Electronic Companion to "Aligning Multiple Inhomogeneous Random Graphs: Fundamental Limits of Exact Recovery"

Taha Ameen and Bruce Hajek

A Maximum Likelihood Estimator for Exactly Matching Multiple Homogeneous Graphs

We show that in the homogeneous setting, the maximum likelihood estimator is given by the permutation profile that minimizes the number of edges in the corresponding union graph, i.e.

$$(\widehat{\pi}_{12}^{\mathrm{MLE}}, \cdots, \widehat{\pi}_{1m}^{\mathrm{MLE}}) \in \underset{\pi_{12},\cdots,\pi_{1m}}{\arg\min} \, |E(G_1 \vee G_2^{\pi_{12}} \vee \cdots \vee G_m^{\pi_{1m}})|.$$

Lemma A.1. *Let $G_1, \cdots, G_m$ be correlated Erdős-Rényi graphs obtained from the subsampling model with parameters p and $s \in (0,1)$. Further, let $\pi_{12}, \cdots, \pi_{1m}$ denote a collection of permutations on $[n]$. Then*

$$\log \mathbb{P}\left(G_1, \cdots, G_m \mid \pi_{12}^* = \pi_{12}, \cdots, \pi_{1m}^* = \pi_{1m}\right) \propto \text{const.} - |E(G_1 \vee G_2^{\pi_{12}} \vee \cdots \vee G_m^{\pi_{1m}})|,$$

where const. *depends only on p, s and $G_1, \cdots, G_m$.*

Proof. Consider the setting where $p_{ij} = p$ for all $i \neq j$. Notice that

$$\begin{aligned}
\mathbb{P}\left(G_1, \cdots, G_m | \pi_{12}^*, \cdots, \pi_{1m}^*\right) &= \prod_{e \in \binom{[n]}{2}} \mathbb{P}\big(G_1(e), G_2(\pi_{12}^*(e)) \cdots, G_m(\pi_{1m}^*(e)) \mid \pi_{12}^*, \cdots, \pi_{1m}^*\big) \\
&= \prod_{e \in \binom{[n]}{2}} \mathbb{P}\left(G_1(e), G_2'(e) \cdots, G_m'(e)\right) \qquad (1)
\end{aligned}$$

where for a node pair $e = \{u, v\}$, the shorthand $\pi(e)$ denotes $\{\pi(u), \pi(v)\}$. The edge status of any node pair e in the graph tuple $(G_1, G_2', \cdots, G_m')$ can be any of the 2^m bit strings of length m, but the corresponding probability in (1) depends only on the number of ones and zeros in the bit string. For $i \in [m]$, let α_i denote the number of node pairs e whose corresponding tuple $(G_1(e), G_2'(e), \cdots, G_m'(e))$ has exactly i 1's:

$$\alpha_i \triangleq \sum_{e \in \binom{[n]}{2}} \mathbb{1}\big\{(G_1(e), G_2'(e), \cdots, G_m'(e)) \text{ has exactly } i \text{ 1's}\big\}.$$

Two key observations are in order. First, it follows by definition that $\alpha_0 + \alpha_1 + \cdots + \alpha_m = \binom{n}{2}$. Second, by definition of α_i, it follows that

$$\sum_{i=0}^{m} i\alpha_i = \sum_{e \in \binom{[n]}{2}} \sum_{j=1}^{m} G_j(e) = \sum_{e \in \binom{[n]}{2}} \sum_{j=1}^{m} G_j'(e) \qquad (2)$$

is constant, independent of $\pi^*_{12}, \cdots, \pi^*_{1m}$. It follows then that

$$\begin{aligned}
(1) &= (1-p+p(1-s)^m)^{\alpha_0} \times \prod_{i=1}^{m} \left(ps^i(1-s)^{m-i}\right)^{\alpha_i} \\
&= (1-p+p(1-s)^m)^{\alpha_0} \times p^{\sum_{i=1}^m \alpha_i} \times \prod_{i=1}^{m} \left(s^i(1-s)^{m-i}\right)^{\alpha_i} \\
&= (1-p+p(1-s)^m)^{\alpha_0} \times p^{\binom{n}{2}-\alpha_0} \times \left(\frac{s}{1-s}\right)^{\sum_{i=1}^m i\alpha_i} \times (1-s)^{m\sum_{i=1}^m \alpha_i} \\
&= \left(\frac{1-p+p(1-s)^m}{p(1-s)^m}\right)^{\alpha_0} \times (p(1-s)^m)^{\binom{n}{2}} \times \left(\frac{s}{1-s}\right)^{\sum_{i=1}^m i\alpha_i} \\
&\propto \left(1 + \frac{1-p}{p(1-s)^m}\right)^{\alpha_0},
\end{aligned}$$

where the last step uses (2). Finally, since $\frac{1-p}{p(1-s)^m} > 0$, it follows that the log-likelihood satisfies

$$\log\left(\mathbb{P}\left(G_1, \cdots, G_m \mid \pi^*_{12}, \cdots, \pi^*_{1m}\right)\right) \propto \text{const.} + \alpha_0,$$

i.e. maximizing the likelihood corresponds to selecting $\pi_{12}, \cdots, \pi_{1m}$ to maximize α_0 - the number of node pairs e for which $G_1(e) = G_2(\pi_{12}(e)) = \cdots = G_m(\pi_{1m}(e)) = 0$. This is equivalent to minimizing the number of edges in the union graph $G_1 \vee G_2^{\pi_{12}} \vee \cdots \vee G_m^{\pi_{1m}}$, as desired. $\square$

Remark A.1. In the case of two graphs, minimizing the number of edges in the union graph $G_1 \vee_\pi G_2$ is equivalent to maximizing the number of edges in the intersection graph $G_1 \wedge_\pi G_2$. This is consistent with existing literature on two graphs (Cullina and Kiyavash 2016, Wu et al. 2022).

B On negative results for exact recovery with two graphs

We build up to the proof of Lemma 1 by presenting first a useful lemma about inhomogeneous random graphs.

Lemma B.1. *Let $G \sim \mathrm{RG}(n, \mathbf{P})$ be an inhomogeneous random graph, and let $N \triangleq \binom{K}{2}$, where K is the number of isolated nodes in G. Then,*

$$\mathbb{P}\left(N \geq 1\right) \geq \exp(-(4n-2)p_{\max}^2 - 6p_{\max}) \times \frac{\mu^2 - 2\mu \max_{1\leq i\leq n} e^{-d_i}}{4(\mu^2+\mu+1)}, \tag{3}$$

where $\mu \triangleq \sum_{i=1}^n e^{-d_i}$.

Proof. Let Z_i denote the indicator random variable that node i is isolated in G. Further, let $\widetilde{Z}_i \sim \mathrm{Bern}(e^{-d_i})$ denote a collection of independent random variables and define $\widetilde{N} \triangleq \sum_{1\leq i<j\leq n} \widetilde{Z}_i \cdot \widetilde{Z}_j$. We use the second moment method by relating the moments of N and $\widetilde{N}$.

Bounding the second moment. Notice that

$$\mathbb{E}[N^2] = \sum_{1\leq i<j\leq n} \sum_{1\leq k<\ell\leq n} \mathbb{E}[Z_i Z_j Z_k Z_\ell]\,, \tag{4}$$

$$\mathbb{E}[\widetilde{N}^2] = \sum_{1\leq i<j\leq n} \sum_{1\leq k<\ell\leq n} \mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell]. \tag{5}$$

Consider the following three exhaustive cases.

– *Case* 1*:* $i = k$ *and* $j = \ell$*.* Notice that

$$\mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell] = \mathbb{E}[\widetilde{Z}_i^2] \cdot \mathbb{E}[\widetilde{Z}_j^2] = e^{-d_i - d_j}.$$

Therefore,

$$\begin{aligned}\mathbb{E}[Z_i Z_j Z_k Z_\ell] &= (1 - p_{ij}) \prod_{\substack{(i,u)\\ u \neq j}} (1 - p_{iu}) \prod_{\substack{(j,v)\\ v \neq i}} (1 - p_{jv}) \\ &\overset{\text{(a)}}{\leq} e^{-d_i - d_j + p_{ij}} \overset{\text{(b)}}{\leq} e^{6p_{\max}} \cdot \mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell], \end{aligned} \tag{6}$$

where (a) uses that $1 - x \leq e^{-x}$ for all x, and (b) uses that $p_{ij} \leq p_{\max} \leq 6p_{\max}$.

– *Case* 2*:* $j = k$*.* Notice that

$$\mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell] = \mathbb{E}[\widetilde{Z}_i] \cdot \mathbb{E}[\widetilde{Z}_j^2] \cdot \mathbb{E}[\widetilde{Z}_\ell] = e^{-d_i - d_j - d_\ell}.$$

Therefore, similarly to case 1, we have

$$\begin{aligned}\mathbb{E}[Z_i Z_j Z_k Z_\ell] &= \prod_{f \in \binom{\{i,j,\ell\}}{2}} (1 - p_f) \times \prod_{\substack{(i,u)\\ u \neq j,\ell}} (1 - p_{iu}) \prod_{\substack{(j,v)\\ v \neq i,\ell}} (1 - p_{jv}) \prod_{\substack{(\ell,w)\\ w \neq i,j}} (1 - p_{\ell w}) \\ &\leq e^{-d_i - d_j - d_\ell + p_{ij} + p_{i\ell} + p_{j\ell}} \\ &\leq e^{6p_{\max}} \cdot \mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell] \end{aligned} \tag{7}$$

– *Case* 3*:* i, j, k, ℓ distinct. Notice that

$$\begin{aligned}\mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell] &= \mathbb{E}[\widetilde{Z}_i] \cdot \mathbb{E}[\widetilde{Z}_j] \cdot \mathbb{E}[\widetilde{Z}_k] \cdot \mathbb{E}[\widetilde{Z}_\ell] \\ &= e^{-d_i - d_j - d_k - d_\ell}.\end{aligned}$$

Therefore, similarly to case 1, we have

$$\begin{aligned}\mathbb{E}[Z_i Z_j Z_k Z_\ell] &= \prod_{f \in \binom{\{i,j,k,\ell\}}{2}} (1 - p_f) \prod_{\substack{(i,u)\\ u \neq j,k,\ell}} (1 - p_{iu}) \prod_{\substack{(j,v)\\ v \neq i,k,\ell}} (1 - p_{jv}) \prod_{\substack{(k,w)\\ w \neq i,j,\ell}} (1 - p_{kw}) \prod_{\substack{(\ell,x)\\ x \neq i,j,k}} (1 - p_{\ell x}) \\ &\leq e^{-d_i - d_j - d_k - d_\ell + p_{ij} + p_{ik} + p_{i\ell} + p_{jk} + p_{j\ell} + p_{k\ell}} \\ &\leq e^{6p_{\max}} \cdot \mathbb{E}[\widetilde{Z}_i \widetilde{Z}_j \widetilde{Z}_k \widetilde{Z}_\ell] \end{aligned} \tag{8}$$

Consequently,

$$\begin{aligned}\mathbb{E}[\widetilde{N}^2] &= \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < \ell \leq n} \left[e^{-d_i - d_j} \mathbf{1}_{\{i=k, j=\ell\}} + e^{-d_i - d_j - d_\ell} \mathbf{1}_{\{j=k\}} + e^{-d_i - d_j - d_k - d_\ell} \mathbf{1}_{\{i \neq j \neq k \neq \ell\}} \right] \\ &\leq \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < \ell \leq n} \left[e^{-d_i - d_j} + e^{-d_i - d_j - d_\ell} + e^{-d_i - d_j - d_k - d_\ell} \right] \\ &\leq \mu^2 + \mu^3 + \mu^4, \end{aligned} \tag{9}$$

where $\mu \triangleq \sum_{1 \leq i \leq n} e^{-d_i}$. Combining equations (4) and (5) with (6), (7), (8) and (9) yields that

$$\mathbb{E}[N^2] \leq e^{6p_{\max}} \cdot \mathbb{E}[\widetilde{N}^2] \leq e^{6p_{\max}} \mu^2 (\mu^2 + \mu + 1). \tag{10}$$

Bounding the first moment. Notice that

$$\mathbb{E}[\widetilde{N}] = \sum_{1\le i<j\le n} \mathbb{E}[\widetilde{Z}_i\widetilde{Z}_j] = \sum_{1\le i<j\le n} e^{-d_i-d_j} = \frac{1}{2}\Big(\mu^2 - \sum_{i=1}^{n} e^{-2d_i}\Big) \ge \frac{\mu}{2}\left(\mu - \max_{1\le i\le n} e^{-d_i}\right), \tag{11}$$

where the last step uses that $e^{-2d_k} \le e^{-d_k}\cdot \max_{1\le i\le n} e^{-d_i}$ for each k. Furthermore,

$$\begin{aligned}
\mathbb{E}[N] &= \sum_{1\le i<j\le n} \mathbb{E}\left[Z_iZ_j\right] \\
&= \sum_{1\le i<j\le n} \Big[(1-p_{ij}) \prod_{\substack{(i,u)\\ u\ne j}} (1-p_{iu}) \prod_{\substack{(j,v)\\ v\ne i}} (1-p_{jv})\Big] \\
&\overset{(a)}{\ge} \sum_{1\le i<j\le n} \Big[e^{-p_{ij}-p_{ij}^2} \prod_{\substack{(i,u)\\ u\ne j}} e^{-p_{iu}-p_{iu}^2} \prod_{\substack{(j,v)\\ v\ne i}} e^{-p_{jv}-p_{jv}^2}\Big] \\
&\overset{(b)}{\ge} e^{-(2n-1)p_{\max}^2} \cdot \sum_{1\le i<j\le n} e^{-d_i-d_j+p_{ij}} \\
&\overset{(c)}{\ge} e^{-(2n-1)p_{\max}^2}\cdot \mathbb{E}[\widetilde{N}],
\end{aligned} \tag{12}$$

where (a) uses that $1-x \ge e^{-x-x^2}$ for all $x\in[0,1/2]$, (b) uses that $e^{-p_{ij}^2} \ge e^{-p_{\max}^2}$, and (c) uses that $e^{p_{ij}} \ge 1$ for all i and j. Finally, combining equations (10), (11) and (12), we have that

$$\frac{\mathbb{E}[N]^2}{\mathbb{E}[N^2]} \ge e^{-(4n-2)p_{\max}^2-6p_{\max}} \frac{\mathbb{E}[\widetilde{N}]^2}{\mathbb{E}[\widetilde{N}^2]} \ge \exp\left(-(4n-2)p_{\max}^2 - 6p_{\max}\right) \times \frac{\mu^2 - 2\mu\max_{1\le i\le n} e^{-d_i}}{4(\mu^2+\mu+1)},$$

where the last step uses that

$$\mathbb{E}[\widetilde{N}]^2 \ge \frac{\mu^2}{4}\left(\mu^2 - 2\mu\cdot\max_{1\le i\le n} e^{-d_i} + \max_{1\le i\le n} e^{-2d_i}\right) \ge \frac{\mu^2}{4}\left(\mu^2 - 2\mu\max_{1\le i\le n} e^{-d_i}\right).$$

Since $\mathbb{P}\left(N\ge 1\right) \ge \frac{\mathbb{E}[N]^2}{\mathbb{E}[N^2]}$, the proof is complete. □

B.1 Proof of Lemma 1

Proof of Lemma 1. Let N be the number of pairs of isolated nodes in H. It follows from Lemma B.1 that

$$\begin{aligned}
\mathbb{P}\left(N\ge 1\right) &\ge \exp\left(-(2n-4)q_{\max}^2 - 6q_{\max}\right) \\
&\qquad \times \frac{\lambda^2 - 2\lambda\max_{1\le i\le n} e^{-r_i}}{4(\lambda^2+\lambda+1)} \\
&\ge (1-o(1))\cdot \frac{\lambda^2 - 2\lambda\max_{1\le i\le n} e^{-r_i}}{4(\lambda^2+\lambda+1)},
\end{aligned} \tag{13}$$

where $\lambda \triangleq \sum_{i=1}^{n} e^{-r_i}$, and the last step in (13) uses that $q_{\max} = o(1/\sqrt{n})$. Let $r_{[1]} \le r_{[2]} \le \cdots \le r_{[n]}$ denote a sorted copy of the sequence $r_1,\cdots,r_n$. Consider the following three cases.

– *Case* 1. $r_{[1]} = \omega(1)$. In this setting, it follows that

$$\lambda \max_{1\le i\le n} e^{-r_i} = \lambda e^{-r_{[1]}} = o(1),$$

since $\lambda = \Omega(1)$. Thus,

$$\mathbb{P}(N \ge 1) \ge (1 - o(1)) \cdot \frac{\lambda^2 - o(1)}{4(\lambda^2 + \lambda + 1)} = \Omega(1).$$

– *Case* 2. $r_{[2]} = O(1)$. In this setting, there exists a constant $C > 0$ such that $C \le e^{-r_{[2]}} \le e^{-r_{[1]}}$. Let A denote the adjacency matrix of G and let $R_{[1]}$ and $R_{[2]}$ denote the degree of two nodes with mean degree $r_{[1]}$ and $r_{[2]}$ respectively. Then,

$$\begin{aligned}
\mathbb{P}\left(R_{[1]} = 0, R_{[2]} = 0\right) &= \mathbb{P}\left(R_{[1]} = 0\right) \cdot \mathbb{P}\left(R_{[2]} = 0 \mid R_{[1]} = 0\right) \\
&\ge \mathbb{P}\left(R_{[1]} = 0\right) \mathbb{P}\left(R_{[2]} = 0\right) \\
&= \prod_{i\ne[1]} (1 - q_{[1]i}) \prod_{j\ne[2]} (1 - q_{[2]j}) \\
&\overset{(a)}{\ge} \prod_{i\ne[1]} e^{-q_{[1]i} - q_{[1]i}^2} \prod_{j\ne[2]} e^{-q[2]j - q_{[2]j}^2} \\
&= e^{-r_{[1]}} e^{-r_{[2]}} \cdot \exp\Big(- \sum_{i\ne[1]} q_{[1]i}^2 - \sum_{j\ne[2]} q_{[2]j}^2 \Big) \\
&\overset{(b)}{\ge} \exp\left(-2(n-1)q_{\max}^2\right) \cdot C^2 \overset{(c)}{=} \Omega(1).
\end{aligned}$$

Here, (a) uses the inequality $1 - x \ge e^{-x-x^2}$ whenever $x \in [0, 1/2]$, (b) uses that $e^{-q_{ij}^2} \ge e^{-q_{\max}^2}$ for all i, j and also that $C \le e^{-r_{[2]}} \le e^{-r_{[1]}}$. Finally, (c) uses that $\exp\left(-2(n-1)q_{\max}^2\right) = \Omega(1)$ whenever $q_{\max} = o(1/\sqrt{n})$.

– *Case* 3. $r_{[2]} = \omega(1)$. Consider the induced subgraph G' of G on the vertex set $V = V - \{[1]\}$. The mean degrees of G' satisfy $r'_{[i]} \le r_{[i]}$, and so it follows that

$$\lambda' \triangleq \sum_{i=2}^{n} e^{-r'_{[i]}} \ge \sum_{i=2}^{n} e^{-r_{[i]}} = \Omega(1),$$

and the smallest mean degree in G' diverges, i.e. $r_{[2]} = \omega(1)$. Equivalently, the graph G' satisfies the conditions for Case 1 above, and so there is a constant lower bound C on the probability that the number of *pairs* of isolated nodes N' in G' is at least one. If node [1] does not connect to an isolated pair of nodes in G', then the pair remains isolated in G as well. Therefore,

$$\begin{aligned}
\mathbb{P}(N \ge 1) &\ge (1 - q_{\max})^2 \cdot \mathbb{P}\left(N' \ge 1\right) \\
&\ge (1 - q_{\max})^2 \cdot C = \Omega(1).
\end{aligned}$$

In all cases, we have that $\mathbb{P}(N \ge 1) = \Omega(1)$, i.e. there is a non-vanishing probability that two isolated nodes exist in H. This concludes the proof. $\square$

C Supplementary proofs for Section 4

The following facts about random variables are used frequently in the analysis. First, we have concentration inequalities for a sum of Bernoulli random variables.

Lemma C.1. *Let X be a sum of n independent Bernoulli random variables, and let $\mathbb{E}[X] = np$. Then,*

1. *For any $\delta > 0$,*

$$\mathbb{P}\left(X \geq (1+\delta)np\right) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{np} \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)np}. \tag{14}$$

2. *For any $\delta > 5$,*

$$\mathbb{P}\left(X \geq (1+\delta)np\right) \leq 2^{-(1+\delta)np}. \tag{15}$$

3. *For any $\delta \in (0,1)$,*

$$\mathbb{P}\left(X \leq (1-\delta)np\right) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{np} = e^{-np} \cdot \left(\frac{e}{1-\delta}\right)^{(1-\delta)np}. \tag{16}$$

Proof. All proofs follow from the Chernoff bound and can be found, or easily derived, from Theorems 4.4 and 4.5 of (Mitzenmacher and Upfal 2017). ☐

Next, we have two lemmas about stochastic ordering of Poisson random variables.

Lemma C.2. *Let $p \in [0,1]$. Consider independent random variables X, Y and Z with distributions*

$$X \sim \mathrm{Bern}(p), \quad Y \sim \mathrm{Pois}(p), \quad Z \sim 2 \cdot \mathrm{Pois}(p/2).$$

Then, for all t, the moment generating functions of X, Y and Z are ordered as

$$M_X(t) \leq M_Y(t) \leq M_Z(t).$$

Proof. We have

$$M_X(t) = 1 + p(e^t - 1) \overset{(a)}{\leq} \exp\left(p(e^t-1)\right) = M_Y(t) \overset{(b)}{\leq} \exp\left(p\left(\frac{e^{2t}-1}{2}\right)\right) = M_Z(t),$$

where (a) uses $1 + x \leq e^x$ and (b) uses that the function $\left(e^{kt} - 1\right)/k$ is increasing in k for all t. ☐

Lemma C.3. *Let $n > 0$ and let $X_1, \cdots, X_n$ be independent random variables with $X_i \sim$ Bern(p_i). Let m be an integer so that $0 \leq m \leq n$, and define*

$$X \triangleq \sum_{i \leq m} X_i + \sum_{j > m} 2X_j.$$

Then, for any $t < \mathbb{E}X$,

$$\mathbb{P}\left(X \leq t\right) \leq \exp\left(-\frac{\mathbb{E}X}{2} + \frac{t}{2}\log\left(\frac{e \cdot \mathbb{E}X}{t}\right)\right)$$

Proof. Let $Y_i \sim \text{Pois}(p_i)$ and $Z_i \sim 2 \cdot \text{Pois}(p_i/2)$. It follows that

$$\begin{aligned} M_X(t) &= \prod_{i=1}^{m} M_{X_i}(t) \cdot \prod_{j=m+1}^{n} M_{X_j}(2t) \\ &\leq \prod_{i=1}^{m} M_{Z_i}(t) \cdot \prod_{j=m+1}^{n} M_{Y_j}(2t) \\ &= \prod_{i=1}^{m} M_{Z_i}(t) \cdot \prod_{j=m+1}^{n} M_{2Y_j}(t) = M_Z(t), \end{aligned}$$

where $Z \triangleq \sum_{i=1}^{m} Z_i + \sum_{j=m+1}^{n} 2Y_j$ is distributed as $Z \sim 2 \cdot \text{Pois}(\mathbb{E}X/2)$.

Therefore, Chernoff tail bounds that hold for Z also hold for X. In particular, for any $t < \mathbb{E}X$,

$$\begin{aligned} \mathbb{P}(X \leq t) &\leq \inf_{\theta \geq 0} e^{\theta t} M_X(-\theta) \\ &\leq \inf_{\theta \geq 0} e^{\theta t} M_Z(-\theta) \\ &= \inf_{\theta \geq 0} \exp\left(\theta t + \frac{\mathbb{E}X}{2}\left(e^{-2\theta} - 1\right)\right). \end{aligned}$$

Setting $\theta = \frac{1}{2}\log(\mathbb{E}X/t)$ gives the desired result. □

C.1 Proof of Lemma 3

If strictly more than n^δ nodes all have degree less than or equal to r, then a subset of exactly n^δ nodes has average degree less than r, so

$$\begin{aligned} \mathbb{P}\left(|Z_r| > n^\delta\right) &= \mathbb{P}\Big(\exists\, T \subseteq [n] \colon \left\{|T| \geq n^\delta\right\} \cap \big\{\max_{u \in T} \delta_G(u) \leq r\big\}\Big) \\ &= \mathbb{P}\Big(\exists\, T \subseteq [n] \colon \left\{|T| = n^\delta\right\} \cap \big\{\max_{u \in T} \delta_G(u) \leq r\big\}\Big) \\ &\leq \mathbb{P}\Big(\exists\, S \subseteq [n] \colon \left\{|S| = n^\delta\right\} \cap \big\{\sum_{u \in S} \delta_G(u) \leq r|S|\big\}\Big) \end{aligned} \tag{17}$$

Let G_{uv} denote the indicator random variable for the presence of edge $\{u, v\}$ in G. For any set $S \subseteq [n]$, define

$$\mu_S \triangleq \sum_{i \in S} \mathbb{E}[\delta_G(i)] = \sum_{i \in S} d_i.$$

Since for any $i \in [n]$, we have that $e^{-d_i} \leq \sum_{j=1}^{n} e^{-d_j} \leq n^{-\alpha}$, it follows that $d_i \geq \alpha \log n$. Therefore, for any $S \subseteq [n]$, we have that $\mu_S \geq \alpha |S| \log n$. Hence,

$$\begin{aligned} \mathbb{P}\left(\sum_{u \in S} \delta_G(u) \leq r|S|\right) &= \mathbb{P}\Big(\sum_{\substack{u \in S \\ v \in [n]\setminus S}} G_{uv} + \sum_{\substack{u,v \in S \\ u < v}} 2G_{uv} \leq r|S|\Big) \\ &\overset{\text{(a)}}{\leq} \exp\left(-\frac{\mu_S}{2} + \frac{r|S|}{2}\log\left(\frac{e \cdot \mu_S}{r|S|}\right)\right) \\ &\overset{\text{(b)}}{\leq} \exp\left(-\frac{\alpha|S|}{2}\log(n) + \frac{r|S|}{2}\log\left(\frac{e \cdot \alpha \log n}{r}\right)\right) \end{aligned} \tag{18}$$

where (a) uses Lemma C.3 and (b) uses that the function on the right hand side of (b) is decreasing in μ_S on the interval $(r|S|, \infty)$, and that $r|S| \leq \alpha|S| \log n \leq \mu_S$ for all sufficiently large n. Finally, combining (17) and (18), and using that $|S| = n^\delta$ yields

$$\begin{aligned}
\mathbb{P}\left(|Z_r| > n^\delta\right) &\leq \sum_{S \subseteq [n]: |S| = n^\delta} \mathbb{P}\Big(\sum_{u \in S} \delta_G(v) \leq r|S|\Big) \\
&\leq \binom{n}{n^\delta} \exp\left(-\frac{\alpha n^\delta}{2} \log n + \frac{r n^\delta}{2} \log\left(\frac{e\alpha \log n}{r}\right)\right) \\
&\overset{\text{(a)}}{\leq} \exp\Big(\big(1-\delta-\frac{\alpha}{2}\big) n^\delta \log n + \frac{r n^\delta}{2} \log\left(\frac{e\alpha \log n}{r}\right)\Big) \\
&= o(1/n),
\end{aligned}$$

whenever $\delta > 1 - \alpha/2$. Here, (a) uses the fact that $\binom{n}{n^\delta} \leq \exp\left((1-\delta) n^\delta \log n\right)$. Hence, (11) follows.

Suppose that $p_{\max} = o(n^{\alpha/2-1})$. Notice that

$$\sum_{u \in S} \delta_G(u) = \sum_{u \in S} \sum_{\substack{v \in [n] \\ v \neq u}} G_{uv} \geq \sum_{u \in S} \sum_{v \in [n] \setminus S} G_{uv} \triangleq Y_S,$$

and so we have that

$$\mathbb{P}\left(\sum_{u \in S} \delta_G(u) \leq r|S|\right) \leq \mathbb{P}\left(Y_S \leq r|S|\right) \overset{\text{(a)}}{\leq} e^{-\mathbb{E}[Y_S]} \cdot \left(\frac{e \cdot \mathbb{E}\left[Y_S\right]}{r|S|}\right)^{r|S|}, \tag{19}$$

where (a) uses that Y_S is a sum of independent Bernoulli random variables, and so the Chernoff tail bound (16) can be applied. On the other hand,

$$\mathbb{E}\left[Y_S\right] = \sum_{u \in S} d_u - \sum_{u \in S} \sum_{\substack{v \in S \\ v \neq u}} 2p_{uv} \overset{\text{(a)}}{\geq} \sum_{u \in S} d_u - \binom{|S|}{2} \cdot (2p_{\max}) \overset{\text{(b)}}{\geq} \alpha|S| \log(n) - |S|^2 \cdot p_{\max} \triangleq \widetilde{\mu}_S,$$

where (a) uses that $p_{uv} \leq p_{\max}$ and (b) uses that $\mu_S \geq \alpha|S| \log n$. Continuing from (19), we have

$$\mathbb{P}\left(Y_S \leq r|S|\right) \leq e^{-\mathbb{E}[Y_S]} \left(\frac{e \cdot \mathbb{E}\left[Y_S\right]}{r|S|}\right)^{r|S|} \leq e^{-\widetilde{\mu}_S} \left(\frac{e \cdot \widetilde{\mu}_S}{r|S|}\right)^{r|S|}, \tag{20}$$

where the last step uses that the function $x \mapsto e^{-x} \cdot \left(\frac{e \cdot x}{r|S|}\right)^{r|S|}$ is decreasing in x whenever $x \geq r|S|$. Thus, for any S such that $|S| = n^\delta$, we have from (20) that

$$\begin{aligned}
\mathbb{P}\left(Y_S \leq r n^\delta\right) &\leq \exp\left(-n^\delta\left(\alpha \log(n) - r - n^\delta p_{\max}\right)\right) \cdot \left(\frac{\alpha \log(n) - n^\delta p_{\max}}{r}\right)^{r n^\delta} \\
&= \exp\left(-n^\delta\left(\alpha \log(n) - r - n^\delta p_{\max} - r \log\left(\frac{\alpha \log(n) - n^\delta p_{\max}}{r}\right)\right)\right).
\end{aligned} \tag{21}$$

Combining this with (17), a union bound yields that

$$\begin{aligned}
\mathbb{P}\left(|Z_r| > n^\delta\right) &\leq \mathbb{P}\Big(\exists\, S \subseteq [n] : \left\{|S| = n^\delta\right\} \cap \Big\{\sum_{u \in S} \delta_G(u) \leq r|S|\Big\}\Big) \\
&\leq \binom{n}{n^\delta} \mathbb{P}\left(Y_S \leq r n^\delta\right) \\
&\overset{\text{(a)}}{\leq} \exp\Big(-n^\delta\Big[(\alpha+\delta-1)\log(n) - n^\delta p_{\max} - r \log\Big(\frac{\alpha \log(n) - n^\delta p_{\max}}{r}\Big) - r - 1\Big]\Big),
\end{aligned} \tag{22}$$

where (a) uses that $\binom{n}{n^\delta} \leq \exp\big((1-\delta)n^\delta \log n\big)$. Finally, note that the right-hand side of (22) is $o(1/n)$ whenever $\alpha+\delta > 1$ and $n^\delta p_{\max} = o(\log n)$. Both conditions are satisfied when $p_{\max} = o(n^{\alpha/2-1})$ by choosing δ such that $1-\alpha < \delta < 1-\alpha/2$.

C.2 Proof of Lemma 5

(a) The proof is by construction: Since U_f is obtained by adding exactly f nodes to U_0, it follows that $U_f^c \subseteq U_0^c = Z_{k+1}^c$, so each node in U_f^c has degree $k+2$ or more in $G-v$. Further, each node in U_f^c has at most 2 neighbors in U_f, else the **for** loop would not have terminated. Thus, the subgraph of $G-v$ induced on the set U_f^c has minimum degree at least k, and the result follows.

(b) Let δ' be such that $1-\alpha < \delta' < \min\{\delta, 1-\alpha+2\varepsilon\}$. If $|U_f| > 3n^\delta$, then either $|U_0| > 3n^{\delta'}$ or there is some M in $\{0,1,\cdots,f\}$ for which $|U_M| = 3n^{\delta'}$. Therefore,

$$\begin{aligned}\mathbb{P}\left(|U_f| > 3n^\delta\right) &\leq \mathbb{P}\left(|U_0| > 3n^{\delta'}\right) + \mathbb{P}\big(\exists\ M \in \{0,1,\cdots,f\} \text{ s.t. } |U_M| = 3n^{\delta'}\big)\\ &= o(1/n) + \underbrace{\mathbb{P}\big(\exists\ M \in \{0,1,\cdots,f\} \text{ s.t. } |U_M| = 3n^{\delta'}\big)}_{(\star)},\end{aligned}$$

where the first term is $o(1/n)$ by Lemma 3. Note that each iteration $i = 0,1,\cdots,M-1$ of the **for** loop adds exactly 1 vertex and at least 3 edges to the subgraph of $G-v$ induced on U_M. Therefore, the induced subgraph $G|_{U_M}$ has $3n^\delta$ vertices and at least $3\left(|U_M|-|U_0|\right)$ edges.

Let us couple the inhomogeneous random graph G with an Erdős-Rényi random graph $\widetilde{G}$ on n nodes with marginal distribution $\widetilde{G} \sim \mathrm{ER}(n, p_{\max})$, such that G is a subgraph of $\widetilde{G}$. Thus,

$$\begin{aligned}(\star) &\leq \mathbb{P}\Big(\exists \text{ subgraph } H\!=\!(W,F) \text{ of } G-v \text{ s.t. } |W| = 3n^{\delta'} \text{ and } |F| \geq 3\big(3n^{\delta'} - |U_0|\big)\Big)\\ &\leq \mathbb{P}\left(|U_0| > n^{\delta'}\right) + \mathbb{P}\big(\exists \text{ subgraph } H = (W,F) \text{ of } G \text{ s.t. } |W| = 3n^{\delta'} \text{ and } |F| \geq 6n^{\delta'}\big)\\ &\leq \mathbb{P}\left(|U_0| > n^{\delta'}\right) + \mathbb{P}\big(\exists \text{ subgraph } \widetilde{H}\!=\!(\widetilde{W},\widetilde{F}) \text{ of } \widetilde{G} \text{ s.t. } |\widetilde{W}| = 3n^{\delta'} \text{ and } |\widetilde{F}| \geq 6n^{\delta'}\big)\\ &\leq o(1/n) + \underbrace{\binom{n}{3n^{\delta'}}\mathbb{P}\Big(\mathrm{Bin}\Big(\binom{3n^{\delta'}}{2}, p_{\max}\Big) > 6n^{\delta'}\Big)}_{(\dagger)},\end{aligned}$$

where the last step uses Lemma 3 and a union bound over all possible choices of W. Finally, using that $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ and the fact that $\mathrm{Bin}(n,p) \preceq \mathrm{Bin}(n',p')$ whenever $n' \geq n$ and $p' \geq p$ yields

$$\begin{aligned}(\dagger) &\leq \left(\frac{n^{1-\delta'}e}{3}\right)^{3n^{\delta'}} \mathbb{P}\Big(\mathrm{Bin}\Big(\frac{9n^{2\delta'}}{2}, n^{\alpha/2-\varepsilon-1}\Big) > 6n^{\delta'}\Big)\\ &\overset{(a)}{\leq} \left(n^{1-\delta'}\right)^{3n^{\delta'}} \cdot \left(\frac{3e}{4}\cdot n^{\alpha/2-\varepsilon-(1-\delta')}\right)^{6n^{\delta'}}\\ &= \left(\frac{3e}{4}\cdot n^{\alpha/2-\varepsilon-(1-\delta')/2}\right)^{6n^{\delta'}} = o(1/n),\end{aligned}$$

since $\alpha/2 - \varepsilon - (1-\delta')/2 < 0$. Note (a) uses the concentration bound (14) from Lemma C.1. This concludes the proof of Lemma 5.

D Supplementary proofs for Section 5

D.1 Proof of Lemma 7

For any vertex cut U,

$$\begin{aligned}\left\{\widetilde{c}_v(U) > \frac{m^2}{4}\left(k+\frac{2}{\alpha}\right)\right\} &\overset{(a)}{\subseteq} \left\{\widetilde{c}_v(U) > |U|\,(m-|U|)\left(k+\frac{2}{\alpha}\right)\right\}\\ &= \left\{\sum_{i\in U}\sum_{j\in U^c}\delta_{G'_i\wedge G'_j}(v) > |U|(m-|U|)\left(k+\frac{2}{\alpha}\right)\right\}\\ &\subseteq \bigcup_{i\in U}\bigcup_{j\in U^c}\left\{\delta_{G'_i\wedge G'_j}(v) > k+\frac{2}{\alpha}\right\},\end{aligned}$$

where (a) uses the fact that the maximum of a set of numbers is greater than or equal to the average. On the other hand,

$$\{c_v(U)=0\} = \bigcap_{i\in U}\bigcap_{j\in U^c}\left\{v\notin \mathrm{core}_k(G'_i\wedge G'_j)\right\}.$$

It follows from the union bound that

$$\begin{aligned}&\mathbb{P}\left(\{c_v(U)=0\}\cap\left\{\widetilde{c}_v(U) > \frac{m^2}{4}\left(k+\frac{2}{\alpha}\right)\right\}\right)\\ &\quad\le \sum_{i\in U}\sum_{j\in U^c}\mathbb{P}\Big(\left\{v\notin \mathrm{core}_k(G'_i\wedge G'_j)\right\}\cap\left\{\delta_{G'_i\wedge G'_j}(v) > k+\frac{2}{\alpha}\right\}\Big) = o(1/n),\end{aligned}$$

since for any choice of i and j, the graph $G'_i\wedge G'_j \sim \mathrm{RG}\left(n,\mathbf{P}s^2\right)$.

D.2 Proof of Lemma 8

We start by making a simple observation about products of Binomial random variables.

Lemma D.1. *Let $X_1,\cdots,X_m\sim\mathrm{Bern}(s)$ be i.i.d. random variables, and let $B = X_1+\cdots+X_m$ denote their sum. For each ℓ in $\{1,2,\cdots,\lfloor m/2\rfloor\}$, define*

$$T_\ell = (X_1+\cdots+X_\ell)\,(X_{\ell+1}+\cdots+X_m).$$

For any $\ell_1,\ell_2\in\{1,2,\cdots,\lfloor m/2\rfloor\}$ such that $\ell_1<\ell_2$, and for any $t\in\mathbb{R}$ and any $b\in\{0,1,\cdots,m\}$,

$$\mathbb{P}\left(T_{\ell_1} > t \mid B=b\right) \le \mathbb{P}\left(T_{\ell_2} > t \mid B=b\right). \tag{23}$$

Proof. Consider overlapping but exhaustive cases:

Case 1: $t<0$. Since $T_\ell\ge 0$ almost surely for all ℓ, the inequality (23) holds.

Case 2: $t\ge b-1$. Note that conditioned on $B=b$, it follows that $T_1\in\{0,b-1\}$. Therefore, the left hand side of (23) equals zero, and the inequality holds.

Case 3: $b=0$ or $b=1$. In this case, T_ℓ is identically zero for all ℓ, so (23) holds.

Case 4: $b \geq 2$ and $0 \leq t < b-1$. For ease of notation, define for any indices $a < b$ the quantity $X_{a:b} \coloneqq \sum_{i=a}^{b} X_i$. For any $\ell \in \{1, 2, \cdots, \lfloor m/2 \rfloor\}$,

$$\begin{aligned}
&\mathbb{P}\left(T_\ell > t \mid B = b\right) &(24)\\
&= \frac{\mathbb{P}\left(\{X_{1:\ell} \cdot X_{\ell+1:m} > t\} \cap \{X_{1:m} = b\}\right)}{\mathbb{P}\left(X_{1:m} = b\right)} \\
&= \frac{\sum_{i:i(b-i)>t} \mathbb{P}\left(\{X_{1:\ell} = i\} \cap \{X_{\ell+1:m} = b-i\}\right)}{\mathbb{P}\left(X_{1:m} = b\right)} \\
&\overset{\text{(a)}}{=} \frac{\sum_{i=1}^{b-1} \mathbb{P}\left(X_{1:\ell} = i\right) \mathbb{P}\left(X_{\ell+1:m} = b-i\right)}{\mathbb{P}\left(X_{1:m} = b\right)} \\
&\overset{\text{(b)}}{=} \frac{\sum_{i=1}^{b-1} \binom{\ell}{i}\binom{m-\ell}{b-i}}{\binom{m}{b}} \\
&= \frac{\sum_{i=0}^{b} \binom{\ell}{i}\binom{m-\ell}{b-i} - \binom{m-\ell}{b} - \binom{\ell}{b}}{\binom{m}{b}} \\
&= \frac{\binom{m}{b} - \binom{m-\ell}{b} - \binom{\ell}{b}}{\binom{m}{b}}, &(25)
\end{aligned}$$

where (a) used the fact that for any t such that $0 \leq t < b-1$, it is true that

$$\{i : i(b-i) > t\} = \{1, 2, \cdots, b-1\}.$$

Here, the notation for binomial coefficients in (b) involves setting $\binom{n}{k} = 0$ whenever $k < 0$ or $k > n$. Let $f_{m,b}(\ell)$ denote the numerator of (25), i.e.

$$f_{m,b}(\ell) \triangleq \binom{m}{b} - \binom{m-\ell}{b} - \binom{\ell}{b}$$

It suffices to show that $f_{m,b}(\ell) - f_{m,b}(\ell-1) \geq 0$ for all $\ell \in \{2, \cdots, \lfloor m/2 \rfloor\}$. Indeed,

$$\begin{aligned}
f_{m,b}(\ell) - f_{m,b}(\ell-1) &= \binom{m-\ell+1}{b} - \binom{m-\ell}{b} - \left(\binom{\ell}{b} - \binom{\ell-1}{b}\right) \\
&\overset{\text{(c)}}{=} \binom{m-\ell}{b-1} - \binom{\ell-1}{b-1} \geq 0,
\end{aligned}$$

whenever $m - \ell \geq \ell - 1$, i.e. $\ell \leq \lfloor m/2 \rfloor$. Here, (c) uses the identity $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, and the fact that $\binom{n_1}{k} \geq \binom{n_2}{k}$ whenever $n_1 \geq n_2$. This concludes the proof. $\square$

Corollary D.1. *Let F be a collection of edges in the parent graph G. For any edge $e_r \in F$, let X_i^r denote the indicator random variable $G_i'(e_r) \sim \mathrm{Bern}(ps)$. For each ℓ in $\{1, \cdots, \lfloor m/2 \rfloor\}$, define*

$$T_\ell^r = (X_1^r + \cdots + X_\ell^r)(X_{\ell+1}^r + \cdots + X_m^r).$$

Then, for any $\ell_1, \ell_2 \in \{1, \cdots, \lfloor m/2 \rfloor\}$ such that $\ell_1 < \ell_2$, the following stochastic ordering holds

$$\sum_{r=1}^{|F|} T_{\ell_1}^r \preceq \sum_{r=1}^{|F|} T_{\ell_2}^r.$$

Proof. It suffices to show that $T^r_{\ell_1} \preceq T^r_{\ell_2}$ for each r, since the edges are independent. For any t,

$$\begin{aligned}\mathbb{P}\left(T^r_{\ell_1} > t\right) &= \sum_{b=0}^{m} \mathbb{P}\left(B = b\right) \mathbb{P}\left(T^r_{\ell_1} > t | B = b\right) \\ &\leq \sum_{b=0}^{m} \mathbb{P}\left(B = b\right) \mathbb{P}\left(T^r_{\ell_2} > t | B = b\right) = \mathbb{P}\left(T^r_{\ell_2} > t\right),\end{aligned}$$

which concludes the proof. $\square$

With this, we are ready to prove Lemma 8. Let $\ell_1, \ell_2 \in \{1, \cdots, \lfloor m/2 \rfloor\}$ such that $\ell_1 < \ell_2$. Let $t \in \mathbb{R}$. Consider the parent graph G and label the set of incident edges on v as $\{e_1, \cdots, e_{\delta_G(v)}\}$. Denote by X^r_i the indicator random variable $G'_i(e_r) \sim \text{Bern}(ps)$. It follows that

$$\begin{aligned}\mathbb{P}\left(\widetilde{c}_v(Z_{\ell_2}) > t\right) &= \mathbb{P}\left(\sum_{i=1}^{\ell_2} \sum_{j=\ell_2+1}^{m} \delta_{G'_i \wedge G'_j}(v) \geq t\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{\ell_2} \sum_{j=\ell_2+1}^{m} \sum_{r=1}^{\delta_G(v)} X^r_i X^r_j > t\right) \\ &= \sum_{d=0}^{n} \mathbb{P}\left(\delta_G(v) = d\right) \times \mathbb{P}\Big(\sum_{r=1}^{d} (X^r_1 + \cdots + X^r_{\ell_2})(X^r_{\ell_2+1} + \cdots + X^r_m) > t\Big) \\ &\overset{(a)}{\geq} \sum_{d=0}^{n} \mathbb{P}\left(\delta_G(v) = d\right) \times \mathbb{P}\Big(\sum_{r=1}^{d} (X^r_1 + \cdots + X^r_{\ell_1})(X^r_{\ell_1+1} + \cdots + X^r_m) > t\Big) \\ &= \mathbb{P}\left(\sum_{i=1}^{\ell_1} \sum_{j=\ell_1+1}^{m} \sum_{r=1}^{\delta_G(v)} X^r_i X^r_j > t\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{\ell_1} \sum_{j=\ell_1+1}^{m} \delta_{G'_i \wedge G'_j}(v) \geq t\right) \\ &= \mathbb{P}\left(\widetilde{c}_v(Z_{\ell_1}) > t\right),\end{aligned}$$

as desired. Here, (a) uses Corollary D.1.

D.3 Proof of Lemma 9

If $s = 1$, then the graphs $G_1, \cdots, G_m$ are all isomorphic, and

$$(\star) = \mathbb{P}\left(\exists\, v \in V \setminus v^* : \delta_{G_1}(v) \leq r\right) \leq \sum_{v \in V \setminus v^*} \mathbb{P}\left(\delta_{G_1}(v) = 0\right) \leq \sum_{i=1}^{n} e^{-d_i} = o(1),$$

and so the result follows. We therefore assume without loss of generality that $s < 1$. For a vertex v, graph k and integer i, let $H^v_k(i)$ denote the event

$$H^v_k(i) \coloneqq \{\delta_{G_1 \wedge G'_k}(v) = i\}.$$

Observe that

$$
\begin{aligned}
(\star) &\le \mathbb{P}\Big(\bigcup_{r_2=0}^{r}\cdots\bigcup_{r_m=0}^{r}\big(\exists\, v\in V\setminus v^* : H_2^v(r_2)\cap\cdots\cap H_m^v(r_m)\big)\Big)\\
&\le \sum_{r_2=0}^{r}\cdots\sum_{r_m=0}^{r}\sum_{v\in V\setminus v^*}\mathbb{E}_{D_v}\big[\mathbb{P}\big(H_2^v(r_2)\cap\cdots\cap H_m^v(r_m)\mid \delta_{G_1}(v)=D_v\big)\big]\\
&\overset{\text{(a)}}{=} \sum_{r_2=0}^{r}\cdots\sum_{r_m=0}^{r}\sum_{v\in V\setminus v^*}\mathbb{E}_{D_v}\left[\prod_{\ell=2}^{m}\binom{D_v}{r_\ell}s^{r_\ell}(1-s)^{D_v-r_\ell}\cdot\mathbf{1}_{\{r_\ell\le D_v\}}\right]\\
&\overset{\text{(b)}}{\le} \sum_{r_2=0}^{r}\cdots\sum_{r_m=0}^{r}\left[\prod_{\ell=2}^{m}\left(\frac{e}{r_\ell}\cdot\frac{s}{1-s}\right)^{r_\ell}\right]\cdot\sum_{v\in V\setminus v^*}\mathbb{E}_{D_v}\left[\prod_{\ell=2}^{m}D_v^{r_\ell}(1-s)^{D_v}\right],
\end{aligned}
\tag{26}
$$

where (a) uses that the degrees of node v in the graphs $G_1\wedge G_2',\cdots,G_1\wedge G_m'$ are conditionally independent of the degree of v in G_1, and that each edge of G_1 is independently present in $G_1\wedge G_j'$ with probability s. Further, (b) uses that $\binom{n}{k}\le(ne/k)^k$ with the implicit definition $(1/0)^0\triangleq 1$. Note that

$$
\begin{aligned}
\sum_{v\in V\setminus v^*}\mathbb{E}_{D_v}\left[\prod_{\ell=2}^{m}D_v^{r_\ell}(1-s)^{D_v}\right] &= \sum_{v\in V\setminus v^*}\left(\sum_{d=0}^{n-1}\mathbb{P}\left(D_v=d\right)\prod_{\ell=2}^{m}d^{r_\ell}(1-s)^d\right)\\
&= \sum_{d=0}^{n-1}\sum_{v\in V\setminus v^*}\mathbb{P}\left(D_v=d\right)d^{\sum_{\ell=2}^m r_\ell}(1-s)^{(m-1)d}.
\end{aligned}
\tag{27}
$$

Proceed by splitting the outer summation in (27) over d at $d^*\triangleq n^{\frac{\alpha}{2(m-1)r}}$. We have

$$
\begin{aligned}
\sum_{d=0}^{d^*}\sum_{v\in V\setminus v^*}\mathbb{P}\left(D_v=d\right)\cdot d^{\sum_{\ell=2}^m r_\ell}\cdot(1-s)^{(m-1)d} &\le \sum_{v\in V\setminus v^*}(d^*)^{\sum_{\ell=2}^m r_\ell}\sum_{d=0}^{d^*}\mathbb{P}\left(D_v=d\right)\cdot(1-s)^{(m-1)d}\\
&\le (d^*)^{\sum_{\ell=2}^m r_\ell}\sum_{v\in V\setminus v^*}\mathbb{E}_{D_v}\left[(1-s)^{(m-1)D_v}\right]\\
&\overset{\text{(c)}}{=} (d^*)^{\sum_{\ell=2}^m r_\ell}\sum_{v\in V\setminus v^*}\prod_{u\ne v}\left(1-p_{uv}s+p_{uv}s(1-s)^{m-1}\right)\\
&\overset{\text{(d)}}{\le} (d^*)^{\sum_{\ell=2}^m r_\ell}\sum_{v\in V\setminus v^*}e^{-\sum_{u\ne v}p_{uv}s\left(1-(1-s)^{m-1}\right)}\\
&\overset{\text{(e)}}{\le} n^{\alpha/2}\cdot\sum_{v\in V\setminus v^*}e^{-d_v s\left(1-(1-s)^{m-1}\right)} = o(1).
\end{aligned}
\tag{28}
$$

Here, (c) uses the moment generating function of D_v, which is distributed as a sum of independent Bernoulli random variables with means p_{uv} respectively. Further, (d) uses that $1-x\le e^{-x}$ for any x, and finally (e) uses the definition of d^* and the assumed condition

$$
\sum_{v\in V\setminus v^*}e^{-d_v s\left(1-(1-s)^{m-1}\right)} = o(n^{-\alpha}).
$$

The other part of the split sum can be bounded as

$$\begin{aligned}
&\sum_{d=d^*}^{n-1} \sum_{v \in V \setminus v^*} \mathbb{P}\left(D_v = d\right) \cdot d^{\sum_{\ell=2}^m r_\ell} \cdot (1-s)^{(m-1)d} \\
&\leq \sum_{v \in V \setminus v^*} \left[\max_{d \in [d^*, n]} d^{\sum_{\ell=2}^m r_\ell} (1-s)^{(m-1)d} \right] \mathbb{P}\left(D_v \geq d^*\right) \\
&\overset{\text{(f)}}{\leq} (d^*)^{\sum_{\ell=2}^m r_\ell} (1-s)^{(m-1)d^*} \cdot \sum_{v \in V \setminus v^*} \mathbb{P}\left(D_v \geq d^*\right) \\
&\overset{\text{(g)}}{\leq} n^{1+\alpha/2} \cdot \left((1-s)^{m-1}\right)^{n^{\frac{\alpha}{2(m-1)r}}} = o(1), \qquad (29)
\end{aligned}$$

whenever $s > 0$. Here, (f) uses that the function $d \mapsto d^{\sum_{\ell=2}^m r_\ell}(1-s)^{(m-1)d}$ is decreasing on the interval $[d^*, n]$ for all sufficiently large n. Finally, (g) uses the definition of d^* and that $\mathbb{P}\left(D_v \geq d^*\right) \leq 1$ for all v. Plugging (28) and (29) in (27) yields that $\sum_{v \in V \setminus v^*} \mathbb{E}_{D_v}\left[\prod_{\ell=2}^m D_v^{r_\ell}(1-s)^{D_v}\right] = o(1)$. Therefore, continuing from (26),

$$\begin{aligned}
(26) &= \sum_{r_2=0}^{r} \cdots \sum_{r_m=0}^{r} \left[\prod_{\ell=2}^m \left(\frac{e}{r_\ell} \cdot \frac{s}{1-s}\right)^{r_\ell}\right] \cdot o(1) \\
&\leq (r+1)^{m-1} \max_{1 \leq r' \leq r} \left(\frac{es}{r'(1-s)}\right)^{r'(m-1)} \cdot o(1) \\
&= o(1),
\end{aligned}$$

since r and m are both constants independent of n. This completes the proof. □

References

Cullina D, Kiyavash N (2016) Improved achievability and converse bounds for Erdős-Rényi graph matching. *ACM SIGMETRICS performance evaluation review* 44(1):63–72.

Mitzenmacher M, Upfal E (2017) *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis* (Cambridge University Press).

Wu Y, Xu J, Yu SH (2022) Settling the sharp reconstruction thresholds of random graph matching. *IEEE Transactions on Information Theory* 68(8):5391–5417.