A detection threshold for correlated multi-graphs

Daniel Ochoa

February 22, 2026

Abstract

We study *correlation detection* across multiple sparse graphs. Under the null, we observe m independent Erdős–Rényi layers $G_i \sim \mathcal{G}(n, ps)$ with $p = n^{-\alpha+o(1)}$ for $\alpha \in (0, 1]$. Under the alternative, layers are generated from a latent parent $G_0 \sim \mathcal{G}(n, p)$ by independently keeping each parent edge with probability s. For $m = 2$, Ding and Du [9] established a sharp threshold for detection. We extend the detection threshold to any number of layers.

1 Introduction

Networks often come in multiple layers: we may observe several graphs on a common vertex set, with edges in different layers exhibiting latent dependence. This paper studies *correlation detection* in such multilayer settings, in the sparse asymptotic regime. Concretely, we observe m layers $G = (G_1, \ldots, G_m)$, each on the same n labeled vertices. Under the alternative, they are edge–correlated noisy copies of a single latent parent graph, each subjected to an unknown vertex permutation. Our goal is to decide, from the observed layers, which hypothesis generated the data.

Models. We fix $p = p(n) = n^{-\alpha+o(1)}$ with $\alpha \in (0, 1]$ and a copying parameter $s \in (0, 1)$.

H_0 : (*independence*): $G_1, \ldots, G_m \overset{\text{i.i.d.}}{\sim} \mathcal{G}(n, ps)$, independent across layers.

H_1 : (*correlated*): sample a parent $G_0 \sim \mathcal{G}(n, p)$. Let $V(G_0)$ and $E(G_0)$ denote the vertex and edge sets of G_0. Draw permutations $\Pi_1^*, \ldots, \Pi_m^* \in S_n$ (where $\Pi_1^* = \text{id}$ for simplicity). Conditional on G_0 and $\Pi^* := (\Pi_1^*, \ldots, \Pi_m^*)$, generate the layers by keeping each parent edge independently with probability s:

$$\mathbb{P}\big[(u,v) \in E(G_i) | G_0, \Pi^*\big] = \begin{cases} s \text{ if } ({\Pi_i^*}^{-1}(u), {\Pi_i^*}^{-1}(v)) \in E(G_0) \\ 0 \text{ if } ({\Pi_i^*}^{-1}(u), {\Pi_i^*}^{-1}(v)) \notin E(G_0) \end{cases}$$

independently over $i \in [m]$ and unordered pairs (u, v). Marginally, each $G_i \sim \mathcal{G}(n, ps)$.

Problem statement. Our goal is to test H_0 versus H_1 given the observed layers $G = (G_1, \ldots, G_m)$, while the latent alignment $\Pi^* = (\mathrm{id}, \Pi_2^*, \ldots, \Pi_m^*)$ remains unknown. It is standard that the optimal testing error is governed by the total variation distance between the null and alternative laws. Our main contribution is to give a strong detection threshold in the sparse regime.

1.1 Background and related work

The graph alignment problem has received sustained interest over the last decade, with applications in network de-anonymization [14], cross-species component matching via protein–protein interactions [3], and textual inference in natural language processing [12]. For the two-layer case $m = 2$, there has been substantial progress on the alignment problem [17, 18, 6, 5, 11, 19, 15, 7, 4]. Cullina and Kiyavash [5, 6] established sharp thresholds for exact alignment, and Cullina [7] obtained a sharp threshold for almost-exact alignment. The exact threshold for partial alignment was resolved by Ding and Du [8], who connected partial recovery to a correlation-detection test based on densest subgraphs [9]. In multigraph settings ($m > 2$), progress is more recent: Ameen and Hajek [1] proved a sharp threshold for exact recovery, and Massoulié–Vassaux [16] established an impossibility threshold for partial alignment.

For $m = 2$, Ding and Du [9] established a sharp threshold in the sparse regime: correlation is detectable iff nps^2 exceeds $\varrho^{-1}(\frac{1}{\alpha})$, where $\varrho(\cdot)$ is the limiting densest–subgraph ratio for $\mathcal{G}(n, \frac{\lambda}{n})$ (see Proposition 1.1). Informally, above this threshold the correctly aligned intersection $G_1 \cap G_2^{{\Pi_2^*}^{-1}}$ contains a linear vertex set with ratio at least $\frac{1}{\alpha}$; under independence, with high probability no permutation Π produces such a witness.

Proposition 1.1. *([2], Theorem 1, Theorem 3) For any constant $\lambda > 0$, there exists a constant $\varrho(\lambda) > 0$ which can be explicitly written via a variational characterization, such that for any Erdős–Rényi graph $G \sim \mathcal{G}(n, \frac{\lambda}{n})$,*

$$\max_{\emptyset \neq U \subset V} \frac{|\mathcal{E}(U)|}{|U|} \;\to\; \varrho(\lambda) \quad \textit{in probability as } n \to \infty, \tag{1}$$

where $\mathcal{E}(U)$ is the collection of edges in $\mathcal{E}(G)$ with both endpoints in U. Moreover, $\varrho(\cdot)$ is continuous, increasing, and satisfies

$$1 \le \frac{\varrho(\beta)}{\varrho(\alpha)} \le \frac{\beta}{\alpha}, \quad \forall\, 0 < \alpha < \beta, \tag{2}$$

We call the largest subgraph that maximizes the left-hand side of (1)(if there are many such subgraphs, pick one of them arbitrarily) the *densest subgraph*.

Proposition 1.2. *([9], Proposition 2.3) For $\lambda > 1$, $\varrho(\lambda) > 1$ and ϱ is strictly increasing. Furthermore, there exists some constant $c_\lambda > 0$, such that with probability tending to 1 as $n \to \infty$, the size of the densest subgraph in an Erdős–Rényi graph $\mathcal{H} \sim \mathbb{G}(n, \frac{\lambda}{n})$ is at least $c_\lambda n$.*

1.2 Our contribution

We extend this picture to *any* number of layers $m \ge 2$. In particular, we quantify how additional layers amplify the effective signal via the union of all pairwise intersections.

Define $q_m(s) := 1 - (1-s)^m - ms(1-s)^{m-1}$. This is the probability that a parent edge appears in at least two observed layers, which is obtained by counting the complement: $(1-s)^m$ is the probability a parent edge does not appear in any graph and $ms(1-s)^{m-1}$ is the probability a parent edge appears in exactly one graph. Let P and Q denote the laws of the observed layers under H_0 and H_1, respectively.

Theorem 1.3. *Suppose $p = p(n)$ satisfies $p = n^{-\alpha+o(1)}$ for some $\alpha \in (0,1]$ as $n \to \infty$. Let $\lambda_* = \varrho^{-1}(\frac{m-1}{\alpha})$, then for any constant $\varepsilon > 0$, the following holds. Suppose there exists a value λ_{**} such that s satisfies $npq_m(s) \ge \lambda_* + \varepsilon$ and $nps^2 \le \lambda_{**}$, then*

$$\mathrm{TV}(P, Q) = 1 - o(1) \quad \text{as} \quad n \to \infty$$

where $\mathrm{TV}(P,Q) = \frac{1}{2}\sum_\omega |P[\omega] - Q[\omega]|$ is the total variation distance between P and Q.

In other words, above this threshold detection is possible: in the correlated model, the permuted pairwise-intersection union

$$\bigcup_{i<j} G_i^{\Pi_i^{*-1}} \cap G_j^{\Pi_j^{*-1}}$$

contains (with high probability) a linear-size subgraph whose ratio exceeds $\frac{m-1}{\alpha}$. Under independence, with high probability there is *no* choice of permutation Π for which the same union contains such a witness.

2 Correlation Detection

We next define our test statistic. For any permutations $\Pi_2, \dots, \Pi_m \in S_n$ and $\Pi = (\mathrm{id}, \Pi_2, \dots, \Pi_m) \in S_n^{m-1}$, let

$$G_i^{\Pi_i} = (V(G), E(G_i^{\Pi_i})), \text{ where } (\Pi_i(u), \Pi_i(v)) \in E(G_i^{\Pi_i}) \text{ iff } (u,v) \in E(G_i)$$

In other words, the graph $G_i^{\Pi_i}$ is the graph G_i after permuting its vertices. We define $\mathcal{H}_\Pi$ as

$$\mathcal{H}_\Pi = (V(G), \mathcal{E}_\Pi), \text{ where } (u,v) \in \mathcal{H}_\Pi \text{ iff } (u,v) \in G_i^{\Pi_i} \cap G_j^{\Pi_j} \text{ for a pair } i < j$$

the union of all pairwise intersection graphs after applying the permutation Π. This extends the test statistic done by Ding and Du. By using the condition $npq_m(s) \geq \lambda_* + \varepsilon$, we force a linear dense subgraph on the correlated case that won't happen w.h.p. on the independent case.

Let $\mathcal{T}(G)$ be the test

$$\mathcal{T}(G) \triangleq \max_{\Pi} \max_{U \subset \mathcal{H}_\Pi : |U| \geq n/\log(n)} \frac{|\mathcal{E}_\Pi(U)|}{|U|}$$

Notice that for the case $m = 2$, this reduces to the densest subgraph on $G_1 \cap G_2^{\Pi_2}$ across the permutations Π_2.

Theorem 2.1. *Suppose $p = p(n)$ satisfies $p = n^{-\alpha + o(1)}$ for some $\alpha \in (0,1]$ as $n \to \infty$. Let $\lambda_* = \varrho^{-1}(\frac{m-1}{\alpha})$, then for any constant $\varepsilon > 0$, the following holds. Suppose there exists a value λ_{**} such that s satisfies $npq_m(s) \geq \lambda_* + \varepsilon$ and $nps^2 \leq \lambda_{**}$, let $\tau = \frac{\varrho(\lambda_*) + \varrho(\lambda_* + \varepsilon)}{2}$, then*

$$Q[\mathcal{T}(G) < \tau] + P[\mathcal{T}(G) \geq \tau] = o(1)$$

Lemma 2.2. *Theorem 2.1 implies Theorem 1.3.*

Proof. Conditioning on the event $\{\mathcal{T} \geq \tau\}$ we have

$$
\begin{aligned}
\mathrm{TV}(P,Q) &= \frac{1}{2}\sum_{w\in\{\tau\geq\mathcal{T}\}} |P(w)-Q(w)| + \frac{1}{2}\sum_{w\in\{\tau<\mathcal{T}\}} |P(w)-Q(w)| \\
&\leq \frac{1}{2}\sum_{w\in\{\tau\geq\mathcal{T}\}} \Big[Q(w)-P(w)\Big] + \frac{1}{2}\sum_{w\in\{\tau<\mathcal{T}\}} \Big[P(w)-Q(w)\Big] \\
&= 1-o(1)
\end{aligned}
$$

□

Let $\lambda = nps^2$, by Proposition 1.1

$$
\frac{m-1}{\alpha} = \varrho(\lambda_*) < \tau < \varrho(\lambda_* + \varepsilon) < \varrho(npq_m(s))
$$

Correlated case: Under H_1, we have that $\mathcal{H}_{\Pi^{*-1}}$ is an Erdős–Rényi $\mathcal{G}(n, \frac{npq_m(s)}{n})$. Since $\varrho(npq_m(s)) > \tau$, by Proposition 1.1 and Proposition 1.2 we have that with probability $1-o(1)$, $\mathcal{H}_{\Pi^{*-1}}$ has a giant dense subgraph with ratio at least τ. Therefore, $Q[\mathcal{T}(G) < \tau] = o(1)$.

Independent case:

The outline of the proof is to divide the edges $\mathcal{E}_\Pi(U)$ of the densest linear subgraph U of $\mathcal{H}_\Pi$ into $\frac{m(m-1)}{2}$ disjoint parts

$$
\mathcal{E}_\Pi(U) = \bigsqcup_{1\leq i<j\leq m} \mathcal{E}_\Pi^{i,j}(U) \tag{3}
$$

$$
\mathcal{E}_\Pi^{i,j}(U) = \left\{ (u,v) \;\middle|\; \begin{array}{l} (u,v) \in E(G_i^{\Pi_i}) \cap E(G_j^{\Pi_j}), \\ (u,v) \notin E(G_{i'}^{\Pi_{i'}}) \cap E(G_{j'}^{\Pi_{j'}}) \text{ for } (i',j') \prec (i,j) \end{array} \right\}, \tag{4}
$$

with $(i',j') \prec (i,j)$ the lexicographic order. Then bound the probability that each $|\mathcal{E}_\Pi^{i,j}(U)|/|U|$ is above certain ratio (for this we use disjointness). After that, we just union bound across all possible $\binom{n}{|U|}^m \cdot |U|!^{m-1}$ intersections of size $|U|$. For this sake, we are going to use the following Chernoff bound for Binomial variables ([13], Theorem 4.4): For $X \sim \mathrm{Bin}(n,p)$, let $\mu = np$, we have:

$$
\mathbb{P}[X \geq (1+\delta)\mu] \leq \exp(-\mu[(1+\delta)\log(1+\delta) - \delta]) \tag{CB}
$$

Lemma 2.3. *For any $\epsilon > 0$, there exists a positive integer $l = l(\epsilon, m)$ such that any permutation Π where $\mathcal{H}_\Pi$ has a giant dense subgraph U of* *ratio larger than* τ *also satisfies the following: there exists an assignment* $t_{\Pi}^{i,j} \in \{0, 1, \ldots \lceil l\tau \rceil\}$ *such that if we divide the edges of* $\mathcal{H}_{\Pi}$ *into* $\frac{m(m-1)}{2}$ *parts as explained in* (3), (4)*:*

1. ***Loss of ratio at most*** ϵ*:* $\sum_{i<j} \frac{t_{\Pi}^{i,j}}{l} \geq \tau - \epsilon$ *(5)*
2. ***Not increase of previous ratio:*** $\frac{|\mathcal{E}_{\Pi}^{i,j}(U)|}{|U|} \geq \frac{t_{\Pi}^{i,j}}{l}$ *for* $i < j$
3. ***Above-mean for Chernoff bounds****:* $|U| \cdot \frac{t_{\Pi}^{i,j}}{l} \geq \binom{|U|}{2} p^2 s^2 \geq \mathbb{E}[\mathcal{E}_{\Pi}^{i,j}(U)]$ *for* $i < j$

If we define $r_{\Pi}^{i,j} = \frac{|\mathcal{E}_{\Pi}^{i,j}(U)|}{|U|}$, the lemma says that we can round down the ratios for each $r_{\Pi}^{i,j}$ to the nearest fraction of the form $\frac{t}{l}$ without losing too much information. The third condition is just to safely use the Chernoff bound for $\mathbb{P}[\text{Bin}()]$ in the calculations below.

Proof. Since $\mathcal{H}_{\Pi}$ has a giant subgraph of ratio at least τ that means that:

$$\sum_{i<j} r_{\Pi}^{i,j} \geq \tau$$

Take $t_{\Pi}^{i,j} = \lfloor l r_{\Pi}^{i,j} \rfloor$ if $|U| \cdot \frac{\lfloor l r_{\Pi}^{i,j} \rfloor}{l} > \binom{|U|}{2} p^2 s^2$, otherwise $t_{\Pi}^{i,j} = 0$. We claim this is the assignment we are looking for. Notice we have

$$\sum_{i<j} r_{\Pi}^{i,j} - \sum_{i<j} \frac{t_{\Pi}^{i,j}}{l} = \sum_{i<j | t_{\Pi}^{i,j} \neq 0} \left(r_{\Pi}^{i,j} - \frac{\lfloor l r_{\Pi}^{i,j} \rfloor}{l} \right) + \sum_{i<j | t_{\Pi}^{i,j} = 0} r_{\Pi}^{i,j}$$

In particular, $r_{\Pi}^{i,j} - \frac{\lfloor l r_{\Pi}^{i,j} \rfloor}{l} \leq \frac{1}{l}$. If $t_{\Pi}^{i,j} = 0$ we have $r_{\Pi}^{i,j} - \frac{\lfloor l r_{\Pi}^{i,j} \rfloor}{l} \leq \max(\frac{1}{l}, \binom{|U|}{2} p^2 s^2 / |U|) = \frac{1}{l}$. The reason is that $\binom{|U|}{2} p^2 s^2 / |U| = o(n^{-\alpha + o(1)}) < \frac{1}{l}$. Therefore

$$\sum_{i<j} \frac{t_{\Pi}^{i,j}}{l} \geq \sum_{i<j} r_{\Pi}^{i,j} - \frac{m(m-1)}{2} \cdot \frac{1}{l} \geq \tau - \epsilon$$

for $l > \lceil \frac{m(m-1)}{2\epsilon} \rceil$. □

By Lemma 2.3 we can bound the probability of existing a Π that creates the giant dense subgraph in $\mathcal{H}_{\Pi}$ by a union bound across all possible assignments of $t_{\Pi}^{i,j}$ satisfying the conditions in Lemma 2.3 and the probability of existing a Π that gives the giant dense subgraph across each of those assignments.

Lemma 2.4. *For $0 < \epsilon \ll \tau - \varrho(\lambda_*)$. For any assignment $0 \le t^{i,j} \le \lceil l\tau \rceil$, $i < j$, satisfying (5), there exists $\nu = \nu(m, \tau, \alpha) > 0$ such that*

$$P\left[\begin{array}{c} \exists\, \Pi \in S_n^{m-1},\ \exists\, U \subseteq \mathcal{H}_\Pi,\ |U| \ge n/\log n \\ \dfrac{|\mathcal{E}_\Pi^{i,j}(U)|}{|U|} \ \ge\ \dfrac{t^{i,j}}{l} \ \textit{ for all } i<j, \quad \textit{and} \quad \dfrac{|\mathcal{E}_\Pi(U)|}{|U|}\ \ge\ \tau \end{array}\right] = o(n^{-\nu}) \tag{6}$$

where $\mathcal{E}_\Pi^{i,j}$ are defined as in (3), (4) *and $l = l(\epsilon, m)$ is the integer found in Lemma 2.3*

Proof. We can upper bound (6) by

$$\sum_{k \ge n/\log n} \Pr\Big[\exists\, U \subseteq V(\mathcal{H}_\Pi),\ |U| = k:\ \frac{|\mathcal{E}_\Pi^{i,j}(U)|}{|U|}\ \ge\ \frac{t^{i,j}}{l}\ \text{ for all } i<j\Big]. \tag{7}$$

By a union bound over candidates Π with U a random subset of size k, we upper bound (7) by

$$\sum_{k \ge n/\log n} \binom{n}{k}^m (k!)^{m-1} \Pr\Big[\frac{|\mathcal{E}_\Pi^{i,j}(U)|}{|U|}\ \ge\ \frac{t^{i,j}}{l}\ \text{ for all } i<j\Big], \tag{8}$$

since there are $\binom{n}{k}$ possible ways to choose the k vertices for each graph and $k!$ ways to permute them across each graph (except the first one). Moreover, once U is fixed, since the variables $|\mathcal{E}_\Pi^{i,j}(U)|$ for $i < j$ are negatively associated we get

$$\Pr\Big[\frac{|\mathcal{E}_\Pi^{i,j}(U)|}{|U|}\ \ge\ \frac{t^{i,j}}{l}\ \text{ for all } i<j\Big]\ \le \prod_{1\le i<j\le m} \Pr\Big[\frac{|\mathcal{E}_\Pi^{i,j}(U)|}{|U|}\ \ge\ \frac{t^{i,j}}{l}\Big]. \tag{9}$$

(see Appendix A for more details). However, $\mathcal{E}_\Pi^{i,j}(U)$ is a binomial random variable which is bounded by.

$$\mathbb{P}\Big[\frac{|\mathcal{E}_\Pi^{i,j}(U)|}{k}\ \ge\ \frac{t^{i,j}}{l}\Big] \le \mathbb{P}\Big[\,\mathrm{Bin}\Big(\frac{k(k-1)}{2}, (ps)^2\Big) \ge \frac{t^{i,j}}{l}k\Big]$$

as the probability of a specific edge appearing in $\mathcal{E}_\Pi^{i,j}(U)$ is upper-bounded by the probability of the same edge appearing in $G_i^{\Pi_i} \cap G_j^{\Pi_j}$. Therefore, we have (8) is bounded by:

$$\begin{aligned}
&\sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \prod_{i<j} \mathbb{P}[\mathrm{Bin}(\frac{k(k-1)}{2},(ps)^2) \geq \frac{t^{i,j}}{l}k] \\
&\overset{\text{(CB)}}{\leq} \sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \prod_{i<j} \exp(-\frac{t^{i,j}}{l}k\log(\frac{2\frac{t^{i,j}}{l}}{kp^2s^2}) + \frac{t^{i,j}}{l}k) \\
&= \sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \prod_{i<j} \exp(-\frac{t^{i,j}}{l}k\log(\frac{2\frac{t^{i,j}}{l}n}{\lambda kp}) + \frac{t^{i,j}}{l}k) \\
&\leq \sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \prod_{i<j} \exp(-\frac{t^{i,j}}{l}\alpha k\log(n) + o(k\log(n)))
\end{aligned}$$

where the last inequality came from $\log(\frac{2\frac{t^{i,j}}{l}n}{\lambda kp}) = \log(\frac{n}{kp}) + o(1) = (\alpha + o(1))\log(n) + o(1)$. Therefore, (6) is bounded by

$$\begin{aligned}
&\sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \exp(-\sum_{i<j}\frac{t^{i,j}}{l}\alpha k\log(n) + o(k\log(n))) \\
&\leq \sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \exp(-(\tau-\epsilon)\alpha k\log(n) + o(k\log(n)))
\end{aligned}$$

By taking ϵ such that $\tau - \epsilon > \frac{m-1}{\alpha} + 2\varsigma$ for a fixed $\varsigma = \varsigma(\alpha,\tau,m)$ we obtain the bound on (6):

$$\begin{aligned}
&\sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \exp(-(\frac{m-1}{\alpha} + 2\varsigma)\alpha k\log(n) + o(k\log(n))) \\
&\leq \sum_{k\geq n/\log(n)} \binom{n}{k}^m k!^{m-1} \exp(-(\frac{m-1}{\alpha} + \varsigma)\alpha k\log(n)) \\
&\leq \sum_{k\geq n/\log(n)} \binom{n}{k} n^{k(m-1)} \exp(-(m-1+\varsigma_*)k\log(n)) \\
&= \sum_{k\geq n/\log(n)} \binom{n}{k} n^{-k\varsigma_*} \leq \sum_{k\geq n/\log(n)} (2e)^n n^{-k\varsigma_*} \leq n(2e)^n n^{-n\varsigma_*/\log(n)} = o(n^{-\nu})
\end{aligned}$$

where $\varsigma_* = \varsigma\alpha$. □

Proof of Theorem 2.1. Now we are ready to bound the probability under the null, such that there exists a Π with a giant, dense subgraph in $\mathcal{H}_\Pi$. By Lemma 2.3 we can divide the probability into at most $\lceil l(\epsilon, m)\tau \rceil^{\frac{m(m-1)}{2}}$ possible scenarios, and by Lemma 2.4 each of those scenarios has $o(n^{-\nu})$ probability. Therefore the total probability is bounded by $\lceil l\tau \rceil^{\frac{m(m-1)}{2}} o(n^{-\nu}) = o(1)$. Therefore, $P[\mathcal{T}(G) \geq \tau] = o(1)$.

This proves $Q[\mathcal{T}(G) < \tau] + P[\mathcal{T}(G) \geq \tau] = o(1)$ as required. □

2.1 Implications

Recall

$$q_m(s) \;=\; 1 - (1-s)^m - ms(1-s)^{m-1}.$$

A Taylor expansion at $s = 0$ gives

$$q_m(s) \;=\; \binom{m}{2} s^2 \;-\; 2\binom{m}{3} s^3 \;+\; O(s^4) \;=\; \binom{m}{2} s^2 \big(1 - O(ms)\big). \tag{10}$$

Hence for $s = o(1/m)$ we have

$$npq_m(s) \;=\; \binom{m}{2} nps^2 \,(1 + o(1)).$$

Therefore the detection condition

$$npq_m(s) \;>\; \varrho^{-1}\big(\tfrac{m-1}{\alpha}\big)$$

is asymptotically equivalent to

$$nps^2 \;>\; \frac{2}{m(m-1)}\, \varrho^{-1}\big(\tfrac{m-1}{\alpha}\big) \quad \text{as } s \to 0. \tag{11}$$

A convenient bound comes from the trivial inequality $\varrho(\lambda) \geq \frac{1}{n}\,|E(G)| = \frac{\lambda}{2}$ (the densest–subgraph ratio dominates the global average), which implies $\varrho^{-1}(x) \leq 2x$. Plugging this into (11) yields, for the sparse regime $\alpha = 1$,

$$nps^2 \;\geq\; \frac{4}{m}$$

an enough threshold for correlation detection. Thus, compared with the two-layer case (which requires a constant order signal level), the m-layer threshold improves the constant by a factor on the order of $1/m$. In particular, for $m \geq 4$ the required constant is strictly below 1.

Finally, recall that in the two-layer setting the sharp detection threshold was key for establishing the partial-alignment threshold [8]. By the same logic, the present m-layer condition

$$npq_m(s) \; > \; \varrho^{-1}\big(\tfrac{m-1}{\alpha}\big)$$

suggests partial alignment should be feasible throughout this region as well.

3 Acknowledgments

This project was carried out as part of the MIT Department of Mathematics UROP+ program. I thank Henry Hu for the regular meetings, suggestions about ideas to try and how to tackle and understand the problem. The problem was proposed by Nike Sun. Support from the UROP+ program is gratefully acknowledged.

References

[1] Taha Ameen and Bruce Hajek. Aligning multiple inhomogeneous random graphs: Fundamental limits of exact recovery. *arXiv preprint arXiv:2405.12293*, 2025. Updated on June 29, 2025.

[2] Venkat Anantharam and Justin Salez. The densest subgraph problem in sparse random graphs. *Annals of Applied Probability*, 26(1):305–327, 2016. arXiv preprint, updated version as of 7 Jan 2016.

[3] Sourav Bandyopadhyay, Roded Sharan, and Trey Ideker. Systematic identification of functional orthologs based on protein network comparison. *Genome Research*, 16(3):428–435, 2006.

[4] Boaz Barak, Chi-Ning Chou, Zhixian Lei, Tselil Schramm, and Yueqi Sheng. (nearly) efficient algorithms for the graph matching problem on correlated random graphs, 2019.

[5] Daniel Cullina and Negar Kiyavash. Improved achievability and converse bounds for Erdős–Rényi graph matching. *ACM SIGMETRICS Performance Evaluation Review*, 44(1):63–72, 2016.

[6] Daniel Cullina and Negar Kiyavash. Exact alignment recovery for correlated Erdős–Rényi graphs, 2018.

[7] Daniel Cullina, Negar Kiyavash, Prateek Mittal, and H. Vincent Poor. Partial recovery of Erdős–Rényi graph alignment via k-core alignment, 2018.

[8] Jian Ding and Hang Du. Detection threshold for correlated Erdős–Rényi graphs via densest subgraphs, 2022.

[9] Jian Ding and Hang Du. Matching recovery threshold for correlated random graphs, 2022.

[10] Devdatt Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *BRICS Report Series*, RS-96(25):1–30, 1996. BRICS Report Series, Basic Research in Computer Science, University of Aarhus.

[11] Luca Ganassali, Marc Lelarge, and Laurent Massoulié. Impossibility of partial recovery in the graph alignment problem. In Mikhail Belkin and Samory Kpotufe, editors, *Proceedings of the 34th Annual Conference*

on Learning Theory, volume 134 of *Proceedings of Machine Learning Research*, pages 1–23. PMLR, 2021.

[12] Aria Haghighi, Andrew Y. Ng, and Christopher D. Manning. Robust textual inference via graph matching. In *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, pages 387–394. Association for Computational Linguistics, 2005.

[13] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, Cambridge, 2005. 1st ed.

[14] Arvind Narayanan and Vitaly Shmatikov. De-anonymizing social networks. In *2009 30th IEEE Symposium on Security and Privacy*, pages 173–187. IEEE, 2009.

[15] Miklós Z. Rácz and Anirudh Sridhar. Matching correlated inhomogeneous random graphs using the k-core estimator, 2023.

[16] Louis Vassaux and Laurent Massoulié. The feasibility of multi-graph alignment: a bayesian approach, 2025.

[17] Yihong Wu, Jiaming Xu, and Sophie H. Yu. Testing correlation of unlabeled random graphs, 2021.

[18] Yihong Wu, Jiaming Xu, and Sophie H. Yu. Settling the sharp reconstruction thresholds of random graph matching, 2022.

[19] Junchi Yan, Xu-Cheng Yin, Weiyao Lin, Cheng Deng, Hongyuan Zha, and Xiaokang Yang. A short survey of recent advances in graph matching. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval (ICMR '16)*, pages 167–174, New York, NY, USA, 2016. Association for Computing Machinery.

A Appendix

First, let's call some definitions and propositions from [10].

Definition A.1 (**Negative association**, [10], Definition 1)**.** *Let* $\mathbf{X} := (X_1, \ldots, X_n)$ *be a vector of random variables.*

1. $(-A)$ *The random variables,* $\mathbf{X}$ *are **negatively associated** if for every two disjoint index sets,* $I, J \subseteq [n]$*,*

$$\mathbb{E}\big[f(X_i,\, i \in I)\, g(X_j,\, j \in J)\big] \leq \mathbb{E}\big[f(X_i,\, i \in I)\big]\, \mathbb{E}\big[g(X_j,\, j \in J)\big]$$

for all functions $f : \mathbb{R}^{|I|} \to \mathbb{R}$ *and* $g : \mathbb{R}^{|J|} \to \mathbb{R}$ *that are both non-decreasing or both non-increasing.*

Proposition A.2. *([10], Proposition 7)*

1. *If* $\mathbf{X}$ *and* $\mathbf{Y}$ *satisfy* $(-A)$ *and are mutually independent, then the augmented vector*

$$(\mathbf{X}, \mathbf{Y}) = (X_1, \cdots, X_n, Y_1, \cdots, Y_m)$$

satisfies $(-A)$*.*

2. *Let* $\mathbf{X} := (X_1, \cdots, X_n)$ *satisfy* $(-A)$*. Let* $I_1, \cdots, I_k \subseteq [n]$ *be disjoint index sets, for some positive integer* k*. For* $j \in [k]$*, let*

$$h_j : \mathbb{R}^{|I_j|} \to \mathbb{R}$$

be functions that are all non-decreasing or all non-increasing, and define

$$Y_j := h_j(X_i,\, i \in I_j).$$

Then the vector $\mathbf{Y} := (Y_1, \cdots, Y_k)$ *also satisfies* $(-A)$*. That is, non-decreasing (or non-increasing) functions of disjoint subsets of negatively associated variables are also negatively associated.*

Lemma A.3 (Zero–One Lemma for $(-A)$, [10], Lemma 8)**.** *If* $X_1, \ldots, X_n$ *are zero-one random variables such that* $\sum_i X_i = 1$*, then* $X_1, \ldots, X_n$ *satisfy* $(-A)$*.*

Proposition A.4 (Marginal Probability Bounds,[10], Proposition 4)**.** *Let* $X_1, \ldots, X_n$ *satisfy* $(-A)$*. Then*

$$\Pr\big[X_i \geq t_i,\ i \in [n]\big] \leq \prod_{i \in [n]} \Pr\big[X_i \geq t_i\big].$$

With this in hand, we can prove (9)

Proof. Under H_0, each $G_i \sim \mathcal{G}(n, ps)$ is invariant under vertex relabeling, hence $(G_1^{\Pi_1}, \ldots, G_m^{\Pi_m}) \stackrel{d}{=} (G_1, \ldots, G_m)$. Therefore, we can assume without loss of generality that $\Pi = \mathrm{id}$. For each edge $e \in \mathcal{E}_\Pi(U)$ and $1 \leq i < j \leq m$, define

$$X_{i,j,e} = \mathbf{1}_{\{e \in \mathcal{E}_\Pi^{i,j}(U)\}},$$

where $\{\mathcal{E}_\Pi^{i,j}(U)\}_{i<j}$ are the disjoint edge sets defined by the lexicographic assignment rules 3 and 4. By construction, for each $e \in \mathcal{E}_\Pi(U)$, we get

$$\sum_{1 \leq i < j \leq m} X_{i,j,e} = \mathbf{1}_{\{e \in \mathcal{E}_\Pi(U)\}} = 1$$

Therefore, we have that $X_e = (X_{i,j,e}\ 1 \leq i < j \leq m)$ satisfies the Zero-One Lemma for (-A). Since $G_1, G_2, \ldots, G_m$ are independent, the sets $\{X_e\}_{e \in \mathcal{E}_\Pi(U)}$ are independent. Proposition A.2(1) implies that $(X_e)_{e \in \mathcal{E}_\Pi(U)}$ is negatively associated, and Proposition A.2(2) that

$$(|\mathcal{E}_\Pi^{i,j}(U)|)_{1 \leq i < j \leq m} = (\sum_{e \in \mathcal{E}_\Pi(U)} X_{i,j,e})_{1 \leq i < j \leq m}$$

is negatively associated, (by taking the disjoint index sets $I_{i,j} = (X_{i,j,e})_{e \in \mathcal{E}(U)}$ and the non-decreasing functions $h_{i,j} : \mathbb{R}^{|I_{i,j}|} \to \mathbb{R}$ defined by $h_{i,j}(v)$ equal the sum of all the coordinates of v). This gives us our desired bound by the *Marginal Probability Bounds.* □