm{d}s = \sum_{k\ge 0} \mathfrak{u}_+^{-k} \int_{\mathbb{R}} g(s)\big(1+\omega\big)^{\frac{s-T}{2\log N}} \mathrm{d}s
\end{aligned}$$

Using a Taylor expansion of the function $\ln(1 + x)$ to the first order we see that

$$(1+\omega)^{\frac{s-T}{2\log N}} = \exp\left(\frac{s\omega}{2\log N} - \frac{\omega T}{2\log N} + O\left(\frac{(s-T)\omega^2}{2\log N}\right)\right)$$

By the monotone convergence theorem (or by computing the Gaussian integral directly), we find

$$\begin{aligned}
\mathbb{E}\left[\Psi\left([T, +\infty)\right)\right] &= \mathrm{e}^{-\omega T/2\log N} \frac{\mathfrak{u}_+}{\mathfrak{u}_+ - 1} \int_{\mathbb{R}} g(s) \exp\left(\frac{s\omega}{2\log N} + O\left(\frac{(s-T)\omega^2}{2\log N}\right)\right) \mathrm{d}s \\
&= \exp\left(-\frac{\omega T}{2\log N} + \frac{1}{2}\log\frac{d}{2\log N}\right)\left(1 + \frac{1}{(d\log N)^{1/4}} + \mathrm{e}^{-(d\log N)^{1/2}}\right),
\end{aligned}$$

where in the last step we split the integral between $|s| \ge (d\log N)^{1/4}$ and $|s| \le (d\log N)^{1/4}$. Setting $T = \beta(t - \alpha)$ we see that

$$\mathbb{E}\left[\Psi\left([T, +\infty)\right)\right] = \mathrm{e}^{-t}(1 + o(1)).$$

This shows the claim. □

The next result is evident from Definition 3.7.

Lemma 3.46. *Let Φ_1 and Φ_2 be two point processes on $\mathbb{R}$ and $a, b \in \mathbb{R}$. Then for any $\kappa \in R$,*

$$\mathcal{D}_\kappa(\Phi_1, \Phi_2) = \mathcal{D}_{b(\kappa - a)}(\Phi_1^{(b,a)}, \Phi_2^{(b,a)}),$$

Combining (3.87) with $\frac{1}{\beta}(\kappa_{\sup} + \alpha) \le -\frac{\log\log N}{2\beta} \ll \kappa$, for κ defined in (3.8), we conclude that

$$\mathcal{D}_\kappa\Big(\Phi^+, \Psi^+\Big) \le \mathcal{D}_{\beta^{-1}(\kappa_{\sup}+\alpha)}\Big(\Phi^+, \Psi^+\Big) = \mathcal{D}_{\kappa_{\sup}}\Big(\big(\Phi^+\big)^{(\alpha,\beta)}, \big(\Psi^+\big)^{(\alpha,\beta)}\Big) \to 0.$$

This proves Theorem 3.8 in the supercritical regime.

Asymptotic behavior of Φ^- for $d \gtrsim \log N$

In this section, we prove (i) of Theorem 3.12. The analysis of Φ^- in the regime $(\log N)^2 \leq d \leq N^{1/3-\varepsilon}$ is completely identical to the analysis of Φ^+ in the same regime. We therefore only focus on the regime $\log N - (\log\log N)^2 \leq d \leq (\log N)^2$.

The analysis of Φ^- starts by combining (i) of Corollary 3.22, (iii) of Lemma 3.38 and the fact that

$$\kappa_- \ll \kappa$$

for κ_- defined below (3.11) as κ defined in (3.15). The proof then unfolds as in the case Φ^+, we skip the details.

Asymptotic behavior of Φ^γ for $\frac{1}{2}\log N \ll d \lesssim \log N$

In this section, we prove (ii) of Theorem 3.12. We begin by recalling (i) of Corollary 3.22, (iii) of Lemma 3.38. Observe that

$$\kappa_- \ll \kappa$$

We define the generic error parameter

$$\eta = \frac{\kappa^2 \log d}{d^{1/4}(v_{(1)} - v_\gamma)^{1+\frac{1}{2}}}$$

Now we have $\tau_\gamma(\varepsilon_x \vee \Delta_x) \leq d^{-1/4}\eta$, where ε_x is defined in (ii) of Corollary 3.22 and Δ_x is the error in (3.80). Similarly to (3.82), we define the process $\Sigma := \sum_{x\in[N]} \delta_{Z_x}$ where

$$Z_x := \begin{cases} \tau_\gamma\theta_\gamma(v_{z_x} - v_\gamma) + d\sqrt{\alpha_{z_x}}(\beta_{z_x} - 1) & \text{if } D_{z_x} - d\mathfrak{u}_\gamma \leq \sqrt{d}\kappa \\ -\infty & \text{otherwise.} \end{cases}$$

Note that we consider $\sigma_\gamma - \lambda$, so that the smallest eigenvalues of Spec $\underline{L}$ correspond to the largest values of the point process Φ^γ. It thus makes sense to send larger eigenvalues to $-\infty$ and not $+\infty$.

We obtain a result similar to Lemma 3.39 and show that

$$\Sigma(E_{s+\eta}) \leq \Phi^\gamma(E_s) \leq \Sigma(E_{s-\eta}).$$

We can then use Proposition 5.29 to obtain a result similar to Lemma 3.41. For $\ell \leq N^{1/3}$ and $a_1, \ldots, a_\ell \in \mathbb{R}$ we have

$$N^\ell \mathbb{P}\Big(\prod_{i\in[l]} Z_i \geq a_i\Big) = \big(\prod_{i\in[l]} \tilde{\rho}^\gamma(E_i)\Big) + N^\ell O\bigg(\frac{d^3 l^2}{N}(\mathrm{e}^{-d}d)^\ell \mathbb{P}\big(\mathcal{P}_d \leq d\mathfrak{u}_\gamma + \sqrt{d}\kappa\big)^\ell\bigg),$$

where $\tilde{\rho}^\gamma$ is defined as $\tilde{\rho}$ in (3.84) but replacing $\mathfrak{u}_+$ by $\mathfrak{u}_\gamma$. We can then conclude that the error term is small, by using the fact that

$$d^3\ell^2\Big(N\mathrm{e}^{-d}d\mathbb{P}(\mathcal{P}_d \leq d\mathfrak{u}_\gamma + \sqrt{d}\kappa)\Big)^\ell \leq \mathfrak{u}_\gamma^{\ell\kappa} d^3\ell^2 \leq N^{1/5},$$

as long as $\ell \leq c\frac{\log N}{\log \mathfrak{u}_\gamma}$ for $c > 0$ small enough.

The rest of the proof is similar, as soon as we observe that $\tilde{\rho}^\gamma(E_{s+\eta}) = O\big(\mathcal{K}\big)$, a bound equivalent to (3.85). This concludes the proof for that last case.

3.8 Eigenvector localization

In this section, we prove Theorems 3.10 and 3.13. We will proceed as in Section 3.7 and detail the steps for Theorem 3.10 and $d \leq (\log N)^2$. We will then explain briefly the adaptations required for $d \geq (\log N)^2$ and Theorem 3.13.

First, we observe that the hypotheses on d are stronger in Theorem 3.10 than they are in Theorem 3.8. In particular, we suppose that

$$(\log \log N)^{1/4} \leq d \leq (\log N)^2 \tag{3.88}$$

Lemma 3.47 (Level spacing for Σ)**.** *Let d be as in Theorem 3.8, κ as in (3.8) and η and Z_x as defined in (3.83) and (3.82). Under the assumptions of Corollary 3.19, for any $a \in \mathbb{R}$ we have*

$$\mathbb{P}\big(\exists x \neq y : Z_x, Z_y \geq -\kappa, |Z_x - Z_y| \leq \eta\big) \leq \mathcal{K}^2$$

Proof. Let us recall that for $1 \leq d \leq (\log N)^2$, we have with high probability $\max_{x \in [N]} \operatorname{Deg}(x) \leq d\mathfrak{u}_+ + C\Big(1 + \sqrt{\frac{d}{\log N}}\Big)$, for $C \geq 1$ large enough. This follows for instance from Bennett's inequality. We deduce that $\max_x Z_x \leq C\frac{v(\mathfrak{u}_+)^2}{\sqrt{\mathfrak{u}_+}} \leq (\log N)^2$ for N large enough.
Thus

$$\mathbb{P}\big(\exists x \neq y : Z_x, Z_y \geq -\kappa, |Z_x - Z_y| \leq \eta\big) \leq q_{\Sigma,2}\Big(\Big\{(s,t) : s,t \in \big[-\kappa, (\log N)^2\big], |s-t| \leq \eta\Big\}\Big) + o(1),$$

where we used the two-point correlation measure $q_{\Sigma,2}$ defined in (3.91) below. By covering the set in the argument of $q_{\Sigma,2}$ by square of the form $[u-\eta, u+\eta]^2$, we find, using Lemmas 3.41 and 3.40

$$\begin{aligned}
&\mathbb{P}\big(\exists x \neq y : Z_x, Z_y \geq -\kappa, |Z_x - Z_y| \leq \eta\big) \\
&\leq 2 \sum_{u \in \eta\mathbb{Z}} \mathbf{1}_{u \in [-\kappa, (\log N)^2]} \Big(\rho([u-\eta, u+\eta]^2) + O\Big(\mathrm{e}^{-(\log N)^{1/5}} + N^{-1/5}\Big)\Big) + o(1) \\
&\leq \mathcal{K}\eta \sum_{u \in \eta\mathbb{Z}} \mathbf{1}_{u \in [-\kappa, (\log N)^2]} \rho([u-\eta, u+\eta]) + \frac{\kappa + (\log N)^2}{\eta} O\Big(\mathrm{e}^{-(\log N)^{1/5}} + N^{-1/5}\Big) + o(1) \\
&\leq \mathcal{K}^2\eta + o(1).
\end{aligned}$$

In the last step, we used (3.88). □

From now on we assume that instead of (3.7), $\mathcal{K}$ satisfies

$$\mathcal{K} \leq (\log \log N)^{1/2}, \tag{3.89}$$

such that $\mathcal{K}^2\eta = o(1)$.

We conclude that under conditions (3.88) and (3.89), with high probability, all points of the process Σ are separated by at least η. Let us recall the definition of $\tilde{\Sigma}$ from the proof of Lemma 3.39. Invoking this result with smaller $\varepsilon > 0$, we conclude the following result.

Lemma 3.48. *With high probability, each interval of the form*

$$\big[Z_x - \eta/4, Z_x + \eta/4\big], \qquad Z_x \geq -\kappa, \tag{3.90}$$

contains exactly one point of $\tilde{\Sigma}$ and one point of Φ. Moreover the complement of the intervals in (3.90) in the region $[-\kappa, \infty)$ contains no point of $\tilde{\Sigma}$ and no point of Φ.

Proof of Theorem 3.10 for $d \leq (\log N)^2$. We work on the intersection of the high-probability events defined in Proposition 3.18, Lemmas 3.47 and 3.48. Let λ be an eigenvalue of $\underline{L}$ satisfying $\tilde{\lambda} := \theta\tau(\lambda - \sigma) \geq -\kappa$. Let x be the unique vertex such that $\tilde{\lambda} \in [Z_x - \eta/4, Z_x + \eta/4]$ and $\mathbf{v} := \mathbf{w}(x)$ defined in Proposition 3.28 and

$\lambda' := \Lambda_x$ (recall the definition of Λ_x from (3.19) and (3.45a)).
Recalling the construction of the orthogonal matrix U in (3.21), we find $\|(\underline{L}-\lambda')\mathbf{v}\| \leq \|E_{\mathcal{W}}\| = O\big(\omega(\log N)^{-4}\big)$. By Lemma 3.48, λ is the only eigenvalue of $\underline{L}$ in the interval $[\lambda'-\Delta, \lambda'+\Delta]$ with $\Delta := \tau\theta\eta/4$. Now since $\tau\theta\eta = O\big((\log N)^2(\log\log N)^{-2}\big)$ we conclude by Lemma A.6

$$\|\mathbf{w}-\mathbf{v}\| = O\Big(\big(\log N\big)^{-3/2}\Big).$$

This concludes the proof. ☐

In order to prove Theorem 3.10 we use (3.87) to show the analog of Lemma 3.48 for $d \geq (\log N)^2$.

Lemma 3.49. *Let Z_x be the variables defined in (3.86). each interval of the form*

$$\big[Z_x - \eta/4, Z_x + \eta/4\big], \qquad Z_x \geq -\kappa,$$

contains exactly one point of Σ_{gumb} and one point of Φ.

Proof of Theorem 3.10 for $d \geq (\log N)^2$. The proof is similar to the regime $d \leq (\log N)^2$ but we do not use the block diagonal approximation of $\underline{L}$. to obtain approximate eigenvectors. Instead, we use the eigenvectors $\mathbf{w}_{x(i)}$ defined in (3.37).
Let λ be an eigenvalue of $\underline{L}$ satisfying $\tilde{\lambda} := \sqrt{2\log N}(v_x - \sqrt{2\log N} - \alpha) \geq -\kappa_{\text{sup}}$.. By Lemma 3.49, the only eigenvalues of $\underline{L}$ in the interval $[\lambda'-\Delta, \lambda'+\Delta]$ with $\Delta = \frac{1}{\sqrt{2\log N}}$ By Lemma A.6, applied for the true eigenvalue λ, the approximate eigenvalue $\sqrt{2\log N}+\alpha+\tilde{\lambda}$, the approximate eigenvector $\mathbf{w}_x$ and $\Delta = \frac{1}{\sqrt{2\log N}}$, we find that if $\mathbf{w}_\lambda$ is the true eigenvector corresponding to λ,

$$\|\mathbf{w}_x - \mathbf{w}_\lambda| = O\Big(\frac{\sqrt{2\log N}}{\log N}\Big) = o(1).$$

In particular since $\mathbf{w}_x|_x = 1 - o(1)$, we conclude that $\mathbf{w}_\lambda$ is localized around x. This concludes the proof. ☐

The proof of Theorem 3.13 is an adaption of the above argument. We do not do it in detail.

3.9 Auxiliary computations

Lemma 3.50. *There exists $\delta > 0$ such that if $1 \leq d \leq N$ and $\mathcal{K} \leq N^\delta$ then $\mathcal{W}^\pm \subseteq \mathcal{U}^\pm$.*

Proof of Lemma 3.50. For $x \in \mathcal{W}^-$ we have $v_x \leq \sqrt{d}(\mathfrak{u}_\gamma - 1) + \delta\frac{\log N}{\sqrt{d}\log d} + 10$. If $d \geq \frac{3}{2}\log N$ then $\sqrt{d}(\mathfrak{u}_\gamma - 1) \leq -\sqrt{\frac{3}{2}\log N}$ and so $v_x \leq \sqrt{(\frac{3}{2}-\delta)\log N}$ which is much smaller than $\sqrt{(1+\varepsilon)\log N}$ if δ is small enough. On the other hand if $d \leq \frac{3}{2}\log N$, then

$$\begin{aligned} v_x &\leq \sqrt{d}(\mathfrak{u}_\gamma - 1) + \delta\frac{\log N}{\sqrt{d}\log d} + 10 \leq -\sqrt{d} + \frac{\log N}{\sqrt{d}}(1+o(1)) + \frac{\log\mathcal{K}}{\log d\sqrt{d}} + 10 \\ &\leq -\sqrt{d} + \frac{1+\delta}{\log d}\frac{\log N}{\sqrt{d}} + 10 \ll -\sqrt{d} + \frac{(1+\varepsilon)\log N}{\log d\sqrt{d}}, \end{aligned}$$

as soon as $\delta > 0$ is small enough.
For $x \in \mathcal{W}^+$, similar. ☐

Proof of Lemma 3.38. The various statements rely on a Taylor expansion in $\Delta t \asymp O(\kappa/\sqrt{d})$. Note that if $t = v_x$ then $\frac{t}{\sqrt{d}} + 1 = \alpha_x$.
The first point is proved by a first order, approximation of the function

$$f(t) = t + \frac{1}{t} + \frac{1}{\sqrt{d}}.$$

around $t_* = \sigma_{\mathrm{sup}}$. (Note that $f(v_x) = v_x + \frac{\alpha_x}{v_x}$.) We compute easily $f'(t) = 1 - \frac{1}{t^2}$ and $f''(t) = -\frac{2}{t^3}$. Since $v_x - t_* = O\big(\frac{\kappa}{\sqrt{d}}\big)$ and $t_* - \kappa \geq t_*/2 \geq \sqrt{\log N}/2$ we conclude that

$$\begin{aligned} v_x + \frac{\alpha_x}{v_x} &= f(\sigma_{\mathrm{sup}}) + \Big(1 - O\big(\frac{1}{\log N}\big)\Big)(v_x - \sigma_{\mathrm{sup}}) + O\Big(\frac{\kappa^2}{d(\log N)^{3/2}}\Big) \\ &= v_x + \frac{1}{\sigma_{\mathrm{sup}}} + \frac{1}{\sqrt{d}} + O\Big(\frac{\kappa}{\log N}\Big). \end{aligned}$$

The proof of (3.78) is similar. Starting from the definition of Λ_x in (3.19) we observe that

$$\Lambda_x = f(v_x) + g(v_x) d\sqrt{\alpha_x}(\beta_x - 1)$$

where

$$f(t) = t + \Big[\frac{1}{t} + \frac{1}{\sqrt{d}}\Big]\Big[1 + \frac{1}{t^2}\Big], \quad g(t) = \frac{\sqrt{td^{-1/2} + 1}}{\sqrt{d}t^2}.$$

Around $t_* = v(\mathfrak{u}_+)$ we expand f to the second order and g to the first order (i.e.å constant). We find $f'(t_*) = 1 - \frac{1}{t_*^2} + O(\frac{1}{t_*^3})$, $f''(t) = O(\frac{1}{t_*^3})$ and

$$g'(t) = O\Big(\frac{1}{dv(\mathfrak{u}_+)^2\sqrt{\mathfrak{u}_+}} + \frac{\sqrt{\mathfrak{u}_+}}{v(\mathfrak{u}_+)^3\sqrt{d}}\Big) = O\Big(\frac{1}{\sqrt{d\mathfrak{u}_+}v(\mathfrak{u}_+)^2}\Big), \quad |t - v(\mathfrak{u}_+)| = O\big(\kappa d^{-1/2}\big).$$

We conclude that

$$\begin{aligned} \Lambda_x &= f(v(\mathfrak{u}_+)) + \Big(1 - \frac{1}{v(\mathfrak{u}_+)^2}\Big)\big(v_x - v(\mathfrak{u}_+)\big) + g(v(\mathfrak{u}_+))d\sqrt{\alpha_x}(\beta_x - 1) + O\Big(\frac{\kappa^2}{d}\|f''\| + \frac{\kappa}{\sqrt{d}}\|g'\|\Big) \\ &= \sigma_{\mathrm{crit},+} + \theta_{\mathrm{crit},+}(v_x - v(\mathfrak{u}_+)) + \frac{1}{\tau_{\mathrm{crit},+}} d\sqrt{\alpha_x}(\beta_x - 1) + O\Big(\frac{\kappa}{\sqrt{d}v(\mathfrak{u}_+)^3}\Big). \end{aligned}$$

For the last estimate we used the assertion from Proposition 3.18 ((iii)) we can bound $(\beta_x - 1) = O(\frac{\kappa \log\log N}{d^{3/2}})$. (3.79) is proved in the same way.

For (3.80), we introduce

$$\begin{aligned} f(t) &= \frac{1}{d(v_{(1)} - t)}\big[1 + \frac{\frac{t}{\sqrt{d}} + 1}{v_{(1)}(v_{(1)} - t)}\big], \quad g(t) = \frac{\frac{t}{\sqrt{d}} + 1}{\sqrt{d}v_{(1)}(v_{(1)} - t)^2}, \\ f'(t) &= \frac{1}{d(v_{(1)} - t)^2}\big[1 + \frac{\frac{t}{\sqrt{d}} + 1}{v_{(1)}(v_{(1)} - t)}\big] + \frac{1}{d(v_{(1)} - t)}\frac{v_{(1)} + 1}{d^{3/2}v_{(1)}(v_{(1)} - t)^3}, \\ f''(t) &= \frac{2}{d(v_{(1)} - t)^3}\big[1 + \frac{\frac{t}{\sqrt{d}} + 1}{v_{(1)}(v_{(1)} - t)}\big] + \frac{v_{(1)} + 1}{d^{3/2}v_{(1)}(v_{(1)} - t)^3} + \frac{3(v_{(1)} + 1)}{d^{3/2}v_{(1)}(v_{(1)} - t)}, \\ g'(t) &= \frac{\frac{t + v_{(1)}}{\sqrt{d}} + 1}{\sqrt{d}v_{(1)}(v_{(1)} - t)^3}, \quad g''(t) = \frac{\frac{2v_{(1)}}{\sqrt{d}} + 6}{\sqrt{d}v_{(1)}(v_{(1)} - t)^4}. \end{aligned}$$

For $t = v_\gamma + O(\kappa/\sqrt{d})$ we get, using $|v_{(1)}| \geq \frac{1}{4}\sqrt{d}$, and using $(\beta_{z_x} - 1) = O\big(\frac{\kappa \log\log N}{\sqrt{\log N}}\big)$ we get

$$\begin{aligned} f(t) =& f(v_\gamma) + \frac{t - v_\gamma}{d(v_{(1)} - v_\gamma)}\big[1 + \frac{\mathfrak{u}_\gamma}{v_{(1)}(v_{(1)} - v_\gamma)} + \frac{1}{\sqrt{d}(v_{(1)} - v_\gamma)^3} + \frac{1}{\sqrt{d}(v_{(1)} - v_\gamma)^4}\big] \\ &+ \frac{(t - v_\gamma)^2}{d^2(v_{(1)} - v_\gamma)^3} + O\big(\frac{\kappa^2}{d^{5/2}}\big), \\ g(t)(\beta_{z_x} - 1) =& g(v_\gamma)(\beta_{z_x} - 1) + O\big(\frac{\kappa^2 \log(d)}{d^{9/4}(v_{(1)} - v_\gamma)^3} + \frac{\kappa^2}{d^{5/2}}\big) \end{aligned}$$

We conclude by observing that

$$g(v_\gamma) + \tau_\gamma = O(\frac{1}{d}\tau_\gamma)$$

by replacing $\frac{1}{v_{(1)}}$ by $-\frac{1}{d}$. □

The rest of this section is a remainder of [7, Appendix B]. For $k \in \mathbb{N}$ and $\Phi := \sum_{x\in[N]} \delta_{Z_x}$ a point process on $\mathbb{R}$ and

$$q_{\Phi,k}(F) := \sum_{x_1,\dots,x_k\in[N]} \mathbb{P}((Z_{x_1},\dots,Z_{x_n}) \in F) = N(N-1)\cdots(N-k+1)\mathbb{P}((Z_1,\dots,Z_k)\in F). \tag{3.91}$$

We have the following inclusion-exclusion principle.

Lemma 3.51. *For any $n, m \in \mathbb{N}^*$, $k_1,\dots,k_n \in \mathbb{N}$, and disjoint measurable $I_1,\dots,I_n \subseteq \mathbb{R}$, we have*

$$\begin{aligned}\mathbb{P}(\Phi(I_1) = k_1,\dots\Phi(I_n) = k_n) = &\frac{1}{k_1!\cdots k_n!}\sum_{l_1,\dots,l_n\in\mathbb{N}} \mathbf{1}_{\sum_i l_i \le m} \frac{(-1)^{\sum_i l_i}}{l_1!\cdots l_n!} q_\Phi(I_1^{k_1\times l_1}\times\cdots\times I_n^{k_n\times l_m})\\ &+O\Big(\frac{1}{k_1!\cdots k_n!}\sum_{l_1,\dots l_n\in\mathbb{N}} \mathbf{1}_{\sum_i l_i = m+1}\frac{1}{l_1!\cdots l_n!} q_\Phi(I_1^{k_1\times l_1}\times\cdots\times I_n^{k_n\times l_m})\Big)\end{aligned}$$

Lemma 3.52 (Closeness between Poisson processes)**.** *Let $\Psi, \tilde\Psi$ be two Poisson processes with intensity ρ and $\tilde\rho$ respectively. Then if $\tilde\rho(E_t) = \rho(E_s)(1+o(1))$ uniformly in $t \ge -\kappa$ we have, for any fix $n\in\mathbb{N}^*$,*

$$\mathbb{P}\Big(\bigcap_{i\in[n]}\{\tilde\Psi(E_{t_i}) \le k_i\}\Big) = \mathbb{P}\Big(\bigcap_{i\in[n]}\{\Psi(E_{s_i}) \le k_i\}\Big) + o(1)$$

uniformly for $k_1,\dots,k_n \in \mathbb{N}^*$ and $s_1,t_1,\dots,s_n,t_n \ge -\kappa$ satisfying $|t_i - s_i| \le |s-t|$

Proof. It suffices to prove

$$\mathbb{P}\Big(\bigcap_{i\in[n]}\{\tilde\Psi(I_i) = k_i\}\Big) = \mathbb{P}\Big(\bigcap_{i\in[n]}\{\Psi(I_i) = k_i\}\Big) + o(1)$$

for I_i disjoint intervals. Now because $\frac{x^k}{k!} \le \mathrm{e}^x$ for $x \ge 0$, we can do a first order approximation

$$\mathrm{e}^{\tilde\rho(E_t)}\frac{\tilde\rho(E_t)^k}{k!} = \mathrm{e}^{\rho(E_s)}\frac{\rho(E_s)^k}{k!} + o(\rho(E_s))$$

and this gives

$$\mathbb{P}\Big(\bigcap_{i\in[n]}\{\tilde\Psi(I_i) = k_i\}\Big) - \mathbb{P}\Big(\bigcap_{i\in[n]}\{\Psi(I_i) = k_i\}\Big) = 2^n o(1)\Big(\prod_{i\in[n]}\mathrm{e}^{\rho(E_{s_i})}\frac{\rho(E_{s_i})^{k_i}}{k_i!}\Big) = o(1).$$

□

Chapter 4

Attached trees and spectral gap

For $d \geq \log N$, the idea of mapping the smallest degree vertices with the smallest eigenvalues works. However, as $d \leq \log N$ this is no longer the case as new *minimal shapes* appear. Smallest eigenvalues are no longer in bijection with smallest degrees (which at this point is 1 in $\mathbb{G}_{cc}$) but rather with maximal trees that are connected to $\mathbb{G}_{cc}$ by exactly one edge.

4.1 Main results

Let $\mathbb{T}_t$, $t \in \mathbb{N}^*$, be the set of all trees on t vertices. By convention, we label the vertices of any tree in $T \in \mathbb{T}_t$ by $V(T) = \{1, \dots, t\}$. The set $\mathbb{T}_t$ is thus a finite set of cardinality t^{t-2} by Cayley's theorem. For instance, the set $\mathbb{T}_1$ consists of the trivial tree on one vertex. Let $\mathbb{T} = \bigcup_{s \geq 0} \mathbb{T}_s$ and $\mathbb{T}_{\leq t} := \bigcup_{1 \leq s \leq t} \mathbb{T}_s$.
We define the line of length $t \in \mathbb{N}^*$ as the graph $L_t = ([t], E)$ with $E = \{(i, i+1) : i \in [t-1]\}$.The line is a tree of size t and its Laplacian matrix is denoted by $L(L_t)$. The spectrum of $L(L_t)$ is real and lies in the interval $[0, \infty)$ with exactly one zero eigenvalue. We define the deterministic quantity

$$\lambda_*(t) := \inf \Big\{ z \in \mathbb{R} : 1 = \Big(\frac{1}{z - L(L_t)} \Big)_{11} \Big\}, \quad t \geq 1, \tag{4.1}$$

as the smallest eigenvalue of the matrix $L(L_t) + \mathbf{1}_1 \mathbf{1}_1^*$ (see Section 4.4). Note that, by a symmetry argument, we could add $\mathbf{1}_t \mathbf{1}_t^*$.

A well-known fact about the Erdős-Rényi graph is that isolated vertices appear when the density d becomes smaller than $\log N$. A natural question is to ask for what regimes connected components and trees of size $t \in \mathbb{N}^*$ appear. This question leads to the following definition

$$d^*(t) := \sup \Big\{ d \geq 0 \colon \frac{1}{t} \big(\log N + (t+1) \log t \big) \geq d - \log d \Big\}, \quad t \in \mathbb{N}^*. \tag{4.2}$$

The next theorem says that the spectral gap of the Laplacian is given by the spectral gap of lines. As the regime d decreases, larger trees appear $\mathbb{G}$ and the spectral gap closes. The different regimes covered by Theorem 4.1 are illustrated in Figure 4.1.

Theorem 4.1. *Let $\varepsilon > 0$ and $t \in \mathbb{N}^*$ and λ_* as defined in (4.1). If $t \geq 2$ and*

$$(1+\varepsilon) d^*(t+1) \leq d \leq (1-\varepsilon) d^*(t), \tag{4.3}$$

and λ_2 denotes the smallest eigenvalue of $\underline{L}$ different from $-\sqrt{d}$, then

$$\lambda_2 = -\sqrt{d} + \frac{\lambda_*(t)}{\sqrt{d}} + o(d^{-1}), \tag{4.4}$$

holds with high probability.

Remark 4.2 (Conventions I)**.** We call a constant universal if it depends only on $\mathbb{T}$. Likewise we call a function or a matrix universal if it depends only $\mathbb{T}$. For instance $\gamma_*(t)$ is a universal constant as it can be computed from the information contained in $\bigcup_{1\le s\le t} \mathbb{T}_s \times [s]$.
For instance, for a fixed tree of finite size $T \in \mathbb{T}$ and for $x \in V(T)$, the *spectral gap* of T is defined as the second smallest eigenvalue of $L(T)$, $\lambda_2(T)$. This is a universal constant. The function $\left(\left(L(T) - z\right)^{-1}\right)_{zz}$ is a universal function. On the other hand $\underline{L}(T)$ is *not* universal, since it depends on d which depends on N.

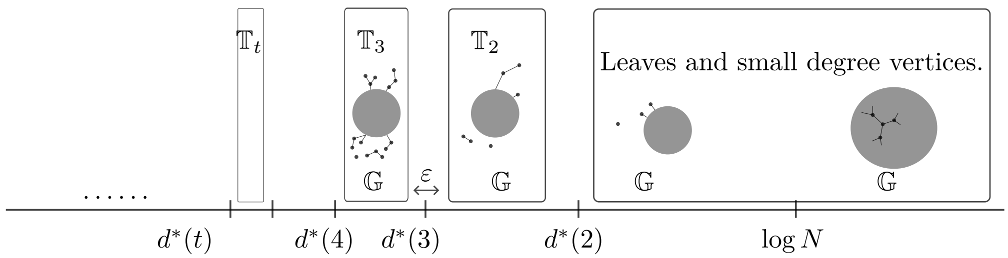


Figure 4.1: Illustration of the different regimes covered by Theorem 4.1.

Remark 4.3 (Localization of the eigenvectors)**.** In the proof of Theorem 4.1, we identify the regions of the graphs that generate the smallest eigenvalue and the proof shares a lot of points with the those of Theorems 3.8 and 3.12 in Chapter 3. We believe the result could be extended into a localization result of the eigenvectors corresponding to the smallest eigenvalues of $\underline{L}$, in an analog to Theorems 3.10 and 3.13. We do not pursue this here.

Remark 4.4 (Conventions II)**.** Throughout this chapter the following conventions hold.

1. Any quantity depends implicitly on N unless mentioned otherwise.

2. For a square matrix M we write $|M|$ the dimension of M and we denote by $\lambda_i(M)$, $i = 1, \dots, |M|$, the eigenvalues of M ordered increasingly

3. For a graph G, we define its Laplacian matrix and its rescaled Laplacian matrix as

$$L(G) := D(G) - A(G), \qquad \underline{L}(G) := \frac{L(G) - d}{\sqrt{d}}. \tag{4.5}$$

4. For $n \in \mathbb{N}^*$ with $n \le N$, any n-by-n matrix can be seen as an N-by-N matrix by embedding.

4.2 Another perspective on leaves

Consider the expression for an approximate eigenvalue generated by a leaf ((3.27) in Definition 3.20) and the way this formula is derived in Proposition 3.35 and Lemma 3.38.

Let us consider the problem of finding the smallest eigenvalue for $\underline{L}|_{B_r(x)}$, where $\underline{L}$ is the Laplacian matrix of an Erdős-Rényi graph with parameter $d \ge (\frac{1}{2} + \varepsilon) \log N$, for some $\varepsilon > 0$ and $x \in [N]$ has degree one and a fairly regular neighborhood $B_r(x)$, in the sense of Proposition 5.19. Consider the trivial graph T given by the $V(T) = \{x\}$ and $E(T) = \emptyset$ and the vertex set $B := B_r(x) \setminus \{x\}$ Now consider the normalized Laplacian matrix of T, L_T and the matrix $L_B := L|_{B_r(x)\setminus\{x\}}$.
By simple perturbation analysis, assuming $\|A\| \le 3$ as in Proposition 3.25 and assuming

$$\max_{y\in B_r(x),\, y\ne x} |v_y| = O(d^\alpha)$$

for some $\alpha > 0$, we get $\|L_B\| = O(d^\alpha)$.
Consider the vector $\mathbf{u} := (\mathbf{1}_x - \mathbf{1}_{z_x})$. Then $\underline{L}|_{B_r(x)}$ is a rank one perturbation of $L_B + L_T$ since

$$\underline{L}|_{B_r(x)} = L_B + L_T + \frac{1}{\sqrt{d}}\mathbf{u}\mathbf{u}^*.$$

Moreover, by interlacing of eigenvalues and basic perturbation theory, we know that, since $\operatorname{Spec} L_T = \{-\sqrt{d}\}$ and $\operatorname{Spec} L_B \subseteq [-O(d^\alpha) + O(d^\alpha)]$, the matrix $\underline{L}|_{B_r(x)}$ has precisely one eigenvalue in $[-\sqrt{d}, -O(d^\alpha)]$ which is satisfies $-\sqrt{d} \leq \lambda \leq -\sqrt{d} + 2d^{-1/2}$. Moreover by Lemma A.8, we have the following equality

$$\sqrt{d} = G_{L_T}(\lambda)_{xx} + G_{L_B}(\lambda)_{z_x z_x}$$

where $G_H(z) := (z - H)^{-1}$. We can use the resolvent formula from Lemma A.7 and the fact that $\min_{y\neq x} |\lambda - v_y| \geq cd^{-1/2}$ for some $c > 0$ small enough to rewrite the above expression as

$$\sqrt{d} = \frac{1}{\lambda + \sqrt{d}} + \frac{1}{\lambda - v_{z_x}} + \frac{1}{d(\lambda - v_{z_x})^2} \sum_{y\in S_1(z),\, y\neq x} \frac{1}{\lambda - v_y} + O\Big(\frac{1}{d^{5/2}}\Big),$$

for some $\lambda \in [-\sqrt{d}, -\sqrt{d} + \frac{2}{\sqrt{d}}]$. In particular, bounding the last three terms by $O(d^{-1/2})$ and inverting the equation we see that

$$\lambda = -\sqrt{d} + \frac{1}{\sqrt{d}} + O(d^{-3/2}).$$

Let us set $\Lambda_{z_x}(\lambda) := \frac{1}{\lambda - v_{z_x}} + \frac{1}{d(\lambda - v_{z_x})^2}\sum_{y\in S_1(z),\, y\neq x} \frac{1}{\lambda - v_y}$ and $t := d^{3/2}(-\sqrt{d} + \frac{1}{\sqrt{d}} - \lambda)$, we find

$$\begin{aligned}\Lambda_{z_x}(\lambda) &= \Lambda_{z_x}\big(-\sqrt{d} + d^{-1/2} + td^{-3/2}\big)\\ &= \frac{1}{v_1 - v_{z_x}} + \frac{1}{d(v_1 - v_{z_x})^2}\sum_y \frac{1}{v_1 - v_y} + O\Big(\frac{t}{d^{5/2}}\Big).\end{aligned}$$

We conclude that

$$\lambda = -\sqrt{d} + \frac{1}{\sqrt{d}} + \frac{1}{d}\Lambda_{z_x}(0) + O\Big(\frac{t^2}{d^{7/2}}\Big)$$

Since $t = O(1)$, we can now proceed as in the proof of Proposition 3.35 to transform the above expression $\Lambda_x^{\mathcal{L}}$ defined in (3.27).

In this chapter, we will show how this argument can be applied to cases where T is a tree of any finite size. We do not prove convergence of the eigenvalue process towards a Poisson Point Process, but the proof could be extended with additional work. This chapter should be thought of as a proof of concept.

4.3 Burned graph

In this section, we study how trees in $\mathbb{G}_{cc}$ (the macroscopic connected component of $\mathbb{G}$) contribute to the spectrum of $\underline{L}$ in the interval $\big[-\sqrt{d}, -\sqrt{d}+d^{-1/2}\big]$. We first give a more precise meaning to what we consider to be a *tree in* $\mathbb{G}$. Indeed any pair $x, y \in [N]$ of neighbors is technically a tree of size two if we restrict the graph $\mathbb{G}$ to the two vertices $\{x, y\}$. However we are interested by configurations that *look like trees* even when embedded in $\mathbb{G}$.

Definition 4.5 (Embedded trees of $\mathbb{G}_{cc}$)**.** Let us fix $\zeta \in \mathbb{N}$ and consider the graph $\mathbb{G}$ restricted to the vertices of degree smaller than ζ,

$$\mathbb{G}^\zeta := \Big(\overline{\mathcal{U}(\zeta)}, \Big\{(x,y) \in E : x, y \in \overline{\mathcal{U}(\zeta)}\Big\}\Big), \quad \overline{\mathcal{U}(\zeta)} := \{x \in [N] : D_x \leq \zeta\}.$$

We define $\mathcal{L} := \mathcal{L}(\mathbb{G},\zeta)$ to be the connected components of size 1. For $t \in \mathbb{N}^*$, $t > 1$, we define $\mathcal{T}_t := \mathcal{T}_t(\mathbb{G},\zeta)$ to be the connected components of $\mathbb{G}^\zeta$ of size t.
We define $\mathcal{T}_{\leq t} := \bigcup_{1<s\leq t} \mathcal{T}_s$. For $U \in \mathbb{T}$, we define

$$\mathcal{T}^{(U)}(\mathbb{G},\zeta) := \{T \in \mathcal{T}(\zeta) \colon \mathbb{G}|_T = U\}. \tag{4.6}$$

We call the parameter $\zeta > 0$ the *cut*. It is usually fixed for the whole graph given some density d. Of course we usually set the cut to be larger than the size of the trees we are studying. We exclude the trees on one vertex on purpose since those are studied separately as small degree vertices. Note that a consequence of Definition 4.5 is that there is no inclusion between trees: if $Y, Y' \in \mathcal{T}_{\leq t}(\zeta)$ and $Y \neq Y'$, then $Y \cap Y' = \emptyset$. An illustration of Definition 4.5 is shown in Figure 4.2.

Definition 4.6 (Tree configurations)**.** For $\ell \in \mathbb{N}$ and $\zeta, t \in \mathbb{N}^*$, we define $\mathcal{T}_t(\zeta,\ell)$ to be the subset of $\mathcal{T}_t(\zeta)$ consisting of all trees $Y \in \mathcal{T}_t(\zeta)$ with precisely ℓ edges connecting Y to $\mathbb{G} \setminus Y$,

$$|E(\mathbb{G} \setminus Y, Y)| = \sum_{x \in Y} \sum_{y \in [N] \setminus Y} A_{xy} = \ell.$$

We call the elements of the set $S_1(Y)$ the *anchors* of Y and the edges $E(Y, S_1(Y))$ the *links*.
The *configuration* of the tree $Y \in \mathcal{T}$ is the tuple $\mathrm{config}^{\mathbb{G}}(Y) := (T, A)$ where $T \in \mathbb{T}$, $A \subseteq [|T|]$, and A is the image of the anchors of Y under the isomorphism $T \sim Y$. In particular $A = \emptyset$ if Y has no anchors.
We define $\mathcal{T}_{\leq t}(\ell)$, $\mathcal{T}_\ell(\leq \ell)$ and $\mathcal{T}_{\leq t}(\leq \ell)$ similarly, by the obvious adaptations.

Note that for $\zeta \geq t-1$, $\mathcal{T}_t(\zeta, 0)$ is the set of all isolated connected components of size t with no cycles.

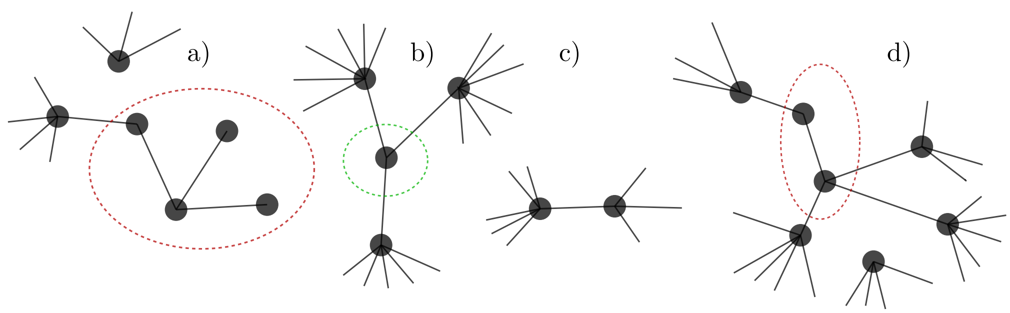


Figure 4.2: Illustration of Definition 4.5 with $\zeta = 2$. Only the configuration that are circled belong to $\mathcal{U}(\zeta)$. The red configurations (a and d) are trees and belong to $\mathcal{T}_{\leq 4}(\zeta)$ while the green configuration (b) is a leaf (or a small degree vertex) and belongs to $\mathcal{T}_1(\zeta)$. Although the restriction of $\mathbb{G}$ to the two vertices in c would be a tree of size two, this does not belong to $\mathcal{U}(\zeta)$.

For $\zeta > 0$, we define the collection of subsets

$$\mathcal{U}(\zeta) := \mathcal{T}(\zeta) \cup \mathcal{L}(\zeta), \quad \mathcal{T}(\zeta) := \bigcup_{t \geq 2} \bigcup_{Y \in \mathcal{T}_t(\zeta)} Y, \quad \mathcal{L}(\zeta) := \bigcup_{Y \in \mathcal{T}_1(\zeta)} Y. \tag{4.7}$$

By the way, recalling Definition 4.5, we see that $\mathcal{L}(\zeta)$ consists of vertices of small degree, which have no small degree neighbors. We define the subset of all vertices belonging to some element of $\mathcal{U}(\zeta)$ as

$$\overline{\mathcal{U}(\zeta)} := \bigcup_{Y \in \mathcal{U}(\zeta)} \bigcup_{x \in Y} \{x\}.$$

The set $\mathcal{T}(\zeta)$ is the set of all trees and the set $\mathcal{L}(\zeta)$ is the set of vertices of small degree that are not trees (i.e. vertices of degree at least 2).

For $\tau \geq 0$, we introduce

$$\mathcal{V}_\tau(\zeta) := \{x \in [N] \colon \; v_x \leq (\tau - 1)\sqrt{d}, \, x \notin \overline{\mathcal{U}(\zeta)}\}. \tag{4.8}$$

Burne The set $\mathcal{V}_\tau$ consist of vertices with small degrees, that is $x \in [N]$ that satisfy $D_x \leq \tau\sqrt{d}$. We remove vertices that belong to some element of $\mathcal{U}(\zeta)$. We usually drop the ζ and simply write $\mathcal{V}_\zeta$.

For $e = (x, y) \in E(\mathbb{G})$ we define the rank-one matrix

$$B(e) := \frac{1}{\sqrt{d}}(\mathbf{1}_x - \mathbf{1}_y)(\mathbf{1}_x - \mathbf{1}_y)^*.$$

We call the subtraction of $B(e)$ from $\underline{L}(\mathbb{G})$ the *burning-off* of edge (xy).

Lemma 4.7 (No cycles around $\mathcal{U}(\zeta)$)**.** *Let $t_* \in \mathbb{N}^*$, d as in (4.2),*

$$10 + t \leq r \leq \log\log N, \qquad \max(t, 10) \leq \zeta \leq 2t. \tag{4.9}$$

There exists $\eta > 0$ such that with probability $1 - O(N^{-\eta})$, the balls $(B_r(U) \colon \, U \in \mathcal{U}(\zeta))$ are trees.

Proof. By Lemma 5.6, we find that

$$\begin{aligned}\mathbb{P}\big[\exists U \in \mathcal{U}(\zeta), B_r(U) \text{ contains a cycle}\big] &\leq \sum_{Y \subseteq [N], \, |Y| = |U|} \mathbb{P}\big[B_r(U) \text{contains a cycle} | S_1(Y)\big] \mathbb{P}[\mathbb{G}|_Y = U] \\ &\leq \frac{1}{N}\big(C(d + |U|\zeta)\big)^{2r+1}(2r)^2\Big(\frac{d}{N}\Big)^{|U|}\big(N\mathrm{e}^{-d}d^\zeta\big)^{|U|-1} \\ &\leq (4Crd)^{2r+1}\mathrm{e}^{-d|U|}d^{\zeta+|U|} \leq (4Crd)^{2r+\zeta+|U|+1}\mathrm{e}^{-d} = O\big(N^{-\eta}\big)\end{aligned}$$

where, using $d \geq \frac{1+\varepsilon}{t_*+1}\log N$ and $r, \zeta, |U| = O(1)$, we chose $\eta \leq \frac{1}{t_*+1}$ in the last equality. □

We need the following observation which ensures that no vertex has too many neighbors in $\mathcal{V}_\tau$.

Lemma 4.8. *Let $t \in \mathbb{N}^*$, $\varepsilon > 0$ and $d \geq (1 + \varepsilon)d(t)$. Then for $k > t$, there exists $\eta > 0$ such that*

$$\mathbb{P}\Big[\inf_{x \in [N], D_x \geq k} \big|S_1(x) \setminus \mathcal{V}_\tau\big| \leq 1\Big] = O\big(N^{-\eta}\big).$$

Proof. By Lemma B.5 $\mathbb{P}[x \in \mathcal{V}_\tau] \leq 2\exp\big(-d(1 + \tau\log\tau)\big)$ we see that for $k \geq 0$

$$\mathbb{P}\big[\exists x \in [N] \, \big|S_1(x) \cap \mathcal{V}_\tau\big| \geq k\big] \leq 2Nd^k\mathrm{e}^{-kd(1+\tau\log\tau)}Nd^k\mathrm{e}^{-(1+\varepsilon)(1+\tau\log\tau)} = O\big(N^{-\eta}\big),$$

where we chose $k > t$ and $\tau, \eta > 0$ small enough. □

Let us recall that for every $t_* \in \mathbb{N}^*$, if d satisfies (4.3), then

$$\mathbb{P}\big[\max_{x \in [N]} D_x \geq C\sqrt{d}\big] \leq N^{-2}, \tag{4.10}$$

for some constant $C = C(t_*) > 0$.

Proposition 4.9 (Burned graph)**.** *Let $t_* \in \mathbb{N}^*$, d as in (4.2),*

$$10 + t \leq r \leq \log\log N, \qquad \max(t, 10) \leq \zeta \leq 2\max(t, 10). \tag{4.11}$$

Then there exists $\tau, \eta > 0$ such that with probability $1 - O(N^{-\eta})$, there exists a graph $\mathbb{G}^{\tau,\zeta}$, $V \subset [N]$ and $E \subset E(\mathbb{G})$, such that satisfies the following properties

(i) The graphs $\mathbb{G}$ and $\mathbb{G}^{\tau,\zeta}$ have the same number of connected components and differ only by E meaning that

$$\underline{L}(\mathbb{G}) = \underline{L}(\mathbb{G}^{\tau,\zeta}) + \sum_{e \in E} B(e).$$

(ii) *For every $Y, U \in \mathcal{U}(\zeta)$, $Y \neq U$, either $d^{\tau,\zeta}(U,Y) = 2$ or $d^{\tau,\zeta}(U.Y) \geq r$.*

(iii) *The set V characterises the overlap of $B^{\tau,\zeta}_{r/2}(T)$ with $\overline{\mathcal{U}(\zeta)} \cup \mathcal{V}_\tau$, in the sense that $V \subseteq \bigcup_{x \in \overline{\mathcal{U}(\zeta)}} S^{\tau,\zeta}_1(x)$ and*

$$\begin{aligned} B^{\tau,\zeta}_r(T) \cap \mathcal{V}_\tau \neq \emptyset \quad &\Leftrightarrow \quad \exists x \in V \cap \mathcal{V}_\tau, \text{ such that } S_1(T) = \{x\}, \quad T \in \mathcal{T}(\zeta), \\ B^{\tau,\zeta}_{r/2}(T) \cap B^{\tau,\zeta}_{r/2}(T') \neq \emptyset \quad &\Leftrightarrow \quad \exists x \in V \text{ such that } S_1(T) = S_1(T') = \{x\}, \quad T, T' \in \mathcal{T}(\zeta, \mathbb{G}). \end{aligned}$$

(iv) *If a subset of vertices $Y \subset [N]$ becomes a tree (in the sense of Definition 4.5) in $\mathbb{G}^{\tau,\zeta}$, either Y is a vertex of degree 1 or there exists $U \in \mathcal{T}(\zeta)$, $|U| < t-2$ such that Y is obtained from extending U by $y \in S_1(U)$.*

(v) *The set V cannot intersect too many large trees, meaning that*

$$\sup_{v \in V} \left|\overline{\mathcal{U}(\zeta)} \cap B^{\tau,\zeta}_r(v)\right| \leq t_*, \quad \sup_{T \in \mathcal{T}_u} \left|B_r(T) \cap E\right| \leq t_* - u \quad \sup_{T \in \mathcal{T}_u} \left|V \cap B_r(T)\right| \leq t_* - u, \quad 1 \leq u \leq t_*.$$

(vi) *In particular, for every $T \in \mathcal{T}_{t_*}$, $B_r(T) \cap E = B_r(T) \cap V = \emptyset$.*

Proof. We work on the event defined by Lemma 4.7 with r as $r + 2t_*$.
We first remove those edges that connect elements of $\mathcal{U}(\zeta)$ to $\mathcal{V}_\tau$. Let $Y \in \mathcal{U}(\zeta)$. For every $x \in \mathcal{V}_\tau \cap B_r(Y)$ we consider the unique path of length at most r such that

$$P := \big\{(x_i, x_{i+1})\colon\ i \in [k],\, x_0 = x,\, x_k \in T,\, (x_i, x_{i+1}) \in E(\mathbb{G})\big\}.$$

If $|P| = 1$, and removing the edge (x_0, x_1) creates a new connected component, we add x to V. Otherwise we add the edge (x_0, x_1) to E.
We now make the elements of $\mathcal{U}(\zeta)$ distant whenever possible. For every $Y, U \in \mathcal{U}(\zeta)$, $U \neq Y$, such that $U \cap B_r(Y) \neq \emptyset$, we consider the unique path of length smaller than r that links the two subsets,

$$P := \big\{(x_i, x_{i+1})\colon\ i \in [k],\, x_0 \in U,\, x_k \in T,\, (x_i, x_{i+1}) \in E(\mathbb{G})\big\}.$$

By Definition 4.5, $|P| \geq 2$, since otherwise U and Y would be equal. There are three cases. First if $|P| > 2$, we can add the edge (x_1, x_2) to E. Second if $|P| = 2$ and removing either (x_0, x_1) or (x_1, x_2) creates a new connected component, we add x_1 to V. Third if $|P| = 2$ and it is possible to remove either (x_0, x_1) or (x_1, x_2) without creating a new connected component, we add that edge to E. Now (i) - (iv) follows from the construction of V and E once (v) and (vi) are established.

Up to now the construction is purely deterministic and does not involve probabilistic estimates. The following formula allows us to measure how likely it is to see high concentrations of elements of $\mathcal{U}(\zeta)$ in a small neighborhood in $\mathbb{G}$. Let $T \in \mathcal{T}_t$, $t \in \mathbb{N}^*$, $l, k \in \mathbb{N}$, $t_1, \dots, t_l \in \mathbb{N}^*$. Let us define the event

$$\Omega := \Big\{\exists T, T_j \in \mathcal{U}(\zeta),\, |T_j| = t_j\, d(T, T_j) \leq r,\, \forall j \in [l] \text{ and } x_i \in \mathcal{V}_\tau(\zeta),\, d(T, x_i), \leq r,\, \forall i \in [k]\Big\}.$$

For $Y \subset [N]$ and $T \in \mathbb{T}$, we define the event

$$\Omega(Y, T) := \big\{\mathbb{G}|_Y = T, \quad \max_{y \in Y} D_y \leq \zeta\big\}.$$

Then we can see that

$$\Omega = \bigcup_{Y, \mathbf{Y}, \mathbf{T}, \mathbf{x}, \mathbf{z}} \Big\{\Omega(Y_j, T_j),\, \Omega(Y, T),\, X_i \in \mathcal{V}_\tau(\zeta),\, Y_j \in \mathbf{Y},\, T_j \in \mathbf{T},\, x_i \in \mathbf{x},\, P \in E(\mathbb{G}), \forall P \in \mathbf{z}\Big\},$$

where the union is taken over all collections of disjoint subsets $\{Y^{(i)}_1, \dots, Y^{(i)}_{t_i}\}$, $i = 0, 1 \dots, l$ (with $t_0 = t$ by convention), collection of trees $\mathbf{T} = (T_0, \dots, T_l)$ with $T_j \in \mathbb{T}_j$, vertices $x_i \in [N]$, $i = 1, \dots, k$ and paths

$P \in \mathbf{z}$ that link Y to Y_i and x_j, $i \in [l]$, $j \in [k]$ respectively.
Then, again using the fact that $d(Y,U) \geq 2$ for any distinct pair of trees $Y, U \in \mathcal{U}(\zeta)$, it is easy to see that

$$\begin{aligned}\mathbb{P}(\Omega) &\leq N d^{(k+l)(r+1)} \prod_{j=1}^{k} \mathbb{P}\left(D_{x_j} \leq \tau d\right) \sum_{T \in \mathbf{T}} \prod_{i=0}^{l} \mathbb{P}\left(\Omega(Y_i, T_i)\right) \\ &\leq N d^{(k+l)(r+1)} \mathrm{e}^{-kd + kd\tau \log \tau} \left(\mathrm{e}^{-d} \frac{\zeta}{d}\right)^{w} \prod_{i=0}^{l} t_i^{t_i - 2} \\ &\leq N \mathrm{e}^{-(w+k)d + kd\tau \log \tau} \zeta^w t_*^w \tau^k (\log N)^{2r(k+l)} \left(1 + O\left(\frac{(\log N)^2}{N}\right)\right)^{w+k},\end{aligned} \tag{4.12}$$

where we wrote $w := t + \sum_{i=1}^{l} t_i$ and used Lemma B.5 in the last step. Using the fact that $d \leq \frac{1-\frac{\varepsilon}{2}}{t_*}$ we find that, for any $r = O(1)$,

$$\mathbb{P}(\Omega) \leq C N^{1 - \frac{(w+k)(1-\varepsilon/2)}{t_*} + kd\tau \log \tau} (\log N)^{3r(k+l)},$$

for some constant $C > 0$. From this estimate and the fact that $\lim_{\tau \to 0} \tau \log \tau = 0$, we deduce that (v) and (vi) hold for τ and η small enough. □

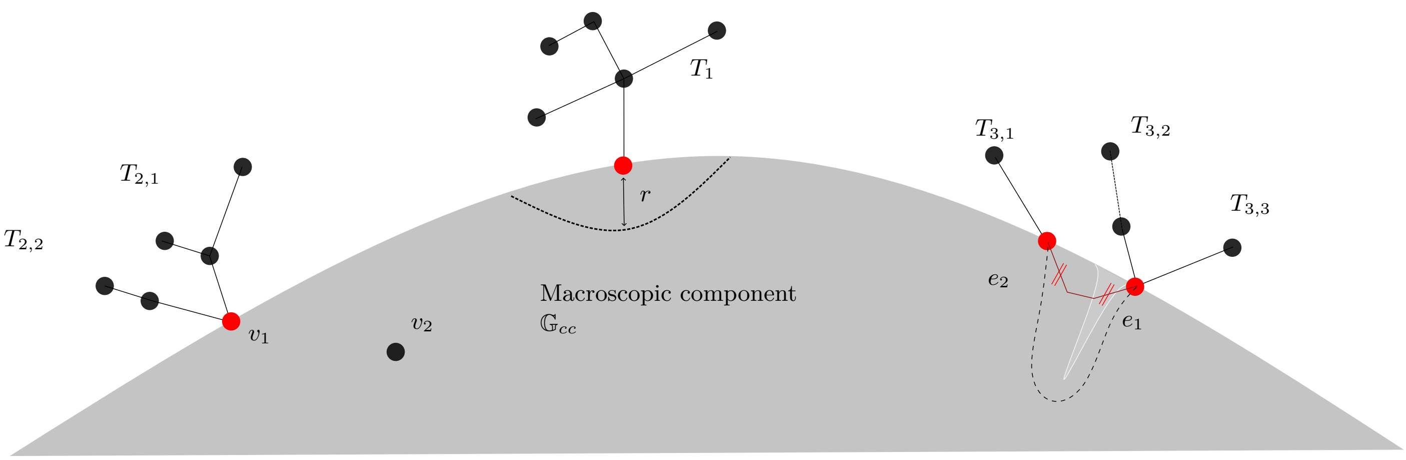


Figure 4.3: Illustration of Proposition 4.9 for $d^*(6) \ll d \leq d^*(5)$. T_1 is a maximal tree of size 5 and belongs to $\mathcal{A}$. The groups $T_{2,\cdot}$ and $T_{3,\cdot}$ constitute bouquets of total size 5 and for 4. The group $T_{3,\cdot}$ can be split into $T_{3,1}$ and $T_{3,2}$, $T_{3,3}$ by cutting either e_1 or e_2. Without loss of generality we cut e_1 and store it in E. In the burned graph $\mathbb{G}^{\tau,\zeta}$, $T_{3,1}$ has a regular neighborhood and thus $T_{3,1} \in \mathcal{C}$. On the other hand, we cannot slit the groups $T_{2,\cdot}$ and $T_{3,2}$ and $T_{3,3}$ because doing so would require to disconnect the graph $\mathbb{G}$ (or rather create a new connected component). Those four trees belong to $\mathcal{B}$. The vertex v_2 represents a vertex with a small degree ($D_x \leq \zeta$) but which is not *per se* a tree. $v_2 \in \mathcal{L}(\zeta)$.

We partition $\mathcal{T}^{\tau,\zeta}(\zeta)$ in the three classes

$$\begin{aligned}\mathcal{A} &:= \{U \in \mathcal{T}(\zeta) \colon\ B_r(U) = B_r^{\tau,\zeta}(U),\ S_1(U) \cap V = \emptyset\}, \\ \mathcal{B} &:= \{U \in \mathcal{T}(\zeta) \colon\ B_r(U) = B_r^{\tau,\zeta}(U),\ \exists v \in V,\ S_1(U) = \{v\}\}, \\ \mathcal{C} &:= \mathcal{T}(\zeta) \setminus \left(\mathcal{A} \cup \mathcal{B}\right), \\ \mathcal{D} &:= \mathcal{T}^{\tau,\zeta}(\zeta) \setminus \mathcal{T}(\zeta).\end{aligned} \tag{4.13}$$

In words, the set $\mathcal{A}$ consists of all trees whose neighborhood is the same in $\mathbb{G}$ and in $\mathbb{G}^{\tau,\zeta}$. Those trees have no *problems* in their vicinity, i.e. $B_r(T)$ contains no element of $\mathcal{V}_\tau$ and no other trees. $\mathcal{B}$ consist of trees $T \in \mathcal{T}(\zeta)$, which have a unique anchor in $\mathbb{G}^{\tau,\zeta}$ which in addition is an element of $\mathcal{V}_\tau$ or is the unique anchor $\mathbb{G}^{\tau,\zeta}$ of another tree $T' \in \mathcal{T}(\zeta)$. Note that in both cases, the anchor of $T \in \mathcal{B}$ lies in the set V. $\mathcal{C}$ consists of trees whose neighborhood might have been altered (they do not belong to $\mathcal{A}$) but which do not have any other tree in their neighborhood in $\mathbb{G}^{\tau,\zeta}$. Finally $\mathcal{D}$ gathers the trees that were created as a result of the burning procedure.

Lemma 4.10. *Under the hypotheses of Proposition 4.9, for any given tree $T \in \mathbb{T}$, most of its representatives in $\mathbb{G}$ did not see their neighborhood changed. There exists $C > 0$ such that*

$$\left|\mathcal{T}^{(T)}(\mathbb{G},\zeta) \cap \big(\mathcal{B} \cup \mathcal{C}\big)\right| \le CN^{-\eta}\left|\mathcal{T}^{(T)}(\mathbb{G},\zeta) \cap \mathcal{A}\right|, \quad \forall T \in \mathbb{T},$$

where $\mathcal{A} := \{U \in \mathcal{T}(\zeta) \colon\ B_r(U) = B_r^{\tau,\zeta}(U)\}$ and $\mathcal{B} := \mathcal{T}(\zeta) \setminus \mathcal{A}$.

Proof. Let us fix $U \in \mathbb{T}$ and abbreviate $A := \left|\mathcal{T}^{(U)}(\mathbb{G},\zeta) \cap \mathcal{A}\right|$ and $B := \left|\mathcal{T}^{(U)}(\mathbb{G},\zeta) \cap \big(\mathcal{B} \cup \mathcal{C}\big)\right|$. Let us recall the event Ω introduced in the proof of Proposition 4.9. Then using the estimate (4.12) as well as Lemma 4.24 we find

$$\mathbb{E}|B| \le \mathbb{P}\Big[v_x \le (\tau-1)\sqrt{d}\Big]^n \mathbb{E}|A|, \quad \mathbb{E}|B|^2 = \mathbb{E}|B|\Big(1 + O\Big(\frac{dt^2}{N}\Big)\Big), \quad \mathbb{E}|A|^2 = \mathbb{E}|A|\Big(1 + O\Big(\frac{dt^2}{N}\Big)\Big)$$

We can conclude using a second-moment argument that with probability $1 - O(u^{-2})$

$$|A| = \mathbb{E}|A|\big(1 + O\big(u^{-2}\big)\big), \quad |B| = \mathbb{E}|B|\big(1 + O\big(u^{-2}\big)\big)$$

and thus, $u = N^\eta = O\big(\mathrm{e}^{-d/4}\big)$, for η small enough and choosing $\eta = \frac{1}{4(t+1)}$, $u = N^\eta \le \mathrm{e}^{-d/4}$ and $\tau > 0$ small enough such that $\mathrm{e}^{d(\tau \log(e/\tau)-1)} \le N^{-1/2(t+1)} \le \mathrm{e}^{-d/2}$ and we deduce that $|B| \le N^{-\eta}|A|$, holds with probability $1 - O\big(N^{-\eta}\big)$. This concludes the proof. □

4.4 Finite trees

In this section, we study the spectrum of trees of finite size and in particular, we establish that if $t \in \mathbb{N}^*$, $\lambda_*(t)$, as defined in (4.1), is the minimizer of some natural quantity defined for all finite trees of size smaller or equal to t.

For $T \in \mathbb{T}$ we write $L(T)$ the Laplacian matrix of the tree T. A few general statements are known about the spectra of the matrices $L(T)$, $T \in \mathbb{T}_t$. For instance if λ_i, $i \in [t]$, denote the eigenvalues of T ordered increasingly, it is known that

$$0 = \lambda_1 < \lambda_2 < \ldots < \lambda_t \le (t-1) + 2\sqrt{t}.$$

A quantity of general interest is the *spectral gap* defined as the smallest non-trivial eigenvalue λ_2. We are interested in a similar quantity. Let us define

$$\gamma(T,i) := \inf\Big\{z \in \mathbb{R} : 1 = \Big(\frac{1}{z - L(T)}\Big)_{ii}\Big\}, \quad i \in [|T|]. \tag{4.14}$$

The parameter $\gamma(T,i)$ is a deterministic value that lies in the interval $\big(0, \lambda_2(L(T))\big)$. It can take different values depending on the choice of $i \in [|T|]$ as is shown in Figure 4.4.

For $t \in \mathbb{N}^*$, let us define

$$\gamma_*(t) = \min_{T \in \mathbb{T}_{\le t}} \min_{i \in [|T|]} \gamma(T,i). \tag{4.15}$$

An important result is that γ_* is a strictly decreasing function of $t \in \mathbb{N}^*$ (see Lemma 4.11). The underlying mechanism is that a tree T' of size $t+1$ can be obtained from a tree T of size t by adding a new vertex to it. The spectral properties of the Laplacian matrix of T' can then be understood as a rank-one perturbation of the Laplacian matrix of T and a zero matrix (see Figure 4.4).

The next lemma shows that $\gamma_*(t)$ is minimized by a tree of size t.

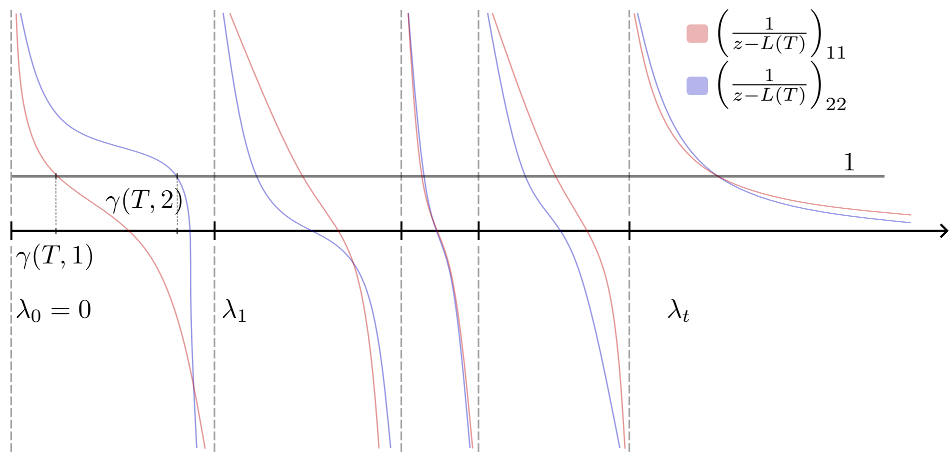


Figure 4.4: Illustration of (4.14). A tree $T \in \mathbb{T}$ might generate different spectral gaps depending on the vertex we use to connect it to the rest of the graph.

Lemma 4.11. *For every $t \in \mathbb{N}^*$, there exists $c_* := c_*(\mathbb{T}_{\leq t}) > 0$ such that*

$$\gamma_*(t+1) \leq \gamma_*(t) - c_*,$$

and for every $T \in \mathbb{T}_{\leq t}$,

$$\max_{i \in [|T|]} \gamma(T,i) \leq \lambda_2(L(T)) - c_*. \tag{4.16}$$

Proof. Let us fix $t \in \mathbb{N}^*$ and $T \in \mathbb{T}_t$. For $i \in [t]$, we define $M(T,i) := L(T) - \mathbf{1}_i\mathbf{1}_i^*$. The matrix $M(T,i)$ is a Hermitian matrix obtained by a rank one perturbation of the Laplacian of T. By Lemma A.3, we have

$$G_{M(T,i)}(z) = G_{L(T)}(z) + G_{L(T)}(z)\mathbf{1}_i\mathbf{1}_i^* G_{M(T,i)}(z).$$

Therefore z is an eigenvalue of $M(T,i)$ if and only if $1 = \big(G_L\big)_{ii}(z)$. We conclude that

$$\gamma(T,i) = \min\{\lambda\colon\ \lambda \in \sigma(M(T,i))\}.$$

If $T \in \mathbb{T}_{\leq T}$, then we can choose a finite constant c_T that satisfies (4.16). Then we can take $c_* = \min_{T \in \mathbb{T}_{\leq T}} c_T$ to get the statement uniform in all trees of size at most t.

Let $T \in \mathbb{T}_{<t}$ and $i \in [|T|]$ be such that $\gamma(T,i) = \gamma_*(t)$. Let us view $L(T)$ as a $(t+1)$-by-$(t+1)$ matrix with the last row and column equal to zero. We can construct a tree T' by adding one neighbor to some vertex of T, for instance 1. Then the matrix $L(T') := L(T) + \mathbf{e}\mathbf{e}^*$, $\mathbf{e} := \mathbf{1}_{t+1} - \mathbf{1}_1$, is the Laplacian matrix of T'. Since $L(T')$ is a rank-one perturbation of $L(T)$, $M(T',i)$ is also a rank-one perturbation of $M(T,i)$. Thus there exists $c_* > 0$ depending on T and i such that

$$\gamma_*(t+1) \leq \min_{\lambda \in \sigma(M(T',i))} \lambda \leq \min_{\lambda \in \sigma(M(T,i))} \lambda - c_* = \gamma_*(t) - c_*,$$

where we used the fact that (T,i) is chosen as the minimizer of $\gamma(T,i)$. This concludes the proof, since in particular □

Note that the proof of Lemma 4.11 gives a way to construct for any $t \in \mathbb{N}^*$, a minimal pair (T,i), in the sense that $\gamma(T,i) = \gamma_*(|T|)$, iteratively in $t-1$ steps.

Let $t \in \mathbb{N}$ and let us define the line on t vertices as the tree

$$L_t := ([t], E), \quad E := \{(i,i+1) : 1 \leq i < t\}.$$

The following lemma says that the tree L_t anchored at one of its extremity is the solution to $\gamma_*(t)$, meaning that $\min_{T,i}\gamma(T,i) = \gamma(L_t, x)$ where the minimum is taken over trees of size t and $x = 1$ or $x = t$.
Before stating the lemma, observe that, for $T \in \mathbb{T}_t$, $i \in [t]$ and $t \in \mathbb{N}^*$, we have

$$\gamma(T,i) = \inf_{u\in\mathbb{R}^t} \frac{\langle u\,, L(T)u\rangle + |u(i)|^2}{\|u\|^2}. \tag{4.17}$$

The first term on the numerator of the right-hand side is just the Dirichlet energy of u on the tree T and is known to be equal to

$$\langle u\,, L(T)u\rangle = \sum_{(x,y)\in E(T)} (u(x)-u(y))^2.$$

Let us denote by $\mathcal{E}(u) := \langle u\,, L(T)u\rangle + |u(i)|^2$

The next lemma shows that $\lambda_*(t) = \gamma_*(t)$.

Lemma 4.12. *Let $t \in \mathbb{N}^*$. The minium of $(\gamma(T,i))$ taken over $T \in \mathbb{T}_{\le t}$ and $i \in [|T|]$ is obtained when T is a line and i an extremal point.*

Proof. We first note that by Lemma 4.11, we can consider $T \in \mathbb{T}_t$. Moreover since

$$\mathcal{E}(u) = \langle u\,, L(T)u\rangle = \sum_{(x,y)\in E} |u(x)-u(y)|^2 \ge \sum_{(x,y)\in E} ||u(x)|-|u(y)||^2,$$

and thus we can assume $u(x) \ge 0$ for every $x \in [t]$.
Let us now change to a dual approach. We will fix $u(i) > 0$ and $(u(x)-u(y)) \in \mathbb{R}$ and thus fix $\mathcal{E}$ and try to maximize $\|u\|$. Then

$$\|u_{T,i}\|^2 = \sum_{x\in T} |u(x)|^2 = \sum_{x\in T} \Big|u(i) + \sum_{(a,b)\in P_{i\to x}} (u(b)-u(a))\Big|^2, \tag{4.18}$$

where we denoted by $P_{i\to x}$ the unique path from i to x in T. Let us denote by $f : (\mathbb{T}_t, [t]) \to \|u_{T,i}\|^2$. We will show that

$$f(L_t,1) = f(L_t,t) = \max_{T,i} f(T,i), \quad f(L_t,1) > \max_{T\neq L_t,\, x\neq 1,t} f(T,i) + c, \tag{4.19}$$

for some constant $c > 0$. Let us denote by Once (4.19) is established, we can write using (4.17)

$$\frac{\mathcal{E}(u)}{\|u_{L_t,1}\|} < \frac{\mathcal{E}(u)}{\|u_{T,i}\|^2 + \varepsilon} \le \frac{\mathcal{E}(u)}{\|u_{T,i}\|^2} - \varepsilon\gamma_*(t),$$

for any valid choice of u and any $(L_t,t) \neq (T,i) \neq (L_t,1)$. Therefore, since by Lemma 4.11 $\gamma_*(t) > c > 0$, for some $c > 0$, we can conclude using (4.17) by seeing that when comparing a tree (T,i) with $(L_t,1)$, the numerator remains the same and the denominator increases.

We turn to the proof of (4.19). For the maximum, we can suppose that $u(a)-u(b) \ge 0$ if $b \in P_{i\to a}$, i.e. if we need to visit b before reaching a when starting from i. In other words, u is increasing when going away from i. Let us write p_n, $n = 1,\dots,t-1$, the values $\{|u(x)-u(y)| : (x,y) \in E(T)\}$ and $q_n := p_{(n)}$ those values sorted decreasingly,

$$q_1 \ge q_2 \ge \ldots \ge q_{t-1} \ge 0.$$

Then $\mathcal{E}(u) = \sum_{n\in[t-1]} p_n^2 = \sum_{n\in[t-1]} q_n^2$. Moreover, suppose $a,b \in [t]$, $(a,b) \in E$ and $l+1 = d(a,i) = d(b,i)+1$ for some $l \ge 0$. Then, by ordering, $u(a)-u(b) \le q_l$. Therefore, for any $x \in [T]$, we have

$$\sum_{(a,b)\in P_{i\to x}} (u(a)-u(b)) \le \sum_{l=1}^{d(i,x)} q_l.$$

Plugging this into (4.18), we ge

$$\|u_{T,i}\|^2 = \sum_{x\in T}\Big|u(i) + \sum_{(a,b)\in P_{i\to x}}(u(a)-u(b))\Big|^2 \le \sum_{x\in T}\Big|u(i) + \sum_{l=1}^{d(i,x)} q_l\Big|^2 \le \sum_{k=0}^{t}\Big|u(i) + \sum_{l=1}^{k} q_l\Big|^2.$$

The only case where we have equality is if T is a line and i an extremal point. This proves the right-hand side of (4.19). □

4.5 Proof of Theorem 4.1

In this section, we prove Theorem 4.1. The strategy is to first study the spectrum of the matrix $\underline{L}(\mathbb{G}^{\tau,\zeta})$ in the interval $\big[-\sqrt{d}, -\sqrt{d}+\zeta d^{-1/2}\big]$. This is done using rank-one perturbation theory (see Lemma A.8) and geometric properties of the burned graph $\mathbb{G}^{\tau,\zeta}$. From that analysis we construct a block diagonal approximation of $\underline{L}(\mathbb{G}^{\tau,\zeta})$. We then use the interlacing of eigenvalues to transfer information from $\operatorname{Spec}\underline{L}(\mathbb{G}^{\tau,\zeta})$ to $\operatorname{Spec}\underline{L}(\mathbb{G})$. Finally we use the fact that the neighborhood of maximal trees generates the smallest spectral gaps (Lemma 4.11) and that they are identical in $\mathbb{G}$ and $\mathbb{G}^{\tau,\zeta}$ ((v) of Proposition 4.9) to conclude.

The two following propositions give estimates on the smallest eigenvalues generated by trees in the graph $\mathbb{G}^{\tau,\zeta}$ described in Proposition 4.9.

Proposition 4.13 (Rigidity for $\mathcal{A}$ and $\mathcal{C}$)**.** *Let $t_* \in \mathbb{N}^*$, $\varepsilon > 0$ and r,ζ,τ as in (4.11). Then there exists $c_* = c_*(\mathbb{T}_{\le t})$ such that for any $\alpha \in (0,1/4)$ the following holds with probability $1 - O(\mathrm{e}^{-d^{2\alpha}})$.*

(i) For each $T \in \mathcal{A}\cup\mathcal{C}$, the matrix $\underline{L}(\mathbb{G}^{\tau,\zeta})$ has at most t eigenvalues, $\mu_1 < \mu_2 < \ldots < \mu_t$ in the interval $(-\sqrt{d}, -\sqrt{d}+2\zeta)$.

(ii) For each $T \in \mathcal{A}\cup\mathcal{C}$, if $|S_1(T)| = 1$, then

$$\mu_1 = -\sqrt{d} + \frac{\gamma_{T,x_T}}{\sqrt{d}} + \begin{cases} \frac{\delta_{T,x_T}}{d} + O\Big(\frac{d^\alpha}{d^{3/2}}\Big), & |T| = t_* \\ O\big(d^{-1}\big), & \textit{else.}\end{cases} \tag{4.20}$$

(iii) For each $T \in \mathcal{A}\cup\mathcal{C}$, if $|S_1(T)| > 1$, then

$$\mu_1 = -\sqrt{d} + \frac{\gamma_{T,x_T(1)} + c_*}{\sqrt{d}} + \Lambda(z_T), \quad \Lambda(z_T) = O(d^{-3/2}) \tag{4.21}$$

(iv) For each $T \in \mathcal{A}\cup\mathcal{C}$, if $\mathbf{w}_n$, denotes the eigenvector corresponding to μ_n then for $1 \le i \le r$

$$\|\mathbf{w}_n|_{S_i(T)}\| = O\left(\frac{1}{d^{i/2}}\right), \quad \langle \mathbf{w}_m\,, (L-\mu_n)\mathbf{w}_n\rangle = O\left(\frac{1}{d^{r/2}}\right), \quad n,m \in [t]. \tag{4.22}$$

The proof of Proposition 4.13 is deferred to Section 4.6.

Proposition 4.14 (Rigidity for $\mathcal{B}$, $\mathcal{D}$ and $\mathcal{L}(\zeta)$)**.** *Let $t_* \in \mathbb{N}^*$, $\varepsilon > 0$ and r,ζ,τ as in (4.11). Then there exists $c_* = c_*(\mathbb{T}_{\le t})$ such that for any $\alpha \in (0,1/4)$ the following holds with probability $1 - O(\mathrm{e}^{-d^{2\alpha}})$.*

(i) For each $U \in \mathcal{B}\cup\mathcal{L}(\zeta)$, the matrix $\underline{L}(\mathbb{G}^{\tau,\zeta})|_{B_r(U)}$ has at most t eigenvalues, $\mu_1 < \mu_2 < \ldots < \mu_t$ in the interval $(-\sqrt{d}, -\sqrt{d}+2\zeta)$.

(ii) For each $U \in \mathcal{B}$ for $n \ge 1$ then

$$\mu_1 \ge -\sqrt{d} + \frac{\gamma_*(t) + c_*}{\sqrt{d}} + o(d^{-1}) \tag{4.23}$$

(iii) *If $x \in \mathcal{L}(\zeta)$ then*

$$\mu_1 \geq -\sqrt{d} + \frac{5}{4\sqrt{d}}. \tag{4.24}$$

(iv) *For each $U \in \mathcal{U}(\zeta)$, if $\mathbf{w}_n$, denotes the eigenvector corresponding to μ_n then for $1 \leq i \leq r$*

$$\|\mathbf{w}_n|_{S_i(U)}\| = O\Big(\frac{1}{d^{(i-t)_+/2}\zeta^{i-t}}\Big), \quad \langle \mathbf{w}_m\,, (L-\mu_n)\mathbf{w}_n\rangle = O\Big(d^{-\frac{r-t}{2}}\Big), \quad n,m \in [t]. \tag{4.25}$$

The proof of Proposition 4.14 is deferred to Section 4.6

Lemma 4.15 (Equi-probability of trees)**.** *Let $1 < d < N$ and $t \in \mathbb{N}^*$. Then for any $T \in \mathbb{T}_s$ and $x \in [s]$, $1 \leq s \leq t$ we have with probability $1 - O\big(N^{1-3\eta}\big)$*

$$\big|\big\{U \in \mathcal{T}_s(1) : U = T,\, U \cap S_1(\mathbb{G} \setminus U) = \{x\}\big\}\big| = \frac{1}{s^{s-3}} N w(t)\big(1 + O(\frac{1}{N^\eta})\big)$$

Proof. This comes from the independence of the entries of A. For $T \subseteq [N]$, conditioned upon the event that $T \in \mathcal{T}_t$ the probability of seeing either of the t^{t-2} trees is the same. □

Proposition 4.16 (Trivial eigenvector and other small connected components)**.** *There are at most $O(N^\eta)$ connected components and the microscopic ones have at most t vertices and are trees. Moreover there exists $\eta > 0$ and a normalized eigenvector $\mathbf{q}$ supported on the complement of $\bigcup_{T \in \mathcal{T}(\zeta)^{\tau,\zeta}} B_r(T)$, such that*

$$\big\|(\underline{L}(\mathbb{G}^{\tau,\zeta}) - d^{-1/2})\mathbf{q}\big\| = O\Big(N^{-\eta}\Big)$$

Proof. Similar to the proof of Proposition 3.27. Use

$$\Big|\bigcup_{T \in \mathcal{T}(\zeta)} B_r(T)\Big| = O\Big(d^{r+1} N d^\zeta \mathrm{e}^{-d}\Big) = O\Big(N^{-\eta}\Big),$$

for $\eta > 0$ small enough, using the fact that $\frac{d}{\log N} > c > 0$ for some small enough constant $c > 0$. □

The following result is a restatement of Proposition 3.25. If d is as in (4.3), then

$$\|\underline{A} - \mathbb{E}\underline{A}\| \leq 2\sqrt{t}, \tag{4.26}$$

with very high probability. Here $\underline{A}$ denotes the adjacency matrix of the graph $\mathbb{G}$.

Proof of Theorem 4.1. We begin by understanding the spectrum of $\underline{L}(\mathbb{G}^{\tau,\zeta})$. Let us set $t \in \mathbb{N}^*$, $r \geq 10 + t$, $c_* = c_*(\mathbb{T}_{\leq t})$ and $\eta > 0$ small enough so that the results of Proposition 4.13, 4.14, 4.16 and Lemma 4.15 hold with probability $1 - O\big(\mathrm{e}^{-d^\eta}\big)$. We work on this event in the rest of the proof.
On this event, there exists an orthogonal matrix U such that

$$\underline{L}^{\tau,\zeta} = U \begin{bmatrix} \nu & 0 & 0 & E_\nu^* & 0 \\ 0 & \mathcal{U} & 0 & E_0^* & 0 \\ 0 & 0 & \mathcal{U}' & E_1^* & 0 \\ E_\nu & E_0 & E_1 & X & 0 \\ 0 & 0 & 0 & 0 & Y \end{bmatrix} U^* \tag{4.27}$$

where $\nu = -\sqrt{d} + O(N^{-\eta})$ for some constant $\eta > 0$ small enough. The diagonal block $\mathcal{U}$ corresponds to the eigenvalues generated by the neighborhood of trees belonging to $\mathcal{A}$ and $\mathcal{C}$. The diagonal block $\mathcal{U}$ corresponds to the eigenvalues generated by the neighborhood of trees belonging to $\mathcal{B}$ and of small degree vertices (i.e.

$\mathcal{L}(\zeta)$). From (4.22) and (4.25), we see that $\|E_i\| = O(d^{-5})$, for $i \in [3]$ and $\|E_\nu\| = O(N^{-\eta})$. From (4.23) and (4.24), we see that

$$\min_{\lambda\in\mathcal{U}} \lambda = -\sqrt{d} + \frac{\gamma_*(t)}{\sqrt{d}} + O\big(d^{-1}\big).$$

Let $\mathbf{w}$ be an eigenvector orthogonal to the eigenvectors from the first $1 + |\mathcal{U}| + |\mathcal{U}'|$ columns of U. Then $\|\mathbf{w}|_{\overline{\mathcal{U}(\zeta)}}\|^2 = O\big(d^{-1}\big)$, by (4.22) and (4.25). If $\lambda \in \operatorname{Spec} X$, we thus have Orthogonal eigenvectors can put maximum weight $\frac{1}{\sqrt{\zeta}}$ on $\mathcal{U}(\zeta)$. Therefore $\sigma(X) \subseteq [\zeta, -\infty)$ since

$$\lambda \geq \zeta - \frac{C}{\sqrt{d}} - \|\underline{A} - \mathbb{E}\underline{A}\| \geq -\sqrt{d} + d^{-1/2}.$$

This shows that the block X does not contribute any eigenvalue to the interval $[-\sqrt{d}, -\sqrt{d} + d^{-1/2}]$. Finally, Y corresponds to the microscopic connected components of $\mathbb{G}^\zeta$. Using (4.16), we find

$$\operatorname{Spec} Y \subseteq \bigcup_{T\in\mathbb{T}_{\leq t}} \bigcup_{\lambda\in\operatorname{Spec}(\underline{L}(T))} \{\lambda\} \subseteq \bigcup_{T\in\mathbb{T}_{\leq t}} \operatorname{Spec} L(T) \subseteq [-\sqrt{d} + \frac{\gamma(t) + c_*}{\sqrt{d}}, +\infty).$$

We can now reconstruct the matrix $L(\mathbb{G})$ via the equality

$$\underline{L}(\mathbb{G}) = \underline{L}(\mathbb{G}^{\tau,\zeta}) + \sum_{e\in E} B(e) = U \begin{bmatrix} \nu & 0 & 0 & E_\nu^* & 0 \\ 0 & \mathcal{U} & 0 & E_0^* & 0 \\ 0 & 0 & \mathcal{U}' & E_1^* & 0 \\ E_\nu & E_0 & E_1 & X & 0 \\ 0 & 0 & 0 & 0 & Y \end{bmatrix} U^* + \sum_{e\in E} B(e).$$

Since $B(e)$ is a non-negative rank one matrix, all the lower inequalities derived for the spectrums of the blocks of $U^*\underline{L}(\mathbb{G}^{\tau,\zeta})U$ remain valid by interlacing of eigenvalues.
We will now show that the spectrum of the block $\mathcal{U}$ is not perturbed too much. Let $\mathbf{w}$ be an eigenvector with corresponding eigenvalue $\lambda \leq -\sqrt{d} + \frac{2t}{\sqrt{d}}$ and $\|\mathbf{w}|_T\| = 1 - o(1)$ for some $T \in \mathcal{A}$. Then by (4.13) and (4.22) we have

$$\left\| \sum_{e\in E} B_e \mathbf{w} \right\| = O(d^{-r}),$$

and therefore $\mathbf{w}$ is an approximate eigenvector of $\underline{L}(\mathbb{G})$, with error $O\big(d^{-r}\big)$. We conclude that there exists an eigenvalue at $\lambda + O\big(d^{-1}\big)$. Using Lemma 4.15 the configuration $(T, x) \in \bigcup_{u\leq t} \mathbb{T}_u \times [u]$ that minimizes $\gamma(T, x)$ appears with high probability as soon as $d \leq (1-\varepsilon)d^*(t, 1)$. Finally, by Lemma 4.12, we can identify precisely the minimizing configuration (T, x) as being the line attached to the rest of the graph by one of its extremity. This concludes the proof. □

4.6 Spectrum around trees

In this section we consider $d \geq 1$ to be the density parameter of the Erdős-Rényi graph. We set $t^* \in \mathbb{N}^*$ and d as in (4.2). We work only on the graph $\mathbb{G}^{\tau,\zeta}$ defined in Proposition 4.9. We drop the superscript τ, ζ in the rest of this section. Let us write $\Xi^{\tau,\zeta}$, $\tau, \zeta > 0$, the event on which the results of Lemma 4.7, Lemma 4.8 and Proposition 4.9 and (4.10) hold. All the results stated in this section hold on the event $\Xi^{\tau,\zeta}$ and can thus be seen as *deterministic* results.

Suppose $Y \in \mathcal{T}(\zeta)$ and Y is isomorphic to some finite tree $T \in \mathbb{T}$. Then there is a natural way to associate the anchors of Y with numbers of $\{1, \ldots, |T|\}$.

We first state a general result about the eigenvalues of finite trees.

Lemma 4.17. *Let $t \in \mathbb{N}^*$, $T \in \mathbb{T}_t$. Then*

$$\operatorname{Spec} L(T) \subseteq [0, 2t], \qquad \frac{2}{t^2} \geq \lambda_2(L(T)) - \lambda_1(L(T)) > 0.$$

Proof. First of all, it is well-known that the Laplacian matrix has positive spectrum, just take the eigenvector $\mathbf{q} := \frac{1}{\sqrt{t}}\mathbf{1}_T$. To show that $\lambda_2 > 0$, observe that if $\mathbf{w} \in \mathbb{R}^t$ is a normalized vector such that

$$\langle \mathbf{w}, L(T)\mathbf{w} \rangle = \sum_{x,y} (\mathbf{w}_x - \mathbf{w}_y)^2 = 0,$$

then $\mathbf{w}_x = \mathbf{w}_y$ for every $x, y \in [t]$. Therefore if $\mathbf{w} \neq \mathbf{q}$, necessarily $\langle \mathbf{w}, L(T)\mathbf{w} \rangle > 0$.
Second of all, $L(T) = D(T) - A(T)$ where $D(T)$ is the diagonal matrix with the degrees on the diagonal and $A(T)$ is the adjacency matrix of T. Now since $\max_{x \in T} \leq t - 1$ and $\|A\| \leq 2\sqrt{\max_{x \in T} D_x} \leq 2\sqrt{t-1}$. Therefore $\|L(T)\| \leq t - 1 + 2\sqrt{t-1} \leq 2t$, as soon as $t \geq 4$. For $t = 2, 3$ we compute explicitly the spectrum of the two trees of size 3 and 2 and find that the bound $\|L(T)\| \leq 2t$ still holds. □

Lemma 4.18 (Spectrum around ideal trees with one anchor)**.** *Let $t \in \mathbb{N}^*$, $T \in \mathbb{T}_t$. There exists a constant $c_{\text{gap}} := c_{\text{gap}}(T) > 0$ such that on the event $\Xi^{\tau,\zeta}$, for every $Y \in (\mathcal{A} \cup \mathcal{C}) \cap \mathcal{T}^{(T)}$, with* $\operatorname{config}^{\mathbb{G}}(Y) = (T, \{a\})$, *$a \in [t]$, the following holds.*
The matrix $\underline{L}|_{B_r(T)}$ has precisely t eigenvalues smaller than $-\sqrt{d} + t$. Moreover its smallest eigenvalue λ_1 and λ_2 satisfy

$$\lambda_1 = -\sqrt{d} + \frac{\gamma(T, a)}{\sqrt{d}} + \frac{\delta(T, a)}{\sqrt{d}(\sqrt{d} + v_z)} + O\big(d^{-3/2}\big), \qquad \lambda_2 \geq \lambda_1 + \frac{c_{\text{gap}}}{\sqrt{d}}, \tag{4.28}$$

where $\{z\} = S_1(Y)$ and γ is defined in (4.14) and δ is a universal constants in the sense of Remark 4.2.
The eigenvectors $(\mathbf{u}(i))_{i \in [t]}$ of $\underline{L}|_{B_r(T)}$ corresponding to the t smallest eigenvalues satisfy

$$\|\mathbf{w}|_{B_r(T) \setminus B_i(T)}\| = O\Big(\Big(\frac{1}{\sqrt{d}}\Big)^{i+1}\Big), \quad i \geq 0. \tag{4.29}$$

Proof. We work on the event $\Xi^{\tau,\zeta}$. Let us write $M := \underline{L}^\zeta|_{B_r(Y)}$. The matrix M can be written as a rank one perturbation of a block diagonal matrix

$$M = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix} + B((x, z)), \quad M_1 := \underline{L}(T), \quad M_2 := M|_{B_r(Y) \setminus Y}. \tag{4.30}$$

Since T is a tree, we know by Lemma 4.17 $\|L(T)\| \leq 2t$. Moreover on $\Xi^{\tau,\zeta}$, the graph $B_r(Y) \setminus T$ is a tree and we deduce

$$\|d^{-1/2} A|_{B_r(Y) \setminus T}\| \leq 2\sqrt{\frac{\max_{y \in [N]} D_y}{d}} \leq C,$$

for some constant $C \geq 0$ that depends only t_*. This allows us to deduce that

$$\operatorname{Spec}\big(M_2\big) \subseteq \Big[\min_{y \in B_r(Y) \setminus T} v_y - C, +\infty\Big) \subseteq \Big[(\tau - 1)\sqrt{d} - C, +\infty\Big). \tag{4.31}$$

Here we used Proposition 4.9 (iii) and the fact that $Y \in \mathcal{A} \cup \mathcal{C}$ to derive the lower bound on $\min_{y \in B_r(T) \setminus T} v_y$. We also observe that $\operatorname{Spec} M_1 \subseteq \big[-\sqrt{d}, -\sqrt{d} + \frac{2t}{\sqrt{d}}\big]$ and that $\lambda_1(M_1) = \lambda_1(L(T)) - \sqrt{d} = -\sqrt{d}$. Since M is a rank-one perturbation of the block matrix $\begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}$, the interlacement of eigenvalues tells us that

$$-\sqrt{d} = \lambda_1(M_1) < \lambda_1(M) < \lambda_2(M_1) < \lambda_2(M) \cdots \lambda_t(M_1) < \lambda_t(M) \leq \lambda_t(M_1) + \|B(x_T, z_T)\| \leq -\sqrt{d} + \frac{4t}{\sqrt{d}},$$

where we used the fact that $\|B(x_T, z_T)\| \le 2d^{-1/2}$. The first claim of the theorem immediately follows.

We will now use perturbation theory to compute the precise location of λ_i, $i \in [t]$. By Lemma A.8 with $H = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}$ and $e = (\mathbf{1}_x - \mathbf{1}_z)d^{-1/2}$ we find that

$$\sqrt{d} = \Big(\frac{1}{t - M_1}\Big)_{xx} + \Big(\frac{1}{t - M_2}\Big)_{zz}, \quad t \in \mathbb{R}. \tag{4.32}$$

By the change of variable $t = -\sqrt{d} + \frac{\theta}{\sqrt{d}}$, for $\theta \in (0, \lambda_2(L(T)))$, (4.32) becomes

$$1 = \Big(\frac{1}{\theta - L(T)}\Big)_{xx} + g\Big(-\sqrt{d} + \frac{\theta}{\sqrt{d}}\Big), \qquad g(t) := \Big(\frac{1}{t - M_2}\Big)_{zz}. \tag{4.33}$$

We will solve (4.33) on the interval $I_T := (0, \lambda_2(L(T)))$.

We first use the spectral gap between $\mathrm{Spec}(M_2)$ and $(-\sqrt{d}, \lambda_2(M_1))$ to linearize the function g around $-\sqrt{d}$. By (4.31), the smallest eigenvalue of M_2 is at distance at least $\frac{\tau}{2}\sqrt{d}$ of $\lambda_2(M_1)$. We deduce that $|g(t)| = O\big(d^{-1/2}\big)$ and $|g'(t)| = \big|\big(\frac{-1}{(z-M_2)^2}\big)_{zz}\big| = O\big(d^{-1}\big)$ for every $t \in [-\sqrt{d}, 2\lambda_t(M_1)]$. By a Taylor expansion to the first order, we find (see the proof of Proposition **??**)

$$g\Big(-\sqrt{d} + \frac{\theta}{\sqrt{d}}\Big) = g\big(-\sqrt{d}\big) + O\big(d^{-3/2}\big) = \frac{-1}{\sqrt{d} + v_z} + O\big(d^{-3/2}\big), \tag{4.34}$$

on Ξ.

Let us introduce the function

$$f : \mathbb{R} \setminus \mathrm{Spec}(L(T)) \to \mathbb{R}, \qquad f(t) = \Big(\frac{1}{t - L(T)}\Big)_{xx}.$$

This function is universal, in the sense of Remark 4.2. The function f is smooth and invertible on the open interval $I := (0, \lambda_2(L(T)))$. By Taylor's theorem, there exists $C, c > 0$, such that $[1 - 2c, 1 + 2c] \subseteq I$,

$$\max_{|u-1| \le c} |(f^{-1})''(u)| \le C.$$

and thus

$$f^{-1}(1 + \varepsilon) = f^{-1}(1) + \frac{(f^{-1})'(1)}{\sqrt{d}}\varepsilon + O(\varepsilon^2), \quad \forall |\varepsilon| \le c. \tag{4.35}$$

We see that the solution θ of (4.33) is given by the self-consistent equation

$$\theta = f^{-1}\Big(1 + g\big(-\sqrt{d} + \theta/\sqrt{d}\big)\Big).$$

Using (4.34) and (4.35), this becomes

$$\theta = f^{-1}(1) + (f^{-1})'(1)\Big[\frac{-1}{\sqrt{d} + v_z} + O\big(d^{-3/2}\big)\Big] + O\Big(d^{-1}\Big) = f^{-1}(1) - \frac{(f^{-1})'(1)}{\sqrt{d} + v_z} + O\big(d^{-1}\big).$$

Setting

$$\gamma(T, x) := f^{-1}(1), \qquad \delta(T, x) := -(f^{-1})'(1), \qquad \Lambda(z) := \frac{1}{\sqrt{d}(\sqrt{d} + v_z)},$$

and deduce that the smallest real solution t of (4.32) lies in the interval $(-\sqrt{d}, \lambda_2(M_1))$ and satisfies

$$t = -\sqrt{d} + \frac{\gamma(T, x)}{\sqrt{d}} + \frac{\delta(T, x)}{\sqrt{d}(v_z + \sqrt{d})} + O\big(d^{-3/2}\big).$$

We conclude the first part of (4.28).
The eigenvalues $\lambda_i(M)$, $2 \le i \le t$, can be computed in the same way, by solving (4.32) on every interval $(\lambda_i(M_1), \lambda_{i+1}(M_1))$, $i = [t]$ (with $\lambda_{t+1}(M_1) = +\infty$ by convention). If θ_i, $i = [t]$, denote the t solutions of (4.33), we see that, as soon as we bound $g(z) = o(1)$, there exists $c > 0$, depending on f, such that $\min_{i \neq j} |\theta_i - \theta_j| > c$. The constant c is universal since it depends only on f and thus the second part of (4.28).

We now turn to the proof of (4.29). The argument is in many ways similar to the one found for the analog statement in the proof of Proposition 3.34 in Chapter 3.

Let $Q := \sum_{u \notin T} \mathbf{1}_u \mathbf{1}_u^*$ be the projection onto the vertices in $B_r(T) \setminus T$. Let $\mu = \lambda_i$ and $\mathbf{w} = \mathbf{w}_i$ for some $i \in [t]$. The eigenvalue eigenvector equation becomes

$$(QMQ - \mu)Q\mathbf{w} = -QA(1 - Q)\mathbf{w}, \tag{4.36}$$

where $A := \underline{A}|_{B_r(T)}$. Again using the fact that $\|A\| \le C$, for some constant $C > 0$, we see that

$$\operatorname{Spec} QMQ \subset \Big[\min_{u \notin T} v_u - \|A\| - \|B((x,z))\|, +\infty\Big) \subset \Big[\big(\frac{\mathcal{T}}{2} - 1\big)\sqrt{d}, +\infty\Big),$$

and thus, using the fact that $\mu \le -\sqrt{d} + 4td^{-1/2}$,

$$\|\mathbf{w}|_{S_1(T)}\| \le \|Q\mathbf{w}\| \le \frac{\|\underline{A}|_{B_r(x)}\|}{\|(QLQ - \mu)\|} = O\big(d^{-1/2}\big). \tag{4.37}$$

The bound for $2 \le i \le r$ is proved similarly. This concludes the proof. □

Lemma 4.19 (Spectrum around ideal trees with many anchors)**.** *Let $t \in \mathbb{N}^*$, $T \in \mathbb{T}_t$. There exists a constant $c_{\text{gap}} := c_{\text{gap}}(T) > 0$ such that on the event $\Xi^{\tau,\zeta}$, for every $Y \in (\mathcal{A} \cup \mathcal{C}) \cap \mathcal{T}^{(T)}$ such that* $\operatorname{config}^{\mathbb{G}}(Y) = (T, A)$, *$|A| > 1$, the following holds.*
The matrix $\underline{L}^{\zeta}|_{B_r(T)}$ has precisely t eigenvalues smaller than $\sqrt{d} + 2t$. The smallest eigenvalue λ_1 satisfies

$$\lambda_1 \ge -\sqrt{d} + \frac{\max_a \gamma(T, a) + c_{\text{gap}}}{\sqrt{d}}, \qquad \lambda_2 \ge \lambda_1 + \frac{c_{\text{gap}}}{\sqrt{d}}, \tag{4.38}$$

where γ is defined in (4.14).
The eigenvectors $(\mathbf{u}(i))_{i \in [t]}$ of $\underline{L}(\mathbb{G}|_{B_r(Y)})$ corresponding to the t smallest eigenvalues satisfy

$$\|\mathbf{w}|_{B_r(T) \setminus B_i(T)}\| = O\Big(\Big(\frac{1}{\sqrt{d}}\Big)^{i+1}\Big).$$

Proof. Let us fix $Y \in \mathcal{A} \cup \mathcal{C}$, $\ell := |S_1(T)|$ and abbreviate $x_i := x_{Y_i}$, $z_i := z_{Y_i}$, for $i \in [\ell]$. We assume that $\gamma(T, x_i) \ge \gamma(T, x_{i+1})$. Writing $M := \underline{L}|_{B_r(T)}$, we introduce

$$M_k := M - \sum_{i=k+1}^{\ell} B((x_i, z_i)), \qquad k \in [\ell].$$

Then M_k corresponds to the graph $B_r(T)$ with the edges (x_i, z_i), $k \ge i$ removed (note that $M_\ell = M$). We will show that there exists $c > 0$, depending only on T, such that

$$\lambda_1(M_2) > \lambda_1(M_1) + \frac{c}{\sqrt{d}}. \tag{4.39}$$

In words, adding a connection between T and $B_r(T) \setminus T$ shifts the smallest eigenvalue of the matrix by a factor of order $d^{-1/2}$. (4.38) will then follow from (4.39) since B is a positive rank-one perturbation and $M_\ell = M$.

We begin by computing the smallest eigenvalue of M_1. We copy the proof of Lemma 4.18 (notice that in the graph corresponding to M_1, the tree Y has one anchor) and we see that

$$\lambda_1(M_1) = -\sqrt{d} + \frac{\gamma(T, x_1)}{\sqrt{d}} + O\big(d^{-1}\big).$$

We apply Lemma A.8 one more time. Here we crucially use the fact that $B_r(Y)$ does not contain any cycle. Setting $H = M_1$ and $e = (\mathbf{1}_{x_2} - \mathbf{1}_{z_2})d^{-1/2}$, we see that H is a block diagonal matrix (one block corresponds to Y and the region around z_1 and the other to the region around z_2) and we find that that $\lambda_1(M_2)$ satisfies the equation

$$\sqrt{d} = \Big(\frac{1}{t - \underline{L}(T)}\Big)_{x_2x_2} + \Big(\frac{1}{t - M_1}\Big)_{z_2z_2}, \qquad t \in \big(\lambda_1(M_1), \lambda_2(M_1)\big).$$

We can use the fact that $B_r(Y)$ is a tree and a spectral gap argument, similar to the one in the proof of Lemma 4.18 to bound $\Big|\frac{1}{t-M_1}\Big|_{z_2z_2} = O\big(d^{-1/2}\big)$. Using the change of variables $z = -\sqrt{d} + \frac{\theta}{\sqrt{d}}$, this equation becomes

$$1 = \Big(\frac{1}{\theta - L(T)}\Big)_{x_2x_2} + O\big(d^{-1/2}\big), \qquad \lambda_1(M_1) < \theta < \lambda_2(M_1). \tag{4.40}$$

Since the function $f(t) := \Big(\frac{1}{\theta-L(T)}\Big)_{x_2x_2}$ is smooth and bijective on the open interval $(\lambda_1(L(T)), \lambda_2(L(T)))$, it admits a smooth inverse f^{-1}. Observe that if $|\theta - \lambda_i(L(T))| \leq 1/10$, $i = 1, 2$, and N is large enough, (4.40) cannot be satisfied. This shows that

$$\lambda_1(M_1) + \frac{1}{10\sqrt{d}} \leq \lambda_1(M_2) \leq \lambda_2(M_1) - \frac{1}{10\sqrt{d}}.$$

Choosing $c_{\text{gap}} < \frac{1}{10}$, we conclude (4.39). The statement about the eigenvectors is proved as its analog of Lemma 4.18. ☐

Proof of Proposition 4.13. Combine Lemmas 4.18 and 4.19. ☐

We now turn to the slightly more tedious proof of Proposition 4.14. We first establish results analogous to Lemma 4.18 for elements of $\mathcal{D}$ and $\mathcal{L}$. The argument for $\mathcal{L}$ is a simple perturbation argument (it is a simplified version of Proposition 3.34 of Chapter 3). The argument for $\mathcal{D}$ relies on (iv), which states that elements of $\mathcal{D}$ are of size $\leq t_* - 1$.

Lemma 4.20. *[Spectrum around small single vertices] On the event $\Xi^{\tau,\zeta}$ the following hold. If $x \in \mathcal{L}(\zeta)$ and $S_1(x) \cap V = \emptyset$ the smallest eigenvalue of $\underline{L}|_{B_{r+1}(x)}$ is greater than*

$$\lambda_1 \geq -\sqrt{d} + \frac{1}{\sqrt{d}}\Big(1 + O\big(d^{-1}\big)\Big).$$

Proof. This follows from Lemma **??**, the smallest eigenvalue of $M := \underline{L}|_{B_{r+1}(x)}$ is given by

$$\mu = v_x + \frac{1}{d}\sum_{y\in S_1(x)} \frac{1}{v_x - v_y} + O\Big(\frac{1}{d^{3/2}}\Big) = v_x + \frac{\alpha_x}{v_x} + O\Big(\frac{1}{d^{3/2}}\Big) \geq v_{(1)} - \frac{1}{\sqrt{d}} + O\Big(\frac{1}{d^{3/2}}\Big).$$

☐

Lemma 4.21. *There exists $c_* > 0$ depending on $\mathbb{T}_{\leq t^*}$ such that on $\Xi^{\tau,\zeta}$ the following hold. For every $Y \in \mathcal{D}$, we have*

$$\lambda_1(\underline{L}|_{B_r(Y)}) \geq -\sqrt{d} + \frac{\gamma_*(t_*) + c_*}{\sqrt{d}}.$$

Proof. This follows from Proposition 4.9 (iv). If $Y \in \mathcal{D}$, this means that either Y is a single vertex with degrees more than t_* or $|Y| < t-1$. In that, it follows from the fact that the eigenvalue is greater than $\gamma_*(t_*-1) > \gamma_*(t_*)$. □

We now turn to the analysis of the spectrum generated by the neighborhoods of elements of $\mathcal{B}$. This is the most technical part of the section. The key observation which is formulated in Lemma 4.22 allows us to compare the spectrum generated by two or more trees that share one common anchor. For such a collection of trees having total weight t, there is a tree of size t that generates a smaller spectral gap.

We partition $\mathcal{B}$ using the equivalence relation

$$Y \sim Y' \Leftrightarrow S_1(Y) \cap S_1(Y') \neq \emptyset.$$

For $\{Y_1, \dots, Y_l\} = [Y] \in \mathcal{B}/\sim$, we define the total size of the equivalence class as $|[Y]| := \sum_i Y_i$. We call an equivalence class a *bouquet* of trees. The trees of any bouquet share a unique common anchor.

The next results state that if the common anchor of an equivalence class of trees is in $\mathcal{V}_\tau$, then the eigenvalue generated by the restriction of $\underline{L}^{\tau,\zeta}$ to the neighborhood of anchor is not minimal. It basically relies on the fact that by Proposition 4.9, this situation is only possible if the size of the equivalence class is strictly smaller than t_*.

Lemma 4.22. *Let $t \in \mathbb{N}^*$ and $v \in V$, $[Y] \in \mathcal{B}/\sim$ such that the common anchor of $[Y]$ is v. On the event $\Xi^{\tau,\zeta}$, if $v \in \mathcal{V}_\tau$, there exists a constant $c > 0$ such that*

$$\lambda_1(\underline{L}|_{B_r(v)}) \geq -\sqrt{d} + \frac{\gamma_*(t_*) + c_*}{\sqrt{d}}. \tag{4.41}$$

If $\overline{Y} := \bigcup_{Y \in [Y]} Y$, the eigenvectors $(\mathbf{u}(i))$, $i \in [|\overline{Y}|]$, corresponding to the $|\overline{Y}|$ smallest eigenvalues satisfy

$$\|\mathbf{w}|_{B_r(\overline{Y}) \setminus B_i(\overline{Y})}\| = O\left(d^{-i/2}\right), \quad i \geq 1.$$

Proof. Consider the trees $T_1, \dots, T_k$ that make up the class $[Y]$. By Proposition 4.9 (iv) we know that $|[Y]| = \sum_i |T_i| < t_*$. The common anchor v has at least two neighbors outside of $\overline{Y}$ since $\zeta > t_* + 2$ and, by Lemma 4.8, the vertex $D_v - D_v^{\tau,\zeta} \leq t_*$. Let $x_1, \dots, x_{D_v^{\tau,\zeta}-1}$ be the neighbors of v and define the matrix

$$L' := \underline{L}|_{B_r(v)} - \sum_{i=3}^{D_v^{\tau,\zeta}-1} B\big((v, x_i)\big).$$

The graph $\mathbb{G}'$ that corresponds to the matrix L' is the graph with a single ideal tree of size $t' := |[Y]| + 1$ with two anchors. Note that $t' \leq t_*$. Let $Y' := \overline{Y} \cup \{v\}$. Then Y' is a tree in $\mathbb{G}'$, with configuration $\mathrm{config}^{\mathbb{G}'}(Y) = (T', A)$, $T' \in \mathbb{T}_{\leq t_*}$ and $A \subset [t']$.
A straightforward adaptation of the proof of Lemma 4.19 (the argument did not use the fact that the anchors were different) shows that there exists $c_* > 0$, depending only on T' and A such that

$$\lambda_1(L') \geq -\sqrt{d} + \frac{\gamma_{t'} + c_*}{\sqrt{d}} + O\big(d^{-3/2}\big) \geq -\sqrt{d} + \frac{\gamma_{t_*} + c_*}{\sqrt{d}} + O\big(d^{-3/2}\big)$$

Since $B(e)$ is a positive rank-one perturbation, we know that $\lambda_1(L') \leq \lambda_1(\underline{L}|_{B_r(v)})$ and (4.41) follows.

The statement about the eigenvectors corresponding to the smallest eigenvalues is proved analogously to (4.29). We can use equation (4.37) but only starting at $S_2(\overline{Y})$, because it might be that $|v_v + \sqrt{d}| = O(1)$. However, we know, by construction, that $B_r(\overline{Y}) \setminus S_1(\overline{Y})$ contains no other vertex in $\mathcal{V}_\tau$. This concludes the proof. □

There remains to rule out the case of a bouquet of trees of total size t_*. Such a configuration might a priori generate an eigenvalue smaller than $\gamma_*(t_*)$. The next lemma is the key result of the section and insures that a bouquet of trees is strictly smaller than a single tree constructed out of the trees of the bouquet.

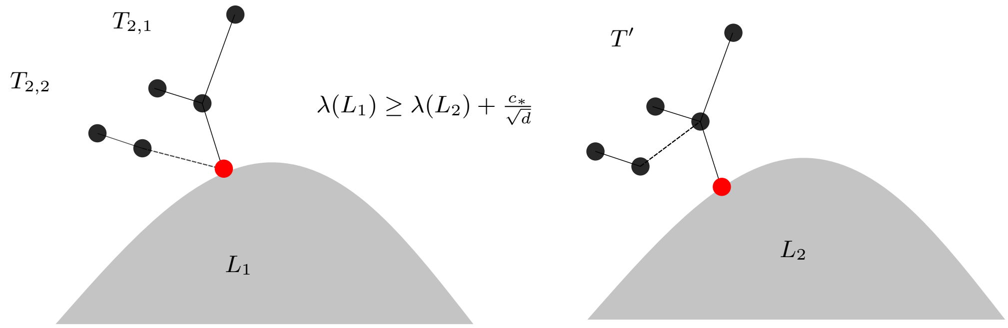


Figure 4.5: Schematic representation of Lemma 4.23

Lemma 4.23. *Let $t \in \mathbb{N}^*$ and $[Y] \in \mathcal{B}/\sim$ be a bouquet of trees with common anchor $v \in V$ that satisfies $v \notin \mathcal{V}_\tau$. There exists a constant $c_* > 0$ depending on the elements of $\mathbb{T}_{\leq t_*}$ such that on the event $\Xi^{\tau,\zeta}$,*

$$\lambda_1(\underline{L}|_{B_r(z)}) \geq -\sqrt{d} + \frac{\gamma_*(t) + c_*}{\sqrt{d}} \geq -\sqrt{d} + \frac{\gamma_*(t_*) + c_*}{\sqrt{d}}. \tag{4.42}$$

If $\overline{Y} := \bigcup_{Y \in [Y]} Y$, the eigenvectors $(\mathbf{u}(i))$, $i \in [|\overline{Y}|]$, corresponding to the $|\overline{Y}|$ smallest eigenvalues satisfy

$$\|\mathbf{w}|_{B_r(\overline{Y}) \setminus B_i(\overline{Y})}\| = O\Big(\Big(\frac{1}{\sqrt{d}}\Big)^i\Big), \quad i \geq 1.$$

Proof. Let $\{Y_1, \dots, Y_n\} = [Y]$, $n \geq 2$ be the trees that make up the bouquet. Let $\mathrm{config}^{\mathbb{G}}(Y_k) = (T_k, A_k)$, $x_k := S_1(v) \cap Y_k$ and $t_k := |T_k|$ $t \in [n]$. Observe that $|A_k| = 1$ (every tree of the bouquet has exactly one anchor) by Proposition 4.9 (iii) and write $\{a_k\} = A_k$. Let us pick $k \in [n]$ such that $\gamma(T_k, a_k)$ and define

$$L_0 := \underline{L}(\mathbb{G})|_{B_r(z)} - B((z, x_k)).$$

The graph that corresponds to L_0 is the graph $\mathbb{G}$ where we disconnected the tree Y_k by burning the edge (z, x_k). Note that the matrix L_0 has an isolated block that corresponds to $\underline{L}(T_k)$. If $Q^{(i)} := \sum_{x \in Y_i} \mathbf{1}_x \mathbf{1}_x^*$, $i \in [k]$, denotes the projection on the vertices of Y_i, then $\underline{L}(T_k) = Q^{(k)} L_0 Q^{(k)}$ corresponds to the (now) disconnected component Y_k . The other block in the matrix L_0 is $M^{(k)} := (1 - Q^{(k)}) L_0 (1 - Q^{(k)})$. We have the following block diagonal representation

$$L_0 = \begin{bmatrix} \underline{L}(T_k) & 0 \\ 0 & M^{(k)} \end{bmatrix}. \tag{4.43}$$

Let us denote by $:\tilde{=} \sum_{i \neq k} t_i$ the total size of the bouquet with Y_k removed. By perturbation theory and Lemma 4.17, we know that L_0, respectively $M^{(k)}$, has exactly $\tilde{t}$ eigenvalues in $\big[-\sqrt{d}, -\sqrt{d} + \frac{3(\tilde{t}+t_k)}{\sqrt{d}}\big]$, respectively $\tilde{t}$. Moreover, since $\underline{L}(T_k)$ is a linear shift of a Laplacian matrix, we know that the spectrum of L_0 contains exactly one eigenvalue at $-\sqrt{d}$ (every connected component generates a trivial eigenvalue).

We will now analyze how the t_k eigenvalues of $\underline{L}(T_k)$ interact with the $\tilde{t}$ smallest eigenvalues of $M^{(k)}$ under a rank-one perturbation.
Let us denote by $\mathbf{w}_i$, $i \in [\tilde{t}]$, the eigenvectors that correspond to $\lambda_i(M^{(k)})$. The first step of the proof is to show that

$$\exists y \in \overline{Y}, \text{and a constant } c > 0, \text{ such that } |\mathbf{w}_1(y)|^2 > c. \tag{4.44}$$

We will first need to derive a localization estimate on the eigenvectors of $M^{(k)}$ that correspond to its smallest eigenvalues. By Proposition 4.9 (ii) we know that $B_r(z) \cap \mathcal{V}_\tau = \emptyset$, meaning that

$$\min_{z \in B_r(\overline{Y}) \setminus \overline{Y}} v_y \geq (1 - \tau)\sqrt{d}. \tag{4.45}$$

Since $\|A|_{B_r(z)}\| = O(1)$ on the event $\Xi^{\tau,\zeta}$, there is a spectral gap of size $\frac{\tau}{3}\sqrt{d}$ between the $\tilde{t}$ first eigenvalues of $M^{(k)}$ and the rest of the spectrum. Let us denote

$$\operatorname{Spec} M^{(k)} = \Lambda_1 \cup \Lambda_2, \quad |\Lambda_1| = \tilde{t}, \quad \Lambda_1 \subset \Big(-\infty, -\sqrt{d} + 3\tilde{t}d^{-1/2}\Big), \, \Lambda_2 \subset \Big(\Big(\frac{\tau}{3} - 1\Big)\sqrt{d}, +\infty\Big) \tag{4.46}$$

Using (4.36) and (4.37), we find

$$\big\|\overline{Q}\mathbf{w}_\lambda\big\| = O\big(d^{-1/2}\big), \qquad \lambda \in \Lambda_1,$$

where $\overline{Q}$ is the projection on the vertices $B_r(v) \setminus \overline{Y}$, $\overline{Q} := 1 - Q$ and $Q := \sum_{i \neq k} Q^{(i)}$.
We deduce that

$$\max_{\lambda \in \Lambda_1} \sum_{u \in B_r(\overline{Y}) \setminus \overline{Y}} \big|\mathbf{w}_\lambda(u)\big|^2 = O\big(d^{-1}\big) \tag{4.47}$$

Since the matrix L_0 is a block diagonal matrix, (4.47) translates immediately to the eigenvectors of $M^{(k)}$ In particular, if $\mu = \lambda_1(M^{(k)})$, there exists $y \in \bigcup_{i \neq k} T_i$ such that $|\mathbf{w}_\mu(y)| \geq \frac{1}{2\tilde{t}}$. Indeed this where not the case. Then we would have

$$\|\mathbf{w}_\mu\|^2 = 1 = \sum_{x \in |M^{(k)}|} |\mathbf{w}_\mu(x)|^2 \leq \frac{\tilde{t}}{2\tilde{t}} + O\big(d^{-1}\big) < \frac{2}{3}.$$

Since $M^{(k)}$ is a sub-block of L_0, this result is identical for the eigenvectors corresponding of L_0 corresponding to $\operatorname{Spec} L_0 \setminus \operatorname{Spec} \underline{L}(T_k)$. This establishes (4.44).

We now look at two different ways to attach the tree Y_k back to the rest of the graph. The first way is to simply recreate the original graph. The second way is to create a new graph by attaching Y_k at the point $y \in \overline{Y}$ constructed in (4.44). In the former case, we add the edge (x_k, z) and in the latter, we add the edge (x_k, y). (see Figure ??? for a visual representation of the procedure.)
These two different graphs correspond to the matrices $L_{0,+} := L_0 + B((x_k, z))$ and $L_{0,-} := L_0 + B((x_k, y))$. Let us denote by T_y the tree that contains y. Those matrices have the following block representations

$$L_0 = \begin{bmatrix} \underline{L}(T_k) & 0 & 0 \\ 0 & \underline{L}(T_y) & E_0 \\ 0 & E_0^* & M \end{bmatrix}, \quad L_{0,-} = \begin{bmatrix} \underline{L}(T_k) & E_y & 0 \\ E_y^* & \underline{L}(T_y) & E_0 \\ 0 & E_0 & M \end{bmatrix}, \quad L_{0,+} = \begin{bmatrix} \underline{L}(T_k) & 0 & E_k \\ 0 & \underline{L}(T_y) & E_0 \\ E_k^* & E_0^* & M \end{bmatrix}, \tag{4.48}$$

where $E_k := -\frac{1}{\sqrt{d}}\mathbf{1}_v\mathbf{1}_{x_k}^*$ and $E_y := -\frac{1}{\sqrt{d}}\mathbf{1}_y\mathbf{1}_{x_k}^*$, and $E_0 := -\frac{1}{\sqrt{d}}\mathbf{1}_v\mathbf{1}_{x_y}^*$ (here $x_y := T_y \cap S_1(v)$ denotes the point of T_y attached to z) and $M := \underline{L}(G^{\tau,\zeta})|_{B_r(z)\setminus(T_y \cup T_k)}$ (compare with (4.43)).

We will now show that there exists a constant $c_* > 0$, depending only on the configurations of the tree in the bouquet (i.e. universal in the sense of Remark 4.2) such that

$$\lambda_1(L_{0,-}) + \frac{c_*}{\sqrt{d}} \leq \lambda_1(L_{0,+}). \tag{4.49}$$

Using $|[Y]| \leq t_*$ and iterating (4.49) $k - 1$ times, we will be able to conclude (4.42).

The rest of the argument is devoted to the proof of (4.49). We denote $\lambda_+ := \lambda_1(L_{0,+})$ and $\lambda_- := \lambda_1(L_{0,-})$ We begin by applying Lemma A.8 to the matrix $H = L_0$ two times, first with $e = E_k$ and then with $e = E_y$. From (4.43) and (A.11), we see that $\lambda_\pm$ satisfy the self-consistent equations

$$\sqrt{d} = \left(\frac{1}{\lambda_+ - \underline{L}(T_k)}\right)_{x_k x_k} + \left(\frac{1}{\lambda_+ - M^{(k)}}\right)_{vv} \quad \text{and} \quad \sqrt{d} = \left(\frac{1}{\lambda_- - \underline{L}(T_k)}\right)_{x_k x_k} + \left(\frac{1}{\lambda_- - M^{(k)}}\right)_{yy}, \tag{4.50}$$

and both lie in the open interval

$$I := \Big(-\sqrt{d}, \, \min\big(\lambda_2(\underline{L}(T_k)), \mu\big)\Big).$$

Recall that μ was defined earlier as being the smallest eigenvalue of $M^{(k)}$. Let us introduce the functions

$$f, \tilde{f} : I \to \mathbb{R}, \quad f(s) = \left(\frac{1}{s - M^{(k)}}\right)_{yy}, \quad \tilde{f}(s) = \frac{1}{2\tilde{t}(s-\mu)}.$$

By definition of I and y (recall (4.44) with $c = 1/2\tilde{t}$), we find, for every $s \in I$,

$$\tilde{f}(s) = \sum_{\lambda \in \operatorname{Spec} M^{(k)}} \frac{|\mathbf{w}_\lambda(y)|^2}{s - \lambda} \le \frac{|\mathbf{w}_\mu(y)|^2}{s-\mu} \le \frac{1}{2\tilde{t}(\mu - s)} = f(s) \le 0. \tag{4.51}$$

We conclude that if $\tilde{\lambda} \in I$ satisfies the implicit equation

$$\sqrt{d} = \left(\frac{1}{\tilde{\lambda} - \underline{L}(T_k)}\right)_{x_k x_k} + \tilde{f}(\tilde{\lambda}),$$

then $\lambda_- \le \tilde{\lambda}$ (see illustration below).
Observe now that $\tilde{\lambda}$ only depends on the spectrum of the $\underline{L}(T_k)$ and on the integer $\tilde{t}$.

We deduce that there exists a constant $c_1 > 0$, depending only on $\mathbb{T}_{\le \tilde{t}}$ such that

$$\left(\tilde{\lambda} - \frac{c_1}{\sqrt{d}}, \tilde{\lambda} + \frac{c_1}{\sqrt{d}}\right) \subset I. \tag{4.52}$$

Indeed if s is too close to $-\sqrt{d}$ then the contribution of $\tilde{f}(s)$ becomes too large to reach for the equality to hold. In particular

$$\lambda_- \le \tilde{\lambda} \le \min\left(\lambda_2(\underline{L}(T_k)), \mu\right) - \frac{c_1}{\sqrt{d}}. \tag{4.53}$$

We now turn to the analysis of λ_+, defined as the solution to the left-hand side equation in (4.50). Either $\lambda_+ \in \left(\tilde{\lambda} + c_1 d^{-1/2}, \min\left(\lambda_2(\underline{L}(T_k)), \mu\right)\right)$, in which case (4.49) holds with $c_* = c_1$. Or

$$\lambda_+ \in I', \qquad I' := \left(-\sqrt{d} + C_1 d^{-1/2}, \min\left(\lambda_2(\underline{L}(T_k)), \mu\right) - c_1 d^{-1/2}\right).$$

Note that we used the same argument as before to prevent λ_+ from being too close to $-\sqrt{d}$.
Let us introduce the change of variable $\lambda_\pm = -\sqrt{d} + \frac{\theta_\pm}{\sqrt{d}}$. Similarly to what was done in the proof of Lemma 4.18 (see (4.32) and (4.33)), we find that $\theta_\pm$ solve the implicit equations

$$1 = f_{T_k, a_k}(\theta_+) + \left(\frac{1}{\theta_+ - (\sqrt{d}M^{(k)} + d)}\right)_{vv}, \quad \text{and} \quad 1 = f_{T_k, a_k}(\theta_-) + \left(\frac{1}{\theta_- - (\sqrt{d}M^{(k)} + d)}\right)_{yy}, \tag{4.54}$$

on the interval

$$I'' := (c_1, J - c_1), \qquad J := \min\left(\lambda_2(L(T_k), \sqrt{d}\mu + d)\right) = O(1),$$

Here we introduced the (universal) function

$$f_{T_k, a_k}(s) := \left(\frac{1}{s - L(T_k)}\right)_{x_k x_k}, \qquad s \in I''.$$

Recalling (4.46), we see that for every $s \in I''$, we have

$$0 \le \left(\frac{1}{(\sqrt{d}M^{(k)} + d) - s}\right)_{vv} = \sum_{\lambda \in \Lambda_1} \frac{|\mathbf{w}_\lambda(v)|^2}{(\sqrt{d}\lambda + d) - s} + \sum_{\lambda \in \Lambda_2} \frac{|\mathbf{w}_\lambda(v)|^2}{(\sqrt{d}\lambda + d) - s} \le \frac{C\tilde{t}}{d} + \frac{3}{\tau d} = O\left(d^{-1}\right). \tag{4.55}$$

Here we used (4.47) and the fact that

$$\inf_{\mu\in\Lambda_1}|\lambda_+-\mu|\geq c_1 d^{-1/2} \quad\Rightarrow\quad \inf_{\mu\in\Lambda_1}|\theta_+-(\sqrt{d}\mu+d)|\geq c_1,$$

to control the contribution of the sum over Λ_1.
Furthermore, recalling (4.51), we know that

$$\left|\left(\frac{1}{\sqrt{d}M^{(k)}+d-s}\right)_{yy}\right| = \tilde{f}\left(\sqrt{d}s+d\right)\geq \frac{1}{2\tilde{t}(s-J)}\geq c_3,\quad s\in I''. \tag{4.56}$$

where $c_3>0$ is a universal constant (it depends only on T_k and $\tilde{t}$).
Inverting (4.54) and using (4.55) and (4.56), we see that $\theta_\pm$ solve the equation

$$\theta_\pm = f^{-1}_{T_k,a_k}(1+\varepsilon_\pm),\qquad |\varepsilon_+-\varepsilon_-|\geq\frac{c_3}{2},\qquad \varepsilon_-\asymp 1,\ \varepsilon_+=O\left(d^{-1/2}\right). \tag{4.57}$$

On the interval I'', there exists $C>0$, that depends on f_{T_k,a_k} such that $\frac{1}{C}\leq f'_{T_k,a_k}\leq C$. Therefore, by the mean value theorem, we deduce that

$$\theta_- - \theta_+ \geq \frac{1}{(\varepsilon_- - \varepsilon_+)C}\geq c_4>0.$$

Since $c_4>0$ is a constant that only depends on c_1 and f_{T_k,a_k}, it is universal. We conclude that

$$\lambda_+\geq\lambda_- + \frac{c_4}{\sqrt{d}},$$

which shows (4.49) and concludes the proof. □

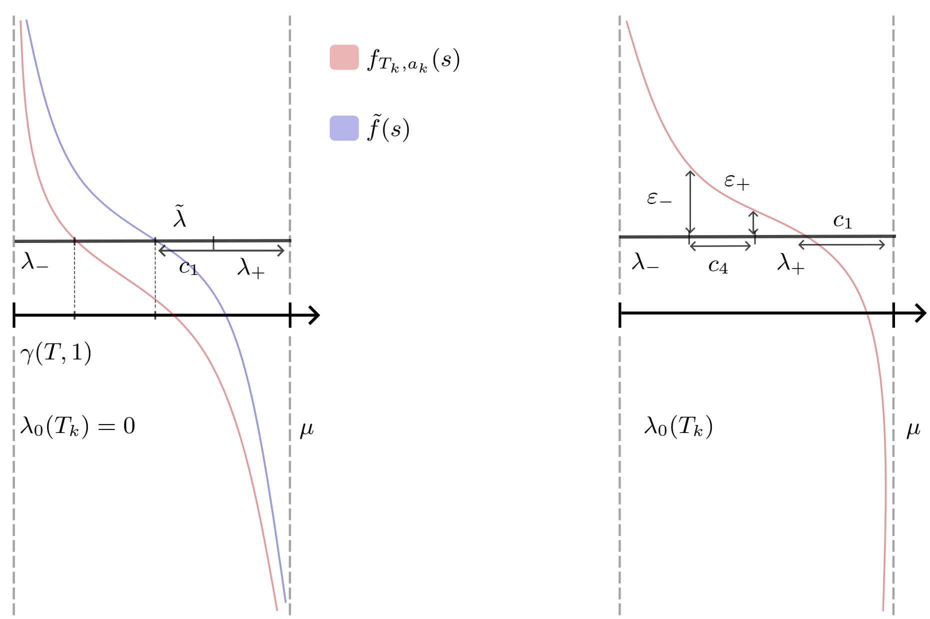


Figure 4.6: Illustration of (4.52) (left-hand scheme) and (4.57)

4.7 Quantitative estimates on trees and bouquets

Let

$$w(t) := t^{t-2} d^t \mathrm{e}^{-td}, \qquad t \in \mathbb{N}^*. \tag{4.58}$$

This quantity describes the probabilistic price to pay to obtain a tree of size t.

Let $n, r, \zeta \in \mathbb{N}^*$ and $\mathbf{t} = (t_1, \dots, t_n)$ where $t_i \in \mathbb{N}$ are fixed integers. Let

$$\mathcal{N}_{\mathbf{t}} := \left|\{T_i \in \mathcal{T}_{t_i}(\zeta) : \operatorname{dist}(T_1, T_i) \le r,\, t_i \in \mathbf{t}\}\right| \tag{4.59}$$

the number of bouquets.

The following heuristics gives us the number of bouquets we can expect to see at a given regime $1 \le d \le N$ (note that we only consider finite collections of finite trees). The price to see a tree of size $t \in \mathbb{N}$ can be approximated by the quantity $w(t)$ and the size of the ball of radius r is d^{r+1} (geometric series in d). Therefore the chance to see a bouquet of size $\mathbf{t}$ should be $d^{r+1} \prod_i w(t_i)$ and thus

$$\mathcal{N}_{\mathbf{t}} \sim N d^{r+1} \prod_i w(t_i).$$

In particular, the number of trees of size t should be approximately (up to factors d^r) equal to the number of bouquets of total size t. This heuristics is formalized in the next lemma.

Lemma 4.24 (Second moment)**.** *Then*

$$\begin{aligned} \mathbb{E}[|\mathcal{T}_t|] &= N w(t) t d \Big(1 + O(\frac{t^2 d}{N})\Big), \\ \mathbb{E}[|\mathcal{T}_t|^2] &= \mathbb{E}[|\mathcal{T}_t|]^2 \Big(1 + O(\frac{d t^2}{N})\Big) \end{aligned} \tag{4.60}$$

as well as

$$\begin{aligned} \mathbb{E}\big[\mathcal{N}_{\mathbf{t},r}\big] &= N d^{r+1} w(\mathbf{t}) 2 \big(1 + O(\dots)\big) \\ \mathbb{E}[\mathcal{N}_{\mathbf{t},r}^2] &= \mathbb{E}\big[\mathcal{N}_{\mathbf{t},r}\big]^2 \Big(1 + O\Big(\frac{t^2 r^2}{N}\Big)\Big) \end{aligned} \tag{4.61}$$

where $w(\mathbf{t}) := \prod_{t \in \mathbf{t}} w(t)$.

Proof of Lemma 4.24. Let $t, c \in \mathbb{N}$, $Y = \{y_1, \dots, y_t\} \subseteq [N]$, $x \in [N] \setminus Y$ and $T \in \mathbb{T}_t$. Let us abbreviate

$$U(Y, x, T) := \{\mathbb{G}|_Y = T,\, S_1(Y) \cap Y^c = \{x\}\}.$$

Then, abbreviating $U = U(Y, x, T)$ we find

$$T(U(Y,x,T)) = \Big(\sum_{i \in [t]} A_{y_i x}\Big) \prod_{(i,j) \in E(T)} A_{y_i y_j} \prod_{(i,j) \notin E(T)} (1 - A_{y_i y_j}) \prod_{y \in Y,\, z \in [N] \setminus Y} (1 - A_{zy}) \tag{4.62}$$

and recalling (4.58), we see that

$$\mathbb{P}(U(Y,x,T)) = t\big(\frac{d}{N}\big)^t \big(1 - \frac{d}{N}\big)^{Nt(1 - O(t/N))} = N^{-t} w(T) \big(1 + O(t^2 d / N)\big).$$

Let $t_i \ge 0$, $T_i \in \mathbb{T}_i$, $Y_i = \{y_1^i, \dots, y_{t_i}^i\} \subseteq [N]$, $i = 1, 2$ disjoint subsets and $x \in [N] \setminus (Y_1 \cup Y_2)$. Let

$$\begin{aligned} &U(\{Y_1, Y_2\}, \{x_1, x_2\}, \{T_1, T_2\}, r) := \\ &\{\mathbb{G}|_{Y_i} = T_i \colon\ S_i(Y_i) \cap [N] \setminus Y_i = \{x_i\} \colon\ i = 1, 2 \text{ and } \operatorname{dist}(x_1, x_2) \le r\} \end{aligned}$$

then similarly as before we can compute

$$w(T_1,T_2) := \mathbb{P}(U(\{Y_1,Y_2\},\{x_1,x_2\},\{T_1,T_2\},0)) = t_1t_2\big(\frac{d}{N}\big)^{t_1+t_2}\mathrm{e}^{-(t_1+t_2)d}\Big(1+O\big(\frac{(t_1+t_2)^2}{N}\big)\Big) \tag{4.63}$$

and a similar but slightly more involved computation gives

$$w(T_1,T_2,r) := \mathbb{P}(U(\{Y_1,Y_2\},\{x_1,x_2\},\{T_1,T_2\},r)) \tag{4.64}$$

$$= \big(\frac{d}{N}\big)^{r+t_1+t_2}t_1t_2\mathrm{e}^{-(t_1+t_2)d}d^{r+1}\Big(1+O\big(\frac{d(t_1+t_2)}{N}\big)\Big) \tag{4.65}$$

Using (4.58) and Cayley theorem for the number of trees on t vertices, we find

$$\mathbb{E}[|\mathcal{T}_t|] = \sum_U w(T) = Ndtw(t)\Big(1+O\big(\frac{1}{t}+\frac{t^2d}{N}\big)\Big).$$

To prove the second moment bound, we extend over all possible pairs (U,V) and use that conditioned on $E(U,V)$ the events $T(U)$ and $T(V)$ are independent. We have

$$\mathbb{E}[|\mathcal{T}_t|^2] = \sum_{U,V}\mathbb{P}[T(U)\cap T(V)] = \sum_{U,V}\mathbb{P}[T(U),T(V)|E(U,V)]\mathbb{E}[E(U,V)] =$$

$$\sum_{U,V}(Ndtw(t))^2+\Big(1+O\big(\frac{dt_1t_2}{N}\big)\Big) = \mathbb{E}[|\mathcal{T}_t|]^2\Big(1+O\big(\frac{dt^2}{N}\big)\Big)$$

Similarly, we compute the first and second moments of $\mathcal{N}_{t_1,t_2}(r)$ using (4.64) and the summation over all possible pairs is given by $w(t_1,t_2)$. □

Chapter 5

Graph geometry

5.1 General properties

We recall from Chapter 2 the notion of very high probability. Most of our statements do not in general hold with very high probability but we nevertheless often use this notion to bound very unlikely events. Very high probability bounds are thus mostly used within proofs.

Definition 5.1 (High and very high probability)**.** An event Ω holds with high probability if $\mathbb{P}[\Omega] \to 1$.
An event holds with very high probability if for every $\nu \geq 0$, there exists $\mathcal{C}_\nu \geq 0$ such that $\mathbb{P}[\Omega] \geq 1 - \mathcal{C}_\nu N^\nu$.

In this chapter we study properties of the Erdős-Rényi graph. In particular probabilistic properties. We define the *degree*, respectively the *normalized degree* of a vertex, as

$$D_x := \sum_{y \neq x} A_{xy}, \quad v_x := \frac{D_x - d}{\sqrt{d}}, \quad x \in [N].$$

Note the random variables D_x, $x \in [N]$, each follow a $\mathcal{B}_{N,d/N}$ distribution and are very weakly correlated

$$\mathbb{E}\big[\big(D_x - d\big)\big(D_y - d\big)\big] = \mathbb{E}\Big[\Big(A_{xy} - \frac{d}{N}\Big)^2\Big] = \mathrm{Var}(Z), \quad x, y \in [N], \tag{5.1}$$

where $Z \sim \mathrm{Bernoulli}(d/N)$.
For $Y \subseteq [N]$, we denote $A_{(Y)}$ as the $|Y|$-by-$|Y|$ matrix defined by $((A_{(Y)})_{xy})_{x,y\in Y} = (A_{xy})_{x,y\in Y}$. We also introduce the number of connections between two sites

$$E(Y, Z) = \sum_{y\in Y} \sum_{z \in Z} A_{yz}, \quad Y, Z \subset [N]$$

and write $E(Y) := E(Y, Y)$.

Lemma 5.2 (High degree probability)**.** *For any $\nu \geq 0$, there exists $\mathcal{C}_\nu \geq 0$ such that we have*

$$\mathbb{P}\big(\max_{x\in[N]} D_x \leq \Delta\big) \geq 1 - \mathcal{C}_\nu N^\nu, \quad \Delta := \Delta(d, N, \mathcal{C}_\nu) = \begin{cases} d + \mathcal{C}_\nu \sqrt{d \log N} & if d \geq \frac{1}{2}\log N \\ \mathcal{C}_\nu \frac{\log N}{\log\log N - \log d} & if d \leq \frac{1}{2}\log N. \end{cases} \tag{5.2}$$

Moreover for $\alpha \geq 0$, we have

$$\mathbb{P}[D_x \geq \alpha d] \leq \frac{2}{N} \mathrm{e}^{(\mathfrak{u}_+ - \alpha) d \log(\mathfrak{u}_+)}. \tag{5.3}$$

Proof. The random variables D_x, $x \in [N]$, each follow a $\mathcal{B}_{N,d/N}$ distribution. Using Lemma B.3 and a union bound gives

$$\mathbb{P}\big[\max_{x\in[N]} D_x \geq \Delta\big] \leq N\mathbb{P}\big[D_1 \geq \Delta\big] \leq N\mathrm{e}^{-dh(\Delta/d)},$$

where h is defined in (B.1). Now since

$$\frac{\Delta}{d} = \begin{cases} \mathcal{C}\big(\frac{\log N}{d}\big)^{1/2} & if d \geq \frac{1}{2}\log N, \\ \mathcal{C}\frac{\log N}{d\log(\log N/d)} & if d \leq \frac{1}{2}\log N, \end{cases}$$

we conclude that if $\mathcal{C}_\nu$ is large enough, the right-hand side of the above equation is bounded by $\mathcal{C}_\nu N^{-\nu}$. To obtain (5.3), we again apply Lemma B.3 and do a first order approximation of h. ☐

Remark 5.3. The bound (5.2) is of course very rough. As hinted at by Lemma B.4 and the very weak correlations between degrees (5.1), the law of extreme values of $(v_x\colon x \in [N])$ should follow the law extreme law of N independent normal distributions. In that case we expect $\max_{x\in[N]} v_x \sim \sqrt{2\log N}$ for $d \gg \log N$. On the other hand for $d \lesssim \log N$, we cna expect $\max_{x\in[N]} D_x$ to behave as the maximum of N independent $\mathcal{P}_d$ variables, for which an explicit formula is derived in Lemma B.7.

The following is a standard result about the connectivity of the Erdős-Rényi graph. This is proved for instance in [16, Theorem 7.3].

Lemma 5.4. *For $d \geq \log N$, the Erdős-Rényi graph is connected and its radius is $\frac{\log N}{\log d}(1 + o(1))$ with high probability.*
For $d < \log N$, the Erdős-Rényi is disconnected with high probability.

The connectivity properties of the Erdős-Rényi graph are not not crucial in our work. However some understanding of the behavior of small connected components is useful. In particular the next result, [5, Corollary A.15].

Lemma 5.5. *There exists $C > 0$ such that if $C \leq d \leq N$, the number of All small components of G have at most $O(\log d/N)$ vertices with very high probability. All small components of G are trees with high probability. The giant component of G has at least $N(1 - \mathrm{e}^{-d/4})$ vertices with high probability.*

An important property of the Erdős-Rényi graph is that it is relatively sparse and in particular contains few cycles in given neighborhoods. The next lemma states that there is a strong relation between the regime of d and the probability to find a cycle in some region of the graph. The result is found in [6, Lemma 5.5]

Lemma 5.6 (Few cycles in small balls)**.** *For $k, r \in \mathbb{N}$, $x \in [N]$, there is $C > 0$ such that*

$$\mathbb{P}\Big(|E(G|_{B_r(x)})| - |B_r(x)| + 1 \geq k|S_1(x)\Big) \leq \frac{1}{N^k}\big(C(d + |S_1(x)|)\big)^{2kr+k}(2kr)^{2k}. \tag{5.4}$$

A useful property of the Erdős-Rényi graph is that it enjoys some regularity in its growth. This a restatement of [6, Lemma 5.4]

Lemma 5.7 (Concentration of $S_i(x)$)**.** *Let $0 \leq d \leq N$, $x \in [N]$. then there are constants $C, c > 0$ such that*

$$\mathbb{P}\Big((1 - \varepsilon - C\mathcal{E}_i)d|S_i(x)| \leq |S_{i+1}(x)| \leq (1 + \varepsilon + C\mathcal{E}_i)d|S_i(x)|\Big|A_{(B_{i-1}(x))}\Big) \geq 1 - 2\exp\big(-cd|S_i(x)|\varepsilon^2\big) \tag{5.5}$$

on the event $|B_i(x)| \leq \sqrt{N}$, where $\mathcal{E}_i := \frac{d|S_i(x)|}{N} + \frac{1}{\sqrt{N}}$.

Proof. This is [6, (5.16)] the proof of which does not require any lower bound on D_x. ☐

The next lemma is an example of how we can accurately control the size of some set of vertices, for instance leaves. A similar proof can be made for any similar subset of vertices. The idea is always to use the small and explicit correlation between pairs of vertices.

Lemma 5.8. *Let $0 < d < N^{1/5}$,*

$$\mathbb{P}\Big[|\mathcal{L}| - \mathbb{E}|\mathcal{L}| \geq t\mathbb{E}|\mathcal{L}|\Big] = O\big(t^{-2}\big).$$

For $d \geq N^{1/5}$, $|\mathcal{L}| = 0$ with high probability.

Proof. The proof is based on a second-moment argument. We first compute the expectation of the number of leaves, using Lemma B.5 and introducing $Z \sim \mathcal{P}_d$, we have

$$\mathbb{E}|\mathcal{L}| = N\mathbb{P}(Z=1)\left(1 + O\left(N^{-1}\right)\right) = N\mathrm{e}^{-d}d\left(1 + O\left(N^{-1}\right)\right).$$

We can already conclude by a union bound that

$$\mathbb{P}[\mathcal{L} > 1] \leq N\mathrm{e}^{-d}d \leq N^{-3/5}, \qquad d > N^{1/5}.$$

The second moment of $|\mathcal{L}|$ can be computed as

$$\begin{aligned}
\mathbb{E}|\mathcal{L}|^2 &= N\mathbb{P}(Z=1)\left(1 + O\left(N^{-1}\right)\right) + \sum_{x \neq y} \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{D_x=1}\mathbf{1}_{D_y=1}|A_{xy}\right]\right] \\
&= N\mathrm{e}^{-d}d\left(1 + O\left(N^{-1}\right)\right) + N(N-1)\left(\frac{d}{N}\left(1 - \frac{d}{N}\right)^{2(N-1)} + \left(1 - \frac{d}{N}\right)\mathbb{P}(Z=1)^2\left(1 + O\left(N^{-1}\right)\right)\right) \\
&= N\mathrm{e}^{-d}d\left(1 + O\left(N^{-1}\right)\right) + N(N-1)\left(\frac{d}{N}\mathrm{e}^{-2d + O\left(d^2/N\right)} + \left(1 - \frac{d}{N}\right)d^2\mathrm{e}^{-2d}\left(1 + O\left(N^{-1}\right)\right)\right) \\
&= N^2\mathrm{e}^{-2d}d^2\Big(1 + O\Big(\frac{d^3}{N}\Big)\Big) = \Big(\mathbb{E}|\mathcal{L}|\Big)^2\Big(1 + O\Big(\frac{d^3}{N}\Big)\Big)
\end{aligned}$$

In the second equality we used the independence of D_x, D_y, $x \neq y$, conditioned on A_{xy} (this simply means that two vertices are only correlated if the edge between them is open).
We can conclude using Chebyshev's inequality

$$\mathbb{P}\Big(|\mathcal{L}| - \mathbb{E}|\mathcal{L}| \geq t\mathbb{E}|\mathcal{L}|\Big) \leq \frac{\mathbb{E}|\mathcal{L}|^2 - \big(\mathbb{E}|\mathcal{L}|\big)^2}{t^2\Big(\mathbb{E}|\mathcal{L}|\Big)^2} \leq 2t^2\Big(1 + O\Big(\frac{d^3}{N}\Big)\Big).$$

This concludes the proof. ☐

Lemma 5.9. *Let $\mathcal{L}$ be defined as the set of vertices of degree 1. Then*

$$\mathbb{P}\Big[|\mathcal{L}| - \mathbb{E}|\mathcal{L}| \geq t\mathbb{E}|\mathcal{L}|\Big] = O\big(t^{-2}\big).$$

Finally, we show a bound on the size of the adjacency matrix of a tree.

Lemma 5.10. *Let A be the adjacency matrix of a graph with maximal degree $q+1$, for some $q \geq 0$. Then $\|A\| \leq q+1$. Moreover if A is the adjacency of a tree then $\|A\| \leq 2\sqrt{q}$.*

Proof. The first claim is obvious by the Schur test on the operator norm or alternatively Gershgorin circle theorem. Let V, E be the vertex and edge sets of the graph in question. To prove the second claim, choose a root vertex o and denote by C_x the set of children of the vertex $x \in V$. Then for any vector $\mathbf{w} = (w_x)$ we have

$$\begin{aligned}
\big|\langle \mathbf{w}, A\mathbf{w}\rangle\big| &= \Big|\sum_{x,y} w_x A_{xy} w_y\Big| = 2\Big|\sum_{x \in V}\sum_{y \in C_x} w_x w_y\Big| \leq \sum_{x \in V}\sum_{y \in C_x}\left(\frac{1}{\sqrt{q}}w_x^2 + \sqrt{q}w_y^2\right) \\
&\leq \frac{q+1}{\sqrt{q}}w_o^2 + \sum_{x \neq o}\left(\frac{q}{\sqrt{q}}w_x^2 + \sqrt{q}w_x^2\right) \leq 2\sqrt{q}\sum_{x \in V} w_x^2.
\end{aligned}$$

Here we used Young's inequality in the third step and in the fourth step the fact that each vertex in the sum appears at most q times as parent. This concludes the proof. ☐

5.2 Exclusion principle

The analysis of extreme eigenvalues of the Erdős-Rényi graph is centered around the study of extreme events, such as high-degree vertices, low-degree vertices, leaves and trees. Such events are called low-probability events. The Erdős-Rényi graph enjoys a probabilistic property that makes its typical realizations well suited to study simultaneously those events. Indeed low probability configurations that happen simultaneously on the graph are with high probability far away from one another on the graph. We call this mechanism the *local exclusion* principle. For our use, it is important that this exclusion mechanism works well on some sufficiently large distance $r \geq 1$.

- For large-degree vertices, Lemma 5.12 as soon d is small enough, the local exclusion principle is applicable. There are two underlying mechanisms behind this fact: first we define large-degree vertices in such a way that their expected number is always smaller than $\sqrt{N}$. Second the radius of the graph grows as $\frac{\log N}{\log d}$.
- For small-degree vertices, Lemma 5.13, the situation is more complex since for values of d smaller than $\log N$, the expected number of small-degree vertices behaves as $N\mathrm{e}^{-d}$. Therefore the local exclusion principle only makes sense down to some regime $d^* \gg \frac{1}{2}\log N$.
- For small- and large-degree vertices in dense regimes, Proposition 5.16, we cannot hope to make their neighborhood disjoint. Indeed the radius of $\mathbb{G}$ is $\frac{\log N}{\log d}$ with high probability and thus if $d = N^{\varepsilon}$, $\varepsilon \geq \frac{1}{n}$, $n \geq 1$, it is impossible to have a local exclusion principle on radius n neighborhoods.

Remark 5.11 (Difference with [6])**.** In [5], the authors use the exclusion principle for distance of the order $r = O(\log N)$. We usually only require $r = O(1)$. This difference explains why perturbation theory makes sense for $\underline{L}$ for a wider range of regimes of d than is makes sense of $\underline{A}$.

Let us recall the definition of $\mathcal{U}^{\pm}(\varepsilon)$, $\varepsilon > 0$ from (3.12). We introduce two events, $\Xi^{\pm}(r, \varepsilon)$, $r \in \mathbb{N}^*$, $\varepsilon > 0$, that check if there exist some pair of elements of $\mathcal{U}^{\pm}$ that are close to each other in $\mathbb{G}$. Let

$$\begin{aligned} \Xi^{+}(r,\varepsilon) &:= \big\{\exists x, y \in \mathcal{U}^{+}(\varepsilon)\colon\ y \in B_{2r}(x)\big\}, \\ \Xi^{-}(r,\varepsilon) &:= \big\{\exists x, y \in \mathcal{U}^{-}(\varepsilon)\colon\ y \in B_{2r}(x)\big\}. \end{aligned} \tag{5.6}$$

Lemma 5.12 (Large vertices are not neighbors)**.** *Let $1 \leq d \leq (\log N)^2$. For $r \in \mathbb{N}^*$, $\varepsilon > 0$, we have*

$$\mathbb{P}\big(\Xi^{+}(r,\varepsilon)\big) = O(N^{-\eta}), \quad 0 < \eta < \varepsilon - \frac{(r+1)\log d}{\log N}. \tag{5.7}$$

Proof. Let us fix $\varepsilon > 0$ and $r \in \mathbb{N}^*$ and introduce

$$\Xi^{(k)} := \Big\{\exists x \in [N]\colon\ x \in \mathcal{U}^{+}(\varepsilon), |\mathcal{U}^{+}(\varepsilon) \cap B_r(x)| \geq k\Big\}, \quad k \in \mathbb{N}^*.$$

Using Lemma B.3 when $d \geq \log N$ and (5.3) for $d \leq \log N$ and using the fact that $|h(\alpha + t/d) - h(\alpha)| = O((t/d)^2)$, uniformly for $t \in \mathbb{R}$ we get, for $t \geq 0$,

$$\begin{aligned} \mathbb{P}\left(v_x \geq \mathfrak{u}^{+}(\varepsilon) - \frac{t}{\sqrt{d}}\right) &\leq \mathbb{P}\left(v_x \geq \mathfrak{u}^{+}(\varepsilon) - \frac{t}{\sqrt{d}}\right)\Big(\mathbf{1}_{d \geq \log N} + \mathbf{1}_{d \leq 2\log N}\Big) \\ &\leq N^{-(1+\varepsilon)/2}\mathrm{e}^{Ct^2/d} + \frac{2}{N}\mathrm{e}^{d\log(\mathfrak{u}_+)(\frac{1-\varepsilon}{2})}\mathrm{e}^{t\log(\mathfrak{u}_+)} \end{aligned} \tag{5.8}$$

for some constant $C > 0$.
For $Y \subseteq [N]$ fixed, let us define

$$\delta_y(Y) := \sum_{z \in Y}\left(A_{yz} - \frac{1}{N}\right), \quad y \in Y. \tag{5.9}$$

Then conditioned on $A_{(Y)}$, the quantities $\{D_y - \delta_y(Y)\colon\ y \in Y\}$ are independent. We have, for $n \in \mathbb{N}^*$,

$$\mathbb{P}\Big[v_x \geq \mathfrak{u}^+(\varepsilon), v_{y_i} \geq \mathfrak{u}^+(\varepsilon),\, i \in [k]\Big] \leq \mathbb{P}\Big[v_x \geq \mathfrak{u}^+(\varepsilon) - \frac{k+n}{\sqrt{d}}\Big]^{(k+1)} + \mathbb{P}\Big[\max_{i\in[k]} \delta_{y_i} \leq n\Big].$$

For $k \leq (\log N)^2$, we have

$$\mathbb{P}\big[\Xi^{(k)}\big] \leq N\bigg(\frac{d^{(r+2)}-1}{d-1}\bigg)^k \bigg[\bigg(\mathbb{P}\Big(v_x \geq \mathfrak{u}^+(\varepsilon) - \frac{k+n}{\sqrt{d}}\Big)\bigg)^{k+1} + \bigg(\frac{(k+1)d}{N}\bigg)^n\bigg]. \tag{5.10}$$

Now setting $k = 1$ and $n \leq 2$ we find, setting $t = 3$ in (5.8) and taking the union bound over all possible open paths between x and y_1, we find

$$\mathbb{P}[\Xi^+(r)] \leq \mathbb{P}[\Xi^{(1)}] \leq CNd^{r+1}\left[N^{-1-\varepsilon'}(1+O(1/d)) + \frac{d^2}{N^2}\right] = O(N^{-\eta}), \tag{5.11}$$

for $\eta \leq \varepsilon - \frac{(r+1)\log d}{\log N}$. □

Note that from (5.10), we could derive results for $r \gg 1$ if we allowed k to be greater than 1. This is what is done for instance in [6, Section 5].

Lemma 5.13 (Small vertices)**.** *For $c > 0$, there exists $0 < \eta < \varepsilon < c$ such that if $(\frac{1}{2}+c)\log N \leq d \leq (\log N)^2$,*

$$\mathbb{P}\big(\Xi^-(r,\varepsilon)\big) = O(N^{-\eta}). \tag{5.12}$$

Moreover for $\varepsilon > 0$ there exists a constant $K := K(\varepsilon, r) \geq 0$ such that if $d \geq \frac{1}{2}\log N + K\log\log N$ then

$$\mathbb{P}\big(\Xi^-(r,\varepsilon)\big) = O(\mathrm{e}^{-K\log\log N/2}) \tag{5.13}$$

Proof. Let us fix $0 < \varepsilon < \frac{1}{2}$ and abbreviate $\mathcal{U}_\varepsilon^- = \mathcal{U}$ and $d \geq \frac{1}{2}\log N + \phi$ for some $\phi \geq 0$ to be set later. Let us first deal with the case $d \geq (\frac{1}{2}+c)\log N$. Then by Lemma B.3 and since $\varepsilon^{-1} \leq \varepsilon\sqrt{d}$ we have

$$\begin{aligned}\mathbb{P}[x \in \mathcal{U}] &\leq \max\big(\mathbb{P}[v_x \leq \sqrt{(1+\varepsilon)\log N}], \mathbb{P}[v_x \leq \varepsilon\sqrt{d}]\big)\\ &= \max\big(-\mathrm{e}^{(1-\varepsilon)^2 d} \vee N^{-\frac{1+\varepsilon}{2}} = O(N^{-\frac{1+\varepsilon}{2}})\big)\end{aligned} \tag{5.14}$$

for $\varepsilon < 1 - \frac{1}{1+2c}$ and where we chose $\varepsilon' < \varepsilon$ in the last step. We can use (5.11) directly, with $k = 1$, without accounting for extra error terms.

$$\mathbb{P}(\Xi^-(r)) \leq Nd^{(r+1)}N^{-1-\varepsilon'} = O(N^{-\eta})$$

for $\eta < \varepsilon' < \varepsilon$. As we are looking at small vertices, removing the δ_i defined in (5.9) does not increase the probability; small degree events are positively correlated. Therefore Since ε' can be made as close as possible to ε by choosing C_1 large, we conclude (5.12).

Turning to (5.13) we first suppose without loss of generality that $\frac{1}{2}\log N + \frac{2}{\varepsilon} \leq d \leq \frac{2}{3}\log N$ and set $\phi := d - \frac{1}{2}\log N$. Let us recall that in that regime $\mathfrak{u}_- = \frac{1}{d}$ for $d \leq \log N$ and that in this regime the condition for $x \in \mathcal{U}^-$ becomes $D_x \leq k_* := \varepsilon d \wedge \frac{\log d}{\varepsilon(d-\frac{1}{2}\log N)\vee 1}$.

Let $Y \sim \mathrm{Poisson}(d)$ and $k_* := \lceil \varepsilon^{-1} \rceil$, for some $c' > 0$. Clearly $k_* \leq (1-c)d$ for some $c > 0$. We can thus use Lemma B.5 and Stirling's formula to find that

$$\begin{aligned}\mathbb{P}\big[D_x \leq \frac{1}{\varepsilon}\big] &\leq \frac{k_*}{d}\mathbb{P}[Y = k_*](1+O(d^2/N)) \leq \mathrm{e}^{-d}\mathrm{e}^{k_*(1+\frac{1}{k_*}+\log(d/k_*))}\left(1+O\left(\frac{1}{k_*}\right)\right)\\ &\leq 2\exp\left(-\frac{1}{2}\log N - \phi + C\frac{\log d}{\varepsilon}\right).\end{aligned} \tag{5.15}$$

for some $C := C(\varepsilon) \geq 0$. Using (5.11) on more time we get

$$\mathbb{P}\big(\Xi^-(r,\varepsilon)\big) \leq N d^{r+2} \left(\mathbb{P}\left[D_x \leq \frac{1}{\varepsilon}\right]\right)^2 = \exp\left((r+2)\log d - \phi + C\frac{\log d}{\varepsilon}\right) \leq O\left(\mathrm{e}^{-K \log\log N}\right)$$

where in the last step we set $\phi \geq 2K(r,\varepsilon)\log\log N$. This concludes the proof. ☐

5.3 Pruning in dense regime

In this section, we focus on the high-density regimes, i.e. $N^{\varepsilon} \leq d \leq N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon > 0$. In the rest of this section we write $\mathcal{U}(\tau) := \mathcal{U}^+(\tau) \cup \mathcal{U}^-(\tau)$ and denote by a superscript τ the objects that relate to the pruned graph $\mathbb{G}^\tau$. We first prove a basic separation result for regimes $d \ll N^{1/2}$.

Proposition 5.14. *Let $\varepsilon > 0$, and $1 \leq d \leq N^{\frac{1}{2}-\varepsilon}$. Then for $2 - \frac{\varepsilon}{2} < \eta < 2$, the balls $\Big(B_2(x) : x \in \mathcal{U}(\eta)\Big)$ are disjoint with probability $1 - O\Big(N^{-\varepsilon}\Big)$.*

Proof. We find that

$$\mathbb{P}\Big[\exists x \in \mathcal{U}(\eta),\, y \in B_1(x) \cap \mathcal{U}(\eta)\Big] \leq CNd^2\Big(\mathbb{P}\big(x \in \mathcal{U}(\eta)\big)\Big)^2 \leq CN^{1+(1-2\varepsilon)-(2-\varepsilon/2)} = O\Big(N^{-\varepsilon}\Big).$$

☐

For $d \ll N^{1/3}$, we can construct prune $\mathbb{G}$ so as to obtain stronger separation properties. We can construct a graph $\mathbb{G}^\tau$ such that on this graph vertices in $\mathcal{U}^+(\tau) \cup \mathcal{U}^-(\tau)$ are at a distant at least 3 from each other and their neighborhoods do not overlap too much.

Proposition 5.15. *Let $\varepsilon > 0$, $\tau \in (1 - \varepsilon/2, 1)$ and $N^{\varepsilon} \leq d \leq N^{\frac{1}{3}-\varepsilon}$. There exists $\eta > 0$ such that with probability $1 - O\big(N^{-\eta}\big)$ there is a graph $\mathbb{G}^\tau$ that satisfies the following conditions*

(i) the balls $\big(B_1(x)\colon x \in \mathcal{U}(\tau)\big)$ are disjoint trees and

$$S_1^\tau(x) \cap S_2^\tau(y) = \emptyset, \qquad x, y \in \mathcal{U}(\tau),\, x \neq y. \tag{5.16}$$

(ii) Each edge in $\mathbb{G} \setminus \mathbb{G}^\tau$ is incident to at least one vertex in $\mathcal{U}(\tau)$ and

$$\max_{x \in E(\mathbb{G}\setminus\mathbb{G}^\tau)} D_x = O(1).$$

(iii) For each $x \in \mathcal{U}(\tau)$

$$\frac{1}{d}\sum_{y \in S_1^\tau(x)} v_y^{(\tau)} = O\big(d^{-1/2}\big)$$

Proof. For (5.16), we use Lemma 5.2 to find

$$\mathbb{P}\Big[\exists x, y \in (\tau)\colon\ |S_1(x) \cap S_2(y)| \geq k\Big] \leq C_1 N^{-\varepsilon}\Big(\frac{C_2 d^3}{N}\Big)^k + C_2 N^{-2}$$

for $C_1, C_2 \geq 0$ coming from (5.17) and Lemma 5.6 respectively. For $k \geq \frac{1}{\tau}$, the right-hand side is bounded by $O\big(N^{-\eta}\big)$, for $\eta > 0$ small enough. If $u \in S_1(x) \cap S_2(y)$, we add the edge (xu) to E^τ. By the previous equation, we see that with probability $1 - O\big(N^{-\eta}\big)$, there are at most $O(1)$ such edges for each $x \in \mathcal{U}(\tau)$.

For (iii), we observe that conditioned on $S_1(x)$ the random variable $\frac{1}{d}\sum_y v_y^{(\tau)}$ follows the same law as $\frac{1}{d^{3/2}}\big(Z-d^2\big)$ where $Z \sim \mathcal{B}_{d(N-|S_1(x)|),d/N}$. Therefore using Lemma B.3, we find

$$\mathbb{P}\left[\frac{1}{d}\sum_y v_y^{(\varepsilon)} \geq t\right] \leq 2\mathrm{e}^{-d^2t^2/2}.$$

Setting $t = 1/\sqrt{d}$ yields a bound $O\big(N^{-1}\big)$. □

For $d \ll N^{1/6}$, we obtain even stronger separation properties that allow us to prove the block diagonal decomposition of $\underline{L}$.

Proposition 5.16 (Pruning). *Let $\varepsilon > 0$, $\tau \in (2-2\varepsilon, 2)$ and $(\log N)^{1+\varepsilon} \leq d \leq N^{\frac{1}{6}-\varepsilon}$. Then there exists $\eta > 0$, such that with probability $1 - O\big(N^{-\eta}\big)$ there exists a graph G^τ such that*

- *(i) The balls $\big(B_3^\tau(x)\colon\ x \in \mathcal{U}(\tau)\big)$ are disjoint trees.*
- *(ii) Each edge in $\mathbb{G}\setminus\mathbb{G}^\tau$ is incident to a least one vertex in $\mathcal{U}(\tau)$ and*

$$\max_{x\in E(\mathbb{G}\setminus\mathbb{G}^\tau)} D_x = O(1).$$

- *(iii) For each $x \in \mathcal{U}(\tau)$,*

$$\big|\big\{y \in B_2^\tau(x)\colon\ |v_x| \geq (\log N)^{1/4}\big\}\big| = O\big(\sqrt{\log N}\big).$$

Proof. Let $E^\tau \subseteq E(\mathbb{G})$ be the set of edges that will be removed to create $\mathbb{G}^\tau$. By Lemma B.3, we find

$$\mathbb{P}\big(x \in \mathcal{U}(\tau)\big) \leq 2N^{-\frac{\varepsilon}{2}\left(1-\sqrt{\log N/d}\right)}, \quad x \in [N], \tag{5.17}$$

and so, following the same logic as in the proof of Lemma 5.12, for $k \geq 0$

$$\mathbb{P}[\exists x \in \mathcal{U}(\tau) : |\mathcal{U}(\tau)\cap B_2(x)| \geq k] \leq Nd^{6k}\left[\exp\Big(-\frac{(k+1)\tau}{2}\log N\Big(1-\frac{k+n}{\sqrt{d\tau\log N}}\Big)\Big) + k\exp\Big(-\frac{n^2N}{dk}\Big)\right],$$

holds for any $n \geq 0$. Setting $n = 10$ and $k \geq \frac{10}{\tau}$, the right-hand side is bounded by $O\big(N^{-1}\big)$. Therefore by deleting at most $O(1)$ edges per vertex, the balls $B_1(x)$, for $x \in \mathcal{U}(\tau)$ can be made disjoint. We add those edges to E^τ.

We will now prove that with high probability,

$$\sum_{y\in S_1(x)} |S_1(x)\cap S_1(y)| = O(1), \quad x \in \mathcal{U}(\tau)$$

By Lemmas 5.2 and 5.6 and (5.17), we see that there exists $C \geq 0$ such that, for $r, k \geq 0$,

$$\begin{aligned}\mathbb{P}\Big[\exists x \in \mathcal{U}(\tau)\colon\ |E(B_r(x))| - |V(B_r(x))| - 1 \geq k|S_1(x)\Big]\\ \leq C_1N^{-\frac{\tau}{2}\left(1-\sqrt{\log N/d}\right)}\frac{\big(2d+C_2\log N\big)^{2kr+k}k^{2k}}{N^{k-1}} + C_2N^{-2},\end{aligned}$$

for $C_1, C_2 \geq 0$ coming from (5.17) and Lemma 5.6 respectively. For $r = 1$ and $k \geq \frac{1}{2\varepsilon}$ we see that the right-hand side is bounded by $O\big(N^{-\eta}\big)$ for $\eta > 0$ small enough. This shows that the cycles in the balls $B_1(x)$, $x \in \mathcal{U}(\tau)$, can be removed by deleting at most $O(1)$ from every $x \in \mathcal{U}(\tau)$. We add those edges to E^τ.

The proof of (iii) follows from Lemma B.3 as we find

$$\mathbb{P}\Big[x \in \mathcal{U}(\tau)\colon\ y_1,\dots,y_k \in B_2(x),\ |v_{y_i}| \geq (\log N)^{1/4},\ i \in [k]\Big] \leq C_1d^3N^{1-\frac{\tau}{2}}\mathrm{e}^{\frac{1}{2}k\sqrt{\log N}\left(1-\sqrt{\log N/d}\right)},$$

for some constant $C_1 \geq 0$. For $k \geq C_2\sqrt{\log N}$, the right-hand side is bounded by $O\big(N^{-\eta}\big)$, for $\eta > 0$ small enough.

The graph $\mathbb{G}^\tau$ is obtained from the graph $\mathbb{G}$ by removing every edge in E^τ. This concludes the proof. □

Remark 5.17. From the proof of Proposition 5.16 it becomes evident why the upper bounds on d are introduced. The bound $d \ll N^{1/3}$ is necessary if we want the graph to be sparse enough. Intuitively we want a distance of 3 between vertices of $\mathcal{U}(\tau)$. But the radius of an Erdős-Rényi with parameter d/N is about $\log N/\log d$. Therefore at $d = N^{1/3}$ we would have a radius precisely equal to 3.
The bound $d \ll N^{\tau/4}$ is necessary to prevent the vertices of $\mathcal{U}(\tau)$ to be too numerous.

5.4 Neighborhood of large-degree vertices

In this section, we analyze the properties of the graph $\mathbb{G}$ in the neighborhood of large vertices. Before stating the technical results that we use in Chapter 3, we give some intuition about the mechanisms we exploit.

Let $X \sim \mathcal{P}_d$. As we mentioned in the introduction of Chapter 3, the Poisson distribution of parameter N can be to some extend well-approximated by a normal law. Indeed, using Stirling's formula, for $k \gg 1$,

$$\mathbb{P}(X = k) = \exp\bigl(k \log d - \log k! - d\bigr) = \exp\left(k \log(d/k) - k - d + \frac{1}{2}\log(2\pi k)(1 + o(1))\right)$$

For $k = d + a\sqrt{d}$, this becomes

$$\begin{aligned}\mathbb{P}(X = k) &= \exp\left(k \log\left(1 + ad^{-1/2}\right) + k - d\right)\\ &= \exp\left(-(d + a\sqrt{d})\left(\frac{a}{\sqrt{d}} - \frac{a^2}{2d} + O\left(\frac{a^3}{d^{3/2}}\right)\right) + a\sqrt{d}\right) = \exp\left(-\frac{a^2}{2} + O\left(\frac{a^3}{\sqrt{d}}\right)\right).\end{aligned}$$

As long as $a^6 \ll \sqrt{d}$, we can morally think of X_0 as a $\mathcal{N}(0,1)$ variable. However, it is clear that the largest-degree vertices of $\mathbb{G}$ do not satisfy this assumption when $d \lesssim \log N$. Indeed, as Lemmas 5.2 and B.7 show, the maximal degree of $\mathbb{G}$ stays is of order $\frac{\log N}{\log(\log N/d)}$, which is much larger than $O(d)$ for small values of d.

However, we can use the sparsity of the Erdős-Rényi graph to say that if the degree of the vertex x is very large (meaning $x \in \mathcal{W}^+(\kappa)$), then the region of $\mathbb{G}$ around x has exhausted its "rare event potential". This is a similar idea that underlines the local exclusion principle introduced in Section 5.2. Therefore the other quantities in $B_r(x)$, $r \geq 1$, that we would be interested in, might be close to their expectation. In particular, we are often interested in the statistics of the sphere of radius 2 around x,

$$\beta_x := \frac{|S_2(x)|}{|S_1(x)|d} - 1, \qquad x \in [N].$$

Then it is clear that conditioned on $|S_1(x)|$, the random variable $d|S_1(x)|\beta_x$ follows a $\mathcal{P}_{|S_1(x)|d}$ distribution. Since $B_r(x)$ has exhausted its "rare event potential", whenever $x \in \mathcal{W}^+(\kappa)$, we can then argue that the normal approximation of Poisson variables is valid for β_x. This idea is used the present section (see Propositions 5.19 and 5.21) to control roughly the statistics of observables in the neighborhood of $B_r(x)$, $x \in \mathcal{W}^+(\kappa)$. The idea is pushed even further in Section 5.6 where the variable β_x is very closely approximated by a $\mathcal{N}(0,1)$ variable.

Let us first make some *a priori* observations. The set $\mathfrak{u}^+(\varepsilon)$ is defined so that $\mathbb{P}[x \in \mathfrak{u}^+(\varepsilon)] \ll N^{-1/2}$. On the other hand, we get directly from (5.3) and from the definition of $\mathcal{W}^+(\kappa)$ in (3.16)

$$\mathbb{P}\bigl[x \in \mathcal{W}^+(\kappa)\bigr] \leq \begin{cases} \frac{2\log N}{N}, & \text{if } d \geq (\log\log N)\log N, \\ \frac{2}{N}\mathrm{e}^{2\kappa\log(\mathfrak{u}_+)} & \text{if } d \leq (\log\log N)\log N. \end{cases} \tag{5.18}$$

In particular we can always write $\mathbb{P}\bigl[x \in \mathcal{W}^+(\kappa)\bigr] \leq \frac{2}{N}\mathrm{e}^{2\kappa\log(\mathfrak{u}_+)}$.

Lemma 5.18 (Behavior of β_x). *For any $\varepsilon > 0$, $1 \leq d \leq N^{\frac{1}{2}-\varepsilon}$, we have with very high probability,*

$$(\beta_x - 1) = \mathcal{C}\begin{cases} \frac{\log(\log N/d)}{\sqrt{d}}, & d \leq \log N \\ \frac{\sqrt{\log N}}{d}, & d \geq \log N, \end{cases} \qquad x \in \mathcal{U}^+(\varepsilon). \tag{5.19}$$

Moreover for κ as in (3.15) and $\alpha \in (0,1)$, there exists $c > 0$ such that,

$$(\beta_x - 1) = O\Big(\frac{v_x^2}{\alpha_x \sqrt{d} \log N}\Big), \quad x \in \mathcal{W}^+(\kappa), \tag{5.20}$$

holds with probability $1 - O\big(\mathrm{e}^{-c\sqrt{\log N}/2(\log\log N)^3}\big)$.

Proof. The first clain is obtained by Lemma 5.7 and union bound. We find

$$\mathbb{P}[\exists x \in \mathcal{U}^+(\varepsilon),\, (\beta_x - 1) \geq t] \leq N^{1/2} \mathrm{e}^{-cd\mathfrak{u}^+(\varepsilon)t^2}, \quad t \geq 0.$$

for some constant $c > 0$. Choosing t as in the right-hand side of (5.19) and using the definition of $\mathfrak{u}^+(\varepsilon)$, we conclude.

To prove (5.20), we use again Lemma 5.7 but this time with the bound (5.3). Let $x \in \mathcal{W}^+(\kappa)$. If $d \geq (\log\log N)\log N$, then by Lemma 5.2, $\frac{v_x^2}{\alpha_x\sqrt{d}\log N} \leq Cd^{-1/2}$ with very high probability, for $C \geq 0$ large enough. A direct application of (5.3) gives

$$\begin{aligned}\mathbb{P}[\exists x \in \mathcal{W}^+,\, (\beta_x - 1) \geq C\frac{\alpha_x v_x^2}{\sqrt{d}\log N}] &\leq \sum_{x\in[N]} \mathbb{P}(x \in \mathcal{W}^+(\kappa))\mathbb{P}\Big[(\beta_x - 1) \geq C\frac{\alpha_x v_x^2}{\sqrt{d}\log N}\Big| S_1(x)\Big] \\ &\leq 2\mathrm{e}^{\log\log N - cd/C^2} \leq \mathrm{e}^{-cd/2},\end{aligned}$$

for some universal $c > 0$.
If $d \leq (\log\log N)\log N$, $\frac{v_x^2}{\alpha_x\sqrt{d}\log N} \leq \frac{1}{\sqrt{d}\log(\log N/d)}$. By Lemma 5.7 we find

$$\mathbb{P}\Big[\exists x \in \mathcal{W}^+,\, (\beta_x - 1) \geq C\frac{v_x^2}{\alpha_x\sqrt{d}\log N}\Big] \leq \exp\left(\kappa\log(\log N/d) - \frac{c(d\mathfrak{u}_+ - \kappa)}{(\log(\log N/d))^2}\right),$$

for some universal $c > 0$. The right-hand side is bounded $O\big(\mathrm{e}^{-c\sqrt{\log N}/2(\log\log N)^3}\big)$ (here we use that $d\mathfrak{u}_+ \geq \sqrt{\log N}(\log\log N)^{-1}$) □

Proposition 5.19 (Neighborhood of $x \in \mathcal{U}^+$)**.** *For any $r \in \mathbb{N}^*$ and $0 < \varepsilon < 1/2$. If $r = O(1)$, there exists $K, c_* > 0$ such that for $0 < \eta < \varepsilon/2$ if*

$$K \leq d \leq (\log N)^2 \tag{5.21}$$

the following holds with probability $1 - O(N^{-\eta})$. Define

$$\psi_+ := \sqrt{\log N} \vee \frac{1}{2}\frac{\log N}{\sqrt{d}\log(\log N/d)} \tag{5.22}$$

(i) The balls $\big(\mathbb{G}|_{B_{r+2}(x)} :\ x \in \mathcal{U}^+(\varepsilon)\big)$ are disjoint trees.

(ii) For each $x \in \mathcal{U}^+(\varepsilon/2)$, for every $y \in B_r(x) \setminus \{x\}$,

$$v_x - v_y \geq c_*\psi_+, \quad |v_y| \leq \psi_+.$$

(iii) For each $x \in \mathcal{U}^+(\varepsilon/2)$, and $\alpha \in (0,1)$,

$$\big|\big\{y \in B_r(x) | |v_y| \geq (\log N)^\alpha\big\}\big| \leq C(\log N)^{1-2\alpha}.$$

(iv) For each $x \in \mathcal{U}^+(\varepsilon/2)$,

$$\left|\frac{|S_{i+1}(x)|}{d|S_i(x)|} - 1\right| = O\bigg(\Big(\frac{\log N}{d|S_i|}\Big)^{-1/2}\bigg), \quad \left|\frac{|S_i(x)|}{D_x d^{i-1}} - 1\right| = O\bigg(\Big(\frac{\log N}{dD_x}\Big)^{-1/2}\bigg), \quad 1 \leq i \leq r. \tag{5.23}$$

(v) For each $x \in \mathcal{U}^+(\varepsilon/2)$,

$$\sum_{y \in S_i(x)} v_y^2 = O\Big(D_x d^{i-1} (\log\log N)^2\Big), \quad 1 \le i \le r.$$

Proof. We work on the event $\Xi^+(r+2,\varepsilon)$. Observe that $\mathfrak{u}^+(\varepsilon)$ defined (3.12) and ψ_+ defined in (5.22) are analogous. By (5.2), (5.8) and Lemma 5.6, there exists $C > 0$, such that

$$\begin{aligned}
&\mathbb{P}[\exists x \in \mathcal{U}^+(\varepsilon),\, \mathbb{G}|_{B_{r+2}(x)} \text{ contains a cycle}] \\
&\le \sum_{x \in [N]} \mathbb{P}\Big(|E(G|_{\mathbb{G}|_{B_{r+2}(x)}})| - |\mathbb{G}|_{B_{r+2}(x)}| + 1 \ge k|S_1(x)\Big)\mathbb{P}(x \in \mathcal{U}^+(\varepsilon)) \\
&\le \bigg(N^{\frac{1-\varepsilon}{2}} + \frac{2}{N}\exp\Big(d\log(\mathfrak{u}_+)\Big(1 - \sqrt{\frac{1-\varepsilon}{2}}\Big)\Big)\bigg)\frac{\big(Cr(d+\log N)\big)^{4r}}{N} = O(N^{-\eta}),
\end{aligned}$$

for $\eta < 1$. This proves that there are no cycles in $\mathbb{G}|_{B_{r+2}(x)}$ on a high probability event. This proves (i).
For (ii) let us fix $\tilde{\varepsilon} \in (0,\varepsilon)$. Then there exists $c > 0$ such that we have

$$\mathfrak{u}^+(\varepsilon) - \mathfrak{u}^+(\tilde{\varepsilon}) > c\psi_+.$$

Moreover by Lemma 5.12, $\Xi^+(r+2,\tilde{\varepsilon})$ holds with high probability. We conclude the first assertion. For the second assertion, observe that if $c > 0$, a reasoning analogous to (5.11) yields

$$\mathbb{P}\Big[\exists x \in \mathcal{U}^+(\varepsilon),\, \exists y \in B_r(x) : |v_y| \ge (1-c)\psi_+\Big] \le CNd^{r+2}N^{-\frac{1+\varepsilon}{2}}N^{-\frac{1-\varepsilon'}{2}} = Cd^{r+2}N^{-\frac{\varepsilon-c}{2}} = O(N^{-\eta/2}),$$

for c and η small enough. This proves (ii).
Similarly for (iii), we find that for $k \in \mathbb{N}^*$ and $t \ge 0$, we have

$$\mathbb{P}[\exists\{y_1,\dots,y_k\} \subseteq B_r(x) : |v_{y_i}| \ge (\log N)^\alpha |S_1(x)] \le CD_x d^{r-1}\mathrm{e}^{-ck(\log N)^{2\alpha}} + \mathbb{P}\Big(\big|\mathbb{G}|_{B_r(x)}\big| \ge CD_x d^{r-1}\Big).$$

The second term on the left-hand side can be made smaller than $O(N^{-\nu})$ by choosing $C = \mathcal{C}_\nu$ large enough and using Lemma 5.7 inductively on $1 \le i \le r$. The first term is smaller than $N^{-\eta}$ as soon as $k = C(\log N)^{1-2\alpha}$ for $C > 0$ large enough.
(iv) follows from [6, Lemma 5.4]. Indeed for $\nu = 2$, there exists $\mathcal{K}_2 \ge 0$ such that on the event

$$\big\{\mathcal{K}_2 \log N/d \le D_x \le \sqrt{N}(2d)^{-r}\big\}$$

then (5.23) hold with probability $O(N^{-2})$. Choosing K in 5.21 large enough, this condition is always satisfied for $x \in \mathcal{U}^+(\varepsilon)$. This prove (iv).
For the last point, let us work only one the event defined by all previous points. Let us fix $x \in \mathcal{U}^+(\varepsilon)$. Then $B_r(x)$ is a tree on which (5.23) hold. Let us define

$$\mathcal{N}_{i,k}(y) := \big\{y \in S_i(x) \colon\, d^2\mathrm{e}^k \le (D_y - d)^2 \le d^2\mathrm{e}^{k+1}\big\}.$$

Using Bennett's inequality we find that if $\ell_{i,k}(t) := \frac{|S_i(x)|+t}{d}(\mathrm{e}^{k/2} \wedge \mathrm{e}^k)$, then for $1 \le j \le r$ and

$$\begin{aligned}
&\mathbb{P}\Big[|\mathcal{N}_{i,k}| \le \ell_{i,k}\, \forall i \le j,\, k \in \mathbb{Z} \text{ with } k \le \log\log N \big| S_1(x)\Big] \\
&\ge \big(1 - r\log\log N \mathrm{e}^{|S_i(x)| - |S_i(x)| - t}\big) \ge 1 - O\big(r\log\log N \mathrm{e}^{-t}\big).
\end{aligned}$$

By (ii), $|v_y| \le C\sqrt{\log N}$ for every $y \in B_r(x)$, $y \ne x$ and for some $C > 0$ large enough. We can thus set $t = 2\log N$ and condition on $|S_1(x)|$ to find

$$\begin{aligned}
\sum_{y \in S_i(x)} (D_y - d)^2 &\le d|S_i(x)| + \sum_{k=-\lfloor \log d \rfloor}^{\lceil \log \psi \rceil} d^2\mathrm{e}^{k+1}\mathcal{N}_{i,k} \\
&\le d|S_i(x)| + (d|S_i(x)| + \log N)(\log d + \log\psi) \\
&\le d|S_i(x)|\bigg(1 + \frac{\log N}{d|S_i(x)|}\bigg)(\log d + \log\psi) \le CD_x d^i \log\log N
\end{aligned}$$

with probability greater than $1 - N^{-2}$. Taking a union bound over $\mathcal{U}^+$ we conclude. This concludes the proof. ☐

We now derive better estimates that are valid in the neighborhood of $\mathcal{W}^+(\kappa)$. We need the following technical result which is [7, Lemma 9.1].

Lemma 5.20. *Let $x \in [N]$, and $X_y := (N_y(x) - d)^2 - d$, for any $x \in [N] \setminus \{x\}$. Then the following holds*

(i) For any $\gamma \in (0, 1/2)$, there is $C > 0$ such that for any $\alpha \in [-d^{-1-2\gamma}, d^{-1-2\gamma}]$, we have

$$\mathbb{E}[\exp(\alpha X_y \mathbf{1}_{X_y \le d^{1+2\gamma}}) | B_k(x)] \le \exp(C\alpha^2 d^2) + C|\alpha| d^2 \mathrm{e}^{-cd^{2\gamma}}$$

if $|B_k(x)| \le \sqrt{N}$.

(ii) For any k, the random variables $(X_y)_{y \in S_k(x)}$ are i.i.d. conditioned on $B_k(x)$.

(iii) For any $\gamma \in (0, 1/2)$, there are $c > 0$ and $C > 0$ such that $\mathbb{P}(X_y \ge d^{1+2\gamma} | B_k(x)) \le C\mathrm{e}^{-cd^{2\gamma}}$.

Proposition 5.21 (Refined of analysis of $\mathbb{G}|_{B_r(x)}$, $x \in \mathcal{W}^+$)**.** *For $0 < \alpha < 1/12$ and $r = O(1)$, there exits $K, c > 0$ depending on α such that if*

$$K \le d \le (\log N)^2$$

then following holds with probability $1 - O\big(\mathrm{e}^{-c\omega(\alpha)}\big)$.

(i) For each $x \in \mathcal{W}^+(\kappa)$

$$(\beta_x - 1) = O\Big(\frac{\kappa \log\log N}{\sqrt{\log N}}\Big).$$

(ii) For each $x \in \mathcal{W}^+(\kappa)$

$$\max_{y \ne x} |v_y| \le C(\log\log N)^2, \quad \big|\big\{y \in B_r(x) \colon |v_y| \ge d^{\alpha}\big\}\big| \le O\Big(\frac{\kappa \log\log N}{d^{2\alpha}}\Big).$$

(iii) For each $x \in \mathcal{W}^+(\kappa)$

$$\sum_{y \in S_i(x)} v_y^2 = D_x d^{i-1}\Big(1 + O\Big(\frac{\omega(\alpha)^2}{\sqrt{D_x}}\Big)\Big) + (\log\log N)^6 \mathbf{1}_{d^{2\alpha} \le \log\log N}, \quad 1 \le i \le r.$$

Proof. Let $\alpha \in (0, \frac{1}{2})$ and $r = O(1)$. We work on the event defined by Propositions 5.19. The proof of (i) is similar to Lemma 5.18. We use (5.3), the fact that $\mathfrak{u}_+ \ge C\frac{\log N}{d \log\log N}$ and Lemma 5.7 with $\varepsilon^2 = \frac{\kappa \log(\mathfrak{u}_+)^2}{\log N}$ and a union bound.
Let us write

$$\Omega_1(x, t, k) := \Big\{\exists y_1, \dots y_k \in B_r(x) \colon |v_{y_i}| \ge t, y_i \ne x, \, i \in [k]\Big\}.$$

Using (5.14) and (B.2), we find

$$\begin{aligned}
\mathbb{P}\Big(\bigcup_{x \in \mathcal{W}^+} \Omega_1\big(x, C(\log\log N)^2, 1\big)\Big) &\le C D_x d^r N \mathbb{P}(x \in \mathcal{W}^+) \mathbb{P}[v_y \ge C(\log\log N)^2] \\
&\le C \exp\Big(\kappa \log(\mathfrak{u}_+) + r \log d + \log(D_x)\Big)\Big[N^{-c^2/2} \vee \exp\Big(-dh\Big(\frac{C(\log\log N)^2}{\sqrt{d}}\Big)\Big)\Big] \\
&\le C \exp\left(2\kappa \log(\mathfrak{u}_+) + (r+1)\log d - \frac{2C^2(\log\log N)^4}{3\left(1 + c\left(\frac{C(\log\log N)^2}{9d}\right)^{1/2}\right)}\right) \\
&\le C \exp\Big(-\frac{C(\log\log N)^3}{2}\Big) = O\big(-\mathrm{e}^{c\omega(\alpha)}\big),
\end{aligned}$$

for c small enough and C large enough so that $\kappa \log\log N \le \frac{C}{2}(\log\log N)^2$. Here we used $\log(D_x) \le \log(\mathfrak{u}_+) + \log d$ in the third inequality and $\kappa = O(\log\log N)$ in the fourth inequality.
We also have

$$\mathbb{P}\bigg(\bigcup_{x\in\mathcal{W}^+}\Omega_1(x, d^\alpha, k)\bigg) \le C\exp\bigg(\kappa\log(\mathfrak{u}_+) + k(r+1)\log d - k\frac{3d^{2\alpha}}{2(3+d^{\alpha-\frac{1}{2}})}\bigg).$$

For $d^{2\alpha} \ge \log\log N$, we can set $k=1$ and the term inside the exponential is dominated by $-c(\log\log N \vee d^{2\alpha})$, for $c>0$ small enough. On the other hand, for $d^{2\alpha} \le \log\log N$ in order to get a dominant negative factor in the exponential, we must impose

$$(r+1)\log d - d^{2\alpha}\frac{3}{8} < 0,$$

which is satisfied for $d \ge K_1(r,\alpha)$. Setting $k \ge C\kappa\frac{\log\log N}{d^\alpha}$, for $C \ge 0$ large enough, the above probability is bounded by $\exp\big(-\frac{C}{2}\kappa\log\log N\big)$. This finishes the proof of (ii).

We now turn to (iii) and write

$$\sum_{y\in S_i(x)} v_y^2 = \sum_{y\in S_i(x)} v_y^2 \mathbf{1}_{|v_y|>d^\alpha} + \sum_{y\in S_i(x)} v_y^2\mathbf{1}_{|v_y|\le d^\alpha}. \tag{5.24}$$

We begin with the first sum. Observe that by (ii), it is zero with high probability when $d^{2\alpha} \ge \log\log N$. To study smaller values of d, let us introduce

$$\mathcal{N}_k(x,i) := \big|\,\big\{y \in S_1(x) : d^2e^k \le (D_y - d)^2 < d^2\mathrm{e}^{k+1}\big\}\,\big|, \quad k\ge 0,\, i\in[r].$$

We find, using Bennett's inequality and (5.2), for any $\nu>0$ there are $C_\nu, c_\nu > 0$ such that, for $1\le i\le r$,

$$\mathbb{P}\Big[\mathcal{N}_k(x,i) \ge n\big||S_i(x)|\Big] \le \binom{|S_i(x)|}{n}\mathrm{e}^{-cdn(\mathrm{e}^k\wedge\mathrm{e}^{k/2})} \le \exp\Big(n\Big[\log\big(|S_i(x)|\big) - cd\big(\mathrm{e}^k\wedge\mathrm{e}^{k/2}\big)\Big]\Big). \tag{5.25}$$

By Lemma 5.2 and (5.23), $r = O(1)$ and $d \le \log\log N$ we have $\log(|S_i(x)|) \le 2\log\log N$. Therefore

$$\mathbb{P}[\mathcal{N}_k(x,i) \ge n||S_i(x)|] \le (\log N)^{-2n} + O\big(N^{-\eta}\big), \quad k \ge k_* := \log\Big(\frac{4\log\log N}{cd}\Big),\, i\in[r],$$

where the term $O\big(N^{-\eta}\big)$ comes from Proposition 5.19.
Writing $\mathcal{N}_k(x) := \sum_{i=1}^r \mathcal{N}_k(x,i)$ and using $r=O(1)$, we find that

$$\begin{aligned}
&\mathbb{P}\Big[\exists x\in\mathcal{W}^+(\kappa),\, i\in[r],\, k_*\le k\le\log\log N:\ \mathcal{N}_k(x) \ge C\log\log N\Big]\\
&\le N\mathbb{P}\big(x\in\mathcal{W}^+(\kappa)\big)\mathbb{P}\Big[\mathcal{N}_k(x,i) \ge C\log\log N/r\Big||S_i(x)|\Big]\\
&\le \mathrm{e}^{\kappa\log(\mathfrak{u}_+)}(\log N)^{-C\log\log N/r} + O\big(N^{-\eta}\big) = O\big(\mathrm{e}^{-c\omega(\alpha)}\big),
\end{aligned}$$

for $C\ge 0$ large enough. Here we used (5.3) to bound $\mathbb{P}(x\in\mathcal{W}^+(\kappa))$. Now since by (ii), $|v_y| = O\big((\log\log N)^2\big)$ for $y\ne x$, we can bound $k \le k_{\max} := \log\log\log N$ and find that with probability $1 - O\big(\mathrm{e}^{-c\omega(\alpha)}\big)$, for each $x\in\mathcal{W}^+(\kappa)$,

$$\begin{aligned}
\sum_{y\in S_i(x)} v_y^2\mathbf{1}_{|v_y|>d^\alpha} &\le \sum_{k=(1-\alpha)\log d}^{k_*} d\mathrm{e}^{k+1}|\mathcal{N}_k(x,i)| + C\sum_{k=k_*}^{k_{\max}}(\log\log N)^5\\
&\le C\frac{\kappa\log\log N}{d^{1+\alpha}} + (\log\log N)^5 k_{\max} \le (\log\log N)^6,
\end{aligned}$$

holds for each $x\in\mathcal{W}^+(\kappa)$ with probability $1 - O\big(\mathrm{e}^{-c\omega(\alpha)}\big)$. Here we used (ii) in the second inequality and $d\mathrm{e}^{k_*} \le Cd(\log\log N)^D$ as well as $\kappa = O((\log\log N)^2)$ and $d \le \log\log N$ in the last inequality.

We now turn to the second sum in (5.24) which is amenable to Lemma 5.20. For $x \in \mathcal{W}^+(\kappa)$, and some universal $C > 0$ and $c > 0$, we get

$$\begin{aligned}
&\mathbb{P}\Big(\Big|\frac{1}{|S_i(x)|}\sum_{y\in S_i(x)}(v_y^2-1)\mathbf{1}_{|v_y|\le d^\alpha}\Big|\ge t\Big||S_1(x)|\Big)+\mathbf{1}_{x\in\mathcal{W}(\kappa)}O(\mathrm{e}^{-cd^{2\alpha}})\\
&=\mathbb{P}\Big(\Big|\sum_{y\in S_i(x)}((D_y-d)^2-d)\mathbf{1}_{(D_y-d)^2\le d^{1+2\alpha}}\Big|\ge t|S_i(x)|d\Big||S_1(x)|\Big)+\mathbf{1}_{x\in\mathcal{W}(\kappa)}O(\mathrm{e}^{-cd^{2\alpha}})\\
&\le 2\inf_{z\in[0,d^{-1-2\gamma}]}\mathrm{e}^{-ztd|S_i(x)|}\big(\mathrm{e}^{Cz^2d^2}+Cd^2z\mathrm{e}^{-cd^{2\alpha}}\big)^{|S_i(x)|}+\mathbf{1}_{x\in\mathcal{W}(\kappa)}O(\mathrm{e}^{-cd^{2\alpha}}).
\end{aligned}$$

Suppose that

$$Cd^{1-2\alpha}\mathrm{e}^{-cd^{2\alpha}}\le 1,$$

which is satisfies as soon as $d \ge K_2(\alpha, c)$. Using the inequality $1 + x \le \mathrm{e}^x$, we find

$$\inf_{z\in[0,d^{-1-2\gamma}]}\mathrm{e}^{-ztd|S_i(x)|}\big(\mathrm{e}^{Cz^2d^2}+Cd^2z\mathrm{e}^{-cd^{2\alpha}}\big)^{|S_i(x)|}\le\inf_{z\in[0,d^{-1-2\gamma}]}\mathrm{e}^{-ztd|S_i(x)|+Cz^2d^2|S_i(x)|}\quad\le\mathrm{e}^{-\frac{t^2|S_i(x)|}{4C}}$$

where we optimized in z at $z_* = \frac{t}{2Cd}$ and set $t = \frac{C\omega(\alpha)^2}{\sqrt{|S_i(x)|}}$ for $C > 0$ large enough. If $\alpha < 1/12$ and $d \le (\log N)^2$ then z^*, which is the minimizer of the expression in the infimum, satisfies $z^* \in [0, d^{-1-2\alpha}]$. Moreover since $|S_i(x)| \ge D_x \ge c\frac{\log N}{\log(\log N/d)}$ for $d \ge K_3 = K(r, \alpha)$, with K defined in Proposition 5.19, we get the correct value for the error term.
Taking a union bound over $x \in [N]$, we conclude using (5.3) and

$$\mathrm{e}^{\kappa\log(\mathfrak{u}_+)}\mathrm{e}^{-c\omega(\alpha)^2}\le\mathrm{e}^{-c\omega(\alpha)^2/2}=O\Big(\mathrm{e}^{-c\omega(\alpha)}\Big)\tag{5.26}$$

since $\kappa \log(\mathfrak{u}_+) \le (\log \log N)^3$.
Taking $d \ge \max(K_1, K_2, K_3)$ concludes the proof. □

5.5 Neighborhood of small-degree vertices

In this section we study the neighborhood of small-degree vertices. At the begin of Section 5.5, we explained how the presence of large-degree vertex can be used to control the neighborhood of that vertex, in particular excluding the occurrence of extreme events (with high probability). While this method works just as well for small-degree vertices when $d \gg \log N$, it fails when $d \le \log N$. The reason for this is easy to understand: if you consider the distribution of all degrees as an histogram on the real axis (see Figure 5.5), then lowering the value of d will push the histogram to the left. However, degrees can only take non-negative values and it is thus only possible to push the histogram so far left before hitting the vertical line $y = 0$. This has the annoying consequence that the number of small-degree vertices becomes polynomially large.

In particular, seeing small-degree becomes less unlikely and thus seeing extreme events in the neighborhood of small-degree vertices becomes less unlikely.

$$\mathbb{P}\big[\mathcal{P}_d\ge d\mathfrak{u}_+-C\big]=\frac{\mathfrak{u}_+^C}{N}(1+o(1))\ll\mathrm{e}^{-d}\frac{d^C}{C!}=\mathbb{P}\big[\mathcal{P}_d\ge d\mathfrak{u}_-+C\big],$$

for $d \le \log N$, since then $d\mathfrak{u}_- = 1$. This is an adversarial effect against which we must fight and it prevents us from reaching the regime $d \ge \frac{1}{2}\log N + C \log \log N$, for $C = O(1)$.

The following lemma is an *a priori* estimate on $\beta_x - 1$ which will be refined later.

Lemma 5.22 (Behavior of β_x, $x \in \mathcal{U}^-$)**.** *Let $\varepsilon > 0$ and $(\frac{1}{2}+\varepsilon)\log N \le d \le N^{\frac{1}{2}-\varepsilon}$. Then, with very high probability for all $x \in [N]$*

$$(\beta_x-1)=\mathcal{C}\frac{\sqrt{\log N}}{d\sqrt{\mathfrak{u}_-}}\tag{5.27}$$

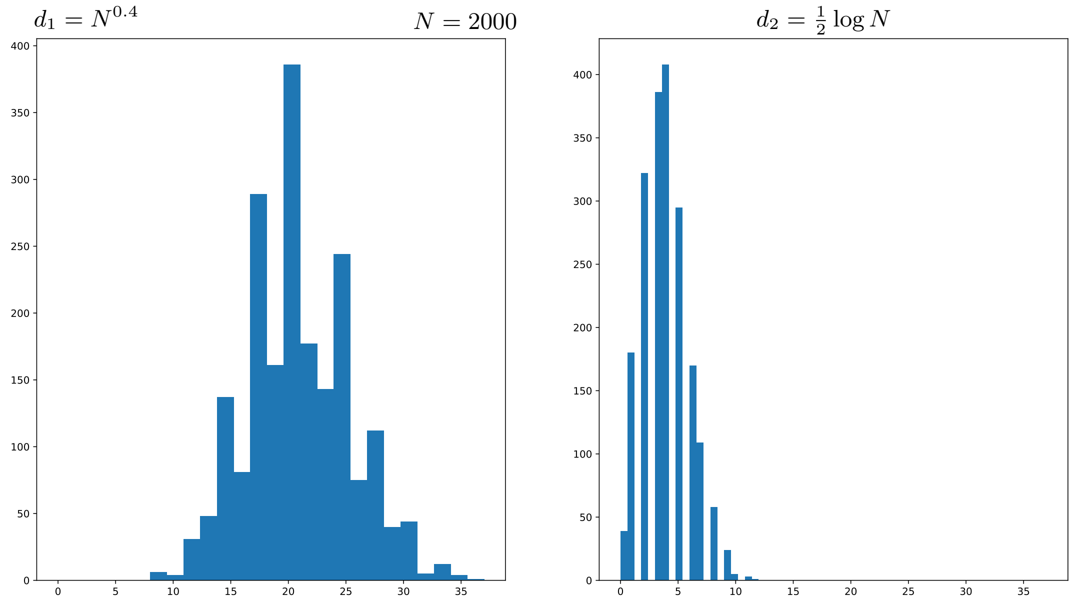


Figure 5.1: We sample $N = 2000$ independent $\mathcal{B}_{N,d}$ random variables. The global statistics of this sampling resembles the global statistics of $(D_x : x[N])$. On the left-hand side, we set $d = N^{0.4} \gg \log N$ and $d = \frac{1}{2}\log N < \log N$ on the right-hand side.

Proof. Same proof as Lemma 5.18. □

Proposition 5.23 (Neighborhood of $x \in \mathcal{U}^-$)**.** *Let $K, \varepsilon, \alpha > 0$, d as in (3.29) and $r \geq 0$. Then if $r = O(1)$, there exists $c_* > 0$ such that the following holds with probability $1 - O\big(\mathrm{e}^{-c_*(\log N)^{c_*}}\big)$.*

(i) The balls $\big(B_r(x) \colon x \in \mathcal{U}^-(\varepsilon)\big)$ are disjoint trees.

(ii) For each $x \in \mathcal{U}^-(\varepsilon)$

$$\left|\frac{|S_{i+1}(x)|}{d|S_i(x)|} - 1\right| = O\bigg(\frac{\log N}{d}|S_i|^{-1/2}\bigg), \qquad \left|\frac{|S_i(x)|}{D_x d^{i-1}} - 1\right| = O\bigg(\frac{\log N}{d} D_x^{-1/2}\bigg), \qquad i \in [r].$$

(iii) For each $x \in \mathcal{U}^-(\varepsilon)$ there exists $y_* \in B_r(x)$, such that

$$\min_{y \in B_r(x)\, y \neq x} |v_x - v_y| \geq \mathbf{1}_{d \leq (1-\varepsilon/2)\log N} c_* \sqrt{\log N} + c_*(\log N)^{c_*}, \qquad \max_{y \in B_r(x)} |v_y| = O\big(\sqrt{\log N}\big),$$

and

$$\Big|\{y \in B_r(x) \colon |v_x - v_y| \leq c_* \sqrt{d}\}\Big| \leq 1. \tag{5.28}$$

(iv) For each $x \in \mathcal{U}^-(\varepsilon)$

$$\Big|\{y \in B_r(x) \colon |v_y| \geq d^{\alpha}\}\Big| = O\big(d^{1-2\alpha}\big).$$

Proof. The fact that the balls $B_r(x)$, $x \in \mathcal{U}^-(\varepsilon)$ are disjoint follows from the stronger Lemma 5.13. The fact that $B_r(x)$ contains no cycles follows from Lemma 5.6 as in the proof of Proposition 5.19. This proves ((i)). ((ii)) is proved exactly as in Proposition 5.19.

By Lemma 5.6, there exists $C > 0$ such that $\max_{x\in\mathcal{U}^-(\varepsilon)}|B_r(x)| \le Cd^r$ with probability $1 - O(N^{-2})$. If $d \ge (1 - \varepsilon/4)\log N$, using Lemma B.3 we find

$$\mathbb{P}[\exists x \in \mathcal{U}^-(\varepsilon),\, \exists y \in B_r(x)\, |v_y| \ge t] \le CN^{\frac{1-\varepsilon}{2}} d^r \mathrm{e}^{-t^2/2} + O(N^{-2}).$$

If $d \ge (1 + \varepsilon/4)$ we can choose $t = \sqrt{(1 - \varepsilon/4)s}$ and if $d \le (1 + \varepsilon/4)$, $t = \sqrt{(1 - \varepsilon/4)\log N}$, to bound the left-hand side is bounded by $O(N^{-\eta})$ for $0 < \eta < \varepsilon/5$. In both cases we can conclude that

$$\min_{y\in B_r(x),\, y\neq x} |v_x - v_y| \ge \sqrt{\varepsilon \log N/4}, \quad x \in \mathcal{U}^-(\varepsilon),$$

with probability $1 - O(N^{-\eta})$. In particular (5.28) follows trivially in that regime.
If $d \le (1-\varepsilon/4)\log N$, then $x \in \mathcal{U}^-(\varepsilon)$ is by definition the set of vertices such that $D_x \le 1/\varepsilon = O(1)$. Denoting $n := \lceil 1/\varepsilon\rceil$ we see that the condition $|v_y-v_x| \le c(\log N)^c$ implies $D_y \le (1-\eta)d$ for any $0 < \eta < 1-\frac{c(\log N)^c+n}{\sqrt{d}}$. Since $c < 1/2$, we can find a constant $\eta > 0$ satisfying this condition and can thus apply Lemma B.5. We find

$$\begin{aligned}\mathbb{P}[\exists x \in \mathcal{U}^-(\varepsilon),\, \exists y \in B_r(x),\, v_y \le c(\log N)^c] &\le \sum_{x\in\mathcal{U}^-(\varepsilon)} \sum_{y\in B_r(x),\, y\neq x} \mathbb{P}\big[v_y \le c(\log N)^c\big]\\ &\le CN\mathrm{e}^{-2d} d^{r+n} \exp\Big(\frac{c}{2}\log(d)(\log N)^c\Big) + O(N^{-2}) = O\big(\mathrm{e}^{-2(\log N)^{1/K} + \frac{c}{2}(\log N)^{2c}}\big) = O\Big(\mathrm{e}^{-c_*(\log N)^{c_*}}\Big),\end{aligned}$$

where in the last inequality we chose $c_* \le \frac{1}{2K} \wedge \frac{1}{2}$. The assertion about $\max_y|v_y|$ is proved analogously.
For (5.28), we find, again using Lemma B.5,

$$\begin{aligned}&\mathbb{P}\Big[\exists x \in \mathcal{U}^-(\varepsilon),\, \exists y_1,\dots,y_k \in B_r(x),\, v_{y_i} \le c\sqrt{d},\, i \in [k]\Big]\\ &\le CN\mathrm{e}^{-3d} d^{r+nk} \mathrm{e}^{2cd\log(1/c)}\Big(1 + O\Big(\frac{d^2}{N}\Big)\Big) + O(N^{-2}) \le C\mathrm{e}^{-3(\log N)^{1/K}} d^{r+nk} N^{-\frac{k-1}{2}+k(c\log(1/c))},\end{aligned}$$

and the right-hand side is bounded by $O\Big(\mathrm{e}^{-c_*(\log N)^{c_*}}\Big)$ as soon as $k \ge 2$ for $0 < c < c_*$ small enough.
The proof of (iv) is analogue Proposition 5.19 (iii). ☐

Lemma 5.24 (Behavior of β_{z_x}, $x \in \mathcal{W}_e(\kappa)$)**.** *Let $\varepsilon > 0$ and $(\frac{1}{2} + \varepsilon)\log N \le d \le N^{\frac{1}{2}-\varepsilon}$. We have with high probability, for every $x \in \mathcal{W}$,*

$$(\beta_{z_x} - 1) = C\frac{\kappa \log d}{(\log N)^{3/4}}$$

Proof. Similar proof as for Lemma 5.18 with the observation that by Lemma 5.13, $\min_x|S_1(z_x)| \ge c\frac{d}{\log d}$, for some $c > 0$ small enough. We have by Lemma 5.7

$$\begin{aligned}\mathbb{P}[\exists x \in \mathcal{W}_e(\kappa),\, (\beta_x - 1) \ge \varepsilon] &\le \mathbb{P}[\exists x \in \mathcal{W},\, (\beta_x - 1) \ge \varepsilon, \Xi_1(r, \varepsilon)] + \mathbb{P}[\Xi_1(r, \varepsilon)^c]\\ &\le \mathrm{e}^{\kappa \log d} \exp\big(-cd^{3/2}(v_1 - v_{z_x})\varepsilon^2\big) + \mathbb{P}[\Xi_1(r, \varepsilon)^c],\end{aligned}$$

where we used $|S_1(z)| = \sqrt{d}(v_1 - v_{z_x}) + 1$.
Since $(v_1 - v_{z_x}) \ge cd^c$ for some $c > 0$ by (ii) of Proposition 5.23, we can make this probability smaller than $C\mathrm{e}^{-d^c}$ by choosing Setting $\varepsilon = C\kappa \log d/d^{3/4}$ for $C \ge 0$ large enough. ☐

Proposition 5.25 (Neighborhood of $x \in \mathcal{W}^-(\kappa)$)**.** *Let $\alpha, \varepsilon > 0$ and*

$$\log N - (\log\log N)^2 \le d \le (\log N)^{3/2}.$$

Then for $c_* > 0$, the following holds with probability $1 - O\Big(\mathrm{e}^{-c_* d^{2\alpha}}\Big)$

(i) For each $x \in \mathcal{W}^-(\kappa)$,

$$\max_{y \in B_r(x),\, y \neq x} |v_y| = O\big(d^\alpha\big)$$

(ii) For each $x \in \mathcal{W}^-(\kappa)$,

$$\sum_{y \in S_i(x)} v_y^2 = D_x d^{i-1}\bigg(1 + O\bigg(\frac{d^\alpha}{(\log N)^{1/2}}\bigg)\bigg), \quad 1 \leq i \leq r.$$

Note that we could improve ((i)) by showing that $\max_{y \in B_r(x),\, y \neq x}|v_y| = O\big((\log\log N)^2\big)$ with probability $1 - O\Big(\mathrm{e}^{(\log\log N)^2}\Big)$. However, since we want to cast all results of Proposition 3.21 under one common bound and since d^α is enough for our purpose, we do not pursue this improvement.

Proof. Using Lemma B.5, we find that

$$\mathbb{P}\big[x \in \mathcal{W}^-(\kappa)\big] \leq \mathrm{e}^{\kappa \log d} \vee N\mathrm{e}^{-d} d^{\tilde{\kappa}}.$$

We can then proceed as in the proof of (ii) Proposition 5.21 to prove that

$$\mathbb{P}\big[\exists x \in \mathcal{W}^-(\kappa),\, y \in B_r(x),\, y \neq x\, |v_y| \geq C d^\alpha\big]$$
$$\Big[\mathrm{e}^{\kappa \log d} \vee N\mathrm{e}^{-d} d^{\tilde{\kappa}}\Big]\mathrm{e}^{-C d^{2\alpha}} = O\Big(\mathrm{e}^{-c_* d^{2\alpha}}\Big),$$

for $c_* > 0$ small enough. We deduce that with probability $1 - O\Big(\mathrm{e}^{-c_* \log\log N}\Big)$, for any $\alpha > 0$,

$$\sum_{y \in S_i(x)} |v_y|^2 = \sum_{y \in S_i(x)} |v_y|^2 \mathbf{1}_{|v_x| \leq d^\alpha}.$$

Therefore we can prove (ii) as we proved (iii) in Proposition 5.21. We skip the details. □

Proposition 5.26 (Neighborhood of $x \in \mathcal{W}^\gamma(\kappa)$)**.** *Let* $K, \varepsilon > 0$*,* d *as in (3.28) and* $r \geq 0$ *and* κ*,* $\tilde{\kappa}$ *as in (3.29). Then there exists* $c_* > 0$ *such that the following holds with probability* $1 - O\big(\mathrm{e}^{-c_*(\log N)^{c_*}}\big)$.

(i) For each $x \in \mathcal{W}^\gamma(\kappa)$,

$$\max_{y \in B_r(x) \setminus B_1(x)} |v_y| = O\big((\log\log N)^2\big)$$

(ii) For each $x \in \mathcal{W}^\gamma(\kappa)$*, if* $y_* := \mathrm{argmin}_{z \in B_r(x)} v_z$*, then*

$$v_{y_*} \geq v_\gamma + O\Big(\frac{d^{c_*}}{\sqrt{d}\log(d - \frac{1}{2}\log N)}\Big), \qquad \min_{z \in B_r(x),\, z \neq x, y_*} |v_x - v_z| \geq c_* \sqrt{\log N}.$$

(iii) For each $x \in \mathcal{W}^\gamma(\kappa)$,

$$\sum_{y \in S_i(x)} v_y^2 = |S_i(x)|\bigg(1 + O\bigg(\frac{d^\alpha}{(\log N)^{3/2}}\bigg)\bigg), \quad 2 \leq i \leq r.$$

Proof. (i) is proved as (ii) of Proposition 5.21, using the fact that we defined $\mathcal{W}^\gamma(\kappa)$ such that

$$\mathbb{P}\big[x \in \mathcal{W}^\gamma(\kappa)\big] = O\Big(\mathrm{e}^{\kappa \log d}\Big).$$

We skip the details.
Turning to (ii) we use Lemma B.5

$$\mathbb{P}\big[\exists x \in \mathcal{U}^-(\varepsilon),\, \exists y \in B_r(x),\, v_y \geq v_e(\gamma) - \frac{t}{\sqrt{d}}\big] = Cd^r N d\mathrm{e}^{-d}\mathbb{P}[\mathcal{P}_d \leq d\mathfrak{u}_\gamma - t] + O\big(N^{-2}\big)$$
$$\leq Cd^{r+1}\mathfrak{u}_\gamma^t \leq C\mathrm{e}^{r+1\log d - t\log(\mathfrak{u}_\gamma)}.$$

Setting $t = \frac{d^c}{\log(\mathfrak{u}_\gamma)}$ and using (B.8) we prove (ii)

The second point is proved by first restricting ourselves to Ω^- the event $\Xi = \big\{\max_{y\in B_r(x),\, y\neq x}|v_y| \leq C(\log\log N)^2\big\}$ which by (i). Then for any $\alpha > 0$, we have

$$\mathbf{1}_\Xi \sum_{y\in S_i(x)} v_y^2 = \sum_{y\in S_i(x)} v_y^2 \mathbf{1}_{v_y^2 \leq d^\alpha}.$$

The sum on the right-hand side can be analyzed using Lemma 5.20 as was done in Proposition 5.19. We skip the computations. □

5.6 Decorrelation results

Let us introduce the $d-$dependent function

$$Q(u,v) \geq \mathbb{P}(\mathcal{P}_{dv} - dv \geq w\sqrt{dv}). \tag{5.29}$$

We recall [7, Proposition 7.1]

Proposition 5.27 (Decorrelation)**.** *Suppose that* $1 \leq d \leq N^{1/12}$ *and* $k \leq N^{1/12}$*. Let* $v_1,\dots,v_k \in \mathbb{N}$ *satisfy* $2 \leq v_1,\dots,v_k \leq N^{1/4}$ *and* $w_1,\dots,w_k \in \mathbb{R}$*. Then*

$$\mathbb{P}\left(\bigcap_{i\in[k]} \{|S_1(i)|\} = v_i,\, |S_2(i)| \geq u_i\right) = \prod_{i\in[k]} \mathbb{P}(\mathcal{P}_d = v_i)\mathbb{P}(\mathcal{P}_{dv_i} \geq u_i) + O\left(N^{-1/3}\prod_{i\in[k]}\mathbb{P}(\mathcal{P}_d = v_i) + N^{-k-1}\right)$$

for any $u_1,\dots,u_k \in \mathbb{N}$.

We need the following variations of [7, Proposition 7.1]. Let us introduce the scaling parameters

$$\mathbf{a} = \sqrt{2\log N} + \frac{1}{2}\log(4\pi\log N), \qquad \mathbf{b} = \sqrt{2\log N}.$$

Proposition 5.28 (Decorrelation in dense regime)**.** *Let* $\varepsilon > 0$ *and* $(\log N)^{1+\varepsilon} \leq d \leq N^{1/3}$ *and* $k \leq (\log N)^{\varepsilon/3}$ *Let* $w_1,\dots,w_k \in \mathbb{R}$*. Then,*

$$\mathbb{P}\Big(\bigcap_{i\in[k]}\{v_i \geq \mathbf{a} + \frac{1}{\mathbf{b}}w_i\}\Big) = \frac{1}{N^k}\prod_{i\in[k]}\mathrm{e}^{-w_i}O\Big(1 + (\log N)^{-\varepsilon/6}\Big)$$
$$+O\Big(N^{-1/4}\prod_{i\in[k]}\mathbb{P}\big(\mathcal{P}_d - d \geq \sqrt{d}\mathbf{b}\big)\Big)$$

Proof. We first recall that the degrees, which are distributed as Binomial $(N, d/N)$ variables, can be very well approximated by Poisson law of parameter d. By Lemma B.5 we have, for $t \geq 0$

$$20\mathbb{P}\left(\mathcal{B}_{N,d/N} \geq t\right) = \mathbb{P}(\mathcal{P}_d \geq t) + \sum_{n\geq t}\mathrm{e}^{-d}\frac{d^n}{n!}\left(1 + O\left(\frac{n^2}{N}\right)\right) = \mathbb{P}(\mathcal{P}_d \geq t) + O\left(\frac{t^3}{N}\right).$$

We define Ξ the event that no vertex in $[k]$ are neighbors and decompose $\Xi^c = \bigsqcup_{\mathbb{U}}\{\mathbb{G}|_{[k]} = \mathbb{U}\}$, where the union ranges over the set of nonempty graphs $\mathbb{U}$ on $[k]$. Thus we estimate

$$\mathbb{P}\left(\bigcap_{i\in[k]}\{v_i \geq \mathbf{a} + \frac{1}{\mathbf{b}}w_i\}\right) = \mathbb{E}\left[\prod_{i\in[k]} \mathbf{1}_{v_i\geq \mathbf{a}+\frac{1}{\mathbf{b}}w_i}\mathbf{1}_{\Xi}\right] + \mathbb{E}\left[\prod_{i\in[k]} \mathbf{1}_{v_i\geq \mathbf{a}+\frac{1}{\mathbf{b}}w_i}\mathbf{1}_{\Xi^c}\right]$$

We write l_i the degree of the vertex i in $\mathbb{U}$ and $u_i = d + \sqrt{d}(\mathbf{a} + \frac{1}{\mathbf{b}}w_i)$. By (B.5) we find

$$\begin{aligned}
\mathbb{E}\left[\prod_{i\in[k]} \mathbf{1}_{v_i\geq \mathbf{a}+\frac{1}{\mathbf{b}}w_i}\mathbf{1}_{\Xi^c}\right] &= \sum_{\mathbb{U}} \mathbb{E}\left[\left[\prod_{i\in[k]} \mathbf{1}_{v_i\geq \mathbf{a}+\frac{1}{\mathbf{b}}w_i}\middle| A_{[k]}\right]\mathbf{1}_{\mathbb{G}|_{[k]}=\mathbb{U}}\right] \\
&= \sum_{\mathbb{U}} \mathbb{E}\left[\prod_{i\in[k]} \mathbb{P}\left(\mathcal{P}_{(N-k+1)d/N} \geq u_i - l_i\right)\left(1 + O\left(\frac{v_i^2+d^2+t^3}{N}\right)|A_{[k]}\right)\right] \\
&= \sum_{\mathbb{U}} \mathbb{E}\left[\prod_{i\in[k]} \mathbb{P}(\mathcal{P}_{(N-k+1)d/N} \geq u_i)\left(\frac{v_i}{d}\right)^{l_i}\left(1 + O\left(\frac{v_i^2+d^2+t^3}{N}\right)|A_{[k]}\right)\right] \\
&= \sum_{l=1}^{\frac{k(k-1)}{2}} \binom{\frac{k(k-1)}{2}}{l}\left(\frac{d}{N}\right)^l \prod_{i\in[k]} \mathbb{P}\left(\mathcal{P}_d \geq u_i\right)\left(1 + O\left(\frac{kd+v_i^2+d^2+t^3}{N}\right)\right) \\
&= O\left(\left(1+\frac{k}{N^{1/3}}\right)\left(1+\frac{k^2d}{N}+\frac{d^kk^2}{N^k}\right)\right)\prod_{i\in[k]} \mathbb{P}\left(\mathcal{P}_d \geq u_i\right)
\end{aligned}$$

where we split the last sum between $l \leq k$ and $k \geq l$ in the last equality.
On the event Ξ, the variables v_i are independent conditioned on $A_{[k]}$ and distributed as $\mathcal{B}_{N-k,d/N}$. But if $X \sim \mathcal{B}_{N-k,d/N}$, we have

$$\mathbb{P}\left[X \geq d + \sqrt{d}\left(\mathbf{a} + \frac{w_i}{\mathbf{b}}\right)\right] = \mathrm{e}^{-d}\sum_{k=d+\sqrt{d}\left(\mathbf{a}+\frac{w_i}{\mathbf{b}}\right)}^{d+C_\nu\sqrt{d}\omega} \frac{d^k}{k!}O\left(\mathrm{e}^{k^2/N}\right) + O(N^{-\nu}),$$

for $C_\nu > 0$ that depend on $\nu > 0$. Choosing $\nu = 10$ and abbreviating $\omega = \frac{1}{\sqrt{d}}(\mathbf{a} + \frac{w_i}{\mathbf{b}})$, we get

$$\begin{aligned}
\sum_{k=d+\sqrt{d}\left(\mathbf{a}+\frac{w_i}{\mathbf{b}}\right)}^{d+C_\nu\sqrt{d}\omega} \frac{d^k}{k!} &= \mathrm{e}^{-d}\sum_{k=0}^{(C-1)\sqrt{d}\omega} \frac{\mathrm{e}^{d(1+\omega)}\mathrm{e}^k}{\sqrt{2\pi d(1+\omega+k/d)}}\exp\left(-d\left(1+\omega+\frac{k}{d}\right)\log\left(1+\omega+\frac{k}{d}\right)\right) \\
&= \mathrm{e}^{-d}\sum_{k=0}^{(C-1)\sqrt{d}\omega}(1+\omega)^{-d(1+\omega)-k}\frac{\mathrm{e}^{d(1+\omega)}\mathrm{e}^k}{\sqrt{2\pi d(1+\omega)}}\left(1+O\left(d^{-1}\right)\right) \\
&\quad \exp\left(-d\left(1+\omega+\frac{k}{d}\right)\left(\frac{k}{1+\omega} - \frac{k^2}{d(1+\omega)^2} + O\left(\frac{1}{\sqrt{d}}\right)\right)\right) \\
&= \frac{(1+\omega)^{-d(1+\omega)}}{\sqrt{2\pi d(1+\omega)}}\sum_{k=0}^{(C-1)\sqrt{d}}(1+\omega)^{-k}\left(1+O\left(d^{-1/2}\right)\right) \\
&= \frac{(1+\omega)^{-d(1+\omega)}}{\sqrt{2\pi d(1+\omega)}}\sum_{k\geq 0}(1+\omega)^{-k}\left(1+O\left(d^{-1/2}\right)\right).
\end{aligned}$$

In the last equality, we used the fact that $(1+\omega)^{(C-1)\sqrt{d}\omega} = O\big(N^{-C/2}\big) = O\big(d^{-1/2}\big)$ to complete the series.

Using $k^2/N = O\big(d^{-1/2}\big)$, we find

$$\begin{aligned}\mathbb{P}\left[X \geq d + \sqrt{d}\left(\mathbf{a} + \frac{w_i}{\mathbf{b}}\right)\right] &= \frac{\mathrm{e}^{d\omega}(1+\omega)^{-d(1+\omega)}}{\sqrt{2\pi d(1+\omega)}} \sum_{k\geq 0} \omega^{-k}\left(1 + O\left(d^{-1/2}\right)\right) \\ &= \exp\left(d\omega - d(1+\omega)\log(1+\omega) - \frac{1}{2}\log(4\pi \log N)\right) \\ &= \exp\left(d\omega - d(1+\omega)\left(\omega + \frac{\omega^2}{2}(1+\varepsilon_i)\right) - \frac{1}{2}\log(4\pi\log N)\right) \\ &= \frac{1}{N}\mathrm{e}^{-w_i(1+\varepsilon_i)},\end{aligned}$$

for $\varepsilon_i = O\left((\log N)^{-\varepsilon/2}\right)$. In the third equality, we used $\log(2\pi d) + \log(\omega) = \log(4\pi \log N)$. Therefore, using the fact that $d \geq \log N$, we find using $k \leq (\log N)^{\varepsilon/3}$ and $\max_i \varepsilon_i = O\big((\log N)^{\varepsilon/2}\big)$,

$$\begin{aligned}\mathbb{E}\left[\prod_{i\in[k]} \mathbf{1}_{v_i \geq \mathbf{a} + \frac{1}{\mathbf{b}}w_i}\mathbf{1}_{\Xi}\right] &= \prod_{i\in[k]}\mathbb{P}\left(\mathcal{P}_d \geq d + \sqrt{d}(\mathbf{a} + \frac{w_i}{\mathbf{b}})\right) \\ &= \frac{1}{N^k}\prod_{i\in[k]}\mathrm{e}^{-w_i}\left(1 + O\left(k\max_i \varepsilon_i\right)\right) = \frac{\prod_{i\in[k]}\mathrm{e}^{-w_i}}{N^k}\left(1 + O\left((\log N)^{-\varepsilon/6}\right)\right).\end{aligned}$$

This concludes the proof. □

Proposition 5.29 (Decorrelation for leaves)**.** *Let $\frac{1}{2}\log N \leq d \leq (\log N)^2$. Let $K \geq 0$, $k \leq (N/Kd)^{1/2}$ and $2 \leq v_1, \dots, v_k \leq Kd$ and $w_1, \dots, w_k \in R$. Then*

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{i\in[k]}\big\{D_i = 1, d\alpha_{z_i} = v_i, d\sqrt{\alpha_{z_i}}(\beta_{zi} - 1) \geq w_i\big\}\Big) &= \prod_{i\in[k]}\mathbb{P}(\mathcal{P}_d = v_i)Q(v_i, w_i) \\ &\quad + O\Big(\frac{d^3k^2}{N}(\mathrm{e}^{-d}d)^k\prod_{i\in[k]}\mathbb{P}(\mathcal{P}_d = v_i)\Big)\end{aligned}$$

Proof. Let us abbreviate by $A(v_i) := \big\{D_i = 1, d\alpha_{z_i} = v_i\big\}$ and $B(w_i) := \big\{d\sqrt{\alpha_{z_i}}(\beta_{zi} - 1) \geq w_i\big\}$.
For $l \geq 1$ we define Ξ_l as the event that there is no geodesic in $\mathbb{G}$ of length l connecting to distinct vertices of $[k]$. We use the abbreviation $\Xi_{\leq j} := \bigcap_{i\leq j}\Xi_j$. We define $\mathcal{M}_i$ for $0 \leq i \leq 5$ and $\mathcal{E}_i$ for $1 \leq i \leq 5$ to be

$$\mathcal{M}_l := \mathbb{E}\left[\prod_{i\in[k]}\mathbf{1}_{A(v_i)\cap B(w_i)}\mathbf{1}_{\Xi_l}\cdots\mathbf{1}_{\Xi_1}\right], \quad \mathcal{E}_l := \mathbb{E}\left[\prod_{i\in[k]}\mathbf{1}_{A(v_i)\cap B(w_i)}\mathbf{1}_{\Xi_l^c}\cdots\mathbf{1}_{\Xi_1}\right].$$

Now since $v_i \geq 2$, Ξ_1^c never happens it would mean that $i, j \in [k]$ form a disconnected component.

We tackle $\Xi_2^c \cup \Xi_3^c$ at the same time. On the event Ξ_1 the possible configurations of $A_{B_1([k])}$ can be decomposed in two steps. First we look for elements of $[k]$ that are at distance 2. Since $[k]$ consist only of leaves, this can be express as a sum over all possible partitions of $[k]$ with at least one non-trivial block, denoted $\mathcal{P}_{<k}$. We say $B_0([k]) \sim \Pi$ if, up to permutation, the two functions $i \mapsto z_i$ and $i \mapsto \pi_i$ are the same.

$$\begin{aligned}&\mathbb{E}\left[\prod_i \mathbf{1}_{A(i)\cap B(i)}\mathbf{1}_{\Xi_2^c\cup\Xi_3^c}\mathbf{1}_{\Xi_1}\right] \\ &= \mathbb{E}\left[\sum_{\Pi\in\mathcal{P}_{<k}}\prod_i \mathbf{1}_{D_i=1}\sum_{s\in([N]\setminus[k])^\Pi}\mathbb{E}\left[\prod_{i\in[|\Pi|]}\mathbf{1}_{v_{s_{\pi_i}} = v_i - |\pi_i| - 1}\prod_{x\in\pi_i}A_{s_{\pi_i}x}\left(\frac{d}{N}\right)^{|\pi_i|-1}\Big|B_0([k])\right]\mathbf{1}_{B_0([k])\sim\Pi}\mathbf{1}_{\Xi_3^c}\mathbf{1}_{\Xi_1}\right] \\ &= \sum_{\Pi\in\mathcal{P}_{<k}}\mathbb{E}\left[\prod_i \mathbf{1}_{D_i=1}\left(\frac{d}{N}\right)^{k-|\Pi|}\prod_{i\in[|\Pi|]}\mathbf{1}_{v_{s_{\pi_i}} = v_i - |\pi_i| - 1}\mathbf{1}_{B_0([k])\sim\Pi}\mathbf{1}_{\Xi_3^c}\mathbf{1}_{\Xi_1}\right].\end{aligned}$$

Then, given $\Pi \in \mathcal{P}_{<k}$, we look among the unique vertices $z_1, \dots, z_k$, which are now in bijection with Π their connections. This can be done by decomposing $\bigsqcup_{\mathbb{U}} \{\mathbb{G}|_{S_1([k])} = \mathbb{U}\}$ where the union ranges over the set of nonempty graphs $\mathbb{U}$ on (Π). Let us denote l_i the degree of z_i for $i \in [k]$ in $A_{(S_1(z))}$ and by $l = |\mathbb{U}|$ the number of edges of $\mathbb{U}$. We have

$$\begin{aligned}
&\mathbb{E}\left[\prod_i \mathbf{1}_{A(i)\cap B(i)} \mathbf{1}_{\Xi_2^c \cup \Xi_3^c} \mathbf{1}_{\Xi_1}\right] \\
&= \sum_{\Pi \in \mathcal{P}_{<k}} \left(\frac{d}{N}\right)^{k-|\Pi|} \mathbb{E}\left[\sum_{\mathbb{U}} \prod_i \mathbf{1}_{D_i=1} \prod_{i\in[|\Pi|]} \mathbf{1}_{v_{s_{\pi_i}}=v_i-|\pi_i|-1} \mathbf{1}_{B_0([k])\sim\Pi} \mathbf{1}_{\mathbb{G}|_{S_1([k])}=\mathbb{U}} \mathbf{1}_{\Xi_3^c} \mathbf{1}_{\Xi_1}\right] \\
&\le \sum_{\Pi \in \mathcal{P}_{<k}} \sum_{\mathbb{U}} \left(\frac{d}{N}\right)^{k-|\Pi|+l} \prod_i \mathbb{P}(\mathcal{B}_{N,d/N}=1) \prod_{i\in[|\Pi|]} \mathbb{P}(\mathcal{B}_{N-2k+1,d/N} = v_i - |s_{\pi_i}| - l_i + 1)
\end{aligned}$$

where we used the fact that conditioned on $\mathbb{U}$ on the event $\{B_0([k]) \sim \Pi\} \cap \{\mathbb{G}|_{S_1([k])} = \mathbb{U}\}$ the degrees of z_i are independent. We can now use Lemma B.5 to approximate the probabilities of the Bernoulli variables. Let n to denote the size of Π and l the number of edges of $\mathbb{U}$. We have

$$\begin{aligned}
\mathbb{E}\left[\prod_i \mathbf{1}_{A(i)\cap B(i)} \mathbf{1}_{\Xi_2^c \cup \Xi_3^c} \mathbf{1}_{\Xi_1}\right] &= (\mathrm{e}^{-d} d)^k \sum_{\Pi \in \mathcal{P}_{<k}} \sum_{\mathbb{U}} \left(\frac{d}{N}\right)^{k-|\Pi|+|\mathbb{U}|} \\
&\qquad \prod_{i\in[n]} \mathbb{P}(\mathcal{B}_{N-2k+1,d/N} = v_i - |s_{\pi_i}| - l_i + 1) \\
&= (\mathrm{e}^{-d} d)^k \sum_{\Pi \in \mathcal{P}_{<k}} \sum_{\mathbb{U}} \left(\frac{d}{N}\right)^{k-|\Pi|+|\mathbb{U}|} \\
&\qquad \prod_{i\in[n]} \mathbb{P}(\mathcal{P}_d = v_i) \left(\frac{v_i}{d}\right)^{|s_{\pi_i}|-1+l_i} \left(1 + O\left(\frac{v_i^2 + d^2 + 4k^2}{N}\right)\right) \\
&\le (\mathrm{e}^{-d} d)^k \sum_{n=1}^{k-1} \sum_{l=1}^{\frac{n(n-1)}{2}} C^{k-n+l} \binom{k}{n} n^{k-n} \binom{n(n-1)/2}{l} \left(\frac{d}{N}\right)^{k-n+l} \\
&\qquad \prod_{i\in[|\Pi|]} \mathbb{P}(\mathcal{P}_d = v_i) \left(1 + O\left(\frac{d^2+k^2+kd}{N}\right)\right)
\end{aligned}$$

where in the last step we used the assumption $v_i \le Cd$, for some $C \ge 0$.
By the binomial theorem we have

$$\begin{aligned}
\sum_{n=1}^{k-1} \sum_{l=1}^{\frac{n(n-1)}{2}} C^{k-n+l} \binom{k}{n} n^{k-n} \binom{n(n-1)/2}{l} \left(\frac{d}{N}\right)^{k-n+l} &\le \sum_{n=1}^{k-1} \binom{k}{n} n^{k-n} \left(\frac{Cd}{N}\right)^{k-n} \left(1 + \frac{Cd}{N}\right)^{n(n-1)/2} \\
&\le \frac{Cdk}{N} \sum_{n=1}^{k-1} \binom{k}{n} \left(\frac{Cdn}{N}\right)^{k-1-n} \left(1 + O\left(\frac{cdn}{N}\right)\right)^n \le \frac{Cdk}{N} \left(1 + \frac{Cdk}{N} + O\left(\frac{Cdk}{N}\right)\right)^k \\
&\le \frac{Cdk}{N} \left(1 + O\left(\frac{dnk^2}{N}\right)\right) = O\left(\frac{(\log N)^4}{\sqrt{N}}\right)
\end{aligned}$$

where in the last equality we used the fact that $\frac{dnk^2}{N} = O(1)$ and $k \le \sqrt{N}$.
The remainder of the proof is very similar to the proof of [7, Proposition 7.1]. We now turn to Ξ_4. We can

decompose $\Xi_4 = \bigcap_{1\le x<y\le k} \Xi_{4,xy}$. By a union bound, we estimate

$$\begin{aligned}\mathcal{E}_4 &\le \mathbb{E}\left[\prod_{i\in[k]} \mathbf{1}_{|S_1(z_i)|=v_i} \mathbf{1}_{\Xi_{\le 3}} \mathbf{1}_{\Xi_4}\right] \le \sum_{1\le x<y\le k} \mathbb{E}\left[\prod_{i\in[k]} \mathbf{1}_{|S_1(z_i)|=v_i} \mathbf{1}_{\Xi_{\le 3}} \mathbf{1}_{\Xi_{4,xy}}\right] \\ &\le \sum_{1\le x<y\le k} \left(\prod_{i\in[k]\setminus\{x,y\}} \mathbb{P}(|S_1(z_i)| = v_i, \Xi_{\le 3}) \mathbb{E}[\mathbf{1}_{|S_1(z_x)=v_x|} \mathbf{1}_{|S_1(z_y)=v_y|} \mathbf{1}_{\Xi_{4,xy}} \mathbf{1}_{\Xi_{\le 3}}]\right)\end{aligned}$$

where in the last step we used that the sets $(S_1(z_i))_{i\in[k]}$ are independent condition on $\Xi_{\le 3}$. We estimate

$$\mathbb{E}[\mathbf{1}_{|S_1(z_x)=v_x|} \mathbf{1}_{|S_1(z_y)=v_y|} \mathbf{1}_{\Xi_{4,xy}} \mathbf{1}_{\Xi_{\le 3}}] \le \mathbb{E}\left[\sum_{w\in S_1(z_x)} \mathbb{E}\left[\mathbf{1}_{|S_1(z_y)=v_y|} A_{z_y w} \middle| S_1(z_x)\right] \mathbf{1}_{|S_1(z_x)|} \mathbf{1}_{A_{z_x z_y}=0}\right],$$

and use that on the event $\{A_{z_x z_y} = 0\}$ and for any $z_y \in S_1(x)$ we have

$$\mathbb{E}\left[\mathbf{1}_{|S_1(z_y)=v_y|} A_{z_y w} \middle| S_1(z_x)\right] \le C\frac{v_{z_y}}{N} \mathbb{P}(\mathcal{P}_d - v_y),$$

using again Lemma B.5. Hence

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{|S_1(z_x)=v_x|} \mathbf{1}_{|S_1(z_y)=v_y|} \mathbf{1}_{\Xi_{4,xy}} \mathbf{1}_{\Xi_{\le 3}}] &\le C\frac{d}{N} \mathbb{P}(\mathcal{P}_d - v_y) \mathbb{E}\left[|S_1(z_x)| \mathbf{1}_{|S_1(z_x)=v_x|}\right] \\ &\le C\frac{d}{N} \mathbb{P}(\mathcal{P}_d = v_y) \mathbb{P}(\mathcal{P}_d = v_x)\end{aligned}$$

and therefore

$$|\mathcal{E}_4| \le \prod_{i\in[k]} \mathbb{P}(\mathcal{P}_d = v_i).$$

Next we estimate $\mathcal{E}_5$. We have

$$\mathbb{P}(\Xi_5^c | A_{(B_1([k]))}) \le \mathbb{E}\left[\sum_{1\le i<j\le k} \sum_{x\in S_2(i)} \sum_{y\in S_2(j)} A_{xy} \middle| A_{(B_1([k]))} \mathbf{1}_{\Xi_{\le 4}}\right] \le \frac{d}{N}\left(\sum_{i\in[k]} |S_2(i)|\right)^2$$

We estimate $|S_2(i)| \le (d + Ck\log N)^2$ with probability N^{-k-2} for some universal constant C. We get, using the independence of $(S_1(v_{z_i}))_{i\in[k]}$ on $\Xi_{\le 4}$

$$\mathcal{E}_5 \le \frac{d^3 k^2}{N} \prod_{i\in[k]} \mathbb{P}(\mathcal{P}_d = v_i).$$

We can now turn to $\mathcal{M}_5$ and use the independence of $(S_2(z_i))_{i\in[k]}$ on the event $\Xi_{\le 5}$ This is done exactly as in [7, Proposition 7.1] and we skip the details. □

Appendix A

Matrix theory

In this appendix, we recall some well-known identities for the Green function and its minors.

A.1 Green function identities

In this section, we set $N \in \mathbb{N}^*$ and $M \in \mathbb{C}^{N\times N}$ to be a general Hermitian matrix. Our results are usually stated first for general M and then for M of the form $M = V - H - R$, where $H \in \mathbb{C}^{N\times N}$ is a general Hermitian matrix with zeroes on the diagonal and the matrices V and R are defined in (2.3). In particular let us abbreviate

$$f := \frac{d}{\sqrt{\gamma}}.$$

We also recall the notation $G^{(T)}$, $\widetilde{G}^{(T)}$ and $\widetilde{G}^{(T,u)}$, $u \notin T \subset [N]$, from Definitions 2.10 and 2.13.

Lemma A.1 (Ward identity)**.** *For a general Hermitian matrix M, we have*

$$\sum_{y\in[N]} |G_{xy}(z)|^2 = \frac{\operatorname{Im} G_{xx}}{\operatorname{Im} z}.$$

For $x \notin T \subset [N]$, we have

$$\sum_y^{(T)} |G^{(T)}_{xy}|^2 = \frac{\operatorname{Im} G^{(T)}_{xx}}{\operatorname{Im} z}, \quad \sum_y^{(T)} |\widetilde{G}^{(T)}_{xy}|^2 = \frac{\operatorname{Im} \widetilde{G}^{(T)}_{xx}}{\operatorname{Im} z},$$

Lemma A.2 (Schur's complement formula)**.** *Provided all of the following inverse matrices exist, the inverse of a block matrix is given by*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} -\mathcal{A}^{-1} & -A^{-1}BD^{-1} \\ -D^{-1}C\mathcal{A}^{-1} & D^{-1}C\mathcal{A}^{-1}BD^{-1} + D^{-1} \end{pmatrix}$$
$$= \begin{pmatrix} A^{-1}B\mathcal{D}^{-1}CA^{-1} + A^{-1} & -A^{-1}B\mathcal{D}^{-1} \\ -\mathcal{D}^{-1}CA^{-1} & \mathcal{D}^{-1} \end{pmatrix},$$

where we defined

$$\mathcal{A} := A - BD^{-1}C, \quad \mathcal{D} := D - CA^{-1}B.$$

For a general M, we have

$$\frac{1}{G_{xx}} = M_{xx} - z - \sum_{a,b}^{(x)} M_{xa} G^{(x)}_{ab} M_{by}. \tag{A.1}$$

In particular for M as in (2.5), we have

$$\frac{1}{G_{xx}} = v_x - \frac{f}{N} - z + \sum_{a,b}^{(x)} H_{xa} G_{ab}^{(x)} H_{by} + \frac{f}{N}\left[\sum_{a,b}^{(x)} G_{ab}^{(x)} H_{by} + \sum_{a,b}^{(x)} H_{xa} G_{ab}^{(x)}\right] + \frac{f^2}{N^2}\sum_{a,b}^{(x)} G_{ab}^{(x)}. \tag{A.2}$$

Lemma A.3 (Second resolvent identity)**.** *Let $A, B \in \mathbb{C}^{N\times N}$, for $N \in \mathbb{N}^*$. Then*

$$G_A - G_B = G_A(A-B)G_B. \tag{A.3}$$

In particular for M as in (2.5), $u \notin T \subset [N]$, we have

$$\widetilde{G}_{xy}^{(Tu)} - \widetilde{G}_{xy}^{(T,u)} = -\sum_{a}^{(Tu)} \widetilde{G}_{xa}^{(Tu)} H_{xa} \widetilde{G}_{ay}^{(T,u)} \tag{A.4}$$

On the resolvent identities see for instance [42, Section VI.3].

Proof. Equation (A.3) is a standard result in complex analysis. Setting $A = \tilde{M}^{(Tu)}$ and $B = \tilde{M}^{(T,u)}$, we find that $(A-B)_{xy} = -\delta_{xy} H_{xu}$ for $x, y \notin T \cup \{u\}$ and the result follows. □

Lemma A.4 (Off-diagonal entries of G)**.** *For any Hermitian matrix $M \in \mathbb{C}^{N\times N}$ we have*

$$G_{xy}^{(u)} = -G_{xx}^{(u)} \sum_{a}^{(x)} M_{xa} G_{ay}^{(x)}, \quad x \neq y. \tag{A.5}$$

In particular for M as in (2.5), for $T \subset [N]$, we have

$$\widetilde{G}_{xy}^{(T)} = -\widetilde{G}_{xx}^{(T)}\left[\sum_{a}^{(Tx)} H_{xa}\widetilde{G}_{ay}^{(T,x)} + \frac{f}{N}\sum_{a}^{(Tx)} \widetilde{G}_{ay}^{(T,x)}\right], \quad x \neq y, \quad x, y \in [N]\setminus T. \tag{A.6}$$

Proof. The first two equalities in (A.5) are classical results and can be found for instance in [14, Lemma 3.5]. The second identity (A.6) follows from the definition of M in (2.5) and by applying (A.5) to the matrix $\tilde{M}^{(T)}$. □

Lemma A.5 (Subblocks of G)**.** *For any Hermitian matrix $M \in \mathbb{C}^{N\times N}$ and for $u \in [N]$ and $x, y \neq u$, we have*

$$\begin{aligned} G_{xy}^{(u)} &= G_{xy} - \frac{G_{xu}G_{uy}}{G_{uu}} \\ &= G_{xy} + G_{xu}\sum_{b}^{(u)} M_{ub} G_{by}^{(u)} = G_{xy} + G_{uu}\sum_{a}^{(u)} G_{xa}^{(u)} M_{ua} \sum_{b}^{(u)} M_{ub} G_{by}^{(u)} \end{aligned} \tag{A.7}$$

In particular for M as in (2.5), we have, for $u \notin T \subset [N]$ and $x, y \notin T \cup \{u\}$,

$$\widetilde{G}_{xy}^{(T,u)} - \widetilde{G}_{xy}^{(T)} = \widetilde{G}_{xu}^{(T)}\left[\sum_{a}^{(Tu)} H_{ua}\widetilde{G}_{ay}^{(T,u)} + \frac{f}{N}\sum_{a}^{(Tu)} \widetilde{G}_{ay}^{(T,y)}\right] \tag{A.8a}$$

$$\begin{aligned} &= \widetilde{G}_{uu}^{(T)} \sum_{a}^{(Tu)} \widetilde{G}_{xa}^{(T,u)} H_{au} \sum_{b}^{(Tu)} H_{ub}\widetilde{G}_{by}^{(T,u)} - \frac{f}{N}\widetilde{G}_{uu}^{(T)}\left[\sum_{a}^{(Tu)} \widetilde{G}_{xa}^{(T,u)} H_{au} \sum_{b}^{(Tu)} \widetilde{G}_{by}^{(T,u)}\right. \\ &\quad \left. - \sum_{a}^{(Tu)} \widetilde{G}_{xa}^{(T,u)} \sum_{b}^{(Tu)} H_{ub}\widetilde{G}_{by}^{(T,u)}\right] + \frac{f^2}{N^2}\widetilde{G}_{uu}^{(T)} \sum_{a,b}^{(Tu)} \widetilde{G}_{ab}^{(T,u)}. \end{aligned} \tag{A.8b}$$

Proof. The first two equalities in (A.7) are classical results and can be found for instance in [14, Lemma 3.5]. The second equality follows by applying (A.5) to $\widetilde{G}^{(T)}_{xu}$, since $x \neq u$.
The equalities in (A.8) follow from the definition of M and $\tilde{M}^{(T)}$ and from (A.7) applied to the matrix $\tilde{M}^{(T)}$. □

Observe that by combining (A.4) with (A.8), we find

$$\widetilde{G}^{(Tu)}_{xy} - \widetilde{G}^{(T)}_{xy} = \widetilde{G}^{(T)}_{xu} \sum_a^{(Tu)} \Big(H_{xa} + \frac{f}{N} \Big) \widetilde{G}^{(T,u)}_{ay} - \sum_a^{(Tu)} \widetilde{G}^{(Tu)}_{xa} H_{ua} \widetilde{G}^{(T,u)}_{ay}, \tag{A.9}$$

and the terms $\widetilde{G}^{(T,u)}$ can then be expanded using (A.4)

A.2 Perturbation theory

The following result is [7, Lemma E.1].

Lemma A.6. *Let M be a Hermitian matrix. Let $\varepsilon, \Delta > 0$ satisfy $5\varepsilon \leq \Delta$. Let $\lambda \in \mathbb{R}$ and suppose M has a unique eigenvalue μ in $[\lambda - \Delta, \lambda + \Delta]$, with corresponding normalized eigenvector $\mathbf{w}$. If there exists a normalized eigenvector $\mathbf{v}$ such that $\|(M - \lambda)\mathbf{v}\| \leq \varepsilon$, then*

$$\mu - \lambda = \langle \mathbf{v}, (M - \lambda)\mathbf{v} \rangle + O\Big(\frac{\varepsilon^2}{\Delta}\Big), \quad \|\mathbf{w} - \mathbf{v}\| = O\Big(\frac{\varepsilon}{\Delta}\Big).$$

The following is a standard result that can be found in many reference books, see for instance [46].

Lemma A.7. *Let A, B be Hermitian matrices. The eigenvalues of A and $A - B$ interlace and differ from at most $\|B\|$. Moreover*

$$\frac{1}{z - (A - B)} = \frac{1}{z - A} - \frac{1}{z - A} B \frac{1}{z - (A - B)}, \qquad z \in \mathbb{C} \setminus \big(\operatorname{Spec} A \cup \operatorname{Spec} A - B\big). \tag{A.10}$$

Lemma A.8. *Let $H \in \mathbb{R}^{n \times n}$, $n \in \mathbb{N}^*$, be an Hermitian matrix with eigenvalues λ_i and corresponding eigenvectors v_i, $i \in [n]$. If $e \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^*$, then $\lambda(\theta) \in \mathbb{R}$ is an eigenvalue of $H + \theta e e^*$ if and only if*

$$\frac{1}{\theta} = \sum_{i=1}^n \frac{\langle e, v_i \rangle^2}{\lambda(\theta) - \lambda_i}. \tag{A.11}$$

Proof. Let $v(\theta)$ be the eigenvector of $H + \theta e e^*$ associated to $\lambda(\theta)$, we have $(\lambda(\theta) - H)v(\theta) = \theta \langle e, v(\theta) \rangle e$. Moreover since $\theta \neq 0$, $\lambda(\theta) \notin \operatorname{Spec}(H)$ and therefore $\lambda(\theta) - H$ is invertible. We find

$$\frac{1}{\theta} v(\theta) = \langle e, v(\theta) \rangle (\lambda(\theta) - H)^{-1} e.$$

By applying e^* on both sides of this equation we obtain $\theta^{-1} = \langle e, (\lambda(\theta) - H)^{-1} e \rangle$ which is precisely (A.11). □

Definition A.9 (Paths and cycles)**.** Let $G = (V, A, x_*)$ be some finite graph with vertex set V, adjacency matrix $A \in \{0,1\}^{V \times V}$ and with one distinguished vertex $x_* \in V$. For $x, y \in V$ and $k \in \mathbb{N}^*$, we define

$$\mathcal{P}(x, y, k) := \big\{ P : \{0, \dots, k\} \to V : \ P(0) = x, \ P(k) = y, \ A_{(P(j-1), P(j))} = 1, \ \forall j \in [k] \big\},$$

to be the set of all paths in G of length k starting at x and finishing at y. A path of length k can be seen as a sequence of $k + 1$ elements of V, written $(P(0), P(1), \dots, P(k))$.
For $P \in \mathcal{P}(x, y, k)$, we define

$$\operatorname{Ran}(P) := \{P(i) : \ 0 \leq i \leq k\}, \quad N(P) := \{i \in \{0, \dots, k\} : \ P(i) = x_*\},$$

the range of the path and the times at which the path P visits the distinguish vertex x_* . We also define $N(P)^* := \{0,\dots,k\} \setminus N(P)$.
We denote by $\mathcal{C}(x,k) := \mathcal{P}(x,x,k)$ the set of cycles of length k rooted in x, i.e. paths such that $P(0) = P(k) = x$ and by

$$\mathcal{T}(k) := \{P \in \mathcal{C}(x_*,k), |N(P)| = 2\}, \quad k \in \mathbb{N}^*, \tag{A.12}$$

the set of cycles of length k rooted in x_* which visit x_* exactly twice.

The following is an adaptation of an argument found in [35].

Proposition A.10. *Let $n \in \mathbb{N}^*$, $V = \mathrm{Diag}(v_0,\dots,v_n)$, $v \in \mathbb{R}^{n+1}$, be a diagonal matrix and $G = (\{0,\dots,n\}, A, 0)$ be a graph as in Definition A.9. Let*

$$H := fA, \quad \lambda := v_0, \quad \psi := \min_{i>0}|v_i - \lambda|,$$

for some $f > 0$. Then if $\psi \geq 4\|H\|$, the matrix $V - H$ has precisely one eigenvalue in the interval $[\lambda - \|H\|, \lambda + \|H\|]$ that we denote μ.
If in addition G is a tree, we have, for $1 \leq k \leq n$,

$$\mu = \lambda + \sum_{l=1}^{k-1}\sum_{e=0}^{l-1} E_{2l}(e) + O\Big(\frac{k\|H\|^{2k}(|\lambda| + \psi)}{\psi^{2k}}\Big), \tag{A.13}$$

where

$$E_l(e) := \frac{(-1)^e}{(e+1)!} f^l \sum_{P \in \mathcal{C}(0,l), |N(P)|=2+e} \Big(\prod_{i \in N(P)^*} \frac{1}{v_0 - v_{P(i)}}\Big) \sum_{i_1,\dots,i_e \in N(P)^*} \prod_{j=1}^{e} \frac{1}{v_0 - v_{P(i_j)}}, \quad l \in \mathbb{N}^* \tag{A.14}$$

The letter e in (A.14) stands for *excess*, as the paths that contribute to $E_l(e)$, for some $l, e \geq 1$ visit 0 more than the minimal amount. A cycle that visits 0 five times has excess $5 - 2 = 3$, for instance.
Before proving Lemma A.10 we introduce the following notions.

Definition A.11 (Shift and equivalence classes)**.** Let $G = (V, A, x_*)$ be as in Definition A.9. For $k \in \mathbb{N}^*$, we define the shift operator T_k which maps cycles onto cycles in the following way. For $P \in \mathcal{C}(x,l)$, $x \in V(G)$, we set

$$T_k\big(P(0), P(1), \dots, P(l)\big) := \big(P(k), P(k+1), \dots, P(l-1), P(0), P(1), \dots, P(k)\big).$$

In words, T_k simply changes the root of the cycle by starting and ending at the $(k+1)$th vertex of the cycle. For $l \geq 0$, we define the equivalence relation $\sim_l$ on $\bigcup_{x\in V} \mathcal{C}(x,l)$ by $P \sim_l P'$ if and only if there exists $0 \leq k \leq l$ such that $T_k P = P'$.

Now observe that any cycle P with $|N(P)| > 0$ can be written as $T_k P'$ for some $k \in [l]$, $P' \in \mathcal{C}(0,l)$ with $|N(P')| = |N(P)| + 1$.
We will now partition the space $\mathcal{P}(l) := \bigcup_{x \in V(G)} \mathcal{C}(x,l)$ according to the equivalence relation $P \sim P'$ if and only if there exists $k \in [l]$ such that $P = T_k P'$. Let $P \in \mathcal{P}(l) \setminus \mathcal{T}(l)$ such that $|N(P)| = 1 + e$, i.e. a loop that does not start at 0 but visits zero $1 + e$ times. Then $|\mathcal{T}(l) \cap [P]_\sim| = e + 1$, since the order of the excursions away from zero is fixed but we have $e + 1$ ways to choose the first excursion.

Lemma A.12. *For $l \in 2\mathbb{N}^*$ and $P \in \mathcal{C}(x,l) \setminus \mathcal{T}(l)$ with $|N(P)| \geq 1$, we have*

$$\big|[P]_{\sim_l} \cap \mathcal{T}(l)\big| = |N(P)| - 1,$$

Proof. For a cycle P with $|N(P)| \geq 1$, we define an *excursion* as the trajectory of P between two times in $N(P)$. Let $P \in \mathcal{C}(x,l) \setminus \mathcal{T}(l)$, and $E_0, E_1, \dots, E_{|N(P)|}$. We define $T_i \in \mathcal{T}(l)$, $i = 1,\dots,|N(P)|$, as the paths $(E_i, E_{i+1}, \dots E_0 \cup E_{|N(P)|}, \dots, E_{i-1})$. Then $T_i \in [P]_{\sim_l}$ for every i. Moreover there are no other elements in $[P]_{\sim_l} \cap \mathcal{T}(l)$. This concludes the proof. ∎

Proof of Lemma A.10. We set $M := V - H$. By Lemma A.7 and the assumption that $\psi \geq 4\|H\|$ we find a unique eigenvalue of M that lies in $[\lambda - \|H\|, \lambda + \|H\|]$ which we call μ. Before we start the proof, let us stress the fact that λ is the eigenvalue of V corresponding to the diagonal entry v_0 and μ is its counterpart in $\mathrm{Spec}(M)$.
Applying the second resolvent identity (A.10) $2k$ times, $k \in \mathbb{N}^*$, we find

$$\begin{aligned}\frac{1}{z-M} &= \frac{1}{z-V} + \frac{1}{z-V}H\frac{1}{z-M} = \sum_{i=0}^{1}\frac{1}{z-V}\Big(H\frac{1}{z-V}\Big)^i + \frac{1}{z-V}\Big(H\frac{1}{z-V}\Big)^1 H\frac{1}{z-M} \\ &= \sum_{i=0}^{2k-1}\frac{1}{z-V}\Big(H\frac{1}{z-V}\Big)^i + \frac{1}{z-V}\Big(H\frac{1}{z-V}\Big)^{2k-1}H\frac{1}{z-M},\end{aligned} \tag{A.15}$$

for any $z \in \mathbb{C} \setminus \big(\mathrm{Spec}(M) \cup \mathrm{Spec}(V)\big)$.
Let us consider the contour in the complex plan $\Gamma := \partial B_{2\psi}(\lambda)$. From the assumption $\psi \geq 4\|H\|$, we conclude

$$\mathrm{Int}(\Gamma) \cap \big(\mathrm{Spec}(M) \cup \mathrm{Spec}(V)\big) = \big\{\lambda, \mu\big\}, \quad \min\Big(\inf_{z\in\Gamma}|z-\lambda|, \inf_{z\in\Gamma}|z-\mu|\Big) \geq \psi. \tag{A.16}$$

Using Cauchy's integral formula we conclude that

$$\mu = \frac{1}{2\pi\mathrm{i}}\oint_\Gamma \mathrm{Tr}\Big(\frac{z}{z-M}\Big)\mathrm{d}z = \sum_{l=0}^{2k-1} P(l) + R(2k-1), \tag{A.17}$$

where we introduced, for $l \in \mathbb{N}$

$$P(l) := \frac{1}{2\pi\mathrm{i}}\oint_\Gamma \mathrm{Tr}\bigg(\frac{z}{z-V}\Big(H\frac{1}{z-V}\Big)^l\bigg)\mathrm{d}z, \quad R(l) := \frac{1}{2\pi\mathrm{i}}\oint_\Gamma \mathrm{Tr}\bigg(\frac{z}{z-V}\Big(H\frac{1}{z-V}\Big)^l H\frac{1}{z-M}\bigg)\mathrm{d}z. \tag{A.18}$$

We will show that

$$P(l) = \mathbf{1}_{l\in 2\mathbb{N}}\bigg(\sum_{P\in\mathcal{T}(l)}\prod_{i\notin N(P)}\frac{f}{\lambda - v_i} + \sum_{e=1}^{l-1}E_{2l}(e)\bigg), \quad R(l) = O\Big(l\frac{\|H\|^{l+1}}{\psi^l}\frac{|\lambda|\vee\psi}{\psi}\Big). \tag{A.19}$$

Owing to the definitions in (A.18), we see that (A.17) becomes $\mu = \sum_{l=0}^{k-1} P(2l) + R(2k-1)$ and, given (A.19), we will be able to conclude (A.13).
We begin with the control of R. Let us denote by $\{e_i\}_{0\leq i\leq n}$ the canonical basis, which is also an eigenvectors basis of V, and by $\{u_\beta\}_{\beta\in\mathrm{Spec}(M)}$ the eigenvectors of M. Using the identities

$$1 = \sum_{i=0}^n e_i e_i^* = \sum_{\beta\in\mathrm{Spec}(M)} u_\beta u_\beta^*, \qquad \mathrm{Tr}\, B = \sum_{i=0}^n e_i^* B e_i = \sum_{\beta\in\mathrm{Spec}(M)} u_\beta^* B u_\beta, \quad B \in \mathbb{C}^{(n+1)\times(n+1)}, \tag{A.20}$$

for $l \geq 1$, we find

$$(2\pi\mathrm{i})R(l) = \sum_{i_0,\dots,i_l=0}^n \sum_{\beta\in\mathrm{Spec}(M)} \oint_\Gamma \frac{z}{z-v_{i_0}}\prod_{j=1}^l \frac{H_{i_{j-1},i_j}}{z-v_{i_j}}\langle e_{i_l}\,, Hu_\beta\rangle\frac{1}{z-\beta}\langle u_\beta\,, e_{i_0}\rangle\mathrm{d}z.$$

By Cauchy's theorem, if the integrand has no poles in the interior of Γ, then the integral vanishes. We conclude that at least one of the summands must create a pole. Therefore one of the $l+2$ sums is actually trivial, as it is reduced for instance to $i_j = 0$ or $\beta = \mu$. Using (A.20) one more time, we can resorb the $l+1$ other sums to conclude

$$\begin{aligned}(2\pi\mathrm{i})R(l) &= \sum_{k=0}^l \oint_\Gamma \frac{z}{z-\lambda}\Big\langle e_0\,,\, \Big(H\frac{1}{z-V}\Big)^{l-k} H\frac{1}{z-M}\Big(\frac{1}{z-V}H\Big)^k e_0\Big\rangle\mathrm{d}z \\ &+ \oint_\Gamma \frac{z}{z-\mu}\Big\langle u_\mu\,,\, H\frac{1}{z-V}\Big(\frac{1}{z-V}H\Big)^l u_\mu\Big\rangle\mathrm{d}z.\end{aligned}$$

Using (A.16), we find

$$|R(l)| \le \frac{l+2}{2\pi} \oint_\Gamma \mathrm{d}z \sup_{z\in\Gamma}\left(\left|\frac{z}{z-\lambda}\right| + \left|\frac{z}{z-\mu}\right|\right)\frac{\|H\|^{l+1}}{\psi^{l+2}} \le 4(l+2)|\Gamma|\frac{\|H\|^{l+1}}{\psi^{l+1}}\frac{|\lambda|\vee\psi}{\psi}.$$

Using the fact that $|\Gamma| = O(\psi)$ proves the second equality in (A.18).

We will now control $P(l)$ using the geometric properties of the graph G. Since the canonical basis is an eigenvector basis of V, we see, using (A.20), that the terms that contribute to the trace are in one-to-one correspondence with elements of $\bigcup_{x\in V(G)} \mathcal{C}(x,l)$. In particular, since G is a tree, there are no odd-length cycles and so $P(l) = 0$ when l is odd. Moreover by (A.16) and Cauchy's theorem, we see that cycles that do not visit the distinguished vertex 0 will have a vanishing contribution after integration around Γ. We conclude

$$(2\pi\mathrm{i})P(l) = \mathbf{1}_{l\in 2\mathbb{N}} f^l \sum_{x\in V(G)} \sum_{P\in\mathcal{C}(x,l), |N(P)|>0} \oint_\Gamma z\prod_{i=0}^{l} h_{P(i)}(z)\mathrm{d}z, \tag{A.21}$$

where we introduced the meromorphic functions

$$h_i(z) := \frac{1}{z-v_i}, \quad 0\le i\le n,\, l\in\mathbb{N},\, z\in\mathbb{C}\setminus\{v_i\}.$$

We will now enumerate all the paths that contribute to the right-hand side of (A.21) using elements of $\mathcal{C}(0,l)$. Let us recall Definition A.11.

Let us write $\mathcal{P}(2l)$, the set of equivalence of $\mathcal{T}(2l)$. For every equivalence class $C\in\mathcal{P}(2l)$, we choose a representative $P_C \in C\cap\mathcal{T}(2l)$. Then by Lemma A.12, we find

$$\bigcup_{x\in V(G)} \bigcup_{C\in\mathcal{C}(x,l), |N(P)|>0} \{P\} = \bigcup_{C\in\mathcal{P}(2l)}\left(\bigcup_{P\in C\cap\mathcal{T}(2l)}\{P\} \cup \bigcup_{k\in[l]\setminus N(P_C)}\{T_kP_C\}\right). \tag{A.22}$$

Therefore to compute (A.21) it suffices to study the contribution of every term on the right-hand side of (A.22) separately. Let $P\in\mathcal{C}(0,l)$. The poles in $\mathrm{Int}(\Gamma)$ of the function $z\prod_i h_{P(i)}(z)$ correspond to elements of $N(P)$. Using Cauchy's integral formula we find, for some $C\in\mathcal{P}(2l)$, denoting by P_C a representative of C which is in $\mathcal{T}(2l)$. Let us write $\nu := |N(P_C)|$. Using Cauchy's integral formula we find

$$\begin{aligned}
&\frac{1}{2\pi\mathrm{i}}\oint_\Gamma \sum_{P\in C} z\prod_{i=0}^{l} h_{P(i)}(z)\mathrm{d}z = \frac{1}{2\pi\mathrm{i}}\oint_\Gamma\left(z\sum_{k\in N(P_C)}\prod_{i=0}^{l} h_{T_kP_C(i)}(z) + z\sum_{k\notin N(P_C)}\prod_{i=0}^{l} h_{T_kP_C(i)}(z)\right)\mathrm{d}z\\
&= \frac{1}{2\pi\mathrm{i}}\oint_\Gamma\left(z\prod_{i=0}^{l} h_{P_C(i)}(z)(\nu-1) + z\sum_{k\notin N(P_C)}\prod_{i=0}^{l} h_{T_kP_C(i)}(z)\right)\mathrm{d}z\\
&= \frac{\nu-1}{(\nu-1)!}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\nu-1}\left(z\prod_{i\notin N(P_C)} h_{P_C(i)}(z)\right)\Bigg|_{z=\lambda} + \frac{1}{(\nu-2)!}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\nu-2}\left(z\sum_{k\notin N(P_C)}\prod_{i\notin N(T_kP_C)} h_{T_kP_C(i)}(z)\right)\Bigg|_{z=\lambda}.
\end{aligned}$$

We used the fact that paths in $\mathcal{T}(2l)\cap C$ have ν poles inside Γ while paths in $C\setminus\mathcal{T}(2l)$ have $\nu-1$ poles. This explains the different orders of the derivatives. Now observe that since there are exactly $|\mathcal{T}(2l)\cap C| = \nu-1$ and since the contribution of any two paths $P, P'\in\mathcal{T}(2l)\cap C$ is the same, we have

$$\begin{aligned}
&\frac{1}{2\pi\mathrm{i}}\oint_\Gamma \sum_{P\in C} z\prod_{i=0}^{l} h_{P(i)}(z)\mathrm{d}z\\
&\quad= \frac{1}{(\nu-2)!}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\nu-2}\left[\frac{\mathrm{d}}{\mathrm{d}z}\left(z\prod_{i\notin N(P_C)} h_{P_C(i)}(z)\right) + z\sum_{k\notin N(P_C)}\prod_{i\notin N(T_kP_C)} h_{T_kP_C(i)}(z)\right]\Bigg|_{z=\lambda}.
\end{aligned}$$

We now use the fact that

$$\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^l h_i(z) = (-1)^l l! h_i(z)^{l+1}, \quad z \in \mathbb{C} \setminus \{v_i\}, \quad 0 \le i \le n, \quad l \ge 0. \tag{A.23}$$

and the fact that the vertex $P(k)$ is visited one more time by $T_k P_C$ than by P_C to pull out a factor $h_{P(k)}(z)$ from the sum on right-hand side. We find

$$\begin{aligned}
&\frac{1}{2\pi \mathrm{i}} \oint_\Gamma \sum_{P\in C} z \prod_{i=0}^{l} h_{P(i)}(z)\mathrm{d}z \\
&= \frac{1}{(\nu-2)!}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\nu-2}\left[\prod_{i\notin N(P_C)} h_{P_C(i)}(z)\Big(1 - z\sum_{k\notin N(P_C)} h_{P(k)}(z)\Big) + z\sum_{k\notin N(P_C)} h_{P_C(k)}(z)\prod_{i\notin N(P_C)} h_{P_C(i)}(z)\right]\Bigg|_{z=\lambda} \\
&= \frac{1}{(\nu-2)!}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\nu-2}\left[\prod_{i\notin N(P_C)} h_{P_C(i)}(z)\right]\Bigg|_{z=\lambda} = \frac{1}{(\nu-1)!}\sum_{P\in C\cap\mathcal{T}(2l)}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\nu-2}\left[\prod_{i\notin N(P_C)} h_{P_C(i)}(z)\right]\Bigg|_{z=\lambda},
\end{aligned}$$

where in the last step, we used the fact that $C \cap \mathcal{T}(2l) = \nu - 1$ and all paths that set have the same contribution.
Now since

$$\sum_{P\in\mathcal{C}(0,l),|N(P)|\ge 2} (\star) = \sum_{P\in\mathcal{C}(0,l),|N(P)|=2} (\star) + \sum_{e\ge 1}\sum_{P\in\mathcal{C}(0,l),|N(P)|=2+e} (\star), \tag{A.24}$$

we see $E_0(2l)$ in (A.19) appears naturally as the contribution of all $\mathcal{T}(l)$, that is of all cycles that visit 0 exactly twice, namely that start and end at 0.
If $|N(P)| = 2 + e$, for some $e \in \mathbb{N}^*$, where the letter e stand for excess, we again use (A.23) to find

$$\frac{1}{(e+1)!}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{e}\left[\prod_{i\notin N(P)} h_{P(i)}(z)\right] = (-1)^e\Big(\prod_{i\notin N(P)} h_{P(i)}(z)\Big)\sum_{i_1,\dots,i_e\notin N(P)}\prod_{j=1}^{e} h_{P(i_j)}(z).$$

Now since there are no loops in the graph, we see that $e \le \frac{l}{2} - 1$. Indeed every path in $P \in \mathcal{C}(0,l)$ can come back to 0 at most $l/2$ times. By accounting for the fact that $P(0) = 0 = P(l)$, $|N(P)| \le \frac{l}{2} + 1$ and so $e \le \frac{l}{2} - 1$. Imposing the last condition in (A.24) and then into (A.21) shows (A.19). This concludes the proof. □

A.3 Rayleigh-Schrödinger coefficients

In Lemma A.10, we provide an elegant geometric interpretation of perturbation theory in the specific case where we perturb a diagonal matrix by the adjacency of a graph with an underlying tree structure. In this section, we explain how this idea generalizes to the so-called Rayleigh-Schrödinger coefficients.

Let V and H be matrices and

$$M(x) := V + xH, \qquad x \in \mathbb{C}.$$

Suppose $\lambda \in \mathrm{Spec}(V)$ is a simple eigenvalue of V with associated eigenvector $\mathbf{v}$.

It is known (see [35]) that there exist an open set $0 \in U \subsetneq \mathbb{C}$ and analytic functions $\lambda(x)$ and $\mathbf{v}(x)$ defined on U such that $\lambda(0) = \lambda$, $\mathbf{v}(0) = \mathbf{v}$ and $(\lambda(x), (v))$ is the eigenvalue-eigenvector pair of $M(x)$ for all $x \in U$. The two natural questions are

1. What is the radius of convergence of the functions $\lambda(x)$ and $\mathbf{v}(x)$?

2. What are the coefficients of power series that describe the two functions?

Let P and Q denote the spectral projections of H onto the eigenspace spanned by $\mathbf{v}$ and the generalized eigenspaces of all other eigenvalues, respectively. Let

$$S := -Q(H-\lambda)^{-1}Q,$$

denote the reduced resolvent. Since the eigenvalue λ is simple, the matrix S is well-defined.

In order to solve the perturbation, it is customary to choose $\mathbf{u}_0$ to be normalized, $\|\mathbf{u}_0\| = 1$, so that $Pu(x) := \mathbf{u}_0$ and $P\mathbf{u}_n = 0$ and $Q\mathbf{u}_n = \mathbf{u}_n$ for all $n \geq 1$.

Proposition A.13. *Let*

$$\lambda_n = \mathbf{u}_0^* H \mathbf{u}_{n-1}, \quad \mathbf{u}_n = SH\mathbf{u}_{n-1} - \sum_{k=1}^{n-1} \lambda_{n-k} S\mathbf{u}_k, \quad n \geq 0, \tag{A.25}$$

with $\lambda_0 = \lambda$ *and* $\mathbf{u}_0 = \mathbf{u}$. *Let*

$$\varepsilon = \|S\|\|H\| \max(1, \|P\|).$$

The power series expansions

$$\lambda(x) = \sum_{k\geq 0} \lambda_k x^k, \qquad \mathbf{u}(x) = \sum_{k\geq 0} \mathbf{u}_k x^k,$$

converge for $|x| \leq \frac{1}{4\varepsilon}$. *Moreover, there is a universal constant* $C > 0$ *such that for* $|x| \leq \frac{1}{15\varepsilon}$

$$\left\| \mathbf{u}(x) - \sum_{m=0}^{n} \mathbf{u}_m x^m \right\| \leq C\big(15\varepsilon|x|\big)^{n+1}, \qquad \left\| \lambda(x) - \sum_{m=0}^{n} \lambda_m x^m \right\| \leq C\|P\|\|H\|(15\varepsilon|x|)^n. \tag{A.26}$$

Before showing Proposition A.13, we make a couple of remarks. In quantum mechanics class, the Rayleigh-Schrödinger are introduced by comparing the coefficients in the eigenvalue-eigenvector equation (see for instance [43, Chapter 11])

$$(V + xH)\big(\mathbf{v}_0 + x\mathbf{v}_1 + x^2\mathbf{v}_2 + \ldots\big) = (\lambda_0 + x\lambda_1 + x^2\lambda_2 + \ldots)\big(\mathbf{v}_0 + x\mathbf{v}_1 + x^2\mathbf{v}_2 + \ldots\big). \tag{A.27}$$

The zeroth order term is just the unperturbed eigenvalue-eigenvectors equation $V\mathbf{v} = \lambda\mathbf{v}$ and the first order equation is

$$H\mathbf{v}_1 + V\mathbf{v}_0 = \lambda_0\mathbf{v}_1 + \lambda_1\mathbf{v}_0.$$

Pushing this analysis to the second order already leads to solving problems like the perturbed quantum harmonic oscillator or the Stark effect. However, the size of the error incurred by stopping at order $k \in \mathbb{N}^*$ is a question that is rarely addressed.

In the context of Proposition A.10, we can identify the matrix H with fA, $\mathbf{v}$ with $\mathbf{1}_0$, λ with v_0 and Q with the projection on all the vertices except x_0. Let us set $f = 1$. In particular, the first and second order terms of the perturbation can be read from (A.25) as

$$\lambda_1 = \mathbf{1}_0^* fA\mathbf{1}_0 = 0, \quad \mathbf{v}_1 = -Q\frac{1}{H - v_0}QA\mathbf{u}_0 = -Q\frac{1}{H - v_0}Q\mathbf{1}_{S_1(0)} = \sum_{x\in S_1(0)} \frac{1}{v_0 - v_x}\mathbf{1}_x,$$
$$\lambda_2 = \mathbf{1}_0^*\Big(\sum_{x\in S_1(0)} \frac{1}{v_0 - v_x}\mathbf{1}_x\Big)\mathbf{1}_0 = \sum_{x\in S_1(0)} \frac{1}{v_0 - v_x}.$$

We recognize in the last line the expression for $E_2(0)$ and in the identity $\lambda_1 = 0$ the fact that the odd rank terms do not contribute. Note how the reduced resolvent S is the analog of (A.15). However, the reader can already appreciate how the formulation in terms of paths from (A.14) is easier to understand than the recursive formulas from (A.25).

Proof of Proposition A.13. Let us abbreviate $a_n := \|\mathbf{u}_n\|$. Recalling $\|\mathbf{u}_0\| = 1$ and the definition of ε, and plugging the left-hand side of (A.25) into the right-hand side, we get the recursion

$$a_0 = 1, \qquad a_{n+1} \leq \varepsilon \sum_{k=0}^{n} a_{n-k} a_k. \tag{A.28}$$

We claim that the power series

$$f(x) := \sum_{n=0}^{\infty} a_n x^n,$$

converges in a neighborhood of the origin.
We give two proofs. First let us consider the function $g(x) := \sum_{n=1}^{\infty} b_n x^n$ where $b_0 = 1$ and $b_{n+1} := \varepsilon \sum_{k=0}^{n} b_k b_{n-k}$. Then by (A.28), we get $a_n \leq b_n$ and thus f converges absolutely whenever g does.
Now observe that formally the recursive definition of the coefficients b_n leads to

$$g(x)^2 = \left(\sum_{n=0}^{\infty} b_n x^n\right)^2 = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} b_k b_{n-k}\right)x^n = \sum_{n=0}^{\infty} \frac{1}{\varepsilon} b_{n+1} x^n = \frac{1}{\varepsilon x}\left(\sum_{n=0}^{\infty} b_n x^n - b_0\right) = \frac{1}{\varepsilon x}\big(g(x) - 1\big).$$

Hence g satisfies the equation $\varepsilon x g^2 - g + 1 = 0$ which can be solved explicitly for the initial condition $g(0) = 1$,

$$g(x) = \frac{1 + \sqrt{1 - 4\varepsilon x}}{2}.$$

This is analytic in the disc of radius $(4\varepsilon)^{-1}$ around the origin.
The second proof proceeds by induction, showing that

$$a_n \leq \frac{r^n}{(n+1)^2}. \tag{A.29}$$

The case $n = 0$ is clear since $a_0 \leq 1$. For the induction we get

$$\begin{aligned} \varepsilon \sum_{k=0}^{n} \frac{r^n}{(n-k+1)^2 (k+1)^2} &\leq 2\varepsilon r^n \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{1}{(n-k+1)^2 (k+1)^2} \\ &\leq \frac{8\varepsilon r^n}{(n+2)^2} \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{1}{(k+1)^2} \leq \frac{8\varepsilon r^n}{(n+2)^2} \frac{\pi^2}{6} \leq \frac{r^{n+1}}{(n+2)^2}, \end{aligned}$$

where we chose

$$r := \frac{4\pi^2}{3}\varepsilon.$$

The second proof gives a slightly worse radius of convergence. However, we can use (A.29) to bound the error term, for $\varepsilon|x| < 1/5$,

$$\left\| \mathbf{u}(x) - \sum_{m=0}^{n} \mathbf{u}_m x^m \right\| \leq \sum_{k \geq n+1} a_k x^k \leq \sum_{k \geq n+1} \frac{(4\pi^2 \varepsilon x)^k}{3^k (k+1)^2} \leq (15\varepsilon|x|)^{n+1} \sum_{k=0}^{\infty} \frac{c^k}{(k+1)^2},$$

for some $c < 1$. Setting C to bound the series on the right-hand side, we conclude the first estimate of (A.26).
The right-hand side of (A.26) is follows from the inequality

$$|\lambda_n| \leq \|P\| \|H\| \|\mathbf{u}_{n-1}\|.$$

This concludes the proof. □

Appendix B

Probability appendix

Proposition B.1 (Central Limit Theorem). *Let X_i, $i \in \mathbb{N}$ be a sequence of independent identically distributed centered random variables with variance 1 and bounded moments. Then, if $S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$, for any $-\infty < a < b < \infty$ we have*

$$\mathbb{P}(S_n \in [a,b]) = (1+o(1)) \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-t^2/2} \mathrm{d}t.$$

Proposition B.2 (Stirling's approximation). *For any $n \geq 1$, we have*

$$\sqrt{2\pi n} \Big(\frac{n}{\mathrm{e}}\Big)^n \mathrm{e}^{\frac{1}{12n+1}} \leq n! \leq \sqrt{2\pi n} \Big(\frac{n}{\mathrm{e}}\Big)^n \mathrm{e}^{\frac{1}{12n}}.$$

B.1 Statistics of Bernoulli and Poisson distributions

Throughout this section, we denote by $\mathcal{B}_{n,p}$ a Bernoulli distribution with parameters $n \in \mathbb{N}^*$ and $p \in [0,1]$ and by $\mathcal{P}_d$ a Poisson distribution with parameter $d \geq 0$.
We introduce the function

$$h : [0,+\infty) \to [0,+\infty), \quad h(\alpha) = (1+\alpha)\log(1+\alpha) - \alpha. \tag{B.1}$$

Note that often we use the Taylor expansion of h to the second order in the neighborhood of zero

$$h(a) = \frac{a^2}{2} - \frac{a^3}{6(1+t(a))^2}, \quad a \geq 0.$$

Lemma B.3 (Benett's inequality). *For $0 \leq \mu \leq n$, $a > 0$, we have*

$$\mathbb{P}(\mathcal{B}_{n.\mu/n} - \mu \geq a\mu) \leq \mathrm{e}^{-\mu h(a)}, \quad \mathbb{P}(\mathcal{B}_{n.\mu/n} - \mu \leq -a\mu) \leq \mathrm{e}^{-\mu a^2/2} \leq \mathrm{e}^{-\mu h(a)}, \tag{B.2}$$

and $\frac{a^2}{2(1+a/3)} \leq h(a) \leq \frac{a^2}{2}$.

Lemma B.4 (Comparaison of Poisson and normal laws). *For $0 \leq \xi \leq 1/6$. Then for $\mu \geq 1$ and $t \leq \mu^{\xi}$ we have*

$$\mathbb{P}[\mathcal{P}_\mu \geq \mu + \sqrt{\mu} t] = G(t)(1 + O(\mu^{3\xi - 1/2})).$$

Lemma B.5 (Poisson approximation of a Bernoulli). *Let $n \in \mathbb{N}^*$, $Y \sim \mathcal{P}_d$, $Z \sim \mathcal{B}_{n,d/n}$. Then there is a universal constant $C > 0$ such that if $d, k \leq \sqrt{n}/C$*

$$\left| \frac{\mathbb{P}[Z=k]}{\mathbb{P}[Y=k]} - 1 \right| = O(k^2/n). \tag{B.3}$$

Moreover if $k < (1-\varepsilon)d$ for $\varepsilon > 0$ then there exists $C = C(\varepsilon) > 0$ such that

$$\mathbb{P}[Z < k] = C\mathbb{P}[Z = k]\frac{k}{d}\big[1 + O(d^2/n)\big]. \tag{B.4}$$

Moreover

$$\mathbb{P}(\mathcal{B}_{n,d/n} = v) = \mathbb{P}(\mathcal{P}_d = v)\Big(1 + O\big(\frac{v^2+d^2}{n}\big)\Big). \tag{B.5}$$

Proof. This is a restatement of [11, Lemma 3.3] for lower tails of Poisson variables, that is for values of k smaller than d. The first statement is proven similarly since the only condition used is $k, d \le \sqrt{N}$. For the second equation

$$\begin{aligned}
\mathbb{P}[Z < k] =&\mathbb{P}[Z = k]\sum_{l=1}^{k-1} d^{-l}\big(1 - \frac{d}{N}\big)^l \prod_{j=0}^{l-1} \frac{k-j}{1 - \frac{k-j}{N}} \\
\le&\mathbb{P}[Z = k]\sum_{l=1}^{k}\Big(\frac{k}{d}\Big)^l \frac{\exp\Big(-\frac{ld}{N} + O\big(\frac{ld^2}{N^2}\big)\Big)}{\exp\big(-\frac{k}{N} + O\big(\frac{k^3}{N}\big)\big)} \\
\le&\mathbb{P}[Z = k]\frac{k}{d}\frac{(k/d)^k}{1-(k/d)}\exp\Big(\frac{k - ld}{N} + O\big(\frac{ld^2 + Nk^3}{N^2}\big)\Big) \\
=&C_\varepsilon\mathbb{P}[Z = k]\frac{k}{d}\big[1 + O(d^2/N)\big].
\end{aligned}$$

□

B.2 Extreme value statistics

Remark B.6. For $d \gg \log N$, the extreme value statistics of $\mathcal{P}(d)$ are well-described by the extreme value statistics of $\mathcal{N}(0,1)$. Indeed for $k = d + \sqrt{d}a$, for some $k \in \mathbb{N}$ we have

$$\log k! = k\log d + k\log\big(1 + ad^{-1/2}\big) - k + \frac{1}{2}\log(2\pi k)(1 + o(1))$$

and thus, writing $h(x) = (x+1)\log(x+1) - x$,

$$\begin{aligned}
\mathbb{P}[Y = k] &= \frac{d^k}{k!}\mathrm{e}^{-d} = \exp\big(k\log d - \log k! - d\big) = \exp\big(dh\big(1 + ad^{-1/2}\big)(1 + o(1))\big) \\
&= \mathrm{e}^{-a^2/2 + O\left(a^3 d\right)}.
\end{aligned}$$

The error term is $o(1)$ as soon as $|k-d| \ll d$. This is guaranteed for $d \gg \log N$ as Lemma B.7 shows.

Recall that if $Y \sim \mathcal{P}_d$

$$\mathbb{P}(Y = k) = \exp\Big(f(k/d) + O(1/k)\Big).$$

where $f = f_d$ was defined in (3.2).

Lemma B.7 (Upper tail of a Poisson distribution)**.** *Let $d \le \log N$, $Y \sim \mathcal{P}_d$ and $\mathfrak{u}_+$ be defined in (3.3). Then there exists $C > 0$ and $c \in (0,1)$ such that*

$$d\mathbf{x}_+ = \frac{\log N - d + \log(\log N/d)}{1 + c\log(\log N/d)}. \tag{B.6}$$

and, for $k \in \mathbb{Z}$ we have

$$\frac{\mathbb{P}[\mathcal{P}_d = \lfloor dm\rfloor + k]}{\mathbb{P}[\mathcal{P}_d = \lfloor dm\rfloor]} = \mathfrak{u}_+^k\big[1 + O\big(\frac{k^2}{d\mathfrak{u}_+}\big)\big] \tag{B.7}$$

Proof. We have

$$x_* \log(x_*/d) + x_* = \log N - d + \log(\sqrt{2\pi x_*})$$

which we estimate from below using $x* \leq C \log N$, for $C > 0$ large enough and from above using $x \geq d$. This proves (B.6).

To prove (B.7) we use Stirling's approximation to find

$$\frac{\mathbb{P}[\mathcal{P}_d = \lfloor d\mathfrak{u}_+ \rfloor + k]}{\mathbb{P}[\mathcal{P}_d = \lfloor d\mathfrak{u}_+ \rfloor]} = \left(\frac{d}{\lfloor d\mathfrak{u}_+ \rfloor}\right)^k \frac{1}{\prod_{i=0}^{k-1}(1+\frac{i}{l})}.$$

A Taylor expansion on the product allows us to conclude. □

In the next lemma, we use n instead of N to distinguish the case where we are looking at the minimal degree on the whole graph or at the minimal degree on the set of neighbors of leaves.

Lemma B.8 (Lower Tail of a Poisson distribution)**.** *Let $n \in \mathbb{N}$ and $d \geq 1$, $Y \sim \mathcal{P}_d$ and define*

$$m_*(n) := f_d^{-1}\left(\frac{\log n}{d}\right) \vee \frac{1}{d}, \quad m_* < 1$$

as well as

$$\alpha(k) := \frac{\mathbb{P}[\mathcal{P}_d = \lfloor dm_* \rfloor + k]}{\mathbb{P}[\mathcal{P}_d = \lfloor dm_* \rfloor]}, \quad k \in \mathbb{Z},$$

and d^* is the solution of the implicit equation $d - \log d = \log n$.

1. *If $d^* \leq d = O(\log n)$ there exists $a \in (0,1)$ depending on n such that*

$$dm_*(n) = \frac{g}{1 + a\log d + O(\log d/d)}, \quad \alpha(k) = (m_*)^k \left[1 + O\left(\frac{k^2}{dm_*}\right)\right]. \tag{B.8}$$

2. *If $d \leq d^*$ then $dm_*(n) = 1$ and $\alpha(k) = d^k(1 + O(d^{-1}))$.*

Proof. According to (B.3) we must solve

$$1 = n\exp(-df_d(k/d)). \tag{B.9}$$

Let us define $g := g(n,d) = d - \log n$. Taking log on both sides we find

$$g(n) = k\left[1 + \log(d) - \log(k) - \frac{1}{2k}\log(2\pi k)\right].$$

We now estimate $h(k) := \log k + \frac{1}{2k}\log(2\pi k)$ from below by $1 - \log(\sqrt{2\pi}) \geq 0.05$ for $k \geq 1$ and from above by $\log d(1 + O(d^{-1}))$. Because $h(k)$ is monotonously increasing for $k \geq 1$ we find

$$\frac{g(n)}{1 + O(\log d/d)} \leq k_* \leq \frac{g(n)}{1 + \log d(1 + O(d^{-1}))}.$$

Observe that $dm_* \geq 1$ and deduce (B.8). To prove the statement regarding $\alpha(k)$ we recall the Stirling approximation, if $Z \sim \mathcal{P}_d$ then

$$\frac{\mathbb{P}[\mathcal{P}_d = \lfloor dm_* \rfloor + k]}{\mathbb{P}[\mathcal{P}_d = \lfloor dm_* \rfloor]} = \left(\frac{d}{l}\right)^k \frac{1}{\prod_{i=0}^{k-1}(1-\frac{i}{l})}.$$

We conclude with a Taylor expansion of the second term.

The second point follows from analysing f and observing that d^* is the threshold below which $ne^{-df(1/d)} \geq 1$. □

Appendix C

Notations

Generic mathematical notations

- $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ usual number sets.
- $[N] := \{1, \dots, N\}$, $N \in \mathbb{N}$.
- $f \ll g$ or $f = o(g)$: $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.
- $f \gg g$: $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty$.
- $f = O(g)$: $\exists C > 0$ such that $f \leq Cg$
- $f \asymp g$: $\exists C, c > 0$ such that $cg \leq f \leq Cg$.
- $\approx$: generic estimation symbol. Has no specific meaning and is used for informal discussions.
- $\|\cdot\| \equiv \|\cdot\|_2$: the ℓ^2 norm. For $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\|^2 := \sum_{k=1}^n |\mathbf{v}(k)|^2$.
- $\|\cdot\|_\infty$: the ℓ^∞ norm. For $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\|_\infty := \sup_{k=1,\dots n} |\mathbf{v}(k)|$.

Generic probability notations

- $\mathrm{Bern}(p)$ Bernoulli random variable of parameter $p \in [0, 1]$.
- $\mathcal{B}_{n,p}$ binomial distribution of parameters $n \in \mathbb{N}^*$ and $p \in [0, 1]$.
- $\mathcal{P}_d$ Poisson random variable of parameter $d \in \mathbb{R}_{\geq 0}$.
- $X \overset{\text{(d)}}{=} Y$: X is equal in distribution to Y.
- $X_1, \dots X_n \overset{\text{i.i.d.}}{\sim} Y$: the variables X_i, $i = [n]$, are independent identically distributed with law Y.
- $X_n \Rightarrow X$: the variables X_i, $i \geq 1$, converge in distribution towards the law of X.

Graph notations

- $d = d_{\mathbb{G}}$: graph distance.
- $S_i(x) := \{y \in V(\mathbb{G}) : d(x, y) = i\}$: sphere of radius i around x, $i \in \mathbb{N}$ and $x \in V(\mathbb{G})$.
- $B_i(x) := \{y \in V(\mathbb{G}) : d(x, y) \leq i\}$: ball of radius i around x, $i \in \mathbb{N}$ and $x \in V(\mathbb{G})$.
- $\mathbb{G}|_T$, $T \subset [N]$: the restriction of the graph to the subset of vertices T,

$$\mathbb{G}|_T := \big(T, \{(x, y) \in E(\mathbb{G}) : x, y \in T\}\big).$$

- $\mathrm{Deg}(x) = D_x$: degree of the vertex $x \in [N]$.
- $v_x := \frac{D_x - d}{\sqrt{d}}$, $x \in [N]$.
- $\alpha_x := \frac{D_x}{d}$, $x \in [N]$.
- $\beta_x := \frac{|S_2(x)|}{|S_1(x)|d} - 1$, $x \in [N]$.

Generic matrix notations.

- For G a graph, $A(G)$ is the adjacency matrix of G, $D(G)$ the matrix of degrees and $L(G) := D(G) - A(G)$ the Laplacian of G.
- $M|_T$ for $M \in \mathbb{R}^{N \times N}$ and $T \subset [N]$: the restriction of the matrix M to T,

$$M|_T := \Big(M_{xy} : x, y \in T\Big).$$

- $M^{(T)}$ for $M \in \mathbb{R}^{N \times N}$ and $T \subset [N]$: the restriction of the matrix M to T^c,

$$M^{(T)} := \Big(M_{xy} : x, y \notin T\Big).$$

- For $M \in \mathbb{R}^{N \times N}$ a Hermitian matrix, we denote by $\lambda_1(M) \leq \lambda_2(M) \leq \cdots \leq \lambda_N(M)$ the eigenvalue of M sorted decreasingly.
- For $M \in \mathbb{R}^{N \times N}$ we denote by $\operatorname{Spec} M$ the spectrum of M. For $\lambda \in \operatorname{Spec} M$, we denote by $\mathbf{w}_\lambda$ the eigenvector associated to λ.

Bibliography

[1] Michael Aizenman, Mira Shamis, and Simone Warzel. Resonances and partial delocalization on the complete graph. In *Annales Henri Poincaré*, volume 16, pages 1969–2003. Springer, 2015.

[2] Michael Aizenman and Simone Warzel. *Random operators*, volume 168. American Mathematical Soc., 2015.

[3] Oskari Ajanki, László Erdős, and Torben Krüger. *Quadratic vector equations on complex upper half-plane*, volume 261. American Mathematical Society, 2019.

[4] Oskari Ajanki, Torben Krüger, and László Erdős. Singularities of solutions to quadratic vector equations on the complex upper half-plane. *Communications on Pure and Applied Mathematics*, 70(9):1672–1705, 2017.

[5] Johannes Alt, Raphael Ducatez, and Antti Knowles. Delocalization transition for critical erdős–rényi graphs. *Communications in Mathematical Physics*, 388(1):507–579, 2021.

[6] Johannes Alt, Raphaël Ducatez, and Antti Knowles. Extremal eigenvalues of critical erdős–rényi graphs. *The Annals of Probability*, 49(3):1347–1401, 2021.

[7] Johannes Alt, Raphael Ducatez, and Antti Knowles. Poisson statistics and localization at the spectral edge of sparse erd\h {o} s–r\'enyi graphs. *arXiv preprint arXiv:2106.12519*, 2021.

[8] Philip W Anderson. Absence of diffusion in certain random lattices. *Physical review*, 109(5):1492, 1958.

[9] Antonio Auffinger, Gérard Ben Arous, and Sandrine Péché. Poisson convergence for the largest eigenvalues of heavy tailed random matrices. In *Annales de l'IHP Probabilités et statistiques*, volume 45, pages 589–610, 2009.

[10] Roland Bauerschmidt, Jiaoyang Huang, Antti Knowles, and Horng-Tzer Yau. Edge rigidity and universality of random regular graphs of intermediate degree. *Geometric and Functional Analysis*, 30(3):693–769, 2020.

[11] Florent Benaych-Georges, Charles Bordenave, and Antti Knowles. Largest eigenvalues of sparse inhomogeneous erdős–rényi graphs. *The Annals of Probability*, 47(3):1653–1676, 2019.

[12] Florent Benaych-Georges, Charles Bordenave, and Antti Knowles. Spectral radii of sparse random matrices. 2020.

[13] Florent Benaych-Georges, Alice Guionnet, and Mylène Maida. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. 2011.

[14] Florent Benaych-Georges and Antti Knowles. Lectures on the local semicircle law for wigner matrices. *arXiv preprint arXiv:1601.04055*, 2016.

[15] Lucas Benigni. Eigenvectors distribution and quantum unique ergodicity for deformed wigner matrices. 2020.

[16] Béla Bollobás and Béla Bollobás. *Random graphs*. Springer, 1998.

[17] Gavin Brown. P. billingsley, probability and measure (wiley, 1979), pp. 532,£ 28 · 95. *Proceedings of the Edinburgh Mathematical Society*, 26(3):398–399, 1983.

[18] Stephen G Brush. History of the lenz-ising model. *Reviews of modern physics*, 39(4):883, 1967.

[19] Włodzimierz Bryc, Amir Dembo, and Tiefeng Jiang. Spectral measure of large random hankel, markov and toeplitz matrices. 2006.

[20] Andrew Campbell, Kyle Luh, Sean O'Rourke, Santiago Arenas-Velilla, and Victor Pérez-Abreu. Extreme eigenvalues of laplacian random matrices with gaussian entries. *arXiv preprint arXiv:2211.17175*, 2022.

[21] Edsger Wybe Dijkstra. A note on two problems in connexion with graphs:(numerische mathematik, 1 (1959), p 269-271). 1959.

[22] Paul Erd6s and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[23] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. Local semicircle law and complete delocalization for wigner random matrices. *Communications in Mathematical Physics*, 287(2):641–655, 2009.

[24] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. Semicircle law on short scales and delocalization of eigenvectors for wigner random matrices. 2009.

[25] László Erdős and Horng-Tzer Yau. *A dynamical approach to random matrix theory*, volume 28. American Mathematical Soc., 2017.

[26] Ronald A Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915.

[27] Jürg Fröhlich and Thomas Spencer. Absence of diffusion in the anderson tight binding model for large disorder or low energy. *Communications in Mathematical Physics*, 88(2):151–184, 1983.

[28] Yukun He, Antti Knowles, and Matteo Marcozzi. Local law and complete eigenvector delocalization for supercritical erdős–rényi graphs. *The Annals of Probability*, 47(5):3278–3302, 2019.

[29] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[30] Jiaoyang Huang and Benjamin Landon. Spectral statistics of sparse erdős–rényi graph laplacians. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, pages 120–154. Institut Henri Poincaré, 2020.

[31] Svante Janson, Andrzej Rucinski, and Tomasz Luczak. *Random graphs*. John Wiley & Sons, 2011.

[32] Tiefeng Jiang. Empirical distributions of laplacian matrices of large dilute random graphs. *Random Matrices: Theory and Applications*, 1(03):1250004, 2012.

[33] Tiefeng Jiang. Low eigenvalues of laplacian matrices of large random graphs. *Probability Theory and Related Fields*, 153:671–690, 2012.

[34] Mark Kac. Can one hear the shape of a drum? *The american mathematical monthly*, 73(4P2):1–23, 1966.

[35] Tosio Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.

[36] Michael Krivelevich and Benny Sudakov. The phase transition in random graphs: A simple proof. *Random Structures & Algorithms*, 43(2):131–138, 2013.

[37] Ji Oon Lee and Kevin Schnelli. Local deformed semicircle law and complete delocalization for wigner matrices with random potential. *Journal of Mathematical Physics*, 54(10):103504, 2013.

[38] Ji Oon Lee and Kevin Schnelli. Extremal eigenvalues and eigenvectors of deformed wigner matrices. *Probability Theory and Related Fields*, 164(1-2):165–241, 2016.

[39] Ji Oon Lee, Kevin Schnelli, Ben Stetler, and Horng-Tzer Yau. Bulk universality for deformed wigner matrices. 2016.

[40] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bring order to the web. Technical report, technical report, Stanford University, 1998.

[41] Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 2(11):559–572, 1901.

[42] Michael Reed. *Methods of modern mathematical physics: Functional analysis*. Elsevier, 2012.

[43] Renato Renner. Quantenmechanik i. *https://edu.itp.phys.ethz.ch/hs11/qm1/*, 2011.

[44] Jun John Sakurai and Eugene D Commins. Modern quantum mechanics, revised edition, 1995.

[45] Sasha Sodin. The spectral edge of some random band matrices. *Annals of mathematics*, pages 2223–2251, 2010.

[46] Terence Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.

[47] Marco Tarzia. Fully localized and partially delocalized states in the tails of erdös-rényi graphs in the critical regime. *Physical Review B*, 105(17):174201, 2022.

[48] Konstantin Tikhomirov and Pierre Youssef. Outliers in spectrum of sparse wigner matrices. *Random Structures & Algorithms*, 58(3):517–605, 2021.

[49] Per von Soosten and Simone Warzel. Non-ergodic delocalization in the rosenzweig–porter model. *Letters in Mathematical Physics*, 109:905–922, 2019.

[50] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327, 1958.

[51] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.

[52] John Wishart. Proofs of the distribution law of the second order moment statistics. *Biometrika*, 35(1/2):55–57, 1948.