Supplementary material to "Multiscale Tests for Point Processes and Longitudinal Networks"

S1. Supplementary material for Sections 2 and 3

S1.1 Concise description of algorithms for longitudinal networks

S1.2 Asymmetric arrays of interactions

So far we've considered testing symmetric interaction processes among a single group of individuals, we can also extend this to the problem of testing the asymmetric interactions between two possibly different groups of individuals. Let V_1 and V_2 be two sets of individuals and suppose $|V_1| = m$ and $|V_2| = n$. Now we let $N_{uv}(\cdot)$ represents the temporal interactions events between individual $u \in V_1$ and individual $v \in V_2$, resulting in a collection of asymmetric point process realizations $\{N_{uv}(\cdot) : u \in V_1,\ v \in V_2\}$ where

$$N_{uv}(\cdot) \sim PP(\Lambda_{uv}), \quad \text{for intensity measure } \Lambda_{uv}.$$

Now we have $m \times n$ realizations in total and similarly we can reduce the the dimensionality of this problem by assuming community structures in both groups. Suppose there are K_1 and K_2 communities respectively in groups V_1 and V_2, we again assume the intensity function of the realization between two individuals only depends on their community memberships. More precisely, let $\boldsymbol{\sigma}_1 : [m] \to [K_1]$ and $\boldsymbol{\sigma}_2 : [n] \to [K_2]$ be two clustering function on groups V_1 and V_2, then we assume

$$\Lambda_{uv} = \Gamma_{\boldsymbol{\sigma}_1(u)\boldsymbol{\sigma}_2(v)}, \text{ for any } u \in V_1 \text{ and } v \in V_2$$

where $\{\Gamma_{st}\}_{s,t\in[K_1]\times[K_2]}$ is a collection of K_1K_2 intensity measures. We can again consider the goodness-of-fit test of the community structure with null hypothesis

$$H_0 : K_1 = K_2 = 1 \quad vs. \quad H_0 : K_1 \cdot K_2 > 1 \tag{S1.1}$$

and with a partition of $\boldsymbol{I}$ of the support $\mathcal{X}$, we can define each discretized local null as

$$\bar{H}_0^{(r,\ell)} : \Lambda_{uv}(I_\ell^{(r)}) = \gamma^{(r,\ell)}, \text{ for some common } \gamma^{(r,\ell)} \geq 0 \text{ and for all } u \in V_1, v \in V_2. \tag{S1.2}$$

Algorithm 1 Computing simultaneously valid p-values for $H_0^{(s,j)}$ in the symmetric array case.

INPUT: Poisson process realizations $\{N_{uv}(\cdot), u < v \in [n]\}$ and a hierarchical partitioning $\boldsymbol{I} = \{I_\ell^{(r)}\}_{r\in[R],\ell\in[2^r]}$ of the domain.

OUTPUT: p-values $\{p_F^{(s,j)}\}$.

1: **for** each $r \in [R]$ **do**

2: **for** each $\ell \in [2^r]$ **do**

3: Define integer matrix $A^{(r,\ell)}$ as (3.16).

4: Set

$$\bar{p}^{(r,\ell)} = 2\min\Big(F_{\mathrm{TW1}}\Big(n^{2/3}\big(\lambda_1(A^{(r,\ell)})\big) - 2\Big), 1 - F_{\mathrm{TW1}}\Big(n^{2/3}\big(\lambda_1(A^{(r,\ell)})\big) - 2\Big)\Big).$$

5: **end for**

6: **end for**

7: Apply Algorithm 2 on $\{\bar{p}^{(r,\ell)}\}$ to obtain $\{\{p_F^{(s,j,r)}\}_{r=s}^{R}\}_{s\in[R],j\in[2^s]}$.

8: Compute $\tilde{p}_F^{(s,j)} = \min\{p_F^{(s,j,r)} : r \in \{s, s+1, \ldots, R\}\}$.

9: Run Metropolis–Hastings described in Section 3.1.3 to generate $M_1^{(b^*)}, \ldots, M_N^{(b^*)}$ for $b^* \in [B]$ and let $\{N_{uv}^{(b^*)}(\cdot)\}$ be the corresponding point process realizations.

10: **for** $b^* \in \{1, 2, \ldots, B\}$ **do**:

11: For each $r \in [R]$ and $\ell \in [2^r]$, construct $A_{uv}^{(r,\ell)\,(b^*)}$ from $\{N_{uv}^{(b^*)}(\cdot)\}$.

12: Repeat lines 1 to 8 to obtain $\tilde{p}_{F,b^*}^{(s,j)}$.

13: **end for**

14: Compute $\check{p}_F^{(s,j)} := \frac{1}{B}\sum_{b^*=1}^{B} \mathbb{1}\{\tilde{p}_{F,b^*}^{(s,j)} \leq \tilde{p}_F^{(s,j)}\}$.

15: Compute $p_F^{(s,j)} = \check{p}_F^{(s,j)} 2^s$.

Algorithm 2 Computing simultaneously valid p-values for $H_0^{(s,j)}$ in the degree-corrected setting.

INPUT: Poisson process realizations $\{N_{uv}(\cdot), u < v \in [n]\}$ and a hierarchical partitioning $\boldsymbol{I} = \{I_\ell^{(r)}\}_{r \in [R], \ell \in [2^r]}$ of the domain.

OUTPUT: p-values $\{p_F^{(s,j)}\}$

1: **for** each $r \in [R]$ **do**

2: **for** each $\ell \in [2^r]$ **do**

3: Set $p^{(r,\ell)} = 2\left[1 - \Phi\left(\left|\frac{T^{(r,\ell)}}{\sqrt{6}(\|\hat{\eta}^{(r,\ell)}\|^2 - 1)^{3/2}}\right|\right)\right]$.

4: **end for**

5: **end for**

6: Apply Algorithm 2 on $\{\bar{p}^{(r,\ell)}\}$ to obtain $\{\{p_F^{(s,j,r)}\}_{r=s}^{R}\}_{s \in [R], j \in [2^s]}$.

7: Compute $\tilde{p}_F^{(s,j)} := \min\{p_F^{(s,j,r)} : r \in \{s, s+1, \ldots, R\}\}$.

8: Run Metropolis–Hastings algorithm described in Section 3.2.3 to generate $\boldsymbol{m}^{(b^*)}$ for $b^* \in [B]$ and let $\{N_{uv}^{(b^*)}(\cdot)\}$ be the corresponding point process realizations.

9: **for** $b^* \in \{1, 2, \ldots, B\}$ **do**:

10: For each $r \in [R]$ and $\ell \in [2^r]$, construct $A_{uv}^{(r,\ell)\,(b^*)}$ from $\{N_{uv}^{(b^*)}(\cdot)\}$.

11: Repeat lines 1 to 8 to obtain $\tilde{p}_F^{(b^*)}$.

12: **end for**

13: Compute the adjusted p-value $\check{p}_F^{(s,j)} := \frac{1}{B}\sum_{b^*=1}^{B} \mathbb{1}\{\tilde{p}_{F,(b^*)}^{(s,j)} \leq \tilde{p}_F^{(s,j)}\}$.

14: Compute $p^{(s,j)} = \check{p}_F^{(s,j)} \cdot 2^{-s}$.

Similar to the local adjacency matrix $A^{(r,\ell)}$ defined in previous section, we let $B^{(r,\ell)}$ be a $m \times n$ matrix with entries being counts of interactions between any two individuals from groups V_1 and V_2 respectively, within interval $I_\ell^{(r)}$

$$B_{uv}^{(r,\ell)} = N_{uv}(I_\ell^{(r)}), \text{ for any } u \in V_1 \text{ and } v \in V_2$$

To test each discretized local null $\bar{H}_0^{(r,\ell)}$ given observed matrix $B^{(r,\ell)}$, we again remove the mean effect and check whether the residual matrix looks like random noise. We define

$$\hat{\gamma}^{(r,\ell)} = \sum_{u \in V_1} \sum_{v \in V_2} \frac{B_{uv}^{(r,\ell)}}{mn}. \tag{S1.3}$$

Moreover, we define

$$\tilde{B}^{(r,l)} = \frac{B^{(r,\ell)} - \hat{\gamma}^{(r,\ell)}}{\sqrt{m \cdot \hat{\gamma}^{(r,\ell)}}} \in \mathbb{R}^{n \times m} \tag{S1.4}$$

as the empirically scaled and centered counterpart of $B^{(r,\ell)}$, with $\hat{\gamma}^{(r,\ell)}$ defined as in (S1.3) and let $\tilde{W}^{(r,\ell)} = (\tilde{B}^{(r,\ell)})^\intercal \tilde{B}^{(r,\ell)}$. Then we have the following limiting distribution of the largest eigenvalues of $\tilde{W}^{(r,\ell)}$.

Theorem S1. *Let $\lambda_1(\tilde{W}^{(r,\ell)})$ be the largest eigenvalue of matrix $\tilde{W}^{(r,\ell)}$ and suppose $\lim_{n\to\infty} n/m \in (0,\infty)$. Then for each $r \in [R]$ and $\ell \in [2^r]$, under the discretized local null hypothesis $\bar{H}_0^{(r,\ell)}$ given in (S1.2), we have, as $n, m \to \infty$,*

$$\frac{m \cdot \lambda_1(\tilde{W}^{(r,\ell)}) - (\sqrt{n} + \sqrt{m})^2}{(\sqrt{n} + \sqrt{m})(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}})^{1/3}} \xrightarrow{d} TW_1. \tag{S1.5}$$

We relegate the proof of Theorem S1 to Section S1.5.3 of the Appendix.

Using Theorem S1, we can let $\lambda_1(\tilde{W}^{(r,\ell)})$ be the test statistics for the local test (S1.2), and derive the p-value for the discretized local null as

$$\begin{aligned} \bar{p}^{(r,\ell)} &\equiv p^{(r,\ell)}(\tilde{W}^{(r,\ell)}) \\ &:= 2\min\Bigg(F_{\text{TW1}}\Big(\frac{m \cdot \lambda_1(\tilde{W}^{(r,\ell)}) - (\sqrt{n} + \sqrt{m})^2}{(\sqrt{n} + \sqrt{m})(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}})^{1/3}}\Big), 1 - F_{\text{TW1}}\Big(\frac{m \cdot \lambda_1(\tilde{W}^{(r,\ell)}) - (\sqrt{n} + \sqrt{m})^2}{(\sqrt{n} + \sqrt{m})(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}})^{1/3}}\Big)\Bigg) \end{aligned}$$

Steps 2, 3, and 4 proceed in the same way as the symmetric case, except that the resampling method changes slightly. To generate samples under the null in this scenario, we can just change the distribution of the random marks to be $\mathbb{P}\big(M_i = (u,v)\big) = \frac{1}{mn}$, $\forall i \in [N]$, $u \in V_1$ and $v \in V_2$. Then we generate sequences of random marks $\{M_1^{(b^*)}, \ldots, M_N^{(b^*)}\}$ under the aforementioned distribution and let the collection $\{N_{uv}^{(b^*)}(\cdot) : u \in V_1, v \in V_2\}$ be a resample of the observed asymmetric array.

S1.3 Randomizing p-value

Let X be a discrete random variable taking value on $\{x_1, \ldots, x_m\} \subset \mathbb{R}$ where we have the ordering $x_1 \le x_2 \le \ldots x_m$. Define $S(x) = \mathbb{P}(X \ge x)$ and

$$q_1 := S(x_1) = 1,\ q_2 := S(x_2), \ldots, q_m := S(x_m),\ q_{m+1} := 0,$$

so that $S(X)$ takes value on $\{q_1, q_2, \ldots, q_m\}$. We define random variable $\tilde{S}$ such that if $S(X) = q_i$, then $\tilde{S} = q_{i+1}$.

Proposition 1. *Let $U \sim$ Unif$[0,1]$ be independent of X. Define*

$$Z := U \cdot S(X) + (1 - U) \cdot \tilde{S}. \tag{S1.6}$$

Then, we have that $Z \le S(X)$ and that $Z \sim$ Unif$[0,1]$.

Proof. Since $\tilde{S} < S(X)$ by definition, it is clear that $Z \le S(X)$ as well. To show that Z has the Unif$[0,1]$ distribution, fix $t \in (0,1)$. Then there exists $i \in [m]$ such that $q_i \ge t > q_{i+1}$. We then have that

$$\begin{aligned}\mathbb{P}(Z \le t) &= \mathbb{P}(S(X) \le q_{i+1}) + \mathbb{P}(Z \le t, S(X) = q_i, \tilde{S} = q_{i+1}) \\ &= q_{i+1} + \mathbb{P}\left(U \le \frac{t - q_{i+1}}{q_i - q_{i+1}}\right)\mathbb{P}(S(X) = q_i) = t,\end{aligned}$$

where the last inequality follows because $\mathbb{P}(S(X) = q_i) = \mathbb{P}(X = x_i) = S(x_i) - S(x_{i+1}) = q_i - q_{i+1}$. The Proposition follows as desired.

□

S1.4 Proof of Proposition 2

Proof. Let $T(\cdot|\cdot)$ be the transition probability of the Markov Chain specified via (3.27), we first verify that the Metropolis-Hastings ratio is 1 by showing that $T(\cdot|\cdot)$ is a symmetric distribution, i.e., for any two sample vector $\boldsymbol{m}^{(1)} \neq \boldsymbol{m}^{(2)} \in \mathcal{M}_{\boldsymbol{d}}$, we have $T(\boldsymbol{m}^{(1)}|\boldsymbol{m}^{(2)}) = T(\boldsymbol{m}^{(2)}|\boldsymbol{m}^{(1)})$. It is obvious that the necessary condition for $T(\boldsymbol{m}^{(1)}|\boldsymbol{m}^{(2)})$ to be positive, is that there must exist exactly two indices $i \neq j \in [N]$ such that $m_i^{(1)} \neq m_i^{(2)}$, $m_j^{(1)} \neq m_j^{(2)}$ while the other elements are all the same for the two vectors. We can see that $T(\boldsymbol{m}^{(1)}|\boldsymbol{m}^{(2)}) = T(\boldsymbol{m}^{(2)}|\boldsymbol{m}^{(1)}) = \frac{1}{5\binom{[N]}{2}}$ regardless of the values of $m_i^{(1)}, m_j^{(1)}, m_i^{(2)}, m_j^{(2)}$.

Next, we show that the Markov Chain is irreducible on the support $\mathcal{M}_{\boldsymbol{d}}$. By definition it suffice to show that for any two sample $\boldsymbol{m}^{(1)} \neq \boldsymbol{m}^{(2)} \in \mathcal{M}_{\boldsymbol{d}}$, there exist a finite steps path $\boldsymbol{m}^{(1)} \longrightarrow \boldsymbol{m}^{(2)}$. Since the vector of all degrees

$\boldsymbol{D(m)}$ are identical for all $\boldsymbol{m} \in \mathcal{M}_{\boldsymbol{d}}$, we can show there exist a path $\boldsymbol{m}^{(1)} \longrightarrow \boldsymbol{m}^{(2)}$ which sequentially matching each element of $\boldsymbol{m^{(1)}}$ to be the same as in $\boldsymbol{m^{(2)}}$. To be specific, we denote the j_{th} elements of vectors $\boldsymbol{m}^{(1)}$ and $\boldsymbol{m}^{(2)}$ as $\boldsymbol{m}_j^{(1)} = (u_j, v_j)$ and $\boldsymbol{m}_j^{(2)} = (u_j', v_j')$ respectively. Let $i = \min\left\{j \in [N] : \boldsymbol{m}_j^{(1)} \neq \boldsymbol{m}_j^{(2)}\right\}$, we first show that we can go from $\boldsymbol{m}^{(1)}$ to an intermediate state $\boldsymbol{m}^{(1,i)} \in \mathcal{M}_{\boldsymbol{d}}$ in finite steps, such that $\boldsymbol{m}_j^{(1,i)} = \boldsymbol{m}_j^{(2)}$ for all $j \leq i$. We can easily see for $\boldsymbol{m}_i^{(1)} \neq \boldsymbol{m}_i^{(2)}$, there could only be two cases

1. $u_i = u_i', v_i \neq v_i'$ or $u_i \neq u_i', v_i = v_i'$

2. $u_i \neq v_i \neq u_i' \neq v_i'$

For the first case, suppose $u_i = u_i', v_i \neq v_i'$, then by the fact $\boldsymbol{D(m^{(1)})} = \boldsymbol{D(m^{(2)})}$ there must exist $s > i$ such that $u_s = v_i'$ or $v_s = v_i'$. Then by (3.27) we can easily check that we can go from $\boldsymbol{m}^{(1)}$ to $\boldsymbol{m}^{(1,i)}$ in one step with $T(\boldsymbol{m}^{(1,i)}|\boldsymbol{m}^{(1)}) = \frac{1}{5\binom{[N]}{2}}$. For the second case, we can go from $\boldsymbol{m}^{(1)}$ to $\boldsymbol{m}^{(1,i)}$ in two steps where in the first step we move to a state $\boldsymbol{m}^{(1,i)\prime}$ such that $\boldsymbol{m}_i^{(1,i)\prime} = (u_i', v_j)$ which is just the state in the first case, so by the same reason we can move to $\boldsymbol{m}^{(1,i)}$ in the second step. Notice that the above paths does not depend on the index i, thus there exist an integer $t \leq 2N$ such that we can sequentially move from $\boldsymbol{m}^{(1)} \longrightarrow \boldsymbol{m}^{(1,i)} \longrightarrow \cdots \longrightarrow \boldsymbol{m}^{(1,N)} = \boldsymbol{m}^{(2)}$ in t steps.

Given any state $\boldsymbol{m}$, if not all edges are the same, i.e., $m_i = (u, v)$ for some $u, v \in [n]$ and all $i \in [N]$, then we can always find $m_i = (u_i, v_i)$ and $m_j = (u_j, v_j)$ with $i < j \in [N]$ such that the nodes u_i, v_i, u_j, v_j satisfy one of the two cases listed above. For case 1, we can see that the five outcomes contains multi-edges, so the Markov Chain can stay at current state with positive probability and thus the period for this state is 1. For 2, we can easily check the state can return in t steps for any $t \geq 2$, thus the period for this state is also 1. Then by irreducibility, we can conclude that the Markov Chain is aperiodic.

With the Markov Chain being irreducible and aperiodic, it converges to its unique stationary distribution. Then by the construction of the Metropolis-Hastings algorithm below (3.27), it is guaranteed that the stationary distribution is the target distribution, i.e., the uniform distribution on $\mathcal{M}_{\boldsymbol{d}}$.

□

S1.5 Supplementary material for Section 3.1

Recall that we let $A^{(r,\ell)}$, defined in (3.16), be the adjacency matrix of a undirected Poisson Stochastic Block Model with K communities. We denote $\boldsymbol{\sigma}$ as the membership vector and $\boldsymbol{\gamma}$ as the connection intensity between different communities, as discussed in Section 3.1.1. Without loss of generality, we omit all the superscripts of $A^{(r,\ell)}$ that represents the partition of the support $\mathcal{X}$ and just use A to denote adjacency matrix generated from Poisson Stochastic Block Model in all subsequent analysis in Section S1.5 for simplicity. For the same reason, we also omit the subscripts of $B^{(r,\ell)}$, the matrix with entries being counts of interactions between two groups of individuals, in the proof of theorem S1.

S1.5.1 Proof of Theorem 2

Proof. Under the null hypothesis, we have that $\boldsymbol{P} = \gamma \mathbf{1}_n \mathbf{1}_n^\intercal$, for some constant $\gamma > 0$.

Let $\tilde{A}'$ be a $n \times n$ matrix such that

$$\tilde{A}'_{uv} = \begin{cases} (A_{uv} - \hat{\gamma})/\sqrt{(n-1)\gamma}, & u \neq v \\ (\gamma - \hat{\gamma})/\sqrt{(n-1)\gamma}, & u = v \end{cases}$$

Where $\hat{\gamma} = \frac{2}{n^2-n}\sum_{u<v} A_{uv}$ is an estimator of γ. Let $C_n = n(\gamma - \hat{\gamma})/\sqrt{(n-1)\gamma}$ and matrix $\tilde{A}^*$ be as defined in (S1.9). Then by definition we have that $\tilde{A}' = \tilde{A}^* + \Delta'$, where $\Delta' = (\gamma - \hat{\gamma})\mathbf{1}_n\mathbf{1}_n^\intercal/\sqrt{(n-1)\gamma} = C_n \mathbf{1}_n \mathbf{1}_n^\intercal / n$. Note that $\hat{\gamma}$ is the sample mean of $n(n-1)/2$ *i.i.d* Poisson random variables with mean γ, we can apply the Poisson tail bound (S1.10) again and get

$$\mathbb{P}\left(|\gamma - \hat{\gamma}| > s\right) \leq 2\exp\big\{-\frac{n(n-1)s^2}{4(\gamma+s)}\big\}$$

And thus we have $|\gamma - \hat{\gamma}| = o_p(\log n/n)$ and that $C_n = o_p(\log n/\sqrt{n})$.

Let μ_i^* be the eigenvector of $\tilde{A}^*$ corresponding to its i_{th} largest eigenvalue. Then by Lemma 2, we have a lower bound on the largest eigenvalue of $\tilde{A}'$:

$$\begin{aligned} \lambda_1(\tilde{A}') &\geq (\mu_1^*)^\intercal \tilde{A}' \mu_1^* \\ &= \lambda_1(\tilde{A}^*) + (\mu_1^*)^\intercal \Delta' \mu_1^* \\ &= \lambda_1(\tilde{A}^*) + C_n (\mu_1^*)^\intercal \mathbf{1}_n \mathbf{1}_n^\intercal \mu_1^*/n \\ &= \lambda_1(\tilde{A}^*) + \tilde{O}_p(1/n) \cdot o_p(\log n/\sqrt{n}) \\ &\geq \lambda_1(\tilde{A}^*) - o_p(n^{-2/3}) \end{aligned}$$

To derive the upper bound of $\lambda_1(\tilde{A}')$, we denote μ_1' as the eigenvector corresponding to the largest eigenvalue of $\tilde{A}'$. Let $\{a_1, \ldots, a_n\}$ be the coordinates of the vector μ_1' with respect to the basis $\{\mu_1^*, \ldots, \mu_n^*\}$, i.e., $\mu_1' = \sum_{i=1}^n a_i \mu_i^*$. Define $\mathcal{S}_{C_n} \subset [n] := \{i \in [n] : \lambda_i(\tilde{A}^*) > \big(\lambda_1(\tilde{A}^*) - |C_n|\big)\}$ as the set of indices of those eigenvalues of $\tilde{A}^*$ that lies in the interval $(\lambda_1(\tilde{A}^*) - |C_n|, \lambda_1(\tilde{A}^*)]$. Then By Lemma 2 and the fact that $|C_n|$ is the largest eigenvalue of Δ', we have

$$\begin{aligned}
\lambda_1(\tilde{A}') &= (\mu_1')^\top \tilde{A}' \mu_1' \\
&= (\mu_1')^\top \tilde{A}^* \mu_1' + (\mu_1')^\top \Delta' \mu_1' \\
&\leq \sum_{i=1}^{n} a_i^2 \lambda_i(\tilde{A}^*) + \big(\sum_{i\in\mathcal{S}_{Cn}} a_i(\mu_i^*)^\top\big)|\Delta'|\big(\sum_{i\in\mathcal{S}_{Cn}} a_i(\mu_i^*)\big) \\
&\quad + \big(\sum_{i\in([n]/\mathcal{S}_{C_n})} a_i(\mu_i^*)^\top\big)|\Delta'|\big(\sum_{i\in([n]/\mathcal{S}_{C_n})} a_i\mu_i^*\big) \\
&\leq \lambda_1(\tilde{A}^*) \sum_{i\in\mathcal{S}_{C_n}}^{m} a_i^2 + \big(\lambda_1(\tilde{A}^*) - |C_n|\big) \sum_{i\in([n]/\mathcal{S}_{C_n})} a_i^2 \\
&+ |\mathcal{S}_{C_n}| \cdot \sum_{i\in([n]/\mathcal{S}_{C_n})}^{m} a_i^2(\mu_i^*)^\top|\Delta'|\mu_i^* \quad + |C_n| \sum_{i\in([n]/\mathcal{S}_{C_n})}^{m} a_i^2 \\
&\leq \lambda_1(\tilde{A}^*) + |\mathcal{S}_{C_n}||C_n| \cdot \big(\sum_{i\in([n]/\mathcal{S}_{C_n})}^{m} a_i^2 \cdot \tilde{O}_p(1/n)\big) \\
&= \lambda_1(\tilde{A}^*) + |\mathcal{S}_{C_n}| \cdot \tilde{O}_p(1/n) \cdot o_p(\log n/\sqrt{n})
\end{aligned}$$

Then we could bound the size of $\mathcal{S}_{C_n}$ by using the results from **?** and **?**, where the main idea is that the empirical counting of the eigenvalues is close to the semicircle counting functions.

Let $N(a, b)$ be the number of eigenvalues of $\tilde{A}^*$ lying in interval $(a, b]$, and define $N_{sc}(a, b) := n \int_a^b \rho_{sc}(x)dx$, where $\rho_{sc} = (1/2\pi)((4 - x^2)_+)^{1/2}$ denote the the density of the semicircle law discussed in **?**. Following Theorem 2.2 in **?** and the discussion in **?**, there exist constant $A_0 > 1$, C, c and $d < 1$, such that for any L satisfying the following:

$$A_0 \mathrm{loglog} n \leq L \leq \log(10n)/\mathrm{loglog} n$$

and for $|a|$, $|b| < 5$, we have :

$$\begin{aligned}
&\mathbb{P}\left(|N(a,b) - N_{sc}(a,b)| \geq 2(\log n)^L\right) \\
\leq &\mathbb{P}\left(|N(-\infty,b) - N_{sc}(\infty,b)| \geq (\log n)^L\right) + \mathbb{P}\left(|N(-\infty,a) - N_{sc}(\infty,a)| \geq (\log n)^L\right) \\
\leq &2C\exp\{-c(\log n)^{(-dL)}\}
\end{aligned}$$

Notice that $\mathcal{S}_{C_n} = N\big(\lambda_1(\tilde{A}^*) - |C_n|, \lambda_1(\tilde{A}^*)\big)$, and from the above inequality we have that:

$$\mathcal{S}_{C_n} = N\big(\lambda_1(\tilde{A}^*) - |C_n|, \lambda_1(\tilde{A}^*)\big) = N_{sc}\big(\lambda_1(\tilde{A}^*) - |C_n|, \lambda_1(\tilde{A}^*)\big)\big) + O_p(\log n)^L \tag{S1.7}$$

And

$$
\begin{aligned}
N_{sc}\big(\lambda_1(\tilde{A^*}) - |C_n|, \lambda_1(\tilde{A^*})\big)\big) &= n\int_{\lambda_1(\tilde{A^*})-|C_n|}^{\lambda_1(\tilde{A^*})} (\frac{1}{2\pi}((4-x^2)_+)^{1/2}dx \\
&\le n\int_{\lambda_1(\tilde{A^*})-|C_n|}^{2} (\frac{1}{2\pi}((4-x^2)_+)^{1/2}dx \\
&= O(n|C_n|^{3/2}) \\
&= o_p\big(n^{1/4}(\log n)^{3/2}\big)\big)
\end{aligned}
$$

Where the second to last equality holds by using the area of a rectangle with side length $(2-|C_n|)$ and $\sqrt{4-(2-|C_n|)^2}$ to cover the actual size of the integral.

Now we can see that

$$
\begin{aligned}
\lambda_1(\tilde{A}') &\le \lambda_1(\tilde{A^*}) + |\mathcal{S}_{C_n}| \cdot \tilde{O}_p(1/n) \cdot o_p(\log n/\sqrt{n}) \\
&= \lambda_1(\tilde{A^*}) + o_p\big(n^{1/4}(\log n)^{3/2}\big)\big) \cdot \tilde{O}_p(1/n) \cdot o_p(\log n/\sqrt{n}) \\
&= \lambda_1(\tilde{A^*}) + \tilde{O}_p((\log n)^{5/2} n^{-5/4}) \\
&\le \lambda_1(\tilde{A^*}) + o_p(n^{-2/3})
\end{aligned}
$$

And combining the lower and upper bound we have that

$$
\lambda_1(\tilde{A}') = \lambda_1(\tilde{A^*}) + o_p(n^{-2/3}) \tag{S1.8}
$$

Now let's get back to the target matrix $\tilde{A} = \sqrt{\frac{\hat{\gamma}}{\gamma}}\Big(\tilde{A}' - \frac{C_n}{n}\boldsymbol{I}_n\Big)$. By triangle inequality of matrix norm we have

$$
\|\tilde{A}'\| - \|\frac{C_n}{n}\boldsymbol{I}_n\| \le \|\tilde{A}' - \frac{C_n}{n}\boldsymbol{I}_n\| \le \|\tilde{A}'\| + \|\frac{C_n}{n}\boldsymbol{I}_n\|
$$

And since $\|\frac{C_n}{n}\boldsymbol{I}_n\| = |C_n/n| = o_p(\log n \cdot n^{-3/2})$, we could easily see that

$$
\begin{aligned}
\lambda_1(\tilde{A}) &= \sqrt{\hat{\gamma}/\gamma} \cdot \|\tilde{A}' - \frac{C_n}{n}\boldsymbol{I}_n\| \\
&= \big(1 + o_p(\log n/n)\big)\big(\lambda_1(\tilde{A}') + o_p(\log n \cdot n^{-3/2})\big) \\
&= \lambda_1(\tilde{A}') + o_p(n^{-2/3}) \\
&= \lambda_1(\tilde{A^*}) + o_p(n^{-2/3})
\end{aligned}
$$

Finally by Lemma 1 and Slutsky's lemma, we have

$$
n^{2/3}\big(\lambda_1(\tilde{A}) - 2\big) \xrightarrow{d} \mathrm{TW}_1.
$$

□

Lemma 1. *Let $\boldsymbol{P}$ be defined in (3.17) and $\tilde{A}^*$ be a matrix such that*

$$\tilde{A}^*_{uv} = \begin{cases} (A_{uv} - \boldsymbol{P}_{uv})/\sqrt{(n-1)\boldsymbol{P}_{uv}}, & u \neq v \\ 0, & u = v \end{cases} \tag{S1.9}$$

Then we have

$$n^{2/3}(\lambda_1(\tilde{A}^*) - 2) \xrightarrow{d} TW_1.$$

Proof. Consider a $n \times n$ real symmetric Wigner matrix

$$G^*_{uv} = \frac{1}{\sqrt{n-1}} x_{uv}, \quad 1 \leq u, v \leq n$$

Where the off-diagonal elements are *i.i.d.* standard normal distributed random variables and the diagonal elements are zeros. Theorem 1.2. in **?** implies that the largest eigenvalue of G^* weakly converges to the Tracy-Widom distribution.

Next, by tail bound of Poisson random variables, for any $s > 0$ and $1 \leq u, v \leq n$ we have

$$\mathbb{P}\left(\left|\tilde{A}^*_{uv}\right| > \frac{s}{\sqrt{n-1}}\right) \leq 2\exp(-\frac{\boldsymbol{P}_{uv}s^2}{2(\boldsymbol{P}_{uv} + \sqrt{\boldsymbol{P}_{uv}}s)}) \tag{S1.10}$$

And thus there exist a constant v independent of n, such that for any $s \geq 1$ we have

$$\mathbb{P}\left(\left|\tilde{A}^*_{uv}\right| > \frac{s}{\sqrt{n-1}}\right) \leq v^{-1}\exp(-s^v)$$

The above inequality shows that the entries of $\tilde{A}^*$ have a uniformly subexponential decay, and thus by Theorem 2.4 in **?**, we have that $n^{2/3}(\lambda_1(\tilde{A}^*) - 2)$ converges to $n^{2/3}(\lambda_1(G^*) - 2)$ in distribution. □

Lemma 2. *For each $1 \leq i \leq n$, let μ^*_i be the eigenvector of $\tilde{A}^*$ corresponding to the i_{th} largest eigenvalue $\lambda_i(\tilde{A}^*)$. Then for any deterministic unit vector $\mathbf{v}$, we have*

$$((\mu^*_i)^\intercal \mathbf{v})^2 = \tilde{O}_p(1/n), \quad \textit{uniformly for all } i \in [n] \tag{S1.11}$$

Where we define $a_n = \tilde{O}_p(b_n)$, if for any $\epsilon > 0$ and $D > 0$, there exists $n_0 = n_0(\epsilon, D)$ such that

$$\mathbb{P}(a_n \geq n^\epsilon b_n) \leq n^{-D} \quad \textit{for all } n \geq n_0.$$

Lemma 2 is a direct application of the eigenvector delocalization theorem proposed in **?**. Note that the conditions of Theorem 2.16 in **?** does not apply to our configuration of $\tilde{A}^*$ since the diagonal entries are made to be all zeros while the original condition requires that all elements of the matrix should have positive variance. However, **?** provides a local semicircle law(Theorem 2.3) which holds even when some entries of a generalized Wigner matrix have zero variance, and as a result of the local semicircle law, the eigenvector delocalization theorem still holds in our setting. See also discussions in **?** and Lei (2016).

S1.5.2 Maximum eigenvalue test statistic under an alternative

We consider the limiting distribution of $\lambda_1(\tilde{A}^{(r,\ell)})$ under some alternative cases. When the adjacency matrix is generated from a Stochastic Block Model with $K > 1$ communities and Bernoulli entries, Bickel and Sarkar (2013) shows that the largest eigenvalue of the scaled and centered adjacency matrix is $O(\sqrt{n})$, given that the community probability matrix $\boldsymbol{\psi}$ is diagonally dominant. Lei (2016) provided a more general result which requires that each community has size at least proportional to n/K. The following proposition is a direct extension of Theorem 3.3 in Lei (2016) to the Poisson network.

Proposition 2. *Let $A^{(r,\ell)}$ be an adjacency matrix from Poisson stochastic model with K communities and let $\mathcal{G}_k = \{u \in [n] : \sigma(u) = k\}$ be the set of vertices that belong to group k for $k \in [K]$.*

Assume there exist a constant $C_K > 0$ such that for all n we have

$$\min_{k\in[K]} |\mathcal{G}_k| \geq C_K \cdot n \tag{S1.12}$$

Then for any $r \in [R]$ and $\ell \in [2^r]$, if $K > 1$ we have

$$\lambda_1(\tilde{A}^{(r,\ell)}) \geq \frac{\sqrt{n}\delta C_K - O_p(1)}{\big(\|\boldsymbol{\gamma}\|_{max} + o_p(\log n/n)\big)^{1/2}} \tag{S1.13}$$

where δ is the minimum ℓ_∞ distance between any two distinct rows of $\boldsymbol{\gamma}$.

Proof. Let $\hat{\boldsymbol{P}} = \hat{\gamma}\mathbf{1}_n\mathbf{1}_n^\intercal$, we have

$$\begin{aligned}\|\tilde{A}\| &= \big((n-1)\hat{\gamma}\big)^{-1/2}\big\|A - \big(\hat{\boldsymbol{P}} - \operatorname{diag}(\hat{\boldsymbol{P}})\big)\big\| \\ &\geq \big((n-1)\hat{\gamma}\big)^{-1/2}\Big(\big\|\boldsymbol{P} - \hat{\boldsymbol{P}} - \operatorname{diag}(\boldsymbol{P} - \hat{\boldsymbol{P}})\big\| - \big\|A - \big(\boldsymbol{P} - \operatorname{diag}(\boldsymbol{P})\big)\big\|\Big)\end{aligned}$$

We can see that the matrix $A - \big(\boldsymbol{P} - \operatorname{diag}(\boldsymbol{P})\big)$ has off-diagonal entries being independent, centered Poisson random variables and diagonal entries being all zeros. By Theorem 2 in **?**, we have that there exist some $C' > 0$ such that

$$\mathbb{E}\big\|A - \big(\boldsymbol{P} - \operatorname{diag}(\boldsymbol{P})\big)\big\| \leq C'\sqrt{n} \tag{S1.14}$$

and thus $\big\|A - \big(\boldsymbol{P} - \mathrm{diag}(\boldsymbol{P})\big)\big\| \leq O_p(\sqrt{n})$.

To derive an upper bound of $\big\|\boldsymbol{P} - \hat{\boldsymbol{P}} - \mathrm{diag}(\boldsymbol{P} - \hat{\boldsymbol{P}})\big\|$, we notice that since $K > 1$, there exist two community $k_1 \neq k_2$. Let $\boldsymbol{g}_{k_1} = \{u \in [n] : \sigma(u) = k_1\}$ and $\boldsymbol{g}_{k_2} = \{u \in [n] : \sigma(u) = k_2\}$ be the set of vertices that belong to k_1 and k_2 respectively. Since we assume the matrix $\boldsymbol{\gamma}$ have pairwise distinct rows, there must exist a group $k_3 \in [K]$ such that $\boldsymbol{\gamma}_{k_1 k_3} \neq \boldsymbol{\gamma}_{k_2 k_3}$, and we can choose

$$k_3 = \underset{k' \in [K]}{\mathrm{argmin}} |\boldsymbol{\gamma}_{k_1 k'} - \boldsymbol{\gamma}_{k_2 k'}|.$$

Note that k_3 can be equal to k_1 or k_2. Now let $\boldsymbol{D}$ be a submatrix of $\boldsymbol{P} - \hat{\boldsymbol{P}} - \mathrm{diag}(\boldsymbol{P} - \hat{\boldsymbol{P}})$, which only consist the rows in $k_1 \cup k_2$ and columns in k_3. We can see that when $k_1 \neq k_2 \neq k_3$, after some row permutaions D could be seen as:

$$\boldsymbol{D} = \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}$$

where D_1 is a $|k_1| \times |k_3|$ matrix with all entries equal to $\boldsymbol{\gamma}_{k_1 k_3} - \hat{\gamma}$ and D_2 is a $|k_2| \times |k_3|$ matrix with all entries equal to $\boldsymbol{\gamma}_{k_2 k_3} - \hat{\gamma}$. Then we have

$$\begin{aligned}\big\|\boldsymbol{D}\big\| &\geq \max\Big((\boldsymbol{\gamma}_{k_1 k_3} - \hat{\gamma})\sqrt{|k_1| \cdot |k_3|}, (\boldsymbol{\gamma}_{k_2 k_3} - \hat{\gamma})\sqrt{|k_2| \cdot |k_3|}\Big) \\ &\geq n\delta C_K\end{aligned}$$

When $k_3 = k_1$ or $k_3 = k_2$, we can see $\boldsymbol{D}$ could still be permuted into a block matrix with blocks D_1 and D_2. However in this case, one of the blocks have all diagonal entries being zeros, since we do not allow self-loops. Without loss of generality, we assume that $k_3 = k_1$, and we still have the same lower bound of $\big\|\boldsymbol{D}\big\|$ by

$$\begin{aligned}\big\|\boldsymbol{D}\big\| &\geq \max\Big((\boldsymbol{\gamma}_{k_1 k_1} - \hat{\gamma})(|k_1| - 1), (\boldsymbol{\gamma}_{k_2 k_1} - \hat{\gamma})\sqrt{|k_2| \cdot |k_1|}\Big) \\ &\geq n\delta C_K - O_p(1)\end{aligned}$$

Finally we have

$$\begin{aligned}\big\|\tilde{A}\big\| &\geq \big((n-1)\hat{\gamma}\big)^{-1/2}\big(n\delta C_K - O_p(\sqrt{n})\big) \\ &\geq \frac{\sqrt{n}\delta C_K - O_p(1)}{\big(\|\boldsymbol{\gamma}\|_{\max} + o_p(\log n/n)\big)^{1/2}}\end{aligned}$$

□

S1.5.3 Proof of Theorem S1

Proof. Under the null hypothesis, we have that $B_{uv} \sim \text{Poisson}(\gamma)$ for some $\gamma > 0$ and for any $u \in V_1, v \in V_2$. Recall that we denote $\hat{\gamma}$ (S1.3) as an estimator of γ, matrix $\tilde{B}$ as the empirically centered and scaled counterpart of B and $\tilde{W} = \tilde{B}^\intercal \tilde{B}$.

Let $\tilde{B}^*$ be as defined in (S1.18) and $\tilde{B}'$ be a $m \times n$ matrix with entries $\tilde{B}'_{uv} = (B_{uv} - \hat{\gamma})/\sqrt{m\gamma}$. Then we have $\tilde{B}' = \tilde{B}^* + \alpha\Delta$, where $\Delta = \mathbf{1}_m \mathbf{1}_n^\intercal$ and $\alpha = \frac{\gamma - \hat{\gamma}}{\sqrt{m\gamma}}$. Denote $\tilde{W}^* = \tilde{B}^{*\intercal}\tilde{B}^*$ and $\tilde{W}' = \tilde{B}'^\intercal \tilde{B}'$. Then we let $\big(\lambda_i(\tilde{W}^*), \mu_i^*\big)_{i=1}^n$ be the pairs of eigenvalue and eigenvector of matrix $\tilde{W}^*$ with the eigenvalues in a non-increasing order, namely $\lambda_1(\tilde{W}^*) \geq \lambda_2(\tilde{W}^*) \cdots \geq \lambda_n(\tilde{W}^*)$. Similarly, we let $(\lambda_i(\tilde{W}'), \mu_i')_{i=1}^n$ be the pairs of eigenvalue and eigenvectors of $\tilde{W}'$, where the eigenvalues are in non-increasing order as well.

First let us derive the a lower bound on $\lambda_1(\tilde{W}')$, the largest eigenvalue of $\tilde{W}' = \tilde{B}'^\intercal \tilde{B}'$:

$$
\begin{aligned}
\lambda_1(\tilde{W}') &\geq (\mu_1^*)^\intercal \tilde{B}'^\intercal \tilde{B}' \mu_1^* \\
&= (\mu_1^*)^\intercal \tilde{B}^{*\intercal} \tilde{B}^* \mu_1^* + \alpha (\mu_1^*)^\intercal (\tilde{B}^{*\intercal}\Delta + \Delta^\intercal \tilde{B}^* + \alpha \Delta^\intercal \Delta)\mu_1^* \\
&= \lambda_1(\tilde{W}^*) + \alpha (\mu_1^*)^\intercal (\tilde{B}^{*\intercal}\Delta + \Delta^\intercal \tilde{B}^* + \alpha \Delta^\intercal \Delta)\mu_1^* \\
&\geq \lambda_1(\tilde{W}^*) - |\alpha (\mu_1^*)^\intercal (\tilde{B}^{*\intercal}\Delta + \Delta^\intercal \tilde{B}^* + \alpha \Delta^\intercal \Delta)\mu_1^*|
\end{aligned}
$$

Let $\tilde{B}^* = \sum_{i=1}^n \sqrt{\lambda_i(\tilde{W}^*)} s_i^* \mu_i^{*\intercal}$ be the singular value decomposition of $\tilde{B}^*$. Then we have:

$$
\mu_i^{*\intercal} \tilde{B}^{*\intercal} = \sum_{i=1}^n \sqrt{\lambda_i(\tilde{W}^*)} \mu_1^{*\intercal} \mu_i^* s_i^{*\intercal} = \sqrt{\lambda_i(\tilde{W}^*)} s_i^{*\intercal}
$$

Notice that s_i^*, μ_i^* are the eigenvectors of the matrix $\tilde{B}^*\tilde{B}^{*\intercal}, \tilde{B}^{*\intercal}\tilde{B}^*$ respectively, and we can easily check the conditions in Lemma 4 hold for both matrix $\tilde{B}^*$ and its transpose, thus we have

$$
\begin{aligned}
\alpha \mu_1^{*\intercal} \tilde{B}^{*\intercal} \Delta \mu_1^* &= \alpha \sqrt{\lambda_1(\tilde{W}^*)} s_i^{*\intercal} \mathbf{1}_m \mathbf{1}_n^\intercal \mu_1^* \\
&= o_p(\log n / n^{3/2}) \cdot \tilde{O}_p(1) \cdot \lambda_1(\tilde{W}^*) \\
&= \tilde{O}_p(\log n / n^{3/2}) \cdot O_p(n^{-1/3})
\end{aligned}
$$

Where the $o_p(\log n/n^{3/2})$ term is derived by noticing that $\hat{\gamma}$ is the sample mean of $m \times n$ independent Poisson random variables, and again by the Poisson tail bound (S1.10) we get $\alpha = \frac{\gamma - \hat{\gamma}}{\sqrt{m\gamma}} = o_p(\log n/n^{3/2})$. On the other hand it is easily seen that $\alpha^2\mu_1^{*\mathsf{T}}\Delta^\mathsf{T}\Delta\mu_1^* = \tilde{O}_p(1)\cdot o_p((\log n)^2/n^2)$, which indicates that:

$$\begin{aligned}
\lambda_1(\tilde{W}') &\geq \lambda_1(\tilde{W}^*) - |\alpha(\mu_1^*)^\mathsf{T}(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\mu_1^*| \\
&\geq \lambda_1(\tilde{W}^*) - \tilde{O}_p(\log n/n^{3/2})\cdot O_p(n^{-1/3}) - \tilde{O}_p((\log n)^2/n^2) \\
&= \lambda_1(\tilde{W}^*) - \tilde{O}_P(\log n\cdot n^{-11/6}) \qquad \text{(S1.15)}
\end{aligned}$$

Next, we derive an upper bound for the largest eigenvalue of $\tilde{W}'$. Let $\{a_1,\ldots,a_n\}$ be the coordinates of μ_1', the eigenvector of $\tilde{W}'$ associated with its largest eigenvalue, with respect to the basis consisting of eigenvector of $\tilde{W}^*$, i.e., $\mu_1' = \sum_{i=1}^n a_i\mu_i^*$. Let $\mathcal{S} = \{i\in[n] : \lambda_i(\tilde{W}^*) > \lambda_1(\tilde{W}^*) - 2\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\|\}$, such that $|\mathcal{S}|$ is the number of $\lambda_i(\tilde{W}^*)'$s in the interval $\big(\lambda_1(\tilde{W}^*) - 2\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\|, \lambda_1(\tilde{W}^*)\big)$. Let $\mathbf{v}_1 = \sum_{i=1}^m a_i\mu_i^*$ and $\mathbf{v}_2 = \sum_{i=m+1}^n a_i\mu_i^*$ so that $\mu_1' = \mathbf{v}_1 + \mathbf{v}_2$. we have:

$$\begin{aligned}
\lambda_1(\tilde{W}') &= (\mu_1')^\mathsf{T}\tilde{W}'\mu_1' \\
&= (\mu_1')^\mathsf{T}\tilde{B}^{*\mathsf{T}}\tilde{B}^*\mu_1' + \alpha(\mu_1')^\mathsf{T}(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\mu_1' \\
&\leq \lambda_1(\tilde{W}^*)\cdot\sum_{j\in\mathcal{S}} a_j^2 + \big(\lambda_1(\tilde{W}^*) - 2\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\|\big)\cdot\sum_{j\in([n]/\mathcal{S})} a_j^2 \\
&\quad + 2\mathbf{v}_1^\mathsf{T}|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)|\mathbf{v}_1 + 2\mathbf{v}_2^\mathsf{T}|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)|\mathbf{v}_2 \\
&\leq \lambda_1(\tilde{W}^*) - 2\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\|\sum_{j\in([n]/\mathcal{S})} a_j^2 \\
&\quad + 2m\sum_{j\in\mathcal{S}} a_j^2(\mu_j^*)^\mathsf{T}|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)|\mu_j^* \\
&\quad + 2\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\|\sum_{j\in([n]/\mathcal{S})} a_j^2 \\
&\leq \lambda_1(\tilde{W}^*) + 2|\mathcal{S}|\sum_{j\in\mathcal{S}} a_j^2(\mu_j^*)^\mathsf{T}\big(|2\alpha(\tilde{B}^{*\mathsf{T}}\Delta| + |\alpha^2\Delta^\mathsf{T}\Delta)|\big)\mu_j^* \\
&\leq \lambda_1(\tilde{W}^*) + 2|\mathcal{S}|\sum_{j\in\mathcal{S}} a_j^2\big(\lambda_j(\tilde{W}^*)\cdot\tilde{O}_p(\log n/n^{3/2}) + \tilde{O}_p((\log n)^2/n^2)\big) \\
&\leq \lambda_1(\tilde{W}^*) + 2|\mathcal{S}|\big(O_p(n^{-1/3})\cdot\tilde{O}_p(\log n/n^{3/2}) + \tilde{O}_p((\log n)^2/n^2)\big) \\
&\leq \lambda_1(\tilde{W}^*) + 2|\mathcal{S}|\cdot\tilde{O}_p(\log n\cdot n^{-11/6})
\end{aligned}$$

Now let $a = \lambda_1(\tilde{W}^*) - 2\|\alpha(\tilde{B}'\Delta + \Delta'\tilde{B} + \alpha\Delta'\Delta)\|$ and $b = \lambda_1(\tilde{W}^*)$. We can see that $|\mathcal{S}| = \mathcal{N}(a) - \mathcal{N}(b)$. Noticing that $|(\mathcal{N}(a) - \mathcal{N}(b)) - (\mathcal{N}_m(a) - \mathcal{N}_m(b))| = |(\mathcal{N}(a) - \mathcal{N}_m(a)) - (\mathcal{N}(b) - \mathcal{N}_m(b))|$ and together with Lemma 5 we have that for any $\varepsilon > 0$

$$
\begin{aligned}
&\mathbb{P}(|(\mathcal{N}(a)-\mathcal{N}(b))-(\mathcal{N}_m(a)-\mathcal{N}_m(b))| \geq 2n^{-1}\log(n)^{C_\varepsilon \log\log(n)}) \\
\leq &\mathbb{P}[(\mathcal{N}(a)-\mathcal{N}_m(a)) \geq n^{-1}\log(n)^{C_\varepsilon \log\log(n)}] + \mathbb{P}[(\mathcal{N}(b)-\mathcal{N}_m(b)) \geq n^{-1}\log(n)^{C_\varepsilon \log\log(n)}] \\
\leq &2n^{C_\varepsilon}\exp(-\log(n)^{\varepsilon\log\log(n)})
\end{aligned}
$$

and which indicates that $|\mathcal{S}| = |\mathcal{N}_m(a)-\mathcal{N}_m(b)| + O_p\left(n^{-1}\log(n)^{C_\varepsilon \log\log(N)}\right)$. Since $\mathcal{N}_m(a)-\mathcal{N}_m(b) = n\int_a^b \varrho_m(x)dx$, and it is easily seen by simple calculus that $\varrho_m(x)$ achieves it's local maximum at $x = \frac{(1-n/m)^2}{1+n/m}$, and $\varrho_m(\frac{(1-n/m)^2}{1+n/m}) = \frac{1}{\pi|1-n/m|\sqrt{n/m}}$. Thus we could have the following bound on the size of $|\mathcal{N}_m(a)-\mathcal{N}_m(b)|$:

$$
|\mathcal{N}_m(a)-\mathcal{N}_m(b)| \leq 2n\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\|\frac{1}{\pi|1-\frac{n}{m}|\sqrt{\frac{n}{m}}}
$$

And we could see that

$$
\begin{aligned}
\left\|\alpha(\tilde{B}^{*\mathsf{T}}\Delta + \Delta^\mathsf{T}\tilde{B}^* + \alpha\Delta^\mathsf{T}\Delta)\right\| &\leq |\alpha|\left(\left\|\tilde{B} * \mathsf{T}\Delta\right\| + \left\|\Delta'\tilde{B}^*\right\| + |\alpha|\cdot\|\Delta'\Delta\|\right) \\
&\leq |\alpha|\left(2\left\|\tilde{B}^*\right\|\cdot\|\Delta\|_F + |\alpha|\cdot\|\Delta\|_F^2\right) \\
&= o_p(\log n\cdot n^{-3/2})\left(2\sqrt{mn\lambda_1(\tilde{W}^*)} + mn\cdot o_p(\log n\cdot n^{-3/2})\right) \\
&= o_p(\log n\cdot n^{-5/6})
\end{aligned}
$$

Thus we have that $|\mathcal{S}| = |\mathcal{N}_m(a)-\mathcal{N}_m(b)| + O_p\left(n^{-1}\log(n)^{C_\varepsilon \log\log(n)}\right) \leq o_p(\log n\cdot n^{1/6})$ and

$$
\begin{aligned}
\lambda_1(\tilde{W}') &\leq \lambda_1(\tilde{W}^*) + |\mathcal{S}|\cdot\tilde{O}_P(\log n\cdot n^{-11/6}) \\
&\leq \lambda_1(\tilde{W}^*) + \tilde{O}_p\left((\log n)^2\cdot n^{-17/6}\right) \qquad \text{(S1.16)}
\end{aligned}
$$

Combine (S1.15) and (S1.16) we conclude that

$$
\lambda_1(\tilde{W}') = \lambda_1(\tilde{W}^*) + o_p(n^{-2/3}) \qquad \text{(S1.17)}
$$

Finally we can look at the target matrix $\tilde{B}$. We see that $\tilde{B} = \tilde{B}'\times\sqrt{\frac{\gamma}{\hat{\gamma}}}$ and we can derive from the Poisson tail bond (S1.10) that $\sqrt{\frac{\gamma}{\hat{\gamma}}} = o_p(\log n\cdot n^{-1/2})$. Thus we can also have

$$
\begin{aligned}
\lambda_1(\tilde{W}') &= \sqrt{\frac{\gamma}{\hat{\gamma}}}\lambda_1(\tilde{W}') \\
&= \left(1 + o_p(\log n\cdot n^{-1/2})\right)\left(\lambda_1(\tilde{W}^*) + o_p(n^{-2/3})\right) \\
&= \lambda_1(\tilde{W}^*) + o_p(n^{-2/3})
\end{aligned}
$$

Then by Lemma 3 and Slutsky's theorem, we get the result of Theorem S1. □

Lemma 3. *Let $\tilde{B}^*_{uv}$ be a matrix with entries:*

$$\tilde{B}^*_{uv} := \frac{B_{uv} - \gamma}{\sqrt{m\gamma}}, \ \forall u \in [m], v \in [n] \tag{S1.18}$$

and let $\tilde{W}^* = \tilde{B}^{*\mathsf{T}}\tilde{B}^*$ with $\lambda_1(\tilde{W}^*)$ being its largest eigenvalue. Suppose that $\lim_{n\to\infty} n/m \in (0,\infty)$, then we have, as m,n $\to \infty$:

$$\frac{m\lambda_1(\tilde{W}^*) - (\sqrt{n} + \sqrt{m})^2}{(\sqrt{m} + \sqrt{n})(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}})^{1/3}} \xrightarrow{d} TW_1, \tag{S1.19}$$

Lemma 4 (Theorem 2.8 in **?**)**.** *Let G be an $m \times n$ random matrix with independent entries satisfying*

$$\mathbb{E}G_{uv} = 0, \quad \mathbb{E}|G_{uv}|^2 = \frac{1}{\sqrt{nm}}.$$

Assume that m and n satisfy the bounds $n^{1/C} \leq m \leq n^C$ for some $C > 0$. Suppose for all $p \in \mathbb{N}$, there exist C_p such that

$$\mathbb{E}|(mn)^{1/4} G_{uv}|^p \leq C_p$$

Let $\boldsymbol{\mu}_i$ be the eigenvalue of $G^\mathsf{T} G$ associated with its i_{th} largest eigenvalue. Then for any $\varepsilon > 0$ we have

$$|\boldsymbol{\mu}_i^\mathsf{T} \mathbf{v}|^2 = \tilde{O}_P(1/n)$$

uniformly for all $i \leq (1-\varepsilon)\min(m,n)$ and any deterministic unit vector $\mathbf{v} \in \mathbb{R}^n$.

Lemma 5 (Theorem 3.3 in **?**)**.** *Let $\xi_\pm = (1 \pm \sqrt{\frac{n}{m}})^2$, and denote the Marchenko-Pastur law by ϱ_m, which is given by*

$$\varrho_m(x) = \frac{m}{2\pi n}\sqrt{\frac{[(\xi_+ - x)(x - \xi_-)]_+}{x^2}}$$

Let $\beta \in \mathbb{R}$, define the empirical spectral distribution of $(\tilde{B}^*)^\mathsf{T}\tilde{B}^*$ by:

$$\mathcal{N}(\beta) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[\beta,\infty)}\big(\lambda_i(\tilde{W}^*)\big)$$

And the distribution given by the Marchenko-Pastur law:

$$\mathcal{N}_m(\beta) := \int_\beta^\infty \varrho_m(x)dx$$

If $\lim_{n\to\infty} \frac{n}{m} \in (0,\infty) \setminus \{1\}$, then for any $\varepsilon > 0$, there exists a constant C_ε such that:

$$\mathbb{P}(|\mathcal{N}(\beta) - \mathcal{N}_m(\beta)| \geq n^{-1}\log(n)^{C_\varepsilon \log\log(n)}) \leq n^{C_\varepsilon}\exp(-\log(n)^{\varepsilon \log\log(n)})$$

S2. Supplementary material for Section 4

For $r \in [R]$ and $\ell \in [2^r]$, define

$$\delta^{(r,\ell)} := \frac{1}{2} \frac{\int_{I_\ell^{(r)}} \lambda_a - \lambda_b \, d\nu}{\int_{I_\ell^{(r)}} \lambda_a + \lambda_b \, d\nu}, \tag{S2.20}$$

so that $\frac{\int_{I_\ell^{(r)}} \lambda_a \, d\nu}{\int_{I_\ell^{(r)}} \lambda \, d\nu} = \frac{1}{2} + \delta^{(r,\ell)}$ and $\frac{\int_{I_\ell^{(r)}} \lambda_b \, d\nu}{\int_{I_\ell^{(r)}} \lambda \, d\nu} = \frac{1}{2} - \delta^{(r,\ell)}$.

Proposition 3. *For every $r \in [R]$ and $l \in [2^r]$, $N_a^{(r,\ell)} \sim Bin\big(\frac{1}{2} + \delta^{(r,\ell)}, N^{(r,\ell)}\big)$ conditional on $N^{(r,\ell)}$ and if $\delta^{(r,\ell)} = 0$, then $p^{(r,\ell)} \sim Unif[0,1]$.*

Moreover, for every $r \in [R]$, conditional on $\{N^{(r,\ell)}\}_{\ell \in [2^r]}$, the collection of random variables $\{N_a^{(r,\ell)}\}_{l \in [2^r]}$ are mutually independent.

Proof. Since $N_a^{(r,\ell)} = N_a(I_\ell^{(r)})$, we have that $N_a^{(r,\ell)}$ has the Poisson distribution with mean $\int_{I_l^{(r)}} \lambda_a d\nu$. Since $N^{(r,\ell)} - N_a^{(r,\ell)} = N_b^{(r,\ell)}$ has the Poisson distribution with mean $\int_{I_l^{(r)}} \lambda_b d\nu$, and is independent of $N_a^{(r,\ell)}$, we have that, for any $s, t \in \mathbb{N}$ where $s \leq t$,

$$\begin{aligned}
\mathbb{P}(N_a^{(r,\ell)} = s \mid N^{(r,\ell)} = t) &= \frac{\mathbb{P}(N_a^{(r,\ell)} = s, N_b^{(r,\ell)} = t - s)}{\mathbb{P}(N^{(r,\ell)} = t)} \\
&= \frac{\frac{1}{s!} e^{-\int_{I_l^{(r)}} \lambda_a d\nu} \left\{\int_{I_l^{(r)}} \lambda_a d\nu\right\}^s \frac{1}{(t-s)!} e^{-\int_{I_l^{(r)}} \lambda_b d\nu} \left\{\int_{I_l^{(r)}} \lambda_b d\nu\right\}^{t-s}}{\frac{1}{t!} e^{-\int_{I_l^{(r)}} \lambda_a + \lambda_b d\nu} \left\{\int_{I_l^{(r)}} \lambda_a + \lambda_b d\nu\right\}^t} \\
&= \binom{t}{s} \left(\frac{1}{2} + \delta^{(r,\ell)}\right)^s \left(\frac{1}{2} - \delta^{(r,\ell)}\right)^{t-s},
\end{aligned}$$

and the first claim follows directly. If $\delta_l^{(k)} = 0$, then $\hat{p}_l^{(k)}$ is uniform by Proposition 1.

The second claim follows from the independent increment property of a Poisson process. □

S2.1 Proof of Theorem 3

Proof. (of Theorem 3)

Let $r^* \in [R]$ denote the resolution level that satisfies (4.28). Recalling that $\delta_l^{(r^*)}$ is defined as (S2.20), we define the event

$$\mathcal{E}_{r^*} := \left\{ \sum_{l=1}^{2^{r^*}} (N^{(r^*,l)} - 1) \delta_l^{(r^*)2} \geq 2^{r^*/2} \left(\frac{C^{1/2}}{\beta^{1/2}} + 2 \log^{1/2} \frac{R}{\alpha} \right) + 2 \log \frac{R}{\alpha} \right\},$$

where C is the universal constant specified in Theorem S2 which we may assume to be greater than 1. Then, by Theorem S2,

$$\begin{aligned}\mathbb{P}\big(p_F \geq \alpha\big) &\leq \mathbb{P}\big(p_F^{(r^*)} \geq \frac{\alpha}{R}\big)\\ &\leq \mathbb{P}\bigg(\Big\{p_F^{(r^*)} \geq \frac{\alpha}{R}\Big\} \cap \mathcal{E}_{r^*}\bigg) + \mathbb{P}(\mathcal{E}_{r^*}^c) \leq \beta + \mathbb{P}(\mathcal{E}_{r^*}^c).\end{aligned}$$

In order to upper bound the probability of $\mathcal{E}_{r^*}^c$, we observe, by our assumption that $\mathbb{E}N^{(r^*,l)} = \int_{I_l^{(r^*)}} \lambda d\nu \geq 2$ for all $l \in [2^{r^*}]$ and the fact that $|\delta_l^{(r^*)2}| \leq \frac{1}{2}$, that

$$\mathbb{E}\sum_{l=1}^{2^{r^*}}(N^{(r^*,l)} - 1)\delta_l^{(r^*)2} \geq \frac{1}{2}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2} \quad \text{and} \tag{S2.21}$$

$$\mathrm{Var}\sum_{l=1}^{2^{r^*}}(N^{(r^*,l)} - 1)\delta_l^{(r^*)2} \leq \frac{1}{4}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2}. \tag{S2.22}$$

As a short hand, we write

$$\begin{aligned}W &\equiv \sum_{l=1}^{2^{r^*}}(N^{(r^*,l)} - 1)\delta_l^{(r^*)2}, \quad \text{and}\\ T_{r^*,R,\alpha,\beta} &\equiv 2^{r^*/2}\big(\frac{C^{1/2}}{\beta^{1/2}} + 2\log^{1/2}\frac{R}{\alpha}\big) + 2\log\frac{R}{\alpha}.\end{aligned}$$

We note that by (4.28), we have $\mathbb{E}W \geq \frac{1}{2}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big) \geq T_{r^*,R,\alpha,\beta}$. By this, Chebyshev' inequality, and (S2.21) and (S2.22), we have

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{r^*}^c) &= \mathbb{P}\Big\{\sum_{l=1}^{2^{r^*}}(N^{(r^*,l)} - 1)\delta_l^{(r^*)2} \leq T_{r^*,R,\alpha,\beta}\Big\}\\ &= \mathbb{P}\big\{W \leq T_{r^*,R,\alpha,\beta}\big\}\\ &= \mathbb{P}\big\{W - \mathbb{E}W \leq T_{r^*,R,\alpha,\beta} - \mathbb{E}W\big\}\\ &\leq \mathrm{Var}(W) \cdot \{\mathbb{E}W - T_{r^*,R,\alpha,\beta}\}^{-2}\\ &\leq \mathrm{Var}\Big\{\sum_{l=1}^{2^{r^*}}(N^{(r^*,l)} - 1)\delta_l^{(r^*)2}\Big\}\Big\{\mathbb{E}\Big(\sum_{l=1}^{2^{r^*}}(N^{(r^*,l)} - 1)\delta_l^{(r^*)2}\Big) - T_{r^*,R,\alpha,\beta}\Big\}^{-2}\\ &\leq \Big\{\frac{1}{4}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2}\Big\}\Big\{\frac{1}{2}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2} - T_{r^*,R,\alpha,\beta}\Big\}^{-2}\\ &\leq \Big\{\frac{1}{4}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2}\Big\}\Big\{\frac{1}{4}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2}\Big\}^{-2}\\ &\leq \Big\{\frac{1}{4}\sum_{l=1}^{2^{r^*}}\Big(\int_{I_l^{(r^*)}} \lambda d\nu\Big)\delta_l^{(r^*)2}\Big\}^{-1} \leq \beta,\end{aligned}$$

where the penultimate inequality follows from (4.28) and the fact that $C \geq 1$. The Theorem follows as desired.

□

S2.2 Proof of Theorem 4

Proof. (of Theorem 4)

We first claim that, writing C as the universal constant specified in Theorem 3,

$$\frac{1}{4n}\int_I\left(\frac{\lambda_a-\lambda_b}{\lambda}\right)^2\lambda\,d\nu \geq \min_{r\in[R]}\frac{2^{r/2}}{n}\left(\frac{C^{1/2}}{\beta}+2\log^{1/2}\frac{R}{\alpha}\right)+\frac{2}{n}\log\frac{R}{\alpha}+\frac{C_HC_d}{2}2^{-\frac{r\gamma}{q}}. \tag{S2.23}$$

To see that this claim is true, define

$$\tilde{r}=\min\left(R,\left\lfloor\frac{\log_2\frac{n}{2}}{\frac{1}{2}+\frac{2\gamma}{q}}-\log_2\frac{c_{\max}}{c_{\min}}\right\rfloor\right),$$

or equivalently,

$$\tilde{r}=\begin{cases}\lfloor\frac{\log_2\frac{n}{2}}{\frac{1}{2}+\frac{2\gamma}{q}}-\log_2\frac{c_{\max}}{c_{\min}}\rfloor & \text{if } \gamma/q\geq 1/4\\ R & \text{if } \gamma/q\leq 1/4\end{cases}.$$

Then, using the fact that $\log_2\frac{n}{2}-\log_2\frac{c_{\max}}{c_{\min}}-1\leq R\leq\log_2\frac{n}{2}-\log_2\frac{c_{\max}}{c_{\min}}$, we have

$$\begin{aligned}&\frac{2^{\tilde{r}/2}}{n}\left(\frac{C^{1/2}}{\beta}+2\log^{1/2}\frac{R}{\alpha}\right)+\frac{2}{n}\log\frac{R}{\alpha}+\frac{C_HC_d}{2}2^{-\frac{2\tilde{r}\gamma}{q}}\\ &\leq\begin{cases}\frac{1}{4}C_1n^{-\frac{4\gamma}{q+4\gamma}}\left(\beta^{-1}+\log\frac{\log n}{\alpha}\right) & \text{if } \gamma/q\geq 1/4\\ \frac{1}{4}C_1n^{-\frac{2\gamma}{q}}\left(\beta^{-1}+\log\frac{\log n}{\alpha}\right) & \text{if } \gamma/q\leq 1/4\end{cases}\end{aligned}$$

for some $C_1>0$ whose value depends only on $\frac{c_{\max}}{c_{\min}}$, C_H, and C_d. Therefore, we have from assumption (4.29) that claim (S2.23) holds. Then, by Lemma 8, we have that for every $r\in[R]$,

$$\begin{aligned}&\frac{1}{n}\sum_{l=1}^{2^r}\left(\frac{\int_{I_l^{(r)}}\lambda_a-\lambda_b d\nu}{\int_{I_l^{(r)}}\lambda d\nu}\right)^2\int_{I_l^{(r)}}\lambda d\nu\\ &\geq\frac{1}{n}\int_I\left(\frac{\lambda_a-\lambda_b}{\lambda}\right)^2\lambda\,d\nu-\frac{C_HC_d}{2}2^{-\frac{2\gamma r}{q}}.\end{aligned}$$

Thus, using (S2.23), we may conclude that there exists a $r\in[R]$ such that

$$\frac{1}{4n}\sum_{l=1}^{2^r}\left(\frac{\int_{I_l^{(r)}}\lambda_a-\lambda_b d\nu}{2\int_{I_l^{(r)}}\lambda d\nu}\right)^2\int_{I_l^{(r)}}\lambda d\nu\geq\frac{2^{r/2}}{n}\left(\frac{C^{1/2}}{\beta}+2\log^{1/2}\frac{R}{\alpha}\right)+\frac{2}{n}\log\frac{R}{\alpha}.$$

From the hypothesis of the theorem, we also have that for all $l\in[2^R]$,

$$\int_{I_l^{(R)}}\lambda d\nu=n\frac{\int_{I_l^{(R)}}\lambda\,\nu}{\int_I\lambda\,d\nu}\geq n\frac{c_{\min}}{c_{\max}}\frac{\nu(I_l^{(R)})}{\nu(I)}\geq n\frac{c_{\min}}{c_{\max}}2^{-R}\geq 2,$$

where, in the final inequality, we use the assumption that $R\leq\log_2\frac{n}{2}-\log_2\frac{c_{\max}}{c_{\min}}$.

Then, from Theorem 3, it holds that $\mathbb{P}(p_F\leq\alpha)\geq 1-2\beta$ and the conclusion of the Theorem follows as desired.

□

S2.3 Proof of Theorem 5

Proof. (of Theorem 5)

The proof is similar to that of Theorem 3. Let $r^* \in [R]$ and $l^* \in [2^{r^*}]$ denote the resolution level and bin such that

$$\frac{1}{4}\left(\frac{\int_{I_{l^*}^{(r^*)}} \lambda_a - \lambda_b d\nu}{2\int_{I_{l^*}^{(r^*)}} \lambda d\nu}\right)^2 \int_{I_{l^*}^{(r^*)}} \lambda d\nu \geq 2r^* + \frac{C^{1/2}}{\beta} + 2\log\frac{K}{\alpha}.$$

Define the event

$$\mathcal{E}_{r^*l^*} := \left\{(m_{l^*}^{(r^*)} - 1)\delta_{l^*}^{(r^*)2} \geq 2r^* + \frac{C^{1/2}}{\beta^{1/2}} + 2\log\frac{K}{\alpha}\right\}.$$

By Theorem S3, we have that

$$\begin{aligned}
\mathbb{P}(\hat{p}_{\min} \geq \frac{\alpha}{K}) &\leq \mathbb{P}(\hat{p}^{(k)} \geq \frac{\alpha}{K}) \\
&\leq \mathbb{P}\left(\left\{\hat{p}^{(k)} \geq \frac{\alpha}{K}\right\} \cap \mathcal{E}_{r^*l^*}\right) + \mathbb{P}(\mathcal{E}_{r^*l^*}^c) \\
&\leq \beta + \mathbb{P}(\mathcal{E}_{r^*l^*}^c).
\end{aligned}$$

To bound $\mathbb{P}(\mathcal{E}_{r^*l^*}^c)$, we use our assumption that $\mathbb{E}m_{l^*}^{(r^*)} = \int_{I_{l^*}^{(r^*)}} \lambda d\nu \geq 2$ and the fact that $\delta_{l^*}^{(r^*)} \leq 1$ to obtain

$$\begin{aligned}
\mathbb{E}(m_{l^*}^{(r^*)} - 1)\delta_{l^*}^{(r^*)2} &\geq \frac{1}{2}\left(\int_{I_{l^*}^{(r^*)}} \lambda d\nu\right)\delta_{l^*}^{(r^*)2} \qquad \text{and} \\
\mathrm{Var}(m_{l^*}^{(r^*)} - 1)\delta_{l^*}^{(r^*)2} &\leq \left(\int_{I_{l^*}^{(r^*)}} \lambda d\nu\right)\delta_{l^*}^{(r^*)2}.
\end{aligned}$$

We have then

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_{l^*r^*}^c) &= \mathbb{P}\left((m_{l^*}^{(r^*)} - 1)\delta_{l^*}^{(r^*)2} \leq \frac{C^{1/2}}{\beta^{1/2}} + 2r^* + 2\log\frac{K}{\alpha}\right) \\
&\leq \left\{\mathrm{Var}(m_{l^*}^{(r^*)} - 1)\delta_{l^*}^{(r^*)2}\right\}\left\{\mathbb{E}(m_{l^*}^{(r^*)} - 1)\delta_{l^*}^{(r^*)2} - \left(2r^* + \frac{C^{1/2}}{\beta^{1/2}} + 2\log\frac{K}{\alpha}\right)\right\}^{-2} \\
&\leq \left\{\frac{1}{4}\left(\int_{I_{l^*}^{(r^*)}} \lambda d\nu\right)\delta_{l^*}^{(r^*)2}\right\}^{-1} \leq \beta.
\end{aligned}$$

□

S2.4 Proof of Theorem 6

Proof. Let $r := \lceil \log_2 \frac{\nu(I)}{\nu(S)} \rceil$ so that

$$\frac{\nu(I)}{2^{r-1}} \geq \nu(S) \geq \frac{\nu(I)}{2^r}.$$

We observe that since $\frac{\nu(S)}{\nu(I)} \geq \frac{c_{\max}}{c_{\min}} \frac{8}{n}$ by assumption,

$$r \leq \lceil \log_2 \frac{\nu(I)}{\nu(S)} \rceil < \lfloor \log_2 \frac{n}{2} - \log_2 \frac{c_{\max}}{c_{\min}} \rfloor = R.$$

Hence, $\{I_l^{(r+1)}\}$ exists in our dyadic partitioning and there exists $l^* \in [2^{r+1}]$ such that the interval $I_{l^*}^{(r+1)} \subset S$. Let C be the universal constant specified in Theorem 5 and let $C_2 := \frac{32 c_{\max}}{c_{\min}} C^{1/2}$. From (4.30), we have that

$$\begin{aligned}
&\frac{1}{4} \max_{l \in [2^{r+1}]} \left(\frac{\int_{I_l^{(r+1)}} \lambda_a - \lambda_b d\nu}{\int_{I_l^{(r+1)}} \lambda d\nu} \right)^2 \int_{I_l^{(r+1)}} \lambda d\nu \\
&\geq \frac{1}{4} \left(\frac{\int_{I_{l^*}^{(r+1)}} \frac{\lambda_a - \lambda_b}{\lambda} \lambda d\nu}{\int_{I_{l^*}^{(r+1)}} \lambda d\nu} \right)^2 \int_{I_{l^*}^{(r+1)}} \lambda d\nu \\
&\geq \frac{1}{4} \delta_S^2 \int_{I_{l^*}^{(r+1)}} \lambda d\nu \overset{(a)}{\geq} \frac{1}{4} \delta_S^2 n \frac{c_{\min}}{c_{\max}} 2^{-(r+1)} \\
&\geq \delta^2 n 2^{-(r-1)} \frac{c_{\min}}{16 c_{\max}} \geq \delta_S^2 n \frac{\nu(S)}{\nu(I)} \frac{c_{\min}}{16 c_{\max}} \geq n \delta_S^2 \frac{\nu(S)}{\nu(I)} \frac{2C^{1/2}}{C_2} \\
&\geq 2 \log n + \frac{C^{1/2}}{\beta} + 2 \log \frac{1}{\alpha} \geq 2r + \frac{C^{1/2}}{\beta} + 2 \log \frac{R}{\alpha},
\end{aligned}$$

where inequality (a) follows from the fact that

$$\int_{I_{l^*}^{(r+1)}} \lambda d\nu = \frac{\int_{I_{l^*}^{(r+1)}} \lambda d\nu}{\int_I \lambda d\nu} \geq n \frac{c_{min}}{c_{max}} 2^{-(r+1)}.$$

The conclusion of the theorem follows from Theorem 5.

□

S2.5 Auxiliary results

Recall that, for a positive integer m, we define

$$S_{\text{Bin}(\frac{1}{2},m)}(t) := \mathbb{P}(|\text{Bin}(\frac{1}{2}, m) - \frac{m}{2}| \geq t) \tag{S2.24}$$

$$S_{\chi_m^2}(t) := \mathbb{P}(\chi_m^2 \geq t). \tag{S2.25}$$

Moreover, define

$$\mathcal{M}_m := \left\{ -\frac{m}{2}, -\frac{m}{2} + 1, \ldots, \frac{m}{2} - 1, \frac{m}{2} \right\} \tag{S2.26}$$

$$\mathcal{M}_m^+ := \begin{cases} \{0, 1, \ldots, \frac{m}{2}\} & \text{if } m \text{ is even,} \\ \{\frac{1}{2}, \frac{3}{2}, \ldots, \frac{m}{2}\} & \text{if } m \text{ is odd.} \end{cases} \tag{S2.27}$$

Theorem S2. *Let d be a positive integer. For each $l \in [d]$, let $m_l \in \mathbb{N}$, $\delta_l \in [0, \frac{1}{2}]$, and let $A_1, \ldots, A_d$ be independent random variables where $A_l \sim Bin\big(\frac{1}{2} + \delta_l, m_l\big)$,*

Let $U_1, \ldots, U_d$ be independent random variables distributed uniform on $[0,1]$ and independent of $A_1, \ldots, A_d$. Define $p_l := U_l \cdot S_{Bin(\frac{1}{2}, m_l)}(|A_l - \frac{m_l}{2}|) + (1 - U_l) S_{Bin(\frac{1}{2}, m_l)}(|A_l - \frac{m_l}{2}| + 1)$, define the set $L := \{l \in [d] \,:\, m_l \geq 2\}$, and define $p := S_{\chi^2_{2|L|}}\big(\sum_{l \in L} -2\log p_l\big)$.

Then, there exists a universal constant $C > 0$ such that, for any $\alpha, \beta \in (0,1)$, if

$$\sum_{l \in L} (m_l - 1)\delta_l^2 \geq |L|^{1/2} \left(\frac{C^{1/2}}{\beta^{1/2}} + 2\log^{1/2} \frac{1}{\alpha} \right) + 2 \log \frac{1}{\alpha}, \tag{S2.28}$$

then $\mathbb{P}(p \leq \alpha) \geq 1 - \beta$.

Proof. Define

$$L_1 := \{l \in L \,:\, (m_l - 1)\delta_l^2 \geq 2\} \quad \text{and} \quad L_2 := \{l \in L \,:\, (m_l - 1)\delta_l^2 < 2\}.$$

For simplicity of presentation, we write $Z_l := -2\log p_l$ and $\tilde{Z}_l := \frac{4}{m_l}\big(A_l - \frac{m_l}{2}\big)^2$ for $l \in L$. By Hoeffding's inequality, it holds that $S_{\mathrm{Bin}(\frac{1}{2}, m)}(t) \leq 2\exp\{-2\frac{t^2}{m}\}$. Therefore, we have that

$$\begin{aligned} Z_l &= -2\log\left\{ U_l \cdot S_{\mathrm{Bin}(\frac{1}{2}, m)}(|A_l - \frac{m_l}{2}|) + (1 - U_l) S_{\mathrm{Bin}(\frac{1}{2}, m)}\left(|A_l - \frac{m_l}{2}| + 1 \right) \right\} \\ &\geq -2\log\left\{ S_{\mathrm{Bin}(\frac{1}{2}, m)}(|A_l - \frac{m_l}{2}|) \right\} \geq \frac{4}{m_l}\left(A_l - \frac{m_l}{2} \right)^2 - 2\log 2 = \tilde{Z}_l - 2\log 2. \end{aligned}$$

By Lemma 9, we have

$$S_{\chi^2_{2|L|}}\left(2|L| + \sqrt{8}|L|^{1/2} \log^{1/2} \frac{1}{\alpha} + 2\log\frac{1}{\alpha} \right) \leq \alpha, \tag{S2.29}$$

By (S2.29), the fact that $S_{\chi^2_{2|L|}}(\cdot)$ is monotone decreasing, and the fact that $2\log 2 \leq 2$, we have

$$\begin{aligned} \mathbb{P}(p \geq \alpha) &= \mathbb{P}\left\{ S_{\chi^2_{2|L|}}\left(\sum_{l \in L} -2\log p_l \right) \geq \alpha \right\} \\ &\leq \mathbb{P}\left\{ \sum_{l \in L} -2\log p_l \leq 2|L| + \sqrt{8}|L|^{\frac{1}{2}} \log^{\frac{1}{2}} \frac{1}{\alpha} + 2\log\frac{1}{\alpha} \right\} \\ &\leq \mathbb{P}\left\{ \sum_{l \in L_1} (\tilde{Z}_l - 2\log 2) + \sum_{l \in L_2} Z_l \leq 2|L| + \sqrt{8}|L|^{\frac{1}{2}} \log^{\frac{1}{2}} \frac{1}{\alpha} + 2\log\frac{1}{\alpha} \right\} \\ &\leq \mathbb{P}\left\{ \sum_{l \in L_1} (\tilde{Z}_l - \mathbb{E}\tilde{Z}_l) + \sum_{l \in L_2} (Z_l - \mathbb{E}Z_l) \leq \sum_{l \in L_1} (4 - \mathbb{E}\tilde{Z}_l) \right. \\ &\qquad \left. + \sum_{l \in L_2} (2 - \mathbb{E}Z_l) + 2|L|^{\frac{1}{2}} \log^{\frac{1}{2}} \frac{1}{\alpha} + 2\log\frac{1}{\alpha} \right\}. \quad (\star) \end{aligned}$$

We now observe that by Lemma 7,

$$\sum_{l\in L_1}(4-\mathbb{E}\tilde{Z}_l)+\sum_{l\in L_2}(2-\mathbb{E}Z_l)+2|L|^{1/2}\log^{1/2}\frac{1}{\alpha}+2\log\frac{1}{\alpha}$$
$$\leq -\sum_{l\in L}2(m_l-1)\delta_l^2+2|L|^{1/2}\log^{1/2}\frac{1}{\alpha}+2\log\frac{1}{\alpha}\leq 0,$$

where the final inequality follows by our assumption (S2.28). Therefore, returning to ($\star$), we may apply Chebyshev's inequality to obtain

$$(\star)\leq\frac{\sum_{l\in L_1}\mathrm{Var}\tilde{Z}_l+\sum_{l\in L_2}\mathrm{Var}Z_l}{\Big\{-\sum_{l\in L}2(m_l-1)\delta_l^2+2|L|^{\frac{1}{2}}\log^{\frac{1}{2}}\frac{1}{\alpha}+2\log\frac{1}{\alpha}\Big\}^2}$$
$$\leq Cd\Big\{-2\sum_{l\in L}(m_l-1)\delta_l^2+2|L|^{\frac{1}{2}}\log^{\frac{1}{2}}\frac{1}{\alpha}+2\log\frac{1}{\alpha}\Big\}^{-2}\leq\beta.$$

where, in the penultimate inequality, we used Lemmas 6 and where $C>0$ is the universal constant specified in Lemma 6. The conclusion of the Theorem follows as desired.

□

Theorem S3. *Let d be a positive integer. For each $l\in[d]$, let $m_l\in\mathbb{N}$, $\delta_l\in[0,1/2]$, and let $A_1,\ldots,A_d$ be independent random variables where $A_l\sim Bin\big(\frac{1}{2}+\delta_l,m_l\big)$.*

Let $U_1,\ldots,U_d$ be independent random variables distributed uniform on $[0,1]$ and independent of $A_1,\ldots,A_d$. Define $p_l:=U_l\cdot S_{Bin(\frac{1}{2},m_l)}(|A_l-\frac{m_l}{2}|)+(1-U_l)S_{Bin(\frac{1}{2},m_l)}(|A_l-\frac{m_l}{2}|+1)$, define the set $L:=\{l\in[d]:m_l\geq 2\}$, and define $p_{min}:=F_{Beta,|L|}\big(\min_{l\in L}p_l\big)$ where $F_{Beta,|L|}(x):=\mathbb{P}(Beta(1,|L|+1)\leq x)$ for any $x\in\mathbb{R}$.

There exists universal constants $C>0$ such that for any $\alpha,\beta\in(0,1)$, if

$$\max_{l\in[d]}(m_l-1)\delta_l^2\geq\frac{C^{1/2}}{\beta^{1/2}}+2\log\frac{|L|}{\alpha},$$

then $\mathbb{P}(p_{min}\leq\alpha)\geq 1-\beta$.

Proof. Let C be the maximum of 4 and the universal constant specified in Lemma 6. By assumption, there exists $l^*\in[d]$ be such that

$$(m_{l^*}-1)\delta_{l^*}^2\geq\frac{C^{1/2}}{\beta^{1/2}}+2\log\frac{|L|}{\alpha}\geq 2,\tag{S2.30}$$

where the last inequality follows since $\beta,\alpha\in(0,1)$.

By Hoeffding's inequality, we have that

$$\begin{aligned}-2\log p_{l^*} &\geq -2\log S_{\mathrm{Bin}(\frac{1}{2},m_{l^*})}(|A_{l^*}-\frac{m_{l^*}}{2}|)\\ &\geq \frac{4}{m_{l^*}}\left(A_{l^*}-\frac{m_{l^*}}{2}\right)^2-2\log 2.\end{aligned}$$

We write $\tilde{Z}_{l^*} := \frac{4}{m_{l^*}}\big(A_{l^*}-\frac{m_{l^*}}{2}\big)^2$ so that $-2\log p_{l^*} \geq \tilde{Z}_{l^*}-2\log 2 \geq \tilde{Z}_{l^*}-4$.

For any $\alpha,\beta\in(0,1)$, we may use the fact that $F_{\mathrm{Beta},|L|}(x)\leq |L|x$ to show that

$$\begin{aligned}\mathbb{P}(p_{\min}\geq\alpha) &\leq \mathbb{P}\big(\min_{l\in L} p_l \geq \frac{\alpha}{|L|}\big) \leq \mathbb{P}\big(p_{l^*}\geq\frac{\alpha}{|L|}\big)\\ &= \mathbb{P}\big(-2\log p_{l^*}\leq 2\log\frac{|L|}{\alpha}\big)\\ &\leq \mathbb{P}\left(\tilde{Z}_{l^*}-\mathbb{E}\tilde{Z}_{l^*}\leq(4-\mathbb{E}\tilde{Z}_{l^*})+2\log\frac{|L|}{\alpha}\right). \quad (\star)\end{aligned}$$

By Lemma 7 and (S2.30), we have that

$$(4-\mathbb{E}\tilde{Z}_{l^*})+2\log\frac{|L|}{\alpha}\leq -2(m_{l^*}-1)\delta_{l^*}^2+2\log\frac{|L|}{\alpha}\leq 0.$$

Therefore, continuing on from $(\star)$, we have by Chebyshev inequality and Lemma 6 that

$$(\star)\leq\frac{\mathrm{Var}(\tilde{Z}_{l^*})}{\big\{-(m_{l^*}-1)\delta_{l^*}^2+2\log\frac{|L|}{\alpha}\big\}^2}\leq C\left\{-(m_{l^*}-1)\delta_{l^*}^2+2\log\frac{|L|}{\alpha}\right\}^{-2}\leq\beta.$$

The conclusion of the Theorem follows as desired.

□

Lemma 6. *Let m be a positive integer and let $\delta\in[0,1/2]$. Let $A\sim Bin(\frac{1}{2}+\delta,m)$ and let $U\sim Unif[0,1]$ be independent of A. Define $Z:=-2\log\{U\cdot S_{Bin(\frac{1}{2},m)}(|A-\frac{m}{2}|)+(1-U)S_{Bin(\frac{1}{2},m)}(|A-\frac{m}{2}|+1)\}$ and $\tilde{Z}:=\frac{4}{m}(A-\frac{m}{2})^2$.*

There exists a universal constant $C>0$ such that

1. *if $(m-1)\delta^2<2$, then $Var(Z)\leq C$,*
2. *and $Var(\tilde{Z})\leq C$.*

Proof. First assume that $(m-1)\delta^2<2$. By increasing the value of the universal constant C if necessary, we may assume without the loss of generality that $m\geq 17$.

Define $\mathcal{M}_m$ as (S2.26). Let P,Q be probability measures on $\mathcal{M}_m$ such that P is the distribution of $|A-\frac{m}{2}|$ and Q is the distribution of $|\mathrm{Bin}(\frac{1}{2},m)-\frac{m}{2}|$.

Then, let $\tilde{S}_0(\cdot)$ be defined as in Lemma 10, we have by the same lemma that

$$
\begin{aligned}
\mathrm{Var} Z &\le \mathbb{E} Z^2 \\
&= \int_0^1 \sum_{s \in \mathcal{M}_m} 4 \log^2 \tilde{S}_0(|s| + u) \cdot \frac{P(s)}{Q(s)} Q(s)\, du \\
&\le \underbrace{\left\{ \sum_{s \in \mathcal{M}_m} \left(\frac{P(s)}{Q(s)} \right)^2 Q(s) \right\}^{1/2}}_{\text{Term 1}} \qquad \text{(S2.31)} \\
&\quad + \underbrace{\left\{ \int_0^1 \sum_{s \in \mathcal{M}_m} \Big(4 \log^2 \big\{ \tilde{S}_0(|s| + u) \big\} \Big)^2 Q(s)\, du \right\}^{1/2}}_{\text{Term 2}}. \qquad \text{(S2.32)}
\end{aligned}
$$

Term 2 of (S2.32) is equal to $16 \cdot \mathbb{E} \log^4 \tilde{S}_0(|\mathrm{Bin}(\frac{1}{2}, m) - \frac{m}{2}| + U)$. Since the random variable $\tilde{S}_0(|\mathrm{Bin}(\frac{1}{2}, m) - \frac{m}{2}| + U)$ is uniformly distributed on $[0, 1]$, we have that Term 2 is upper bounded by a universal constant.

For Term 1, we define $r := \frac{1+2\delta}{1-2\delta}$ and observe that for any $s \in \mathcal{M}_m$,

$$
\begin{aligned}
\frac{P(s)}{Q(s)} &= \frac{1}{2} 2^m \left\{ \left(\frac{1}{2} + \delta \right)^{\frac{m}{2} + s} \left(\frac{1}{2} - \delta \right)^{\frac{m}{2} - s} + \left(\frac{1}{2} + \delta \right)^{\frac{m}{2} - s} \left(\frac{1}{2} - \delta \right)^{\frac{m}{2} + s} \right\} \\
&= \frac{1}{2} (1 - 4\delta^2)^{\frac{m}{2}} (r^s + r^{-s}) \le r^s.
\end{aligned}
$$

Since we assume $(m-1)\delta^2 \le 2$ and since we assume that $m \ge 17$, we have that $\delta^2 \le \frac{1}{8}$ and thus $0 \le \log r \le 8\delta$. Let W be a random variable distributed as $|\mathrm{Bin}(\frac{1}{2}, m) - \frac{m}{2}|$. Then

$$
\begin{aligned}
\sum_{s \in \mathcal{M}} r^{2s} Q(s) = \mathbb{E} r^{2W} &= \int_1^\infty \mathbb{P}(r^{2W} \ge t)\, dt \\
&= \int_1^\infty \mathbb{P}\left(W \ge \frac{\log t}{2 \log r} \right) dt \\
&\le \int_1^\infty \exp\left(-\frac{\log^2 t}{4m \log^2 r} \right) dt \\
&\le \int_1^\infty t^{-\frac{\log t}{2^{12}}}\, dt \le C,
\end{aligned}
$$

where $C > 0$ is a universal constant.

Now we turn to the second claim. Write $A = \sum_{i=1}^m \epsilon_i$ where $\epsilon_1, \ldots, \epsilon_m$ are independent and identically distributed $\mathrm{Ber}(\frac{1}{2} + \delta)$ random variables.

For any $i \in [m]$, we have

$$
\mathrm{Var}_{\cdot \mid \{\epsilon_{-i}\}} \left[\left(A - \frac{m}{2} \right)^2 \right] \le \sup_{z \in [-\frac{m}{2}, \frac{m}{2}]} \mathrm{Var}\big[(z + \epsilon_i)^2 \big] \le m.
$$

Thus, by the Efron–Stein inequality,

$$\begin{aligned}\mathrm{Var}\tilde{Z} &= \frac{16}{m^2}\mathrm{Var}\left[\left(A-\frac{m}{2}\right)^2\right] \\ &\le \frac{16}{m^2}\mathbb{E}\sum_{i=1}^{m}\mathrm{Var}_{\cdot\,|\,\{\epsilon_{-i}\}}\left[\left(A-\frac{m}{2}\right)^2\right] \le 16.\end{aligned}$$

The conclusion of the lemma follows as desired.

□

Lemma 7. *Let m be a positive integer and let $\delta \in [0, 1/2]$. Let $A \sim Bin(m, \frac{1}{2}+\delta)$ and let $U \sim Unif[0,1]$ be independent of A. Define $Z := -2\log\{U \cdot S_{Bin(\frac{1}{2},m)}(|A-\frac{m}{2}|) + (1-U)S_{Bin(\frac{1}{2},m)}(|A-\frac{m}{2}|+1)\}$ and $\tilde{Z} := \frac{4}{m}(A-\frac{m}{2})^2$.*

We have that

1. $\mathbb{E}Z - 2 \ge 8(m-1)\delta^2$,
2. *and if $(m-1)\delta^2 \ge 2$, then $\mathbb{E}\tilde{Z} - 4 \ge 2(m-1)\delta^2$.*

Proof. Define $\mathcal{M}_m$ as (S2.26) and note that $\mathcal{M}_m = -\mathcal{M}_m$. For $s \in \mathcal{M}_m$, write $P_m(s,\delta) = \binom{m}{\frac{m}{2}+s}(\frac{1}{2}+\delta)^{\frac{m}{2}+s}(\frac{1}{2}-\delta)^{\frac{m}{2}-s}$ as the probability that Bin$(m, \frac{1}{2}+\delta)$ random variable is equal to $\frac{m}{2}+s$ and $Q_m(s) := \binom{m}{\frac{m}{2}+s}(\frac{1}{2})^m = P_m(s,0)$.

Define $W = |A-\frac{m}{2}| + U$. We also define

$$\begin{aligned}F_m(\delta) = \mathbb{E}Z &= \mathbb{E}\big[-2\log\tilde{S}_0(W)\big] \\ &= \sum_{s\in\mathcal{M}_m} P_m(s,\delta)\int_0^1 \big\{-2\log\tilde{S}_0(|s|+u)\big\}\,du,\end{aligned}$$

where the definition of $\tilde{S}_0(\cdot)$ and the second equality follow from Lemma 10. We note then that

$$\mathbb{E}Z - 2 \ge F_m(\delta) - F_m(0).$$

Moreover, since the function $\delta \mapsto P_m(s,\delta)$ is equal to its Taylor series expansion for all $\delta \in (-\frac{1}{2}, \frac{1}{2})$, the same holds for $F_m(\delta)$, that is,

$$F_m(\delta) = F_m(0) + \sum_{j=1}^{\infty} F_m^{(j)}(0)\frac{\delta^j}{j!}, \qquad \text{for all } \delta \in (-\frac{1}{2}, \frac{1}{2}).$$

By symmetry, $P_m(s,\delta) = P_m(-s,-\delta)$ and thus, $F_m(\delta) = F_m(-\delta)$ and $F_m^{(j)}(0) = 0$ when j is an odd integer. When j is an even integer, we have that, by Lemma 11,

$$F_m^{(j)}(0) = \sum_{s\mathcal{M}_m}\left(\partial_\delta^{(j)} P_m(s,\delta)\big|_{\delta=0}\right)\int_0^1 \big\{-2\log\tilde{S}_0(|s|+u)\big\}\,du \ge 0.$$

We now claim that $F_m^{(2)}(0) \geq 8(m-1)$. To see this, first observe that, by Hoeffding's inequality, it holds that $-2\log\{S_{\mathrm{Bin}(\frac{1}{2},m)}(|s|)\} \geq \frac{4}{m}s^2 - 2\log 2$. Moreover, since $\sum_{s\in\mathcal{M}_m} P_m(s,\delta) = 1$, writing $P_m^{(2)}(s,\delta)$ as second derivative of $P_m(s,\delta)$ with respect to δ, we have $\sum_{s\in\mathcal{M}_m} P_m^{(2)}(s,\delta) = 0$. Thus, using the fact that $P_m^{(2)}(s,0) \geq 0$ for all $s \in \mathcal{M}_m$ (by Lemma 11), we have that, for any $\delta \in (-1/2, 1/2)$,

$$
\begin{aligned}
F_m^{(2)}(\delta) &= \sum_{s\in\mathcal{M}_m} P_m^{(2)}(s,\delta) \int_0^1 \{-2\log \tilde{S}_0(|s|+u)\}\,du \\
&\geq \sum_{s\in\mathcal{M}_m} P_m^{(2)}(s,\delta)\big(-2\log\{S_{\mathrm{Bin}(\frac{1}{2},m)}(|s|)\}\big) \\
&\geq \sum_{s\in\mathcal{M}_m} P_m^{(2)}(s,\delta)\frac{4}{m}s^2 \\
&\geq \frac{4}{m}\left(\frac{d}{d\delta}\right)^2 \underbrace{\Big\{\sum_{s\in\mathcal{M}_m} P_m(s,\delta)s^2\Big\}}_{\mathbb{E}(A-m/2)^2 \text{ where } A\sim\mathrm{Bin}(1/2+\delta,m)} \\
&= \frac{4}{m}\left(\frac{d}{d\delta}\right)^2 \left\{\frac{m}{4}(1-4\delta^2) + m^2\delta^2\right\} = 1 + 8(m-1).
\end{aligned}
$$

We may conclude then that $F_m^{(2)}(0) \geq 8(m-1)$ as desired. Therefore, we have that

$$
\mathbb{E}Z - 2 = F_m(\delta) - F_m(0) \geq 8(m-1)\delta^2.
$$

For the second claim of the Lemma, we note that $A \sim \mathrm{Bin}(\frac{1}{2}+\delta, m)$. Therefore, assuming $(m-1)\delta^2 \geq 2$, we have that

$$
\begin{aligned}
\mathbb{E}\tilde{Z} - 4 &= \frac{4}{m}\mathbb{E}\big(A - \frac{m}{2}\big)^2 - 4 \\
&= \frac{4}{m}\Big\{\mathbb{E}\big(A - \frac{m}{2} - m\delta\big)^2 + m^2\delta^2\Big\} - 4 \\
&= (1-4\delta^2) + 4m\delta^2 - 4 \geq 2m\delta^2
\end{aligned}
$$

as desired. The conclusion of the lemma thus follows. □

Lemma 8. *Let $I \subset \mathbb{R}^q$ and let $I_1,\ldots,I_L$ be a partition of I such that diam$(I_l) \leq C_d L^{-1/q}$ for all $l \in [L]$ for some $C_d > 0$. Write $\delta := \frac{\lambda_a - \lambda_b}{\lambda}$ and suppose that δ is γ-Holder continuous for $\gamma \in (0,1]$, i.e., $|\delta(x) - \delta(y)| \leq C_H\|x-y\|_2^\gamma$ for all $x, y \in I$, for some $C_H > 0$.*

Then, we have that

$$
0 \leq \int_I \left(\frac{\lambda_a - \lambda_b}{\lambda}\right)^2 \lambda\,d\nu - \sum_{l=1}^L \left\{\left(\frac{\int_{I_l}\lambda_a - \lambda_b\,d\nu}{\int_{I_l}\lambda\,d\nu}\right)^2 \int_{I_l}\lambda\,d\nu\right\} \leq 2C_H C_d d^{-\frac{2\gamma}{q}}\left(\int_I \lambda\,d\nu\right).
$$

Proof. Fix an arbitrary $l \in [L]$ and define $\mathbb{E}^{(l)}[\,\cdot\,]$ as expectation with respect to the probability measure with density $\frac{\lambda}{\int_{I_l} \lambda d\nu}$. We then have that

$$\frac{\int_{I_l} \left(\frac{\lambda_a - \lambda_b}{\lambda}\right)^2 \lambda \, d\nu}{\int_{I_l} \lambda \, d\nu} = \mathbb{E}^{(l)}[\delta^2] \geq \{\mathbb{E}^{(l)}\delta\}^2 = \left\{ \frac{\int_{I_l} \lambda_a - \lambda_b d\nu}{\int_{I_l} \lambda d\nu} \right\}^2.$$

For the other direction, we observe that

$$\begin{aligned}
&\frac{\int_{I_l} \left(\frac{\lambda_a - \lambda_b}{\lambda}\right)^2 \lambda \, d\nu}{\int_{I_l} \lambda \, d\nu} - \left\{ \frac{\int_{I_l} \lambda_a - \lambda_b d\nu}{\int_{I_l} \lambda d\nu} \right\}^2 \\
&= \mathbb{E}^{(l)}[\delta^2] - \{\mathbb{E}^{(l)}\delta\}^2 = \mathrm{Var}^{(l)}(\delta) \\
&\overset{(a)}{=} \frac{1}{2}\mathbb{E}^{(l)}\left[(\delta(X) - \delta(Y))^2\right] \\
&\leq \frac{1}{2} C_H \mathbb{E}^{(l)} \|X - Y\|_2^{2\gamma} \leq \frac{1}{2} C_H \sup_{x,y \in I_l} \|x - y\|_2^{2\gamma} \overset{(b)}{\leq} \frac{C_H C_d}{2} L^{-2\gamma/q}.
\end{aligned}$$

where in inequality (a), the random variables X, Y are independent and distributed with density $\frac{\lambda}{\int_{I_l} \lambda \, d\nu}$ and where in inequality (b), we use the assumption that $\mathrm{diam}(I_l) \leq C_d L^{-1/q}$.

In summary, we have that, for each $l \in [L]$,

$$0 \leq \int_{I_l} \left(\frac{\lambda_a - \lambda_b}{\lambda} \right)^2 \lambda \, d\nu - \left(\frac{\int_{I_l} \lambda_a - \lambda_b d\nu}{\int_{I_l} \lambda \, d\nu} \right)^2 \int_{I_l} \lambda \, d\nu \leq \frac{C_H C_d}{2} L^{-2\gamma/q} \int_{I_l} \lambda \, d\nu.$$

By summing over $l \in [L]$, the conclusion of the theorem follows as desired. □

S2.6 Technical lemmas

Lemma 9. *Let $X \sim \chi^2_{2k}$. Then, we have that for all $t > 0$,*

$$\mathbb{P}(X \geq 2k + 2\sqrt{2kt} + 2t) \leq e^{-t}.$$

Proof. If $X \sim \chi^2_{2k}$, then $\mathbb{E}X = 2k$ and X is also a Gamma$(k, 2)$ random variable and hence its moment generation function is bounded by $\mathbb{E}e^{\lambda(X - \mathbb{E}X)} \leq \frac{4k\lambda^2}{2(1-2\lambda)}$ for all $\lambda \in (0, \frac{1}{c})$ by **?**, Section 2.4. Then, it holds by **?**, Theorem 2.3 that $\mathbb{P}(X - \mathbb{E}X \geq \sqrt{8kt} + ct) \leq e^{-t}$ for all $t > 0$. The Lemma immediately follows. □

Lemma 10. *Let $m \in \mathbb{N}$ and let $\tilde{A} \sim Bin(\frac{1}{2}, m)$. Define $\tilde{W} = |\tilde{A} - \frac{m}{2}| + U$ where $U \sim Unif[0, 1]$ is independent of $\tilde{A}$. We let $\mathcal{M}_m^+$ be as defined in (S2.27).*

Write $\tilde{S}_0(z) := \mathbb{P}(\tilde{W} \geq z)$*. We have that, for any* $z \in \mathbb{R}$*,*

$$\tilde{S}_0(z) = \begin{cases} (1-(z-k_1))S_{Bin(\frac{1}{2},m)}(k_1) + (z-k_1)S_{Bin(\frac{1}{2},m)}(k_1+1) & \text{if } z \in \big[\min \mathcal{M}_m^+, 1+\max \mathcal{M}_m^+\big) \\ 1 & \text{if } z < \min \mathcal{M}_m^+ \\ 0 & \text{if } z \geq 1+\max \mathcal{M}_m^+, \end{cases} \tag{S2.33}$$

where in the first case, k_1 *is defined as* $k_1 := \max\{k \in \mathcal{M}_m^+ : k \leq z\}$*.*

Moreover, we have that

$$\tilde{S}_0'(z) = \begin{cases} -\mathbb{P}\big(|\tilde{A} - \frac{m}{2}| = k_1\big) & \text{if } z \in \big[\min \mathcal{M}_m^+, 1+\max \mathcal{M}_m^+\big) \\ 0 & \text{else }. \end{cases}$$

Finally, let $A \sim Bin(\frac{1}{2}+\delta, m)$ *and let* $W = |A - \frac{m}{2}| + U$ *where* $U \sim Unif[0,1]$ *is independent of* A*, we have that*

$$\tilde{S}_0(W) \overset{d}{=} (1-U)S_{Bin(\frac{1}{2},m)}\Big(|A - \frac{m}{2}|\Big) + U \cdot S_{Bin(\frac{1}{2},m)}\Big(|A - \frac{m}{2}| + 1\Big).$$

Proof. To establish the first claim, let $z \in \big[\min \mathcal{M}_m^+, 1+\max \mathcal{M}_m^+\big)$ and let $k_1 := \max\{k \in \mathcal{M}_m^+ : k \leq z\}$. Define the event

$$\mathcal{E}_{k_1} = \Big\{|\tilde{A} - \frac{m}{2}| = k_1\Big\}.$$

Then, we have that

$$\begin{aligned} \tilde{S}_0(z) &= \mathbb{P}(\tilde{W} \geq z) \\ &= \mathbb{P}(\mathcal{E}_{k_1})\mathbb{P}(\tilde{W} \geq z \,|\, \mathcal{E}_{k_1}) + \mathbb{P}\Big(|\tilde{A} - \frac{m}{2}| > k_1\Big). \\ &= \big\{S_{\text{Bin}(\frac{1}{2},m)}(k_1) - S_{\text{Bin}(\frac{1}{2},m)}(k_1+1)\big\}(1-(z-k_1)) + S_{\text{Bin}(\frac{1}{2},m)}(k_1+1). \end{aligned}$$

The first claim (S2.33) follows immediately.

The second claim follows by direct differentiation, and the third claim follows directly from the first claim. The whole Lemma thus follows as desired. ☐

Lemma 11. *For* $m \in \mathbb{N}$*,* $s \in \mathcal{M}_m$ *(defined as (S2.26)), and* $\delta \in (-\frac{1}{2}, \frac{1}{2})$*, define* $P_m(s,\delta) = \binom{m}{\frac{m}{2}+s}(\frac{1}{2}+\delta)^{\frac{m}{2}+s}(\frac{1}{2}-\delta)^{\frac{m}{2}-s}$*. We then have that, for any integer* $j \geq 1$*,*

$$\frac{\partial^{(2j)}}{\partial \delta^{(2j)}} P_m(s,\delta)\Big|_{\delta=0} \geq 0.$$

Proof. First suppose $s \geq 0$. Since $|2\delta| < 1$, we have that,

$$P_m(s,\delta) = \binom{m}{\frac{m}{2}+s} 2^{-m} \left(\frac{1+2\delta}{1-2\delta}\right)^s$$
$$= \binom{m}{\frac{m}{2}+s} 2^{-m} (1+2\delta)^s \left(1 + \sum_{k=1}^{\infty} (2\delta)^k\right)^s.$$

It is thus clear that in Taylor series expansion of $\delta \mapsto P_m(s,\delta)$, all the coefficients are non-negative and and thus, $\frac{\partial^{(2j)}}{\partial \delta^{(2j)}} P_m(s,\delta) \geq 0$.

If $s \leq 0$ on the other hand, the same claim follows by the fact that

$$P_m(-s,\delta) = P_m(s,-\delta).$$

The lemma thus immediately follows. □

S3. Supplementary material for Section 5

S3.1 Two sample test Type I error

We consider the following three intensities

1. $\lambda_a(x) = \lambda_b(x) = 40 \cdot \mathbf{1}_{[0,1]}(x)$
2. $\lambda_a(x) = \lambda_b(x) = 40 \cdot \big(\sin(2\pi x) + 1\big)$
3. $\lambda_a(x) = \lambda_b(x) = 40 \cdot \frac{x(1-x)^4}{\int_0^1 x(1-x)^4 dx} \mathbf{1}_{[0,1]}(x)$

The first function is uniform, while the other two are not, indicating the intensities changes on the support. Note that the third function is the scaled beta density function with parameters (2,5). For each of the three cases under the null hypothesis, we conduct 2000 simulations of two independent Poisson processes with the intensities functions given in the corresponding case and present the proportions of rejections out of all simulations based on the adjusted p-value of each test. We generate 500 bootstrap resamples of each of the 2000 pairs of Poisson processes conditional on the total number of observations of the pooled process $N = N_a + N_b$, and use the same resamples to derive adjusted p-values for all tests. We provide the percentage of rejections at level $\alpha = 0.05, 0.1$ and 0.25 for the five test procedures under 3 different intensities, the results are given in Table 1. We can see from the results that these five tests all attains the corresponding nominal levels, which is not a surprise due to the Monte Carlo Approximation of the exact rejection threshold.

Table 1: The empirical level (% of rejections) of different tests under the null

Test	case 1			case 2			case 3		
	5%	10%	25%	5%	10%	25%	5%	10%	25%
MF	4.9	9.8	25.7	5.5	8.9	23.7	5.1	10.2	24.5
MM	4.6	8.9	22.4	5.4	10.4	23.9	4.9	10.4	26.0
KN_1	4.8	9.6	23.5	5.1	9.5	24.2	6.1	11.3	27.2
KN_2	4.7	9.7	25.9	4.4	9.3	24.1	5.9	10.9	25.8
KS	5.1	9.9	24.9	4.5	9.1	24.6	6.2	11.4	25.9

S3.2 Testing homogeneous array

As an empirical verification of Theorem 2, in Figure 1, we plot the finite sample distributions of the largest eigenvalue of the adjacency matrix $A^{(r,\ell)}$ under the null hypothesis. We give the details of the experimental set-up in Section S3.3 of the appendix; in that section, we also discuss the bootstrap correction method proposed by **?** to improve the Tracy-Widom approximation.

Next, we consider two alternative Poisson SBM models with $K = 2$ and $K = 3$ equally sized communities respectively. We let the probability distribution of the interactions between two nodes u, v only depends on whether they are in the same community and we denote the intensity function of realizations between individuals within the same community as $\lambda_{\text{same}}(\cdot)$ and from different communities as $\lambda_{\text{diff}}(\cdot)$. We then define

$$\lambda_{\text{same}}(x) = s \cdot \mathbf{1}_{[0,1]}(x), \quad \lambda_{\text{diff}}(x) = s \cdot \frac{x(1-x)^4}{\int_0^1 x(1-x)^4 dx} \mathbf{1}_{[0,1]}(x)$$

for both the two alternative Poisson SBM models, where s is a parameter that controls the sparsity levels of the networks in this experiments. We again have $n = 200$ and either $K = 2$ or $K = 3$ equally sized communities. We then generate 200 sample collections of realizations on the same support for each of the two models and for each value of $s \in \{0.1, 0.175, 0.25, 0.5, 1\}$ and conduct our proposed test on these samples where the bootstrap sample size and partitioning of the support are exactly the same as in the preceding experiment. The proportions of rejections for the two SBM models under different sparsity levels are recorded in Table 2.

	$K=2$					$K=3$				
s	1.0	0.5	0.25	0.175	0.1	1.0	0.5	0.25	0.175	0.1
$\alpha = 0.01$	1	1	0.98	0.41	0.055	1	1	1	0.785	0.05
$\alpha = 0.05$	1	1	0.99	0.575	0.115	1	1	1	0.905	0.145
$\alpha = 0.10$	1	1	0.995	0.63	0.165	1	1	1	0.97	0.24
$\alpha = 0.25$	1	1	1	0.74	0.36	1	1	1	1	0.38

Table 2: The proportion of rejections of the proposed array test out of 200 simulated samples of networks of 200 nodes at different sparsity and confidence levels $\alpha \in \{0.01, 0.05, 0.10, 0.25\}$.

S3.3 Empirical verification of Tracy-Widom approximation and bootstrap correction

To see how fast the largest eigenvalues converge to the limiting distribution, we consider two cases with the numbers of nodes $n = 300$ and $n = 1600$ respectively. For each case we simulate 1000 adjacency matrix A whose entries $\{A_{ij} : i \neq j \leq n\}$ are independent and identically distributed Poisson random variables with mean equals to 20. Then we plot the sample distribution of the test statistics, i.e., $n^{2/3}\big(\lambda_1(\tilde{A}) - 2\big)$ against the Tracy-Widom distribution, where $\tilde{A}$ is the empirically centered and scaled version of A.

We can see from the first two graphs in Figure 1 that when $n = 300$ the sample distribution deviates in location compared with the target distribution and when $n = 1600$ the location is corrected but there still is some difference in scale. Though there are some differences in location and scale, we can see the sample distribution does have similar shape with the Tracy-Widom distribution even when the number of nodes is as small as 300. In similar experiments where adjacency matrices have Bernoulli distributed entries, **?** proposed to apply bootstrap correction to the largest eigenvalue, where they generate parametric bootstrap samples of the adjacency matrices and use the bootstrapped mean and variance of the largest eigenvalues to shift and scale the test statistics to have a better match with the Tracy-Widom distribution. Here we adapted the same bootstrap correction technique to the eigenvalues of adjacency matrix with $n = 300$ nodes, where we generate 50 bootstrap samples for each sample adjacency matrix. We plot the empirical distribution of the test statistics after bootstrap correction as the third graph in Figure 1. We can see that even with just 50 bootstrap samples, the sample distributions of the test statistics looks much closer to the target distribution.

Remember that here we are using the Tracy-widom distribution to compute p-values for every local null hypotheses $\bar{H}_0^{(r,\ell)}$ and we generate bootstrap samples to estimate the exact critical threshold for the global null hypothesis

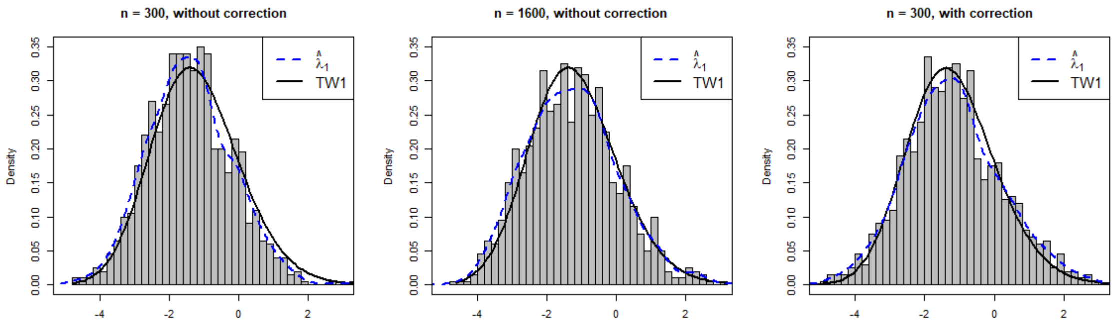

Figure 1: The empirical distribution of 1000 simulated samples of centered and scaled largest eigenvalues of $\tilde{A}$, compared with the Tracy-Widom distribution.

$\bar{H}_0$, thus we could simply use the same bootstrap samples generated for testing the global null to correct the location and scale of the largest eigenvalue of each local adjacency matrix $\tilde{A}^{(r,\ell)}$. Given observed collection of Poisson process realizations $\{N_{uv}(\cdot) : u < v \in [n]\}$, we describe the procedure to derive the local p-values with bootstrap correction of the location and scale of the largest eigenvalue in the following steps:

1. For $b^* = 1, 2, \ldots, B$, generate bootstrap sample collections $\{N_{uv}^{b^*}(\cdot) : u < v \in [n]\}$ as described in Section 3.1.3.

2. For the observed realization, estimate the Poisson mean $\hat{\lambda}^{(r,\ell)}$ of each discretized interval as $\hat{\lambda}^{(r,\ell)} = \frac{1}{n^2-n}\sum_{u\neq v} N_{uv}^{(r,\ell)}$ and let $\tilde{A}^{(r,\ell)}$ be the centered and re-scaled adjacency matrix for interval $I^{(r,\ell)}$
$$\tilde{A}_{uv}^{(r,\ell)} := \begin{cases} \frac{N_{uv}^{(r,\ell)} - \hat{\lambda}^{(r,\ell)}}{\sqrt{(n-1)\hat{\lambda}^{(r,\ell)}}}, & u \neq v, \\ 0, & u = v. \end{cases}$$
and let $\lambda_1(\tilde{A}^{(r,\ell)})$ be the largest eigenvalue of adjacency matrix $\tilde{A}^{(r,\ell)}$.

3. Do step 2 for every bootstrap resamples to derive their largest eigenvalues $\lambda_1(\tilde{A}^{(r,\ell,b^*)})$ at every discretized interval. Then we calculate the sample mean and standard deviation of $\left\{\lambda_1(\tilde{A}^{(r,\ell,b^*)}) : b^* \in \{1, 2, \ldots, B\}\right\}$ for each $r \in [R], \ell \in [2^r]$ and denote them as $\hat{\mu}_1^{(r,\ell)}, \hat{s}_1^{(r,\ell)}$ respectively.

4. Denote μ_{tw} and s_{tw} as the mean and standard deviation of Tracy-Widom distribution with $\beta = 1$ and let
$$\lambda_{bc}^{(r,\ell)} = \mu_{\text{tw}} + s_{\text{tw}} \frac{\lambda_1(\tilde{A}^{(r,\ell)}) - \hat{\mu}_1^{(r,\ell)}}{\hat{s}_1^{(r,\ell)}}$$
be the test statistic after bootstrap correction.

5. Finally we compute the p-value for the discretized local null $\bar{H}_{)}^{(r,\ell)}$ as

$$p^{(r,\ell)} \equiv p^{(r,\ell)}\left(\lambda_{bc}^{(r,\ell)}\right) := 2\min\left(F_{\mathrm{TW1}}\left(\lambda_{bc}^{(r,\ell)}\right), 1 - F_{\mathrm{TW1}}\left(\lambda_{bc}^{(r,\ell)}\right)\right)$$