Semilocalization for inhomogeneous random graphs

Thomas Buc–d'Alché Antti Knowles

January 8, 2026

Abstract

1. Introduction

Conventions and notations. Every quantity that is not explicitly *constant* depends on N. We omit this dependence in our notation. We use C, c to denote generic positive constants, which may change from step to step. We write $X = O(Y)$ to mean $|X| \leqslant CY$. We write $\mathbb{N} = \{0, 1, 2, \dots\}$. We set $[n] := \{1, \dots, n\}$ for any $n \in \mathbb{N}^*$. We write $\#X$ for the cardinality of a finite set X.

1.1. Model and assumptions. Let G be a graph on the vertex set $[N]$, whereby we identify G with its set of edges. It is characterized by its adjacency matrix $A = (A_{xy})_{x,y \in [N]}$, where $A_{xy} := \mathbb{1}_{\{x,y\} \in G}$. We only consider simple graphs, so that $A_{xx} = 0$ for all $x \in [N]$. We write $x \sim y$ whenever $A_{xy} = 1$. We endow G with the usual graph distance: the distance between x and y is the number of edges in the shortest path in the graph joining x and y. For $x \in [N]$ and $r \in \mathbb{N}$, we denote by $B_r(x)$ the ball of radius r around x, i.e. the set of vertices at distance at most r from x, as well as by $S_r(x)$ the sphere of radius r around x, i.e. the set of vertices at distance r from x. We denote by $D_x = \sum_{y \in [N]} A_{xy}$ the degree of x.

In this paper we consider random graphs G whose law is given by the Generalized Random Graph (GRG) model (see e.g. [Hof16, Chapter 6]): the family $(A_{xy} : x < y)$ is independent with

$$\mathbb{P}(A_{xy} = 1) = p_{xy} := \frac{w_x w_y}{\sum_z w_z + w_x w_y}\,, \tag{1.1}$$

where $(w_x)_{x \in [N]} \in (0, \infty)^N$ is a family of strictly positive weights.

For $k \in \mathbb{N}^*$, we define the empirical moment

$$m_k := \frac{1}{N} \sum_{x=1}^{N} w_x^k\,. \tag{1.2}$$

We fix two constant exponents $0 < \varepsilon < 1/2$ and $0 < \delta < 1/3$ and make the following two assumptions on the weights (w_x).

Assumption 1.1. For all $x \in [N]$,

$$w_x \leqslant N^{1/2-\varepsilon}\,.$$

Assumption 1.2. The first and second empirical moments satisfy

$$m_1 \geqslant N^{-\varepsilon}\,, \qquad \frac{m_2}{m_1} = O((\log N)^{\delta})\,.$$

Remark 1.3. Another commonly used model of an inhomogeneous random graph is the Chung-Lu model [CL02], where, instead of (1.1), we set

$$\mathbb{P}(A_{xy} = 1) = \tilde{p}_{xy} := \frac{w_x w_y}{\sum_z w_z} \wedge 1 \,.$$

It follows immediately from Assumptions 1.1 and 1.2 that $p_{xy} = \tilde{p}_{xy}(1 + O(N^{-\varepsilon}))$, so that, in the regime that we consider, the GRG and Chung-Lu models are asymptotically equivalent. In particular, all of our results easily carry over to the Chung-Lu model.

Remark 1.4. A natural way to construct the weights (w_x) is to choose a fixed probability measure μ on $(0, \infty)$ and to choose the weights as either (i) independent random variables with law μ or (ii) the $(N+1)$-quantiles of μ. We give two standard examples of μ, to which our main result is applicable.

Example 1.5 (Exponential distribution)**.** Let μ be the exponential distribution with parameter $\alpha > 0$. Then Assumptions 1.1 and 1.2 hold (in case (i) of Remark 1.4, they hold with high probability). In case (ii) of Remark 1.4, the weights are

$$w_x = \frac{1}{\alpha} \log \frac{N+1}{x} \,,$$

for $x \in [N]$.

Example 1.6 (Power law)**.** Suppose that μ is a power law with exponent $\alpha > 2$. That is,

$$\mu([t, \infty)) = L(t)\, t^{-\alpha} \,, \tag{1.3}$$

where L is a slowly varying function, i.e. there exists $u_0 > 0$ such that

$$\lim_{t \to \infty} \frac{L(tu)}{L(u)} = 1 \quad \text{for all } u \geqslant u_0 \,.$$

It is easy to check that Assumptions 1.1 and 1.2 hold (in case (i) of Remark 1.4, they hold with high probability). For instance, for $L(t) = \alpha c^{\alpha} \mathbb{1}_{t \geqslant c}$ for some $c > 0$, the quantile weights from Remark 1.4 (ii) are given by

$$w_x = x \left(\frac{N+1}{x} \right)^{1/\alpha} .$$

If the weights (w_x) have a power law distribution with exponent α, then with high probability so do the degrees (D_x) of the random graph; see e.g. [Hof16, Theorem 6.12]. Hence, the choice (1.3) yields random graphs with power-law degree distribution (often also called scale-free graphs).

We conclude this section with the following quantitative notion of high probability that we use throughout the paper.

Definition 1.7. Let $\nu > 0$ be a constant. An event $E \equiv E_N$ holds with *ν-high probability* if there exists a constant $C > 0$ such that, for all N,

$$\mathbb{P}(E) \geqslant 1 - CN^{-\nu}.$$

1.2. Semilocalization. We now state a semilocalization result on the eigenvectors of A. It pertains to vertices whose associated energies lie in a interval of width $\eta > 0$ centered around an energy $\lambda \in \mathbb{R}$. We define the set of *resonant vertices* by

$$\mathcal{W}_{\lambda,\eta} := \Big\{x \in [N] : \Big|\sqrt{D_x} - \lambda\Big| \leqslant \eta\Big\}.$$

Theorem 1.8 (Semilocalization)**.** *Suppose that Assumptions 1.1 and 1.2 hold. For any $\nu > 0$ there exists $C_\nu > 0$ such that the following holds with ν-high probability. For any normalized eigenvector $\mathbf{q}$ of A with associated eigenvalue λ and for any $\eta \leqslant |\lambda|/2$ we have*

$$\sum_{x \in \mathcal{W}_{\lambda,\eta}} \langle \mathbf{q}, \mathbf{u}_{\mathrm{sgn}(\lambda)}(x)\rangle^2 \geqslant 1 - \frac{C_\nu}{\eta^2}\frac{\log N}{\log\log N}\,,$$

where $\mathbf{u}_\pm(x)$ is an explicit vector supported in $B_2(x)$ defined in Proposition 3.3 below.

The next result, Proposition 1.9, shows that in the cases of interest the set of resonant vertices is of negligible size compared to the size of the graph. More precise results leading to a localization phenomenon when looking at the greatest eigenvalues are discussed in Section 5.

Proposition 1.9. *For any $\nu > 0$ there exists $C_\nu > 0$ such that with ν-high probability*

$$\#\mathcal{W}_{\lambda,\eta} \leqslant 2\mathbb{E}[\#\mathcal{W}_{\lambda,\eta+1}] \vee \frac{2\nu\log N}{\log\log N}$$

whenever

$$C_\nu\sqrt{\frac{\log N}{\log\log N}} < \eta < \frac{\lambda}{2}.$$

The expectation on the right-hand side can be estimated in general by a second moment method as

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta+1}] \leqslant \frac{m_2}{(\lambda-\eta)^4}N\,.$$

Furthermore, when the weights are taken to be the quantile (see case (ii) of Remark 1.4) of the exponential or the heavy tail distribution as in Examples 1.5 and 1.6, we get the sharper bounds

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta+1}] = O\bigg(\frac{N}{(\alpha+1)^{(\lambda-\eta)^2}} + (\log N)^{2\delta}\bigg)\,,$$

and for all $\iota > 0$ if $\lim_{t\to\infty} L(t) = \infty$ and $\iota = 0$ if L is bounded,

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta+1}] = O\bigg(\frac{N\eta}{\lambda^{2\alpha+3-\iota}} + (\log N)^{2\delta}\bigg)\,,$$

respectively.

Remark 1.10. If the weights are distributed as the quantiles of an exponential distribution of parameter $\alpha > 0$ (Example 1.5), then when $\alpha < 1/(2\nu)$, Lemma 1.12 implies that the greatest degree in the graph is of order $\log N$, and Theorem 1.8 is not vacuous.

If the weights are distributed as the quantiles of a heavy-tailed distribution with parameter $\alpha > 2$ (Example 1.6), the greatest degree in the graph is of order $N^{1/\alpha}$, and Theorem 1.8 is not vacuous as well.

Remark 1.11. In the large N limit, most of the random variables we consider have Poisson tails. The $\log N$ scale thus appears naturally when deriving results that hold with ν-high probability. In particular, our argument relies on the fact that the degrees $(D_x)_{x\in[N]}$ have Poisson tails whose parameters are the w_x's. This is a consequence of Bennett's inequality.

Furthermore, if X is a Poisson variable with parameter $\lambda > 0$, we have the sharp bound

$$\mathbb{P}\left(X \geqslant \lambda\nu \frac{\log N}{\log\log N}\right) \leqslant N^{-\nu(1-o(1))} \text{ as } N \to \infty.$$

This bound is essentially the reason of the $\frac{\log N}{\log\log N}$ factor in our Theorem 1.8. To get results below the scale $\frac{\log N}{\log\log N}$, we would need to understand finer properties of the structure of the graph, and depart from the approximation that the degrees are independent Poisson variables.

Note that if the weights are distributed as the quantiles of the exponential distribution – see Example 1.5 – both the biggest weight and biggest degree in the graph are of order $\log N$. Theorem 1.8 with $\frac{\log N}{\log\log N} \ll \eta^2 \ll \log N$ gives a non-trivial result in this case. If we had used the more naïve bound

$$\mathbb{P}(X \geqslant \lambda\nu\log N) \leqslant N^{-\nu},$$

the random graph with exponential weights would not have been covered by Theorem 1.8.

1.3. The degrees and the weights. We conclude the introduction with a few basic facts and tools on the relationship between weights and degrees, which we shall use throughout the proofs. In the sequel, we will mainly consider vertices of high degree. In particular, vertices with weights greater than $c\log N$, for a fixed $c > 0$, are easier to describe.

We denote by

$$d_x := \mathbb{E}[D_x]$$

the expectation of the degree of x. Using

$$p_{xy} = \frac{w_x w_y}{m_1 N + w_x w_y} \leqslant \frac{w_x w_y}{m_1 N}\,. \tag{1.4}$$

we find

$$d_x = \sum_{y\in[N]\setminus\{x\}} \frac{w_x w_y}{\sum_z w_z + w_x w_y} \leqslant w_x\,.$$

An asymptotically matching lower bound for d_x follows from Remark 1.3, which yields

$$d_x = \sum_{y\in[N]\setminus\{x\}} \frac{w_x w_y}{m_1 N}(1+O(N^{-\varepsilon})) = w_x\left(1 - \frac{w_x}{m_1 N}\right)(1+O(N^{-\varepsilon})) = w_x(1+O(N^{-\varepsilon}))\,. \tag{1.5}$$

As soon as the weight of a vertex x is at least of order $\log N$, the degree D_x is as well at least of order $\log N$, with ν-high probability.

Lemma 1.12. *Let $\nu > 0$ and $x \in [N]$. With ν-high probability,*

$$d_x - \sqrt{2\nu d_x \log N} \leqslant D_x \leqslant d_x + 2\sqrt{\nu\log N\left(d_x \vee \frac{4\nu}{9}\log N\right)}\,.$$

Proof. This is an application of Bennett's inequality (see [BLM13, Theorem 2.9]). We have for $M > 0$,

$$\mathbb{P}(D_x \geqslant M + d_x) = \mathbb{P}(D_x - d_x \geqslant M) \leqslant \exp(-\frac{M^2}{2(w_x + M/3)}).$$

If we choose $M = 2\sqrt{\nu \log N (w_x \vee (2/3)\nu \log N)}$, we get the upper bound.

For the lower bound, we use the slightly sharper bound (see [Hof16, Theorem 2.21]),

$$\mathbb{P}(D_x \leqslant d_x - \sqrt{2\nu d_x \log N}) \leqslant \exp(-\frac{(\sqrt{2\nu d_x \log N})^2}{2d_x}) \leqslant N^{-\nu}. \qquad \square$$

Remark 1.13 (Lower bound for D_x)**.** Lemma 1.12 implies that as soon as $D_x \geqslant 1$ and $d_x \geqslant 4\nu \log N$, we have that

$$d_x \leqslant \frac{\sqrt{2}}{\sqrt{2}-1} D_x \quad \text{with } \nu\text{-high probability.}$$

Otherwise, if $D_x \geqslant C \geqslant 1$ and $d_x < 4\nu \log N$, we only have the crude bound

$$d_x \leqslant \frac{4\nu}{C} \log N D_x.$$

Hence, we have in any case

$$d_x \leqslant \left(\frac{\sqrt{2}}{\sqrt{2}-1} \vee \frac{4\nu}{C}\log N\right) D_x.$$

Remark 1.14 (Upper bound for D_x)**.** Lemma 1.12 implies that if $d_x > \frac{4\nu}{9}\log N$, we have

$$D_x \leqslant d_x + 2\sqrt{\nu \log N d_x} \leqslant d_x + 2\sqrt{\frac{9}{4}d_x^2} = 4d_x$$

with ν-high probability. However, if $d_x \leqslant \frac{4\nu}{9}\log N$, we have

$$D_x \leqslant d_x + \frac{4\nu}{3}\log N \leqslant \frac{16\nu}{9}\log N$$

with ν-high probability. Hence, we have in all cases

$$D_x \leqslant 4d_x \vee \frac{16\nu}{9}\log N \quad \text{with } \nu\text{-high probability}.$$

In the sequel, it will be convenient to order the vertices in terms of their degree.

Definition 1.15. We define the strict order relation $\prec$ on the set of vertices $[N]$ as follows. For any two vertices $x, y \in [N]$,

$$x \prec y \text{ if and only if } \Big((D_x < D_y) \text{ or } (D_x = D_y \text{ and } x > y)\Big).$$

We define the (random) permutation $\pi \in \mathfrak{S}_N$ to be the unique permutation such that

$$\pi(N) \prec \pi(N-1) \prec \cdots \prec \pi(2) \prec \pi(1).$$

Note that in particular $D_{\pi(N)} \leqslant D_{\pi(N-1)} \leqslant \cdots \leqslant D_{\pi(2)} \leqslant D_{\pi(1)}$. This order allows us to define two notions of neighborhood and degree:

$$\begin{aligned} S_1^+(x) &= \{y \in [N] : x \sim y, x \prec y\} \quad \text{and} \quad D_x^+ = \# S_1^+(x)\,, \\ S_1^-(x) &= \{y \in [N] : x \sim y, x \succ y\} \quad \text{and} \quad D_x^- = \# S_1^-(x)\,. \end{aligned} \tag{1.6}$$

Thus, $S_1^+(x)$ and $S_1^-(x)$ partition $S_1(x)$ with $D_x = D_x^+ + D_x^-$.

Notation. Finally, we introduce some standard notations that are used throughout the proofs. We use the following notations for vectors. Vectors in $\mathbb{R}^N$ are denoted by boldface lowercase Latin letters; we use the notation $\mathbf{v} = (v_x)_{x \in [N]} \in \mathbb{R}^N$ for the entries of a vector. We denote by $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{x \in [N]} v_x w_x$ the Euclidean scalar product on $\mathbb{R}^N$ and by $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ the induced Euclidean norm. We denote by $\|A\|$ the induced operator norm on $N \times N$ matrices A. For any $x \in [N]$, we define the standard basis vector $\mathbf{1}_x := (\delta_{xy})_{y \in [N]} \in \mathbb{R}^N$. To any subset $S \subset [N]$ we assign the vector $\mathbf{1}_S \in \mathbb{R}^N$ given by $\mathbf{1}_S := \sum_{x \in S} \mathbf{1}_x$. For $S \subset [N]$ we denote by $G|_S$ the subgraph of G induced by S.

2. Pruning the graph

Similarly as in [ADK21b, ADK21a], it is more convenient to work on a pruned version of the graph. The GRG model is inhomogeneous, compared to the Erdős-Rényi model: since the laws of the degrees in the graph are governed by the weights (w_x), there are greater differences of degrees in the graph. Because of this greater heterogeneity, the pruning procedure has to be more subtle. As in the Erdős-Rényi case, we first prune the graph to remove cycles in small balls. Then, instead of removing all edges connecting two vertices of high degree, we remove edges appearing in a very specific pattern. This procedure is key to simplifying the computations in Section 3.

Throughout the following we fix a constant $r \geqslant 6$. For convenience, we will sometimes omit it from the notation.

Remark 2.1. The pruning of [ADK21a] (besides removing cycles) amounts to removing all the edges between vertices of high degree. Because of the inhomogeneity in the GRG case, this would mean cutting a number of vertices proportional to w_x around a vertex x. This would prevent us from obtaining good bounds on the error we make when replacing the adjacency matrix of the original graph by the one of the pruned graph.

The new pruning presented below is asymmetric, and based on the order $\prec$ introduced in Definition 1.15. We orient each edge $\{x, y\}$ in G from x to y if $x \prec y$. We then remove special paths we call *down-up paths*, paths of length 2 between vertices x and z, going through a vertex y so that $y \prec x \prec z$.

In the Erdős-Rényi model, one can prune the graph so that in a ball of small radius around any vertex, the graph is a tree and contains at most one vertex of high degree. In the GRG model, by removing down-up paths, we get a pruned graph which is globally a forest. Furthermore, each connected component is a tree that is naturally rooted at the vertex of greatest degree in the connected component.

2.1. The pruning procedure. We now explain precisely the pruning procedure. It produces a pruned graph G^{p}. To give the construction, we introduce notation pertaining to paths.

Definition 2.2. A *path* γ in a graph G is a sequence of vertices $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_l)$, with $\{\gamma_{i-1}, \gamma_i\} \in G$ for $i \in [n]$. The length of the path is $l(\gamma) = l$. A path is said to be *simple* if $\gamma_i \neq \gamma_j$ for $i \neq j, \{i, j\} \neq \{0, l\}$.

The set of paths γ in G satisfying $\gamma_0 = x$ and $\gamma_{l(\gamma)} = y$ is denoted by $\mathcal{P}_{xy}(G)$, and its subset of simple paths is denoted by $\mathcal{P}^*_{xy}(G)$.

We mentioned that we will consider a particular set of paths, the down-up paths, which we define below.

Definition 2.3. A *down-up path* between two distinct vertices $x \in [N]$ and $z \in [N]$ is a path (x, y, z) with $y \prec x \prec z$.

Consider a vertex $x \in [N]$. We define the two sets

$$\begin{aligned} S_1^{\mathrm{cyc}}(x) &= \{y \in S_1(x) : \exists \gamma \in \mathcal{P}^*_{xx}(G), l(\gamma) \leqslant 2r+1, \gamma_1 = y\}, \\ S_1^{\mathrm{du}}(x) &= \{y \in S_1(x) \setminus S_1^{\mathrm{cyc}}(x) : \exists z \in S_1(y) \setminus S_1^{\mathrm{cyc}}(y), y \prec x \prec z\}. \end{aligned} \tag{2.1}$$

i.e. $S_1^{\mathrm{cyc}}(x)$ is the set of vertices connected to x that are part of a cycle which is a simple loop, and $S_1^{\mathrm{du}}(x)$ is the set of vertices part of an up-down path starting at x.

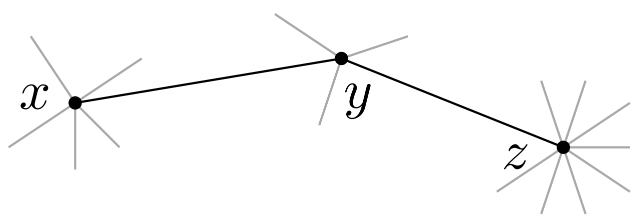

Figure 2.1. A down-up path.

To construct the pruned graph G^{p}, we proceed in two steps. Firstly, we do a first pruning procedure to remove some cycles in balls of radius r around all vertices. Then, we remove the down-up path, see Definition 2.3.

Definition 2.4 (Pruning procedure)**.** The graph G^{p} is defined by the following procedure.

1. for each $x \in [N]$, and each $y \in S_1^{\mathrm{cyc}}(x)$, we remove from G the edge $\{x, y\}$, and then
2. for each $x \in [N]$, and each $y \in S_1^{\mathrm{du}}(x)$, we remove from G the edge $\{x, y\}$.

We denote the graph obtained after the first step by G^{nc}.

Note that the definition of G^{nc} and G^{p} do not depend on the order in which the edges are removed. We indicate by the superscript p (respectively nc) that the adjacency matrix, degrees, spheres, balls, ... correspond to the pruned graph G^{p} (respectively G^{nc}). For instance, D_x^{p} is the degree of the vertex x in G^{p}. The event that $\{x, y\} \in G^{\mathrm{nc}}$ is denoted by $\{x \overset{\mathrm{nc}}{\sim} y\}$, and that $\{x, y\} \in G^{\mathrm{p}}$ is denoted by $\{x \overset{\mathrm{p}}{\sim} y\}$.

Proposition 2.5. *Let $\nu > 0$. We define the threshold*

$$\xi = \xi_\nu = \frac{4(\nu+1)(2-3\delta)}{(1-\delta)(1-2\delta)} \frac{\log N}{\log \log N} .$$

The graph G^{p} satisfies the following properties.

1. *With ν-high probability, for all $x \in [N]$, $D_x - D_x^{\mathrm{p}} \leqslant \xi/2$.*
2. *There are no down-up paths in the graph G^{p}.*
3. *The graph G^{p} is a forest.*

The proof of Proposition 2.5 is delayed to the end of Section 2.4, and relies on the next two sections: Section 2.2 explains how many edges are removed when constructing G^{nc} and Section 2.4 explains how many edges are removed when constructing G^{p}.

Remark 2.6. Note that we prune the graph around every vertex in the graph, and not only those of high degree as in [ADK21b, ADK21a]. This ensures that the graph G^{p} is globally, and not just locally, a forest. This proves useful in Section 4.2 when bounding the operator norm of the adjacency matrix of the graph G^{p} restricted to the vertices of low degrees: we can immediately say that this operator norm is of the order of the square root of the maximal degree in this restricted graph.

2.2. Removing the cycles. In this section, we estimate the number of edges to prune around each vertex to remove all cycles in small balls centered around each vertex. This will prove the part of Proposition 2.5 concerning step 1 of the pruning procedure. Recall that we fixed $r \geqslant 6$.

In Section 2.1, we explained that we prune the edges in the set $\{\{x,y\} : y \in S_1^{\mathrm{cyc}}(x)\}$ so as to remove all cycles in all small balls. In this section, we give an upper bound for the cardinality of the set $S_1^{\mathrm{cyc}}(x)$ defined in (2.1).

Proposition 2.7. *Fix $x \in [N]$ and $\nu > 0$. There exists a constant $C_\nu \geqslant 0$ such that with ν-high probability,*

$$\# S_1^{\mathrm{cyc}}(x) \leqslant C_\nu .$$

To prove Proposition 2.7, we will use the notion of edge-disjoint paths.

Definition 2.8. Two paths γ and γ' are *edge-intersecting* if there exists i and j such that $\gamma_i = \gamma'_j$ and either $\gamma_{i-1} = \gamma'_{j-1}$ or $\gamma_{i-1} = \gamma'_{j+1}$. Two paths γ and γ' are *edge-disjoint* if they are not edge-intersecting.

We shall consider the following event concerning edge-disjoint paths

$$\mathrm{ED}_{k,r}(x) = \left\{ \exists \gamma^{(1)}, \dots, \gamma^{(k)} \in \mathcal{P}_{xx}^*(G|_{B_r(x)}) : \begin{array}{c} \forall i \neq j, \gamma^{(i)} \text{ and } \gamma^{(j)} \text{ are edge-disjoint,} \\ \sum_{i=1}^k l(\gamma^{(i)}) \leqslant 3k^2(2r+1) \end{array} \right\} .$$

It is depicted in Figure 2.2.

We argue that $\# S_1^{\mathrm{cyc}}(x)$ can be bounded by the number of non edge-intersecting paths in the graph $G|_{B_r(x)\setminus\{x\}}$ between pairs of points of $S_1(x)$.

Lemma 2.9. *Let $k \geqslant 1$ and $x \in [N]$. We have the inclusion of events*

$$\{\# S_1^{\mathrm{cyc}}(x) \geqslant 3k\} \subset \mathrm{ED}_{k,r}(x) .$$

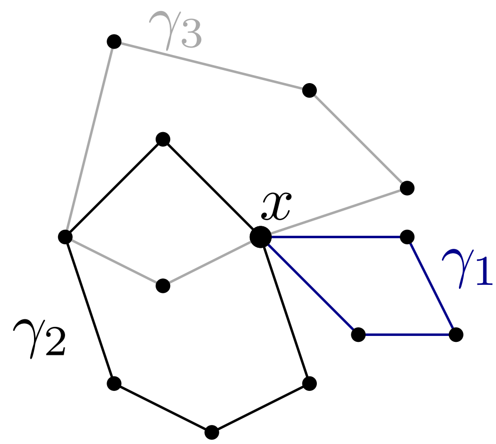

Figure 2.2. A representation of the event $\mathrm{ED}_{k,r}$. The simple paths can share a vertex, but cannot be edge-intersecting.

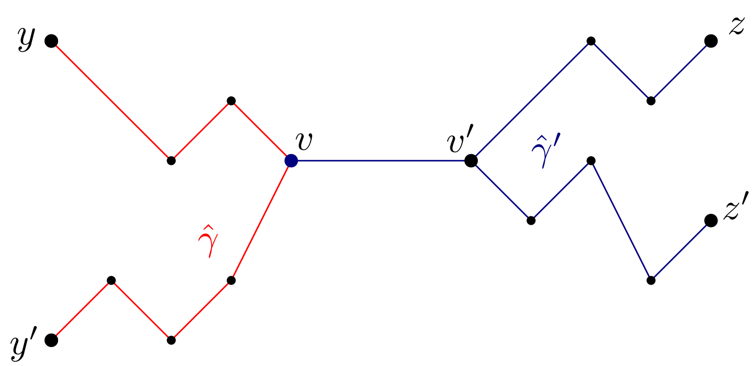

Figure 2.3. Construction of the paths $\hat\gamma$ and $\hat\gamma'$.
The paths γ and γ' are edge-intersecting, respectively from y to z and y' to z'.

Proof. Let $y_1, \dots, y_{3k} \in S_1^{\mathrm{cyc}}(x)$ be distinct vertices. By definition, there exists $z_1, \dots, z_{3k} \in S_1^{\mathrm{cyc}}(x)$, with $z_i \neq y_i$, and simple paths $\gamma^{(i)}, i = 1, \dots, 3k$, each of them connecting y_i to z_i (without going through x). We now argue that we can find a family of vertices $\{\tilde y_i, \tilde z_i\}_{1 \leqslant i \leqslant k}$ and a family of paths $(\tilde\gamma_1, \dots, \tilde\gamma_k)$ such that $\tilde\gamma^{(i)}$ is a path from $\tilde y_i \in S_1(x)$ to $\tilde z_i \in S_1(x)$ and the vertices $\{\tilde y_i, \tilde z_i\}$ are all distinct. To see this, consider the graph $\tilde G$ whose vertex set is $\tilde V = \{y_i, z_i; 1 \leqslant i \leqslant 3k\}$ and the edge set is made of the pairs $\{v_1, v_2\}$ such that $v_1 \in \tilde V$ and $v_2 \in \tilde V$ are the endpoints of one of the path in the family (γ_i). In each connected component of $\tilde G$ made of the vertices $v_1, \dots, v_m$ we choose a perfect matching of the vertices $\{v_1, \dots, v_m\}$ if m is even, and a perfect matching of the vertices $\{v_1, \dots, v_{m-1}\}$ if m is odd. Note that $m \geqslant 2$ by definition of the γ_i's. For each couple $\{v, v'\}$ in a matching we can choose a simple path in G going from v to v'. Each such path has length at most $3k(2r+1)$. Hence the total number of pairs in the matching is

$$\frac{1}{2} \sum_{\tilde G_c \text{ connected component of } \tilde G} \#\tilde V_c - \mathbb{1}_{\{\#\tilde V_c \text{ odd }\}} = \frac{3k}{2} - \frac{1}{2} \#\Big\{ \tilde G_c \text{ connected component with } \tilde V_c \text{ odd } \Big\},$$

where $\tilde V_c$ is the vertex set of the connected component $\tilde G_c$ of $\tilde G$. Since each connected component has at least two vertices, we have

$$\#\Big\{\tilde{G}_c \text{ connected component with } \tilde{V}_c \text{ odd }\Big\} \leqslant k.$$

This means that doing so, we find at least k pairs: $2k$ distinct vertices $\{\tilde{y}_i, \tilde{z}_i\}$ and k simple paths $(\tilde{\gamma}_i)$, with $\tilde{\gamma}_i$ going from $\tilde{y}_i$ to $\tilde{z}_i$.

We now argue that from $(\tilde{\gamma}^{(i)})_{1\leqslant i\leqslant k}$ we can produce a family $(\hat{\gamma}^{(i)})_{1\leqslant i\leqslant k}$ of edge-disjoint paths. To see this, we explain how to produce from two paths $\tilde{\gamma}$ and $\tilde{\gamma}'$ that do not share an endpoint a pair of paths $\hat{\gamma}$ and $\hat{\gamma}'$ that are edge-disjoint and do not share an endpoint. Assume that $\{v, v'\}$ is an edge shared by $\tilde{\gamma}$ and $\tilde{\gamma}'$. Without loss of generality, we can assume that $\tilde{\gamma}$ and $\tilde{\gamma}'$ both first encounter v and then v'. Otherwise we replace $\tilde{\gamma}'$ by

$$(\tilde{\gamma}'_{l(\tilde{\gamma}')}, \tilde{\gamma}'_{l(\tilde{\gamma}')-1}, \dots, \tilde{\gamma}'_1, \tilde{\gamma}'_0).$$

In that case, we can write

$$\begin{cases} \tilde{\gamma} & = (\tilde{\gamma}_0 = y, \tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_i = v, \tilde{\gamma}_{i+1} = v', \tilde{\gamma}_{i+2}, \dots, \tilde{\gamma}_{l-1}, z) \\ \tilde{\gamma}' & = (\tilde{\gamma}'_0 = y', \tilde{\gamma}'_1, \tilde{\gamma}'_2, \dots, \tilde{\gamma}'_j = v, \tilde{\gamma}'_{j+1} = v', \tilde{\gamma}'_{j+2}, \dots, \tilde{\gamma}'_{l'-1}, z'). \end{cases}$$

We then set

$$\begin{cases} \hat{\gamma} & = (y, \tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_i = v = \tilde{\gamma}'_j, \tilde{\gamma}'_{j-1}, \dots, \tilde{\gamma}'_1, y') \\ \hat{\gamma}' & = (z, \tilde{\gamma}_{l-1}, \tilde{\gamma}_{l-2}, \dots, \tilde{\gamma}_{i+1} = v' = \tilde{\gamma}'_{j+1}, \tilde{\gamma}'_{j+2}, \dots, \tilde{\gamma}'_{l'-1}, z'). \end{cases}$$

These new paths are depicted in Fig. 2.3. The new paths $\hat{\gamma}$ and $\hat{\gamma}'$ do not share the edge $\{v, v'\}$, have disjoint endpoints, and all the edges in $\hat{\gamma}$ and $\hat{\gamma}'$ appear in $\tilde{\gamma}$ or $\tilde{\gamma}'$ the same number of times.

To construct a family of non edge-intersecting paths from the paths $(\tilde{\gamma}^{(i)})_{1\leqslant i\leqslant k}$, we consider first the path $\tilde{\gamma}^{(1)}$. We consider the first edge e in $\gamma^{(1)}$ that intersect another path $\gamma^{(i)}$, $i \neq 1$. We apply the procedure described above, and obtain a new family of paths of size k:

$$(\hat{\gamma}^{(1)}, \tilde{\gamma}^{(2)}, \dots, \tilde{\gamma}^{(i-1)}, \hat{\gamma}^{(i)}, \tilde{\gamma}^{(i+1)}, \dots, \tilde{\gamma}^{(k)}),$$

such that the endpoints of all the paths appearing in the family are distinct, and there is one less edge appearing in two paths than in $(\tilde{\gamma}^{(i)})_{1\leqslant i\leqslant k}$.

We keep applying this procedure on the first path of the family until it is no longer edge-intersecting with any other path of the family. We then consider the second path, and proceed as previously until all the paths are edge-disjoint. Notice that this procedure terminates as there is a finite number of edges that are part of two or more paths. We end up with a family of k edge-disjoint paths. Note that the paths thus created are not necessarily simple but contain a simple path between their endpoints (as their endpoints are distinct). Replacing each $\hat{\gamma}^{(i)}$ by the simple path it contains yield the result.

Notice that we have not added any edges in the procedure, thus the total length of the non edge-intersecting paths thus created is less than $3k^2(2r-1)$. This shows that the required inclusion holds. $\square$

Lemma 2.9 then implies Proposition 2.7.

Proof of Proposition 2.7. Let $k \geqslant 1$. Lemma 2.9 implies that

$$\mathbb{P}(\#S_1^{\text{cyc}}(x) \geqslant 3k) \leqslant \mathbb{P}(\mathrm{ED}_{k,r}(x))\,.$$

The union bound then implies

$$\mathbb{P}(\mathrm{ED}_{k,r}(x)) \leqslant \sum_{\substack{l_1,\dots,l_k \geqslant 1 \\ \sum_i l_i \leqslant 3k^2(2r+1)}} \prod_{i=1}^{k} \Bigg(\sum_{y_1,\dots,y_{l_i-1}} p_{xy_1}p_{y_1y_2}\cdots p_{y_{l_i-2}y_{l_i-1}}p_{y_{l_i-1}x} \Bigg) .$$

Notice that we have independence of the edges because we ensured that the paths are edge-disjoint.

We have using (1.4) and then (1.2) that

$$\sum_{y_1,\dots,y_{l_i-1}} p_{xy_1}p_{y_1y_2}\cdots p_{y_{l_i-2}y_{l_i-1}}p_{y_{l_i-1}x} \leqslant \sum_{y_1,\dots,y_{l_i-1}} \frac{w_x^2 w_{y_1}^2 \cdots w_{l_i-1}^2}{m_1^{l_i} N^{l_i}} = \frac{w_x^2}{m_1 N}\left(\frac{m_2}{m_1}\right)^{l_i-1} .$$

Assumptions 1.1 and 1.2 then imply

$$\sum_{y_1,\dots,y_{l_i-1}} p_{xy_1}p_{y_1y_2}\cdots p_{y_{l_i-2}y_{l_i-1}}p_{y_{l_i-1}x} = O(N^{-\varepsilon/2})$$

Finally, we have

$$\mathbb{P}(\# S_1^{\mathrm{cyc}}(x) \geqslant 3k) \leqslant O(N^{-k\varepsilon/2}).$$

Choosing k big enough gives the result. □

2.3. Coupling with a tree. The problem of the graph G^{nc} is that its edges are not independent. To solve this problem, we are going to introduce a new graph $\mathcal{T}_x$, whose edges are independent. This graph $\mathcal{T}_x$ will be a forest. Actually, we introduce two versions of the forest, $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$ with $\check{\mathcal{T}}_x \subset \mathcal{T}_x$. The reason is that while $\mathcal{T}_x$ has independent edges and is more convenient, it may have too many edges for some purposes (see Corollary 2.16 below). The graph $\check{\mathcal{T}}_x$ has less edges but its edge are independent conditionally to an appropriate σ-algebra. The balls of small radius in the graph G^{nc} around a fixed vertex $x \in [N]$ may be coupled with $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$.

Let us explain breifly the construction of the enlarged graphs $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$. Their vertices are indexed by families of vertices $\gamma = (x, y_1, \dots, y_d)$ for any length $d \geqslant 0$ of G. To define the edge set, we introduce Bernoulli random variables Z_γ such that there is an edge between $(x, y_1, \dots, y_{d-1})$ and γ in $\mathcal{T}_x$ (respectively, in $\check{\mathcal{T}}_x$) if and only if $Z_\gamma = 1$ (respectively, $\check{Z}_\gamma = 1$). For $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$ to be coupled in a convenient way with the balls centered on x in G^{nc}, we will have to make choice as to which edges of G we keep in $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$. The precise construction is as follows.

Let $x \in [N]$. We introduce the family of independent random variables $(\hat{Z}_{xy_1\cdots y_k}; k \geqslant 2, y_1, \dots, y_k \in [N])$ such that for all $k \geqslant 2$ and $y_1, \dots, y_k \in [N]$, $\hat{Z}_{xy_1\cdots y_k}$ is a Bernoulli random variable with parameter $p_{y_{k-1}y_k}$. For $k \geqslant 2$ and $y_k \in S_k(x)$ we define the path

$$\gamma^*(y_k) = \min\Big\{ \gamma \in \mathcal{P}^*_{xy_k}(G) \colon l(\gamma) = k \Big\} ,$$

where the minimum is with respect to the lexicographic order (for the usual order on $[N]$). Said otherwise, it is the path of length k between x and y_k that is minimal for the lexicographic order. We are now ready to introduce the variables describing the edges of the random tree. We define $\check{Z}_{xy_1}$ and Z_{xy_1} by

$$\check{Z}_{xy_1} = Z_{xy_1} = \mathbb{1}_{\{x \sim y_1\}} \quad \text{for } y_1 \in [N] \setminus \{x\} .$$

and for $k \geqslant 1$ and $y_1, \dots, y_{k+1} \in [N]$, $Z_{xy_1\cdots y_{k+1}}$ is defined by

$$Z_{xy_1\cdots y_{k+1}} := \begin{cases} \mathbb{1}_{\{y_k \sim y_{k+1}\}} & \text{if } y_{k+1} \notin B_k(x), y_k \in S_k(x), \text{ and } (x, y_1, \dots, y_k) = \gamma^\star(y_k) \\ \hat{Z}_{xy_1\cdots y_k y_{k+1}} & \text{otherwise.} \end{cases}$$

The corresponding version for $\check{\mathcal{T}}_x$ is

$$\check{Z}_{xy_1\cdots y_{k+1}} := \begin{cases} \mathbb{1}_{\{y_k \sim y_{k+1}\}} & \text{if } y_{k+1} \notin B_k(x), y_k \in S_k(x), \text{ and } (x, y_1, \dots, y_k) = \gamma^\star(y_k) \\ \hat{Z}_{xy_1\cdots y_k y_{k+1}} & \text{if } y_{k+1} \notin B_k(x), y_k \in S_k(x), \text{ and } (x, y_1, \dots, y_k) \neq \gamma^\star(y_k) \\ 0 & \text{otherwise.} \end{cases}$$

We introduce the filtration $(\mathcal{F}_k(x))_{k\geqslant 1}$ given by

$$\mathcal{F}_k(x) = \sigma\left(\left\{\{\gamma \text{ is a path in } G\} \colon \gamma \in \bigcup_{i=1}^{k} \{x\} \times [N]^i\right\} \cup \left\{\hat{Z}_\gamma \colon \gamma \in \bigcup_{i=1}^{k} \{x\} \times [N]^i\right\}\right), \tag{2.2}$$

i.e. it the σ-algebra generated by the events that paths starting from x of length smaller that k belong in G, and the random variables $\hat{Z}_\gamma$ with γ of size at most k. Note that given $\gamma = (x, y_1, \dots, y_k)$, the event $\{y_k \in S_k(x), \gamma = \gamma^*(y_k)\}$ belongs to $\mathcal{F}_k(x)$. Hence, we see that given $k \geqslant 1$, the random variables $Z_{xy_1\dots y_{k+1}}, y_1, \dots, y_{k+1} \in [N]$ (respectively, the random variables $\check{Z}_{xy_1\dots y_{k+1}}, y_1, \dots, y_{k+1} \in [N]$) are independent conditionally to $\mathcal{F}_k(x)$. Furthermore, given $y_1, \dots, y_{k+1} \in [N]$ and conditionally on $\mathcal{F}_k(x)$,

- $Z_{xy_1\cdots y_{k+1}}$ is a Bernoulli random variable of parameter $p_{y_k y_{k+1}}$.
- $\check{Z}_{xy_1\cdots y_{k+1}}$ is a Bernoulli random variable of parameter $p_{y_k y_{k+1}} \mathbb{1}_{\{y_{k+1} \notin B_k(x), y_k \in S_k(x)\}}$.

Definition 2.10 (Forests $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$)**.** The graph $\mathcal{T}_x$ (respectively $\check{\mathcal{T}}_x$) is the graph

- with vertex set $V_x = \{x\} \cup \bigcup_{d\geqslant 1} \{x\} \times [N]^d$, and
- such that given $\gamma = (x, y_1, \dots, y_d) \in V_x$ and $\gamma' = (x, y_1', \dots, y_{d'}') \in V_x$ with $d \leqslant d'$, we have $\{\gamma, \gamma'\} \in \mathcal{T}_x$ if and only if $d' = d + 1$, $y_1 = y_1', \dots, y_d = y_d'$, and $Z_{xy_1'\cdots y_{d'}'} = 1$ (respectively $\check{Z}_{xy_1'\cdots y_{d'}'} = 1$).

Note that both $\check{\mathcal{T}}_x$ and $\mathcal{T}_x$ are infinite forests, while the connected components of $\check{\mathcal{T}}_x$ and $\mathcal{T}_x$ containing x are trees naturally rooted at x. We indicate by an exponent $\mathcal{T}_x$ (or $\check{\mathcal{T}}_x$) the spheres, degree, …associated to $\mathcal{T}_x$ (or $\check{\mathcal{T}}_x$). For instance, we write $D_\gamma^{\mathcal{T}_x}$ the degree of $\gamma \in V_x$ in $\mathcal{T}_x$. For convenience, we also define for all $\gamma = (x, y_1, \dots, y_d, y_{d+1}) \in V_x$ the number of children of γ:

$$D_\gamma^{\mathcal{T}_x \uparrow} := D_\gamma^{\mathcal{T}_x} - \mathbb{1}_{\left\{\gamma \neq (x) \overset{\mathcal{T}_x}{\sim} \gamma\right\}}. \tag{2.3}$$

Remark 2.11. Introducing the graph $\mathcal{T}_x$ is motivated by the fact that the edges of G^{nc} are not independent. At the cost of introducing a small number of additional edges, the graph $\check{\mathcal{T}}_x$ has edges at depth k that are independent conditionally to the σ-algebra $\mathcal{F}_k(x)$. The graph $\mathcal{T}_x$ enjoys a stronger property: the edges of $\mathcal{T}_x$ are all independent as we show in Lemma 2.12.

Lemma 2.12. *The random variables Z_γ for $\gamma \in V_x$ are all independent.*

Proof. It suffices to show that for every $k \geqslant 2$ and $\gamma_1, \dots, \gamma_k \in V_x$ distinct vertices of $\mathcal{T}_x$, we have

$$\mathbb{P}\Big(Z_{\gamma_1} = 1; \cdots ; Z_{\gamma_k} = 1\Big) = \prod_{i=1}^{k} p_{y_{i,d_i} y_{i,d_i+1}},$$

where for all i, $\gamma_i = (x, y_{i,1}, \ldots, y_{i,d_i}, y_{i,d_i+1})$. We may assume that $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_k$, and that $d_p = d_{p+1} = \cdots = d_k$ for some $1 \leqslant p \leqslant k$. Since $\{Z_{\gamma_1} = 1; \cdots ; Z_{\gamma_{p-1}} = 1\}$ is $\mathcal{F}_{d_k}(x)$-measurable, we have by conditional independence

$$\mathbb{P}\Big(Z_{\gamma_1} = 1; \cdots ; Z_{\gamma_k} = 1 \mid \mathcal{F}_{d_k}(x)\Big) = \mathbb{1}_{\{Z_{\gamma_1}=1;\cdots;Z_{\gamma_{p-1}}=1\}} \prod_{i=p}^{k} p_{y_{i,d_i} y_{i,d_i+1}} \,.$$

Proceeding by induction, we get the result. $\square$

Remark 2.13. The forests $\mathcal{T}_x$ and $\check{\mathcal{T}}_x$ are constructed in such a way that if γ and γ are two distinct elements of V_x of the same length $d = d(x,y)$ from x to y that are present in G, we have that

- conditionally to $\mathcal{F}_d(x)$, $D_{\gamma}^{\mathcal{T}_x \uparrow}$ and $D_{\gamma'}^{\mathcal{T}_x \uparrow}$ are independent and identically distributed;
- $D_{\gamma}^{\mathcal{T}_x \uparrow}$ and $D_{\gamma'}^{\mathcal{T}_x \uparrow}$ are independent and identically distributed.

Lemma 2.14. *Let $y \in B_r^{\mathrm{nc}}(x)$. Then, there exists a unique path $\gamma \in \mathcal{P}_{xy}^*(G)$ with $l(\gamma) \leqslant r$. Furthermore, $y \in S_{l(\gamma)}^{\mathrm{nc}}(x)$ and $\gamma = \gamma^*(y)$.*

Proof. The fact that there exists a unique path $\gamma \in \mathcal{P}_{xy}^*(G)$ with $l(\gamma) \leqslant r$ follows by construction of G^{nc}. Indeed, if there existed $\gamma' \in \mathcal{P}_{xy}^*(G)$ with $\gamma' \neq \gamma$, we could consider the concatenation $\tilde{\gamma}$ of γ and γ' with its order reversed and construct a path contained in $B_r(x)$ from x to x. While $\tilde{\gamma}$ would not be necessarily simple, there would exists x' such that $\tilde{\gamma}$ contains a simple path in $\mathcal{P}_{x'x'}^*(G)$ which is contained in $B_r(x')$. Hence, one vertex y' in γ would belong to $S_1^{\mathrm{cyc}}(x')$ and thus the edge $\{x', y'\}$ would not appear in G^{nc}. Hence, the subset of $\mathcal{P}_{xy}^*(G^{\mathrm{nc}})$ comprising paths of length at most r must be empty, contradicting $y \in B_r^{\mathrm{nc}}$.

Hence, we get that $d(x,y) = l(\gamma)$, i.e. $y \in S_{l(\gamma)}(x)$ and $y \in S_{l(\gamma)}^{\mathrm{nc}}(x)$. It then follows that $\gamma = \gamma^*(y)$. $\square$

The graph $G^{\mathrm{nc}}|_{B_r^{\mathrm{nc}}(x)}$ can be embedded in $\check{\mathcal{T}}_x$, and hence in $\mathcal{T}_x$, as follows. We define a mapping $\iota \colon B_r^{\mathrm{nc}}(x) \to V_x$ by

$$\iota(x) = (x)\,.$$

Then, for all $y \in B_r^{\mathrm{nc}}(x) \setminus \{x\}$ we use Lemma 2.14 and set

$$\iota(y) = \gamma^*(y)\,.$$

Proposition 2.15. *The mapping ι defines a graph embedding of $G^{\mathrm{nc}}|_{B_r^{\mathrm{nc}}(x)}$ into $\check{\mathcal{T}}_x$ and $\mathcal{T}_x$.*

Proposition 2.15 will be used several times in the sequel. An important implication of this result is the following corollary.

Corollary 2.16. *Let $y \in B_{r-1}^{\mathrm{nc}}(x)$. Then, we have*

$$D_y^{\mathrm{nc}} \leqslant D_{\iota(y)}^{\check{\mathcal{T}}_x} \leqslant D_{\iota(y)}^{\mathcal{T}_x}\,.$$

Furthermore, there exists a constant $C_\nu > 0$ such that with ν-high probability

$$D_{\iota(y)}^{\check{\mathcal{T}}_x} \leqslant D_y^{\mathrm{nc}} + C_\nu\,.$$

Proof of Corollary 2.16. Let $y \in B^{\mathrm{nc}}_{r-1}(x)$. The first inequality is a direct consequence of Proposition 2.15: since ι is a graph embedding the set of neighbors of $\iota(y)$ in $B^{\mathrm{nc}}_{r-1}(x)$ contains the image of the set of neighbors of y.

Let us prove the second inequality. By Lemma 2.14, we have that $d := d(x, y) = l(\gamma^*(y))$. Hence, $y \in S_d(x)$. We have

$$D^{\check{\mathcal{T}}_x}_{\iota(y)} = \mathbb{1}_{\left\{\left(\gamma^*(y)\right)_{d-1} \sim y\right\}} + \sum_{z \notin B_d(x)} \mathbb{1}_{\{y \sim z\}} \leqslant D_y \leqslant D^{\mathrm{nc}}_y + C_\nu \,,$$

where the last inequality holds with ν-high probability by Proposition 2.7. □

Proof of Proposition 2.15. It is clear that ι is injective. Let $\{y, y'\} \in G^{\mathrm{nc}}|_{B^{\mathrm{nc}}_r(x)}$, and let us show that $\{\gamma(y)^*, \gamma(y')^*\} \in \mathcal{T}_x$. We can assume without loss of generality that $d := l(\gamma^*(y)) \leqslant l(\gamma^*(y'))$. Since $G^{\mathrm{nc}}|_{B^{\mathrm{nc}}_r(x)}$ is a tree,

$$\gamma^*(y') = (x, \gamma^*(y)_1, \ldots, \gamma^*(y)_{l(\gamma(y))}, y') \,,$$

and in particular $l(\gamma^*(y')) = d + 1$. It remains to show that $Z_{\gamma^*(y')} = \check{Z}_{\gamma^*(y')} = 1$. This follows from Lemma 2.14: indeed, we have $y \in S_d(x)$ and $y' \notin B_d(x)$ so $Z_{\gamma^*(y')} = \check{Z}_{\gamma^*(y')} = \mathbb{1}_{\{y \sim y'\}} = 1$. □

2.4. Removing the down-up paths. As explained in Section 2.1, step 2 of the pruning procedure consists in removing the edges $\{x, y\} \in G$ where $x \in [N]$ and $y \in S^{\mathrm{du}}_1(x) \setminus S^{\mathrm{cyc}}_1(x)$. When removing such an edge, we change the degree of both x and y. Controlling how these two degrees change is the content of Lemmas 2.17 and 2.18 below. More precisely, if we consider a vertex $x \in [N]$ and $y \in S_1(x) \setminus S^{\mathrm{cyc}}_1(x)$, then

- either $y \prec x$, and if $y \in S^{\mathrm{du}}_1(x)$, we remove $\{x, y\}$ from G;
- or $y \succ x$, and if there exists $y' \in S^+_1(x) \setminus S^{\mathrm{cyc}}_1(x)$ with $y \prec y'$, we remove $\{x, y\}$ from G.

It implies that

$$D^{\mathrm{p}}_x = D_x - \#S^{\mathrm{cyc}}_1(x) - \#S^{\mathrm{du}}_1(x) - \Big(\#\Big(S^+_1(x) \setminus S^{\mathrm{cyc}}_1(x)\Big) - 1\Big) \vee 0.$$

We have upper bounded the second term in Proposition 2.7. Lemma 2.17 bounds the third term, and Lemma 2.18 bounds the fourth one, as $\#S^+_1(x) = D^+_x$.

Lemma 2.17. *Let $\nu > 0$ and $x \in [N]$. With ν-high probability,*

$$\#S^{\mathrm{du}}_1(x) \leqslant \frac{2\nu}{1 - 2\delta} \frac{\log N}{\log \log N} \,.$$

Lemma 2.18. *Let $x \in [N]$ and $\nu > 0$. Introduce*

$$D^{\mathrm{nc}+}_x := \#\Big(S^+_1(x) \setminus S^{\mathrm{cyc}}_1(x)\Big) \,. \tag{2.4}$$

With ν-high probability,

$$D^{\mathrm{nc}+}_x \leqslant \frac{3}{2} \frac{\nu}{1 - \delta} \frac{\log N}{\log \log N} \,.$$

Thus, Lemmas 2.17 and 2.18 show that we remove roughly $\log N / \log \log N$ edges around each vertex when removing the down-up paths.

Proof of Lemma 2.17. We start by noticing that if $D_x < \frac{\log N}{\log\log N}$ the result is immediate as $\# S_1^{\mathrm{du}}(x) \leqslant D_x$. We thus assume that $D_x \geqslant \frac{\log N}{\log\log N}$. Assuming this, we introduce $\chi = 4\nu \log\log N$ and use Remark 1.13 to get

$$\chi D_x \geqslant w_x$$

with ν-high probability. This gives us

$$\# S_1^{\mathrm{du}}(x) = \sum_{y \in [N]} \mathbb{1}_{\left\{x \overset{\mathrm{nc}}{\sim} y\right\}} \mathbb{1}_{\left\{\exists z \neq x, y \overset{\mathrm{nc}}{\sim} z, D_z \geqslant D_x \geqslant D_y\right\}} \leqslant \sum_{y \in [N]} \mathbb{1}_{\left\{x \overset{\mathrm{nc}}{\sim} y\right\}} \mathbb{1}_{\left\{\exists z \neq x, y \overset{\mathrm{nc}}{\sim} z, \chi D_z \geqslant w_x\right\}},$$

with ν-high probability.

We now relate the quantity $\# S_1^{\mathrm{du}}(x)$ to the tree $\mathcal{T}_x$. To do so we use Proposition 2.7: there exists a constant $C_\nu > 0$ such that with ν-high probability

$$\# S_1^{\mathrm{du}}(x) \leqslant \sum_{y \in [N]} \mathbb{1}_{\left\{x \overset{\mathrm{nc}}{\sim} y\right\}} \mathbb{1}_{\left\{\exists z \neq x, y \overset{\mathrm{nc}}{\sim} z, \chi\left(D_z^{\mathrm{nc}} + C_\nu\right) \geqslant w_x\right\}}.$$

Proposition 2.15 implies that

$$\# S_1^{\mathrm{du}}(x) \leqslant \sum_{y \in [N]} \mathbb{1}_{\left\{x \overset{\mathcal{T}_x}{\sim} y\right\}} \mathbb{1}_{\left\{\exists z \neq x, y \overset{\mathcal{T}_x}{\sim} z, \chi\left(D_z^{\mathcal{T}_x} + C_\nu\right) \geqslant w_x\right\}}.$$

To get the result, we are going to use Bennett's inequality [BLM13, Theorem 2.9]. The Bernoulli random variables

$$\left(\mathbb{1}_{\left\{x \overset{\mathcal{T}_x}{\sim} y\right\}} \mathbb{1}_{\left\{\exists z \neq x, y \overset{\mathcal{T}_x}{\sim} z, \chi\left(D_z^{\mathcal{T}_x} + C_\nu\right) \geqslant w_x\right\}}\right)_{y \in [N]}$$

are independent, since by Lemma 2.12 the edges in $\mathcal{T}_x$ are independent. We compute

$$v := \sum_{y \in [N]} \mathbb{E}\left[\mathbb{1}_{\left\{x \overset{\mathcal{T}_x}{\sim} y\right\}} \mathbb{1}_{\left\{\exists z \neq x, y \overset{\mathcal{T}_x}{\sim} z, \chi\left(D_z^{\mathcal{T}_x} + C_\nu\right) \geqslant w_x\right\}}\right] = \sum_{y \in [N]} p_{xy} \mathbb{P}\left(\exists z \neq x, y \overset{\mathcal{T}_x}{\sim} z, \chi(D_z^{\mathcal{T}_x} + C_\nu) \geqslant w_x\right).$$

The union bound followed by Markov's inequality yield

$$\begin{aligned} \mathbb{P}\left(\exists z \neq x, y \overset{\mathcal{T}_x}{\sim} z, \chi(D_z^{\mathcal{T}_x} + C_\nu) \geqslant w_x\right) &\leqslant \sum_{z \neq x} p_{yz} \mathbb{P}\left(\chi(D_z^{\mathcal{T}_x} - \mathbb{1}_{\left\{y \overset{\mathcal{T}_x}{\sim} z\right\}} + 1 + C_\nu) \geqslant w_x\right) \\ &\leqslant \sum_{z \neq x} p_{yz} \chi \frac{d_z + 1 + C_\nu}{d_x}. \end{aligned}$$

Using (1.2), (1.4), and the crude bound $w_x/2 \leqslant d_x \leqslant w_x$, we get

$$\mathbb{P}\left(\exists z \neq x, y \overset{\mathcal{T}_x}{\sim} z, \chi(D_z^{\mathcal{T}_x} + C_\nu) \geqslant d_x\right) \leqslant \sum_z \frac{w_y w_z^2 + (1 + C_\nu) w_y w_z}{m_1 N} \frac{\chi}{d_x} \leqslant 2\chi \frac{w_y}{w_x} \frac{m_2 + (C_\nu + 1) m_1}{m_1}.$$

Using (1.2) and (1.4) again, we get

$$v \leqslant 2\chi \frac{m_2}{m_1} \left(\frac{m_2}{m_1} + C_\nu + 1\right).$$

Bennett's inequality then implies the result:

$$\begin{aligned}
&\mathbb{P}\Big(\#S_1^{\mathrm{du}} \geqslant \frac{2\nu}{1-2\delta}\frac{\log N}{\log\log N}\Big) \\
&\qquad\leqslant \mathbb{P}\Big(\sum_{y\in[N]} \mathbb{1}_{\{x\overset{\mathcal{T}_x}{\sim}y\}}\mathbb{1}_{\{\exists z\neq x, y\overset{\mathcal{T}_x}{\sim}z, \chi\left(D_z^{\mathcal{T}_x}+C_\nu\right)\geqslant w_x\}} \geqslant \frac{2\nu}{1-2\delta}\frac{\log N}{\log\log N}\Big) + O(N^{-\nu}) \\
&\qquad\leqslant \exp\Big(-(v+\frac{2\nu}{1-2\delta}\frac{\log N}{\log\log N})\ln(1+\frac{2\nu\log N}{v(1-2\delta)\log\log N})(1+o(1))\Big) + O(N^{-\nu}) \\
&\qquad= \exp(-2\nu\log N(1+o(1))) + O(N^{-\nu}) = O(N^{-\nu})\,. \qquad\square
\end{aligned}$$

Proof of Lemma 2.18. Notice that if $D_x < \frac{\log N}{\log\log N}$ the result is immediate. Hence, we assume that $D_x \geqslant \frac{\log N}{\log\log N}$. We introduce $\chi = 4\nu\log\log N$. By Remark 1.13 we have that with ν-high probability that

$$\chi D_x \geqslant d_x\,.$$

Consider the random variable

$$D_x^{\mathrm{nc}+} = \#\Big(S_1^+(x)\setminus S_1^{\mathrm{cyc}}(x)\Big) = \sum_{y\in[N]} \mathbb{1}_{\{x\overset{\mathrm{nc}}{\sim}y, D_x\leqslant D_y\}}\,.$$

By the previous discussion, we have with ν-high probability

$$D_x^{\mathrm{nc}+} \leqslant \sum_{y\in[N]} \mathbb{1}_{\left\{x\overset{\mathrm{nc}}{\sim}y, \chi D_y\geqslant d_x\right\}}\,.$$

By Proposition 2.7, there exists a constant C_ν such that with ν-high probability, we have

$$D_x^{\mathrm{nc}+} \leqslant \sum_{y\in[N]} \mathbb{1}_{\left\{x\overset{\mathrm{nc}}{\sim}y, \chi\left(D_y^{\mathrm{nc}}+C_\nu\right)\geqslant d_x\right\}}\,.$$

Proposition 2.15 allow us to bound this using the tree $\mathcal{T}_x$:

$$D_x^{\mathrm{nc}+} \leqslant D_x^{\mathcal{T}_x+} := \sum_{y\in V_x} \mathbb{1}_{\left\{x\overset{\mathcal{T}_x}{\sim}y, \chi\left(D_y^{\mathcal{T}_x\uparrow}+1+C_\nu\right)\geqslant d_x\right\}}\,.$$

As in the proof of Lemma 2.17, we conclude using Bennett's inequality (see [BLM13, Theorem 2.9]). The random variables

$$\left(\mathbb{1}_{\left\{x\overset{\mathcal{T}_x}{\sim}y\right\}}\mathbb{1}_{\left\{\chi\left(D_y^{\mathcal{T}_x\uparrow}+1+C_\nu\right)\geqslant d_x\right\}}\right)_{y\in V_x}$$

are independent by Lemma 2.12 and bounded by 1. Markov's inequality gives us:

$$v := \sum_{y\in V_x}\mathbb{E}\Big[\mathbb{1}_{\left\{x\overset{\mathcal{T}_x}{\sim}y, \chi\left(D_y^{\mathcal{T}_x\uparrow}+1+C_\nu\right)\geqslant d_x\right\}}\Big] \leqslant \sum_{y\in V_x} p_{xy}\chi\frac{d_y+1+C_\nu}{d_x}\,.$$

Using (1.2), (1.4), and the crude bound $w_x/2 \leqslant d_x \leqslant w_x$, we get

$$v \leqslant 2\chi\frac{m_2+(C_\nu+1)m_1}{m_1}\,.$$

Bennett's inequality then implies

$$\mathbb{P}\left(D_x^{\mathrm{nc}+} \leqslant \frac{3}{2}\frac{\nu}{1-\delta}\frac{\log N}{\log\log N}\right) \leqslant \mathbb{P}\left(D_x^{\mathcal{T}_x+} \leqslant \frac{3}{2}\frac{\nu}{1-\delta}\frac{\log N}{\log\log N}\right) + O(N^{-\nu}) \leqslant O(N^{-\nu})\,. \qquad\square$$

With the preceding results, Proposition 2.7 and Lemmas 2.17 and 2.18, we are ready to prove Proposition 2.5.

Proof of Proposition 2.5. Let $x \in [N]$. During the pruning procedure, we remove at most

$$D_x^{\mathrm{nc}+} + \# S_1^{\mathrm{du}}(x) + \# S_1^{\mathrm{cyc}}(x)$$

edges around x. Proposition 2.7 together with Lemmas 2.18 and 2.17 yield claim 1. The two other claims are consequence of the construction of G^{p}. Let us detail the third one. Assume that there is a simple loop in G^{p} composed of the vertices $(\gamma_0, \gamma_1, \ldots, \gamma_k = \gamma_0)$ with $\{\gamma_{i-1}, \gamma_i\} \in G^{\mathrm{p}}$ for all $i \in [k]$. Then there is a vertex, say γ_0, which is minimal for the total order $\prec$. Then $\gamma_{k-1} \succ \gamma_k = \gamma_0 \prec \gamma_1$, and either $(\gamma_{k-1}, \gamma_0, \gamma_1)$ or $(\gamma_1, \gamma_0, \gamma_{k-1})$ is a down-up path. As there is no such path in G^{p}, there are no cycle in G^{p}. $\square$

2.5. Estimate of $\|A - A^{\mathrm{p}}\|$. We now give estimates for the error we make when working with the adjacency matrix of the pruned graph A^{p} rather than the adjacency matrix of the original graph A.

Proposition 2.19. *Let $\nu > 0$. There exists a constant $C_\nu > 0$ such that with ν-high probability,*

$$\|A - A^{\mathrm{p}}\| \leqslant C_\nu \sqrt{\frac{\log N}{\log \log N}} .$$

The proof of Proposition 2.19 relies on Lemma 2.20 and Lemma 2.21 stated below. They are stated using a partition of the set of vertices $[N] = \mathcal{V}^{(\mathrm{l})} \sqcup \mathcal{V}^{(\mathrm{i})} \sqcup \mathcal{V}^{(\mathrm{h})}$, where

$$\begin{aligned} \mathcal{V}_\nu^{(\mathrm{l})} &= \{x \in [N] : D_x < \xi_\nu, w_x \leqslant 4\xi_\nu\} && \text{(vertices of low degree)} \\ \mathcal{V}_\nu^{(\mathrm{i})} &= \{x \in [N] : D_x < \xi_\nu, w_x > 4\xi_\nu\} && \text{(vertices of intermediate degree)} \\ \mathcal{V}_\nu^{(\mathrm{h})} &= \{x \in [N] : \xi_\nu \leqslant D_x\} && \text{(vertices of high degree)} \ . \end{aligned} \tag{2.5}$$

Recall that the threshold ξ_ν was defined in Proposition 2.5. The reason why we partition $[N]$ into three set of vertices will become apparent in the proof of Proposition 2.19. The key point is that treating the vertices belonging in different sets $\mathcal{V}_\nu^{(\circ)}, \circ \in \{\mathrm{l}, \mathrm{i}, \mathrm{h}\}$ require different techniques. In particular, treating vertices in $\mathcal{V}_\nu^{(\mathrm{h})}$ requires using in a fine way the properties of down-up paths, while treating vertices in $\mathcal{V}_\nu^{(\mathrm{l})}$ is done by using results in [BGBK20] about sparse random matrices with small weights. The two lemmas concerning vertices of high and intermediate degree are the following.

Lemma 2.20. *Let $\nu > 0$. There exists $C_\nu > 0$ such that for all $x_1 \in \mathcal{V}_\nu^{(\mathrm{h})}$,*

$$\sum_{\substack{x_2, x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \succ x_1, x_3}} \left\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \right\rangle \left\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2) \setminus S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \right\rangle \leqslant C_\nu \left(\frac{\log N}{\log \log N} \right)^2 .$$

Lemma 2.21. *There exists a constant $C_\nu > 0$ such that for all $x_1 \in \mathcal{V}_\nu^{(\mathrm{i})}$:*

$$\sum_{\substack{x_2 \in \mathcal{V}_\nu^{(\mathrm{i})} \\ x_2 \neq x_1}} \left\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \right\rangle \leqslant 2\nu \frac{\log N}{\log \log N}$$

with ν-high probability.

The two previous lemmas are based on probabilistic estimates, while the proof of Proposition 2.19 below contains mainly algebraic arguments.

Proof of Proposition 2.19. The matrix $A - A^{\mathrm{p}}$ is the adjacency matrix of the graph made of the edges removed during the pruning. Consider the adjacency matrix A^{nc} of the graph G^{nc} obtained after step 1 of the pruning procedure Definition 2.4. By Proposition 2.7, the maximum degree of a vertex in the graph described by $A - A^{\mathrm{nc}}$ is bounded by a constant C_ν, with ν-high probability. It implies that $\|A - A^{\mathrm{nc}}\| \leqslant C_\nu$, with ν-high probability. Thus, it suffices to bound $\|A^{\mathrm{nc}} - A^{\mathrm{p}}\|$.

Let us introduce for convenience the matrices

$$\tilde{A}_\circ = \sum_{x \in \mathcal{V}^{(\circ)}} \mathbf{1}_x \mathbf{1}^*_{S_1^{\mathrm{du}}(x)} \quad \text{for } \circ \in \{\mathrm{l}, \mathrm{i}, \mathrm{h}\},$$

so that

$$A^{\mathrm{p}} - A^{\mathrm{nc}} = \sum_{\circ \in \{\mathrm{l},\mathrm{i},\mathrm{h}\}} \tilde{A}_\circ + \tilde{A}_\circ^*.$$

Let us now bound the norms of the operators $\tilde{A}_\circ$, starting with $\circ = \mathrm{h}$. Let us consider the matrix $\tilde{A}_{\mathrm{h}} \tilde{A}_{\mathrm{h}}^*$. We have

$$\begin{aligned}
\tilde{A}_{\mathrm{h}} \tilde{A}_{\mathrm{h}}^* &= \sum_{x_1, x_2 \in \mathcal{V}_\nu^{(\mathrm{h})}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \mathbf{1}_{x_1} \mathbf{1}^*_{x_2} \\
&= \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \Big(\# S_1^{\mathrm{du}}(x)\Big) \mathbf{1}_x \mathbf{1}_x^* + \sum_{\substack{x_1, x_2 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_1 \neq x_2}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \mathbf{1}_{x_1} \mathbf{1}^*_{x_2}.
\end{aligned}$$

The first term, a diagonal matrix, has its operator norm bounded by $\frac{2\nu}{1-2\delta} \frac{\log N}{\log \log N}$ with ν-high probability by Lemma 2.17. Let us concentrate on the second term, which we call B. We introduce $B_\succ$:

$$B_\succ := \sum_{\substack{x_1, x_2 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_1 \succ x_2}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \mathbf{1}_{x_1} \mathbf{1}^*_{x_2} = \sum_{\substack{x_1, x_2 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_1 \succ x_2}} \Big(\# S_1^{\mathrm{du}}(x_1) \cap S_1^{\mathrm{du}}(x_2)\Big) \mathbf{1}_{x_1} \mathbf{1}^*_{x_2},$$

so that $B = B_\succ + B_\succ^*$. We then have

$$\begin{aligned}
B_\succ B_\succ^* &= \sum_{\substack{x_1, x_2, x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \prec x_1, x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \Big\rangle \mathbf{1}_{x_1} \mathbf{1}^*_{x_3} \\
&= \sum_{\substack{x_1, x_2, x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \prec x_1, x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2) \setminus S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \Big\rangle \mathbf{1}_{x_1} \mathbf{1}^*_{x_3} \\
&\quad + \sum_{\substack{x_1, x_2, x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \prec x_1, x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1) \cap S_1^{\mathrm{du}}(x_2)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \Big\rangle \mathbf{1}_{x_1} \mathbf{1}^*_{x_3}.
\end{aligned}$$

We thus have

$$\begin{aligned}
\|B_\succ\|^2 \leqslant \max_{\|\mathbf{u}\|=1} & \sum_{\substack{x_1, x_2, x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \prec x_1, x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2) \setminus S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \Big\rangle u_{x_1} u_{x_3} \\
& + \max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1, x_2, x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \prec x_1, x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1) \cap S_1^{\mathrm{du}}(x_2)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \Big\rangle u_{x_1} u_{x_3}.
\end{aligned} \tag{2.6}$$

The first term of (2.6) can be bounded using Young's inequality:

$$\max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1,x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2)\setminus S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)}\Big\rangle u_{x_1}u_{x_3}$$
$$\leqslant \max_{\|\mathbf{u}\|=1} \sum_{x_1\in\mathcal{V}^{(\mathrm{h})}} u_{x_1}^2 \sum_{\substack{x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2)\setminus S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)}\Big\rangle\,.$$

Then, Lemma 2.20 allows to bound this by $2\nu\Big(\frac{\log N}{\log\log N}\Big)^2$. The second term of (2.6) can be treated as follows:

$$\max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1,x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_2)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)}\Big\rangle u_{x_1}u_{x_3}$$
$$= \max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1,x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \sum_{y_1,y_2} \mathbb{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_2)}(y_1)\mathbb{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_2)\cap S_1^{\mathrm{du}}(x_3)}(y_2)u_{x_1}u_{x_3}$$
$$= \max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1,x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \sum_{y} \mathbb{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_2)\cap S_1^{\mathrm{du}}(x_3)}(y)u_{x_1}u_{x_3}\,,$$

where we could remove one of the sum on y as if $y_1\neq y_2$ we would be considering the case of having a cycle $x_1 \overset{\mathrm{nc}}{\sim} y_1 \overset{\mathrm{nc}}{\sim} x_2 \overset{\mathrm{nc}}{\sim} y_2 \overset{\mathrm{nc}}{\sim} x_1$ in G^{nc}, which contradicts the definition of G^{nc}. We then notice that $\#\Big\{x_2\colon y\in S_1^{\mathrm{du}}(x_2)\Big\}\leqslant D_y^{\mathrm{nc}+}$, so that with ν-high probability Lemma 2.18 gives

$$\max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1,x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_2)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)}\Big\rangle u_{x_1}u_{x_3}$$
$$\leqslant \max_{\|\mathbf{u}\|=1} \sum_{x_1,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}} \sum_{y} D_y^{\mathrm{nc}+}\mathbb{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_3)}(y)u_{x_1}u_{x_3}$$
$$\leqslant \frac{3\nu}{2(1-\delta)}\frac{\log N}{\log\log N} \max_{\|\mathbf{u}\|=1} \sum_{x_1,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}} \#\Big(S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_3)\Big)u_{x_1}u_{x_3}\,.$$

Note that since the entries of the matrices we consider are non-negative, we can assume that $\mathbf{u}$ has positive coefficients. This allows the upper bound in the latter expression. We then recognize the norm of the symmetric matrix B whose definition is

$$B = \sum_{\substack{x_1,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_1\neq x_3}} \langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}\rangle \mathbf{1}_{x_1}\mathbf{1}_{x_3}^* = \sum_{\substack{x_1,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_1\neq x_3}} \#\Big(S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_3)\Big)\mathbf{1}_{x_1}\mathbf{1}_{x_3}^*\,,$$

so that with ν-high probability

$$\max_{\|\mathbf{u}\|=1} \sum_{\substack{x_1,x_2,x_3\in\mathcal{V}_\nu^{(\mathrm{h})}\\ x_2\prec x_1,x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)\cap S_1^{\mathrm{du}}(x_2)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)}\Big\rangle u_{x_1}u_{x_3}$$
$$\leqslant \frac{3\nu}{2(1-\delta)}\frac{\log N}{\log\log N}\|B\| + \frac{3\nu}{2(1-\delta)}\frac{\log N}{\log\log N}\max_{\|\mathbf{u}\|=1}\sum_{x\in\mathcal{V}^{(\mathrm{h})}}\Big(\#S_1^{\mathrm{du}}(x)\Big)u_x^2$$
$$\leqslant \frac{3\nu}{2(1-\delta)}\frac{\log N}{\log\log N}\|B\| + \frac{6\nu^2}{2(1-\delta)(1-2\delta)}\Big(\frac{\log N}{\log\log N}\Big)^2\,.$$

where we used a second time Lemma 2.17. Putting the two bounds together in (2.6) we have:

$$\|B\|^2 \leqslant 4\|B_{\succ}\|^2 \leqslant \Big(8\nu + \frac{12\nu^2}{(1-\delta)(1-2\delta)}\Big)\Big(\frac{\log N}{\log\log N}\Big)^2 + \frac{6\nu}{1-\delta}\frac{\log N}{\log\log N}\|B\|\,.$$

After solving a quadratic equation, it implies that for some constant $C_\nu > 0$, we have with ν-high probability

$$\|B\| \leqslant C_\nu \frac{\log N}{\log\log N}$$

Hence, we have finally that there exists a constant $C_\nu > 0$ such that with ν-high probability

$$\|\tilde{A}_{\mathrm{h}}\|^2 \leqslant C_\nu \frac{\log N}{\log\log N}\,.$$

Let us now consider $\tilde{A}_{\mathrm{i}}$. We have

$$\Big\|\tilde{A}_{\mathrm{i}}\Big\|^2 = \max_{\|\mathbf{u}\|=1} \sum_{x_1,x_2\in\mathcal{V}_\nu^{(\mathrm{i})}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle u_{x_1}u_{x_2}\,.$$

Young's inequality followed by a use of Lemma 2.21 then yields:

$$\Big\|\tilde{A}_{\mathrm{i}}\Big\|^2 \leqslant \max_{\|\mathbf{u}\|=1} \sum_{x_1,x_2\in\mathcal{V}_\nu^{(\mathrm{i})}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)}\Big\rangle u_{x_1}^2 \leqslant 2\nu\frac{\log N}{\log\log N}$$

with ν-high probability.

It remains to bound the norm of $\tilde{A}_{\mathrm{l}}$. First, notice that this is the adjacency matrix of a graph $\tilde{G}_l$. Let us introduce the graph G_{l}, the sub-graph of G containing only edges between vertices x and y whose weights satisfy $w_x, w_y \leqslant 4\xi_\nu$. Denote by A_{l} the adjacency matrix of G_{l}. We also introduce $\hat{A}_{\mathrm{l}}$, the adjacency matrix of the sub-graph $\hat{G}_{\mathrm{l}}$ of G_{l} in which we only kept edges between vertices $x, y \in \mathcal{V}_\nu^{(\mathrm{l})}$. We see that we have the inclusion of graphs $\tilde{G}_{\mathrm{l}} \subset \hat{G}_{\mathrm{l}} \subset G_{\mathrm{l}}$. It implies in particular

$$\|\tilde{A}_{\mathrm{l}}\| \leqslant \|\hat{A}_{\mathrm{l}}\| \leqslant \|A_{\mathrm{l}}\|.$$

We are going to bound $\|\hat{A}_{\mathrm{l}}\|$, the bound on $\|\tilde{A}_{\mathrm{l}}\|$ will immediately follow.

To bound $\|\hat{A}_{\mathrm{l}}\|$, we use the results of [BGBK20]. First, we observe that

$$\|\mathbb{E}\hat{A}_{\mathrm{l}}\|^2 \leqslant \max_{\|\mathbf{u}\|=1}\sum_{x_1,x_2,y} p_{x_1y}p_{x_2y}u_{x_1}u_{x_2} \leqslant \max_{\|\mathbf{u}\|=1}\sum_{x_1,x_2}\frac{w_{x_1}w_{x_2}}{m_1N}\frac{m_2}{m_1}u_{x_1}u_{x_2} \leqslant \max_{\|\mathbf{u}\|=1}\sum_{x_1,x_2}\frac{w_{x_2}^2}{m_1N}\frac{m_2}{m_1}u_{x_1}^2 \leqslant \Big(\frac{m_2}{m_1}\Big)^2,$$

where we used Young's inequality and (1.2). It remains to bound $\|\hat{A}_{\mathrm{l}} - \mathbb{E}\hat{A}_{\mathrm{l}}\|$. Set $H = \frac{1}{\sqrt{4\xi}}(\hat{A}_{\mathrm{l}} - \mathbb{E}\hat{A}_{\mathrm{l}})$ and denote by $\hat{B}_{\mathrm{l}}$ and B_{l} the non-backtracking matrices of $\hat{A}_{\mathrm{l}}$ and A_{l} respectively. We refer to [BGBK20] for its definition and only notice that the inclusion of graphs implies $\rho(\hat{B}_{\mathrm{l}}) \leqslant \rho(B_{\mathrm{l}})$, and since

$$\begin{aligned}
\max_{x;w_x\leqslant 4\xi_\nu}\sum_y \mathbb{E}|H_{xy}|^2 &\leqslant \max_x \frac{w_x}{4\xi_\nu} \leqslant 1\\
\max_{x,y;w_x,w_y\leqslant 4\xi_\nu} \mathbb{E}|H_{xy}|^2 &\leqslant \max_{x,y;w_x,w_y\leqslant 4\xi_\nu}\frac{w_xw_y}{m_1N4\xi_\nu} \leqslant \frac{4\xi_\nu}{m_1}\frac{1}{N}\\
\max_{x,y;w_x,w_y\leqslant 4\xi_\nu}|H_{xy}| &\leqslant \frac{1}{\sqrt{4\xi_\nu}}\,,
\end{aligned}$$

we have that [BGBK20, Assumption 2.4] is satisfied. Hence, [BGBK20, Theorem 2.5] implies that with ν-high probability, the spectral radius $\rho(B_{\mathrm{l}})$ of B_{l} is bounded by 2. This implies that

$$\rho(\hat{B}_{\mathrm{l}}) \leqslant 2 \text{ with } \nu\text{-high probability.}$$

Then, introducing the norms

$$\begin{aligned}
\|H\|_{2\to\infty} &:= \max_x \sqrt{\sum_y |H_{xy}|^2} \leqslant \max_x \sqrt{\frac{D_x(1+O(N^{-\varepsilon}))}{4\xi_\nu}} \\
\|H\|_{1\to\infty} &:= \max_{x,y} |H_{xy}| \leqslant 1\,,
\end{aligned}$$

we have by [BGBK20, Theorem 2.2]:

$$\|H\| \leqslant \|H\|_{2\to\infty} f\left(\frac{\rho(\hat{B}_{\mathrm{l}})}{\|H\|_{2\to\infty}}\right) + 7\|H\|_{1\to\infty}\,,$$

where $f(x) = 2$ if $0 \leqslant x \leqslant 1$ and $f(x) = x + 1/x$ if $x > 1$. Since the degrees in $\hat{G}_{\mathrm{l}}$ are bounded by ξ_ν, we get that $\|H\|_{2\to\infty} \leqslant 1$ so that

$$\|\hat{A}_{\mathrm{l}} - \mathbb{E}\hat{A}_{\mathrm{l}}\| \leqslant \begin{cases} 16\sqrt{\xi_\nu} & \text{if } \rho(\hat{B}_{\mathrm{l}}) \leqslant \|H\|_{2\to\infty} \\ 2\sqrt{\xi_\nu}\left(\rho(\hat{B}_{\mathrm{l}}) + \frac{\|H\|_{2\to\infty}^2}{\rho(\hat{B}_{\mathrm{l}})} + 7\right) \leqslant 2\sqrt{\xi_\nu}(2+1+7) & \text{if } \rho(\hat{B}_{\mathrm{l}}) > \|H\|_{2\to\infty}. \end{cases}$$

Hence, we get that

$$\|\hat{A}_{\mathrm{l}} - \mathbb{E}\hat{A}_{\mathrm{l}}\| \leqslant 20\xi\,.$$

This concludes the proof that

$$\|\tilde{A}_{\mathrm{l}}\| = O\Big(\sqrt{\frac{\log N}{\log\log N}}\Big)\,.$$

Putting the three bounds for $\circ \in \{\mathrm{h}, \mathrm{i}, \mathrm{l}\}$ together, we get the result. □

Proof of Lemma 2.20. We are going to bound the quantity

$$\mathcal{P}_{x_1}^{(1)} := \sum_{\substack{x_2,x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_2 \prec x_1, x_3}} \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_2)\setminus S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_3)} \Big\rangle\,.$$

with ν-high probability. By using the definition (2.5) of the sets $S_1^{\mathrm{du}}(x_i), i = 1,2,3$, $\mathcal{P}_{x_1}$ may be rewritten as

$$\begin{aligned}
\mathcal{P}_{x_1}^{(1)} = \sum_{\substack{x_2,x_3 \in \mathcal{V}_\nu^{(\mathrm{h})} \\ x_1,x_2,x_3 \text{ distinct}}} \sum_{\substack{y_1,y_2 \in [N] \\ y_1 \neq y_2}} & \mathbb{1}_{\{x_1 \overset{\mathrm{nc}}{\sim} y_1 \overset{\mathrm{nc}}{\sim} x_2 \overset{\mathrm{nc}}{\sim} y_2 \overset{\mathrm{nc}}{\sim} x_3\}} \mathbb{1}_{\{D_{y_1} \leqslant D_{x_2} \leqslant D_{x_1}, D_{y_2} \leqslant D_{x_2} \leqslant D_{x_3}\}} \\
& \times \mathbb{1}_{\{\exists z_1 \notin \{x_1,x_2\}, y_1 \overset{\mathrm{nc}}{\sim} z_1, D_{x_1} \leqslant D_{z_1}\}} \mathbb{1}_{\{\exists z_2 \notin \{x_2,x_3\}, y_2 \overset{\mathrm{nc}}{\sim} z_2, D_{x_3} \leqslant D_{z_2}\}}\,.
\end{aligned}$$

Note that it is so in particular because if $y_1 \in S_1^{\mathrm{du}}(x_1) \cap S_1^{\mathrm{du}}(x_2)$ and $y_2 \in (S_1^{\mathrm{du}}(x_2) \setminus S_1^{\mathrm{du}}(x_1)) \cap S_1^{\mathrm{du}}(x_3)$, we have necessarily $y_1 \neq y_2$ and x_1, x_2, and x_3 must be distinct: indeed if we had $x_1 = x_3$ there would be a cycle $x_1 \overset{\mathrm{nc}}{\sim} y_1 \overset{\mathrm{nc}}{\sim} x_2 \overset{\mathrm{nc}}{\sim} y_2 \overset{\mathrm{nc}}{\sim} x_3 = x_1$ in G^{nc}, in contradiction with the properties of G^{nc}.

Introduce the notation $\chi = \log\log N$. Remark 1.13 implies that with ν-high probability, we have for every $x \in [N]$:

$$d_x \leqslant \chi(D_x \vee \xi_\nu)\,.$$

We use this fact and discard some unneeded events. We get that with ν-high probability,

$$\mathcal{P}^{(1)}_{x_1} \leqslant \sum_{\substack{x_2,x_3\in[N]\\ x_1,x_2,x_3 \text{ distinct}}} \sum_{\substack{y_1,y_2\in[N]\\ y_1\neq y_2}} \mathbb{1}_{\left\{x_1 \overset{\mathrm{nc}}{\sim} y_1 \overset{\mathrm{nc}}{\sim} x_2 \overset{\mathrm{nc}}{\sim} y_2 \overset{\mathrm{nc}}{\sim} x_3\right\}} \mathbb{1}_{\left\{d_{y_1}\vee\xi_\nu\leqslant\chi D_{x_2},\, d_{y_1}\vee d_{y_2}\vee d_{x_2}\leqslant \chi D_{x_3}\right\}}$$
$$\times\, \mathbb{1}_{\left\{\exists z\neq x_1,x_2,y_1 \overset{\mathrm{nc}}{\sim} z,\, d_{y_1}\vee d_{x_1}\leqslant \chi D_z\right\}}\,.$$

The indicator function on the last line can be further simplified using Remark 1.14. Indeed, if we assume that $d_1 > 8\nu\log N$ and $D_{x_1} \leqslant D_z$, Lemma 1.12 implies that with ν-high probability

$$4\nu\log N < \frac{d_{x_1}}{2} < d_{x_1} - \sqrt{2\nu\log N d_1} \leqslant D_{x_1} \leqslant D_z \leqslant d_z + 2\sqrt{\nu\log N\left(d_z \vee \frac{4\nu}{9}\log N\right)} \leqslant 4d_z \vee \frac{16\nu}{9}\log N\,.$$

The last inequality has been discussed in Remark 1.14. This implies that with ν-high probability, we have $D_z \leqslant 4d_z$. We thus have with ν-high probability

$$\mathcal{P}^{(1)}_{x_1} \leqslant \sum_{\substack{x_2,x_3\in[N]\\ x_1,x_2,x_3 \text{ distinct}}} \sum_{\substack{y_1,y_2\in[N]\\ y_1\neq y_2}} \mathbb{1}_{\left\{x_1 \overset{\mathrm{nc}}{\sim} y_1 \overset{\mathrm{nc}}{\sim} x_2 \overset{\mathrm{nc}}{\sim} y_2 \overset{\mathrm{nc}}{\sim} x_3\right\}} \mathbb{1}_{\left\{d_{y_1}\vee\xi_\nu\leqslant\chi D_{x_2},\, d_{y_1}\vee d_{y_2}\vee d_{x_2}\leqslant \chi D_{x_3}\right\}}$$
$$\times \left(\mathbb{1}_{\left\{d_1\leqslant 8\nu\log N\right\}} + \mathbb{1}_{\left\{d_1 > 8\nu\log N\right\}} \mathbb{1}_{\left\{\exists z\neq x_1,y_1 \overset{\mathrm{nc}}{\sim} z,\, d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\right\}}\right).$$

Recall that by Proposition 2.7, there exists $C_\nu > 0$ such that with ν-high probability, we have for all $x \in [N]$, $D_x \leqslant D^{\mathrm{nc}}_x + C_\nu$. We thus get that with ν-high probability

$$\mathcal{P}^{(1)}_{x_1} \leqslant \sum_{\substack{x_2,x_3\in[N]\\ x_1,x_2,x_3 \text{ distinct}}} \sum_{\substack{y_1,y_2\in[N]\\ y_1\neq y_2}} \mathbb{1}_{\left\{x_1 \overset{\mathrm{nc}}{\sim} y_1 \overset{\mathrm{nc}}{\sim} x_2 \overset{\mathrm{nc}}{\sim} y_2 \overset{\mathrm{nc}}{\sim} x_3\right\}} \mathbb{1}_{\left\{d_{y_1}\vee\xi_\nu\leqslant\chi (D^{\mathrm{nc}}_{x_2}+C_\nu),\, d_{y_1}\vee d_{y_2}\vee d_{x_2}\leqslant \chi (D^{\mathrm{nc}}_{x_3}+C_\nu)\right\}}$$
$$\times \left(\mathbb{1}_{\left\{d_1\leqslant 8\nu\log N\right\}} + \mathbb{1}_{\left\{d_1 > 8\nu\log N\right\}} \mathbb{1}_{\left\{\exists z\neq x_1,y_1 \overset{\mathrm{nc}}{\sim} z,\, d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\right\}}\right).$$

We can now use the coupling with a rooted tree introduced in Section 2.3. To make notation lighter we write $\mathcal{T} := \mathcal{T}_{x_1}$. Proposition 2.15 implies that for all $\tilde{x} \in B^{\mathrm{nc}}_r(x)$,

$$D^{\mathrm{nc}}_{\tilde{x}} \leqslant D^{\mathcal{T}}_{\tilde{x}} \leqslant D^{\mathcal{T}\uparrow}_{\tilde{x}} + 1\,,$$

so up to replacing C_ν by $C_\nu + 1$ we have with ν-high probability:

$$\mathcal{P}^{(1)}_{x_1} \leqslant \sum_{\substack{x_2,x_3\in V_{x_1}\\ x_1,x_2,x_3 \text{ distinct}}} \sum_{\substack{y_1,y_2\in V_{x_1}\\ y_1\neq y_2}} \mathbb{1}_{\left\{x_1 \overset{\mathcal{T}}{\sim} y_1 \overset{\mathcal{T}}{\sim} x_2 \overset{\mathcal{T}}{\sim} y_2 \overset{\mathcal{T}}{\sim} x_3\right\}} \mathbb{1}_{\left\{d_{y_1}\vee\xi_\nu\leqslant\chi (D^{\mathcal{T}\uparrow}_{x_2}+C_\nu),\, d_{y_1}\vee d_{y_2}\vee d_{x_2}\leqslant \chi (D^{\mathcal{T}\uparrow}_{x_3}+C_\nu)\right\}}$$
$$\times \left(\mathbb{1}_{\left\{d_1\leqslant 8\nu\log N\right\}} + \mathbb{1}_{\left\{d_1 > 8\nu\log N\right\}} \mathbb{1}_{\left\{\exists z\neq x_1,y_1 \overset{\mathcal{T}}{\sim} z,\, d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\right\}}\right).$$

In the sequel, we are going to use the filtration $(\mathcal{F}_k(x_1))_{k\geqslant 1}$ defined in (2.2).

Bounding $\mathcal{P}^{(1)}_{x_1}$ is based on a Chernoff-type bound. We have for all $\lambda > 0$ and $k \geqslant 0$:

$$\mathbb{P}\left(\mathcal{P}^{(1)}_{x_1} \geqslant k\right) \leqslant \mathrm{e}^{-\lambda k}\mathbb{E}\left[\exp(\lambda\mathcal{P}^{(1)}_{x_1})\right].$$

Let us bound the Laplace transform in the right-hand term. We write

$$\mathcal{P}^{(1)}_{x_1} = \sum_{y_1 \in V_x} X^{(1)}_{x_1 y_1} ,$$

with

$$\begin{aligned}
X^{(4)}_{x_1y_1x_2y_2x_3} &:= \mathbb{1}_{\left\{y_2 \overset{\mathcal{T}}{\sim} x_3, d_{x_2} \vee d_{y_1} \vee d_{y_2} \leqslant \chi\left(D^{\mathcal{T}\uparrow}_{x_3} + C_\nu\right)\right\}} \\
X^{(3)}_{x_1y_1x_2y_2} &:= \mathbb{1}_{\left\{x_2 \overset{\mathcal{T}}{\sim} y_2\right\}} \sum_{x_3 \neq x_1, x_2} X^{(4)}_{x_1y_1x_2y_2x_3} \\
X^{(2)}_{x_1y_1x_2} &:= \mathbb{1}_{\left\{y_1 \overset{\mathcal{T}}{\sim} x_2\right\}} \sum_{y_2 \neq y_1} X^{(3)}_{x_1y_1x_2y_2} \\
X^{(1)}_{x_1y_1} &:= \mathbb{1}_{\left\{x_1 \overset{\mathcal{T}}{\sim} y_1\right\}} \left(\mathbb{1}_{\left\{d_1 \leqslant 8\nu \log N\right\}} + \mathbb{1}_{\left\{d_1 > 8\nu \log N\right\}} \mathbb{1}_{\left\{\exists z \neq x_1, y_1 \overset{\mathcal{T}}{\sim} z, d_{y_1} \vee d_{x_1} \leqslant 4\chi d_z\right\}}\right) \sum_{x_2 \neq x_1} X^{(2)}_{x_1y_1x_2} .
\end{aligned}$$

We will consider these different random variables starting from $X^{(4)}_{x_1y_1x_2y_2x_3}$ and use the independence properties of $\mathcal{T}$ to estimate the Laplace transform of each of those. In this computation, the tree structure of $\mathcal{T}$ is critical.

Let us fix y_1, x_2, y_2, x_3 for now, and consider $X^{(4)}_{x_1y_1x_2y_2x_3}$. Conditionally to $\mathcal{F}_4(x_1)$, we have

$$\mathbb{E}\left[\mathrm{e}^{\lambda X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_4(x_1)\right] = 1 + \mathbb{P}\left(y_2 \overset{\mathcal{T}}{\sim} x_3, d_{x_2} \vee d_{y_1} \vee d_{y_2} \leqslant \chi(D^{\mathcal{T}\uparrow}_{x_3} + C_\nu) \mid \mathcal{F}_4(x_1)\right)\left(\mathrm{e}^{\lambda} - 1\right).$$

Markov's inequality gives us

$$\mathbb{E}\left[\mathrm{e}^{\lambda X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_4(x_1)\right] = 1 + \mathbb{1}_{\left\{y_2 \overset{\mathcal{T}}{\sim} x_3\right\}} \chi \frac{d_{x_3} + C_\nu}{d_{x_2} \vee d_{y_1} \vee d_{y_2}} \left(\mathrm{e}^{\lambda} - 1\right).$$

We choose for λ the small value $\lambda = \frac{\chi^2}{\log N}$. Hence, we have for $1 < c < 2$ that for N big enough:

$$\mathrm{e}^{\lambda} - 1 \leqslant c \frac{\chi^2}{\log N}.$$

This means that

$$\mathbb{E}\left[\mathrm{e}^{\lambda X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_4(x_1)\right] \leqslant 1 + \mathbb{1}_{\left\{y_2 \overset{\mathcal{T}}{\sim} x_3\right\}} c \frac{\chi^3}{\log N} \frac{d_{x_3} + C_\nu}{d_{x_2} \vee d_{y_1} \vee d_{y_2}}.$$

Still keeping x_2, y_1, y_2 fixed, we take the product on $x_3 \neq x_1, x_2$ and condition on $\mathcal{F}_3(x_1)$. The conditional independence property of $\mathcal{T}$ gives us

$$\mathbb{E}\left[\mathrm{e}^{\lambda \sum_{x_3 \neq x_1, x_2} X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_3(x_1)\right] \leqslant \prod_{x_3 \neq x_1, x_2} \left(1 + c p_{y_2 x_3} \frac{\chi^3}{\log N} \frac{d_{x_3} + C_\nu}{d_{x_2} \vee d_{y_1} \vee d_{y_2}}\right).$$

Using (1.4), and $\ln(1+u) \leqslant u$, we get

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{\lambda \sum_{x_3 \neq x_1, x_2} X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_3(x_1)\right] &\leqslant \prod_{x_3 \neq x_1, x_2} \left(1 + c \frac{\chi^3}{m_1 N \log N} \frac{w_{y_2} w_{x_3} (w_{x_3} + C_\nu)}{d_{x_2} \vee d_{y_1} \vee d_{y_2}}\right) \\
&\leqslant \exp\left(\sum_{x_3 \neq x_1, x_2} c \frac{\chi^3}{m_1 N \log N} \frac{w_{y_2} w_{x_3} (w_{x_3} + C_\nu)}{d_{x_2} \vee d_{y_1} \vee d_{y_2}}\right).
\end{aligned}$$

We can make the ratio of empirical moments m_2/m_1 appear using (1.2) and get with Assumption 1.2

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \sum_{x_3\neq x_1,x_2} X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_3(x_1)\Big] \leqslant \exp\bigg(\frac{c\chi^3}{\log N}\frac{m_2+C_\nu m_1}{m_1}\frac{w_{y_2}}{d_{x_2}\vee d_{y_1}\vee d_{y_2}}\bigg)$$
$$\leqslant \exp\bigg(\frac{c^2\chi^3}{(\log N)^{1-\delta}}\frac{w_{y_2}}{w_{x_2}\vee w_{y_1}\vee w_{y_2}}\bigg).$$

We could replace the average degrees $d_{x_2}\vee d_{y_1}\vee d_{y_2}$ by the weights at the cost of adding a small constant c in the product. Finally, we can linearize the exponential again and get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \sum_{x_3\neq x_1,x_2} X^{(4)}_{x_1y_1x_2y_2x_3}} \mid \mathcal{F}_3(x_1)\Big] \leqslant 1+\frac{c^3\chi^3}{(\log N)^{1-\delta}}\frac{w_{y_2}}{w_{x_2}\vee w_{y_1}\vee w_{y_2}}. \tag{2.7}$$

We are going to apply a similar argument to the one we used on $X^{(4)}_{x_1y_1x_2y_2x_3}$ on $X^{(3)}_{x_1y_1x_2y_2}$. We fix y_1, x_2, y_2. We have

$$\mathrm{e}^{\lambda X^{(3)}_{x_1y_1x_2y_2}} = 1+\mathbb{1}_{\{x_2\overset{\mathcal{T}}{\sim}y_2\}}\Big(\mathrm{e}^{\lambda \sum_{x_3\neq x_1,x_2} X^{(4)}_{x_1y_1x_2y_2x_3}}-1\Big).$$

Taking the expectation conditional to $\mathcal{F}_3(x_1)$ and using (2.7), we get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(3)}_{x_1y_1x_2y_2}} \mid \mathcal{F}_3(x_1)\Big] \leqslant 1+\mathbb{1}_{\{x_2\overset{\mathcal{T}}{\sim}y_2\}}\frac{c^3\chi^3}{(\log N)^{1-\delta}}\frac{w_{y_2}}{w_{x_2}\vee w_{y_1}\vee w_{y_2}}.$$

Again, we take the product on $y_2\neq y_1$ and the expectation conditionally on $\mathcal{F}_2(x_1)$. We get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \sum_{y_2\neq y_1} X^{(3)}_{x_1y_1x_2y_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant \prod_{y_2\neq y_1}\bigg(1+\frac{c^3\chi^3}{(\log N)^{1-\delta}}p_{x_2y_2}\frac{w_{y_2}}{w_{x_2}\vee w_{y_1}\vee w_{y_2}}\bigg).$$

Using $\log(1+u)\leqslant u$, (1.4), and (1.2), we get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \sum_{y_2\neq y_1} X^{(3)}_{x_1y_1x_2y_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant \exp\Big(\sum_{y_2\neq y_1}\frac{c^3\chi^3}{(\log N)^{1-\delta}}p_{x_2y_2}\frac{w_{y_2}}{w_{x_2}\vee w_{y_1}\vee w_{y_2}}\Big)$$
$$\leqslant \exp\Big(\sum_{y_2\neq y_1}\frac{c^3\chi^3}{m_1N(\log N)^{1-\delta}}\frac{w_{x_2}w_{y_2}^2}{w_{x_2}\vee w_{y_1}\vee w_{y_2}}\Big)$$
$$\leqslant \exp\Big(\frac{c^3\chi^3}{(\log N)^{1-2\delta}}\frac{w_{x_2}}{w_{x_2}\vee w_{y_1}}\Big).$$

By linearizing the exponential, we get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \sum_{y_2\neq y_1} X^{(3)}_{x_1y_1x_2y_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant 1+\frac{c^3\chi^3}{(\log N)^{1-2\delta}}\frac{w_{x_2}}{w_{x_2}\vee w_{y_1}}.$$

Since we have

$$\mathrm{e}^{\lambda X^{(2)}_{x_1y_1x_2}} = 1+\mathbb{1}_{\{y_1\overset{\mathcal{T}}{\sim}x_2\}}\Big(\mathrm{e}^{\lambda \sum_{y_2\neq y_1} X^{(3)}_{x_1y_1x_2y_2}}-1\Big),$$

We finally get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(2)}_{x_1y_1x_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant 1+\mathbb{1}_{\{y_1\overset{\mathcal{T}}{\sim}x_2\}}\frac{c^3\chi^3}{(\log N)^{1-2\delta}}\frac{w_{x_2}}{w_{x_2}\vee w_{y_1}}. \tag{2.8}$$

We may take the product on $x_2 \neq x_1$ and the expectation conditionally on $\mathcal{F}_1(x_1)$. We get using the conditional independence of the $\mathbb{1}_{\{y_1 \overset{\mathcal{T}}{\sim} x_2\}}$:

$$\mathbb{E}\Big[\mathrm{e}^{\sum_{x_2\neq x_1} \lambda X^{(2)}_{x_1y_1x_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant \prod_{x_2\neq x_1} \Big(1 + p_{y_1x_2}\frac{c^3\chi^3}{(\log N)^{1-2\delta}}\frac{w_{x_2}}{w_{x_2}\vee w_{y_1}}\Big).$$

Using again $\log(1+u)\leqslant u$, (1.4), and (1.2), we get

$$\mathbb{E}\Big[\mathrm{e}^{\sum_{x_2\neq x_1} \lambda X^{(2)}_{x_1y_1x_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant \exp\Big(\frac{c^3\chi^3}{(\log N)^{1-3\delta}}\Big).$$

Linearizing the exponential, we get

$$\mathbb{E}\Big[\mathrm{e}^{\sum_{x_2\neq x_1} \lambda X^{(2)}_{x_1y_1x_2}} \mid \mathcal{F}_2(x_1)\Big] \leqslant 1 + \frac{c^3\chi^3}{(\log N)^{1-3\delta}}. \tag{2.9}$$

The computation of the Laplace transform of $X^{(1)}_{x_1y_1}$ is more involved than the preceding ones. By definition, we have for x_1 and y_1 fixed that if $d_1\leqslant 8\nu\log N$

$$\mathrm{e}^{\lambda X^{(1)}_{x_1y_1}} = 1 + \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}}\Big(\mathrm{e}^{\lambda\sum_{x_2\neq x_1} X^{(2)}_{x_1y_1x_2}} - 1\Big).$$

In this case, it is straightforward to compute the expectation conditionally on $\mathcal{F}_1(x_1)$ using (2.9):

$$\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(1)}_{x_1y_1}} \mid \mathcal{F}_1(x_1)\Big] \leqslant 1 + \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}}\frac{c^3\chi^3}{(\log N)^{1-3\delta}}.$$

The more involved case is when $d_{x_1} > 8\nu\log N$. In this case:

$$\mathrm{e}^{\lambda X^{(1)}_{x_1y_1}} = 1 + \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}}\mathbb{1}_{\{\exists z\neq x_1, y_1 \overset{\mathcal{T}}{\sim} z, d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\}}\Big(\mathrm{e}^{\lambda\sum_{x_2\neq x_1} X^{(2)}_{x_1y_1x_2}} - 1\Big).$$

At this point, we notice the union bound

$$\mathbb{1}_{\{\exists z\neq x_1, y_1 \overset{\mathcal{T}}{\sim} z, d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\}} \leqslant \sum_{z\neq x_1}\mathbb{1}_{\{y_1 \overset{\mathcal{T}}{\sim} z, d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\}}.$$

Together with $\mathbb{1}_{\{d_{y_1}\vee d_{x_1}\leqslant 4\chi d_z\}} \leqslant \frac{4\chi d_z}{d_{y_1}\vee d_{x_1}}$, this allows the bound

$$\mathrm{e}^{\lambda X^{(1)}_{x_1y_1}} \leqslant 1 + \sum_{z\neq x_1}\mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1 \overset{\mathcal{T}}{\sim} z\}}\frac{4\chi d_z}{d_{y_1}\vee d_{x_1}}\Big(\mathrm{e}^{\lambda\sum_{x_2\neq x_1} X^{(2)}_{x_1y_1x_2}} - 1\Big).$$

We first take the expectation conditionally on $\mathcal{F}_2(x_1)$ and use (2.8)

$$\begin{aligned}\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(1)}_{x_1y_1}} \mid \mathcal{F}_2(x_1)\Big] &\leqslant 1 + \sum_{z\neq x_1}\mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1 \overset{\mathcal{T}}{\sim} z\}}\frac{4\chi d_z}{d_{y_1}\vee d_{x_1}}\Big(\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(2)}_{x_1y_1z}+\lambda\sum_{x_2\neq x_1,z} X^{(2)}_{x_1y_1x_2}} \mid \mathcal{F}_2(x_1)\Big] - 1\Big)\\ &\leqslant 1 + \sum_{z\neq x_1}\mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1 \overset{\mathcal{T}}{\sim} z\}}\frac{4\chi d_z}{d_{y_1}\vee d_{x_1}}\Big(\mathbb{E}\Big[\mathrm{e}^{\lambda\sum_{x_2\neq x_1,z} X^{(2)}_{x_1y_1x_2}}(1+\frac{c^3\chi^3}{(\log N)^{1-2\delta}}) \mid \mathcal{F}_2(x_1)\Big] - 1\Big).\end{aligned}$$

We can now take the expectation conditionally on $\mathcal{F}_1(x_1)$ and use the conditional independence and (2.9) to get

$$\begin{aligned}
&\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(1)}_{x_1 y_1}} \mid \mathcal{F}_1(x_2)\Big] \\
&\qquad \leqslant 1 + \sum_{z \neq x_1} \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}} p_{y_1 z} \frac{4\chi d_z}{d_{y_1} \vee d_{x_1}} \Big((1 + \frac{c^3\chi^3}{(\log N)^{1-3\delta}})(1 + \frac{c^3\chi^3}{(\log N)^{1-2\delta}}) \mid \mathcal{F}_1(x_1)\Big] - 1\Big) \\
&\qquad \leqslant 1 + \sum_{z \neq x_1} \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}} p_{y_1 z} \frac{d_z}{d_{y_1} \vee d_{x_1}} \frac{4c^4\chi^4}{(\log N)^{1-3\delta}} \,.
\end{aligned}$$

We use (1.4) and (1.2) to get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(1)}_{x_1 y_1}} \mid \mathcal{F}_1(x_1)\Big] \leqslant 1 + \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}} \frac{4c^4\chi^4}{(\log N)^{1-4\delta}} \frac{w_{y_1}}{d_{y_1} \vee d_{x_1}} \,.$$

Up to adding a factor c, we can replace $d_{y_1} \vee d_{x_1}$ by $w_{y_1} \vee w_{x_1}$. We finally have

$$\mathbb{E}\Big[\mathrm{e}^{\lambda X^{(1)}_{x_1 y_1}} \mid \mathcal{F}_1(x_1)\Big] \leqslant 1 + \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y_1\}} \frac{4c^5\chi^4}{(\log N)^{1-4\delta}} \frac{w_{y_1}}{w_{y_1} \vee w_{x_1}} \,. \tag{2.10}$$

It remains to take the product on y_1 and the expectation conditional to $\mathcal{F}_1(x_1)$. We consider first the case $d_{x_1} \leqslant 8\nu \log N$. We have by independence

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}^{(1)}_{x_1}}\Big] \leqslant \prod_{y_1} \Big(1 + p_{x_1 y_1} \frac{c^3\chi^3}{(\log N)^{1-3\delta}}\Big) \,.$$

Using as before (1.4) and (1.2) we get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}^{(1)}_{x_1}}\Big] \leqslant \exp\Big(w_{x_1} \frac{c^3\chi^3}{(\log N)^{1-3\delta}}\Big) \leqslant \exp\Big(c^4 \nu \chi^3 (\log N)^{3\delta}\Big) \,.$$

In the case $d_{x_1} > 8\nu \log N$, we have by independence,

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}^{(1)}_{x_1}}\Big] \leqslant \prod_{y_1} \Big(1 + p_{x_1 y_1} \frac{4c^5\chi^4}{(\log N)^{1-4\delta}} \frac{w_{y_1}}{w_{y_1} \vee w_{x_1}}\Big) \,.$$

We proceed as in the previous case and get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}^{(1)}_{x_1}}\Big] \leqslant \exp\Big(\frac{4c^5\chi^4}{(\log N)^{1-5\delta}}\Big) \,.$$

We have shown in both cases that

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}^{(1)}_{x_1}}\Big] \ll N \,.$$

This implies

$$\mathbb{P}\bigg(\mathcal{P}^{(1)}_{x_1} \geqslant 2\nu \Big(\frac{\log N}{\log \log N}\Big)^2\bigg) = O(N^{-\nu}) \,.$$

□

Proof of Lemma 2.21. Let $x_1 \in \mathcal{V}_\nu^{(\mathrm{i})}$. We are going to bound the quantity

$$\mathcal{P}_{x_1}^{(2)} := \sum_{\substack{x_2 \in \mathcal{V}_\nu^{(\mathrm{i})} \\ x_2 \neq x_1}} \left\langle \mathbf{1}_{S_1^{\mathrm{du}}(x_1)}, \mathbf{1}_{S_1^{\mathrm{du}}(x_2)} \right\rangle .$$

Let us start by noticing that by Lemma 1.12, we have the following property with ν-high probability: for all $z \in \mathcal{V}_\nu^{(\mathrm{l})} \cup \mathcal{V}_\nu^{(\mathrm{i})}$ then $w_z < 4\nu \log N$.

We thus observe that with ν-high probability, we have $w_{x_1} < 4\nu \log N$ and

$$\mathcal{P}_{x_1}^{(2)} \leqslant \sum_{\substack{x_2 \in [N] \\ x_2 \neq x_1 \\ 4\xi \leqslant w_{x_2} \leqslant 4\nu \log N}} \sum_{\substack{y \in [N] \\ w_y \leqslant 4\nu \log N}} \mathbb{1}_{\left\{x \overset{\mathrm{nc}}{\sim} y \overset{\mathrm{nc}}{\sim} z\right\}} \mathbb{1}_{\left\{D_z < \xi\right\}} .$$

Let us relate the right-hand side to the rooted tree $\check{\mathcal{T}} := \check{\mathcal{T}}_{x_1}$ of Section 2.3. We have by Proposition 2.15 and Corollary 2.16 that there exists a constant $C_\nu > 0$ such that

$$\begin{aligned} \mathcal{P}_{x_1}^{(2)} &\leqslant \sum_{\substack{x_2 \in [N] \\ x_2 \neq x_1 \\ 4\xi \leqslant w_{x_2} \leqslant 4\nu \log N}} \sum_{\substack{y \in [N] \\ w_y \leqslant 4\nu \log N}} \mathbb{1}_{\left\{x \overset{\mathrm{nc}}{\sim} y \overset{\mathrm{nc}}{\sim} z\right\}} \mathbb{1}_{\left\{D_z^{\mathrm{nc}} < \xi\right\}} \\ &\leqslant \sum_{\substack{x_2 \in V_{x_1} \\ x_2 \neq x_1 \\ 4\xi \leqslant w_{x_2} \leqslant 4\nu \log N}} \sum_{\substack{y \in V_{x_1} \\ w_y \leqslant 4\nu \log N}} \mathbb{1}_{\left\{x \overset{\check{Tree}}{\sim} y \overset{\mathcal{T}}{\sim} z\right\}} \mathbb{1}_{\left\{D_z^{\check{Tree}\uparrow} - C_\nu < \xi\right\}} , \end{aligned}$$

with ν-high probability.

Before computing $\mathbb{E}[\exp(\lambda \mathcal{P}_{x_1}^{(2)})]$, let us now compute $\mathbb{P}\Big(D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\Big)$ using the version of Bennett's inequality given in [Hof16, Theorem 2.21]. We have

$$\sum_{z \in V_{x_1}} \mathbb{E}\Big[\mathbb{1}_{\left\{x \overset{\mathcal{T}}{\sim} z\right\}}\Big] \leqslant w_x .$$

Hence,

$$\mathbb{P}\Big(D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\Big) = \mathbb{P}\Big(D_{x_2}^{\check{\mathcal{T}}\uparrow} - w_{x_2} \leqslant -(w_{x_2} - \xi - C_\nu)\Big) \leqslant \exp\Big(-\frac{(w_{x_2} - \xi - C_\nu)^2}{2w_{x_2}}\Big) .$$

Hence, we get that

$$\mathbb{P}\Big(D_{x_2}^{\check{\mathcal{T}}\uparrow} + C_\nu \leqslant \xi\Big) \leqslant \exp(-\xi(1 + o(1))) .$$

For us it suffices that for all $\beta > 0$ we have for N big enough

$$\mathbb{P}\Big(D_{x_2}^{\check{\mathcal{T}}\uparrow} + C_\nu \leqslant \xi\Big) \leqslant \frac{1}{(\log N)^\beta} . \tag{2.11}$$

We have that for x_1 and y fixed

$$\exp\Big(\lambda \sum_{\substack{x_2 \neq x_1 \\ 4\xi_\nu \leqslant w_{x_2} \leqslant 4\nu \log N}} \mathbb{1}_{\left\{y \overset{\mathcal{T}}{\sim} x_2\right\}} \mathbb{1}_{\left\{D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\right\}}\Big) = \prod_{\substack{x_2 \neq x_1 \\ 4\xi_\nu \leqslant w_{x_2} \leqslant 4\nu \log N}} \left(1 + \mathbb{1}_{\left\{y \overset{\mathcal{T}}{\sim} x_2\right\}} \mathbb{1}_{\left\{D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\right\}} \left(\mathrm{e}^{\lambda} - 1\right)\right) .$$

Taking the expectation conditionally on $\mathcal{F}_2(x_1)$, using the conditional independence, and injecting (2.11) yields

$$\mathbb{E}\Big[\exp\Big(\lambda \sum_{x_2 \neq x_1} \mathbb{1}_{\{y \overset{\mathcal{T}}{\sim} x_2\}} \mathbb{1}_{\{D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\}}\Big) \mid \mathcal{F}_2(x_1)\Big] \leqslant \prod_{x_2 \neq x_1} \Big(1 + \mathbb{1}_{\{y \overset{\mathcal{T}}{\sim} x_2\}} \frac{1}{(\log N)^\beta} \big(\mathrm{e}^\lambda - 1\big)\Big),$$

where the sum on the left-hand size and product on the right-hand side are over $x_2 \in V_{x_1}$ satisfying $x_2 \neq x_1$ and $4\xi \leqslant w_{x_2} \leqslant 4\nu \log N$.

We choose $\lambda = \log \log N$ and take the expectation conditionally on $\mathcal{F}_1(x_1)$:

$$\mathbb{E}\Big[\exp\Big(\lambda \sum_{x_2 \neq x_1} \mathbb{1}_{\{y \overset{\mathcal{T}}{\sim} x_2\}} \mathbb{1}_{\{D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\}}\Big) \mid \mathcal{F}_1(x_1)\Big] \leqslant \prod_{x_2 \neq x_1} \Big(1 + p_{y x_2} \frac{1}{(\log N)^{\beta - 1}}\Big).$$

Using $\log(1+u) \leqslant u$ and (1.4), we get

$$\mathbb{E}\Big[\exp\Big(\lambda \sum_{x_2 \neq x_1} \mathbb{1}_{\{y \overset{\mathcal{T}}{\sim} x_2\}} \mathbb{1}_{\{D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\}}\Big) \mid \mathcal{F}_1(x_1)\Big] \leqslant \exp\Big(\frac{w_y}{(\log N)^{\beta-1}}\Big).$$

Assuming $w_y \leqslant 4\nu \log N$ and $\beta > 2$, we can linearize the exponential and get

$$\mathbb{E}\Big[\exp\Big(\lambda \sum_{x_2 \neq x_1} \mathbb{1}_{\{y \overset{\mathcal{T}}{\sim} x_2\}} \mathbb{1}_{\{D_{x_2}^{\check{\mathcal{T}}\uparrow} - C_\nu \leqslant \xi\}}\Big) \mid \mathcal{F}_1(x_1)\Big] \leqslant 1 + c \frac{w_y}{(\log N)^{\beta-1}}$$

for $1 < c < 2$. We thus have

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}_{x_1}^{(2)}}\Big] \leqslant \mathbb{E}\Big[\prod_y \Big(1 + \mathbb{1}_{\{x_1 \overset{\mathcal{T}}{\sim} y\}} c \frac{w_y}{(\log N)^{\beta-1}}\Big)\Big].$$

By independence, we get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}_{x_1}^{(2)}}\Big] \leqslant \mathbb{E}\Big[\prod_y \Big(1 + p_{x_1 y} c \frac{w_y}{(\log N)^{\beta-1}}\Big)\Big].$$

Using $\log(1+u) \leqslant u$, (1.4), and Assumption 1.2, we get

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}_{x_1}^{(2)}}\Big] \leqslant \exp\Big(c \frac{w_{x_1}}{(\log N)^{\beta-1-\delta}}\Big).$$

Since we assumed that $w_{x_1} \leqslant 4\nu \log N$, we have for $\beta > 3$,

$$\mathbb{E}\Big[\mathrm{e}^{\lambda \mathcal{P}_{x_1}^{(2)}}\Big] \leqslant \exp\Big(\frac{4c\nu}{(\log N)^{\beta-2-\delta}}\Big) \ll N.$$

Chernoff's bound implies

$$\mathbb{P}\Big(\mathcal{P}_{x_1}^{(2)} \leqslant 2\nu \frac{\log N}{\log \log N}\Big) = O(N^{-\nu}),$$

which is the wanted result. □

3. Finding the eigenvectors

In this section, we construct a family of orthonormal vectors that will be close to eigenvectors of A. The main intuition to show that the eigenvectors associated to the greatest eigenvalues are (semi-)localized, is that these eigenvectors are probably close to the vectors

$$\mathbf{v}_\sigma(x) = \frac{1}{\sqrt{2}}\left(\mathbf{1}_x + \frac{\sigma}{\sqrt{D_x^{\mathrm{p}-}}}\mathbf{1}_{S_1^{\mathrm{p}-}(x)}\right),$$

where $x \in [N]$ and $\sigma \in \{\pm 1\}$. This particular choice is motivated by the fact that such vectors are eigenvectors of the adjacency matrix of a star of degree $D_x^{\mathrm{p}-}$, i.e. a tree with one vertex connected to $D_x^{\mathrm{p}-}$ leaves.

Notice that these vectors are normalized, but in the pruned graph G^{p}, we have (recall Definition 1.15)

$$\langle \mathbf{v}_\rho(x), \mathbf{v}_\sigma(y)\rangle = \frac{\delta_{xy}}{2}(1+\rho\sigma) + \frac{\mathbb{1}_{\{x\sim y\}}}{2}\left(\frac{\rho\mathbb{1}_{\{x\prec y\}}}{\sqrt{D_x^{\mathrm{p}-}}} + \frac{\sigma\mathbb{1}_{\{x\succ y\}}}{\sqrt{D_y^{\mathrm{p}-}}}\right),$$

that is, we only have $\langle \mathbf{v}_-(x), \mathbf{v}_+(x)\rangle = 0$ for all $x \in [N]$, but in general $\langle \mathbf{v}_\rho(x), \mathbf{v}_\sigma(y)\rangle$ is not zero for $x \sim y$.

We are going to define a family of pseudo-eigenvectors that are supported near the vertices in $\mathcal{V}_\nu^{(\mathrm{h})}$, with the greatest degrees in the graph G.

Remark 3.1. In the pruned graph G^{p}, for each vertex x, there is at most one vertex in $S_1^+(x)$. Indeed, if it were not the case, i.e. if there existed two distinct vertices $y, z \in S_1^+(x)$, then one of (y, x, z) or (z, x, y) would be a down-up path. There are no such paths in the pruned graph.

This remark allows the following definition.

Definition 3.2. Let $x \in [N]$. The unique element $y \in S_1^{\mathrm{p}+}(x)$, if it exists, is called the *parent* of x, and denoted by $\hat{x}$. Conversely, the *children* of x are the vertices in $S_1^{\mathrm{p}-}(x)$. The elements of $S_1^{\mathrm{p}-}(\hat{x})$ are called the *siblings* of x. The set of siblings of a vertex x is denoted by $\mathrm{Sib}(x)$ and is empty if x has no parent. The set of smaller siblings is

$$\#\,\mathrm{Sib}^-(x) = \mathrm{Sib}(x) \cap \{y \in [N] : y \prec x\}.$$

We use the following conventions.

1. $\frac{\mathbf{1}_\emptyset}{0} = 0$,
2. Any term where the symbol $\hat{x}$ appears is 0 if the vertex x has no parent.

The orthonormal family we shall use is defined in the following proposition.

Proposition 3.3. *For all $x \in \mathcal{V}_\nu^{(\mathrm{h})}$, $\sigma \in \{\pm 1\}$, define*

$$Z_x = 2 + 2/\#\,\mathrm{Sib}^-(x),$$

and

$$\mathbf{u}_\sigma(x) = \frac{1}{\sqrt{Z_x}}\left(\mathbf{1}_x + \sigma\frac{\mathbf{1}_{S_1^{\mathrm{p}-}(x)}}{\sqrt{D_x^{\mathrm{p}-}}} - \frac{\mathbf{1}_{\mathrm{Sib}^-(x)}}{\#\,\mathrm{Sib}^-(x)}\right).$$

The family $(\mathbf{u}_\sigma(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})},\sigma\in\{\pm1\}}$ is orthonormal.

The family $(\mathbf{u}_\sigma(x))$ is convenient as it is an orthonormal family of vectors that are localized around the vertices of $\mathcal{V}_\nu^{(\mathrm{h})}$. In the sequel, we will consider the operator

$$\sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm 1\}}} \sigma\sqrt{D_x^{\mathrm{p}-}}\,\mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^*\,. \tag{3.1}$$

In fact, Proposition 4.1 stated in the sequel implies that this is a good approximation of A^{p} restricted to the eigenvectors of its greatest eigenvalues. The operator (3.1) will be further approximated by

$$\sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm 1\}}} \sigma\sqrt{D_x^{\mathrm{p}-}}\,\mathbf{v}_\sigma(x)\mathbf{v}_\sigma(x)^*\,, \tag{3.2}$$

in Proposition 3.4. It is convenient to define (3.2) as it has good geometric properties: it is the adjacency matrix of the pruned graph restricted to the neighborhood of the vertices in $\mathcal{V}^{(\mathrm{h})}$. This geometric insight will be useful in Section 4.2. We now show that the matrices (3.1) and (3.2) are similar.

Proposition 3.4. *Let $\nu>0$. With ν-high probability,*

$$\left\|\sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm 1\}}} \sigma\sqrt{D_x^{\mathrm{p}-}}\left(\mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^*-\mathbf{v}_\sigma(x)\mathbf{v}_\sigma(x)^*\right)\right\| \leqslant \sqrt{9\nu\frac{\log N}{\log\log N}}\,.$$

We now turn to the proofs. Before showing that the vectors $(\mathbf{u}_\sigma(x))$ form an orthogonal family, we explain heuristically how their expression was found. Below, we discuss a Gram-Schmidt orthonormalization procedure. The vector obtained by this procedure will provide us with the ansatz that allowed us to define the $(\mathbf{u}_\sigma(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})},\sigma\in\{\pm1\}}$.

Introduce the families of vectors $(\mathbf{V}_0(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$, $(\mathbf{V}_1(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$, defined by

$$\mathbf{V}_0(x)=\mathbf{1}_x\,,\qquad\text{and}\qquad \mathbf{V}_1(x)=\frac{\mathbf{1}_{S_1^{\mathrm{p}-}(x)}}{\sqrt{D_x^{\mathrm{p}-}}}\,.$$

In particular, we have $\mathbf{v}_\sigma(x)=(\mathbf{V}_0(x)+\sigma\mathbf{V}_1(x))/\sqrt{2}$.

The family of vectors we shall consider is the family $(\mathbf{U}_1(x),\mathbf{U}_0(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$ obtained after applying the Gram-Schmidt orthonormalization procedure on $(\mathbf{V}_1(x),\mathbf{V}_0(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$, starting with the vectors of $(\mathbf{V}_1(x))$, and ordering the vectors of $(\mathbf{V}_0(x))$ decreasingly according to the strict order $\prec$. Working in the pruned graph makes this procedure simpler. Proposition 2.5 implies that the family $(\mathbf{V}_1(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$ is orthonormal. Indeed, if $x\neq y$, then $\langle\mathbf{V}_1(x),\mathbf{V}_1(y)\rangle$ is nonzero if and only if there is a down-up path between $x\in\mathcal{V}^{(\mathrm{h})}$ and $y\in\mathcal{V}^{(\mathrm{h})}$, and in the pruned graph there are no such paths. Thus, we set $\mathbf{U}_1(x)=\mathbf{V}_1(x)$ for all $x\in\mathcal{V}^{(\mathrm{h})}$.

The vectors $\mathbf{U}_0(x), x\in\mathcal{V}_\nu^{(\mathrm{h})}$ resulting from the Gram-Schmidt procedure, are defined by

$$\mathbf{U}_0(x):=\frac{\tilde{\mathbf{U}}_0(x)}{\|\tilde{\mathbf{U}}_0(x)\|}\,,\qquad \tilde{\mathbf{U}}_0(x):=\mathbf{V}_0(x)-\sum_{y\in\mathcal{V}^{(\mathrm{h})}}\langle\mathbf{V}_0(x),\mathbf{U}_1(y)\rangle\mathbf{U}_1(y)-\sum_{\substack{y\in\mathcal{V}^{(\mathrm{h})}\\ y\succ x}}\langle\mathbf{V}_0(x),\mathbf{U}_0(y)\rangle\mathbf{U}_0(y)\,.$$

We then have

$$
\begin{aligned}
\tilde{\mathbf{U}}_0(x) &= \mathbf{1}_x - \sum_{y\in\mathcal{V}^{(\mathrm{h})}} \frac{\left\langle \mathbf{1}_x, \mathbf{1}_{S_1^{\mathrm{p}-}(y)} \right\rangle}{D_y^{\mathrm{p}-}} \mathbf{1}_{S_1^{\mathrm{p}-}(y)} - \sum_{\substack{y\in\mathcal{V}^{(\mathrm{h})}\\ y\succ x}} \langle \mathbf{1}_x, \mathbf{U}_0(y)\rangle \mathbf{U}_0(y) \\
&= \mathbf{1}_x - \sum_{y\in\mathcal{V}^{(\mathrm{h})}} \frac{\mathbb{1}_{\{x\sim y, x\prec y\}}}{D_y^{\mathrm{p}-}} \mathbf{1}_{S_1^{\mathrm{p}-}(y)} - \sum_{\substack{y\in\mathcal{V}^{(\mathrm{h})}\\ y\succ x}} \langle \mathbf{1}_x, \mathbf{U}_0(y)\rangle \mathbf{U}_0(y) \\
&= \mathbf{1}_x - \frac{1}{D_{\hat{x}}^{\mathrm{p}-}} \mathbf{1}_{S_1^{\mathrm{p}-}(\hat{x})} - \sum_{y\succ x} \langle \mathbf{1}_x, \mathbf{U}_0(y)\rangle \mathbf{U}_0(y)\,.
\end{aligned} \tag{3.3}
$$

Lemma 3.5. *For all $x \in \mathcal{V}_\nu^{(\mathrm{h})}$, the vector $\mathbf{U}_0(x)$ is supported on $\{x\} \cup \mathrm{Sib}(x)$.*

Proof. We proceed by induction, starting from the vertices that are the greatest for the order $\prec$. We first notice that for all $x \in \mathcal{V}_\nu^{(\mathrm{h})}$ which has no parent, we have $\tilde{\mathbf{U}}_0(x) = \mathbf{1}_x$.

Then, considering (3.3), we see that

$$
\mathbf{1}_x - \frac{1}{D_{\hat{x}}^{\mathrm{p}-}} \mathbf{1}_{S_1^{\mathrm{p}-}(\hat{x})}
$$

is supported on $\{x\} \cup S_1^{\mathrm{p}-}(\hat{x})$.

By the induction hypothesis $\langle \mathbf{1}_x, \mathbf{U}_0(y)\rangle$ with $y \succ x$, is non-zero only if $x \in S_1^{\mathrm{p}-}(\hat{y})$, i.e. x and y are siblings. Thus,

$$
\sum_{\substack{y\in\mathcal{V}_\nu^{(\mathrm{h})}\\ y\succ x}} \langle \mathbf{1}_x, \mathbf{U}_0(y)\rangle \mathbf{U}_0(y)
$$

is supported on $\{x\} \cup S_1^{-}(\hat{x})$. Indeed, the siblings of the siblings y of x are the siblings of x. □

Lemma 3.5 serves as heuristics to prove Proposition 3.3.

Proof of Proposition 3.3. Motivated by Lemma 3.5, we look for vectors $\hat{\mathbf{U}}_0(x), x \in \mathcal{V}_\nu^{(\mathrm{h})}$ of the form

$$
\hat{\mathbf{U}}_0(x) = a_x \mathbf{1}_x + \sum_{y\in S_1^{\mathrm{p}-}(\hat{x})} b_x(y) \mathbf{1}_y\,,
$$

and set

$$
\hat{\mathbf{U}}_1(x) = \frac{1}{\sqrt{D_x^{\mathrm{p}-}}} \mathbf{1}_{S_1^{\mathrm{p}-}(x)}\,.
$$

We are going to find necessary and sufficient conditions for the family $\left(\hat{\mathbf{U}}_0(x), \hat{\mathbf{U}}_1(x)\right)_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$ to be orthogonal. If it were the case then, for all $x, y \in \mathcal{V}_\nu^{(\mathrm{h})}$,

$$
0 = \left\langle \hat{\mathbf{U}}_0(x), \hat{\mathbf{U}}_1(y) \right\rangle = \delta_{y\hat{x}} \left(\frac{a_x}{\sqrt{D_y^{\mathrm{p}-}}} + \sum_{z\in S_1^{\mathrm{p}-}(\hat{x})} \frac{b_x(z)}{\sqrt{D_y^{\mathrm{p}-}}} \right). \tag{3.4}
$$

Furthermore, for all $x, y \in \mathcal{V}_\nu^{(\mathrm{h})}$,

$$
\delta_{xy} = \left\langle \hat{\mathbf{U}}_0(x), \hat{\mathbf{U}}_0(y) \right\rangle = a_x^2 \delta_{xy} + \delta_{\hat{x}\hat{y}}(a_x b_y(x) + a_y b_x(y)) + \delta_{\hat{x}\hat{y}} \sum_{z\in S_1^{\mathrm{p}-}(\hat{x})} b_x(z) b_y(z)\,. \tag{3.5}
$$

This implies than if $x = y$

$$1 = a_x^2 + 2a_x b_x(x) + \sum_{z \in S_1^{\mathrm{p}-}(\hat{x})} b_x(z)^2 \,,$$

and if $x \neq y$ with $\hat{x} = \hat{y}$,

$$0 = a_x b_y(x) + a_y b_x(y) + \sum_{z \in S_1^{\mathrm{p}-}(\hat{x})} b_x(z) b_y(z) \,.$$

Consider the particular choice

$$a_x = \frac{1}{\sqrt{Z_x 2}} \quad \text{and} \quad b_x(y) = -\frac{\mathbb{1}_{\{y \prec x\}}}{\sqrt{Z_x/2 \# \operatorname{Sib}^-(x)}},$$

if x has a parent and $b_y(x) = 0$ otherwise. With this choice of coefficients, (3.4) and (3.5) are satisfied. The family

$$(\hat{\mathbf{U}}_1(x), \hat{\mathbf{U}}_0(x))_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} = \left(\frac{\mathbf{1}_{S_1^-(x)}}{\sqrt{D_x^{\mathrm{p}-}}} \,,\, a_x \mathbf{1}_x + \sum_{y \in S_1^{\mathrm{p}-}(\hat{x})} b_x(y) \mathbf{1}_y \right)_{x \in \mathcal{V}_\nu^{(\mathrm{h})}}$$

is then orthonormal. Note however that we do not claim that it is the family obtained from the orthonormalization of $(\mathbf{V}_1(x), \mathbf{V}_0(x))$. This does not matter: defining

$$\mathbf{u}_\sigma(x) = \frac{\hat{\mathbf{U}}_0(x) + \sigma \hat{\mathbf{U}}_1(x)}{\sqrt{2}}$$

yields an orthonormal family. $\square$

We now turn to the proof of Proposition 3.4. We first need to prove Lemmas 3.6 and 3.7 to bound the number of vertices in a ball around a vertex x, and of siblings of a vertex x. Lemma 3.6 will also be used in the proof of Proposition 4.1.

Lemma 3.6. *Let $x \in [N]$, and $\eta > 0$ and $\nu > 0$ two constants. With ν-high probability, for all $x \in \mathcal{V}^{(\mathrm{h})}$ we have*

$$\frac{1}{D_x^{\mathrm{p}}} \sum_{y \in S_1^{\mathrm{p}-}(x)} D_y^{\mathrm{p}-} \leqslant 3\nu \frac{\log N}{\log \log N} \,.$$

Lemma 3.7. *Let $\nu > 0$. With ν-high probability, for all $x \in \mathcal{V}_\nu^{(\mathrm{h})}$ that has a parent, we have*

$$\# \operatorname{Sib}^-(x) \geqslant \frac{1}{2} D_{\hat{x}}^{\mathrm{p}} \,.$$

Proof of Lemma 3.6. We start by noticing that by the union bound

$$\begin{aligned} \mathbb{P}\left(\forall x \in \mathcal{V}^{(\mathrm{h})}, \frac{1}{D_x^{\mathrm{p}}} \sum_{y \in S_1^{\mathrm{p}-}(x)} D_y^{\mathrm{p}-} \leqslant (\log N)^{\eta+\delta} \right) &= 1 - \mathbb{P}\left(\exists x \in \mathcal{V}^{(\mathrm{h})}, \frac{1}{D_x^{\mathrm{p}}} \sum_{y \in S_1^{\mathrm{p}-}(x)} D_y^{\mathrm{p}-} > (\log N)^{\eta+\delta} \right) \\ &\geqslant 1 - \sum_{x \in [N]} \mathbb{P}\left(x \in \mathcal{V}^{(\mathrm{h})}, \frac{1}{D_x^{\mathrm{p}}} \sum_{y \in S_1^{\mathrm{p}-}(x)} D_y^{\mathrm{p}-} > (\log N)^{\eta+\delta} \right). \end{aligned}$$

It suffices to upper bound the term in the sum by $CN^{-\nu-1}$ to get the result.

We introduce the notation

$$\mathcal{P}_x^{(3)} := \frac{1}{D_x} \sum_{y \in S_1^-(x) \setminus S_1^{\mathrm{cyc}}(x)} (D_y^{\mathrm{nc}} - 1) \,.$$

It suffices to bound $\mathcal{P}_x^{(3)}$ to bound $\sum_{y \in S_1^{\mathrm{p}-}(x)} D_y^{\mathrm{p}-}$. We prove below that

$$\mathcal{P}_x^{(3)} \leqslant \frac{5\nu}{4(1-\delta)} \frac{\log N}{\log\log N} \quad \text{with } \nu\text{-high probability.} \tag{3.6}$$

Assuming (3.6), by Proposition 2.5 we get

$$\frac{1}{D_x^{\mathrm{p}}} \sum_{y \in S_1^{\mathrm{p}-}(x)} D_y^{\mathrm{p}-} \leqslant 1 + \mathcal{P}_x^{(3)} \frac{D_x}{D_x^{\mathrm{p}}} \leqslant 1 + \frac{5\nu \log N}{4(1-\delta)\log\log N} \frac{D_x}{D_x - \xi/2} \leqslant 1 + 2\nu \frac{\log N}{\log\log N} \,.$$

which gives the result.

We turn to the proof of (3.6). Firstly, recall that using Remark 1.13, we have with ν-high probability that

$$w_x \leqslant \chi D_x \,,$$

where $\chi = \log\log N$. If furthermore $y \prec x$, we have

$$\frac{D_y^{\mathrm{nc}}}{D_x} \leqslant 1 \wedge \frac{\chi D_y^{\mathrm{nc}}}{D_x} \,.$$

Secondly, using Proposition 2.15 and Corollary 2.16, we have

$$\mathcal{P}_x^{(3)} \leqslant \sum_{y \neq x} \mathbb{1}_{\{x \overset{\mathcal{T}_x}{\sim} y\}} \left(1 \wedge \frac{\chi D_y^{\mathcal{T}_x \uparrow}}{w_x}\right) ,$$

with ν-high probability.

We are going to prove (3.6) using Bennett's inequality [BLM13, Theorem 2.9]. Note that the variables

$$\mathbb{1}_{\{x \overset{\mathcal{T}_x}{\sim} y\}} \left(1 \wedge \frac{\chi D_y^{\mathcal{T}_x \uparrow}}{w_x}\right)$$

are independent by Lemma 2.12, and bounded by 1. Furthermore, we have

$$v := \sum_{y \in [N]} \mathbb{E}\left[\mathbb{1}_{\{x \overset{\mathcal{T}_x}{\sim} y\}} \left(1 \wedge \frac{\chi D_y^{\mathcal{T}_x \uparrow}}{w_x}\right)\right] \leqslant \sum_{y \in [N]} p_{xy} \frac{\chi w_y}{w_x} \leqslant \chi \frac{m_2}{m_1} \,.$$

Hence, we have for $t > 0$

$$\mathbb{P}\left(\mathcal{P}_x^{(3)} \geqslant t + v\right) \leqslant \exp\left(-(v+t)\log\frac{t}{v} + t\right) + O(N^{-\nu}) \,.$$

Taking $t = \frac{5\nu}{4(1-\delta)} \frac{\log N}{\log\log N}$ proves the claim (3.6), and thus the result. $\square$

Proof of Lemma 3.7. To show this, we shall rather show that with ν high probability, we have the following property. For all $x \in \mathcal{V}_\nu^{(h)}$, the random variable

$$D_x^{\mathrm{nc},\mathcal{V}} := \#\Big(S_1^{\mathrm{nc}}(x) \cap \mathcal{V}_\nu^{(h)}\Big) = \sum_{y \neq x} \mathbb{1}_{\{x \overset{\mathrm{nc}}{\sim} y, \xi < D_y\}},$$

is bounded by $D_x/4$.

We see that by using Proposition 2.7 and then Proposition 2.15, we get

$$D_x^{\mathrm{nc},\mathcal{V}} \leqslant \sum_{y \neq x} \mathbb{1}_{\{x \overset{\mathrm{nc}}{\sim} y, \xi < D_y^{\mathrm{nc}} + C_\nu\}} \leqslant \sum_{y \neq x} \mathbb{1}_{\{x \overset{\mathcal{T}_x}{\sim} y, \xi \leqslant D_y^{\mathcal{T}_x \uparrow} + C_\nu\}},$$

with ν-high probability for some constant $C_\nu > 0$.

Let us use Bennett's inequality to bound the right-hand side. We proceed as in the proof of Lemma 2.18. The random variables

$$\Big(\mathbb{1}_{\{x \overset{\mathcal{T}_x}{\sim} y\}} \mathbb{1}_{\{\xi \leqslant D_y^{\mathcal{T}_x \uparrow} + C_\nu\}}\Big)_{y \in [N]}$$

are independent by Lemma 2.12, and bounded by 1. We compute

$$v := \sum_{y \neq x} \mathbb{E}\Big[\mathbb{1}_{\{x \overset{\mathcal{T}_x}{\sim} y, \xi \leqslant D_y^{\mathcal{T}_x \uparrow} + C_\nu\}}\Big] \leqslant \sum_{y \neq x} p_{xy} \frac{d_y}{\xi - C_\nu} \leqslant 2 \frac{w_x}{\xi} \frac{m_2}{m_1}.$$

We get that

$$\mathbb{P}\Big(D_x^{\mathrm{nc},\mathcal{V}} \geqslant t\Big) \leqslant \exp\Big(-t \ln \frac{t}{v} + t\Big) + O(N^{-\nu}).$$

Take $t = \frac{w_x}{4\chi} \vee \frac{2\nu \log N}{\log\log N}$.

If $\frac{w_x}{4\chi} \geqslant \frac{2\nu \log N}{\log\log N}$,

$$\mathbb{P}\Big(D_x^{\mathrm{nc},\mathcal{V}} \geqslant t\Big) \leqslant \exp\Big(-2\nu(1-\delta) \ln N(1+o(1))\Big) + O(N^{-\nu}) = O(N^{-\nu}),$$

as $2(1-\delta) \geqslant 4/3 > 1$. Recall that by Remark 1.13 we have with ν-high probability that

$$w_x \leqslant \chi D_x.$$

We then have

$$D_x^{\mathrm{nc},\mathcal{V}} \leqslant \frac{w_x}{4\chi} \leqslant \frac{D_x}{4}$$

with ν-high probability.

If $\frac{w_x}{4\chi} \leqslant \frac{2\nu \log N}{\log\log N}$,

$$\mathbb{P}\Big(D_x^{\mathrm{nc},\mathcal{V}} \geqslant t\Big) \leqslant \exp\Big(-2\nu \ln N(1+o(1))\Big) + O(N^{-\nu}) = O(N^{-\nu}),$$

and then

$$D_x^{\mathrm{nc},\mathcal{V}} \leqslant 2\nu \frac{\log N}{\log\log N} \leqslant \frac{\xi}{4} \leqslant \frac{D_x}{4},$$

with ν-high probability. We have proved the claim.

For $x \in \mathcal{V}_\nu^{(\mathrm{h})}$, the neighbors of $\hat{x}$ are in $\mathrm{Sib}^-(x)$, or in $S_1(\hat{x}) \cap \mathcal{V}_\nu^{(\mathrm{h})}$. Thus,

$$\#\,\mathrm{Sib}^-(x) + D_{\hat{x}}^{\mathrm{nc},\mathcal{V}} \geqslant D_{\hat{x}}^{\mathrm{p}} .$$

The previous result on $D_{\hat{x}}^{\mathrm{nc},\mathcal{V}}$ implies that with ν-high probability,

$$\#\,\mathrm{Sib}^-(x) \geqslant D_{\hat{x}}^{\mathrm{p}} - D_{\hat{x}}^{\mathrm{nc},\mathcal{V}} \geqslant D_{\hat{x}}^{\mathrm{p}} - \frac{1}{4} D_{\hat{x}} \geqslant \frac{1}{2} D_{\hat{x}}^{\mathrm{p}} ,$$

where we used that when $x \in \mathcal{V}_\nu^{(\mathrm{h})}$, by Proposition 2.5, $D_x - D_x^{\mathrm{p}} \leqslant \xi/2$ which implies $D_{\hat{x}}/2 \leqslant D_{\hat{x}} - \xi/2 \leqslant D_{\hat{x}}^{\mathrm{p}}$. □

Proof of Proposition 3.4. First, notice that

$$\sum_{\sigma \in \{\pm 1\}} \sigma \sqrt{D_x^{\mathrm{p}-}} \mathbf{u}_\sigma(x) \mathbf{u}_\sigma(x)^* = \frac{2}{Z_x} \left(\sum_{\sigma \in \{\pm 1\}} \sigma \sqrt{D_x^{\mathrm{p}-}} \mathbf{v}_\sigma(x) \mathbf{v}_\sigma(x)^* - \frac{\mathbf{1}_{\mathrm{Sib}^-(x)} \mathbf{1}_{S_1^-(x)}^* + \mathbf{1}_{S_1^-(x)} \mathbf{1}_{\mathrm{Sib}^-(x)}^*}{\#\,\mathrm{Sib}^-(x)} \right) ,$$

because of the cancellations occurring when summing contribution of $\sigma = +1$ and $\sigma = -1$.

We consider first

$$\sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \left(\frac{2}{Z_x} - 1 \right) \sum_{\sigma \in \{\pm 1\}} \sigma \sqrt{D_x^-} \mathbf{v}_\sigma(x) \mathbf{v}_\sigma(x)^* ,$$

whose operator norm satisfies

$$\begin{aligned} \left\| \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \left(\frac{2}{Z_x} - 1 \right) \sum_{\sigma \in \{\pm 1\}} \sigma \sqrt{D_x^-} \mathbf{v}_\sigma(x) \mathbf{v}_\sigma(x)^* \right\|^2 &= \left\| \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{\mathbf{1}_x \mathbf{1}_{S_1^-(x)}^* + \mathbf{1}_{S_1^-(x)} \mathbf{1}_x^*}{\#\,\mathrm{Sib}^-(x) + 1} \right\|^2 \\ &\leqslant \max_{\|\mathbf{u}\|=1} \sum_{x, x' \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{1}{(\#\,\mathrm{Sib}^-(x) + 1)^2} \Big(\delta_{xx'} \Big\langle \mathbf{1}_{S_1^-(x)}, \mathbf{u} \Big\rangle^2 + \delta_{xx'} D_x^- u_x^2 + 2 \mathbb{1}_{\{x \sim x', x' \prec x\}} u_x \Big\langle \mathbf{1}_{S_1^-(x')}, \mathbf{u} \Big\rangle \Big). \end{aligned}$$

Lemma 3.7 implies that with ν-high probability, for all $x \in \mathcal{V}_\nu^{(\mathrm{h})}$,

$$\frac{1}{2} D_x^- \leqslant \#\,\mathrm{Sib}^-(x) .$$

Together with Young's Lemma and the fact that $\mathbf{1}_{S_1^-(x)}$ form an orthogonal family in the pruned graph, it implies

$$\left\| \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \left(\frac{2}{Z_x} - 1 \right) \sum_{\sigma \in \{\pm 1\}} \sigma \sqrt{D_x^{\mathrm{p}-}} \mathbf{v}_\sigma(x) \mathbf{v}_\sigma(x)^* \right\|^2 = O(1/\xi).$$

Consider now the operator norm

$$\begin{aligned} \left\| \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{\mathbf{1}_{\mathrm{Sib}^-(x)} \mathbf{1}_{S_1^{\mathrm{p}-}(x)}^*}{\#\,\mathrm{Sib}^-(x)} \right\|^2 &\leqslant \max_{\|\mathbf{u}\|=1} \sum_{x, x' \in \mathcal{V}_\nu^{(\mathrm{h})}} \delta_{xx'} \frac{D_x^{\mathrm{p}-} \Big\langle \mathbf{1}_{\mathrm{Sib}^-(x)}, \mathbf{u} \Big\rangle \Big\langle \mathbf{1}_{\mathrm{Sib}^-(x')}, \mathbf{u} \Big\rangle}{(\#\,\mathrm{Sib}^-(x))^2} \\ &\quad + \max_{\|\mathbf{u}\|=1} 2 \sum_{x, x' \in \mathcal{V}_\nu^{(\mathrm{h})}} \delta_{\hat{x}\hat{x}'} \mathbb{1}_{\mathrm{Sib}^-(x)}(x') \frac{\Big\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x)}, \mathbf{u} \Big\rangle \Big\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x')}, \mathbf{u} \Big\rangle}{\#\,\mathrm{Sib}^-(x)} . \end{aligned}$$

The first term is bounded by

$$\sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \delta_{xx'} \frac{D_x^{\mathrm{p}-}\left\langle \mathbf{1}_{\mathrm{Sib}^-(x)},\mathbf{u}\right\rangle\left\langle \mathbf{1}_{\mathrm{Sib}^-(x')},\mathbf{u}\right\rangle}{(\#\,\mathrm{Sib}^-(x))^2} = \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{D_x^{\mathrm{p}-}}{\#\,\mathrm{Sib}^-(x)^2} \sum_{y,y'\in\mathrm{Sib}^-(x)} u_y u_{y'}$$
$$\leqslant 2 \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{D_x^{\mathrm{p}-}}{D_{\hat{x}}^{\mathrm{p}-}} \sum_{y\in\mathrm{Sib}^-(x)} u_y^2\,,$$

with ν-high probability. We used Lemma 3.7 and Young's inequality. This can be bounded as follows:

$$\sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \delta_{xx'} \frac{D_x^{\mathrm{p}-}\left\langle \mathbf{1}_{\mathrm{Sib}^-(x)},\mathbf{u}\right\rangle\left\langle \mathbf{1}_{\mathrm{Sib}^-(x')},\mathbf{u}\right\rangle}{(\#\,\mathrm{Sib}^-(x))^2} \leqslant 2\sum_y u_y^2 \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}\cap S_1^-(\hat{y})} \frac{D_x^{\mathrm{p}-}}{D_{\hat{y}}^{\mathrm{p}-}}\,.$$

Lemma 3.6 then gives that this is upper bounded by $2\nu\frac{\log N}{\log\log N}$ with ν-high probability.

Similarly, using Young's inequality, Lemma 3.7, and then Lemma 3.6 we see that the second term is bounded by

$$\sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \delta_{\hat{x}\hat{x}'} \mathbb{1}_{\mathrm{Sib}^-(x)}(x') \frac{\left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x)},\mathbf{u}\right\rangle\left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x')},\mathbf{u}\right\rangle}{\#\,\mathrm{Sib}^-(x)} \leqslant 2\sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \left(\frac{\left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x)},\mathbf{u}\right\rangle}{\sqrt{D_x^{\mathrm{p}-}}}\right)^2 \sum_{x'\in S_1^-(\hat{x})\cap\mathcal{V}_\nu^{(\mathrm{h})}} \frac{D_{x'}^{\mathrm{p}-}}{D_{\hat{x}}^{\mathrm{p}-}}$$
$$\leqslant 2\nu\frac{\log N}{\log\log N} \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \left(\frac{\left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x)},\mathbf{u}\right\rangle}{\sqrt{D_x^{\mathrm{p}-}}}\right)^2,$$

with ν-high probability. The fact that $(V_1(x))_{x\in\mathcal{V}_\nu^{(\mathrm{h})}}$ is an orthonormal family allow us to conclude that

$$\left\|\sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{\mathbf{1}_{\mathrm{Sib}^-(x)}\mathbf{1}^*_{S_1^{\mathrm{p}-}(x)} + \mathbf{1}_{S_1^{\mathrm{p}-}(x)}\mathbf{1}^*_{\mathrm{Sib}^-(x)}}{\#\,\mathrm{Sib}^-(x)}\right\| \leqslant 2\left\|\sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{\mathbf{1}_{\mathrm{Sib}^-(x)}\mathbf{1}^*_{S_1^{\mathrm{p}-}(x)}}{\#\,\mathrm{Sib}^-(x)}\right\| \leqslant \sqrt{8\nu\frac{\log N}{\log\log N}}$$

with ν-high probability. Putting together the two bounds, we get the result. □

4. The spectral gap and the semilocalization phenomenon

In this Section, we prove Theorem 1.8. To do so, we construct a block-diagonal approximation $\hat{A}$ of A, whose eigenvectors associated to the biggest eigenvalues are the $\mathbf{u}_\sigma(x)$ for $\sigma\in\{\pm1\}$ and $x\in\mathcal{V}_\nu^{(\mathrm{h})}$. A spectral gap property is the proved for $\hat{A}$, and can be transferred to a spectral gap property for A.

4.1. The block-diagonal approximation. We introduce some notation. We consider in the sequel the orthogonal projections

$$\Pi^{\mathrm{p}} = \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^* \quad\text{and}\quad \overline{\Pi}^{\mathrm{p}} = \mathrm{Id}-\Pi^{\mathrm{p}}\,, \tag{4.1}$$

and the block-diagonal approximation of A^{p},

$$\hat{A} = \sum_{\substack{x \in \mathcal{V}_\nu^{(\mathrm{h})} \\ \sigma \in \{\pm 1\}}} \sigma \sqrt{D_x^{\mathrm{p}-}} \mathbf{u}_\sigma(x) \mathbf{u}_\sigma(x)^* + \overline{\Pi} A^{\mathrm{p}} \overline{\Pi} . \tag{4.2}$$

Proposition 4.1. *Let $\nu > 0$. There exists $C_\nu > 0$ such that with ν-high probability,*

$$\|\hat{A} - A^{\mathrm{p}}\| \leqslant C_\nu \sqrt{\frac{\log N}{\log \log N}} \quad \text{and} \quad \|\hat{A} - A\| \leqslant C_\nu \sqrt{\frac{\log N}{\log \log N}} .$$

To prove Proposition 4.1 we use the following Lemma.

Lemma 4.2. *Define for all $x \in \mathcal{V}_\nu^{(\mathrm{h})}, \sigma \in \{\pm 1\}$,*

$$\delta_\sigma(x) = A^{\mathrm{p}} \mathbf{u}_\sigma(x) - \sigma \sqrt{D_x^{\mathrm{p}-}} \mathbf{u}_\sigma(x) .$$

This vector can also be expressed as

$$\delta_\sigma(x) = \frac{1}{\sqrt{Z_x}} \left(\sigma \sum_{y \in S_1^{\mathrm{p}-}(x)} \frac{\mathbf{1}_{S_1^{\mathrm{p}}(y) \setminus \{x\}}}{\sqrt{D_x^{\mathrm{p}-}}} - \sum_{y \in S_1^{\mathrm{p}-}(\hat{x})} \frac{\mathbb{1}_{\{y \prec x\}}}{\# \operatorname{Sib}^-(x)} \mathbf{1}_{S_1^{\mathrm{p}-}(y)} + \sigma \sqrt{D_x^{\mathrm{p}-}} \frac{\mathbf{1}_{\operatorname{Sib}^-(x)}}{\# \operatorname{Sib}^-(x)} \right) .$$

Proof of Lemma 4.2. We compute $A^{\mathrm{p}} \mathbf{u}_\sigma(x)$:

$$\begin{aligned} A^{\mathrm{p}} \mathbf{u}_\sigma(x) &= \frac{1}{\sqrt{Z_x}} \left(\mathbf{1}_{S_1^{\mathrm{p}}(x)} + \frac{\sigma}{\sqrt{D_x^{\mathrm{p}-}}} \left(D_x^{\mathrm{p}-} \mathbf{1}_x + \sum_{y \in S_1^{\mathrm{p}-}(x)} \mathbf{1}_{S_1^{\mathrm{p}}(y) \setminus \{x\}} \right) - \sum_{y \in S_1^{\mathrm{p}-}(\hat{x})} \frac{\mathbb{1}_{\{y \prec x\}}}{\# \operatorname{Sib}^-(x)} \mathbf{1}_{S_1^{\mathrm{p}}(y)} \right) \\ &= \sigma \sqrt{D_x^-} \mathbf{u}_\sigma(x) \\ &\quad + \frac{\sigma}{\sqrt{Z_x}} \left(\sum_{y \in S_1^-(x)} \frac{\mathbf{1}_{S_1^{\mathrm{p}}(y) \setminus \{x\}}}{\sqrt{D_x^-}} - \sigma \sum_{y \in S_1^{\mathrm{p}-}(\hat{x})} \frac{\mathbb{1}_{\{y \prec x\}}}{\# \operatorname{Sib}^-(x)} \mathbf{1}_{S_1^{\mathrm{p}}(y) \setminus \{\hat{x}\}} + \sqrt{D_x^{\mathrm{p}-}} \frac{\mathbf{1}_{\operatorname{Sib}^-(x)}}{\# \operatorname{Sib}^-(x)} \right) . \end{aligned}$$

The result follows when noticing that in the pruned graph $S_1^{\mathrm{p}}(y) \setminus \{\hat{y}\} = S_1^{\mathrm{p}-}(y)$. □

Proof of Proposition 4.1. The triangular inequality yields

$$\|\hat{A} - A\| \leqslant \|\hat{A} - A^{\mathrm{p}}\| + \|A^{\mathrm{p}} - A\| .$$

Proposition 2.19 allows us to bound the second part. The first part can be bounded as follows

$$\|\hat{A} - A^{\mathrm{p}}\| \leqslant \Bigg\| \sum_{\substack{x \in \mathcal{V}_\nu^{(\mathrm{h})} \\ \sigma \in \{\pm 1\}}} \sigma \sqrt{D_x^-} \mathbf{u}_\sigma(x) \mathbf{u}_\sigma(x)^* - \Pi^{\mathrm{p}} A^{\mathrm{p}} \Pi^{\mathrm{p}} \Bigg\| + \|\overline{\Pi}^{\mathrm{p}} A^{\mathrm{p}} \Pi^{\mathrm{p}} + \Pi^{\mathrm{p}} A^{\mathrm{p}} \overline{\Pi}^{\mathrm{p}}\| .$$

Recall that in Lemma 4.2, we defined the vector

$$\delta_\sigma(x) = A^{\mathrm{p}} \mathbf{u}_\sigma(x) - \sigma \sqrt{D_x^{\mathrm{p}-}} \mathbf{u}_\sigma(x) \quad \text{for all } x \in \mathcal{V}_\nu^{(\mathrm{h})}, \sigma \in \{\pm 1\} .$$

We have

$$A^{\mathrm{p}}\Pi^{\mathrm{p}} = \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} A^{\mathrm{p}}\mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^* = \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \sigma\sqrt{D_x^{\mathrm{p}-}}\mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^* + B\,,$$

where $B = \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \delta_\sigma(x)\mathbf{u}_\sigma(x)^*$. We now bound the operator norm of B. Using the expression of $\delta_\sigma(x)$ from Lemma 4.2, we write the operator as a sum $B = B_1 + B_2 + B_3 + B_4$, with

$$B_1 = \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x} \sum_{y\in S_1^-(x)} \frac{\mathbf{1}_{S_1^{\mathrm{p}-}(y)}\mathbf{1}^*_{S_1^{\mathrm{p}-}(x)}}{D_x^{\mathrm{p}-}}$$

$$B_2 = \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x} \sum_{\substack{y\in S_1^{\mathrm{p}-}(\hat{x})\\ y\prec x}} \frac{\mathbf{1}_y\mathbf{1}^*_{S_1^{\mathrm{p}-}(x)}}{\#\,\mathrm{Sib}^-(x)}$$

$$B_3 = -\sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x} \sum_{\substack{y\in S_1^{\mathrm{p}-}(\hat{x})\\ y\prec x}} \frac{\mathbf{1}_{S_1^{\mathrm{p}-}(y)}\mathbf{1}^*_x}{\#\,\mathrm{Sib}^-(x)}$$

$$B_4 = \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x} \sum_{\substack{y,z\in S_1^{\mathrm{p}-}(\hat{x})\\ y,z\prec x}} \frac{\mathbf{1}_{S_1^{\mathrm{p}-}(y)}\mathbf{1}^*_z}{(\#\,\mathrm{Sib}^-(x))^2}\,.$$

We now bound the operator norm of each of the four operators. The first one is

$$\begin{aligned}\|B_1\|^2 &= \max_{\|\mathbf{u}\|=1}(\mathbf{u}^*B_1^*B_1\mathbf{u})\\ &= \max_{\|\mathbf{u}\|=1} \sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{4}{Z_xZ_{x'}} \frac{\left\langle\mathbf{u},\mathbf{1}_{S_1^{\mathrm{p}-}(x)}\right\rangle\left\langle\mathbf{1}_{S_1^{\mathrm{p}-}(x')},\mathbf{u}\right\rangle}{D_x^{\mathrm{p}-}D_{x'}^{\mathrm{p}-}} \sum_{\substack{y\in S_1^{\mathrm{p}-}(x)\\ y'\in S_1^{\mathrm{p}-}(x')}} \left\langle\mathbf{1}_{S_1^{\mathrm{p}-}(y)},\mathbf{1}_{S_1^{\mathrm{p}-}(y')}\right\rangle.\end{aligned}$$

The orthogonality of the vectors $(\mathbf{1}_{S_1^{\mathrm{p}-}(y)})_{y\in[N]} = (\sqrt{D_y^{\mathrm{p}-}}V_1(y))_{y\in[N]}$ implies that we must take $y = y'$ in the sum above. As $x = \hat{y}$ and $x' = \hat{y}'$ this means that the contributions for $x \neq x'$ vanish. We are left with

$$\|B_1\|^2 = \max_{\|\mathbf{u}\|=1} \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{4}{Z_x^2} \frac{\left\langle\mathbf{u},\mathbf{1}_{S_1^{\mathrm{p}-}(x)}\right\rangle^2}{D_x^{\mathrm{p}-}} \leqslant \max_{\|\mathbf{u}\|=1} \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \langle\mathbf{u},\mathbf{V}_1(x)\rangle^2 \leqslant 1\,.$$

The other bounds are proved similarly, in some cases with the help of Lemma 3.6, in Section A. Thus, there exists a constant $C'_\nu > 0$ such that with ν-high probability,

$$\|A^{\mathrm{p}}\Pi^{\mathrm{p}}\| \leqslant C'_\nu\sqrt{\frac{\log N}{\log\log N}}\,.$$

The result follows from the fact that Π^{p} and $\overline{\Pi}^{\mathrm{p}}$ are orthogonal projections. $\square$

4.2. Bounds on $\overline{\Pi}A^{\mathrm{p}}\overline{\Pi}$. The last step before proving Theorem 1.8 is to bound the operator norm of $\overline{\Pi}A^{\mathrm{p}}\overline{\Pi}$. To do so, we use the fact that the pruned graph is actually a forest.

Let us consider $\overline{\Pi}A^{\mathrm{p}}\overline{\Pi}$. By definition, see (4.1), we have

$$\begin{aligned}\overline{\Pi}A^{\mathrm{p}}\overline{\Pi} = \hat{A} - \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \sigma\sqrt{D_x^-}\,\mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^* = (A^{\mathrm{p}} - \hat{A}) \\ - \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \sigma\sqrt{D_x^-}\,(\mathbf{u}_\sigma(x)\mathbf{u}_\sigma(x)^* - \mathbf{v}_\sigma(x)\mathbf{v}_\sigma(x)^*) \\ + (A^{\mathrm{p}} - \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \sigma\sqrt{D_x^-}\,\mathbf{v}_\sigma(x)\mathbf{v}_\sigma(x)^*)\,.\end{aligned}$$

Proposition 4.1 implies that the error $\|A^{\mathrm{p}} - \hat{A}^{\mathrm{p}}\|$ is at most of order $\sqrt{\frac{\log N}{\log\log N}}$, and Proposition 3.4 that the second part is at most of order $\sqrt{\frac{\log N}{\log\log N}}$. We now prove that the third part is of order $\sqrt{\frac{\log N}{\log\log N}}$.

Lemma 4.3. *We have*

$$\Bigg\|A^{\mathrm{p}} - \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \sigma\sqrt{D_x^-}\,\mathbf{v}_\sigma(x)\mathbf{v}_\sigma(x)^*\Bigg\| \leqslant 2\sqrt{\xi_\nu}\,.$$

Proof. The matrix

$$A' = A^{\mathrm{p}} - \sum_{\substack{x\in\mathcal{V}_\nu^{(\mathrm{h})}\\ \sigma\in\{\pm1\}}} \sigma\sqrt{D_x^-}\,\mathbf{v}_\sigma(x)\mathbf{v}_\sigma(x)^*$$

is the adjacency matrix of the graph $\tilde{G} = G^{\mathrm{p}}|_{\left(\mathcal{V}_\nu^{(\mathrm{h})}\right)^c}$. The degree of the vertices of this graph are bounded by ξ_ν, and the graph is actually a forest by Proposition 2.5.

Thus, a standard estimate – see for instance [ADK21a, Lemma A.4] – implies that

$$\|A'\| \leqslant 2\sqrt{\xi_\nu}\,.$$

$\square$

Together with the previous discussion, Lemma 4.3 yields the following bound on $\|\overline{\Pi}^{\mathrm{p}}A^{\mathrm{p}}\overline{\Pi}^{\mathrm{p}}\|$.

Proposition 4.4. *Let $\nu > 0$. There exists a constant $C_\nu > 0$ such that with ν-high probability, we have*

$$\|\overline{\Pi}^{\mathrm{p}}A^{\mathrm{p}}\overline{\Pi}^{\mathrm{p}}\| \leqslant C_\nu\sqrt{\frac{\log N}{\log\log N}}\,.$$

4.3. Semilocalization. We now prove the main result. First, we introduce some notation. Let $\eta > 0$ and $\lambda > 0$. We define the sets

$$\mathcal{W}^{\mathrm{p}}_{\lambda,\eta} = \left\{ x \in [N] : \left| \sqrt{D^{\mathrm{p}-}_x} - \lambda \right| \leqslant \eta \right\} \quad \text{and} \quad \mathcal{W}_{\lambda,\eta} = \left\{ x \in [N] : \left| \sqrt{D_x} - \lambda \right| \leqslant \eta \right\},$$

and the orthogonal projections

$$\Pi^{\mathrm{p}}_{\lambda,\eta} = \sum_{x \in \mathcal{W}^{\mathrm{p}}_{\lambda,\eta}} \mathbf{u}_+(x) \mathbf{u}_+(x)^* \quad \text{and} \quad \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} = 1 - \Pi^{\mathrm{p}}_{\lambda,\eta}\,.$$

The following semilocalization result, and its proof, are very close to [ADK21a, Theorem 3.4].

Theorem 4.5. *Let $\nu > 0$. There exists $c_\nu, C_\nu > 0$ such that with ν-high probability, the following statement is true. For all eigenvalue $|\lambda| > C_\nu \sqrt{\frac{\log N}{\log \log N}}$ with associated normalized eigenvector q, for all $0 < \eta \leqslant |\lambda|/2$, we have*

$$\sum_{x \in \mathcal{W}^{\mathrm{p}}_{\lambda,\eta}} \langle \mathbf{q}, \mathbf{u}_+(x) \rangle^2 \geqslant 1 - \left(\frac{c_\nu}{\eta} \sqrt{\frac{\log N}{\log \log N}} \right)^2 .$$

Proof. As in [ADK21a, Theorem 3.4], the core of the proof is the spectral gap property of the form

$$\mathrm{Spec}\Big(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} A \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \Big) \subset \mathbb{R} \setminus [\lambda - \eta, \lambda + \eta]\,.$$

Consider first the block-diagonal approximation $\hat{A}$. The orthogonal projections Π^{p} and $\Pi^{\mathrm{p}}_{\lambda,\eta}$ commute and we have the inclusion property

$$\Pi^{\mathrm{p}} \Pi^{\mathrm{p}}_{\lambda,\eta} = \Pi^{\mathrm{p}}_{\lambda,\eta}\,.$$

Note that we also have

$$\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} = \mathrm{Id} - \Pi^{\mathrm{p}} \Pi^{\mathrm{p}}_{\lambda,\eta} = \overline{\Pi}^{\mathrm{p}} + \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \Pi^{\mathrm{p}}\,.$$

These properties allow us to rewrite $\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \hat{A} \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}$ as

$$\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \hat{A} \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} = \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \Pi^{\mathrm{p}} \hat{A} \Pi^{\mathrm{p}} \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} + \overline{\Pi}^{\mathrm{p}} \hat{A} \overline{\Pi}^{\mathrm{p}}\,. \tag{4.3}$$

The spectral gap property can be shown for $\Pi^{\mathrm{p}}_{\lambda,\eta} \hat{A} \Pi^{\mathrm{p}}_{\lambda,\eta}$ by showing it for each of the two terms in (4.3): they are the two blocks of a block decomposition of the operator. By definition, we immediately have

$$\mathrm{Spec}\Big(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \Pi^{\mathrm{p}} \hat{A} \Pi^{\mathrm{p}} \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} \Big) \subset \mathbb{R} \setminus [\lambda - \eta, \lambda + \eta]\,.$$

For the second part, Proposition 4.4 implies that there exists a constant $C_\nu > 0$ such that with ν-high probability,

$$\mathrm{Spec}\Big(\overline{\Pi}^{\mathrm{p}} \hat{A} \overline{\Pi}^{\mathrm{p}} \Big) \subset \left[-\frac{C_\nu}{2} \sqrt{\frac{\log N}{\log \log N}}, \frac{C_\nu}{2} \sqrt{\frac{\log N}{\log \log N}} \right] \subset \mathbb{R} \setminus [\lambda - \eta, \lambda + \eta]\,,$$

as we chose η such that $|\lambda \pm \eta| \geqslant |\lambda|/2 > C_\nu \sqrt{\frac{\log N}{\log \log N}}/2$.

We have proved
$$\mathrm{Spec}\Big(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\hat{A}\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\Big) \subset \mathbb{R} \setminus [\lambda-\eta, \lambda+\eta]\,.$$
Proposition 4.1 allows to upgrade this to a spectral gap property for $\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}$: there exists $c_\nu > 0$ such that with ν-high probability
$$\mathrm{Spec}\Big(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\Big) \subset \mathbb{R} \setminus \left[\lambda - \left(\eta - c_\nu\sqrt{\frac{\log N}{\log\log N}}\right), \lambda + \left(\eta - c_\nu\sqrt{\frac{\log N}{\log\log N}}\right)\right].$$
By convention, if $\eta - c_\nu\sqrt{\frac{\log N}{\log\log N}} > 0$, then the interval is $\emptyset$. Assume that $\eta > c_\nu\sqrt{\frac{\log N}{\log\log N}}$, otherwise the result is vacuous.

We now conclude as follows. Let λ be an eigenvalue associated to a normalized eigenvector $\mathbf{q}$, we have $(A-\lambda)\mathbf{q} = 0$. Multiplying by $\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}$ and introducing $\mathrm{Id} = \Pi^{\mathrm{p}}_{\lambda,\eta} + \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}$, we get
$$\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}(A-\lambda)\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\mathbf{q} + \overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}(A-\lambda)\Pi^{\mathrm{p}}_{\lambda,\eta}\mathbf{q} = 0\,,$$
which simplifies to
$$(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} - \lambda)\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\mathbf{q} = -\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\Pi^{\mathrm{p}}_{\lambda,\eta}\mathbf{q}\,.$$
Finally, we have
$$\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\mathbf{q} = -(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} - \lambda)^{-1}\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\Pi^{\mathrm{p}}_{\lambda,\eta}\mathbf{q}\,.$$
The spectral gap property and Proposition 4.1 imply the bounds
$$\|(\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta} - \lambda)^{-1}\| \leqslant \frac{1}{\eta - c_\nu\sqrt{\frac{\log N}{\log\log N}}}$$
$$\|\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}A\Pi^{\mathrm{p}}_{\lambda,\eta}\| \leqslant \|\hat{A} - A\| \leqslant c_\nu\sqrt{\frac{\log N}{\log\log N}}\,.$$
They allow us to deduce
$$\|\overline{\Pi}^{\mathrm{p}}_{\lambda,\eta}\mathbf{q}\| \leqslant \frac{c_\nu\sqrt{\frac{\log N}{\log\log N}}}{\eta - c_\nu\sqrt{\frac{\log N}{\log\log N}}} \wedge 1 \leqslant \frac{2c_\nu\sqrt{\frac{\log N}{\log\log N}}}{\eta}\,,$$
which is the wanted result. □

Proof of Theorem 1.8. Notice that by Proposition 2.5, with ν-high probability, $D^{\mathrm{p}}_x \leqslant D_x - \xi/2$. Thus, assuming $\lambda, \eta \geqslant \sqrt{\xi}/2$ we get
$$\mathcal{W}^{\mathrm{p}}_{\lambda,\eta-\sqrt{\xi/2}} \subset \mathcal{W}_{\lambda,\eta}\,. \tag{4.4}$$
Set $\tilde{\eta} = \eta - \sqrt{\frac{\xi}{2}}$. The inclusion (4.4) implies that for all vector u,
$$\left\|\left(1 - \sum_{x\in\mathcal{W}_{\lambda,\eta}} \mathbf{u}_+(x)\mathbf{u}_+(x)^*\right)u\right\| \leqslant \|\overline{\Pi}^{\mathrm{p}}_{\lambda,\tilde{\eta}}u\|\,.$$
We have by Theorem 4.5 that with ν-high probability,
$$\left\|\left(1 - \sum_{x\in\mathcal{W}_{\lambda,\eta}} \mathbf{u}_+(x)\mathbf{u}_+(x)^*\right)\mathbf{q}\right\| \leqslant \|\overline{\Pi}^{\mathrm{p}}_{\lambda,\tilde{\eta}}\mathbf{q}\| \leqslant \frac{c_\nu\sqrt{\frac{\log N}{\log\log N}}}{\tilde{\eta}} = \frac{c_\nu\sqrt{\frac{\log N}{\log\log N}}}{\eta - \sqrt{\xi/2}}\,,$$
□

5. The biggest eigenvalues and the localization phenomenon

We now turn to the biggest eigenvalues. We show that they are close to the square roots of the degrees of some vertices in G. Furthermore, assuming that the average degrees d_x are well separated, we obtain the localization of some eigenvectors around single vertices.

Theorem 5.1. *There exists a constant $C_\nu > 0$ such that with ν-high probability, for all $i \in [N]$, if $\lambda_i(A) > C_\nu \sqrt{\frac{\log N}{\log\log N}}$, then*

$$\left|\lambda_i(A) - \sqrt{D_{\pi(i)}}\right| \leqslant C_\nu \sqrt{\frac{\log N}{\log\log N}}\,.$$

or if $\lambda_i(A) < -C_\nu \sqrt{\frac{\log N}{\log\log N}}$, then

$$\left|\lambda_i(A) + \sqrt{D_{\pi(i)}}\right| \leqslant C_\nu \sqrt{\frac{\log N}{\log\log N}}\,.$$

Proof. This is a consequence of Proposition 4.1. There exists $C'_\nu > 0$ such that with ν-high probability

$$|\lambda_i(A) - \lambda_i(\hat{A})| \leqslant \|A - \hat{A}\| \leqslant C'_\nu \sqrt{\frac{\log N}{\log\log N}}\,,$$

for all $i \in [N]$.

The i-th eigenvalue of $\hat{A}$ satisfies $\lambda_i(\hat{A}) = \pm\sqrt{D^{\mathrm{p}-}_{\pi(i)}}$ if $\lambda_i(\hat{A}) > \|\overline{\Pi}^{\mathrm{p}} A^{\mathrm{p}} \overline{\Pi}^{\mathrm{p}}\|$. Furthermore, Proposition 2.5 implies that with ν-high probability, for all $x \in \mathcal{V}^{(\mathrm{h})}$,

$$|D_x - D^{\mathrm{p}-}_x| \leqslant \xi/2\,.$$

Hence, there exists a constant C_ν such that with ν-high probability: for all i such that $\lambda_i(\hat{A}) > \|\overline{\Pi}^{\mathrm{p}} A^{\mathrm{p}} \overline{\Pi}^{\mathrm{p}}\|$,

$$|\lambda_i(A) - \sqrt{D_x}| \leqslant C_\nu \sqrt{\frac{\log N}{\log\log N}}\,.$$

We have a similar result when $\lambda_i(\hat{A}) < -\|\overline{\Pi}^{\mathrm{p}} A^{\mathrm{p}} \overline{\Pi}^{\mathrm{p}}\|$. □

We now consider the phenomenon of localization around a single vertex, in general a stronger result than the semilocalization. According to Theorem 4.5, it occurs when $\#\mathcal{W}_{\lambda,\eta} = 1$ for an appropriate pair (λ, η). We fix $\nu > 0$ and $\eta > 0$.

We introduce the set of isolated vertices:

$$\mathcal{V}^*_{\nu,\eta} = \left\{ x \in [N] : \begin{array}{l} \bullet\ \forall y \in [N], y \neq x, |d_x - d_y| \geqslant \left(4\sqrt{\nu \log N d_x} + 4\sqrt{\nu \log N d_y}\right) \vee 16\eta^2 \\ \bullet\ d_x \geqslant \dfrac{4\nu}{9} \log N \end{array} \right\}. \tag{5.1}$$

We shall show that the eigenvectors associated to the vertices in $\mathcal{V}^*_{\nu,\eta}$ are localized with ν-high probability.

Theorem 5.2. *There exists $C_\nu > 0$ such that with ν-high probability, for all eigenvalue $\lambda > C_\nu\sqrt{\frac{\log N}{\log\log N}}$ of A, with associated eigenvector $\mathbf{q}$, and all $\eta \leqslant \lambda/2$, we have the following property.*

If $\mathcal{W}_{\lambda,\eta} \cap \mathcal{V}^*_{\nu,\eta} \neq \emptyset$ then there exists $x \in \mathcal{V}^*_{\nu,\eta}$ such that

$$\langle \mathbf{q}, \mathbf{u}_+(x)\rangle^2 \geqslant 1 - \left(\frac{1}{\eta}\frac{C_\nu\sqrt{\log N}}{\sqrt{\log\log N}}\right)^2 .$$

Proof. We shall show that if $\mathcal{W}_{\lambda,\eta} \cap \mathcal{V}^*_{\nu,\eta} \neq \emptyset$, then $\#\mathcal{W}_{\lambda,\eta} = 1$. Theorem 5.2 is then a consequence of Theorem 4.5.

Let $x \in \mathcal{W}_{\lambda,\eta} \cap \mathcal{V}^*_{\nu,\eta}$ and $y \in \mathcal{W}_{\lambda,\eta}$ with $d_y \geqslant \frac{4\nu}{9}\log N$. Note that this second point holds with ν-high probability as soon as we take $C_\nu > 0$ big enough. Lemma 1.12 implies that with ν-high probability,

$$d_x - d_y - 2\sqrt{\nu\log N d_x} - 2\sqrt{\nu\log N d_y} \leqslant D_x - D_y \leqslant d_x - d_y + 2\sqrt{\nu\log N d_x} + 2\sqrt{\nu\log N d_y} .$$

Using (5.1), we have

$$d_x - d_y - \frac{1}{2}|d_x - d_y| \leqslant D_x - D_y \leqslant d_x - d_y + \frac{1}{2}|d_x - d_y| .$$

Notice that if $D_x > D_y$ then $d_x \geqslant d_y$ as otherwise we would have

$$D_x - D_y \leqslant -\frac{1}{2}|d_x - d_y| \leqslant 0 ,$$

a contradiction. We can show similarly that if $D_x < D_y$ then $d_x \leqslant d_y$. Thus, we have

$$|D_x - D_y| \geqslant \frac{1}{2}|d_x - d_y| . \tag{5.2}$$

Finally, if $x \neq y$, we have by (5.2) and (5.1):

$$|\sqrt{D_x} - \sqrt{D_y}| \geqslant \sqrt{\frac{|D_x - D_y|}{2}} \geqslant \frac{\sqrt{|d_x - d_y|}}{2} \geqslant 2\eta .$$

Thus, either $x = y$ or $y \notin \mathcal{W}_{\lambda,\eta}$. Hence, $\mathcal{W}_{\lambda,\eta}$ is a singleton. $\square$

Example 5.3. Consider Example 1.6. The weights are chosen as the quantiles of a law with heavy tail. Let $\alpha > 2$. We choose for simplicity

$$w_i = \left(\frac{N}{i}\right)^{1/\alpha} .$$

In that case, if $i \leqslant N^{1/(2\alpha+2)}$, we have

$$|w_i - w_{i-1}| \geqslant \left(\frac{N}{i}\right)^{1/\alpha}\frac{c}{i} ,$$

for some constant $c > 0$. As the sequence $(|w_{i+1} - w_i|)$ is increasing, we get that $\{1, \dots, \lfloor N^{1/(2\alpha+2)}\rfloor\} \subset \mathcal{V}^*_{\nu,\eta}$ with $\eta = \frac{\sqrt{\alpha}}{4}N^{1/(4\alpha)}$. On the other hand, Theorem 5.1 implies that for $i \leqslant \lfloor N^{1/(2\alpha+2)}\rfloor$,

$$\left|\lambda_i(A) - \sqrt{D_{\pi(i)}}\right| \ll \eta$$

with ν-high probability, that is for all such i,

$$\pi(i) \in \mathcal{W}_{\lambda_i(A),\eta}.$$

Bennett's inequality shows that $D_i > D_{i+1}$ with ν-high probability for $i \leqslant \lfloor N^{1/(2\alpha+2)} \rfloor$: thus $\pi(i) = i$ for such i's.

It then follows from Theorem 5.2 that with ν-high probability, the eigenvectors corresponding to the $N^{1/(2\alpha+2)}$ first eigenvalues are localized.

A. Bounds for the proof of Proposition 4.1

Consider the operator

$$B_2 = \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x} \sum_{\substack{y \in S_1^{\mathrm{p}-}(\hat{x}) \\ y \prec x}} \frac{\mathbf{1}_{S_1^{\mathrm{p}-}(x)} \mathbf{1}_y^*}{\sqrt{D_x^{\mathrm{p}-} \# \mathrm{Sib}^-(x)}}.$$

Its operator norm is

$$\begin{aligned} \|B_2\|^2 = \max_{\|\mathbf{u}\|=1} (\mathbf{u}^* B_2 B_2^* \mathbf{u}) &= \max_{\|\mathbf{u}\|=1} \sum_{x,x' \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{4 \left\langle \mathbf{u}, \mathbf{1}_{S_1^{\mathrm{p}-}(x)} \right\rangle \left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x')}, \mathbf{u} \right\rangle}{Z_x Z_{x'} \sqrt{D_x^{\mathrm{p}-} D_{x'}^{\mathrm{p}-} \# \mathrm{Sib}^-(x) \,\mathrm{Sib}_{x'}^-}} \sum_{\substack{y \in S_1^{\mathrm{p}-}(\hat{x}) \\ y \prec x \\ y' \in S_1^{\mathrm{p}-}(\hat{x}') \\ y' \prec x'}} \mathbf{1}_y^* \mathbf{1}_{y'} \\ &= \max_{\|\mathbf{u}\|=1} \sum_{x,x' \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{4 \left\langle \mathbf{u}, \mathbf{1}_{S_1^{\mathrm{p}-}(x)} \right\rangle \left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x')}, \mathbf{u} \right\rangle}{Z_x Z_{x'} \sqrt{D_x^{\mathrm{p}-} D_{x'}^{\mathrm{p}-}} \sqrt{\# \mathrm{Sib}^-(x) \,\mathrm{Sib}_{x'}^-}} \delta_{\hat{x}\hat{x}'} \\ &= \max_{\|\mathbf{u}\|=1} \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{4 \left\langle \mathbf{u}, \mathbf{1}_{S_1^{\mathrm{p}-}(x)} \right\rangle \left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(x)}, \mathbf{u} \right\rangle}{Z_x^2 D_x^{\mathrm{p}-} \# \mathrm{Sib}^-(x)}. \end{aligned}$$

At this point, we use the orthonormal family $(\mathbf{V}_1(x))_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} = \left(\frac{1}{\sqrt{D_x^{\mathrm{p}-}}} \mathbf{1}_{S_1^{\mathrm{p}-}(x)} \right)_{x \in \mathcal{V}_\nu^{(\mathrm{h})}}$. There exists a constant $C_\nu > 0$ such that $\frac{1}{\# \mathrm{Sib}^-(x)} \leqslant \frac{C_\nu}{\xi_\nu}$ so that

$$\|B_2\|^2 \leqslant \frac{C_\nu}{\xi_\nu} \max_{\|\mathbf{u}\|=1} \sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \langle \mathbf{u}, \mathbf{V}_1(x) \rangle^2 \leqslant \frac{C_\nu}{\xi_\nu}.$$

We now turn to the operator

$$B_3 = -\sum_{x \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x^2} \sum_{y \in \mathrm{Sib}^-(x)} \frac{\mathbf{1}_{S_1^{\mathrm{p}-}(y)} \mathbf{1}_x^*}{\# \mathrm{Sib}^-(x)}.$$

Its operator norm is

$$\|B_3\|^2 = \max_{\|\mathbf{u}\|=1} \mathbf{u}^* B_3^* B_3 \mathbf{u} = \sum_{x,x' \in \mathcal{V}_\nu^{(\mathrm{h})}} \frac{4 u_x u_{x'}}{Z_x Z_{x'} \# \mathrm{Sib}^-(x) \,\mathrm{Sib}_{x'}^-} \sum_{\substack{y \in \mathrm{Sib}^-(x) \\ y' \in \mathrm{Sib}^-(x')}} \left\langle \mathbf{1}_{S_1^{\mathrm{p}-}(y)}, \mathbf{1}_{S_1^{\mathrm{p}-}(y')} \right\rangle.$$

The orthogonality of the vectors $(\mathbf{1}_{S_1^{\mathrm{p}-}(y)})_{y\in[N]} = (\sqrt{D_y^{\mathrm{p}-}}V_1(y))_{y\in[N]}$ yields

$$\|B_3\|^2 = \sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{4\delta_{\hat{x}\hat{x}'}u_x u_{x'}}{Z_x Z_{x'}\#\operatorname{Sib}^-(x)\#\operatorname{Sib}^-(x')} \sum_{y\in\operatorname{Sib}^-(x)\cap\operatorname{Sib}^-(x')} D_y\,.$$

We apply Young's inequality to replace $\frac{u_x u_{x'}}{\#\operatorname{Sib}^-(x)\#\operatorname{Sib}^-(x')}$ by $\frac{u_x^2}{(\#\operatorname{Sib}^-(x))^2}$:

$$\|B_3\|^2 \leqslant \sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{4\delta_{\hat{x}\hat{x}'}u_x^2}{Z_x Z_{x'}\big(\#\operatorname{Sib}^-(x)\big)^2} \sum_{y\in\operatorname{Sib}^-(x)\cap\operatorname{Sib}^-(x')} D_y \leqslant \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{4u_x^2 D_{\hat{x}}}{Z_x\big(\#\operatorname{Sib}^-(x)\big)^2} \sum_{y\in\operatorname{Sib}^-(x)} D_y^{\mathrm{p}}\,.$$

Lemmas 3.7 and 3.6 give that with ν-high probability

$$\|B_3\|^2 \leqslant 8\nu \sum_{x,x'\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{4u_x^2}{Z_x}\frac{\log N}{\log\log N} = O\Big(\frac{\log N}{\log\log N}\Big)\,.$$

Finally, consider the operator

$$B_4 = \sum_{x\in\mathcal{V}_\nu^{(\mathrm{h})}} \frac{2}{Z_x} \sum_{y,z\in\operatorname{Sib}^-(x)} \frac{\mathbf{1}_{S_1^{\mathrm{p}-}(y)}\mathbf{1}_z^*}{(\#\operatorname{Sib}^-(x))^2}\,,$$

and define for $x\in\mathcal{V}_\nu^{(\mathrm{h})}$,

$$\mathbf{u}(x) = \frac{1}{\#\operatorname{Sib}^-(x)}\mathbf{1}_{\operatorname{Sib}^-(x)}$$
$$\mathbf{v}(x) = \frac{1}{\#\operatorname{Sib}^-(x)}\sum_{y\in\operatorname{Sib}^-(x)}\mathbf{1}_{S_1^{\mathrm{p}-}(y)}\,.$$

Notice that

$$B_4\mathbf{u}(x) = \frac{2}{Z_x(\#\operatorname{Sib}^-(x))^3}\sum_{y,z\in\operatorname{Sib}^-(x)}\mathbf{1}_{S_1^{\mathrm{p}-}(y)} = \frac{2}{Z_x(\#\operatorname{Sib}^-(x))^2}\sum_{y\in\operatorname{Sib}^-(x)}\mathbf{1}_{S_1^{\mathrm{p}-}(y)} = \frac{2}{Z_x\#\operatorname{Sib}^-(x)}\mathbf{v}(x)\,.$$

Furthermore,

$$\mathbf{v}(x)^*B_4 = \frac{2}{Z_x(\#\operatorname{Sib}^-(x))^3}\sum_{y,z\in\operatorname{Sib}^-(x)} D_y^{\mathrm{p}-}\mathbf{1}_z^* = \Bigg(\frac{2}{Z_x(\#\operatorname{Sib}^-(x))^2}\sum_{y\in\operatorname{Sib}^-(x)} D_y^{\mathrm{p}-}\Bigg)\mathbf{u}(x)^*\,.$$

Thus, Lemmas 3.6 and 3.7, and the Schur test imply that

$$\|B_4\| = O(1)\,.$$

B. Estimation of the size of $\mathcal{W}_{\lambda,\eta}$

Let $\lambda,\eta>0$ such that $2\sqrt{\frac{\log N}{\log\log N}} \leqslant \eta \leqslant \lambda/2$. We consider the expectation $\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}]$, that we rewrite as

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] = \sum_{\substack{x\in[N]\\ w_x\leqslant\sqrt{\log N}}} \mathbb{P}\Big((\lambda-\eta)^2 \leqslant D_x \leqslant (\lambda+\eta)^2\Big) + \sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \mathbb{P}\Big((\lambda-\eta)^2 \leqslant D_x \leqslant (\lambda+\eta)^2\Big)\,.$$

The first part can be bounded using Bennett's inequality [BLM13, Theorem 2.9]:

$$\mathbb{P}\Big(D_x \geqslant (\lambda-\eta)^2\Big) \leqslant \exp\Bigg(-(\lambda-\eta)^2\log(\frac{((\lambda-\eta)^2)}{d_x}) - (\lambda-\eta)^2 + d_x\Bigg) = O(N^{-2})\,,$$

so that

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] = \sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \mathbb{P}\Big((\lambda-\eta)^2 \leqslant D_x \leqslant (\lambda+\eta)^2\Big) + O(N^{-1})\,.$$

We start be recalling the following Lemma of approximation of the degrees by Poisson variables.

Lemma B.1 (Approximation of degrees by a Poisson variable [Hof16, Theorem 6.7])**.** *There exists a coupling $(\hat{D}_x, \hat{P}_x)$ of the degree D_x of vertex x and a Poisson variable x with parameter w_x, such that*

$$\mathbb{P}\Big(\hat{D}_x \neq \hat{P}_x\Big) \leqslant \frac{w_x^2}{m_1 N}\Big(1 + 2\frac{m_2}{m_1}\Big)\,.$$

This Lemma will be key in estimating $\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}]$ for some $\lambda, \eta > 0$. Indeed, we have

$$\begin{aligned}\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] &= \sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \mathbb{P}\Big((\lambda-\eta)^2 \leqslant D_x \leqslant (\lambda+\eta)^2\Big) O(N^{-1}) \\ &\leqslant \sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \mathbb{P}\Big((\lambda-\eta)^2 \leqslant \hat{P}_x \leqslant (\lambda+\eta)^2\Big) + \frac{m_2}{m_1}\Big(1 + 2\frac{m_2}{m_1}\Big) + O(N^{-1})\,.\end{aligned}$$

By Assumption 1.2, the last term is of order at most $(\log N)^{2/3}$. The first term can be written in term of incomplete Gamma functions

$$\Gamma(s,x) = \int_x^\infty t^{s-1}\mathrm{e}^{-t}\mathrm{d}t\,.$$

Indeed, we have

$$\sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \mathbb{P}\Big((\lambda-\eta)^2 \leqslant \hat{P}_x \leqslant (\lambda+\eta)^2\Big) = \sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \left(\frac{\Gamma(U_N, w_x)}{\Gamma(U_N)} - \frac{\Gamma(L_N - 1, w_x)}{\Gamma(L_N - 1)}\right),$$

where we set for convenience $L_N = \lfloor(\lambda-\eta)^2\rfloor$ and $U_N = \lceil(\lambda+\eta)^2\rceil$.

We shall use the two following properties of incomplete Gamma function.

Lemma B.2. *Let $s \geqslant 1$ and $x > 0$. Then, $\Gamma(s) - x^{s-1} \leqslant \Gamma(s,x) \leqslant \Gamma(s)$. Furthermore, $\Gamma(s,x) \sim x^{s-1}\mathrm{e}^{-x}$ as $x \to \infty$.*

Proof. It is immediate that $\Gamma(s,x) \leqslant \Gamma(s)$. For the other bound, we have

$$\Gamma(s,x) = \Gamma(s) - \int_0^x t^{s-1}\mathrm{e}^{-t}\mathrm{d}t \geqslant \Gamma(s) - x^{s-1}\int_0^x \mathrm{e}^{-t}\mathrm{d}t \leqslant \Gamma(s) - x^{s-1}\,.$$

To prove the asymptotic estimate, we remark that

$$\frac{\Gamma(s,x)}{x^{s-1}\mathrm{e}^{-x}} = \int_x^\infty \left(\frac{t}{x}\right)^{s-1}\mathrm{e}^{-(t-x)}\mathrm{d}t = \int_0^\infty \left(\frac{t}{x}+1\right)^{s-1}\mathrm{e}^{-t}\mathrm{d}t\,.$$

The dominated convergence theorem then implies that the limit of the left-hand term is 1 as $x \to \infty$. ☐

We kept in the sum only terms x such that $w_x \to \infty$ as $N \to \infty$, hence by Lemma B.2, we have

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] \leqslant \sum_{\substack{x\in[N]\\ w_x>\sqrt{\log N}}} \mathrm{e}^{-w_x}\left(\frac{w_x^{U_N-1}}{(U_N-1)!} - \frac{w_x^{L_N-2}}{(L_N-2)!}\right)(1+o(1)) + O((\log N)^{2\delta})\,.$$

We now consider two cases:

- The weights (w_x) are the $(N+1)$-quantiles of an exponential law as in Example 1.5.
- The weights (w_x) are the $(N+1)$-quantiles of a law with heavy tails as in Example 1.6.

In the exponential case, we then have

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] \leqslant N\int_0^\infty \alpha\mathrm{e}^{-\alpha t-t}\left(\frac{t^{U_N-1}}{(U_N-1)!} - \frac{t^{L_N-2}}{(L_N-2)!}\right)\mathrm{d}t(1+o(1)) + O((\log N)^{2\delta})\,.$$

Using that the k-th moment of an exponential law of parameter $\alpha+1$ is $k!/(\alpha+1)^k$ we get

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] \leqslant N\frac{\alpha}{\alpha+1}\left(\frac{1}{(\alpha+1)^{L_N-2}} - \frac{1}{(\alpha+1)^{U_N-1}}\right)(1+o(1)) + O((\log N)^{2\delta})\,.$$

That is,

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] \leqslant N\frac{\alpha}{(\alpha+1)^{\lfloor(\lambda-\eta)^2\rfloor-1}}(1+o(1)) + O((\log N)^{2\delta})\,. \tag{B.1}$$

In the heavy tail case, we denote the heavy tail measure by μ and recall that

$$\mu([t,+\infty)) = L(t)t^{-\alpha},$$

where L is a positive, slowly varying function. We notice that

$$\int_0^\infty \frac{t^{U_N-1}}{(U_N-1)!}\mathrm{e}^{-t}\mathrm{d}\mu(t) = \int_0^\infty L(t)t^{-\alpha}\left(\frac{t^{U_N-2}}{(U_N-2)!} - \frac{t^{U_N-1}}{(U_N-1)!}\right)\mathrm{e}^{-t}\mathrm{d}t.$$

Then, for all $\iota>0$, there exists $C>0$ such that $0\leqslant L(t)\leqslant Ct^{\iota}$. Note that if L is bounded we can choose $\iota=0$. We observe that on $(0,\frac{U_N-2}{\mathrm{e}})$, the integrand is positive. Hence, we have

$$\int_0^\infty L(t)t^{-\alpha}\left(\frac{t^{U_N-2}}{(U_N-2)!} - \frac{t^{U_N-1}}{(U_N-1)!}\right)\mathrm{e}^{-t}\mathrm{d}t \leqslant C\int_0^{+\infty}\left(\frac{t^{U_N-\alpha-2+\iota}}{(U_N-2)!} - \frac{t^{U_N-\alpha-1+\iota}}{(U_N-1)!}\right) + O\left(\mathrm{e}^{-\frac{U_N-2}{2}}\right),$$

where the error $O(\mathrm{e}^{-\frac{U_N-2}{2}})$ is a bound for the integral on $(\frac{U_N-2}{\mathrm{e}},+\infty)$, with and without L. Finally, we have

$$\int_0^\infty \frac{t^{U_N-1}}{(U_N-1)!}\mathrm{e}^{-t}\mathrm{d}\mu(t) \leqslant C\left(\frac{\Gamma(U_N-\alpha-1+\iota)}{\Gamma(U_N-1)} - \frac{\Gamma(U_N-\alpha+\iota)}{\Gamma(U_N)}\right).$$

Lemma B.3 (Stirling's approximation)**.** *Let $t>0$. We have as $x\to\infty$*

$$\frac{\Gamma(x-t)}{\Gamma(x)} = \frac{1}{x^t}\mathrm{e}^{\frac{-t}{2x}+o(1/x)}\,.$$

Proof. Stirling approximation to the first order

$$\log\Gamma(x) = x\log x - x + \frac{1}{2}\log\frac{2\pi}{x} + \frac{12}{x} + O\Big(\frac{1}{x^2}\Big)$$

gives

$$\begin{aligned}\log\frac{\Gamma(x-t)}{\Gamma(x)} &= -t\log x - (x-t)\Big(\frac{t}{x} + o\Big(\frac{1}{x}\Big)\Big) + t + \frac{1}{2}\Big(\frac{t}{x} + o\Big(\frac{1}{x}\Big)\Big) + o\Big(\frac{1}{x}\Big) \\ &= -t\log x + \frac{t}{2x} + o\Big(\frac{1}{x}\Big).\end{aligned}$$

□

Lemma B.3 immediately gives

$$\frac{\Gamma(U_N - 1 - \alpha + \iota)}{\Gamma(U_N - 1)} = \frac{1}{(U_N-1)^{\alpha-\iota}} \mathrm{e}^{\frac{t}{2U_N} + o(1/U_N)} = \frac{1}{U_N^{\alpha-\iota}}\Big(1 + \frac{\alpha-\iota}{U_N} + o(\frac{1}{U_N})\Big).$$

Similarly, we have

$$\frac{\Gamma(U_N - \alpha + \iota)}{\Gamma(U_N)} = \frac{1}{U_N^{\alpha-\iota}}\Big(1 + o(\frac{1}{U_N})\Big).$$

This means that

$$\int_0^\infty \frac{t^{U_N-1}}{(U_N-1)!}\mathrm{e}^{-t}\mathrm{d}\mu(t) = \frac{\alpha-\iota}{U_N^{\alpha-\iota+1}} + o\Big(\frac{1}{U_N^{\alpha+1-\iota}}\Big).$$

Similarly,

$$\int_0^\infty \frac{t^{L_N-2}}{(L_N-2)!}\mathrm{e}^{-t}\mathrm{d}\mu(t) = \frac{\alpha-\iota}{L_N^{\alpha-\iota+1}} + o\Big(\frac{1}{L_N^{\alpha+1-\iota}}\Big).$$

Putting everything together, we get that there is a constant $C > 0$ such that

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] \leqslant C\bigg(\frac{\alpha-\iota}{L_N^{\alpha-\iota+1}} - \frac{\alpha-\iota}{U_N^{\alpha-\iota+1}}\bigg)\Big(1 + o(\frac{1}{L_N})\Big) + O\Big((\log N)^{2\delta}\Big).$$

This means

$$\mathbb{E}[\#\mathcal{W}_{\lambda,\eta}] = O\Big(\frac{N}{\lambda^{2\alpha+3-2\iota}}\Big) + O\Big((\log N)^{2\delta}\Big). \tag{B.2}$$

We now turn to the proof of Proposition 1.9. The proof will use the following variant of Lemma 2.18.

Lemma B.4. *For each vertex $x \in [N]$, define*

$$\hat{D}_x^+ = \#\{y \in S_1(x) : w_y \geqslant w_x\} = \sum_{y \neq x} \mathbb{1}_{\{x\sim y, w_x \leqslant w_y\}}.$$

Let $\nu > 0$. With ν-high probability, we have

$$\hat{D}_x^+ \leqslant \frac{2\nu}{1-\delta}\frac{\log N}{\log\log N}.$$

Proof of Lemma B.4. Let $k \geqslant 1$ be an integer. The union bound implies

$$\mathbb{P}\Big(\hat{D}_x^+ \geqslant k\Big) \leqslant \frac{1}{k!} \sum_{\substack{x_1,\dots,x_k \\ \text{distinct}}} \mathbb{P}(\forall i \in [k], x \sim x_i, w_x \leqslant w_{x_i})\,.$$

By independence and (1.4), we have

$$\mathbb{P}\Big(\hat{D}_x^+ \geqslant k\Big) \leqslant \frac{1}{k!} \sum_{\substack{x_1,\dots,x_k \\ \text{distinct}}} \prod_{i=1}^{k} \left(\frac{w_x w_{x_i}}{m_1 N} \mathbb{1}_{\{w_x \leqslant w_{x_i}\}}\right) \leqslant \frac{1}{k!} \sum_{\substack{x_1,\dots,x_k \\ \text{distinct}}} \prod_{i=1}^{k} \left(\frac{w_{x_i}^2}{m_1 N}\right),$$

where in the last line, we used that $\mathbb{1}_{\{w_x \leqslant w_{x_i}\}} \leqslant w_{x_i}/w_x$. By definition of the second empirical moment, we have

$$\mathbb{P}\Big(\hat{D}_x^+ \geqslant k\Big) \leqslant \frac{1}{k!} \left(\frac{m_2}{m_1}\right)^k (1 + o(1))\,.$$

Taking $k = \lfloor \frac{2\nu}{1-\delta} \frac{\log N}{\log\log N} \rfloor$ allows us to conclude. □

Proof of Proposition 1.9. Let $k \geqslant 1$ be an integer. We use the union bound to write

$$\mathbb{P}(\#\mathcal{W}_{\lambda,\eta} \geqslant k) \leqslant \frac{1}{k!} \sum_{\substack{x_1,\dots,x_k \\ \text{distinct}}} \mathbb{P}\Big(\forall i \in [k], (\lambda - \eta)^2 \leqslant D_{x_i} \leqslant (\lambda + \eta)^2\Big)\,.$$

Set $\hat{D}_x^- = D_x - \hat{D}_x^+$ for all $x \in [N]$. By Lemma B.4, we have for all $x \in [N]$ that

$$D_x = \hat{D}_x^+ + \hat{D}_x^- \leqslant \hat{D}_x^- + \frac{2\nu + 2}{1 - \delta} \frac{\log N}{\log\log N}\,,$$

with ν-high probability. For convenience, write $c_{\nu,N} = \frac{2\nu+2}{1-\delta} \frac{\log N}{\log\log N}$.

Now, notice that the random variables $\hat{D}_x^-$ are independent. It implies

$$\begin{aligned}
\mathbb{P}(\#\mathcal{W}_{\lambda,\eta} \geqslant k) &\leqslant \frac{1}{k!} \sum_{\substack{x_1,\dots,x_k \\ \text{distinct}}} \prod_{i=1}^{k} \mathbb{P}\Big((\lambda - \eta)^2 - c_{\nu,N} \leqslant \hat{D}_{x_i}^- \leqslant (\lambda + \eta)^2\Big) + O(N^{-\nu}) \\
&\leqslant \frac{1}{k!} \left(\sum_x \mathbb{P}\Big((\lambda - \eta)^2 - c_{\nu,N} \leqslant \hat{D}_x^- \leqslant (\lambda + \eta)^2\Big)\right)^k + O(N^{-\nu})\,.
\end{aligned}$$

Notice that we have

$$\begin{aligned}
\mathbb{P}(\#\mathcal{W}_{\lambda,\eta} \geqslant k) &\leqslant \frac{1}{k!} \left(\sum_x \mathbb{P}\Big((\lambda - \eta)^2 - c_{\nu,N} \leqslant D_x \leqslant (\lambda + \eta)^2 + c_{\nu,N}\Big)\right)^k + O(N^{-\nu}) \\
&\leqslant \frac{\Big(\mathbb{E}\#\mathcal{W}_{\lambda,\eta + c_{\nu,N}/2\lambda}\Big)^k}{k!} + O(N^{-\nu})\,.
\end{aligned}$$

Markov's inequality gives the crude bound

$$\mathbb{P}(\#\mathcal{W}_{\lambda,\eta} \geqslant k) \leqslant \frac{1}{k!} \left(\sum_x \frac{w_x^2}{((\lambda - \eta)^2 - c_{\nu,N})^2}\right)^k = \frac{1}{k!} \left(N \frac{m_2}{((\lambda - \eta)^2 - c_{\nu,N})^2}\right)^k.$$

Choosing $k = \lfloor \frac{2m_2}{(\lambda-\eta)^4} N \vee \frac{2\nu \log N}{\log\log N} \rfloor$ gives the result. The result can be improved in our two examples using expressions (B.1) and (B.2), derived above. □

Acknowledgements. TBdA was supported by ERC Project LDRAM : ERC-2019-ADG Project 884584 in the initial stages of this projet, and by ERC Project InSpeGMos 101096550 in the late stages of this project.

References

[ADK21a] J. Alt, R. Ducatez, and A. Knowles. Delocalization transition for critical Erdős-Rényi graphs. *Comm. Math. Phys.*, 388(1):507–579, 2021.

[ADK21b] J. Alt, R. Ducatez, and A. Knowles. Extremal eigenvalues of critical Erdős-Rényi graphs. *Ann. Prob.*, 49(3):1347–1401, 2021.

[BGBK20] Florent Benaych-Georges, Charles Bordenave, and Antti Knowles. Spectral radii of sparse random matrices. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 56(3):2141–2161, 2020.

[BLM13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: A nonasymptotic theory of independence.* Oxford university press, 2013.

[CL02] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, 6(2):125–145, 2002.

[Hof16] Remco van der Hofstad. *Random Graphs and Complex Networks.* Cambridge University Press, 2016.