Spectral Detection and Recovery of a Planted Biclique in a Random Bipartite Graph

Zachary Doll Simon Barras

December 14, 2025

Abstract

We study the problem of detecting and recovering a planted biclique in a random bipartite graph. The graph consists of two vertex parts of sizes n_1 and n_2 with i.i.d. Bernoulli(p) edges, into which we plant a complete biclique on unknown subsets $K_L \subseteq L$ and $K_R \subseteq R$ of sizes k_1 and k_2. We show that a simple spectral algorithm based on the leading singular vectors of the centered adjacency matrix succeeds in recovering the planted biclique with high probability whenever

$$(1-p)\sqrt{k_1 k_2} \gtrsim \sqrt{p(1-p)}\big(\sqrt{n_1} + \sqrt{n_2}\big),$$

which is the natural analogue of the Alon–Krivelevich–Sudakov threshold for planted clique in $G(n, 1/2)$. The proof proceeds via operator norm bounds for random rectangular matrices and a perturbation analysis of singular vectors. We also discuss extensions to partially dense planted bipartite subgraphs.

1 Introduction

The planted clique problem in $G(n, 1/2)$ is a central result in spectral graph theory and randomized algorithms. In its classical form, one first samples an Erdos-Renyi graph $G \sim G(n, 1/2)$, then "plants" a clique of size k by making all edges present within a subset $K \subseteq [n]$ of distinct vertices. The algorithmic challenge is then to recover K from the observed adjacency matrix.

Alon, Krivelevich, and Sudakov (AKS) introduced a spectral algorithm which recovers the planted clique in polynomial time when k is at least on the order of $\sqrt{n}$ [1].

In many applications, however, the natural graph structure is bipartite rather than unipartite. Examples include:

- user–item interaction graphs (e.g., accounts and posts),
- account–transaction graphs in fraud detection,
- IP addresses and destinations in network traffic.

In such settings, anomalous behavior often manifests as an unusually dense bipartite subgraph, or in the simplest case, as a biclique. This motivates a bipartite planted biclique model and the detection and recovery problems associated to it.

Our goal: We consider a random bipartite graph on vertex sets L and R of sizes n_1 and n_2 with i.i.d. Bernoulli(p) edges, in which we plant a complete biclique on unknown subsets $K_L \subseteq L$ and $K_R \subseteq R$. We show that:

- The centered ajacency matrix can be decomposed as a random noise matrix, plus a rank one signal supported on $K_L \times K_R$.
- Under a natural signal-to-noise condition, the principal singular vectors are correlated with the planted indicator vectors.

- A simple rounding step based on these singular vectors recovers the planted biclique up to $o(k_1)$ and $o(k_2)$ errors with high probability.

Our main theorem provides an explicit condition in terms of (n_1, n_2, k_1, k_2, p). In the symmetric case $n_1 = n_2 = n$ and $k_1 = k_2 = k$, the result reduces to a threshold $k \gtrsim \sqrt{n}$, in accordance to the AKS result.

2 Model and Problem Statement

2.1 Random bipartite graph

Let L and R be disjoint vertex sets of respective sizes $|L| = n_1$ and $|R| = n_2$. We represent a bipartite graph $G = (L \cup R, E)$ by its adjacency matrix $A \in \{0,1\}^{n_1 \times n_2}$, where

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \text{ with } i \in L, j \in R, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.1 (Null model)**.** *Under the null model, we assume*

$$A_{ij} \sim \mathit{Bernoulli}(p), \quad \textit{independently for all } i \in L, j \in R,$$

for some fixed $p \in (0,1)$. We denote this distribution by $G_{\mathrm{bip}}(n_1, n_2, p)$.

2.2 Planted biclique model

We now define the planted model of interest.

Definition 2.2 (Planted biclique model)**.** *Let $k_1 \le n_1$ and $k_2 \le n_2$. Pick subsets $K_L \subseteq L$ and $K_R \subseteq R$ of sizes $|K_L| = k_1$ and $|K_R| = k_2$, uniformly at random and independently.*

Conditional on (K_L, K_R), generate A as follows:

- *For all $(i,j) \in K_L \times K_R$, set $A_{ij} = 1$ deterministically, to plant a complete biclique.*
- *For all $(i,j) \notin K_L \times K_R$, sample $A_{ij} \overset{\text{i.i.d.}}{\sim} \mathit{Bernoulli}(p)$.*

We denote the resulting distribution by

$$A \sim G_{\mathrm{bip}}^{\mathrm{planted}}(n_1, n_2, p, k_1, k_2).$$

In what follows, we treat the subsets K_L and K_R as fixed, and analyze probabilities over the edge randomness of A, conditional on (K_L, K_R).

2.3 Detection and recovery tasks

We consider two related tasks.

Detection. Given A, we wish to test

$$H_0 : A \sim G_{\mathrm{bip}}(n_1, n_2, p) \quad \text{vs} \quad H_1 : A \sim G_{\mathrm{bip}}^{\mathrm{planted}}(n_1, n_2, p, k_1, k_2)$$

Recovery. Given A under H_1, we want to output estimated subsets $(\widehat{K}_L, \widehat{K}_R)$ with such that the overlaps fractions

$$\frac{|K_L \cap \widehat{K}_L|}{k_1}, \quad \frac{|K_R \cap \widehat{K}_R|}{k_2}$$

tend to 1 with high probability.

We note that the studied spectral algorithm addresses both tasks: recovery directly implies detection, by testing whether the recovered substructure is significantly denser than expected under the null model.

3 Spectral Decomposition of the Centered Adjacency Matrix

3.1 Decomposition into signal and noise

Let J denote the $n_1 \times n_2$ all-ones matrix, and consider the centered matrix B obtained by centering A:

$$B \coloneqq A - pJ.$$

Further define the indicator vectors $x^\star \in \{0,1\}^{n_1}$ and $y^\star \in \{0,1\}^{n_2}$ by

$$x_i^\star = \mathbb{1}\{i \in K_L\}, \qquad y_j^\star = \mathbb{1}\{j \in K_R\},$$

that we shall consider as fixed. We further decompose B as

$$B = N + S,$$

where:

- $N_{ij} \coloneqq A_{ij} - \mathbb{E}[A_{ij}]$ is a zero-mean noise term,
- $S_{ij} \coloneqq \mathbb{E}[A_{ij}] - p = (1-p)\mathbb{1}\{(i,j) \in K_L \times K_R\}$ is the signal.

Importantly, because we have

$$x^\star (y^\star)^T = \mathbb{1}\{(i,j) \in K_L \times K_R\},$$

it follows that $S = (1-p)x^\star (y^\star)^T$ is rank one.

3.2 Singular structure of the signal

Let

$$u^\star \coloneqq \frac{x^\star}{\sqrt{k_1}}, \qquad v^\star \coloneqq \frac{y^\star}{\sqrt{k_2}}.$$

Then $u^\star$ and $v^\star$ are unit vectors, and

$$S = (1-p)\sqrt{k_1 k_2}\, u^\star (v^\star)^T.$$

It follows that S has a unique non-trivial singular value

$$\sigma_{\text{sig}} = (1-p)\sqrt{k_1 k_2}$$

with left singular vector $u^\star$ and right singular vector $v^\star$.

3.3 Noise operator norm

The noise matrix N has independent entries with mean zero and variance

$$\mathrm{var}(N_{ij}) \le \mathrm{var}(A_{ij}) = p(1-p),$$

and $|N_{ij}| \le \max\{p, 1-p\} \le 1$.

We will use the following standard bound for the operator norms of a random rectangular matrix with idependent, zero-mean, sub-Gaussian entries.

Lemma 3.1 (Operator norm bound)**.** *There exist absolute constants $C, c > 0$ such that the following holds. Let $N \in \mathbb{R}^{n_1 \times n_2}$ have independent entries with mean zero and variance at most σ^2, and assume $|N_{ij}| \le 1$ almost surely. Then, for every $t \ge 0$,*

$$\mathbb{P}\left(\|N\|_{\mathrm{op}} \ge C\sigma\left(\sqrt{n_1} + \sqrt{n_2}\right) + t\right) \le 2\exp\left(-c\frac{t^2}{\sigma^2}\right).$$

In particular, with probability at least $1 - n^{-10}$ (for large enough n_1, n_2),

$$\|N\|_{\mathrm{op}} \le C'\sqrt{p(1-p)}\left(\sqrt{n_1} + \sqrt{n_2}\right)$$

for some absolute constant C'.

Proof. This follows from standard results on random matrices (e.g. a matrix Bernstein bound, followed by an ε-net argument to union bound over the unit sphere). We omit the full proof and refer to, for example, Vershynin's notes on non-asympotic random matrix theory [2]. □

Thus, with high probability, $B = N + S$ is a rank-one perturbation of a random noise matrix whose operator norm is of order

$$\|N\|_{\mathrm{op}} \asymp \sqrt{p(1-p)}(\sqrt{n_1} + \sqrt{n_2}).$$

4 Spectral Algorithm and Main Theorem

4.1 Spectral Algorithm and Main Theorem

We initially assume the parameters (p, k_1, k_2) to be known, and discuss the scenario in which they are unknown in Remark 4.6.

Algorithm 1 (Spectral biclique recovery).

1. *Centering.* Given $A \in \{0,1\}^{n_1 \times n_2}$, generated the centered matrix

$$B := A - pJ.$$

2. *SVD.* Compute the top singular triplet of B:

$$Bv_1 = \sigma_1 u_1, \qquad B^T u_1 = \sigma_1 v_1$$

where $\sigma_1 = \|B\|_{\mathrm{op}}$, and $\|u_1\| = \|v_1\| = 1$.

3. *Rounding on L.* Let $\widehat{K}_L$ be the set of indices corresponding to the k_1 largest entries of u_1 in absolute value.

4. *Rounding on R.* Let $\widehat{K}_R$ be the set of indices corresponding to the k_2 largest entries of v_1, again in absolute value.

The main challenge here is to show steps 3 and 4, namely that, under an appropriate signal-to-noise condition, the leading singular vectors (u_1, v_1) are sufficiently close to $(u^\star, v^\star)$ such that the above defined sets $(\widehat{K}_L, \widehat{K}_R)$ achieve high overlap with (K_L, K_R).

4.2 Correlation of singular vectors

We start by showing that the leading singular vectors are correlated with the planted indicators.

Theorem 4.1 (Singular vector correlation). *Let* $A \sim G_{\text{bip}}^{\text{planted}}(n_1, n_2, p, k_1, k_2)$*. Let* $B = A - pJ$*, and write* $B = N + S$*, as above. Assume that, for sufficiently large absolute constant* $C > 0$*,*

$$(1-p)\sqrt{k_1 k_2} \geq C\sqrt{p(1-p)}(\sqrt{n_1} + \sqrt{n_2}). \tag{1}$$

Let (σ_1, u_1, v_1) *be the top singular triplet of* B*. Then, with probability at least* $1 - n^{-10}$*,*

$$\langle u_1, u^\star \rangle^2 \geq 1 - \varepsilon, \qquad \langle v_1, v^\star \rangle^2 \geq 1 - \varepsilon,$$

where $\varepsilon = O\left(\left(\frac{\sqrt{p(1-p)}(\sqrt{n_1} + \sqrt{n_2})}{(1-p)\sqrt{k_1 k_2}}\right)^2\right)$ *tends to* 0 *under condition* (1)*.*

Proof sketch. We consider B as a perturbation of S be N. The matrix S has top singular value

$$\sigma_{\text{sig}} = (1-p)\sqrt{k_1 k_2},$$

with singular vectors $(u^\star, v^\star)$. By Lemma 3.1, with high probability

$$\|N\|_{\text{op}} \leq C_0\sqrt{p(1-p)}(\sqrt{n_1} + \sqrt{n_2}).$$

Under the signal-to-noise condition (1), we have

$$\|N\|_{\text{op}} \leq \frac{1}{2}\sigma_{\text{sig}}$$

for C large enough. We then apply a singular vector perturbation bound of the likes of Wedin's theorem [3] to the pair $(S, B = S + N)$. In particular, the Davis-Kahan-Wedin $\sin\Theta$ theorem implies that

$$\sin\angle(u_1, u^\star) \lesssim \frac{\|N\|_{\text{op}}}{\sigma_{\text{sig}}} = O\left(\frac{\sqrt{p(1-p)}(\sqrt{n_1} + \sqrt{n_2})}{(1-p)\sqrt{k_1 k_2}}\right),$$

and similarly for v_1 and $v^\star$. Finally, taking the square yields the claimed correlation bound

$$\langle u_1, u^\star \rangle^2 = \cos^2\angle(u_1, u^\star) = 1 - \sin^2\angle(u_1, u^\star) \geq 1 - \varepsilon.$$

□

4.3 From correlation to recovery

We now translate the ℓ_2-correlation into a statement about set recovery under the rounding step.

Lemma 4.2 (Entrywise seperation). *Under the assumption of Theorem 4.1, the following holds with high probability. There exists constants* $c_1, c_2 > 0$ *such that:*

- *For all but at most* $o(k_1)$ *indicies* $i \in K_L$*,*

$$u_1(i) \geq \frac{c_1}{\sqrt{k_1}},$$

- *For all but at most* $o(k_1)$ *indicies* $i \notin K_L$*,*

$$u_1(i) \leq \frac{c_2}{\sqrt{k_1}}$$

with $c_2 < c_1$*.*

An analogous statement holds for v_1 on the right side.

Proof sketch. Let us decompose u_1 as an orthogonal projection on $u^\star$:

$$u_1 = \alpha u^\star + z, \qquad z \perp u^\star,$$

with $\alpha \coloneqq \langle u_1, u^\star \rangle$. It follows directly from Theorem 4.1 that $\alpha^2 \geq 1 - \varepsilon$ and thus $\alpha \geq 1 - O(\varepsilon)$. Applying this same bound to the Pythagorean theorem also yields $\|z\|^2 \leq \varepsilon$. For $i \in K_L$, we have $u_i^\star = 1/\sqrt{k_1}$, so

$$u_1(i) = \frac{\alpha}{\sqrt{k_1}} + z_i \geq \frac{1 - O(\varepsilon)}{\sqrt{k_1}} - |z_i|.$$

Since z has squared norm at most ε, a standard first-moment bound implies that at most $o(k_1)$ of the coordinates z_i on K_L can be larger in magnitude than a threshold $t \coloneqq \varepsilon^{1/4}/\sqrt{k_1}$, chosen for convenience (see Appendix A.1). Thus, for all but $o(k_1)$ indicies in K_L,

$$u_1(i) \geq \frac{1 - O(\varepsilon)}{\sqrt{k_1}} - \frac{\varepsilon^{1/4}}{\sqrt{k_1}}.$$

We choose c_1 slightly under 1 and pick n_1, n_2 large enough such that ε is small.

For $i \notin K_L$, we have $u_i^\star = 0$ and hence $u_1(i) = z_i$. Since $\|z\|^2 \leq \varepsilon$, we similarly have at most $o(k_1)$ coordinates of z outside of K_L that may exceed a level such as $c_2/\sqrt{k_1}$ for a suitably small c_2. This yields the second inequality. An analogous argument holds for v_1. □

Theorem 4.3 (Recovery guarantee)**.** *Assume the conditions of Theorem 4.1. Let $(\widehat{K}_L, \widehat{K}_R)$ be the output of Algorithm* 1*. Then, with probability tending to* 1 *as* $\min(n_1, n_2) \to \infty$*,*

$$\frac{|K_L \cap \widehat{K}_L|}{k_1} \to 1, \qquad \frac{|K_R \cap \widehat{K}_R|}{k_2} \to 1.$$

In particular, the fraction of misclassified vertices on each side is $o(1)$.

Proof. By Lemma 4.2, the k_1 largest entries of u_1 include all but $o(k_1)$ vertices of K_L and at most $o(k_1)$ outside of K_L. Thus, $|\widehat{K}_L \Delta K_L| = o(k_1)$, which directly implies

$$\frac{|\widehat{K}_L \Delta K_L|}{k_1} \to 0 \iff \frac{|K_L \cap \widehat{K}_L|}{k_1} \to 1.$$

The same holds for $\widehat{K}_R$ and K_R. □

Corollary 4.4 (Detection)**.** *Under the assumptions of Theorem 4.3, there exists a polynomial-time test which distinguishes H_1 from H_0* w.h.p *as* $\min\{n_1, n_2\} \to \infty$*.*

Proof. Under H_1, Theorem 4.3 yields candidate sets $(\widehat{K}_L, \widehat{K}_R)$ of approximate sizes (k_1, k_2) and edge density tending to 1.

Under H_0, Lemma A.2 implies that, for some fixed $\delta > 0$, with high probability, no fixed pair of subsets of sizes (k_1, k_2) has density greater than $p + \delta$. Thus, it suffices to define a test which accepts H_1 if and only if

$$\mathrm{dens}(\widehat{K}_L, \widehat{K}_R) \geq p + \delta.$$

□

Remark 4.5 (Symmetric case)**.** *We note that in the symmteric case $n_1 = n_2 = n$ and $k_1 = k_2 = k$, condition* (1) *simplifies to*

$$(1-p)k \gtrsim \sqrt{p(1-p)}\, 2\sqrt{n},$$

or equivalently

$$k \gtrsim C'\sqrt{n}$$

for some constant C' depending on p. In particular, for $p = 1/2$, we obtain $k \gtrsim C''\sqrt{n}$, matching the classical AKS planted clique bound up to constants.

Remark 4.6 (Unknown parameters)**.** *If p is unknown, one may estimate it using the empirical edge density*

$$\hat{p} := \frac{1}{n_1 n_2} \sum_{i,j} A_{ij}$$

and proceed with $B := A - \hat{p}J$ instead. Standard concentration bounds imply $|\hat{p} - p| = o(1)$ (see Appendix A.3), hence this presents only minor perturbations in the above analysis. If (k_1, k_2) are unknown, one can either iterate over a range of candidate sizes. We do not pursue this extension in detail here.

5 Partially Dense Planted Bipartite Subgraphs

We briefly outline an extension to the particular case in which the planted subgraph is only partially dense.

Definition 5.1 (Partially dense planted model)**.** *Fix $p < q \leq 1$, and choose (K_L, K_R) as before. Conditional on (K_L, K_R), we generate A as:*

- *For $(i, j) \in K_L \times K_R$, draw $A_{ij} \sim$ Bernoulli(q).*
- *For $(i, j) \notin K_L \times K_R$, draw $A_{ij} \sim$ Bernoulli(p).*

In this case, the signal matrix has entries

$$S_{ij} = \mathbb{E}[A_{ij}] - p = \begin{cases} q - p & \text{if } i \in K_L, j \in K_R, \\ 0 & \text{otherwise,} \end{cases}$$

so that

$$S = (q - p)x^\star (y^\star)^T, \qquad \sigma_{\text{sig}} = (q - p)\sqrt{k_1 k_2}.$$

The analysis proceeds as before with $(q - p)$ replacing $(1 - p)$, and the detectability condition becomes

$$(q - p)\sqrt{k_1 k_2} \gtrsim \sqrt{p(1 - p)}(\sqrt{n_1} + \sqrt{n_2}).$$

6 Conclusion

We analyzed a spectral algorithm for detecting and recovering a planted biclique in a random bipartite graph. By centering the adjacency matrix and applying singular value decomposition, we reduced the problem to recovering a rank-one signal perturbed by a random noise matrix. Using operator norm bounds for the noise and singular vector perturbation theory, we established a detectability threshold of the form

$$\sqrt{k_1 k_2} \gtrsim \sqrt{n_1} + \sqrt{n_2},$$

which, in the symmetric case, matches the classical planted clique threshold up to constants.

References

[1] Noga Alon, Micheal Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3–4):457–466, 1998.

[2] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing*, Cambridge University Press, 2012.

[3] Per-Åke Wedin. Perturbation bounds in connection with singular value decomposition. *BIT Numerical Mathematics*, 12(1):99–111, 1972.

A Standard Bound Derivations

Lemma A.1. *Let $u_1 = \alpha u^\star + z$ with $z \perp u^\star$ and $\|z\|^2 \le \varepsilon$. Then for any threshold $t > 0$, the number of coordinates i such that $|z_i| > t$ satisfies*

$$|\{i : |z_i| > t\}| \le \frac{\varepsilon}{t^2}.$$

In particular, choosing $t = \varepsilon^{1/4}/\sqrt{k_1}$ implies that at most $k_1\sqrt{\varepsilon} = o(k_1)$ coordinates exceed t.

Proof. Let

$$S := \{i \in K_L : |z_i| > t\}.$$

Then by definition,

$$\sum_{i \in S} z_i^2 \ge |S| t^2.$$

But it is also know from Theorem 4.1 that $\|z\|^2 = \sum_i z_i^2 \le \varepsilon$. Hence:

$$|S| t^2 \le \varepsilon \Longrightarrow |S| \le \frac{\varepsilon}{t^2}.$$

Taking $t = \varepsilon^{1/4}/\sqrt{k_1}$ yields

$$|S| \le \frac{\varepsilon}{\varepsilon^{1/2}/k_1} = k_1\sqrt{\varepsilon}.$$

Under the SNR conditions imposed by Theorem 4.1, $\varepsilon \to 0$ hence

$$k_1\sqrt{\varepsilon} = o(k_1)$$

□

Lemma A.2 (No unusually dense bicliques under H_0). *Assume $H_0 : A \sim G_{\mathrm{bip}}(n_1, n_2, p)$ and fix integers k_1, k_2. For subsets $S \subseteq L$, $T \subseteq R$ with $|S| = k_1$, $|T| = k_2$, let*

$$X_{S,T} := \sum_{i \in S} \sum_{j \in T} A_{ij}$$

denote the number of edges between S and T. Then for all $\delta > 0$, there exist constants $c, C > 0$ such that

$$\begin{aligned}\mathbb{P}\Big(\exists\, S \subseteq L,\, T \subseteq R,\, |S| = k_1,\, |T| = k_2 : X_{S,T} \ge (p+\delta)k_1k_2\Big) \\ \le \exp\Big(-c\,\delta^2 k_1 k_2 + C\big(k_1 \log \tfrac{en_1}{k_1} + k_2 \log \tfrac{en_2}{k_2}\big)\Big).\end{aligned}$$

In particular, if

$$c\,\delta^2 k_1 k_2 \gg k_1 \log \frac{en_1}{k_1} + k_2 \log \frac{en_2}{k_2},$$

then w.h.p., there is no pair of subsets (S, T) of sizes (k_1, k_2) whose density exceeds $p + \delta$.

Proof. Under H_0, the entries $(A_{ij})_{i \in L, j \in R}$ are i.i.d. Bernoulli(p) rnadom variables.

Fix subsets $S \subseteq L$, $T \subseteq R$ with $|S| = k_1$ and $|T| = k_2$. Define $X_{S,T}$ as the number of edges between both subsets:

$$X_{S,T} = \sum_{i \in S} \sum_{j \in T} A_{ij}$$

It immediately follows that

$$\mathbb{E}[X_{S,T}] = k_1 k_2 p := mp.$$

By independence, Hoeffding's inequality yields:

$$\mathbb{P}\left[X_{S,T} \geq (p+\delta)m\right] = \mathbb{P}\left[X_{S,T} - \mathbb{E}X_{S,T} \geq \delta m\right] \leq \exp\big(-2\delta^2 m\big) = \exp\big(-2\delta^2 k_1 k_2\big).$$

Thus, there exists a constant $c > 0$ (namely $c = 2$) such that

$$\mathbb{P}\left[X_{S,T} \geq (p+\delta)k_1 k_2\right] \leq \exp\big(-c\,\delta^2 k_1 k_2\big). \tag{2}$$

There are

$$N = \binom{n_1}{k_1}\binom{n_2}{k_2}$$

choices of subset pairs (S,T) with $|S| = k_1$ and $|T| = k_2$. Using the known bound

$$\binom{N}{K} \leq \left(\frac{eN}{K}\right)^K$$

we get

$$N \leq \left(\frac{en_1}{k_1}\right)^{k_1}\left(\frac{en_2}{k_2}\right)^{k_2} = \exp\left(k_1 \log\frac{en_1}{k_1} + k_2 \log\frac{en_2}{k_2}\right).$$

Thus, union bounding (2) over all possible pairs yields

$$\begin{aligned}&\mathbb{P}\Big(\exists\, S \subseteq L,\, T \subseteq R,\, |S| = k_1,\, |T| = k_2 : X_{S,T} \geq (p+\delta)k_1 k_2\Big)\\ &\qquad \leq \exp\Big(-c\,\delta^2 k_1 k_2 + C\big(k_1 \log\tfrac{en_1}{k_1} + k_2 \log\tfrac{en_2}{k_2}\big)\Big).\end{aligned}$$

For any $C \geq 1$. Finally, if

$$c\delta^2 k_1 k_2 \gg k_1 \log\frac{en_1}{k_1} + k_2 \log\frac{en_2}{k_2},$$

the exponent tends to $-\infty$, and thus with high probability, there exists no pairs (S,T) of sizes (k_1, k_2) such that

$$X_{S,T} > (p+\delta)k_1 k_2.$$

□

Lemma A.3 (Concentration of the empirical edge density)**.** *Let $A \sim G_{\text{bip}}(n_1, n_2, p)$, and define*

$$\hat{p} = \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} A_{ij}.$$

Then for any $t > 0$,

$$\mathbb{P}\big(|\hat{p} - p| \geq t\big) \leq 2\exp\big(-2t^2 n_1 n_2\big).$$

In particular, when $n_1 n_2 \to \infty$,

$$|\hat{p} - p| = o(1).$$

Proof. Under H_0, the A_{ij} are i.i.d. Bernoulli(p) and are in $[0,1]$. Hoeffding's inequality yields

$$\mathbb{P}\left(\left|\sum_{i,j} A_{ij} - pn_1 n_2\right| \geq tn_1 n_2\right) \leq 2\exp\big(-2t^2 n_1 n_2\big).$$

Dividing by $n_1 n_2$ gives the stated bound. □