Shotgun Threshold for Sparse Erdős–Rényi Graphs

Jian Ding, Yiyang Jiang, and Heng Ma

***Abstract*— In the shotgun assembly problem for a graph, we are given the empirical profile for rooted neighborhoods of depth r (up to isomorphism) for some $r \geq 1$ and we wish to recover the underlying graph up to isomorphism. When the underlying graph is an Erdős–Rényi $\mathcal{G}(n, \frac{\lambda}{n})$, we show that the shotgun assembly threshold r_* satisfies that $r_* \approx \frac{\log n}{\log(\lambda^2\gamma_\lambda)^{-1}}$ where γ_λ is the probability for two independent Poisson–Galton–Watson trees with parameter λ to be rooted isomorphic with each other. Our result sharpens a constant factor in a previous work by Mossel and Ross (2019) and thus solves a question therein.**

***Index Terms*— Shotgun assembly, Erdős–Rényi graphs, phase transition, Poisson–Galton–Watson trees.**

I. Introduction

THE shotgun assembly problems aim for recovering a global structure from local observations; these problems have substantial interests in applications such as DNA sequencing [3], [11], [27] and recovering neural networks [19]. In [25], the precise formulation and general mathematical framework was proposed together with numerous inspiring open questions on a number of concrete shotgun models. Since (the circulation of) [25], there has been extensive study on shotgun assembly questions including on random jigsaw problems [5], [9], [18], [24], on random coloring models [29], on some extension of DNA sequencing models [30] and on lattice labeling models [10].

An example of shotgun assembling problems of particular interest is for random graph models [1], [13], [17], [26]. For random regular graphs with fixed degree, the asymptotic shotgun threshold was implied in [7] and was improved in [26] to the precision of up to additive constant. For Erdős–Rényi graphs with polynomially growing average degree, it was determined in [13] and [17] whether recovery is possible from neighborhoods of depth 1. For Erdős–Rényi graphs with constant average degree, the shotgun threshold was known to have order $\log n$ from [25], and the main contribution of this paper is to determine its sharp asymptotics. (In fact, order $\log n$ was obtained except for the critical case, i.e, when the average degree is 1, where only a polynomial upper bound was obtained in [25].)

Manuscript received 2 October 2022; revised 26 May 2023; accepted 19 July 2023. Date of publication 24 July 2023; date of current version 20 October 2023. The work of Jian Ding was supported in part by the NSFC Key Program under Project 12231002. *(Corresponding author: Jian Ding.)*

The authors are with the School of Mathematical Sciences, Peking University, Beijing 100871, China (e-mail: dingjian@math.pku.edu.cn; jyy.0916@stu.pku.edu.cn; maheng@stu.pku.edu.cn).

Communicated by X. Wang, Associate Editor for Networking and Computation.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2023.3298515.

Digital Object Identifier 10.1109/TIT.2023.3298515

Before stating our result, we first define our model more formally. Fix $\lambda > 0$. Let $\mathcal{G} \sim \mathcal{G}(n, \frac{\lambda}{n})$, i.e., let $\mathcal{G}$ be an Erdős–Rényi graph on n vertices where there is an edge between each unordered pair of vertices independently with probability λ/n. For $v \in \mathcal{G}$ and $r \geq 1$, let $\mathsf{N}_r(v)$ be the depth r-neighborhood rooted at v viewed modulo isomorphism (equivalently, all other vertices in $\mathsf{N}_r(v)$ except v are unlabeled). A couple of comments are in order: (1) here $\mathsf{N}_r(v)$ contains all vertices whose graph distances to v are at most r and all edges among these vertices; (2) here the notion of isomorphism for rooted graphs is as follows: a graph $G = (V, E)$ rooted at o is isomorphic to a graph $G' = (V', E')$ rooted at o' (denoted as $G \sim G'$) if there exists a bijection $\phi : V \mapsto V'$ with $\phi(o) = o'$ such that $(u, v) \in E$ if and only if $(\phi(u), \phi(v)) \in E'$. In the shotgun assembly problem, we are given the empirical profile for all rooted depth r-neighborhoods, i.e., we are given $\{\mathsf{N}_r(v) : v \in \mathcal{G}\}$ and we wish to recover $\mathcal{G}$ up to isomorphism. We say the problem is *non-identifiable* if there exist two graphs which are not isomorphic and both have empirical neighborhood profile $\{\mathsf{N}_r(v) : v \in \mathcal{G}\}$; otherwise we say the problem is *identifiable*.

Theorem 1: Fix $\lambda > 0$. Define

$$\gamma_\lambda = \mathbb{P}(\mathbf{T} \sim \mathbf{T}') \tag{I.1}$$

where $\mathbf{T}, \mathbf{T}'$ are two independent Poisson–Galton–Watson trees with parameter λ (where the root of the tree is naturally the initial ancestor). Let $\mathcal{G} \sim \mathcal{G}(n, \frac{\lambda}{n})$. Then the following hold for any fixed $\epsilon_0 > 0$:

(i) for $r \leq \frac{(1-\epsilon_0)\log n}{\log(\lambda^2\gamma_\lambda)^{-1}}$, the shotgun problem is non-identifiable with probability tending to 1 as $n \to \infty$;

(ii) for $r \geq \frac{(1+\epsilon_0)\log n}{\log(\lambda^2\gamma_\lambda)^{-1}}$, the shotgun problem is identifiable with probability tending to 1 as $n \to \infty$ and in addition $\mathcal{G}$ can be recovered via a polynomial time algorithm.

Remark 2: The function $\lambda \mapsto (\lambda^2\gamma_\lambda)$ is continuous on $(0, \infty)$, and is increasing on $(0, 1]$ as well as decreasing on $[1, \infty)$. See Lemma 4.

From the statement of Theorem 1, we see that a key novelty in this work is a connection to the isomorphic probability for Poisson–Galton–Watson (PGW) trees (isomorphism between random trees has been recently studied in [28] although the isomorphic probabilities considered in [28] are different from what we need here). Therefore, the driving mechanism for the shotgun threshold of random regular graphs and Erdős–Rényi graphs is completely different: as argued in [26], for random regular graphs "tree neighborhoods are all alike, but every non-tree neighborhood is filled with cycles in its own way" and as a result cycle structures are essential in distinguishing neighborhoods in regular graphs; in the contrast, for Erdős–Rényi graphs, local neighborhoods behave like PGW trees and loosely speaking these trees will all look different from each other with a suitably chosen depth.

That being said, cycles do appear in some of the neighborhoods of Erdős–Rényi graphs and potentially this may incur an issue for our scheme of approximation by trees. The good news is that, by (II.4) a typical neighborhood with depth near the threshold is a tree so that conceptually our reduction to isomorphism between PGW trees is valid. However, on the technical side, the occurrence of cycles forms a substantial obstacle for the analysis which is why our proof for identifiability is fairly involved. For non-identifiability, while we do investigate neighborhoods with depth twice the critical threshold, we focus on some special type of structures (see, e.g., (II.14)) which are trees.

Finally, it is non-trivial how exactly the isomorphic probability is related to the shotgun threshold. As we will see in Section II-A, what is of fundamental importance to us is the probability $\mathfrak{p}_r$ that two independent PGW trees survive for at least r levels and at the same time their first r levels are isomorphic with each other (see (II.1)): this probability decays exponentially in r where the exponential rate is governed by γ_λ (see Lemma 5). In order to further elaborate this connection, we feel it would be more useful to simply present the proof for non-identifiability (as incorporated in Section II), where tree isomorphism plays a role in constructing blocking configurations (which is inspired by and also an improvement upon what was considered in [25]). After defining the blocking configurations, we can then compute that the expected number of such blocking configurations is of order

$$n^2\mathfrak{p}_{2r} \approx n^2(\lambda^2\gamma_\lambda)^{2r}.$$

It is then natural to suspect that our recovery threshold is such that the above expectation tends to infinity when r is above the threshold, and tends to 0 when r is below the threshold. However, on the one hand, the expectation tending to infinity on its own does not guarantee the existence of a blocking configuration, and this will be made rigorous via a second moment computation as in Section II. On the other hand, the non-existence of our blocking configurations does not directly imply the identifiability. The proof for the reconstruction, as mentioned above, is substantially more challenging and will be carried out in Sections III and IV.

A. Notation Convention

We denote by $\mathbb{N}$ the collection of all natural numbers. We use *with high probability* for with probability tending to 1 as $n \to \infty$. For non-negative sequences f_n and g_n, we write $f_n \lesssim g_n$ if there exists a constant $C > 0$ such that $f_n \le C g_n$ for all $n \ge 1$. We write $f_n \lesssim_\lambda g_n$ in order to stress that the constant C depends on λ. For events A, B and random variable X, we write $\mathbb{P}(A, B)$ or $\mathbb{P}(A; B)$ for $\mathbb{P}(A \cap B)$, and write $\mathbb{E}[X; B]$ for $\mathbb{E}[X \cdot 1_B]$. We use $\mathrm{Bin}(n, p)$ to denote a binomial distribution/variable with parameter (n, p) and we write $X \sim \mathrm{Bin}(n, p)$ if X is a binomial variable with parameter (n, p). As we will reiterate in the main text, we assume that there is a pre-fixed (but arbitrarily chosen) ordering on $V = V(\mathcal{G})$. We write $\mathcal{G}_{uv} = 1$ if and only if there is an edge in $\mathcal{G}$ between u and v. For a graph G, we let $\mathrm{Comp}(G)$ be the *complexity*, that is, $\mathrm{Comp}(G)$ is the minimal number of edges that one has to remove from G so that no cycle remains. For a rooted tree T, we say it survives ℓ levels if its ℓ-th level contains at least one vertex. For a rooted depth r-neighborhood $\mathsf{N}_r(v)$, we say $\mathsf{N}_r(v)$ survives if $\mathsf{N}_{r-1}(v) \neq \mathsf{N}_r(v)$. Also, we write depth r-neighborhood as r-neighborhood for short. Let $\mathbf{T}$ and $\mathbf{T}'$ be two independent Poisson–Galton–Watson trees with parameter λ (PGW(λ)-trees). We summarize the various quantities of PGW-trees in a glossary.

- $\gamma_\lambda = \mathbb{P}(\mathbf{T} \sim \mathbf{T}')$: The probability that $\mathbf{T}, \mathbf{T}'$ are (rooted) isomorphic.
- $g_r = \mathbb{P}(\mathbf{T}|_r \sim_r \mathbf{T}'|_r)$: The probability that first r levels of $\mathbf{T}, \mathbf{T}'$ are (rooted) isomorphic.
- $\mathfrak{p}_r = \mathbb{P}(\mathbf{T} \sim_r \mathbf{T}')$: The probability that $\mathbf{T}, \mathbf{T}'$ survive r levels and $\mathbf{T}|_r \sim_r \mathbf{T}|'_r$.
- $\alpha_\lambda = \lambda^2\gamma_\lambda$: Lemma 5 shows that $\exists c, C > 0$ s.t. $c\alpha_\lambda^r \le \mathfrak{p}_r \le C\alpha_\lambda^r$

II. Proof of Non-Identifiability

As mentioned earlier, we need a particular version of isomorphic probability which we now introduce. Along the way, we also introduce some useful notations.

- For a rooted tree T, we denote by $|T|$ the size of T, and by $H(T)$ the height of tree T. For each positive integer r, let $T|_r$ be the restriction of the tree T to its first r levels. For an individual $u \in T$, we denote by $|u|$ the level of u, i.e., the distance from u to the root of T.
- Let T, T' be two rooted trees. Recall that we have defined isomorphism between T and T'. In addition, for each $r \in \mathbb{N} \cup \{\infty\}$, we write $T \sim_r T'$ if $\min\{H(T), H(T')\} \ge r$ and $T|_r \sim T'|_r$; this is the key notion of isomorphism for our analysis later.
- Let $\mathbf{T}$ and $\mathbf{T}'$ be two independent Poisson–Galton–Watson trees with parameter λ (PGW(λ)-trees). Define

$$\mathfrak{p}_r = \mathfrak{p}_r(\lambda) = \mathbb{P}\left(\mathbf{T} \sim_r \mathbf{T}'\right), \text{ for } r \in \mathbb{N} \cup \{\infty\}. \quad \text{(II.1)}$$

The importance of $\mathfrak{p}_r$ lies in the following fact: if there exist two isomorphic $2r$-neighborhoods which are disjoint trees, then there should also exist v, u such that their $2r$-neighborhoods are two disjoint isomorphic trees with some decoration (see Figure 1 where the line segment between w and w' in the figure is the decoration) and also with some additional structural properties (in fact, we also will need to adjust the value of r slightly in order to pose the additional structural properties). In this case, we can then construct two non-isomorphic graphs which have the same empirical profile for r-neighborhoods.

A. Isomorphism for Galton-Watson Trees

In this subsection we prove a number of lemmas on isomorphism of PGW-trees. Unless specified otherwise, we will denote by $\mathbf{T}, \mathbf{T}'$ two independent PGW(λ) trees.

Lemma 3: The sequence $(\mathfrak{p}_r)_{r\ge1}$ is non-increasing in r, and $\mathfrak{p}_r \to \mathfrak{p}_\infty = 0$ as $r \to \infty$.

Proof: The monotonicity and convergence are obvious since $\{\mathbf{T} \sim_{r+1} \mathbf{T}'\} \subset \{\mathbf{T} \sim_r \mathbf{T}'\}$ and $\cap_{r=1}^\infty\{\mathbf{T} \sim_r \mathbf{T}'\} = \{\mathbf{T} \sim_\infty \mathbf{T}'\}$. It remains to show that $\mathfrak{p}_\infty = 0$.

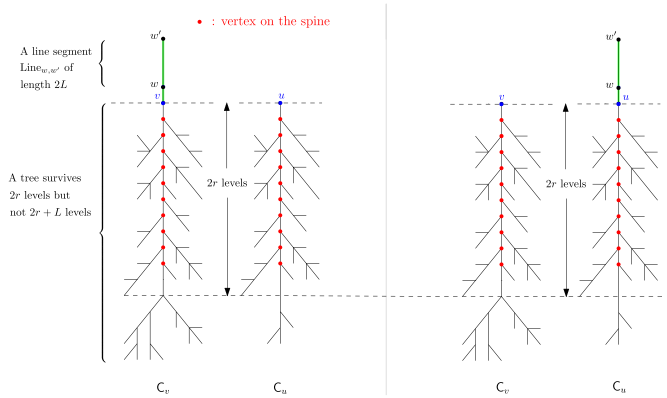


Fig. 1. On the left hand side is an example of blocking subgraph $(\mathsf{C}_v, \mathsf{C}_u)$. The pair $(\mathsf{C}'_v, \mathsf{C}'_u)$ on the right hand side is obtained by cutting Line$_{w,w'}$ from C_v and attaching it to u (that is, in C_u). Both graphs (i.e., $(\mathsf{C}_v, \mathsf{C}_u)$ and $(\mathsf{C}'_v, \mathsf{C}'_u)$) have the same profile for r-neighborhoods.

Let Z_n and Z'_n be the numbers of the vertices in the level n of the tree $\mathbf{T}$, $\mathbf{T}'$ respectively. Then applying the Kesten-Stigum theorem [20] (see also [23]) we have that

$$\begin{aligned}&\mathbb{P}(\mathbf{T} \sim \mathbf{T}' \text{ and } |H(\mathbf{T})| = |H(\mathbf{T}')| = \infty)\\ &\le \mathbb{P}(Z_n = Z'_n \ge 1, \forall\, n \ge 1) \le \mathbb{P}(W = W' > 0)\,,\end{aligned}$$

where W, W' are the limits of L^2-bounded martingales Z_n/λ^n and Z'_n/λ^n. By [15] (see also [16, Theorem 8.3] and [32]) W has a probability density on the set $\{W > 0\}$. Since W and W' are independent, we get that $\mathbb{P}(W = W' > 0) = 0$. $\square$

We next estimate the decay rate for $\mathfrak{p}_r$. To this end, define

$$\alpha_\lambda = \lambda^2 \gamma_\lambda \text{ and } q_\lambda = \mathbb{P}(|\mathbf{T}| < \infty)\,. \tag{II.2}$$

We claim that

$$q_\lambda < \lambda^{-2} \text{ for } \lambda > 1\,. \tag{II.3}$$

There is nothing original in (II.3) and we supply a proof merely for completeness. It is well-known (by the method of conditioning on the number of children for the root, see e.g. [16, Theorem 6.1]) that q_λ is the minimal zero for the equation $x = e^{-\lambda + \lambda x}$. Therefore, (II.3) can be reduced to $\exp(-\lambda + \lambda^{-1}) < \lambda^{-2}$. Let $f(\lambda) = \lambda^2 \exp(-\lambda + \lambda^{-1})$, then

$$f'(\lambda) = -(\lambda - 1)^2 \exp(-\lambda + \lambda^{-1}) < 0\,, \text{ for all } \lambda > 1\,.$$

Thus $f(\lambda) < f(1) = 1$ for all $\lambda > 1$, completing the verification. The following lemma will be useful in controlling $\mathfrak{p}_r$ in Lemma 5.

Lemma 4: For any $\lambda > 0$, we have $\alpha_\lambda = \alpha_{\lambda q_\lambda} < 1$. Furthermore, there is a power series A with non-negative coefficients such that $\alpha_\lambda = A(\lambda e^{-\lambda})$.

Proof: Since $\mathbb{P}(\mathbf{T} \sim_\infty \mathbf{T}') = 0$ (by Lemma 3), we have

$$\gamma_\lambda = \mathbb{P}(\mathbf{T} \sim \mathbf{T}') = \mathbb{P}(\mathbf{T} \sim \mathbf{T}' \,\big|\, |\mathbf{T}|, |\mathbf{T}'| < \infty) q_\lambda^2\,.$$

It is well-known that the law of $\mathbf{T}$ under $\mathbb{P}(\cdot \,\big|\, |\mathbf{T}| < \infty)$ is the same as the law for PGW(λq_λ). Since $\mathbf{T}, \mathbf{T}'$ are independent, we get $\mathbb{P}(\mathbf{T} \sim \mathbf{T}' \,\big|\, |\mathbf{T}|, |\mathbf{T}'| < \infty) = \gamma_{\lambda q_\lambda}$. This implies that $\gamma_\lambda = \gamma_{\lambda q_\lambda} q_\lambda^2$ and $\alpha_\lambda = \alpha_{\lambda q_\lambda}$. In addition, note $\gamma_\lambda \in (0,1)$ for all $\lambda > 0$. Thus, $\alpha_\lambda < 1$ for all $\lambda \le 1$. For $\lambda > 1$, thanks to (II.3) we have $\gamma_\lambda < q_\lambda^2 \le \frac{1}{\lambda^4}$, implying $\alpha_\lambda < 1$.

We now prove the second assertion. By Lemma 3, we have $\gamma_\lambda = \sum_\tau \mathbb{P}(\mathbf{T} \sim \tau)^2$, where τ is summed over all the equivalence classes for finite rooted ordered trees and the equivalence is given by rooted isomorphism. We set $b(\tau) = \prod_{j=1}^{|\tau|} b_j(\tau)!$, where $\{b_j(\tau) : j = 1, \ldots, |\tau|\}$ is the collection of the numbers of children for vertices in the tree τ. Then we have

$$\mathbb{P}(\mathbf{T} \sim \tau) = \frac{\#\tau}{b(\tau)} e^{-\lambda|\tau|} \lambda^{|\tau|-1}\,,$$

where $\#\tau = \#\{T \text{ is a rooted ordered tree} : T \sim \tau\}$. Let

$$A(s) = \sum_\tau \left[\frac{\#\tau}{b(\tau)} s^{|\tau|}\right]^2 \text{ for all } s \in [0, e^{-1}]\,.$$

Since $A(e^{-1}) = \gamma_1 < 1$, we see that A is well-defined on $[0, e^{-1}]$. Then a simple computation yields that $\alpha_\lambda = A(\lambda e^{-\lambda})$. $\square$

By Lemma 4 and the fact that $\lambda e^{-\lambda}$ is increasing in λ on $[0, 1]$ and decreasing in λ on $[1, \infty)$, we obtain the desired monotonicity of α_λ as in Remark 2. In addition, for $\lambda > 1$, we have from Lemma 4, (II.2) and (II.3) that $\alpha_\lambda = (\lambda q_\lambda)^2 \gamma_{\lambda q_\lambda} < \lambda^{-2}$, and hence

$$\frac{1}{\log(1/\alpha_\lambda)} < \frac{1}{2\log(\lambda)} \text{ for } \lambda > 1. \tag{II.4}$$

This implies that when the depth r is near the shotgun threshold, a typical r-neighborhood has at most $O(n^{1/2-\delta})$ vertices for some $\delta > 0$ and as a result is a tree (see Lemma 24).

For convenience, we let μ_k's be Poisson probabilities given by

$$\mu_k = \mu_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \text{ for } k \ge 0\,. \tag{II.5}$$

The following simple identity will be used repeatedly in our proof:

$$k\mu_k = \lambda \mu_{k-1} \text{ for all } k \ge 1\,. \tag{II.6}$$

Lemma 5: We have $\mathfrak{p}_r \asymp \alpha_\lambda^r$, i.e., there exist two constants $c, C > 0$ (possibly depending on λ) such that $c\alpha_\lambda^r \le \mathfrak{p}_r \le C\alpha_\lambda^r$.

Proof: For two independent PGW(λ) trees $\mathbf{T}$ and $\mathbf{T}'$, define

$$g_r = \mathbb{P}(\mathbf{T}|_r \sim \mathbf{T}'|_r)\,. \tag{II.7}$$

Note that

$$\begin{aligned}&\{\mathbf{T} \sim \mathbf{T}'\} \subset \{\mathbf{T}|_r \sim \mathbf{T}'|_r\}\\ &= \{\mathbf{T} \sim \mathbf{T}', H(\mathbf{T}) \le r\} \cup \{\mathbf{T} \sim_r \mathbf{T}', H(\mathbf{T}) > r\}\,.\end{aligned}$$

Thus, by Lemma 3 we have

$$\gamma_\lambda \le g_r \le \gamma_\lambda + \mathfrak{p}_r \text{ and } g_r \downarrow \gamma_\lambda \text{ as } r \to \infty\,. \tag{II.8}$$

Let D, D' be the numbers of children for the roots of $\mathbf{T}, \mathbf{T}'$ respectively. Let $\mathbf{T}_i$ be the subtree rooted at the i-th child for the root of $\mathbf{T}$ (similar notation applies for $\mathbf{T}'_i$), that is, the tree that consists of the i-th child as well as all of its descendants. Then, on the event $\{D = D' = k\}$ we have

$$\{\mathbf{T} \sim_r \mathbf{T}'\} = \bigcup_{1\le i,j\le k} \{\mathbf{T}_i \sim_{r-1} \mathbf{T}'_j\} \cap \{(\mathbf{T}\backslash\mathbf{T}_i)|_r \sim (\mathbf{T}'\backslash\mathbf{T}'_j)|_r\}\,. \quad \text{(II.9)}$$

On the one hand, we get from (II.9) and (II.6) that

$$\begin{aligned}\mathfrak{p}_r &\le \sum_{k=1}^{\infty} \mu_k^2 k^2 \mathfrak{p}_{r-1}\mathbb{P}\left(\mathbf{T}|_r \sim \mathbf{T}'|_r \mid D = D' = k-1\right)\\ &= \sum_{k=1}^{\infty} \lambda^2 \mathfrak{p}_{r-1}\mu_{k-1}^2 \mathbb{P}\left(\mathbf{T}|_r \sim \mathbf{T}'|_r \mid D = D' = k-1\right)\\ &= \lambda^2 g_r \mathfrak{p}_{r-1}\,. \end{aligned} \quad \text{(II.10)}$$

On the other hand, using the inequality $\mathbb{P}(\cup_i A_i) \ge \sum_i \mathbb{P}(A_i) - \frac{1}{2}\sum_{i\ne j}\mathbb{P}(A_i \cap A_j)$ we get from (II.9) that

$$\begin{aligned}\mathfrak{p}_r &\ge \sum_{k=1}^{\infty} k^2 \mathfrak{p}_{r-1}\mathbb{P}\left(\mathbf{T}|_r \sim \mathbf{T}'|_r \mid D = D' = k-1\right)\mu_k^2 - \frac{\mathrm{err}}{2}\\ &= \lambda^2 g_r \mathfrak{p}_{r-1} - \frac{\mathrm{err}}{2} \end{aligned} \quad \text{(II.11)}$$

where err equals

$$\sum_{k=2}^{\infty} \sum_{(i_1,j_1)\ne(i_2,j_2)} \mathbb{P}(\mathbf{T}_{i_l} \sim_{r-1} \mathbf{T}'_{j_l}, (\mathbf{T}\backslash\mathbf{T}_{i_l})|_r \sim (\mathbf{T}'\backslash\mathbf{T}'_{j_l})|_r \text{ for } l \in \{1,2\} \mid D = D' = k)\mu_k^2\,.$$

If $i_1 \ne i_2$ and $j_1 \ne j_2$, then the conditional probability above is equal to

$$\mathfrak{p}_{r-1}^2 \mathbb{P}\left(\mathbf{T}|_r \sim \mathbf{T}'|_r \mid D = D' = k-2\right)\,. \quad \text{(II.12)}$$

Otherwise if $i_1 \ne i_2$ but $j_1 = j_2$ (the other case is symmetric), on the event $\mathbf{T}_{i_l} \sim_{r-1} \mathbf{T}'_{j_l}, (\mathbf{T}\backslash\mathbf{T}_{i_l})|_r \sim (\mathbf{T}'\backslash\mathbf{T}'_{j_l})|_r$ for $l \in \{1,2\}$ there must exist $1 \le j'_1 < j'_2 \le k$ such that $\mathbf{T}_{i_l} \sim_{r-1} \mathbf{T}'_{j'_l}$ and $(\mathbf{T}\backslash\mathbf{T}_{i_l})|_r \sim (\mathbf{T}'\backslash\mathbf{T}'_{j'_l})|_r$ for $l \in \{1,2\}$. Since the probability for each such event with respect to a fixed pair (j'_1, j'_2) is bounded by (II.12), by a union bound we can bound the aforementioned conditional probability. Altogether,

$$\begin{aligned}&\mathrm{err}\\ &\le 4\sum_{k\geqslant 2}(k(k-1))^2 \mathfrak{p}_{r-1}^2 \mathbb{P}\left(\mathbf{T}|_r \sim \mathbf{T}'|_r \mid D{=}D'{=}k-2\right)\mu_k^2\\ &= 4\sum_{k\geqslant 2}\lambda^4 \mathfrak{p}_{r-1}^2 \mathbb{P}\left(\mathbf{T}|_r \sim \mathbf{T}'|_r \mid D = D' = k-2\right)\mu_{k-2}^2\\ &= 4\lambda^4 g_r \mathfrak{p}_{r-1}^2\end{aligned}$$

where we have used (II.6) twice. Combined with (II.10) and (II.11), it yields that

$$\lambda^2 g_r \mathfrak{p}_{r-1} - 4\lambda^4 g_r \mathfrak{p}_{r-1}^2 \le \mathfrak{p}_r \le \lambda^2 g_r \mathfrak{p}_{r-1}\,. \quad \text{(II.13)}$$

By (II.8) and $\mathfrak{p}_r \to 0$ (Lemma 3), we have that $\mathfrak{p}_r = \alpha_\lambda^{(1+o(1))r}$ where $o(1)$ vanishes in r. Since $\alpha_\lambda < 1$ (Lemma 4), this implies that $\mathfrak{p}_r$ decays exponentially. Combining this with (II.8) and (II.13), we see that $\sum_{i=1}^{\infty} |\frac{\mathfrak{p}_{i+1}}{\mathfrak{p}_i} - \alpha_\lambda| < \infty$, yielding the desired bound. $\square$

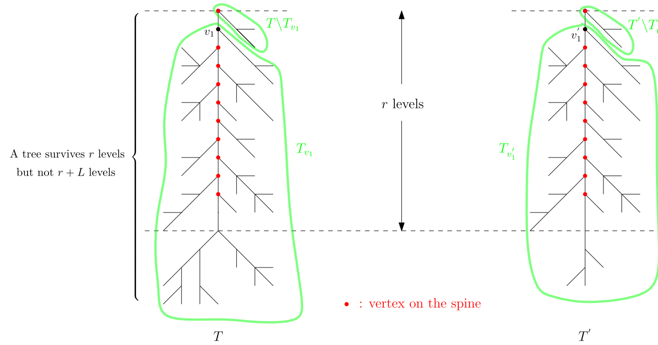


Fig. 2. The event $\mathcal{E}(T, T'; r, L)$.

In order to prove non-identifiability, we need to pose some additional structure on two isomorphic trees. To this end, we say a sequence of vertices $(v_i)_{i=0}^m$ is a *spine* of T if it is a path (we say a sequence of distinct vertices is a path if each neighboring pair is connected by an edge) of T started at v_0 (i.e., the root of T). Furthermore, for $v \in T$, we denote by T_v the subtree rooted at v. For $r, L \ge 1$, define $\mathfrak{p}_{r,L} = \mathbb{E}(\mathcal{E}(\mathbf{T}, \mathbf{T}'; r, L))$ where $\mathcal{E}(T, T'; r, L) = 1$ if

$$\begin{aligned}&T \sim_r T', H(T), H(T') \le r + L; H(T) \ne H(T');\\ &\text{every vertex in } r'\text{-th level of } T, T' \text{ has either 0 or 2 children}\\ &\qquad \text{for all } r' > r;\\ &\exists \text{ spine } (v_i)_{i=0}^{r-L} \text{ in } T, \text{ such that } |T_{v_{i-1}}\backslash T_{v_i}| \le L \text{ for all } i\,.\end{aligned} \quad \text{(II.14)}$$

Lemma 6: For every $\epsilon > 0$, there exists $L = L_\epsilon$ depending only on ϵ such that $\mathfrak{p}_{r,L} \ge (\alpha_\lambda - \epsilon)^r$ for sufficiently large r.

Proof: For two rooted trees T and T', define the event $A(T, T') = \{T \sim T', |T| \le L\}$. Note that $\mathcal{E}(T, T'; r, L)$ can be defined inductively. For $r > L$, we have that $\mathcal{E}(T, T'; r, L) = 1$ if and only if there exists a *unique* pair $(v_1, v'_1) \in (T, T')$ where v_1 (respectively v'_1) is a child of the root in T_1 (respectively T'_1), such that $\mathcal{E}(T_{v_1}, T'_{v'_1}; r-1, L) = 1$ and that $A(T\backslash T_{v_1}, T'\backslash T'_{v'_1})$ occurs (see Figure 2). Therefore, denoting by D, D' the numbers of children for the roots of $\mathbf{T}, \mathbf{T}'$ and recalling (II.5), we have that

$$\begin{aligned}\mathfrak{p}_{r,L} &= \sum_{k=1}^{\infty} \sum_{1\le v_1, v'_1 \le k} \mu_k^2 \mathbb{E}(\mathcal{E}(\mathbf{T}_{v_1}, \mathbf{T}'_{v'_1}; r-1, L))\\ &\qquad \times \mathbb{P}(A(\mathbf{T}\setminus\mathbf{T}_{v_1}, \mathbf{T}'\setminus\mathbf{T}'_{v'_1}) \mid D = D' = k)\\ &= \mathfrak{p}_{r-1,L}\lambda^2 \sum_{k=1}^{\infty} \mu_{k-1}^2 \mathbb{P}(A(\mathbf{T}, \mathbf{T}') \mid D = D' = k-1)\\ &= \mathfrak{p}_{r-1,L}\lambda^2 \mathbb{P}(A(\mathbf{T}, \mathbf{T}'))\,,\end{aligned} \quad \text{(II.15)}$$

where the second equality follows from (II.6) and the fact that the conditional law of $(\mathbf{T}\setminus\mathbf{T}_{v_1}, \mathbf{T}'\setminus\mathbf{T}'_{v'_1})$ given $D = D' = k$ is the same as the conditional law of $(\mathbf{T}, \mathbf{T}')$ given $D = D' = k-1$. By Lemma 3, we see that $\mathbb{P}(A(\mathbf{T}, \mathbf{T}') \mid \mathbf{T} \sim \mathbf{T}') \to 1$ as $L \to \infty$ and thus $\mathbb{P}(A(\mathbf{T}, \mathbf{T}')) \to \mathbb{P}(\mathbf{T} \sim \mathbf{T}') = \gamma_\lambda$. Therefore, for any $\epsilon > 0$, there exists $L = L_\epsilon$ such that $\mathbb{P}(A(\mathbf{T}, \mathbf{T}')) \ge \gamma_\lambda - \epsilon/(2\lambda^2)$. Combined with (II.15), this gives that

$$\mathfrak{p}_{r,L} \geq (\alpha_\lambda - \epsilon/2)^{r-L}\mathfrak{p}_{L,L} \geq c_L(\alpha_\lambda - \epsilon/2)^{r-L}$$

where $c_L > 0$ depending only on (L, λ). This completes the proof. □

We also need an estimate that compares local neighborhoods of Erdős–Rényi graphs to PGW trees. This has been well-understood and fairly straightforward. For instance, a straightforward extension of [31, Lemma 2.2] leads to the following lemma (so we omit the proof). For each vertex $v \in \mathcal{G}$, let C_v be the component of v in $\mathcal{G}$ with root v.

Lemma 7: Let $\mathbf{T}_i$'s be independent PGW(λ)-trees. For $k \geq 1$ and $v_1, \ldots, v_k \in \mathcal{G}$, and for rooted trees $\tau_1, \ldots, \tau_k$ with $\sum_{i=1}^k |\tau_i| = o(\sqrt{n})$, we have that as $n \to \infty$,

$$\begin{aligned}&\mathbb{P}(\mathsf{C}_{v_i} \sim \tau_i \text{ for } 1 \leq i \leq k \text{ and } \mathsf{C}_{v_i}\text{'s are disjoint})\\ &= (1 + o(1))\mathbb{P}(\mathbf{T}_i \sim \tau_i \text{ for } 1 \leq i \leq k)\,.\end{aligned}$$

B. Proof of Non-Identifiability

This whole subsection is devoted to the proof of non-identifiability. We set $r = \frac{(1-\epsilon_0)\log n}{\log \alpha_\lambda^{-1}}$ so that

$$\alpha_\lambda^{2r} = n^{-2(1-\epsilon_0)} \text{ for an arbitrary fixed small } \epsilon_0 > 0\,. \quad \text{(II.16)}$$

Moreover, we choose $\epsilon = \alpha_\lambda(1 - \alpha_\lambda^{2\epsilon_0/[3(1-\epsilon_0)]})$ and $L = L_\epsilon$ in Lemma 6 so that

$$(\alpha_\lambda - \epsilon)^{2r} = n^{-2(1-\epsilon_0/3)}\,. \quad \text{(II.17)}$$

In order to prove non-identifiability, we will construct a *blocking configuration* as in [25] whose existence certifies non-identifiability, and then we need to show that with high probability such a blocking configuration exists.

1) Construction of Blocking Configuration: We refer to Figure 1 for an illustration of the construction. We say $(\mathsf{C}_v, \mathsf{C}_u)$ is a blocking configuration if it satisfies the following property:

- $\mathsf{C}_v, \mathsf{C}_u$ are disjoint trees and there is a line segment $\mathrm{Line}_{w,w'}$ of length $2L$ with endpoints w, w' and w is connected to v.
- The event $\mathcal{E}(\mathsf{C}_v \setminus \mathrm{Line}_{w,w'}, \mathsf{C}_u; 2r, L) = 1$ holds where $\mathcal{E}$ is defined as in (II.14).

We next show that when a blocking configuration exists for some pair $(\mathsf{C}_v, \mathsf{C}_u)$, the graph is non-identifiable from the empirical profile for r-neighborhoods. Indeed, if we remove the edge (w, v) and add an edge (w, u), i.e., cut $\mathrm{Line}(w, w')$ from v and attach it to u, then we claim that:

(i) the empirical profile for r-neighborhoods is unchanged;
(ii) the empirical profile for $(2r + 4L)$-neighborhoods is changed.

Assuming the claim, we see that the empirical profile of r-neighborhoods does not determine the whole graph up to isomorphism since it does not even determine a unique empirical profile for $(2r + 4L)$-neighborhoods.

We now prove (i). Note that the 'cut-attach' procedure only changes r-neighborhoods for vertices in $\mathsf{N}_r(v)$ and $\mathsf{N}_r(u)$. Since $(\mathsf{C}_v \setminus \mathrm{L}_{w,w'}) \sim_{2r} \mathsf{C}_u$, we can let ϕ be an isomorphism between these two trees. For a vertex z, we let $\widetilde{\mathsf{N}_r}(z)$ be the r-neighborhood of z after the 'cut-attach' procedure. Then it is clear that $\mathsf{N}_r(z) \sim \widetilde{\mathsf{N}_r}(z)$ for all $z \in \mathrm{Line}_{w,w'}$, $\mathsf{N}_r(v') \sim \widetilde{\mathsf{N}_r}(\phi(v'))$ for all $v' \in \mathsf{N}_r(v)\backslash \mathrm{L}_{w,w'}$ and $\mathsf{N}_r(u') \sim \widetilde{\mathsf{N}_r}(\phi^{-1}(u'))$ for all $u' \in \mathsf{N}_r(u)$. This implies (i).

We next prove (ii). This is where we need the additional structure in the definition of (II.14). Suppose without loss of generality that $H(\mathsf{C}_v \setminus \mathrm{Line}_{w,w'}) > H(\mathsf{C}_u)$. By properties in (II.14), we see that the diameter of C_v is $2L + H(\mathsf{C}_v \setminus \mathrm{Line}_{w,w'}) > 2L + H(\mathsf{C}_u)$ which is the diameter of $\mathsf{C}_u \cup \mathrm{Line}_{w,w'} \cup \{(u, w)\}$. This implies (ii), since after the 'cut-attach' procedure the maximal diameter in the components of u and v is changed.

2) Existence of Blocking Configuration: For $v, u \in \mathcal{G}$, let $X_{v,u}$ be the indicator function that $(\mathsf{C}_v, \mathsf{C}_u)$ is a blocking configuration. Let $N = \sum_{v,u} X_{v,u}$. We need to show that with high probability $N \geq 1$. To this end, we need some facts from PGW trees. Denote by BConf the collection of all pairs of rooted unlabeled trees which are blocking configurations. That is, BConf consists of the equivalence classes of blocking configurations and the equivalence is given by rooted isomorphism. Let $\mathbf{T}, \mathbf{T}'$ be two independent PGW(λ)-trees. For $(\tau, \tau') \in \mathrm{BConf}$, the roots of τ, τ' have degree at most L. As a result, under the law of PGW(λ)-tree, cutting a line of size $2L$ from the root of τ only changes its probability density up to a factor depending on (L, λ). Therefore, by Lemma 6 and (II.17) we have that for some constant $c = c(L, \lambda)$

$$\sum_{(\tau,\tau')\in \mathrm{BConf}} \mathbb{P}(\mathbf{T} \sim \tau, \mathbf{T}' \sim \tau') \geq c(\alpha - \epsilon)^{2r} = cn^{-2(1-\epsilon_0/3)}\,. \quad \text{(II.18)}$$

We also note that $|\tau|, |\tau'| = O(\log n)$ whenever $(\tau, \tau') \in \mathrm{BConf}$ and we will use this fact repeatedly (e.g., together with Lemma 7).

We are now ready to employ the second moment method in order to show $N \geq 1$. We have

$$\begin{aligned}\mathbb{E}[N^2] &= \mathbb{E}[N] + \sum_{(u_1,u_2)\neq(u_3,u_4)} \mathbb{E}(X_{u_1,u_2}X_{u_3,u_4})\\ &\leq \mathbb{E}N + n^4\mathbb{E}(X_{v_1,v_2}X_{v_3,v_4})\\ &\quad + n^3\mathbb{E}(X_{v_1,v_2}X_{v_1,v_4}) + n^3\mathbb{E}(X_{v_1,v_2}X_{v_3,v_2})\,,\end{aligned}$$

where $v_1, \ldots, v_4$ are pairwise different. Thus, it suffices to show that

(a) $\lim_{n\to\infty} n^2\mathbb{E}(X_{v_1,v_2}) = \infty$;

(b) $\limsup_{n\to\infty} \frac{\mathbb{E}(X_{v_1,v_2}X_{v_3,v_4})}{(\mathbb{E}X_{v_1,v_2})^2} \leq 1$

(c) $\lim_{n\to\infty} \frac{\mathbb{E}(X_{v_1,v_2}X_{v_1,v_4})}{n(\mathbb{E}X_{v_1,v_2})^2} = \lim_{n\to\infty} \frac{\mathbb{E}(X_{v_1,v_2}X_{v_3,v_2})}{n(\mathbb{E}X_{v_1,v_2})^2} = 0\,.$

Indeed, having verified (a), (b) and (c), we can then apply Cauchy-Schwarz inequality and get that $\mathbb{P}(N \geq 1) \geq \frac{[\mathbb{E}N]^2}{\mathbb{E}[N^2]} \to 1$ as $n \to \infty$. In what follows we denote by $\mathbf{T}_i$'s as independent PGW(λ)-trees.

Proof of (a): By Lemma 7 and (II.18), we get that

$$\mathbb{E}X_{v_1,v_2} = (1+o(1)) \sum_{(\tau_1,\tau_2)\in \mathrm{BConf}} \mathbb{P}(\mathbf{T}_1 \sim \tau_1, \mathbf{T}_2 \sim \tau_2)$$
$$\text{and } \mathbb{E}N \geq (c+o(1))n^{2\epsilon_0/3}\,, \tag{II.19}$$

which implies (a).

Proof of (b): Write Ω as the event that $\mathsf{C}_{v_1},\dots,\mathsf{C}_{v_4}$ are mutually disjoint. By Lemma 7 we get that

$$\begin{aligned}\mathbb{E}(X_{v_1,v_2}X_{v_3,v_4}\mathbf{1}_{\Omega}) &= (1+o(1))\mathbb{E}(X_{v_1,v_2})\mathbb{E}(X_{v_3,v_4})\\ &= (1+o(1))(\mathbb{E}X_{v_1,v_2})^2.\end{aligned}$$

Therefore, in order to prove (b) it suffices to show that

$$\mathbb{E}(X_{v_1,v_2}X_{v_3,v_4}\mathbf{1}_{\Omega^c}) = o((\mathbb{E}X_{v_1,v_2})^2). \tag{II.20}$$

By the definition of a blocking configuration (see properties in (II.14)), for $(\tau_1,\tau_2) \in \mathrm{BConf}$, there is only one vertex $w' \in \tau_1 \cup \tau_2$ whose $2L$-neighborhood is a line $\mathrm{Line}_{w,w'}$ of length $2L$ (with w' being one endpoint), and thus the root of τ_1 is the only vertex which is connected to w (the other endpoint of $\mathrm{Line}_{w,w'}$). Therefore, when $X_{v_1,v_2}X_{v_3,v_4} = 1$, we have $v_3 \notin \mathsf{C}_{v_1}$ since otherwise there would exist two lines of length $2L$ attached to two vertices $v_1, v_3 \in \mathsf{C}_{v_1}$, which is a contradiction. In addition, we have $v_3 \notin \mathsf{C}_{v_2}$ since otherwise there would exist a line of length $2L$ attached to $v_3 \in \mathsf{C}_{v_2}$, which is a contradiction (because such v_3 only exists in C_{v_1} but not in C_{v_2}). For the same reason, we have $v_1 \notin \mathsf{C}_{v_4}$ (which is equivalent to $v_4 \notin \mathsf{C}_{v_1}$). Therefore, $\mathsf{C}_{v_1}, \mathsf{C}_{v_2}, \mathsf{C}_{v_3}$ are disjoint. Thus, if in addition Ω does not occur, we must have that $v_4 \in \mathsf{C}_{v_2}$. By Lemma 7 and the fact that any blocking configuration has size $O(\log n)$, we get that

$$\begin{aligned}&\mathbb{E}[X_{v_1,v_2}X_{v_3,v_4}\mathbf{1}_{\Omega^c}]\\ &\leq \sum_{\substack{(\tau_1,\tau_2)\in\mathrm{BConf}\\(\tau_3,\tau_4)\in\mathrm{BConf}}} \mathbb{P}[\mathsf{C}_{v_j} \sim \tau_j, 1\leq j\leq 4;\\ &\qquad\qquad \mathsf{C}_{v_1},\mathsf{C}_{v_2},\mathsf{C}_{v_3} \text{ disjoint }; v_4 \in \mathsf{C}_{v_2}]\\ &\leq (1+o(1)) \sum_{(\tau_1,\tau_2)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_1\sim\tau_1,\mathbf{T}_2\sim\tau_2)\\ &\quad\times \sum_{(\tau_3,\tau_4)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_3\sim\tau_3)1_{\{\tau_4\in\mathcal{E}(\tau_2)\}}\frac{O(\log n)}{n}\,,\end{aligned}$$

where $\mathcal{E}(\tau_2)$ consists of all (equivalence classes for) rooted trees that arise from τ_2 by re-choosing the root as an arbitrary vertex in τ_2. Let us explain in details the procedure of computation which leads to the final bound above.

- We first check the component containing v_1: the event $\mathsf{C}_{v_1} \sim \tau_1$ has probability $(1+o(1))\mathbf{P}(\mathbf{T}\sim\tau_1)$ by Lemma 7.
- The event $v_2, v_3, v_4 \notin C_{v_1}$ has probability $1-o(1)$.
- We then check the component containing v_2: the event $\mathsf{C}_{v_2} \sim \tau_2, v_3 \notin \mathsf{C}_{v_2}$ has probability $(1-o(1))\mathbf{P}(\mathbf{T}\sim\tau_2)$.
- We then check the component of v_4: since $v_4 \in \mathsf{C}_{v_2}$, we must have $\tau_4 \in \mathcal{E}(\tau_2)$ (which leads to $1_{\{\tau_4\in\mathcal{E}(\tau_2)\}}$) since otherwise it will have contribution 0 to the sum. In addition v_4 has to be at some correct position in C_{v_2}, which has probability at most $O(\log n/n)$.

Noting that $|\mathcal{E}(\tau_2)| = O(\log n)$ (since the number of equivalence classes is bounded by the number of ways to choosing a new root in τ_2, which in turn is bounded by the size of τ_2), we see that the number of valid choices for τ_3 in the preceding sum is $O(\log n)$. Thus,

$$\begin{aligned}\mathbb{E}[X_{v_1,v_2}X_{v_3,v_4}\mathbf{1}_{\Omega^c}] &\leq \sum_{(\tau_1,\tau_2)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_1\sim\tau_1,\mathbf{T}_2\sim\tau_2)\\ &\quad\times \max_{(\tau_3,\tau_4)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_3\sim\tau_3)\frac{O((\log n)^2)}{n}\,.\end{aligned}$$

In addition,

$$\max_{(\tau_3,\tau_4)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_3\sim\tau_3) \leq \sqrt{\mathfrak{p}_{2r}} = O(\alpha_\lambda^r) \tag{II.21}$$

where the last equality follows from Lemma 5. Altogether, we get that

$$\begin{aligned}&\mathbb{E}[X_{v_1,v_2}X_{v_3,v_4}\mathbf{1}_{\Omega^c}]\\ &\leq \mathbb{P}((\mathbf{T}_1,\mathbf{T}_2)\in\mathrm{BConf})\frac{O((\log n)^2)\alpha_\lambda^r}{n}.\end{aligned}$$

Combined with (II.16) and (II.19), this yields (II.20) as required.

Proof of (c): When $X_{v_1,v_2}X_{v_1,v_4} = 1$, (using the uniqueness of the line graph of length $2L$ in the blocking configuration as argued in Proof of (b)) we must have either $\mathsf{C}_{v_1}, \mathsf{C}_{v_2}, \mathsf{C}_{v_4}$ are disjoint (which we denote by the event $\tilde{\Omega}$) or $v_4 \in \mathsf{C}_{v_2}$. By Lemma 7, we get

$$\begin{aligned}&\mathbb{E}[X_{v_1,v_2}X_{v_1,v_4}\mathbf{1}_{\tilde{\Omega}}]\\ &= (1+o(1)) \sum_{\substack{(\tau_1,\tau_2)\in\mathrm{BConf}\\(\tau_1,\tau_4)\in\mathrm{BConf}}} \mathbb{P}(\mathbf{T}_j\sim\tau_j \text{ for } j=1,2,4)\\ &= \sum_{(\tau_1,\tau_2)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_1\sim\tau_1,\mathbf{T}_2\sim\tau_2)\mathbb{P}(\mathbf{T}_4\sim\tau_2)\end{aligned}$$

where we have used the fact that the valid configurations in the summation satisfy $\tau_4 = \tau_2$. Therefore, combined with (II.19) and (II.21), it yields that

$$\begin{aligned}&\mathbb{E}[X_{v_1,v_2}X_{v_1,v_4}\mathbf{1}_{\tilde{\Omega}}]\\ &\leq \mathbb{P}((\mathbf{T}_1,\mathbf{T}_2)\in\mathrm{BConf})\sqrt{\mathfrak{p}_{2r}} = o(n(\mathbb{E}X_{v_1,v_2})^2)\,.\end{aligned}$$

We next estimate the expectation on the event $\tilde{\Omega}^c$. We have

$$\begin{aligned}&\mathbb{E}[X_{v_1,v_2}X_{v_1,v_4}\mathbf{1}_{\tilde{\Omega}^c}]\\ &\leq \sum_{(\tau_1,\tau_2)\in\mathrm{BConf}} \mathbb{P}(\mathbf{T}_1\sim\tau_1,\mathbf{T}_2\sim\tau_2)O(\frac{\log n}{n})\\ &= o(n(\mathbb{E}X_{v_1,v_2})^2)\,,\end{aligned}$$

where the $O(\frac{\log n}{n})$ comes from the event that $v_4 \in \mathsf{C}_{v_2}$ (and $|\mathsf{C}_{v_2}| = O(\log n)$ by definition of blocking configuration), and the last equality uses (II.19).

When $X_{v_1,v_2}X_{v_3,v_2} = 1$, we must have that $\mathsf{C}_{v_1}, \mathsf{C}_{v_2}, \mathsf{C}_{v_3}$ are disjoint (again using the uniqueness of the line graph of length L in the blocking configuration). Then the bound on $\mathbb{E}(X_{v_1,v_2}X_{v_3,v_2})$ can be derived in the same way as that for $\mathbb{E}[X_{v_1,v_2}X_{v_1,v_4}\mathbf{1}_{\tilde{\Omega}}]$, completing the proof of (c) and thus completing the proof of non-identifiability.

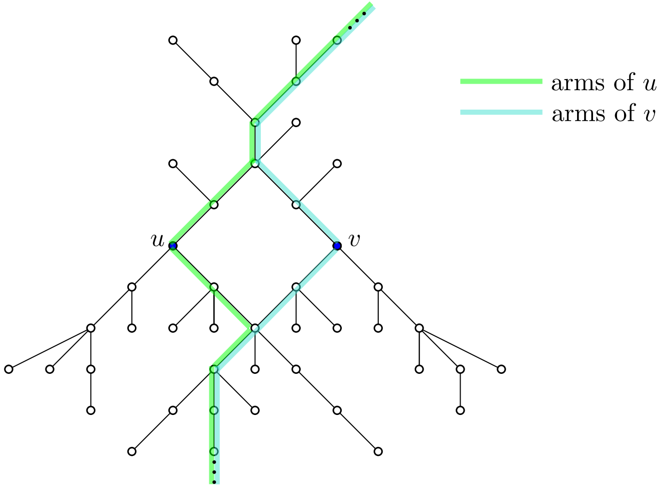


Fig. 3. An illustration for isomorphic neighborhoods with two arms centered on a cycle.

III. Proof of Identifiability

The rest of the paper is devoted to the proof of identifiability when

$$r = \frac{(1+\epsilon_0)\log n}{\log \alpha_\lambda^{-1}}\,,\ \alpha_\lambda^r < \alpha_\lambda^{\rho r} \le n^{-1-\frac{\epsilon_0}{2}}$$
$$\text{and } \frac{\log(\alpha_\lambda^{-1})}{2(1+\epsilon_0)} > \log(\lambda)\,, \tag{III.1}$$

for an arbitrarily fixed $\epsilon_0 > 0$ and $\rho = \rho(\epsilon_0) < 1$. The third inequality is possible for small ϵ_0 since $\frac{\log(\alpha_\lambda^{-1})}{2} > \log(\lambda)$ by (II.4) (note that identifiability is harder for smaller r, so it is fine to assume ϵ_0 to be small as incorporated by the third inequality), and we made such an assumption for convenience of controlling the volume of the r-neighborhood as in Lemma 24. We now explain how to relate Lemma 5 to the above choice of r. We say a neighborhood $\mathsf{N}_r(v)$ has *two r-arms* (or we say v has *two r-arms*) if there are two paths of length r which are both rooted at v and intersect only at v. By Lemma 5, we would essentially get that the probability for two r-neighborhoods with two r-arms to be isomorphic is at most $o(n^{-2})$ and as a result essentially each r-neighborhood with two arms is unique. (Here the word "essentially" refers to omitting the consideration for the scenario of $\mathsf{N}_r(v) \sim \mathsf{N}_r(u)$ when both neighborhoods have two r-arms but u, v are contained in a short cycle; see Figure 3.) Intuitively, this would be sufficient for recovery since a vertex v is either contained in some r-neighborhood (not necessarily rooted at v) with two r-arms or the component of v is contained in some r-neighborhood.

We now describe how we treat small components as our preprocessing procedure. We say a vertex v is *degenerate* if $\mathsf{N}_r(v) = \mathsf{N}_{r-1}(v)$. Note that for a degenerate v, we have that

$$\mathsf{N}_r(v) = \mathsf{C}_v \text{ and } \mathsf{N}_r(u) \sim \mathsf{N}_r(u;\mathsf{C}_v) \text{ for all } u \in \mathsf{N}_r(v)\,, \tag{III.2}$$

where $\mathsf{N}_r(u;\mathsf{C}_v)$ is the r-neighborhood of u in C_v. Recall that we have assumed an ordering on V (in the paragraph for notation convention). We set $U_1 = U_2 = \emptyset$ and iteratively apply the following procedure:

- If $V \setminus U_2 = \emptyset$, stop. Otherwise, take the minimal degenerate $v \in V \setminus U_2$, and choose a subset $A \subset V$ such that $\{\mathsf{N}_r(u;\mathsf{C}_v) : u \in \mathsf{C}_v\} = \{\mathsf{N}_r(u) : u \in A\}$ (this is possible due to (III.2));
- Add vertex v to U_1 and add vertices in A to U_2.

At the end of our procedure, we have that $\cup_{v\in U_1}\mathsf{C}_v$ contains all degenerate vertices (it may also contain some non-degenerate vertices as well) and

$$\{\mathsf{N}_r(u) : u \in \cup_{v\in U_1}\mathsf{C}_v\} = \{\mathsf{N}_r(u) : u \in U_2\}\,.$$

Therefore, if we can identify the graph from $\{\mathsf{N}_r(u) : u \in V \setminus U_2\}$ up to isomorphism, then adding disjoint components $\{\mathsf{C}_v : v \in U_1\}$ to it yields the original graph up to isomorphism. For this reason, in what follows we assume without loss of generality that all vertices are not degenerate (or equivalently, we have removed components of degenerate vertices using the preprocessing procedure above).

We next describe our recovery procedure, whose success relies on certain structural properties for the Erdős–Rényi graph which we discuss later. We say a vertex v is *good* if $\mathsf{N}_{\rho r}(v)$ is unique (among $\{\mathsf{N}_{\rho r}(u) : u \in \mathcal{G}\}$), and we let $V_{\rm g}$ be the collection of all good vertices (crucially, a good vertex can be regarded as labeled since its neighborhood is unique). Write $V_{\rm b} = V \setminus V_{\rm g}$ as the collection of *bad* vertices. For each bad component $\mathsf{C}_{\rm b}$, i.e., a component in the induced subgraph of $\mathcal{G}$ on $V_{\rm b}$, let $\partial_{\rm e}\mathsf{C}_{\rm b} = \{w \in V \setminus \mathsf{C}_{\rm b} : (w,u) \in \mathcal{G} \text{ for some } u \in \mathsf{C}_{\rm b}\}$ be the *external boundary* of $\mathsf{C}_{\rm b}$ (we comment that by definition $\partial_{\rm e}\mathsf{C}_{\rm b} \subset V_{\rm g}$) and we let $\mathsf{D}(\mathsf{C}_{\rm b})$ be the graph on $\mathsf{C}_{\rm b} \cup \partial_{\rm e}\mathsf{C}_{\rm b}$ which contains all edges within $\mathsf{C}_{\rm b}$ and all edges between $\mathsf{C}_{\rm b}$ and $\partial_{\rm e}\mathsf{C}_{\rm b}$. We would like to consider the empirical profile for $\mathsf{D}(\mathsf{C}_{\rm b})$ with $\mathsf{C}_{\rm b}$ ranging over all bad components. To this end, we let Ψ be a mapping that maps each graph to its equivalence class where the equivalence is given by isomorphism that preserves good vertices. That is to say, we view $\Psi(\mathsf{D}(\mathsf{C}_{\rm b}))$ as a graph with vertices in $\partial_{\rm e}\mathsf{C}_{\rm b}$ labeled but with vertices in $\mathsf{C}_{\rm b}$ unlabeled. Define the profile $\mathcal{D}_{\rm b} = \{\Psi(\mathsf{D}(\mathsf{C}_{\rm b})) : \mathsf{C}_{\rm b} \text{ is a bad component}\}$. Note that $\mathcal{D}_{\rm b}$ may be a set with multiplicity since we may have multiple copies for a certain $\Psi(\mathsf{D}(\mathsf{C}_{\rm b}))$ due to the fact that vertices in $\mathsf{C}_{\rm b}$ are not labeled.

Our key intuition is to recover bad components from good vertices. To formalize this, for each $x \in V_{\rm g}$ let U_x be the collection of vertices whose ρr-neighborhoods are contained in $\mathsf{N}_{r-1}(x)$. We remark that in the definition above we require to be contained in $\mathsf{N}_{r-1}(x)$ instead of in $\mathsf{N}_r(x)$ for the reason that the former event is measurable with respect to $\mathsf{N}_r(x)$ but the latter is not. Write $U_{x,{\rm g}} = U_x \cap V_{\rm g}$ and $U_{x,{\rm b}} = U_x \cap V_{\rm b}$. We say that $\mathsf{C}_{x,{\rm b}}$ is a bad component in $\mathsf{N}_r(x)$ if

- $\mathsf{C}_{x,{\rm b}}$ is connected in the subgraph induced on $U_{x,{\rm b}}$;
- $\partial_{\rm e}\mathsf{C}_{x,{\rm b}} \subset U_{x,{\rm g}}$ and $x \in \partial_{\rm e}\mathsf{C}_{x,{\rm b}}$ where

$$\partial_{\rm e}\mathsf{C}_{x,{\rm b}} = \{w \in \mathsf{N}_r(x) \setminus \mathsf{C}_{x,{\rm b}} : w \text{ is neighboring some } w' \in \mathsf{C}_{x,{\rm b}}\}\,.$$

As above $\mathsf{D}(\mathsf{C}_{x,{\rm b}})$ is the graph on $\mathsf{C}_{x,{\rm b}} \cup \partial_{\rm e}\mathsf{C}_{x,{\rm b}}$ which contains all edges within $\mathsf{C}_{x,{\rm b}}$ and all edges between $\mathsf{C}_{x,{\rm b}}$ and $\partial_{\rm e}\mathsf{C}_{x,{\rm b}}$. Let $\mathfrak{D}_{x,{\rm b}}$ be the collection of $\mathsf{D}(\mathsf{C}_{x,{\rm b}})$ for all bad component $\mathsf{C}_{x,{\rm b}}$ in $\mathsf{N}_r(x)$. We need to be careful about what we can recover exactly from a rooted neighborhood where all other vertices are not labeled. Since from $\mathsf{N}_r(v)$ we know the ρr-neighborhoods for vertices in $\partial_{\mathrm{e}}\mathsf{C}_{x,\mathrm{b}}$ and since these vertices are good, we can assume that we know the labels for vertices in $\partial_{\mathrm{e}}\mathsf{C}_{x,\mathrm{b}}$; but we do not know labels for vertices in $\mathsf{C}_{x,\mathrm{b}}$. This echoes the definition of Ψ from above. Similarly, $\mathfrak{D}_{x,\mathrm{b}}$ may be a set with multiplicity. Furthermore, we define $\mathfrak{D}'_{x,\mathrm{b}}$ to be the collection of $\mathsf{D}(\mathsf{C}_{x,\mathrm{b}}) \in \mathfrak{D}_{x,\mathrm{b}}$ such that $\mathsf{D}(\mathsf{C}_{x,\mathrm{b}}) \notin \mathfrak{D}_{y,\mathrm{b}}$ for any $y \in \partial_{\mathrm{e}}\mathsf{C}_{x,\mathrm{b}}$ which is less than x, where we recall that we have fixed an arbitrary ordering on V.

We are now ready to recover our original graph by adding edges between V_{g} and then adding small components incident to V_{g} as follows:

- For any pair of good vertices, whether there is an edge can be determined by the r-neighborhood for either of them and we then add an edge if there is one.
- For each good vertex x and each $\mathsf{D}_{x,\mathrm{b}}(= \mathsf{D}(\mathsf{C}_{x,\mathrm{b}})) \in \mathfrak{D}'_{x,\mathrm{b}}$, we add a copy of $\mathsf{D}_{x,\mathrm{b}}$ where bad vertices (i.e., those in $\mathsf{C}_{x,\mathrm{b}}$) in each such added copy are disjoint.

We denote by $\mathcal{G}' = (V', E')$ as the graph obtained from the preceding construction.

A. Running Time Analysis

In our procedure, most operations are standard and can be performed in polynomial time except for the algorithm of testing isomorphism between two rooted neighborhoods. So far there is no polynomial-time algorithm known to test isomorphism for general graphs, and the best result is a quasi-polynomial-time algorithm [4]. However, for r-neighborhoods under consideration in our problem, they are trees or tree-like graphs and efficient algorithms are known for isomorphism. More precisely, polynomial-time algorithms have been proposed to test isomorphism for graphs with bounded tree-width which in particular include graphs with bounded complexity; see [6], [12], [14], [21], and [22] (we also note that there is a classic linear-time algorithm to test isomorphism for rooted trees [2]). Thanks to Lemma 24 (below), with high probability all r-neighborhoods under consideration have bounded complexity and as a result isomorphism can be tested via polynomial-time algorithms (we may also simply stop the algorithm and declare failure if on the rare event the algorithm detects that some neighborhood has a complexity exceeding a prescribed bound, so that the algorithm stops in polynomial-time deterministically). Altogether, our recovery procedure has a polynomial running time. It is an interesting question to design an algorithm that achieves the "optimal" running time.

The much more challenging task is to prove that the preceding procedure succeeds to recover the original graph $\mathcal{G}$ with high probability. Reall that $\mathcal{G}'$ is a graph obtained from the preceding construction. To this end, we need the following *admissibility* condition for the Erdős–Rényi graph.

Definition 8: We say $\mathcal{G}$ is (r, ρ)-admissible if

$$\mathcal{D}_{\mathrm{b}} = \cup_x \mathfrak{D}'_{x,\mathrm{b}}\,.$$

Lemma 9: If $\mathcal{G}$ is (r, ρ)-admissible, then $\mathcal{G}'$ is isomorphic to $\mathcal{G}$.

Remark 10: Note that Lemma 9 holds for all $r \geq 1$ and $\rho < 1$. The assumption (III.1) made at the beginning of this section is for the purpose of verifying admissibility.

Proof of Lemma 9: In order to prove the lemma, we will define a vertex bijection φ and we will prove that φ is an isomorphism between $\mathcal{G}$ and $\mathcal{G}'$. From our construction we see that $V_{\mathrm{g}} \subset V'$ and we define φ to be the identical map on V_{g}. It remains to define φ on V_{b}. Let $\mathcal{D}'$ be the isomorphic copies of $\cup_x \mathfrak{D}'_{x,\mathrm{b}}$ in $\mathcal{G}'$. By admissibility, there exists a bijection $\Gamma : \mathcal{D}_{\mathrm{b}} \mapsto \mathcal{D}'$ such that D_{b} is isomorphic to $\Gamma(\mathsf{D}_{\mathrm{b}})$ for each $\mathsf{D}_{\mathrm{b}} \in \mathcal{D}_b$. In addition, we let $\varphi_{\mathsf{D}_{\mathrm{b}}}$ be an isomorphism; we remind the reader that $\varphi_{\mathsf{D}_{\mathrm{b}}}$ preserves good vertices. Since bad vertices in D_{b}'s are disjoint, we can then define

$$\varphi(v) = \varphi_{\mathsf{D}_{\mathrm{b}}}(v) \text{ for } v \in V_{\mathrm{b}} \cap \mathsf{D}_{\mathrm{b}}\,.$$

Clearly φ is a bijection. It remains to prove that φ preserves edges. It is obvious that φ preserves edges within V_{g}, and thus it remains to check edges that are incident to at least one bad vertex. For each $\mathsf{D}_{\mathrm{b}} \in \mathcal{D}_{\mathrm{b}}$, we can write $\mathsf{D}_{\mathrm{b}} = \mathsf{D}(\mathsf{C}_{\mathrm{b}})$. In addition, for $\mathsf{D}'_{\mathrm{b}} = \Gamma(\mathsf{D}_{\mathrm{b}})$ we let C'_{b} be the collection of vertices in D'_{b} but not in V_{g}. From our construction, it is clear that C_{b} is not neighboring any vertex outside the vertex set of D_{b}, and also C'_{b} is not neighboring any vertex outside the vertex set of D'_{b}. For edges within D_{b}, φ preserves them since the restriction of φ on D_{b} is the same as $\varphi_{\mathsf{D}_{\mathrm{b}}}$ (which is an isomorphism between D_{b} and D'_{b}). This completes the proof. $\square$

In light of Lemma 9, the main remaining task is to prove that $\mathcal{G}$ is admissible with high probability. To this end, we present a sufficient condition for admissibility in this section and we verify this condition in Section IV.

In light of our preprocessing, all remaining components are non-degenerate, i.e., each component does not contain a degenerate vertex. In addition, the admissibility for each connected component implies admissibility for the whole graph. As a result, in this section we consider a connected non-degenerate graph $G = (V(G), E(G))$, and we define $V_{\mathrm{g}}(G)$ and $V_{\mathrm{b}}(G)$ as V_{g} and V_{b} above but with respect to graph G. For notation convenience, in many cases we drop the dependence on G when there is no ambiguity.

We next define *cycle* and *simple cycle*: we say a sequence of (not necessarily distinct) vertices is a cycle if each of the neighboring pairs (including the pair for the starting and ending vertices) is connected by an edge in G and in addition all these edges are distinct; we say a cycle is a simple cycle if each vertex has degree 2 in this cycle. We say an edge $e \in E(G)$ is a *bridge* if it is not contained in any cycle in G. Let $E_{\mathrm{br}}(G)$ be the collection of bridges in G. Note that if T is a connected component for the subgraph induced by $E_{\mathrm{br}}(G)$, then T must be a tree; in this case we say T is a *bridging-tree*. Also, we denote by T_v the bridging-tree containing v (this is well-defined since different bridging-trees are vertex disjoint). In addition, we let $\partial_{\mathrm{i}}\mathsf{T}$ be the *internal boundary* of T, consisting of vertices in T which are neighboring to some vertex outside of T. Furthermore, if B is a connected component for the subgraph induced by $E(G) \setminus E_{\mathrm{br}}(G)$, we say B is a *block* of G. We note that u, v are in the same block if and only if there is a cycle (not necessarily simple) containing u and v. Partly for the purpose of facilitating our understanding, we make some simple observations: (1) different blocks are vertex disjoint (so are different bridging-trees as we pointed out earlier); (2) a bridging-tree and a block share no common edge; (3) For a bridging-tree T, we have that $u \in \partial_{\mathrm{i}}\mathsf{T}$ if and only if there is a (unique) block B such that $u = V(\mathsf{T}) \cap V(\mathsf{B})$ (this is implied by Lemma 13 (i) below).

We are now ready to define strong-admissibility which guarantees admissibility (as shown in Proposition 12).

Definition 11: Let $L, r \geq 1$ and $\rho \in (0,1)$. We say G is (r, ρ, L)-strongly-admissible if the following properties hold.

(1) For every $v \in V_{\mathrm{b}}(G)$, if $\mathsf{N}_{\rho r}(v)$ has two ρr-arms or there is a cycle in $\mathsf{N}_{\rho r}(v)$ containing v, then there exists a unique simple cycle O in $\mathsf{N}_{\rho r}(v)$ containing v and moreover $\mathrm{Length}(\mathsf{O}) \leq L$. In addition, the connected component of v in the subgraph induced by $E(G) \backslash E(\mathsf{O})$ is a bridging-tree in G (namely T_v) satisfying $H(\mathsf{T}_v) \leq L$ and $\partial_{\mathrm{i}}\mathsf{T}_v = \{v\}$.

(2) There are at most $\log r$ vertices in G which are contained in cycles of lengths less than L.

Proposition 12: There exists $r_0 = r_0(\rho, L) \geq 1$ such that if G is (r, ρ, L)-strongly-admissible for some $r \geq r_0$, then G is (r, ρ)-admissible.

The rest of this section is devoted to the proof of Proposition 12. To this end, we need the following two lemmas whose proofs are postponed until the end of this section. Recall that we have assumed G is connected and non-degenerate.

Lemma 13: The following hold for a bridging-tree T of G:

(i) For each $u \in \partial_{\mathrm{i}}\mathsf{T}$, let G_u be the connected component of u in the subgraph of G induced by edges in $E(G) \backslash E(\mathsf{T})$. Then

$$V(G) = \bigcup_{u \in \partial_{\mathrm{i}}\mathsf{T}} V(G_u) \cup V(\mathsf{T})$$
$$E(G) = \bigcup_{u \in \partial_{\mathrm{i}}\mathsf{T}} E(G_u) \cup E(\mathsf{T}) \qquad \text{(III.3)}$$

and $V(G_u) \cap V(\mathsf{T}) = \{u\}$, $V(G_{u_1}) \cap V(G_{u_2}) = \emptyset$ for $u_1, u_2 \in \partial_{\mathrm{i}}\mathsf{T}$ and $u_1 \neq u_2$.

(ii) For $v, w \in \mathsf{T}$, there exists a unique path (denoted by $[v, w]$) from v to w in G. For $y \in [v, w]$, let $V_{v;y}$ be the collection of vertices u such that there is a path from u to v without visiting y. Then $V_{v;y} \cap V_{w;y} = \emptyset$.

For $v \in V_{\mathrm{b}}$, let $\mathsf{D}_v = \mathsf{D}(\mathsf{C}_{\mathrm{b}}(v))$ where $\mathsf{C}_{\mathrm{b}}(v)$ is the bad component containing v.

Lemma 14: There exists $r_0 = r_0(\rho, L) \geq 1$ such that for all $r \geq r_0$ the following hold provided that G is an (r, ρ, L)-strongly-admissible graph:

(i) There exists $x \in V_{\mathrm{g}}$ such that $\mathsf{N}_{\rho r}(x)$ has two ρr-arms.

(ii) If $x \in V_{\mathrm{g}} \cap \mathsf{D}_v$ and $\mathsf{N}_{\rho r}(x)$ has two ρr-arms, then $\mathsf{N}_{\rho r}(u) \subset \mathsf{N}_{r-1}(x)$ for all $u \in \mathsf{D}_v$.

We introduce yet another notation. Given two intersecting paths $P^1 = (u_0, u_1, \ldots, u_m)$ and $P^2 = (v_0, v_1, \ldots, v_\ell)$, let $k = \min\{j : u_j \in P^2\}$ and let k' be such that $u_k = v_{k'}$. We define $g(P^1, P^2)$ to be the path $(u_0, u_1, \ldots, u_k = v_{k'}, v_{k'+1}, \ldots, v_\ell)$. We are now ready to prove Proposition 12.

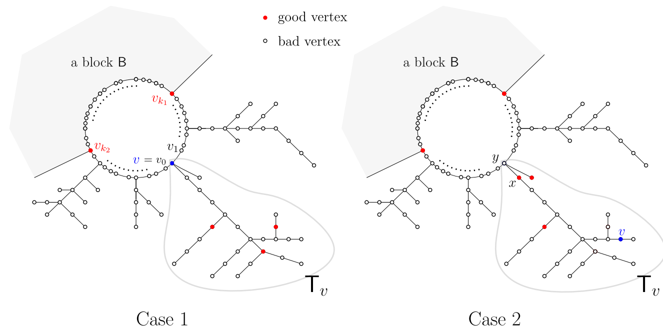


Fig. 4. Case 1 and Case 2 in Proposition 12.

Proof of Proposition 12: It suffices to show that for all $v \in V_{\mathrm{b}}(G)$, there exists $x \in V_{\mathrm{g}} \cap \mathsf{D}_v$ such that

$$\mathsf{N}_{\rho r}(u) \subset \mathsf{N}_{r-1}(x) \quad \text{for all } u \in \mathsf{D}_v \,. \qquad \text{(III.4)}$$

The proof of (III.4) proceeds as analysis by cases. The easier case is that v is contained in a cycle. If not, in order to facilitate our analysis we will further consider the scenarios regarding to the bridging-tree containing v. To this end, we claim that for any bridging-tree T, either $\partial_{\mathrm{i}}\mathsf{T} = \{y\}$ for some $y \in V_{\mathrm{b}}$ or $\partial_{\mathrm{i}}\mathsf{T} \subset V_{\mathrm{g}}$. To see this, note that if there is $y \in V_{\mathrm{b}} \cap \partial_{\mathrm{i}}\mathsf{T}$, then there is a simple cycle containing y. If the length of this cycle is $> 2\rho r$, then y has two ρr-arms; if not, then this cycle is contained in $\mathsf{N}_{\rho r}(y)$. Thus, by Definition 11 we have $\mathsf{T} = \mathsf{T}_y$ and $\partial_{\mathrm{i}}\mathsf{T} = \{y\}$, verifying the claim. In light of this claim, we only need to consider the following three cases for v (see Figures 4 and 5 for illustrations).

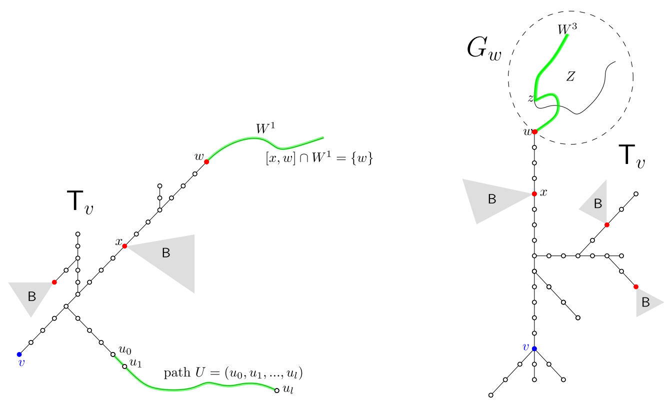
Case 1: v is contained in a cycle. By Definition 11, v is contained in a simple cycle $\mathsf{O} = (v_0 = v, v_1, \ldots, v_m = v)$ with $m \leq L$. We claim that there must exist $u \in \mathsf{O} \cap V_{\mathrm{g}}$. Otherwise, we have $H(\mathsf{T}_w) \leq L$ for all $w \in \mathsf{O}$ by Definition 11. This would imply that the diameter of G is at most $3L$ and thus G is degenerate (assuming that r is sufficiently large), arriving at a contradiction.

Now, let $k_1 = \inf\{k \geq 1 : v_k \in V_{\mathrm{g}}\}$ and $k_2 = \sup\{k \geq 1 : v_k \in V_{\mathrm{g}}\}$. Clearly $v_{k_1} \in V_{\mathrm{g}} \cap \mathsf{D}_v$. By Definition 11, if a path starting from v_j (with $j < k_1$ or $j > k_2$) does not visit v_{k_1} or v_{k_2}, then the path is contained in $\cup_{k<k_1, k>k_2} V(\mathsf{T}_{v_k})$ where each tree satisfies $H(\mathsf{T}_{v_k}) \leq L$. Thus, we have $\mathsf{D}_v \subset \cup_{k<k_1, k>k_2} V(\mathsf{T}_{v_k})$. Since $m \leq L$, we see that $\mathsf{D}_v \subset \mathsf{N}_{2L}(v_{k_1})$ and hence (III.4) holds with $x = v_{k_1}$ (assuming that r is sufficiently large).

Case 2: v is not contained in any cycle and $\partial_{\mathrm{i}}\mathsf{T}_v = \{y\}$ for some $y \in V_{\mathrm{b}}$. If $\mathsf{D}_y = \mathsf{D}_v$, then this reduces to Case 1. Thus, we may assume in addition that $\mathsf{D}_v \neq \mathsf{D}_y$. This implies that on the unique path from v to y, there exists $x \in V_{\mathrm{g}}(\mathsf{D}_v) \cap V(\mathsf{T}_v)$, and hence $\mathsf{D}_v \subset \mathsf{T}_v = \mathsf{T}_y$. Since $H(\mathsf{T}_y) \leq L$ (as $y \in \partial_{\mathrm{i}}\mathsf{T}_v$ is contained in a cycle), we have $\mathsf{D}_v \subset \mathsf{N}_{2L}(x)$ and hence (III.4) holds.

Case 3: v is not contained in any cycle and $\partial_{\mathrm{i}}\mathsf{T}_v \subset V_{\mathrm{g}}$. In this case, we have $\mathsf{D}_v \subset \mathsf{T}_v$, since by (III.3) for $u \in \partial_{\mathrm{i}}\mathsf{T}_v$ one has that every path from v to some $w \in G_u \backslash \{u\}$ visits u (Recall that G_u is defined in Statement (i) of Lemma 13). We further divide Case 3 into two subcases.

Case 3.A: there exists $w \in \mathsf{T}_v$ with two ρr-arms. If $w \in V_{\mathrm{b}}$, then by Definition 11 w must be contained in a cycle and thus



Case 3.A. Case 3.B.

Fig. 5. Case 3 in Proposition 12.

$w \in \partial_{\rm i}\mathsf{T}_v$, contradicting the assumption that $\partial_{\rm i}\mathsf{T}_v \subset V_{\rm g}$. As a result, w must be good.

Let $[v,w]$ be the unique path from v to w (using Lemma 13 (ii)), and let x be the good vertex on $[v,w]$ which is closest to w. By the choice of "closest", we have $x \in \mathsf{D}_v$. Let W^1, W^2 be the two ρr-arms of w. By Lemma 13 (ii), we have either $W^1 \cap [w,x] = \{w\}$ or $W^2 \cap [w,x] = \{w\}$ since otherwise W^1 and W^2 will not be edge-disjoint. Without loss of generality we assume that $W^1 \cap [w,x] = \{w\}$.

If (III.4) fails, there exists a path $U = (u_0, u_1, \cdots, u_\ell)$ with $\ell < \rho r$ such that $u_0 \in \mathsf{D}_v \subset \mathsf{T}_v$ and $u_\ell \notin \mathsf{N}_{r-1}(x)$. On the one hand, using the notation in Lemma 13 (ii) we have $u_0 \in V_{v;x}$. Noting that $u_j \neq x$ for all j, we then get $u_j \in V_{v;x}$ for all j. Let $P_{x\to u_0}$ be a path from x to u_0 and let $P_{x\to u_\ell} = g(P_{x\to u_0}, U)$. Then the length of $P_{x\to u_\ell}$ satisfies $\mathrm{Length}(P_{x\to u_\ell}) > \rho r$ since $u_\ell \notin \mathsf{N}_{r-1}(x)$. In addition, $P_{x\to u_\ell} \setminus \{x\} \subset V_{v;x}$. On the other hand, the path $[x,w] \cup W^1$ has length $> \rho r$ and the vertices on this path (except x) are all in $V_{w;x}$. Thanks to Lemma 13, $V_{v;x} \cap V_{w;x} = \emptyset$. Altogether, we get that x has two ρr-arms; combining with $x \in V_{\rm g}$ yields (III.4) by Lemma 14.

<u>Case 3.B</u>: there exists no vertex in T_v with two ρr-arms. By Lemma 14, there exists $z \in G$ such that $\mathsf{N}_{\rho r}(z)$ has two ρr-arms. Then $z \notin \mathsf{T}_v$. Let $P_{v\to z}$ be a path from v to z and let w be the last vertex on $P_{v\to z}$ such that $w \in \mathsf{T}_v$. Then clearly $w \in \partial_{\rm i}\mathsf{T}_v$ and the subpath $P_{w\to z} \subset G_w$. Since z has two ρr-arms, there is a path Z with $2\rho r$ edges and with z the middle point. Let w' be the first point at which $P_{w\to z}$ intersects Z, and let W^3 be the path obtained by concatenating the subpath of $P_{w\to z}$ from w to w' and the longer subpath of Z separated by w'. Then $W^3 \subset G_w$ and has length $> \rho r$. For our v, there is a unique path $[v,w]$ from v to w and we let x be the closest good vertex to v on $[v,w]$. Then $x \in \mathsf{D}_v$ by definition. Using the same argument in the Case 3.A (replacing W^1 by W^3) we can show that (III.4) holds. $\square$

Proof of Lemma 13: We will employ proof by contradiction.

We first prove (i). By definition, $E(G_u) \cap E(\mathsf{T}) = \emptyset$ and $u \in V(G_u) \cap V(\mathsf{T})$. If $w \in V(G_u) \cap V(\mathsf{T})$ for some $w \neq u$, then there exists a path connecting u and w with edges in $E(\mathsf{T})$ and another path connecting u and w with edges in $E(G_u)$. Thus, these two paths altogether form a cycle, contradicting the definition of bridging-tree. Therefore, $V(G_u) \cap V(\mathsf{T}) = \{u\}$.

For $u_1 \neq u_2$ in $\partial_{\rm i}\mathsf{T}$, if $V(G_{u_1}) \cap V(G_{u_2}) \neq \emptyset$, then $V(G_{u_1}) = V(G_{u_2})$ (since G_{u_1} and G_{u_2} are connected components). Since we have shown that $V(G_{u_i}) \cap V(\mathsf{T}) = \{u_i\}$ for $i = 1,2$, this yields a contradiction. Therefore, $V(G_{u_1}) \cap V(G_{u_2}) = \emptyset$ and consequently $E(G_{u_1}) \cap E(G_{u_2}) = \emptyset$.

For $e \in E(G) \setminus E(\mathsf{T})$, let z_0 be an end-vertex of e with $z_0 \notin V(\mathsf{T})$. Since G is connected, there exists a path $(z_0, z_1, \ldots, z_m)$ so that z_m is the only vertex in T. This implies that $u = z_m \in \partial_{\rm i}\mathsf{T}$. In addition, by definition $e \in E(G_u)$ and $z_0 \in V(G_u)$. Hence the decomposition (III.3) holds.

We next prove (ii). Let $[v,w]$ be the path from v to w in the bridging-tree T. If $(v = v_0, v_1, \ldots, v_m = w)$ is another path in G, let $k_1 = \inf\{j : v_j \notin [v,w]\}$ and let $k_2 = \inf\{j > k_1 : v_j \in [v,w]\}$. Then we can see that the path $[v_{k_2}, v_{k_1}]$ in T and $(v_{k_1}, v_{k_1+1}, \ldots, v_{k_2})$ form a simple cycle, contradicting with the definition of bridging-tree. Finally, for $y \in [v,w]$, if $V_{v;y} \cap V_{w;y} \neq \emptyset$, then there exists a path from v to w without visiting y, contradicting with the uniqueness of the path $[v,w]$. $\square$

Proof of Lemma 14: We first prove (i). Let P be the longest path in G. Since G is non-degenerate, we see that the length of P is at least $2r-2$ since otherwise $G \subset \mathsf{N}_{r-1}(u)$ (and thus G is degenerate) where u is the mid-point of P (or one of the two mid-points of P if P has an even number of vertices). Therefore, there are at least $(r - \rho r - 1) \geq \log r$ (assuming that r is sufficiently large) many vertices on P which have two ρr-arms. By Definition 11, one of them must be good.

We now prove (ii). We claim that, for every $y \notin \mathsf{N}_{r-1}(x)$ and every path $P_{y\to x} = (y_0 = y, y_1, \ldots, y_m = x)$ from y to x of length m, there exists $j \in [m - \log r, m)$ such that $y_j \in V_{\rm g}$. Provided with this claim we now show that $\mathsf{N}_{\rho r+1}(v) \subset \mathsf{N}_{r-1}(x)$ if $x \in V_g \cap \mathsf{D}_v$. Denote by dist the graph distance on G. When $\mathrm{dist}(v,x) < \frac{(1-\rho)r}{4}$, for $u \in N_{\rho r+1}(v)$ we apply the triangle inequality and get that $\mathrm{dist}(u,x) \leq \mathrm{dist}(u,v) + \mathrm{dist}(v,x) < \rho r + 1 + (1-\rho)r/4 < r-1$ for r large enough (recall $\rho < 1$), yielding that $\mathsf{N}_{\rho r+1}(v) \subset \mathsf{N}_{r-1}(x)$. Thus we may assume $\mathrm{dist}(v,x) \geq \frac{(1-\rho)r}{4}$. Since $x \in \mathsf{D}_v$, there exists a path $(x_0 = x, x_1, \ldots, x_{m'} = v)$ such that $x_i \in V_{\rm b}$ for $1 \leq i \leq m'$. If there exists $w \in \mathsf{N}_{\rho r+1}(v) \setminus \mathsf{N}_{r-1}(x)$, then there exists a geodesic $(v_0 = v, v_1, \ldots, v_\ell = w)$ from v to w. Let $k = \inf\{i : x_i \in \{v_j\}\}$ and let k' be such that $x_k = v_{k'}$. By the triangle inequality, we have $k \geq \mathrm{dist}(x,w) - \mathrm{dist}(x_k,w)$. Since $w \in N_{\rho r+1}(v) \setminus N_{r-1}(x)$, we have $\mathrm{dist}(x,w) \geq r$. Since x_k is on the geodesic $(v = v_0, v_1, \ldots, v_\ell = w)$, we have $\mathrm{dist}(x_k,w) \leq \mathrm{dist}(v,w) \leq \rho r + 1$. Thus, by the triangle inequality again $\mathrm{dist}(x,w) - \mathrm{dist}(x_k,w) \geq r - (\rho r + 1) \geq \frac{(1-\rho)r}{4}$. Note that $(x_0, \ldots, x_k = v_{k'}, v_{k'+1}, \ldots, v_\ell)$ is a path connecting x and $w \notin \mathsf{N}_{r-1}(x)$. Applying the claim to the reverse of this path, we see that $\{x_i : 0 < i \leq k\} \cap V_{\rm g} \neq \emptyset$ (assuming r to be sufficiently large), arriving at a contradiction. Therefore, $\mathsf{N}_{\rho r+1}(v) \subset \mathsf{N}_{r-1}(x)$. As a consequence, for all $u \in \mathsf{D}_v \cap V_{\rm b}$ we have $\mathsf{N}_{\rho r+1}(u) \subset \mathsf{N}_{r-1}(x)$ since $\mathsf{D}_v = \mathsf{D}_u$ (and thus we can apply the above reasoning with v replaced by u). For $u \in \mathsf{D}_v \cap V_{\rm g}$, there exists $y \in \mathsf{D}_v \cap V_{\rm b}$ neighboring to u, so $\mathsf{N}_{\rho r}(u) \subset \mathsf{N}_{\rho r+1}(y) \subset \mathsf{N}_{r-1}(x)$.

It remains to prove the claim made at the beginning of the proof. Suppose X^1 and X^2 are two paths of length ρr and have the unique common vertex x. If the path $P_{y\to x}\setminus\{x\}$ does not intersect with X^2, then y_j has two ρr-arms $(y_0,\dots,y_j)$ and $(y_j,\dots,y_m = x_0^{(2)},\dots,x_{\rho r}^{(2)})$ for all $j \in [m-\log r, m)$. By Definition 11, one of these y_j's must be good (since the number of such y_j's is $\log r$), as desired. The case for $P_{u\to x}\setminus\{x\}$ does not intersect with X^1 can be treated similarly.

Finally, we consider the case when $P_{y\to x}\setminus\{x\}$ intersects with X^1 and X^2. Let $k_j = \sup\{i : y_i \in X^j \setminus \{x\}\}$ for $j = 1, 2$. Without loss of generality, assume that $k_2 > k_1$. Let ℓ be such that $u_{k_1} = x_\ell^{(1)}$. Note that $\mathsf{O} = (x_0^{(1)},\dots,x_\ell^{(1)} = y_{k_1},\dots,y_m = x_0^{(1)})$ is a simple cycle. Our proof proceeds by dividing into three cases depending on $\mathrm{Length}(\mathsf{O})$, i.e., the length of O.

Case 1: $\mathrm{Length}(\mathsf{O}) > 2\rho r$. Then all the vertices on O (which include y_j with $j \in [m-\log r, m)$) has two ρr-arms. By Definition 11, one of those y_j's must be good.

Case 2: $L < \mathrm{Length}(\mathsf{O}) < 2\rho r$. Then by Definition 11 y_{m-1} must be good (since y_{m-1} is on a cycle of length larger than L).

Case 3: $\mathrm{Length}(\mathsf{O}) \le L$. Then $k_1 > m - \log r$ and $(y_{k_1},\dots,y_0)$ is a path of length $> \rho r$ which has disjoint edges with O. By Definition 11, y_{k_1} must be good. □

IV. Admissibility for Erdős–Rényi Graphs

In order to complete the proof of identifiability in Theorem 1, it suffices to prove the following result in light of Lemma 9 and Proposition 12. In what follows, we say a graph is strongly-admissible, if each connected non-degenerate component of it is strongly-admissible.

Proposition 15: Fix $\lambda, \epsilon_0 > 0$ and assume that ρ, r satisfy (III.1). For any $\epsilon > 0$, there exist $N_\epsilon \in \mathbb{N}$ and $L = L_\epsilon \ge 1$ (both may depend on λ and ϵ_0) such that for all $n \ge N_\epsilon$,

$$\mathbb{P}(\mathcal{G}_{n,\frac{\lambda}{n}} \text{ is } (r,\rho,L)\text{-strongly-admissible}) \ge 1-\epsilon.$$

In the rest of the paper, unless otherwise specified, we assume that (III.1) holds.

A. Proof of Proposition 15

In this subsection, we prove Proposition 15 with postponing the proof for Lemma 16 (below) to later subsections. To this end, we will employ breadth-first-search (BFS) process simultaneously from a pair of vertices $u, v \in \mathcal{G}$ (which we refer to as *reduced BFS*). In order to approximate their r-neighborhoods by independent PGW(λ) trees, we will need some kind of "cut off" and "graft" operations as we describe in what follows.

We now describe our reduced BFS (with respect to $u, v \in \mathcal{G}$) which is a modification of the standard BFS. As a comment on notation below, we will denote by A_t for *active* vertices, i.e., we are going to explore their neighbors; we denote R_t for *removed* vertices, i.e., we will not explore their neighbors; we denote U_t for *unexplored* vertices, i.e., these vertices have not

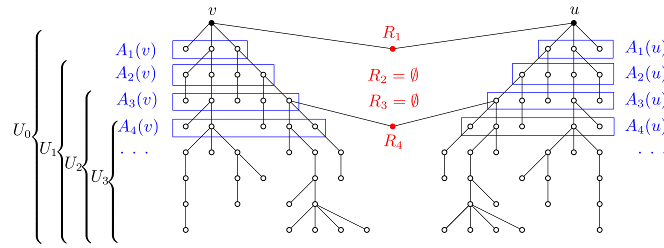


Fig. 6. Example of reduced BFS.

be explored as neighbors of some active vertices. Initially we set $R_0 = \emptyset$, $A_0(v) = \{v\}$, $A_0(u) = \{u\}$, $A_0 = A_0(u) \cup A_0(v)$ and $U_0 = V \setminus A_0$. For $t \ge 0$, as long as $A_t(u) \cup A_t(v) \ne \emptyset$, we make the following recursive definition (which corresponds to a "search process" from $A_t(u) \cup A_t(v)$) where the set R_t corresponds to the aforementioned "cut off" operation:

$$\begin{aligned}
&R_{t+1} = \{w \in U_t : \exists x \in A_t(u), y \in A_t(v), \mathcal{G}_{xw} = \mathcal{G}_{yw} = 1\};\\
&A_{t+1}(u) = \{w \in U_t : \exists x \in A_t(u), \mathcal{G}_{xw} = 1\}\setminus R_{t+1};\\
&A_{t+1}(v) = \{w \in U_t : \exists y \in A_t(v), \mathcal{G}_{yw} = 1\}\setminus R_{t+1};\\
&U_{t+1} = U_t \setminus (A_{t+1}(u) \cup A_{t+1}(v) \cup R_{t+1}). \qquad \text{(IV.1)}
\end{aligned}$$

(We note that $A_t(u), A_t(v), R_t$ for $t \ge 0$ are pairwise disjoint.)

We next inductively construct two rooted trees $\mathcal{T}_{\rm cut}(u)$ and $\mathcal{T}_{\rm cut}(v)$ with vertex sets $\cup_t A_t(u)$ and $\cup_t A_t(v)$ respectively. Recall that we have a pre-fixed ordering on V. For $t \ge 0$, we assume the first t-levels of $\mathcal{T}_{\rm cut}(v)$ and $\mathcal{T}_{\rm cut}(u)$ have been defined (and their t-th levels are $A_t(v)$ and $A_t(u)$, respectively). We arrange the vertices in $A_t(v)$ as $v_{t,1},\dots,v_{t,|A_t(v)|}$ according to our pre-fixed order. Then for each j, we add the edges

$$\big\{(v_{t,j}, y) : \mathcal{G}_{v_{t,j}y} = 1, y \in U_t \setminus R_{t+1} \text{ and } \mathcal{G}_{v_{t,i}y} = 0 \text{ for } 1 \le i < j\big\}$$

to $E(\mathcal{T}_{\rm cut}(v))$. Note that the $(t+1)$-th level of $\mathcal{T}_{\rm cut}(v)$ is exactly $A_{t+1}(v)$. This gives the tree $\mathcal{T}_{\rm cut}(v)$ and we define $\mathcal{T}_{\rm cut}(u)$ similarly.

Furthermore, we define the auxiliary tree $\mathcal{T}_{\rm aux}(v)$ as an enlargement of $\mathcal{T}_{\rm cut}(v)$ as follows. For each $y \in \mathcal{T}_{\rm cut}(v) \cap A_t(u)$, if there is $w \in R_{t+1}$ such that $\mathcal{G}_{yw} = 1$, we add an independent $\mathrm{Bin}(n, \lambda/n)$-Galton-Watson tree rooted at a *copy* of w (where all other vertices are labeled as ∞) to $\mathcal{T}_{\rm cut}(v)$ by connecting this copy of w to y (as the child of y). This corresponds to the aforementioned "graft" operation. (Note that it may be slightly more natural to add a PGW(λ) tree, and we chose to add a $\mathrm{Bin}(n, \lambda/n)$-GW tree just for the slight convenience of exposition in the proof of Lemma 25.)

Then we get a rooted tree $\mathcal{T}_{\rm aux}(v)$. Similarly, we can define $\mathcal{T}_{\rm aux}(u)$. Indeed, $\mathcal{T}_{\rm cut}(v)$ can be obtained as a subtree of $\mathcal{T}_{\rm aux}(v)$ by deleting all the vertices whose labels also appear in $\mathcal{T}_{\rm aux}(u)$ (and similarly for $\mathcal{T}_{\rm cut}(u)$).

Suppose that $u, v \in \mathcal{G}$ and $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$. Let ϕ be an isomorphism from $\mathsf{N}_{r+1}(u)$ to $\mathsf{N}_{r+1}(v)$ such that $\phi(u) = v$. We claim that

$$\phi(A_t(u)) = A_t(v) \text{ and } \phi(R_t) = R_t \text{ for all } t \le r. \qquad \text{(IV.2)}$$

Clearly (IV.2) holds for $t = 0$. Assume that (IV.2) holds for $0,\dots,t-1$, and $t \le r$. Then for any $y \in A_t(u)$, by (IV.1) and by our choice of ϕ we have $\phi(y) \in A_t(v) \cup R_t$. If $\phi(y) \in R_t$, then there is $z \in A_{t-1}(u)$ such that $\mathcal{G}_{z,\phi(y)} = 1$. Since $\phi(y) \in R_t$ and $t \leq r$, we have $z \in \mathsf{N}_{r+1}(v)$, and hence $(z, \phi(y)) \in E(\mathsf{N}_{r+1}(v))$. Hence $(\phi^{-1}(z), y) \in E(\mathsf{N}_{r+1}(u))$. By our induction hypothesis $\phi^{-1}(z) \in A_{t-1}(v)$, so $y \in R_t$, which contradicts with $y \in A_t(u)$. Thus $\phi(y) \in A_t(v)$. This implies that $\phi(A_t(u)) = A_t(v)$, and as a result we also have $\phi(R_t) = R_t$.

We need some more quantities to describe the structure of the r-neighborhoods of u, v. Let $A_{\leq r}(u) = \cup_{t=0}^r A_t(u)$, $R_{\leq r} = \cup_{t=0}^r R_t$, and $A_{\leq r} = A_{\leq r}(u) \cup A_{\leq r}(v)$.

- Let $\Xi_1 = \sum_{t=0}^r \sum_{x \in A_t(u), y \in A_t(v)} \mathcal{G}_{xy}$. If $x \in A_t(u), y \in A_{t'}(v)$ and $\mathcal{G}_{xy} = 1$, by the definition of reduced BFS (see (IV.1)) we have $t = t'$. Hence $\Xi_1 = \sum_{x \in A_{\leq r}(u), y \in A_{\leq r}(v)} \mathcal{G}_{xy}$.
- Let $\Xi_2 = \sum_{t=0}^{r-1} \sum_{w \in U_t} \mathcal{G}_{w,A_t(u)} \mathcal{G}_{w,A_t(v)}$, where $\mathcal{G}_{w,A} = \sum_{y \in A} \mathcal{G}_{wy}$ for a subset A. We can see that $|R_{\leq r}| = \sum_{t=0}^{r-1} \sum_{w \in U_t} 1_{\{\mathcal{G}_{w,A_t(u)} \geq 1\}} 1_{\{\mathcal{G}_{w,A_t(v)} \geq 1\}}$, and hence $|R_{\leq r}| \leq \Xi_2$.
- Let $\Lambda_1(u) = \mathrm{Comp}(\mathcal{G}[A_{\leq r}(u)])$, where $\mathcal{G}[A_{\leq r}(u)]$ is the subgraph on $\mathcal{G}$ induced by $A_{\leq r}(u)$ and (we recall that) $\mathrm{Comp}(G)$ is the *complexity* for a graph G (that is, $\mathrm{Comp}(G)$ is the minimal number of edges that one has to remove from G so that no cycle remains). Similarly let $\Lambda_1(v) = \mathrm{Comp}(\mathcal{G}[A_{\leq r}(v)])$. Let $\Lambda_1 = \Lambda_1(u) + \Lambda_1(v)$. By (IV.2), if $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$, we have $\Lambda_1(u) = \Lambda_1(v)$.
- Write $A_{[t,r]}(u) = \cup_{s=t}^r A_s(u)$ for $t \leq r$ and define
$$\Lambda_2(u) = \sum_{t=0}^{r} \sum_{w \in R_t, x \in A_{[t,r]}(u)} \mathcal{G}_{w,x} + \sum_{x \in A_{\leq r}(u)} \sum_{w \in R_{\leq r}} \sum_{y \in \mathsf{N}_r(w) \cap \mathsf{N}_r(u) \setminus (A_{\leq r} \cup R_{\leq r})} \mathcal{G}_{xy}\,.$$
Similarly we define $\Lambda_2(v)$. We can see that when $\Xi_2 = |R_{\leq r}|$ and $\Lambda_1(u) = \Lambda_2(u) = 0$, for each $w \in R_{\leq r}$ there exists a unique path in $\mathcal{G}[A_{\leq}(u)]$ from u to w. Let $\Lambda_2 = \Lambda_2(u) + \Lambda_2(v)$. By (IV.2), if $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$, then $\Lambda_2(u) = \Lambda_2(v)$.
- Let Λ_3 be the indicator function of the event that there are $w_1 \neq w_2$ in $R_{\leq r}$ which are connected by a path in $(\mathsf{N}_r(u) \cup \mathsf{N}_r(v)) \setminus A_{\leq r}$.

Finally, let $\Xi(u, v) = \Xi = \Xi_1 + \Xi_2$ and let $\Lambda(u, v) = \Lambda = \Lambda_1 + \Lambda_2 + \Lambda_3$.

Lemma 16: Assume (III.1). For any $\epsilon > 0$, there exist $N_\epsilon \in \mathbb{N}$ and $L_\epsilon \geq 1$ (which may depend on λ and ϵ_0) such that for every $n \geq N_\epsilon$ the following holds with probability at least $1 - \epsilon$. For any two vertices u, v in $\mathcal{G}_{n,\frac{\lambda}{n}}$ such that $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$ and $\mathsf{N}_r(v)$ survives (recall that this means $\mathsf{N}_r(v) \neq \mathsf{N}_{r-1}(v)$), we have

(i) $\Xi \leq 2$. As a consequence, $\Xi_2 = |R_{\leq r}|$.
(ii) If $\Xi = 0$, then $\Lambda = 0$, and u does not have two r-arms.
(iii) If $\Xi = 1$, then $\Lambda = 0$, and $H(\mathcal{T}_{\mathrm{cut}}(v)) < \rho' r$, where $\rho' = \rho'(\epsilon_0) = \frac{1+\epsilon_0/2}{1+\epsilon_0}$.
(iv) If $\Xi = 2$, then $\Lambda = 0$, and $H(\mathcal{T}_{\mathrm{cut}}(v)) \leq L_\epsilon$.

Proof of Proposition 15: Assuming that $\mathcal{G}$ satisfies the properties in the statement of Lemma 16, we claim that for every non-degenerate v with non-unique $(r+1)$-neighborhood, if v has two r-arms or there is a cycle in $\mathsf{N}_{r+1}(v)$ containing v, then there exists a unique simple cycle O in $\mathsf{N}_{r+1}(v)$ containing v with $\mathrm{Length}(\mathsf{O}) \leq 4L_\epsilon$ and in addition the connected component of v in the subgraph induced by $E(\mathcal{G}) \backslash E(\mathsf{O})$ is a bridging-tree in $\mathcal{G}$ (which is T_v) satisfying $H(\mathsf{T}_v) \leq L_\epsilon$ and $\partial_{\mathrm{i}} \mathsf{T}_v = \{v\}$. Since $r = \frac{(1+\epsilon_0)\log n}{\log \alpha_\lambda^{-1}}$ and we choose ϵ_0 arbitrarily, the claim above also holds if we replace r by $\rho r - 1$ (the replacement also occurs in Lemma 16), which is (1) in Definition 11 (with L replaced by $4L_\epsilon$). In addition, it is well known that the number of cycles of length ℓ in $\mathcal{G}$ converges to a Poisson random variable (see e.g., [8, Corollary 4.9]). Thus, (2) in Definition 11 holds. Therefore, it remains to prove the above claim.

Assume that $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$ for some $u \neq v$ in $\mathcal{G}$. Then we do the reduced BFS for u, v. By Lemma 16, we have $\Xi = \Xi(u, v) \leq 2$. We next show that Ξ is not 0 or 1.

- If $\Xi(u, v) = 0$, then by Lemma 16 $\Lambda(u, v) = 0$ and v does not have two r-arms. Also, $\Xi = \Lambda = 0$ implies that v is not on any cycle, arriving at a contradiction.
- If $\Xi(u, v) = 1$, by Lemma 16 we have $\Lambda(u, v) = 0$ and $H(\mathcal{T}_{\mathrm{cut}}(v)) \leq \rho' r$. When $\Xi = \Xi_2 = 1$, let $P_{v \to w}$ be the unique path from v to the vertex $w \in R_{\leq r}$; when $\Xi = \Xi_1 = 1$, let $P_{v \to w}$ be the unique path to the vertex $w \in A_{\leq r}(v)$ such that there exists $y \in A_{\leq r}(u)$ with $\mathcal{G}_{wy} = 1$. In both scenarios, every path starting from v and not contained in $\mathcal{T}_{\mathrm{cut}}(v)$ must contain $P_{v \to w}$. If there is a simple cycle containing v, then this cycle also contains a vertex $x \notin \mathcal{T}_{\mathrm{cut}}(v)$ as $\Lambda_1 = 0$. Thus, there are two edge-disjoint paths from v to x both of which contain $P_{v \to w}$, arriving at a contradiction. Therefore, v is not contained in any simple cycle. In addition, since $\Lambda = 0$ and $H(\mathcal{T}_{\mathrm{cut}}(v)) \leq \rho' r$, every path from v with length r must contain a vertex not in $\mathcal{T}_{\mathrm{cut}}(v)$. By the same argument, v does not have two r-arms either, arriving at a contradiction.

Therefore it must be $\Xi = 2$ and $\Lambda = 0$ by Lemma 16. We may assume that $\Xi = \Xi_2 = |R_{\leq r}| = 2$ and write $R_{\leq r} = \{w_1, w_2\}$ (the other case can be proved in the same manner). For $i = 1, 2$, let $P_{v \to w_i}$ be the path from v to w_i consisting of vertices in $\mathcal{T}_{\mathrm{cut}}(v) \cup \{w_i\}$, which is unique as $\Lambda = 0$. We similarly define $P_{u \to w_i}$ (note that here for instance $P_{u \to w_1}$ is the reverse of $P_{w_1 \to u}$). Then $\mathsf{O} = P_{v \to w_1} \cup P_{w_1 \to u} \cup P_{u \to w_2} \cup P_{w_2 \to v}$ is a cycle in $\mathsf{N}_r(v)$ containing v. Since $H(\mathcal{T}_{\mathrm{cut}}(v)) \leq L_\epsilon$, we have $\mathrm{dist}(w_i, v) \leq L_\epsilon$ for $i = 1, 2$. Similarly we have $\mathrm{dist}(w_i, u) \leq L_\epsilon$ for $i = 1, 2$. Therefore, $\mathrm{Length}(\mathsf{O}) \leq 4L_\epsilon$.

We now show that O is simple. It suffices to show that $P_{v \to w_1} \cap P_{v \to w_2} = \{v\}$ and $P_{u \to w_1} \cap P_{u \to w_2} = \{u\}$. We will prove the statement for v (and that for u can be proved similarly), which is divided into two cases.

- If v has two r-arms, each of the two r-arms must visit a vertex not in $\mathcal{T}_{\mathrm{cut}}(v)$ since $H(\mathcal{T}_{\mathrm{cut}}(v)) \leq L_\epsilon$. Since $\Lambda = 0$, these two r-arms must contain $P_{v \to w_1}$, $P_{v \to w_2}$, respectively. So $P_{v \to w_1} \cap P_{v \to w_2} = \{v\}$.
- If v is contained by a simple cycle in $\mathsf{N}_{r+1}(v)$, from $\Lambda = 0$ we see there exists a vertex $y \in A_{\leq}(u)$ on this

simple cycle. Thus from v to y there are two paths only intersecting at v and y. Since $\Lambda = 0$, these two paths must contain $P_{v\to w_1}$, $P_{v\to w_2}$, respectively. So $P_{v\to w_1} \cap P_{v\to w_2} = \{v\}$.

Combining with $\Lambda = 0$ and $\Xi = 2$, we can then further deduce that O is the unique simple cycle containing v in $\mathsf{N}_{r+1}(v)$.

In addition, the connected component of v in $E(\mathcal{G})\backslash E(\mathsf{O})$ is $\mathcal{T}_{\mathrm{cut}}(v)\backslash(P_{v\to w_1}\cup P_{v\to w_2})$. Thus this component is a tree and has height at most L_ϵ. Since $\Lambda = 0$, every edge in this tree is not contained by any cycle in $\mathcal{G}$, yielding that this tree is a bridging-tree in $\mathcal{G}$ (i.e, it is T_v). This completes the proof of (1) in Definition 11. □

It remains to prove Lemma 16. To this end, we prove additional properties for PGW trees in Section IV-B and then provide the proof for Lemma 16 in Section IV-C.

B. Additional Properties for Galton-Watson Trees

We need to control the volume growth for a Galton-Watson tree as incorporated in Lemma 17; this is fairly standard and we include a proof only for completeness. The new ingredient of significance to us is the conditioning on isomorphism as in Lemma 19.

Lemma 17: Fix $\lambda \geq 1$. Let $(Z_\ell)_{\ell\geq 0}$ be the number of vertices in the ℓ-th level of a PGW(λ)-tree. Then for every $m \geq 1$, there is a constant $C_{m,\lambda} > 0$ depending only on λ and m (we denote by $C_m = C_{m,1}$ for short) such that

(i) when $\lambda = 1$, $\mathbb{E}[Z_\ell^m] \leq C_m \ell^{m-1}$ for all $\ell \geq 1$;
(ii) when $\lambda > 1$, $\mathbb{E}[Z_\ell^m] \leq C_{m,\lambda}\lambda^{\ell m}$ for all $\ell \geq 1$.

As a consequence, for every $\theta > \log(\lambda) \geq 0$,

$$\lim_{r\to\infty}\frac{1}{r}\log\mathbb{P}\big(\textstyle\sum_{\ell=0}^r Z_\ell > e^{\theta r}\big) = -\infty\,. \tag{IV.3}$$

Remark 18: The same proof for Lemma 17 below easily gives the same result for a Galton-Watson tree with offspring distribution as $\mathrm{Bin}(n,\lambda/n)$. As a result, in what follows we also apply Lemma 17 in this case.

Proof of Lemma 17: We prove (i) and (ii) by induction on m. The base case for $m = 1$ holds obviously. Assume they hold for $1,\cdots,m-1$. Since the conditional law of Z_ℓ given $Z_{\ell-1}$ is Poisson with mean $\lambda Z_{\ell-1}$, using the factorial moments of Poisson random variables we have that

$$\mathbb{E}[Z_\ell(Z_\ell-1)\dots(Z_\ell-m+1)|Z_{\ell-1}] = \lambda^m Z_{\ell-1}^m\,.$$

Let $C'_m = \sum_{j=0}^m\binom{m}{j}(m+1)^m$ we have

$$\mathbb{E}[Z_\ell^m] \leq \lambda^m\mathbb{E}[Z_{\ell-1}^m] + C'_m\sum_{1\leq j\leq m-1}\mathbb{E}[Z_\ell^j]\,.$$

Then by our induction hypothesis we get that for all $\ell \geq 1$

$$\mathbb{E}[Z_\ell^m] \leq \mathbb{E}[Z_{\ell-1}^m] + (mC'_m\max_{j\leq m-1}C_j)\ell^{m-2}\ \text{for}\ \lambda = 1\,,$$
$$\mathbb{E}[(\tfrac{Z_\ell}{\lambda^\ell})^m] \leq \mathbb{E}[(\tfrac{Z_{\ell-1}}{\lambda^{\ell-1}})^m] + (mC'_m\max_{j\leq m-1}C_{j,\lambda})\tfrac{1}{\lambda^\ell}\ \text{for}\ \lambda > 1\,.$$

This completes the proof of (i) and (ii) by induction and by choosing C_m and $C_{m,\lambda}$ appropriately.

We next prove (IV.3). For every $m \geq 1$, writing $\beta_\ell = \frac{\ell^{-2}}{\sum_{i\geq 1}i^{-2}}$ we have that

$$\mathbb{P}\left(\sum_{\ell=0}^r Z_\ell > e^{\theta r}\right) \leq \sum_{\ell=0}^r\mathbb{P}\left(Z_\ell > e^{\theta r}\beta_\ell\right) \leq e^{-m\theta r}\sum_{\ell=0}^r\mathbb{E}[Z_\ell^m\beta_\ell^{-m}]\,.$$

Then using (i) and (ii), we have

$$\begin{aligned}&\limsup_{r\to\infty}\frac{1}{r}\log\mathbb{P}\Big(\sum_{\ell=0}^r Z_\ell > e^{\theta\ell}\Big)\\ &\leq -m\theta + \limsup_{r\to\infty}\frac{1}{r}\log\left(\sum_{\ell=0}^r\mathbb{E}[Z_\ell^m\beta_\ell^{-m}]\right)\\ &= -m\theta + \limsup_{r\to\infty}\frac{1}{r}\log\left(\sum_{\ell=0}^r\frac{C_{m,\lambda}\lambda^{\ell m}+C_m\ell^{m-1}}{\beta_\ell^m}\right)\\ &\leq -m[\theta - \log(\lambda)]\,,\end{aligned}$$

which implies (IV.3) by sending $m\to\infty$. □

In what follows, for a rooted tree T we write $Z_m(T)$ to denote the number of the vertices in the m-th level of T and write $Z_{\leq m}(T)$ to denote the number of vertices in the first m-levels. Let $\mathbf{T}$, $\mathbf{T}'$ be independent PGW(λ) trees. We next control the volume growth conditioned on isomorphism and also heights of the trees.

Lemma 19: Fix $\lambda > 0$. For $m \geq 0$, there exists a constant $C_m = C_m(\lambda) > 0$ such that

$$\mathbb{E}([Z_\ell(\mathbf{T})]^m|\mathbf{T}\sim_\ell\mathbf{T}') \leq C_m\ell^m\ \text{for all}\ \ell\geq 1\,. \tag{IV.4}$$

As a consequence, $\mathbb{E}([Z_{\leq\ell}(\mathbf{T})]^m|\mathbf{T}\sim_\ell\mathbf{T}') \leq C_m\ell^{2m+2}$.

Proof: The main task is to prove (IV.4) by induction. The case of $m = 0$ is trivial. Assume that (IV.4) holds for all $1,\dots,m-1$. Using the same argument (and notations) in the proof of Lemma 5, we have

$$\begin{aligned}&a_\ell(m) := \mathbb{E}[Z_\ell(\mathbf{T})^m;\mathbf{T}\sim_\ell\mathbf{T}']\\ &= \sum_{k\geq 1}\mathbb{E}[Z_\ell(\mathbf{T})^m;\mathbf{T}\sim_\ell\mathbf{T}'|D=D'=k]\mu_k^2\\ &\leq \sum_{k\geq 1}k^2\mu_k^2\mathbb{E}\{[Z_{\ell-1}(\mathbf{T}_1)+Z_\ell(\mathbf{T}\backslash\mathbf{T}_1)]^m;\\ &\qquad \mathbf{T}_1\sim_{\ell-1}\mathbf{T}'_1,(\mathbf{T}\backslash\mathbf{T}_1)|_\ell\sim(\mathbf{T}'\backslash\mathbf{T}'_1)|_\ell|D=D'=k\}\,.\end{aligned} \tag{IV.5}$$

Here in the inequality, recalling (II.9),and applying a union bound on (i,j) gives the estimate (in particular, the union bound leads to the factor of k^2). By independence among different subtrees of PGW tree, we get that for each $0\leq j\leq m$,

$$\begin{aligned}&\mathbb{E}\Big[Z_{\ell-1}(\mathbf{T}_1)^jZ_\ell(\mathbf{T}\backslash\mathbf{T}_1)^{m-j};\\ &\qquad\mathbf{T}_1\sim_{\ell-1}\mathbf{T}'_1,(\mathbf{T}\backslash\mathbf{T}_1)|_\ell\sim(\mathbf{T}'\backslash\mathbf{T}'_1)|_\ell\Big|D=D'=k\Big]\\ &= \mathbb{E}[Z_{\ell-1}(\mathbf{T})^j;\mathbf{T}\sim_{\ell-1}\mathbf{T}']\\ &\quad\times\mathbb{E}[Z_\ell(\mathbf{T})^{m-j};\mathbf{T}|_\ell\sim\mathbf{T}'|_\ell|D=D'=k-1]\,.\end{aligned}$$

Therefore, recalling that $k\mu_k = \lambda\mu_{k-1}$ in (II.6) and using the definition of $a_\ell(m)$ from (IV.5), we get from straightforward computations that

$$\begin{aligned} a_\ell(m) &\le \lambda^2 \sum_{0\le j\le m} \binom{m}{j} \mathbb{E}[Z_{\ell-1}(\mathbf{T})^j; \mathbf{T} \sim_{\ell-1} \mathbf{T}'] \\ &\qquad \times \mathbb{E}[Z_\ell(\mathbf{T})^{m-j}; \mathbf{T}|_\ell \sim \mathbf{T}'|_\ell] \\ &\le \lambda^2 g_\ell a_{\ell-1}(m) \\ &\quad + \lambda^2 \sum_{1\le j\le m-1} \binom{m}{j} a_{\ell-1}(j) \mathbb{E}[Z_\ell(\mathbf{T})^{m-j}; \mathbf{T}|_\ell \sim \mathbf{T}'|_\ell] \\ &\quad + \lambda^2 \mathfrak{p}_{\ell-1} \left(\mathbb{E}[Z_\ell(\mathbf{T})^m; \mathbf{T} \sim \mathbf{T}'] + a_\ell(m) \right), \end{aligned} \tag{IV.6}$$

where g_ℓ is defined as in (II.7), and in the case $j = 0$ we upper-bound $\mathbb{E}[Z_\ell(\mathbf{T})^m; \mathbf{T}|_\ell \sim \mathbf{T}'|_\ell]$ by $\mathbb{E}[Z_\ell(\mathbf{T})^m; \mathbf{T} \sim \mathbf{T}'] + a_\ell(m)$ since $\{\mathbf{T}|_\ell \sim \mathbf{T}'|_\ell\} \subset \{\mathbf{T} \sim \mathbf{T}'\} \cup \{\mathbf{T} \sim_\ell \mathbf{T}'\}$.

Note that our goal is to provide an upper bound of $a_\ell(m)/\mathfrak{p}_\ell$. To this end, note that

(1) $a_{\ell-1}(j) \lesssim_\lambda (\ell-1)^j \mathfrak{p}_{\ell-1}$ by the induction hypothesis for $1 \le j \le m-1$;

(2) $\mathbb{E}[Z_\ell(\mathbf{T})^{m-j}; \mathbf{T}|_\ell \sim \mathbf{T}'|_\ell] \lesssim_\lambda \ell^{m-j-1}$ for $1 \le j \le m-1$. In order to see this, note that $\{\mathbf{T}|_\ell \sim \mathbf{T}'|_\ell\} \subset \{\mathbf{T} \sim_\ell \mathbf{T}'\} \cup \{\mathbf{T} \sim \mathbf{T}'\}$. Thus, we can combine the bound that $\mathbb{E}[Z_\ell(\mathbf{T})^{m-j}; \mathbf{T} \sim_\ell \mathbf{T}'] \to 0$ as $\ell \to \infty$ by the induction hypothesis and Lemma 5, as well as the bound that $\mathbb{E}[Z_\ell(\mathbf{T})^{m-j}; \mathbf{T} \sim \mathbf{T}'] \le \mathbb{E}[Z_\ell(\mathbf{T})^{m-j}; \mathbf{T} \text{ finite}] \lesssim \ell^{m-j-1}$ by Lemmas 3 and 17 (i).

(3) $\mathbb{E}[Z_\ell(\mathbf{T})^m; \mathbf{T} \sim \mathbf{T}'] \le \mathbb{E}[Z_\ell(\mathbf{T})^m; \mathbf{T} \text{ finite}] \lesssim \ell^{m-1}$ by Lemmas 3 and 17 (i).

Combining (1), (2), (3) with the inequality (IV.6), we get that for some constant $C > 0$ depending on λ

$$\begin{aligned} a_\ell(m) &\le C\lambda^2 \sum_{1\le j\le m-1} \binom{m}{j} (\ell-1)^{m-1} \mathfrak{p}_{\ell-1} \\ &\qquad + C\lambda^2 g_\ell a_{\ell-1}(m) + \lambda^2 \mathfrak{p}_{\ell-1} \left(\ell^m + a_\ell(m) \right). \end{aligned}$$

Thus, there exists a constant $C_0 > 0$ depending on λ and m such that

$$\frac{a_\ell(m)}{\mathfrak{p}_\ell} \le \frac{1}{1 - \lambda^2 \mathfrak{p}_{\ell-1}} \frac{\lambda^2 g_\ell \mathfrak{p}_{\ell-1}}{\mathfrak{p}_\ell} \frac{a_{\ell-1}(m)}{\mathfrak{p}_{\ell-1}} + C_0 \ell^{m-1}. \tag{IV.7}$$

By Lemma 5, $\prod_{\ell=1}^\infty (1 - \lambda^2 \mathfrak{p}_\ell)^{-1} < \infty$. Note that $\frac{\lambda^2 g_\ell \mathfrak{p}_{\ell-1}}{\mathfrak{p}_\ell} \ge 1$ as in (II.13) and $g_\ell \le \gamma_\lambda + \mathfrak{p}_\ell$ as in (II.8). Combined with the definition that $\alpha_\lambda = \lambda^2 \gamma_\lambda$, it yields that

$$\begin{aligned} \prod_{\ell=1}^\infty \frac{\lambda^2 g_\ell \mathfrak{p}_{\ell-1}}{\mathfrak{p}_\ell} &\le \limsup_{L\to\infty} \mathfrak{p}_L^{-1} \prod_{\ell=1}^L (\alpha_\lambda + \lambda^2 \mathfrak{p}_{\ell-1}) \\ &\lesssim \limsup_{L\to\infty} \prod_{\ell=0}^L (1 + \frac{\lambda^2 \mathfrak{p}_\ell}{\alpha_\lambda}) < \infty. \end{aligned}$$

Combined with (IV.7), it yields that there is a constant $C'_m > 0$ depending on m and λ, such that $a_\ell(m)/\mathfrak{p}_\ell \le C'_m \ell^m$. This completes the proof for (IV.4).

We next show how to derive the consequence from (IV.4). If $\mathbf{T} \sim_\ell \mathbf{T}$, then $\mathbf{T} \sim_j \mathbf{T}'$ and there exist v, v' in the j-th levels of $\mathbf{T}$, $\mathbf{T}'$ respectively such that $\mathbf{T}_v \sim_{\ell-j} \mathbf{T}'_{v'}$. Therefore we have, for $j \le \ell$

$$\begin{aligned} &\mathbb{E}[Z_j(\mathbf{T})^m 1_{\{\mathbf{T}\sim_\ell \mathbf{T}'\}} \mid \mathbf{T}|_j, \mathbf{T}'|_j] \\ &\le Z_j(\mathbf{T})^m 1_{\{\mathbf{T}\sim_j \mathbf{T}'\}} \times Z_j(T)^2 \times \mathbb{P}(\mathbf{T} \sim_{\ell-j} \mathbf{T}), \end{aligned}$$

which implies that $\mathbb{E}([Z_j(\mathbf{T})]^m; \mathbf{T} \sim_\ell \mathbf{T}') \le C'_m \mathfrak{p}_j j^{m+2} \cdot \mathfrak{p}_{\ell-j} \le C_m \ell^{m+2} \mathfrak{p}_\ell$ for $j \le \ell$. Here $C_m = C'_m \sup_{\ell\ge 1, j\le \ell} \frac{\mathfrak{p}_j \mathfrak{p}_{\ell-j}}{\mathfrak{p}_\ell} < \infty$ by Lemma 5. Therefore,

$$\begin{aligned} &\mathbb{E}([Z_{\le\ell}(\mathbf{T})]^m; \mathbf{T} \sim_\ell \mathbf{T}') \\ &= \sum_{1\le j_1,\ldots,j_m\le \ell} \mathbb{E}[Z_{j_1}(\mathbf{T}) \cdots Z_{j_m}(\mathbf{T}); \mathbf{T} \sim_\ell \mathbf{T}'] \\ &\le \sum_{1\le j_1,\ldots,j_m\le \ell} \prod_{k=1}^m \mathbb{E}[Z_{j_k}^m(\mathbf{T}); \mathbf{T} \sim_\ell \mathbf{T}']^{1/m} \\ &\le \sum_{1\le j_1,\ldots,j_m\le \ell} C_m \ell^{m+2} \mathfrak{p}_\ell = C_m \ell^{2m+2} \mathfrak{p}_\ell, \end{aligned}$$

yielding that $\mathbb{E}([Z_{\le\ell}(\mathbf{T})]^m | \mathbf{T} \sim_\ell \mathbf{T}') \le C_m \ell^{2m+2}$. □

We next control the volume of PGW trees conditioned on isomorphism without constraints on heights of the trees. Our bound is likely far from being sharp, as suggested by [28, Theorem 2] for a bound of exponential decay when the offspring distribution has finite support.

Lemma 20: Let $\{\xi_i : 1 \le i \le m\}$ be i.i.d. Poisson (λ) random variables. Then for sufficiently large m,

$$\begin{aligned} &\mathbb{P}\left(\{\xi_k : 1 \le k \le m\} = \{\xi'_k : 1 \le k \le m\} \right) \\ &\le \exp\{-(\log m)^{1.7}\}. \end{aligned}$$

Proof: Let M and M' be two independent Poisson variables with mean m. In addition, let $N_k = \sum_{i=1}^M 1_{\{\xi_i = k\}}$ and $N'_k = \sum_{i=1}^{M'} 1_{\{\xi'_i = k\}}$ for all $k \ge 0$. By the Poisson thinning property, we see that N_k and N'_k for $k = 0, 1, \ldots$ are mutually independent Poisson variables with $\mathbb{E}N_k = \mathbb{E}N'_k = m\mu_k$ (recall that $\mu_k = \mathbb{P}(\xi_1 = k)$). A simple computation gives that $\mathbb{P}(M = M' = m) \ge c/m$ for a positive constant $c > 0$. Therefore,

$$\begin{aligned} &\mathbb{P}\left(\{\xi_k : 1 \le k \le m\} = \{\xi'_k : 1 \le k \le m\} \right) \\ &= \mathbb{P}\left(N_k = N'_k \text{ for } k \ge 0 \mid M = M' = m \right) \\ &\le O(m) \mathbb{P}\left(N_k = N'_k \text{ for } k \ge 0 \right) \\ &= O(m) \prod_{k=0}^\infty \mathbb{P}(N_k = N'_k) \le O(m) e^{-(\log m)^{1.8}} \end{aligned}$$

where the last inequality follows from a straightforward bound on $\mathbb{P}(N_k = N'_k) \le m^{-0.01}$ for $k \le \frac{\log m}{100 \log\log m}$ and $\mathbb{P}(N_k = N'_k) \le 1$ for $k > \frac{\log m}{100 \log\log m}$ (the power 1.8 is chosen as a arbitrary number less than 2). This completes the proof of the lemma. □

Lemma 21: For $\lambda \ge 1$, let $\mathbf{T}, \mathbf{T}'$ be two independent PGW(λ) trees. Then we have for sufficiently large m,

$$\mathbb{P}(\mathbf{T} \sim \mathbf{T}'; |\mathbf{T}| \ge m) \le \exp\{-(\log m)^{3/2}\}.$$

Proof: Indeed, for $\lambda > 1$ we have exponential decay in m, since $\{\mathbf{T} \sim \mathbf{T}'\}$ implies that $\mathbf{T}$ is a finite tree (Lemma 3). However, we prove the above results for all $\lambda \ge 1$.

Note that $\mathbf{T} \sim \mathbf{T}'$ implies that the degree sequence $(\xi_i)_{i=1}^{|\mathbf{T}|}$ and $(\xi_i')_{i=1}^{|\mathbf{T}|}$ have the same empirical distribution, where ξ_i is the number of the descendants in the tree $\mathbf{T}$ of the i-th vertex in the tree $\mathbf{T}$ (here we may use the breadth-first order, but the ordering is irrelevant anyway since we are only interested in the empirical distribution). Applying Lemma 20, we have,

$$\begin{aligned}&\mathbb{P}(\mathbf{T} \sim \mathbf{T}'; |\mathbf{T}| \geq m)\\&= \sum_{t \geq m} \mathbb{P}(\mathbf{T} \sim \mathbf{T}'; |\mathbf{T}| = t)\\&\leq \sum_{t \geq m} \mathbb{P}(\{\xi_k : 1 \leq k \leq t\} = \{\xi_k' : 1 \leq k \leq t\})\\&\leq \sum_{t \geq m} \exp\{-(\log t)^{1.7}\} \leq \exp\{-(\log m)^{3/2}\}\end{aligned}$$

for large m. This completes the proof of the lemma. □

In the end of this subsection, we prove some tail estimates for binomial variables. There is nothing novel, and we only record the proof for completeness.

Lemma 22: Let $X = \sum_{i=1}^m X_i$ where X_i's are independent and X_i is a Bernoulli variable with parameter p_i for $1 \leq i \leq m$. Then for all $x > 0$, we have

$$\mathbb{P}(X \geq x) \leq \left(\frac{e\mathbb{E}[X]}{x}\right)^x .$$

Proof: For any $\theta > 0$, a direct computation yields

$$\begin{aligned}&\mathbb{P}(X > x) \leq e^{-\theta x}\mathbb{E}[e^{\theta X}]\\&= \exp\left\{\sum_{i=1}^m \log(1 + p_i(e^\theta - 1)) - \theta x\right\}\\&\leq \exp\left\{\mathbb{E}X\left(e^\theta - 1\right) - \theta x\right\} .\end{aligned}$$

Setting $\theta = \log(1+x/\mathbb{E}[X])$ in the previous inequality, we get that

$$\mathbb{P}(X > x) \leq \left(\frac{e\mathbb{E}[X]}{x + \mathbb{E}[X]}\right)^x \leq \left(\frac{e}{x}\right)^x (\mathbb{E}X)^x .$$

□

□

Lemma 23: Assume $1 \leq f(n) = o(\sqrt{n})$. Let $X_1 \sim \mathrm{Bin}(f(n)^2, \frac{\lambda}{n})$. Let $X_2 = \sum_{j=1}^{rn} \xi_{1,j}\xi_{2,j}$, where $(\xi_{i,j})$ are i.i.d. $\mathrm{Bin}(f(n), \frac{\lambda}{n})$ variables and are independent of X_1. Then $\mathbb{P}(X_1 + X_2 \geq k) \lesssim_\lambda \left(\frac{rf(n)^2}{n}\right)^k$ for $k = 1, 2, 3$.

Proof: In the proof we write $\xi = \xi_{1,1}$ for short. Let $X_3 = \sum_{j=1}^{rn} 1_{\{\xi_{1,j} \geq 1\}} 1_{\{\xi_{2,j} \geq 1\}}$. Then $X_3 \sim \mathrm{Bin}(rn, p_n^2)$ where $p_n = \mathbb{P}(\xi \geq 1) = (1 + o(1))\frac{\lambda f(n)}{n}$. If $X_1 + X_2 \geq 3$ but $X_1 + X_3 \leq 2$, we have $X_3 = 1$ or $X_3 = 2$ (Note that if $X_3 = 0$ then $X_2 = 0$, and thus $X_1 \geq 3$, contradicting with $X_1 \leq 2$). Also, if in addition $X_3 = 2$, we then have $X_1 = 0$ and thus $X_2 \geq 3$; if in addition $X_3 = 1$, we then have either $X_1 = 1, X_2 \geq 2$ or $X_1 = 0, X_2 \geq 3$. Therefore,

$$\begin{aligned}\mathbb{P}(X_1 + X_2 \geq 3) &\leq \mathbb{P}(X_1 + X_3 \geq 3) + \mathbb{P}(X_2 \geq 3, X_3 = 2)\\&+ \mathbb{P}(X_2 \geq 3, X_3 = 1) + \mathbb{P}(X_2 \geq 2, X_3 = 1)\mathbb{P}(X_1 \geq 1) .\end{aligned} \tag{IV.8}$$

By Lemma 22, we have $\mathbb{P}(X_1 + X_3 \geq 3) \lesssim_\lambda (f(n)^2/n)^3$ and in addition $\mathbb{P}(\xi \geq k | \xi \geq 1) \lesssim_\lambda (f(n)/n)^{k-1}$ for $k \geq 1$. Then,

$$\begin{aligned}&\mathbb{P}(X_2 \geq 3 | X_3 = 1)\\&= \mathbb{P}(\xi_{1,1}\xi_{2,1} \geq 3 | \xi_{1,1} \geq 1, \xi_{2,1} \geq 1)\\&\leq 2\mathbb{P}(\xi \geq 3 | \xi \geq 1) + \mathbb{P}(\xi \geq 2 | \xi \geq 1)\mathbb{P}(\xi \geq 2 | \xi \geq 1)\\&\lesssim_\lambda f(n)^2/n^2 .\end{aligned}$$

Similarly we have $\mathbb{P}(X_2 \geq 2 | X_3 = 1) \lesssim \frac{f(n)}{n}$. Moreover, $\mathbb{P}(X_3 = 2) \lesssim_\lambda ((f(n))^2/n)^2$ and

$$\begin{aligned}&\mathbb{P}(X_2 \geq 3 | X_3 = 2)\\&= \mathbb{P}(\xi_{1,1}\xi_{2,1} + \xi_{1,2}\xi_{2,2} \geq 3 | \xi_{i,j} \geq 1; i, j \leq 2)\\&\leq 4\mathbb{P}(\xi \geq 2 | \xi \geq 1) \lesssim_\lambda f(n)/n .\end{aligned}$$

Plugging all these estimates into (IV.8) (together with straightforward bound on $\mathbb{P}(X_1 \geq 1)$ and $\mathbb{P}(X_3 = 1)$ as well as $\mathbb{P}(X_3 = 2)$), we have $\mathbb{P}(X_1 + X_2 \geq 3) \lesssim_\lambda \frac{r^3 f(n)^6}{n^3}$. The same argument shows $\mathbb{P}(X_1 + X_2 \geq 2) \lesssim_\lambda \frac{r^2 f(n)^4}{n^2}$ (and a much simpler argument proves the bound for $\mathbb{P}(X_1 + X_2 \geq 1)$). □

C. Proof of Lemma 16

This subsection is devoted to the proof of Lemma 16. To this end, we need the following four lemmas whose proofs are presented at the end of this subsection.

Lemma 24: For the Erdős-Rényi graph $\mathcal{G} = \mathcal{G}_{n,\frac{\lambda}{n}}$, there exist constants $\delta = \delta_{\lambda,\epsilon_0}$ and $s_\lambda = s_{\lambda,\epsilon_0}$ depending only on λ and ϵ_0 such that for all $u, v \in \mathcal{G}$,

$$\mathbb{P}(|\mathsf{N}_r(v)| > n^{\frac{1-\delta}{2}}) = o(n^{-2}) ; \tag{IV.9}$$

$$\mathbb{P}(\mathrm{Comp}(\mathsf{N}_r(v)) > s_\lambda) = o(n^{-2}) ; \tag{IV.10}$$

$$\mathbb{P}(\Xi_2(u, v) > s_\lambda) = o(n^{-2}) . \tag{IV.11}$$

The next lemma needs some notations. If we only keep the relative ordering for labels of the vertices in $\mathcal{T}_{\mathrm{aux}}(u)$, we get a rooted ordered tree. More precisely, let $\mathcal{U} = \cup_{n=0}^\infty \mathbb{N}^n$. We map $y \in \mathcal{T}_{\mathrm{aux}}(u)$ to a label $\mathbf{i} = i_1 i_2 \ldots i_m \in \mathcal{U}$, if the path from u to y is $(y_0 = u, y_1, \ldots, y_m = y)$ and y_k is the i_k-th smallest one among all the children of y_{k-1}. Then we get a rooted order tree which we denote by τ (we write $\mathcal{T}_{\mathrm{aux}}(u) = \tau$ for short).

Lemma 25: Let δ be chosen as in Lemma 24. For $u, v \in \mathcal{G}_{n,\frac{\lambda}{n}}$, we have

$$\mathbb{P}(|\mathcal{T}_{\mathrm{aux}}(u)|_r| > n^{\frac{1-\delta}{2}} \text{ or } |\mathcal{T}_{\mathrm{aux}}(v)|_r| > n^{\frac{1-\delta}{2}}) = o(n^{-2}) .$$

Moreover, there exists Δ_n depending only on δ with $\Delta_n \overset{n\to\infty}{\longrightarrow} 0$ such that for all rooted ordered trees τ, τ' with $|\tau|, |\tau| \leq n^{\frac{1-\delta}{2}}$, we have (denote by $\mathbf{T}, \mathbf{T}'$ two independent PGW(λ) trees)

$$\left|\frac{\mathbb{P}(\mathcal{T}_{\mathrm{aux}}(u)|_r = \tau, \mathcal{T}_{\mathrm{aux}}(v)|_r = \tau')}{\mathbb{P}(\mathbf{T}|_r = \tau)\mathbb{P}(\mathbf{T}|_r = \tau')} - 1\right| \leq \Delta_n .$$

Lemma 26: Let s_λ be chosen as in Lemma 24. For two vertices u and v, let $\Omega_a = \{\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v), \mathrm{Comp}(\mathsf{N}_r(u)) \leq s_\lambda, \Xi_2 \leq s_\lambda\}$. Then, there exists a constant $C_\lambda > 0$ depending on λ and s_λ such that for any event Ω_b which is measurable with respect to the σ-field generated by $\{\Xi_i, \Lambda_i$ for $i = 1, 2$, $|\mathcal{T}_{\mathrm{cut}}(u)|, |\mathcal{T}_{\mathrm{cut}}(v)|,$

$H(\mathcal{T}_{\text{cut}}(u)), H(\mathcal{T}_{\text{cut}}(v)), |\mathsf{N}_r(w)|$ for $w \in \mathcal{G}\}$, we have that

$$\mathbb{P}(\mathcal{T}_{\text{aux}}(u)|_r \sim \mathcal{T}_{\text{aux}}(v)|_r \big| \Omega_a \cap \Omega_b) \geq \frac{1}{C_\lambda}\,.$$

Lemma 27: Let δ be chosen as in Lemma 24. For any two vertices $u, v \in \mathcal{G}_{n,\frac{\lambda}{n}}$, we have that

$$\mathbb{P}(\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v) \text{ and } |\mathcal{T}_{\text{cut}}(u)|_r| = |\mathcal{T}_{\text{cut}}(v)|_r| \geq n^{\frac{\epsilon_0 \wedge \delta}{9}}) = o(n^{-2})\,.$$

Proof of Lemma 16: Let Ω_{typ} be the intersection of the following events: $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$, $\mathsf{N}_r(u)$ survives, $|\mathcal{T}_{\text{cut}}(u)|_r| \leq f(n)$ where $f(n) = n^{\frac{\epsilon_0 \wedge \delta}{9}}$, $\mathrm{Comp}(\mathsf{N}_r(u)) \leq s_\lambda$ and $|\mathsf{N}_r(w)| \leq n^{\frac{1-\delta}{2}}$ for all $w \in \mathcal{G}$. Then by Lemmas 24 and 27, we have

$$\mathbb{P}(\{\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v), \mathsf{N}_r(u) \text{ survives}\} \cap \Omega_{\text{typ}}^c) = o(n^{-2})\,. \tag{IV.12}$$

For an event A, we define $\mathbb{P}_{\text{typ}}(A) = \mathbb{P}(A \cap \Omega_{\text{typ}})$. Next, we prove Lemma 16 item by item.

(i). On the event $|A_{\leq r}(v)| = |A_{\leq r}(u)| \leq f(n)$, we have that $\Xi_1 = \sum_{x \in A_{\leq r}(u), y \in A_{\leq r}(v)} \mathcal{G}_{xy}$ is stochastically dominated by a binomial variable $X_1 \sim \mathrm{Bin}(f(n)^2, \frac{\lambda}{n})$. Similarly, since $\Xi_2 = \sum_{t=0}^{r-1} \sum_{w \in U_t} \mathcal{G}_{w,A_t(u)} \mathcal{G}_{w,A_t(v)}$, we have Ξ_2 is stochastically dominated by $X_2 = \sum_{j=1}^{rn} \xi_{1,j}\xi_{2,j}$, where $\xi_{i,j}$'s are i.i.d. binomial variables $\mathrm{Bin}(f(n), \frac{\lambda}{n})$. Observing that Ξ_1 and Ξ_2 are independent since they are measurable functions of different edges, Ξ is stochastically dominated by $X_1 + X_2$ with X_1 independent of X_2. Applying Lemma 23, we have

$$\mathbb{P}_{\text{typ}}(\Xi \geq 3) \leq \mathbb{P}(X_1 + X_2 \geq 3) \lesssim_\lambda \frac{r^3 f(n)^6}{n^3} = o(\frac{1}{n^2})\,.$$

Combined with (IV.12), this proves the first assertion in (i) via a simple union bound.

Furthermore, when $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$, if $w \in U_t$ and $\mathcal{G}_{w,A_t(u)}\mathcal{G}_{w,A_t(v)} \geq 1$, then by (IV.2) we have either $\mathcal{G}_{w,A_t(u)} = \mathcal{G}_{w,A_t(v)}$ or there exists $w' \neq w$ in U_t such that $\mathcal{G}_{w,A_t(u)}\mathcal{G}_{w,A_t(v)} = \mathcal{G}_{w',A_t(u)}\mathcal{G}_{w',A_t(v)}$. Thus when $\Xi_2 \leq 2$, we have $\mathcal{G}_{w,A_t(u)}\mathcal{G}_{w,A_t(v)} \in \{0,1\}$ for all $w \in U_t$. Hence $|R_{\leq r}| = \Xi_2$.

(ii). On the event Ω_{typ} and $\Xi = 0$, by definition we have $\Lambda_i = 0$ for $i = 2, 3$. The trees $\mathcal{T}_{\text{cut}}(v)$, $\mathcal{T}_{\text{cut}}(u)$ are exactly the auxiliary trees $\mathcal{T}_{\text{aux}}(v)$, $\mathcal{T}_{\text{aux}}(u)$, and $H(\mathcal{T}_{\text{aux}}(v)) \geq r$. Applying Lemma 26, we have

$$\mathbb{P}(\mathcal{T}_{\text{aux}}(v) \sim_r \mathcal{T}_{\text{aux}}(u) | \Omega_{\text{typ}}, \Xi = 0, \Lambda_1 \geq 1) \geq \frac{1}{C_\lambda}. \tag{IV.13}$$

On the other hand, note that $\mathrm{Comp}(\mathcal{G}[A_{\leq r}(u)])$ comes from edges within $A_{\leq r}(u)$ but not in $E(\mathcal{T}_{\text{cut}}(u)|_r)$. Conditioned on $\mathcal{T}_{\text{cut}}(u)|_r$ and $\mathcal{T}_{\text{cut}}(v)|_r$ with $|\mathcal{T}_{\text{cut}}(u)|_r| \leq f(n)$ and $|\mathcal{T}_{\text{cut}}(v)|_r| \leq f(n)$, we have that $\Lambda_1(u)$ and $\Lambda_1(v)$ are independent and are both stochastically dominated by a binomial variable $Y_1 \sim \mathrm{Bin}(f(n)^2, \frac{\lambda}{n})$. By the fact that $\Lambda_1(u) = \Lambda_1(v)$ on the event $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$ and by (IV.13), we have

$$\begin{aligned}
&\mathbb{P}_{\text{typ}}(\Xi = 0, \Lambda_1 \geq 1) \\
&\leq C_\lambda \mathbb{P}_{\text{typ}}(\Xi = 0, \mathcal{T}_{\text{aux}}(u) \sim_r \mathcal{T}_{\text{aux}}(v), \Lambda_1(u) \geq 1, \Lambda_1(v) \geq 1) \\
&\leq C_\lambda \mathbb{P}(\mathcal{T}_{\text{aux}}(u) \sim_r \mathcal{T}_{\text{aux}}(v)) \mathbb{P}(Y_1 \geq 1)^2 \\
&\lesssim_\lambda \alpha_\lambda^r \frac{f(n)^2}{n} = o(\frac{1}{n^2})\,,
\end{aligned} \tag{IV.14}$$

where the second inequality follows from independence and the aforementioned stochastic dominance, and the third inequality follows from Lemmas 25, 5 and 22.

When $\Xi = \Lambda = 0$, we have $\mathsf{N}_r(u) = \mathcal{T}_{\text{aux}}(u)|_r$. If u has two r-arms, then the tree $\mathcal{T}_{\text{aux}}(u)$ has two subtrees both of which survive $(r-1)$ levels. Hence by Lemmas 25 and 5,

$$\begin{aligned}
&\mathbb{P}_{\text{typ}}(\Lambda = \Xi = 0, u \text{ has two } r\text{-arms}) \\
&\leq \mathbb{P}(\mathcal{T}_{\text{aux}}(u) \sim_r \mathcal{T}_{\text{aux}}(v), \mathcal{T}_{\text{aux}}(u) \text{ has two subtrees} \\
&\qquad \text{surviving } (r-1) \text{ levels}) \\
&\lesssim_\lambda \alpha_\lambda^r \times \alpha_\lambda^r = o(\frac{1}{n^2})\,.
\end{aligned}$$

Combined with (IV.14), this proves (ii) via a simple union bound.

(iii). When $\Xi = \Xi_1 = 1$, by definition $\Lambda_2 = 0$. We next consider the case for $\Xi = \Xi_2 = 1$. Since $\Lambda_2(u)$, $\Lambda_2(v)$ and Ξ are measurable functions of different edges, conditioned on $\{\Xi = \Xi_2 = 1, |A_{\leq r}(u)| = |A_{\leq r}(v)| \leq f(n), |\mathsf{N}_r(w)| \leq n^{\frac{1-\delta}{2}}$ for $w \in R_{\leq r}\}$, we have that $\Lambda_2(u)$ and $\Lambda_2(v)$ are independent and are both stochastically dominated by a binomial variable $Y_2 \sim \mathrm{Bin}(f(n)n^{\frac{1-\delta}{2}}, \frac{\lambda}{n})$. Noting that $\Lambda_2(u) = \Lambda_2(v)$ when $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$, by Lemmas 23, 22, we have for $i = 1, 2$,

$$\begin{aligned}
&\mathbb{P}_{\text{typ}}(\Xi = 1, \Lambda_i \geq 1) = \mathbb{P}_{\text{typ}}(\Xi = 1, \Lambda_i(u) \geq 1, \Lambda_i(v) \geq 1) \\
&\leq \mathbb{P}(X_1 + X_2 \geq 1)\mathbb{P}(Y_i \geq 1)^2 = o(\frac{1}{n^2})\,.
\end{aligned}$$

Write $\rho' = \frac{1+\epsilon_0/2}{1+\epsilon_0}$. Note that $\mathcal{T}_{\text{cut}}(u)$ is a subtree of $\mathcal{T}_{\text{aux}}(u)$ by deleting at most 1 vertex when $\Xi = 1$. Therefore,

$$\begin{aligned}
&\mathbb{P}_{\text{typ}}(\Xi = 1, \Lambda = 0, H(\mathcal{T}_{\text{cut}}(v)) > \rho' r) \\
&\leq C_\lambda \mathbb{P}_{\text{typ}}(\Xi = 1, \mathcal{T}_{\text{aux}}(v) \sim_{\rho' r} \mathcal{T}_{\text{aux}}(u)) \\
&\leq C_\lambda \mathbb{P}(\mathcal{T}_{\text{aux}}(v) \sim_{\rho' r} \mathcal{T}_{\text{aux}}(u))\mathbb{P}(X_1 + X_2 \geq 1) \\
&\lesssim_\lambda \alpha_\lambda^{\rho' r} \frac{r f(n)^2}{n} = o(\frac{1}{n^2})\,.
\end{aligned}$$

Here the first inequality follows from Lemma 26; the second inequality follows from the independence and the stochastic dominance; the third inequality follows from Lemmas 25, 23 and 5. This proves (iii) via a simple union bound.

(iv). On Ω_{typ}, when $|R_{\leq r}| \leq 1$ we have $\Lambda_3 = 0$. When $\Xi = \Xi_2 = |R_{\leq r}| = 2$ (and we write $R_{\leq r} = \{w_1, w_2\}$), we have $\Lambda_3 \geq 1$ only if there exists $\ell \leq r$ such that $\mathsf{N}_k(w_1; \mathcal{G}[V \setminus A_{\leq r}]) \cap \mathsf{N}_k(w_2; \mathcal{G}[V \setminus A_{\leq r}]) = \emptyset$ holds for $k = \ell - 1$ but not $k = \ell$. Here $\mathcal{G}[A]$ is the subgraph on $\mathcal{G}$ induced by the vertex set A. Conditioned on $\Xi = \Xi_2 = 2$ and $|\mathsf{N}_k(w_i; \mathcal{G}[V \setminus A_{\leq r}])| \leq n^{\frac{1-\delta}{2}}$ for $i \in \{1, 2\}$, we have that Λ_3 is stochastically dominated by some $Y_3 \sim \mathrm{Bin}(n^{1-\delta}, \frac{\lambda}{n})$. Then by the aforementioned stochastic dominance and Lemmas 23, 22, we have that for $i = 1, 2, 3$,

$$\mathbb{P}_{\text{typ}}(\Xi = 2, \Lambda_i \geq 1) \leq \mathbb{P}(X_1 + X_2 \geq 2)\mathbb{P}(Y_i \geq 1) = o(\frac{1}{n^2})\,. \tag{IV.15}$$

In addition, by Lemma 26, we have that for every $\ell \leq r$,

$$\begin{aligned}
&\mathbb{P}(\mathcal{T}_{\text{aux}}(u) \sim_\ell \mathcal{T}_{\text{aux}}(v) | \Omega_{\text{typ}}, \Xi = 2, \Lambda = 0, H(\mathcal{T}_{\text{cut}}(v)) = \ell) \\
&\geq \frac{1}{C_\lambda}\,.
\end{aligned} \tag{IV.16}$$

Let

$$\Gamma_\ell = \sum_{t=0}^{\ell} \sum_{x\in A_t(u), y\in A_t(v)} \mathcal{G}_{xy} + \sum_{t=0}^{\ell-1} \sum_{w\in U_t} 1_{\{\mathcal{G}_{w,A_t(u)}\geq 1\}} 1_{\{\mathcal{G}_{w,A_t(v)\geq 1}\}}$$

By the fact that $\Gamma_\ell = 2$ (when $\Xi = 2$ and $H(\mathcal{T}_{\text{cut}}(v)) = \ell$) and the fact that $H(\mathcal{T}_{\text{aux}}(v)) \geq H(\mathcal{T}_{\text{cut}}(v))$, we get from (IV.16) that

$$\mathbb{P}_{\text{typ}}(\Xi = 2, \Lambda = 0, H(\mathcal{T}_{\text{cut}}(v)) > L) \leq C_\lambda \sum_{\ell=L}^{r} \mathbb{P}(\mathcal{T}_{\text{aux}}(v) \sim_\ell \mathcal{T}_{\text{aux}}(u); \Gamma_\ell = 2)\,, \tag{IV.17}$$

Given $\mathcal{T}_{\text{cut}}(u)$ and $\mathcal{T}_{\text{cut}}(v)$ as well as $\mathcal{T}_{\text{aux}}(u)$ and $\mathcal{T}_{\text{aux}}(v)$, we have that Γ_ℓ is stochastically dominated by the sum of $\text{Bin}(|\mathcal{T}_{\text{cut}}(u)|_\ell||\mathcal{T}_{\text{cut}}(u)|_\ell|, \frac{\lambda}{n})$ and $\text{Bin}(n, \frac{\lambda|\mathcal{T}_{\text{cut}}(u)|_\ell||\mathcal{T}_{\text{cut}}(u)|_\ell|}{n^2})$ where these two binomial variables are independent. Then by Lemmas 25 and 22,

$$\begin{aligned}
&\sum_{\ell=L}^{r} \mathbb{E}\left[\mathbb{P}(\Gamma_\ell = 2|\mathcal{T}_{\text{aux}}(u), \mathcal{T}_{\text{aux}}(v)); \mathcal{T}_{\text{aux}}(u) \sim_\ell \mathcal{T}_{\text{aux}}(v)\right] \\
&\lesssim_\lambda \sum_{\ell=L}^{r} \mathbb{E}\left[\frac{|Z_{\leq\ell}(\mathcal{T}_{\text{aux}}(u))|^4}{n^2}; \mathcal{T}_{\text{aux}}(u) \sim_\ell \mathcal{T}_{\text{aux}}(v);\right. \\
&\qquad\qquad \left. |\mathcal{T}_{\text{aux}}(u)|_r| \leq n^{\frac{1-\delta}{2}}\right] + o(n^{-2}) \\
&\lesssim_\lambda \frac{1}{n^2} \sum_{\ell=L}^{\infty} \mathbb{E}\left[|Z_{\leq\ell}(\mathbf{T})|^4; \mathbf{T} \sim_\ell \mathbf{T}'\right] + o(n^{-2})\,,
\end{aligned}$$

where $Z_{\leq\ell}$ (as before) denotes for the number of vertices in the first ℓ-levels of a tree and $\mathbf{T}, \mathbf{T}'$ are two independent PGW(λ)-trees. Thanks to Lemmas 19 and 5, we have that the sum on the right-hand side above vanishes in L. Combined with (IV.15) and (IV.17), it yields (iv) by a simple union bound. □

It remains to provide the postponed proofs for Lemmas 24, 25, 25 and 26.

Proof of Lemma 24: We first prove (i). When $\lambda < 1$, with high probability every component in $\mathcal{G}$ has size $O(\log n)$ and has at most one cycle (see, e.g., [8]). The desired quantitative bounds are also straightforward in this case and can be proved via standard methods (or by an easy adaption for the arguments below when $\lambda \geq 1$). Thus, in what follows we focus on $\lambda \geq 1$. We first control the volume of $\mathsf{N}_r(v)$ for $v \in \mathcal{G}$. Employing a standard breadth-first-search algorithm, we see that $|\mathsf{N}_r(v)|$ is stochastically dominated by the number of vertices in the first r-levels of a $\text{Bin}(n, \frac{\lambda}{n})$-GW branching process (denoted as $Z_{\leq r}$). By (the second inequality in) (III.1), there exists $\delta = \delta(\lambda; \epsilon_0)$ so that $\frac{(1-\delta)\log(\alpha_\lambda^{-1})}{2(1+\epsilon_0)} > \log(\lambda)$. As $n^{\frac{1-\delta}{2}} = \exp\{\frac{(1-\delta)\log(\alpha_\lambda^{-1})}{2(1+\epsilon_0)} r\}$, applying Lemma 17 (and Remark 18) we deduce (IV.9) as follows:

$$\mathbb{P}(|\mathsf{N}_r(v)| > n^{\frac{1-\delta}{2}}) \leq \mathbb{P}\big(Z_{\leq r} > n^{\frac{1-\delta}{2}}\big) = o(n^{-2})\,.$$

A cycle in $\mathsf{N}_r(v)$ is created by a "self-intersection", that is, an edge from v_t (the vertex we are exploring in the breadth-first-search process at time t) to some vertex we have explored before t. When $|\mathsf{N}_r(v)| \leq n^{\frac{1-\delta}{2}}$, it is easy to see that $\text{Comp}(\mathsf{N}_r(v))$ is stochastically dominated by $\text{Bin}\left(n^{1-\delta}, \frac{\lambda}{n}\right)$. By Lemma 22,

$$\begin{aligned}
&\mathbb{P}\left(\text{Comp}(\mathsf{N}_r(v)) \geq \frac{4}{\delta}\right) \\
&\leq \mathbb{P}\left(\text{Bin}(n^{1-\delta}, \frac{\lambda}{n}) \geq \frac{4}{\delta}\right) + o(\frac{1}{n^2}) = o(\frac{1}{n^2})\,.
\end{aligned}$$

Taking $s_\lambda > 1 + \frac{4}{\delta}$ and combining (IV.9), we obtain (IV.10).

We next prove (ii). When $|\mathsf{N}_r(u)|, |\mathsf{N}_r(v)| \leq n^{\frac{1-\delta}{2}}$, since $\Xi_2 = \sum_{t=0}^{r-1} \sum_{w\in U_t} \mathcal{G}_{w,A_t(u)} \mathcal{G}_{w,A_t(v)}$ we have Ξ_2 is stochastically dominated by $\sum_{j=1}^{rn} \xi_{j,1}\xi_{j,2}$ where $\xi_{j,i}$'s are i.i.d. binomial variables $\text{Bin}(n^{\frac{1-\delta}{2}}, \frac{\lambda}{n})$. Applying Lemma 22, there exists a constant S_1 depending only on $\delta = \delta(\lambda; \epsilon_0)$ such that

$$\mathbb{P}\left(\sum_{j=1}^{rn} 1_{\{\xi_{j,1}\geq 1\}} 1_{\{\xi_{j,2}\geq 1\}} \geq S_1\right) = o(\frac{1}{n^2})\,.$$

Take S_2 large enough such that $\mathbb{P}(\xi_{1,1} \geq S_2 \mid \xi_{1,1} \geq 1) = o(\frac{1}{n^2})$, then we have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{j=1}^{rn} \xi_{j,1}\xi_{j,2} \geq (S_2)^2 S_1 \middle| \sum_{j=1}^{rn} 1_{\{\xi_{j,1}\geq 1\}} 1_{\{\xi_{j,2}\geq 1\}} < S_1\right) \\
&\leq 2S_1 \mathbb{P}(\xi_{1,1} \geq S_2 \mid \xi_{1,1} \geq 1) = o(\frac{1}{n^2})\,.
\end{aligned}$$

Let $s_\lambda > S_2^2 S_1$. Then by (i), we deduce (IV.11), as required. □

Proof of Lemma 25: For $u, v \in \mathcal{G}$, note that on the event $\{|\mathsf{N}_r(v)| \leq \frac{1}{2} n^{\frac{1-\delta}{2}}, \Xi_2(u,v) \leq s_\lambda\}$, we have that $|\mathcal{T}_{\text{aux}}(v)|_r|$ is stochastically dominated by $\frac{1}{2} n^{\frac{1-\delta}{2}} + \sum_{i=1}^{s_\lambda} Z_{\leq r}^{(i)}$ where $Z_{\leq r}^{(i)}$ are independent and have the same distribution as $Z_{\leq r}$ (which, as in the proof of Lemma 24, is the number of vertices in the first r-levels of a $\text{Bin}(n, \lambda/n)$-GW tree). Thus by Lemmas 24, 17 (and Remark 18)

$$\begin{aligned}
&\mathbb{P}(|\mathcal{T}_{\text{aux}}(v)|_r| > n^{\frac{1-\delta}{2}}) \\
&\leq s_\lambda \mathbb{P}(Z_{\leq r} > n^{\frac{1}{2}-\delta}) + o(\frac{1}{n^2}) = o(\frac{1}{n^2})\,.
\end{aligned}$$

By symmetry, one can get the same bound for u and this proves the first assertion for Lemma 25.

We next prove the second assertion, which is similar to [31, Lemma 2.2]. We provide a complete proof here since our auxiliary tree is defined in a slightly non-standard manner. For two rooted trees τ, τ' with heights at most r and with $|\tau|, |\tau| \leq n^{\frac{1-\delta}{2}}$, let $(b_1, \ldots, b_{|\tau|_{r-1}|})$ and $(b'_1, \ldots, b'_{|\tau'|_{r-1}|})$ be the number of children for vertices obtained along with the breadth-first-search process for $\tau|_{r-1}$ and $\tau'|_{r-1}$ respectively (we do not care about vertices in the r-th level as they are surely leaves). For notation convenience, in this proof we write $t = |\tau|_{r-1}|$ and $t' = |\tau'|_{r-1}|$. Note that $\sum_{j=1}^{t} b_j = |\tau| - 1$ (and similarly for the prime version). When $\mathcal{T}_{\text{aux}}(u)|_r = \tau$, we define by σ the map such that $\sigma(\mathbf{i})$ is the label of the vertex on $\mathcal{T}_{\text{aux}}(u)|_r$ corresponding to $\mathbf{i}$ for $\mathbf{i} \in \tau$. Then we can regard $\mathcal{T}_{\text{aux}}(u)|_r$ as a labeled rooted ordered tree, and write $\mathcal{T}_{\text{aux}}(u)|_r = (\tau, \sigma)$. Therefore, we have

$$\begin{aligned}
&\mathbb{P}((\mathcal{T}_{\text{aux}}(u))|_r = \tau, (\mathcal{T}_{\text{aux}}(v))|_r = \tau') \\
&= \sum_{\sigma,\sigma'} \mathbb{P}((\mathcal{T}_{\text{aux}}(u))|_r = (\tau, \sigma), (\mathcal{T}_{\text{aux}}(v))|_r = (\tau', \sigma'))
\end{aligned}$$

where the sum is over all possible configurations for σ, σ'. Given (τ, σ) and (τ', σ'), when we are exploring the j-th vertex in the tree τ (or τ'), we know that it has b_j (or b'_j) children and does not connect to $n - f_{\sigma,\sigma'}(j)$ (or $n - f'_{\sigma,\sigma'}(j)$) vertices, where $|f_{\sigma,\sigma'}(j)| \le |\tau|$ and $|f'_{\sigma,\sigma'}(j)| \le |\tau'|$ for all j. Therefore,

$$\mathbb{P}((\mathcal{T}_{\mathrm{aux}}(u))|_r = (\tau,\sigma), (\mathcal{T}_{\mathrm{aux}}(v))|_r = (\tau',\sigma'))$$
$$= \prod_{j=1}^{t} \big(\frac{\lambda}{n}\big)^{b_j}\big(1-\frac{\lambda}{n}\big)^{n-f_{\sigma,\sigma'}(j)} \prod_{j=1}^{t'} \big(\frac{\lambda}{n}\big)^{b'_j}\big(1-\frac{\lambda}{n}\big)^{n-f'_{\sigma,\sigma'}(j)}$$
$$\in \big(\frac{\lambda}{n}\big)^{|\tau|+|\tau'|-2}\Big[\big(1-\frac{\lambda}{n}\big)^{n(t+t')}, \big(1-\frac{\lambda}{n}\big)^{(n-|\tau|-|\tau'|)(t+t')}\Big],$$

where as we will see the upper and lower bounds above are very close to each other. In addition, the total number of all the configurations for (σ,σ') is upper-bounded by

$$\binom{n-2}{b_1,\ldots,b_t,n-|\tau|-1}\binom{n-2}{b'_1,\ldots,b'_{t'},n-|\tau'|-1} \le \frac{n^{|\tau|+|\tau'|-2}}{\prod_j b_j! \prod_j b'_j!}$$

and lower-bounded by

$$\binom{n-2}{b_1,\ldots,b_t,b'_1,\ldots,b'_{t'},n-|\tau|-|\tau'|} = \frac{(n-2)\cdots[n-(|\tau|+|\tau'|-1)]}{\prod_j b_j! \prod_j b'_j!}.$$

Furthermore, we can compute the analogous probability regarding to two independent PGW(λ)-trees as follows:

$$\mathbb{P}(\mathbf{T}|_r = \tau)\mathbb{P}(\mathbf{T}|_r = \tau')$$
$$= \lambda^{|\tau|+|\tau'|-2} e^{-\lambda(t+t')} \prod_{j=1}^{t}\frac{1}{b_j!}\prod_{j=1}^{t'}\frac{1}{b'_j!}.$$

Therefore, noting that $|\tau|, |\tau'| \le n^{\frac{1-\delta}{2}}$ and using a straightforward computation, we see that the ratio $\frac{\mathbb{P}((\mathcal{T}_{\mathrm{aux}}(u))|_r=\tau,(\mathcal{T}_{\mathrm{aux}}(v))|_r=\tau')}{\mathbb{P}(\mathbf{T}|_r=\tau)\mathbb{P}(\mathbf{T}|_r=\tau')}$ converges to 1 as $n \to \infty$, proving the second assertion as required. □

Next we prove Lemma 26. Assume that $\mathsf{N}_{r+1}(u) \sim \mathsf{N}_{r+1}(v)$ and ϕ is an isomorphism. Note that (IV.2) implies that $\mathcal{G}[A_{\le r}(u)]$ and $\mathcal{G}[A_{\le r}(v)]$ are isomorphic to each other. However, when there exist cycles in $\mathcal{G}[A_{\le r}(u)]$ and $\mathcal{G}[A_{\le r}(v)]$, our trees $\mathcal{T}_{\mathrm{cut}}(u)|_r$ and $\mathcal{T}_{\mathrm{cut}}(u)|_r$ (which are spanning trees of $\mathcal{G}[A_{\le r}(u)]$ and $\mathcal{G}[A_{\le r}(v)]$ respectively) may not be isomorphic since there are some edges deleted. In this case, the event $\mathcal{T}_{\mathrm{cut}}(u)|_r \sim \mathcal{T}_{\mathrm{cut}}(v)|_r$ occurs or not depends on the labeling configuration on $\mathcal{G}$.

Proof of Lemma 26: For fixed $u, v \in V$, we say two graphs G_1 and G_2 (on V) are equivalent if there exists an isomorphism φ from G_1 to G_2 such that $\varphi(u) = u$ and $\varphi(v) = v$. We write $[G]$ to denote an equivalence class for this equivalence relation. We can sample the graph $\mathcal{G}$ as follows: first we sample an equivalence class and then we sample the labels uniformly from all labelings that yield the sampled equivalence class. Now we have

$$\mathbb{P}(\mathcal{T}_{\mathrm{aux}}(u)|_r \sim \mathcal{T}_{\mathrm{aux}}(v)|_r | \Omega_a \cap \Omega_b)$$
$$= \sum_{[G]\in\Omega_a\cap\Omega_b} \mathbb{P}(\mathcal{T}_{\mathrm{aux}}(u)|_r \sim \mathcal{T}_{\mathrm{aux}}(v)|_r | \mathcal{G} \in [G])$$
$$\times \mathbb{P}(\mathcal{G} \in [G] | \Omega_a \cap \Omega_b).$$

Let ϕ be an isomorphism between $\mathsf{N}_{r+1}(u)$ and $\mathsf{N}_{r+1}(v)$. Recall our reduced BFS. We say a label configuration is successful, if for each vertex y in $A_{t+1}(u)$ with neighbors $z_1(y), \ldots, z_{\ell(y)}(y)$ in $A_t(u)$ such that $z_1(y), \ldots, z_{\ell(y)}(y)$ is increasing in the (prefixed) ordering on V, then $\phi(z_1(y)), \ldots, \phi(z_{\ell(y)}(y))$ is also increasing in the ordering on V. Clearly, with a successful labeling configuration we have $\mathcal{T}_{\mathrm{cut}}(u)|_r \sim \mathcal{T}_{\mathrm{cut}}(v)|_r$. When $\mathrm{Comp}(\mathcal{G}[A_{\le r}(v)]) \le s_\lambda$, we claim that

$$\#\text{successful configurations} \ge \frac{\#\text{all configurations}}{((s_\lambda+1)!)^{s_\lambda}}. \tag{IV.18}$$

In order to see this, we observe that $|Y| \le s_\lambda$ where $Y = \{y \in A_{\le r}(u) : \ell(y) \ge 2\}$ are the only vertices we need to investigate in order for the labeling configuration to be successful. In addition, once the labels are fixed elsewhere except at $\phi(z_1(y)), \ldots, \phi(z_{\ell(y)}(y))$ for $y \in Y$, the number of valid completions for the labeling configuration is at most $\prod_{y\in Y} \ell_y$ and out of which at least one of them is successful. Since $\ell(y) \le \mathrm{Comp}(\mathcal{G}[A_{\le r}(v)]) + 1 \le s_\lambda + 1$, this implies (IV.18).

At this point, we note that the event $\mathcal{T}_{\mathrm{aux}}(u)|_r \sim \mathcal{T}_{\mathrm{aux}}(v)|_r$ occurs if the following hold:

- the labeling configuration is successful;
- for each of the corresponding pairs of the independent $\mathrm{Bin}(n, \frac{\lambda}{n})$-GW trees we grafted they are isomorphic.

When $\Xi_2 \le s_\lambda$, we have grafted at most s_λ pairs of trees. Combined with (IV.18), it gives that for all $[G] \in \Omega_a \cap \Omega_b$

$$\mathbb{P}(\mathcal{T}_{\mathrm{aux}}(u)|_r \sim \mathcal{T}_{\mathrm{aux}}(v)|_r | \mathcal{G} \in [G]) \ge \frac{\gamma_\lambda^{s_\lambda}}{((s_\lambda+1)!)^{s_\lambda}}.$$

□

Proof of Lemma 27: In the case $\lambda < 1$, the statement follows from a simple tail estimate on subcritical branching process. Thus, in what follows we assume that $\lambda \ge 1$. Let $f(n) = n^{\frac{\epsilon_0\wedge\delta}{9}}$. Applying Lemma 26, we have

$$\mathbb{P}(\mathsf{N}_{r+1}(v) \sim \mathsf{N}_{r+1}(u), \mathrm{Comp}(\mathsf{N}_r(v)) \le s_\lambda, \tag{IV.19}$$
$$\Xi_2 \le s_\lambda, |\mathcal{T}_{\mathrm{cut}}(u)|_r| > f(n))$$
$$\le C_\lambda \mathbb{P}(\mathcal{T}_{\mathrm{aux}}(v)|_r \sim \mathcal{T}_{\mathrm{aux}}(u)|_r; |\mathcal{T}_{\mathrm{aux}}(v)|_r| > f(n)). \tag{IV.20}$$

Since $\mathcal{T}_{\mathrm{aux}}(u)$ and $\mathcal{T}_{\mathrm{aux}}(v)$ behaves like independent PGW(λ) trees (Lemma 25), we have

$$\mathbb{P}(\mathcal{T}_{\mathrm{aux}}(v)|_r \sim \mathcal{T}_{\mathrm{aux}}(u)|_r, f(n) < |\mathcal{T}_{\mathrm{aux}}(v)|_r| \le n^{\frac{1-\delta}{2}})$$
$$= \sum_{\tau\sim\tau', f(n)<|\tau|\le n^{\frac{1-\delta}{2}}} \mathbb{P}(\mathcal{T}_{\mathrm{aux}}(v)|_r = \tau, \mathcal{T}_{\mathrm{aux}}(u)|_r = \tau')$$
$$\lesssim \sum_{\tau\sim\tau', |\tau|>f(n)} \mathbb{P}(\mathbf{T}|_r = \tau, \mathbf{T}'|_r = \tau')$$
$$= \mathbb{P}(\mathbf{T}|_r \sim \mathbf{T}'|_r; |\mathbf{T}|_r| > f(n))$$
$$\le \mathbb{P}(\mathbf{T} \sim_r \mathbf{T}', |\mathbf{T}|_r| > f(n)) + \mathbb{P}(\mathbf{T} \sim \mathbf{T}', |\mathbf{T}| > f(n)),$$

where the last inequality follows from $\{\mathbf{T}|_r \sim \mathbf{T}'|_r\} \subset \{\mathbf{T} \sim_r \mathbf{T}'\} \cup \{\mathbf{T} \sim \mathbf{T}'\}$. Applying Lemma 19 with $m = \frac{10}{\epsilon_0\wedge\delta}$ and applying Markov's inequality, we have

$$\mathbb{P}(\mathbf{T} \sim_r \mathbf{T}', |\mathbf{T}|_r| > f(n)) \leq \mathbb{E}\left[\frac{|Z_{\leq r}(\mathbf{T})|^m}{n^{m(\epsilon_0 \wedge \delta)/9}}; \mathbf{T} \sim_r \mathbf{T}'\right]$$
$$\lesssim \frac{r^{2m}}{n^{10/9}} \alpha_\lambda^r = o(n^{-2}).$$

Applying Lemma 21, we have $\mathbb{P}(\mathbf{T} \sim \mathbf{T}', |\mathbf{T}| > f(n)) = o(n^{-2})$. Altogether, this implies that

$$\mathbb{P}(\mathcal{T}_{\text{aux}}(v)|_r \sim \mathcal{T}_{\text{aux}}(u)|_r, f(n) < |\mathcal{T}_{\text{aux}}(v)|_r| \leq n^{\frac{1-\delta}{2}})$$
$$= o(n^{-2}).$$

Combined with (IV.19) and Lemma 25, this implies that

$$\mathbb{P}(\mathsf{N}_{r+1}(v) \sim \mathsf{N}_{r+1}(u), \mathrm{Comp}(\mathsf{N}_r(v)) \leq s_\lambda,$$
$$\Xi_2 \leq s_\lambda, |\mathcal{T}_{\text{cut}}(u)|_r| > f(n)) = o(n^{-2}).$$

Combined with (IV.10) and (IV.11), this competes the proof of the lemma. □

Acknowledgment

The authors warmly thank Hang Du, Haojie Hou, Haoyu Liu, Yanxia Ren, and Fan Yang for helpful discussions. They also warmly thank anonymous reviewers for their detailed and helpful comments.

References

[1] K. Adhikari and S. Chakraborty, "Shotgun assembly of random geometric graphs," 2022, *arXiv:2202.02968*.
[2] A. Aho, J. Hoperoft, and J. Ullman, *The Design and Analysis of Computer Algorithms*. Reading, MA, USA: Addison-Wesley, 1975.
[3] R. Arratia, D. Martin, G. Reinert, and M. Waterman, "Poisson process approximation for repeats in one sequence and its application to sequencing by hybridization," in *Combinatorial Pattern Matching* (Lecture Notes in Computer Science), vol. 1075. Berlin, Germany: Springer, 1996, pp. 209–219.
[4] L. Babai, "Graph isomorphism in quasipolynomial time [extended abstract]," in *Proc. 48th Annu. ACM Symp. Theory Comput.*, New York, NY, USA, Jun. 2016, pp. 684–697.
[5] P. Balister, B. Bollobás, and B. Narayanan, "Reconstructing random jigsaws," in *Multiplex and Multilevel Networks*. Oxford, U.K.: Oxford Univ. Press, 2019, pp. 31–50.
[6] H. L. Bodlaender, "Polynomial algorithms for graph isomorphism and chromatic index on partial k-trees," *J. Algorithms*, vol. 11, no. 4, pp. 631–643, Dec. 1990.
[7] B. Bollobás, "Distinguishing vertices of random graphs," in *Graph Theory* (North-Holland Mathematics Studies), vol. 62. Amsterdam, The Netherlands: North Holland, 1982, pp. 33–49.
[8] B. Bollobás, *Random Graphs* (Cambridge Studies in Advanced Mathematics), vol. 73, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2001.
[9] C. Bordenave, U. Feige, and E. Mossel, "Shotgun assembly of random jigsaw puzzles," *Random Struct. Algorithms*, vol. 56, no. 4, pp. 998–1015, Jul. 2020.
[10] J. Ding and H. Liu, "Shotgun assembly threshold for lattice labeling model," *Probab. Theory Rel. Fields*, vol. 187, pp. 1–20, Apr. 2023.
[11] M. Dyer, A. Frieze, and S. Suen, "The probability of unique solutions of sequencing by hybridization," *J. Comput. Biol.*, vol. 1, no. 2, pp. 105–110, Jan. 1994.
[12] M. Elberfeld and P. Schweitzer, "Canonizing graphs of bounded tree width in logspace," *ACM Trans. Comput. Theory*, vol. 9, no. 3, pp. 1–29, Sep. 2017.
[13] J. Gaudio and E. Mossel, "Shotgun assembly of Erdős–Rényi random graphs," *Electron. Commun. Probab.*, vol. 27, pp. 1–14, Jan. 2022.
[14] M. Grohe, D. Neuen, P. Schweitzer, and D. Wiebking, "An improved isomorphism test for bounded-tree-width graphs," *ACM Trans. Algorithms*, vol. 16, no. 3, pp. 1–31, Jul. 2020.
[15] T. E. Harris, "Branching processes," *Ann. Math. Statist.*, vol. 19, no. 4, pp. 474–494, 1948.
[16] T. E. Harris, *The theory of Branching Processes*. Berlin, Germany: Springer-Verlag, 1963.
[17] H. Huang and K. Tikhomirov, "Shotgun assembly of unlabeled Erdős–Rényi graphs," 2021, *arXiv:2108.09636*.
[18] V. Huroyan, G. Lerman, and H.-T. Wu, "Solving jigsaw puzzles by the graph connection Laplacian," *SIAM J. Imag. Sci.*, vol. 13, no. 4, pp. 1717–1753, Jan. 2020.
[19] D. Soudry, S. Keshri, P. Stinson, M.-H. Oh, G. Iyengar, and L. Paninski, "A shotgun sampling solution for the common input problem in neural connectivity inference," 2013, *arXiv:1309.3724*.
[20] H. Kesten and B. P. Stigum, "A limit theorem for multidimensional Galton–Watson processes," *Ann. Math. Statist.*, vol. 37, no. 5, pp. 1211–1223, Oct. 1966.
[21] D. Lokshtanov, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, "Fixed-parameter tractable canonization and isomorphism test for graphs of bounded treewidth," *SIAM J. Comput.*, vol. 46, no. 1, pp. 161–189, Jan. 2017.
[22] E. M. Luks, "Isomorphism of graphs of bounded valence can be tested in polynomial time," *J. Comput. Syst. Sci.*, vol. 25, no. 1, pp. 42–65, Aug. 1982.
[23] R. Lyons, R. Pemantle, and Y. Peres, "Conceptual proofs of $L\log L$ criteria for mean behavior of branching processes," *Ann. Probab.*, vol. 23, no. 3, pp. 1125–1138, 1995.
[24] A. Martinsson, "A linear threshold for uniqueness of solutions to random jigsaw puzzles," *Combinatorics, Probab. Comput.*, vol. 28, no. 2, pp. 287–302, Mar. 2019.
[25] E. Mossel and N. Ross, "Shotgun assembly of labeled graphs," *IEEE Trans. Netw. Sci. Eng.*, vol. 6, no. 2, pp. 145–157, Apr. 2019.
[26] E. Mossel and N. Sun, "Shotgun assembly of random regular graphs," 2015, *arXiv:1512.08473*.
[27] A. S. Motahari, G. Bresler, and D. N. C. Tse, "Information theory of DNA shotgun sequencing," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6273–6289, Oct. 2013.
[28] C. Olsson, "The probability of random trees being isomorphic," 2022, *arXiv:2205.06625*.
[29] M. Przykucki, A. Roberts, and A. Scott, "Shotgun reconstruction in the hypercube," *Random Struct. Algorithms*, vol. 60, no. 1, pp. 117–150, Jan. 2022.
[30] J. Raymond, R. Bland, and K. McGoff, "Shotgun identification on groups," *Involve, J. Math.*, vol. 14, no. 4, pp. 631–682, Oct. 2021.
[31] O. Riordan and N. Wormald, "The diameter of sparse random graphs," *Combin. Probab. Comput.*, vol. 19, nos. 5–6, pp. 835–926, 2010.
[32] B. P. Stigum, "A theorem on the Galton–Watson process," *Ann. Math. Statist.*, vol. 37, no. 3, pp. 695–698, Jun. 1966.

Jian Ding is currently a Chair Professor with the School of Mathematical Sciences, Peking University. His main research interests include probability theory, with an emphasis on interactions with statistical physics, theoretical computer science, and theoretical statistics. His work has been recognized by a number of international awards, including the Sloan Fellowship in 2015, the Rollo Davidson Prize in 2017 (shared with Nike Sun), an ICM Invited Lecturer in 2022 (joint with Ewain Gwynne and Julien Dubédat), the ICCM Gold Medal in 2022, an IMS Medallion Lecturer in 2024, and a Plenary Lecturer for the International Congress of Mathematical Physics in 2024. He also serves on the editorial boards for *Annals of Applied Probability*, *Annals of Probability*, *Communications in Mathematical Physics*, *Frontiers of Mathematics*, *Journal of the American Mathematical Society*, *Science China Mathematics*, and *Transactions of the London Mathematical Society*.

Yiyang Jiang is currently pursuing the Ph.D. degree with the School of Mathematical Sciences, Peking University. Her research interests include probability theory, in particular branching processes and related models.

Heng Ma is currently pursuing the Ph.D. degree with the School of Mathematical Sciences, Peking University. His research interests include probability theory and its applications, with a particular focus on probability models that incorporate hierarchical structures.