SUPPLEMENTARY MATERIAL FOR "FITTING NETWORKS WITH A CANCELLATION TRICK"

Jiashun Jin
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213, USA
jiashun@andrew.cmu.edu

Jingming Wang
Department of Statistics
University of Virginia
Virginia, VA 22903, USA
pdw9qv@virginia.edu

A NOTATIONS

Throughout this supplementary material, we will adopt the following notions. (1) Let $\{e_k\}_{k=1}^K$ be the standard basis of $\mathbb{R}^K$. To distinguish, we use $\{e_{n,i}\}_{i=1}^n$ to denote the standard basis of $\mathbb{R}^n$. (2) We write $\mathbf{1}_m$ the all-one vector of dimension m. (3) For two sequence of numbers $a_n, b_n > 0$ depending on n, we write $a_n \gg b_n$ or $b_n \ll a_n$ if $b_n/a_n = o(1)$ as $n \to \infty$; and $a_n \asymp b_n$ is there exists constants $C, c > 0$ such that $cb_n < a_n \leq Cb_n$. (4) Let $O(K-1)$ be group of all $(K-1)\times(K-1)$ orthogonal matrices. (5) For any matrix $M \in \mathbb{R}^{m\times m}$, let its SVD be $M = UDV'$. We adopt the notion that $\mathrm{sgn}(M) = UV'$. (6) We denote the (i,j)-th entry of a matrix M as $M(i,j)$ or M_{ij}, and the i-th component of a vector u as $u(i)$ or u_i. (7) We denote c, C the generic constants which may vary from line to line.

B THE ERROR RATE OF SCORE

B.1 PROOF OF THEOREM 3.1

The proof of Theorem 3.1 can be separated into two parts. First, we connect the Hamming error with the error rates of SCORE vectors R where

$$R = \mathrm{diag}(\xi_1)^{-1}(\xi_2, \cdots, \xi_K)$$

with ξ_k denoted as the eigenvector associated with the k-th largest eigenvalue (in magnitude) of $\widetilde{\Omega}$, for $1 \leq k \leq K$. The result is collected in the following lemma and the proof is postponed to next subsection.

Lemma B.1 *Let $\widehat{R}$ be the SCORE vectors obtained from the observed network (either A or $A \oslash \widehat{N}$). Denote by R, the counterpart for $\widetilde{\Omega}$. Suppose that $\min_{\mathcal{O}\in O(K-1)} \|\widehat{R}\mathcal{O} - R\|_F^2 = o(n)$. Then, the Hamming error r_n satisfies*

$$r_n = n^{-1}\sum_{i=1}^n \|\hat{\pi}_i - \pi_i\|_1 \leq n^{-1} \min_{\{\mathcal{O}\in O(K-1)\}} \|\widehat{R}\mathcal{O} - R\|_F^2$$

Next, we claim the error rate $n^{-1}\min_{\{\mathcal{O}\in O(K-1)\}} \|\widehat{R}\mathcal{O} - R\|_F^2$ by applying SCORE algorithm. The key technical component is to conduct delicate eigenvector analysis and especially employ leave-one-out technique to derive sharp entry-wise eigenvector bounds for ξ_1. We present the result below, and the proof is relegated into Section B.3.

Lemma B.2 *Let $\widetilde{R}$ denote the SCORE vectors by employing SCORE directly on A. Under the assumptions in Theorem 3.1, it holds that with probability $1 - o(n^{-3})$,*

$$n^{-1}\min_{\mathcal{O}\in O(K-1)} \|\widetilde{R}\mathcal{O} - R\|_F^2 \leq C\frac{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|^2 + \lambda_1(\widetilde{\Omega})}{|\lambda_K(\widetilde{\Omega})|^2}$$

Therefore, Theorem 3.1 follows directly from Lemmas B.1 and B.2. In particular, if A satisfies DCBM, by definition, $N - \mathbf{1}_n\mathbf{1}_n'$. Then, $\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| = 0$, which yields that

$$r_n(\widehat{\Pi}^{score}) \leq C\frac{\lambda_1(\widetilde{\Omega})}{|\lambda_K(\widetilde{\Omega})|^2}.$$

To complete the proof, we show the proofs of Lemmas B.1 and B.2 in the subsequent two subsections.

B.2 Proof of Lemma B.1

The proof is similar to the proof of Theorem 2.2 SCORE (Jin, 2015), we provide the details below for readers' convenience. Without loss of generality, let us assume $\mathcal{O} = I_K$ for simplicity. According to (Jin, 2015), R contains exactly K distinct rows. Let $r_{(1)}, \cdots, r_{(K)}$ be the K distinct rows in R. To claim the bound, we first show that

$$\|r_{(k)} - r_{(\ell)}\| \geq c_1$$

for some constant $c_1 > 0$. To see this, we note that $(\xi_1, \xi_2, \cdots, \xi_K) = (\xi_1, \Xi_1) = \Theta\Pi B$ for some $B = (b_1, b_2, \cdots, b_K) \in \mathbb{R}^{K\times K}$. Then, it follows that $BB' = (\Pi'\Theta^2\Pi)^{-1} = \mathcal{P}\text{diag}([\sum_{i\in\mathcal{C}_1}\theta_i^2]^{-1}, \cdots, [\sum_{i\in\mathcal{C}_K}\theta_i^2]^{-1})\mathcal{P}'$ for some permutation matrix $\mathcal{P}$. Thanks to the conditions that $\theta_i \asymp \bar{\theta}$ and $n_k \asymp n$ for all $1 \leq k \leq K$, the conditional number of BB' is constant and

$$\lambda_{\min}(BB') \asymp \lambda_{\max}(BB') \asymp \frac{1}{n\bar{\theta}^2}$$

In particular, as $\xi_1 = \Theta\Pi b_1$, it is not hard to derive from $\Theta\Pi P\Pi'\Theta\xi_1 = \lambda_1\Theta\Pi b_1$ that

$$P\Pi'\Theta^2\Pi b_1 = \lambda_1 b_1$$

Therefore, b_1 is the first right eigenvector of $P(\Pi'\Theta^2\Pi)$. Under the condition in (17), $|b_1(k)| \asymp 1/\sqrt{n\bar{\theta}^2}$ for all $1 \leq k \leq K$. As a result,

$$|\xi_1(i)| \asymp \theta_i/\sqrt{n\bar{\theta}^2} \asymp 1/\sqrt{n}, \qquad \|e_{n,i}'\Xi_1\| \leq C/\sqrt{n} \tag{1}$$

and

$$\lambda_{\min}(\text{diag}(b_1)^{-1}B) \geq c_0$$

for some $c_0 > 0$. Notice that

$$R = \Pi[\text{diag}(b_1)^{-1}(b_2, \cdots, b_K)] = \Pi(r_{(1)}', \cdots, r_{(K)}')'$$

Therefore,

$$\|r_{(i)} - r_{(j)}\| = \|e_i'\text{diag}(b_1)^{-1}B - e_j'\text{diag}(b_1)^{-1}B\| \geq \sqrt{2}\lambda_{\min}(\text{diag}(b_1)^{-1}B) \geq \sqrt{2}c_0$$

for some $c_0 > 0$. We define $c_1 = 2c_0$. Let $V_1, \cdots, V_K$ denote the disjoint index sets corresponding to $r_{(1)}, \cdots, r_{(K)}$. The K-means algorithm aims to find a partition of the nodes $S^* = (S_1, S_2, \cdots, S_K)$ such that

$$S^* = \text{argmin}\sum_{k=1}^{K}\sum_{i\in S_k}\|\hat{r}_i - m_k\|^2, \quad m_k = \sum_{i\in S_k}\hat{r}_i \text{ for } 1 \leq k \leq K.$$

Define the output centers are $m_1^*, \cdots, m_K^*$. We introduce a matrix $M = (m_1', \cdots, m_n')'$ such that

$$m_i = m_k^*, \qquad \text{if} \quad i \in S_k$$

Thus,

$$\|\widehat{R} - M\|_F^2 \leq \|\widehat{R} - R\|_F^2 \quad \text{and} \quad \|R - M\|_F^2 \leq 4\|\widehat{R} - R\|_F^2$$

Let $\mathcal{I} := \{i : \|\hat{r}_i - r_i\| \leq \sqrt{2}c_0/8, \|m_i - r_i\| \leq \sqrt{2}c_0/8\}$ and $\mathcal{I}_k = \mathcal{I} \cap V_k$ for $1 \leq k \leq K$. We first prove that the nodes in $\mathcal{I}$ are correctly recovered. It suffices to show that for any $i \in \mathcal{I}_k, j \in \mathcal{I}_\ell$,

$$m_i = m_j \quad \text{if and only if} \quad k = \ell. \tag{2}$$

To see this, consider $k \neq \ell$, then $\|r_k - r_\ell\| \geq \sqrt{2}c_0$. It further yields that for $i \in \mathcal{I}_k, j \in \mathcal{I}_\ell$

$$\|m_i - m_j\| \geq \|r_i - r_j\| - \|m_i - r_i\| - \|m_j - r_j\| \geq \sqrt{2}c_0 \cdot 3/4$$

Suppose that $\mathcal{I}_k \neq \emptyset$ for all $1 \leq k \leq K$, then for every k, we select a point i_k and its corresponding m_{i_k}. It follows that

$$\|m_{i_k} - m_{i_\ell}\| \geq \|r_{i_k} - r_{i_\ell}\| - \|m_{i_k} - r_{i_k}\| - \|m_{i_\ell} - r_{i_\ell}\| \geq \sqrt{2}c_0 \cdot 3/4$$

By doing so, we fix the K distinct rows in M. Thus, based on the above arguments, for any two rows in M, their ℓ_2 norm distance is either 0 or larger than $\sqrt{2}c_0 \cdot 3/4$. For any $i, j \in \mathcal{I}_k$, since

$$\|m_i - m_j\| \leq \|m_i - r_k\| + \|m_j - r_k\| \leq \sqrt{2}c_0/4,$$

it must hold that $m_i = m_j$.

To complete the proof of (2), we need to claim $\mathcal{I}_k \neq \emptyset$ for all $1 \leq k \leq K$. We will prove by contradiction. Suppose there exist k_0 such that $\mathcal{I}_{k_0} = \emptyset$. It follows that

$$\sum_{i \in V_{k_0}} \|\hat{r}_i - r_i\|^2 + \|m_i - r_i\|^2 \geq |V_{k_0}|c_0/32 \geq \tilde{c}n$$

under the assumption that $n_k \asymp n$ for all $1 \leq k \leq K$. This implies that $\|\widehat{R}-R\|_F^2 + \|M-R\|_F^2 \geq \tilde{c}n$. Moreover,

$$\|\widehat{R} - R\|_F^2 \geq \tilde{c}n/5$$

which contradicts to $\|\widehat{R} - R\|_F^2 \ll n$. As a result, $\mathcal{I}_k \neq \emptyset$ for all $1 \leq k \leq K$. We thus finish the proof that nodes in $\mathcal{I}$ are exactly recovered.

Next, to finish the proof, we show that

$$|\mathcal{I}^c| \leq \|\widehat{R} - R\|_F^2$$

Note that for $i \in \mathcal{I}^c$, either $\|\hat{r}_i - r_i\| > \sqrt{2}c_0/8$ or $\|m_i - r_i\| > \sqrt{2}c_0/8$. Since $\|M - R\|_F^2 \leq 4\|\widehat{R} - R\|_F^2$, we can obtain that

$$|\mathcal{I}^c| \leq \frac{\|\widehat{R} - R\|_F^2}{(\sqrt{2}c_0/8)^2} + \frac{\|M - R\|_F^2}{(\sqrt{2}c_0/8)^2} \leq \frac{160}{c_0^2}\|\widehat{R} - R\|_F^2$$

Consequently,

$$r_n = \sum_{i \in \mathcal{I}} \|\hat{\pi}_i - \pi_i\|_1 + \sum_{i \in \mathcal{I}^c} \|\hat{\pi}_i - \pi_i\|_1 \leq 2|\mathcal{I}^c| \leq C_1\|\widehat{R} - R\|_F^2$$

We thereby conclude the proof.

B.3 Proof of Lemma B.2

We define $(\tilde{\lambda}_k, \tilde{\xi}_k)$ be the k-th largest eigen-pair of A (in magnitude) for $1 \leq k \leq K$. For simplicity, write $\widetilde{\Xi}_1 := (\tilde{\xi}_2, \cdots, \tilde{\xi}_K)$. Without loss of generality, we assume that $\mathrm{sgn}(\tilde{\xi}_1'\xi_1) = 1$. Let $\mathcal{O} := \mathrm{sgn}(\widetilde{\Xi}_1'\Xi_1)$. By definition,

$$\begin{aligned}
\|\widetilde{R}\mathcal{O} - R\|_F^2 &= \|\mathrm{diag}(\tilde{\xi}_1)^{-1}(\widetilde{\Xi}_1\mathcal{O} - \Xi_1) - \big[\mathrm{diag}(\tilde{\xi}_1)^{-1} - \mathrm{diag}(\xi_1)^{-1}\big]\Xi_1\|_F^2 \\
&\leq C\Big(\|\mathrm{diag}(\tilde{\xi}_1)^{-1}(\widetilde{\Xi}_1\mathcal{O} - \Xi_1)\|_F^2 + \|\big[\mathrm{diag}(\tilde{\xi}_1)^{-1} - \mathrm{diag}(\xi_1)^{-1}\big]\Xi_1\|_F^2\Big) \\
&\leq C\Big(\|\mathrm{diag}(\tilde{\xi}_1)^{-1}(\widetilde{\Xi}_1\mathcal{O} - \Xi_1)\|_F^2 + \|\tilde{\xi}_1 - \xi_1\|^2\|\mathrm{diag}(\tilde{\xi}_1)^{-1}\mathrm{diag}(\xi_1)^{-1}\Xi_1\|_{2\to\infty}^2\Big)
\end{aligned} \tag{3}$$

According to the RHS, we need to prove an upper bounds for $\|\widetilde{\Xi}_1\mathcal{O} - \Xi_1\|_F$ and $\|\tilde{\xi}_1 - \xi_1\|$, and further show that $|\tilde{\xi}_1(i)| \asymp 1/\sqrt{n}$ for $1 \leq i \leq n$.

First, we claim upper bounds for $\|\widetilde{\Xi}_1\mathcal{O} - \Xi\|_F$ and $\|\tilde{\xi}_1 - \xi_1\|$. Using sine-theta theorem (Davis & Kahan, 1970; Yu et al., 2015), we have

$$\min\{\|\tilde{\xi}_1 - \xi_1\|, \|\Xi_1'\tilde{\xi}_1\|\} \le C\frac{\|A-\widetilde{\Omega}\|}{\lambda_1(\widetilde{\Omega})} \qquad \|\widetilde{\Xi}_1 O - \Xi_1\|_F \le C\frac{\|A-\widetilde{\Omega}\|}{|\lambda_K(\widetilde{\Omega})|}$$

Since $A = \widetilde{\Omega} + (N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega} - \mathrm{diag}(N \circ \widetilde{\Omega}) + W = \widetilde{\Omega} + \widetilde{W}$, we thus bound

$$\|A - \widetilde{\Omega}\| \le \|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \|\mathrm{diag}(N \circ \widetilde{\Omega})\| + \|W\| \le \|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + C\sqrt{n\bar{\theta}^2}$$

with probability $1 - o(n^{-3})$. Here we used the derivation

$$\|\mathrm{diag}(N \circ \widetilde{\Omega})\| \le \|\mathrm{diag}(\widetilde{\Omega})\| \le C\bar{\theta}^2, \qquad \|W\| \le C\sqrt{n\bar{\theta}^2}$$

by the non-asymptotic bounds on the norm of random matrices in (Bandeira & Van Handel, 2016). It is worth mentioning that $\lambda_1(\widetilde{\Omega}) = \lambda_1(P(\Pi'\Theta^2\Pi)) \asymp n\bar{\theta}^2$. As a result,

$$\min\{\|\tilde{\xi}_1 - \xi_1\|, \|\Xi_1'\tilde{\xi}_1\|\} \le C\frac{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})}$$

$$\|\widetilde{\Xi}_1 O - \Xi_1\|_F \le C\frac{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \sqrt{\lambda_1(\widetilde{\Omega})}}{|\lambda_K(\widetilde{\Omega})|} \tag{4}$$

Next, we aim to show that $|\tilde{\xi}_1(i)| \asymp 1/\sqrt{n}$ for $1 \le i \le n$. Given that $|\xi_1(i)| \asymp 1/\sqrt{n}$ for $1 \le i \le n$, it suffices to show that $\|\tilde{\xi}_1 - \xi_1\|_{\max} \ll 1/\sqrt{n}$. To see this, we consider the eigen-perturbation that

$$\tilde{\xi}_1 - \xi_1 = (\tilde{\lambda}_1^{-1}\lambda_1\xi_1'\tilde{\xi}_1 - 1)\xi_1 + \tilde{\lambda}_1^{-1}\Xi_1\mathrm{diag}(\lambda_2, \cdots, \lambda_K)\Xi_1'\tilde{\xi}_1 + \tilde{\lambda}_1^{-1}\widetilde{W}\tilde{\xi}_1\,.$$

By the first inequality in (4) and the Weyl's inequality, we bound

$$\begin{aligned}|\tilde{\lambda}_1^{-1}\lambda_1\xi_1'\tilde{\xi}_1 - 1| &\le C\Big(\frac{|\tilde{\lambda}_1 - \lambda_1|}{\lambda_1} + |\xi_1'\tilde{\xi}_1 - 1|\Big) \le C\Big(\frac{\|A-\widetilde{\Omega}\|}{\lambda_1} + \|\tilde{\xi}_1 - \xi_1\|^2\Big)\\ &\le C\frac{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})}\end{aligned}$$

and

$$\|\tilde{\lambda}_1^{-1}\mathrm{diag}(\lambda_2, \cdots, \lambda_K)\Xi_1'\tilde{\xi}_1\| \le \|\Xi_1'\tilde{\xi}_1\| \le C\frac{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})}$$

Based on these, we arrive at

$$|\tilde{\xi}_1(i) - \xi_1(i)| \le \frac{C}{\sqrt{n}}\frac{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})} + \frac{|e_{n,i}'\widetilde{W}\tilde{\xi}_1|}{n\bar{\theta}^2}$$

for $1 \le i \le n$, due to the fact that $|\tilde{\lambda}_1 - \lambda_1| \le \|A - \widetilde{\Omega}\| \ll \lambda_1 \asymp n\bar{\theta}^2$ and $\max\|\xi_1\|_{\max}, \max_i \|e_{n,i}'\Xi\| \le C/\sqrt{n}$. Then, it suffices to derive an upper bound for $e_{n,i}'\widetilde{W}\tilde{\xi}_1$. We first decompose

$$|e_{n,i}'\widetilde{W}\tilde{\xi}_1| \le |e_{n,i}'(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\tilde{\xi}_1| + |e_{n,i}'\mathrm{diag}(N \circ \widetilde{\Omega})\tilde{\xi}_1| + |e_{n,i}'W\tilde{\xi}_1| \le \bar{\theta}^2|\tilde{\xi}_1(i)| + |e_{n,i}'W\tilde{\xi}_1|$$

Let $\tilde{\xi}_1^{(i)}$ be the first eigenvector of $A^{(i)} = \Omega - \mathrm{diag}(\Omega) + W^{(i)}$ where $W^{(i)}$ is obtained by zeroing out the i-th row and column of W. Then,

$$|e_{n,i}'W\tilde{\xi}_1| \le |e_{n,i}'W\tilde{\xi}_1^{(i)}| + \sqrt{n\bar{\theta}^2}\|\tilde{\xi}_1 - \tilde{\xi}_1^{(i)}\| \tag{5}$$

By Bernstein inequality, we bound

$$\begin{aligned}|e_{n,i}'W\tilde{\xi}_1^{(i)}| &\le C(\sqrt{\bar{\theta}^2\|\tilde{\xi}_1^{(i)}\|^2\log(n)} + \|\tilde{\xi}_1^{(i)}\|_{\max}\log(n))\\ &\le C(\bar{\theta}\sqrt{\log(n)} + \|\tilde{\xi}_1\|_{\max}\log(n) + \|\tilde{\xi}_1^{(i)} - \tilde{\xi}_1\|\log(n))\end{aligned} \tag{6}$$

simultaneously for all $1 \le i \le n$, with probability $1 - o(n^{-3})$.

To proceed, we analyze $\|\tilde{\xi}_1^{(i)} - \tilde{\xi}_1\|$ below. By sine-theta theorem,

$$\|\tilde{\xi}_1^{(i)} - \tilde{\xi}_1\| \le C\frac{\|(A^{(i)} - A)\tilde{\xi}_1\|}{n\bar{\theta}^2} \le C\frac{\|e_{n,i}e_{n,i}'W\tilde{\xi}_1\|}{n\bar{\theta}^2} + C\frac{\|We_{n,i}e_{n,i}'\tilde{\xi}_1\|}{n\bar{\theta}^2}$$
$$\le C\frac{|e_{n,i}'W\tilde{\xi}_1|}{n\bar{\theta}^2} + C\frac{\sqrt{n\bar{\theta}^2}|\tilde{\xi}_1(i)|}{n\bar{\theta}^2} \tag{7}$$

Combining (5)-(7) gives

$$|e_{n,i}'W\tilde{\xi}_1| \le C(\bar{\theta}\sqrt{\log(n)} + \|\tilde{\xi}_1\|_{\max}\log(n))$$

Consequently,

$$|\tilde{\xi}_1(i) - \xi_1(i)| \le \frac{C}{\sqrt{n}}\frac{\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\| + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})} + \frac{C|\tilde{\xi}_1(i)|}{n}$$
$$+ \frac{C}{n\bar{\theta}^2}(\bar{\theta}\sqrt{\log(n)} + \|\tilde{\xi}_1\|_{\max}\log(n))$$

By decomposing $|\tilde{\xi}_1(i)| \le |\xi_1(i)| + |\tilde{\xi}_1(i) - \xi_1(i)|$ and $\|\tilde{\xi}_1\|_{\max} \le \|\xi_1\|_{\max} + \|\tilde{\xi}_1 - \xi_1\|_{\max}$, we further have

$$|\tilde{\xi}_1(i) - \xi_1(i)| \le \frac{C}{\sqrt{n}}\frac{\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|}{n\bar{\theta}^2} + \frac{C\sqrt{\log(n)}}{n\bar{\theta}} + \|\tilde{\xi}_1 - \xi_1\|_{\max}\frac{\log(n)}{n\bar{\theta}^2}$$

Taking maximum and rearranging both sides, it follows that with probability $1 - o(n^{-3})$,

$$\|\tilde{\xi}_1 - \xi_1\|_{\max} \le \frac{C}{\sqrt{n}}\frac{\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|}{n\bar{\theta}^2} + \frac{C\sqrt{\log(n)}}{n\bar{\theta}} \ll 1/\sqrt{n}$$

under the condition that $\sqrt{n\bar{\theta}^2} \ge C\log(n)$ and $\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\| \ll n\bar{\theta}^2$. This completes the proof of $|\tilde{\xi}_1(i)| \asymp 1/\sqrt{n}$ for $1 \le i \le n$.

Therefore, we deduce from (3), (4) that with probability $1 - o(n^{-3})$,

$$\|\widetilde{R}\mathcal{O} - R\|_F^2 \le Cn\Big(\|(\widetilde{\Xi}_1\mathcal{O} - \Xi)\|_F^2 + \|\tilde{\xi}_1 - \xi_1\|^2\Big) \le Cn\frac{\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|^2 + \lambda_1(\widetilde{\Omega})}{|\lambda_K(\widetilde{\Omega})|^2}$$

We thus finish the proof.

B.4 A REMARK ON $\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|$

We discuss the relation of $\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|$ with the eigenvalues of $\widetilde{\Omega}$ and Ω in the following lemma.

Lemma B.3 *The following inequalities hold.*

$$\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\| \le \|N - \mathbf{1}_n\mathbf{1}_n'\|_{\max}\lambda_1(\widetilde{\Omega})$$
$$|\lambda_{K+1}(\Omega)| \le \|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|$$

Proof B.1 *Notice that $c < \min_{i,j} N(i,j) \le \|N\|_{\max} < 1$ for some constant $c > 0$. Then, $(\mathbf{1}_n\mathbf{1}_n' - N)\circ\widetilde{\Omega}$ is a symmetric matrix with positive entries. By Perron's theorem (see (Horn & Johnson, 1985) for example), the first eigenvector, denoted by u_1, is a positive vector and $\lambda_1((\mathbf{1}_n\mathbf{1}_n' - N)\circ\widetilde{\Omega}) = \|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|$. It follows that*

$$\|(N - \mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\| = u_1'(\mathbf{1}_n\mathbf{1}_n' - N)\circ\widetilde{\Omega}u_1 = \sum_{i,j}u_1(i)u_1(j)(1 - N_{ij})\widetilde{\Omega}_{ij}$$
$$\le \|N - \mathbf{1}_n\mathbf{1}_n'\|_{\max}\cdot\sum_{i,j}u_1(i)u_1(j)\widetilde{\Omega}_{ij}$$
$$\le \|N - \mathbf{1}_n\mathbf{1}_n'\|_{\max}\lambda_1(\widetilde{\Omega})\,.$$

Next, we show the second inequality. Recall the decomposition

$$\Omega = N \circ \widetilde{\Omega} = \widetilde{\Omega} + (N - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\,.$$

By Weyl's inequality (see (Horn & Johnson, 1985) for example),

$$|\lambda_{K+1}(\Omega) - \lambda_{K+1}(\widetilde{\Omega})| \leq \|\Omega - \widetilde{\Omega}\| \leq \|(N - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\|$$

Since $\lambda_{K+1}(\widetilde{\Omega}) = 0$*, we conclude that*

$$|\lambda_{K+1}(\Omega)| \leq \|(N - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\|$$

C THE ERROR RATE OF R-SCORE

In this section, we mainly prove Lemma 3.1 and Theorem 3.2. We streamline the proofs as follows:

(1) We show the error rate of SCORE vectors by R-SCORE, i.e., $\|\widehat{R} - R\|_F^2$ up to some orthogonal transformation. This, together with Lemma B.1 concludes the proof of Lemma 3.1 (see Section C.1);
(2) We prove the error rate of refitting θ and P (see Sections C.2 and C.3);
(3) Third, we investigate the error rate of N, more precisely, $\|(N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\|$ and $\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F$ (see Section C.4);
(4) Combining all the previous results, together with Lemma B.1, we complete the proof of Theorem 3.2 (see Section C.5);
(5) We also provide the brief proof of the Corollary 3.1, as it follows simply from Theorem 3.2.

The details are provided in the subsequent subsections.

C.1 PROOF OF LEMMA 3.1

Recall the assumption that $\widehat{N}$ satisties

$$\mathbf{1}_n \mathbf{1}_n' \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n' = \widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}$$

such that with probability $1 - o(n^{-3})$,

$$\|\widehat{P} - P\|_{\max} \ll \min\{1, \lambda_{\min}(P)\bar{\theta}^{-1}\}, \quad \|\widehat{\Pi} - \Pi\|(\sqrt{n}\,\lambda_{\min}(P))^{-1}\bar{\theta} \to 0,$$

and

$$\hat{\theta}_i \leq C\bar{\theta}, \quad \bar{\theta} = o(1)$$

for some constant $c, C > 0$. It follows that $\widehat{N}_{ij} = (1 + \hat{\theta}_i\hat{\theta}_j\hat{\pi}_i'\widehat{P}\hat{\pi}_j)^{-1} > C$ for some constant $0 < C < 1$ and $\widehat{N}_{ij} \leq 1$ for all $0 < i, j < n$.

Let $(\hat{\lambda}_k, \hat{\xi}_k)$ be the k-th largest eigen-pair of $A \oslash \widehat{N}$ (in magnitude) for $1 \leq k \leq K$. For brevity, write $\widehat{\Xi}_1 := (\hat{\xi}_2, \cdots, \hat{\xi}_K)$. Denote by (λ_k, ξ_k) and Ξ_1 the counterparts for the low-rank matrix $\widetilde{\Omega} = \Theta\Pi P\Pi'\Theta$. Without loss of generality, we assume both $\hat{\xi}_1$ and ξ_1 are positive. Under these notations, the SCORE vectors of A and $\widetilde{\Omega}$ are defined as

$$\widehat{R} = (\hat{r}_1, \hat{r}_2, \cdots, \hat{r}_n)' = \mathrm{diag}(\hat{\xi}_1)^{-1}\widehat{\Xi}_1, \qquad R = (r_1, r_2, \cdots, r_n)' = \mathrm{diag}(\xi_1)^{-1}\Xi_1$$

We bound the error of $\widehat{R} - R$ by the eigenvalues of A. Consider the SVD of $\widehat{\Xi}_1'\Xi = UDV'$. Define $O := \mathrm{sgn}(\widehat{\Xi}_1'\Xi_1) = UV'$. Our model assumptions gives that $\lambda_1(\widetilde{\Omega}) - |\lambda_2(\widetilde{\Omega})| \geq c\lambda_1(\widetilde{\Omega})$. Applying sine-theta theorem (Davis & Kahan, 1970; Yu et al., 2015), we have

$$\min\{\|\hat{\xi}_1 - \xi_1\|, \|\Xi_1'\hat{\xi}_1\|\} \leq C\frac{\|A \oslash \widehat{N} - \widetilde{\Omega}\|}{\lambda_1(\widetilde{\Omega})} \qquad \|\widehat{\Xi}_1 O - \Xi_1\|_F \leq C\frac{\|A \oslash \widehat{N} - \widetilde{\Omega}\|}{|\lambda_K(\widetilde{\Omega})|}$$

We write

$$A \oslash \widehat{N} = \widetilde{\Omega} + (N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega} - \mathrm{diag}(N \circ \widetilde{\Omega} \oslash \widehat{N}) + W \oslash \widehat{N} := \widetilde{\Omega} + \widetilde{W}$$

It follows that

$$\begin{aligned}\|\widetilde{W}\| &\leq \|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \|\mathrm{diag}(N \circ \widetilde{\Omega} \oslash \widehat{N})\| + \|W \oslash \widehat{N}\| \\ &\leq \|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\| + C\sqrt{n\bar{\theta}^2}\end{aligned}$$

To obtain the RHS, we bound

$$\|\mathrm{diag}(N \circ \widetilde{\Omega} \oslash \widehat{N})\|_F \leq C\|\widetilde{\Omega}\|_{\max} \leq C\bar{\theta}^2$$

and

$$\begin{aligned}\|W \oslash \widehat{N}\| &\leq \|W \oslash N\| + \|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\| \\ &\leq C\sqrt{n\bar{\theta}^2} + \|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\|\end{aligned}$$

where we used non-asymptotic bounds on the norm of random matrices in (Bandeira & Van Handel, 2016) since $W \oslash N$ is a symmetric random matrix with independent upper triangular entries and each entry in N of constant order. We further study $\|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\|$ as follows. Notice that

$$N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n' = (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta) \circ N$$

by the definition of N and $\widehat{N}$. Therefore, it suffices to bound

$$\|W \circ (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)\|$$

Next, we decompose

$$\begin{aligned}&\|W \circ (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)\| \\ &= \|W \circ \widehat{\Theta}\widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\widehat{\Theta}\| + \|W \circ \widehat{\Theta}(\widehat{\Pi} - \Pi)P\widehat{\Pi}'\widehat{\Theta}\| + \|W \circ \widehat{\Theta}\Pi P(\widehat{\Pi} - \Pi)'\widehat{\Theta}\| \\ &\quad + \|W \circ (\widehat{\Theta} - \Theta)\Pi P\Pi'\widehat{\Theta}\| + \|W \circ \Theta\Pi P\Pi'(\widehat{\Theta} - \Theta)\| \\ &=: \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_4 + \mathcal{T}_5\end{aligned}$$

We bound each term separately below. For $\mathcal{T}_1$, we have

$$\begin{aligned}\mathcal{T}_1 = \|\widehat{\Theta}(W \circ \widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}')\widehat{\Theta}\| \leq C\bar{\theta}^2\|W \circ \widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\| &\leq C\bar{\theta}^2\|\widehat{P} - P\|_{\max}\|W\|_F \\ &\leq n\bar{\theta}^3\|\widehat{P} - P\|_{\max}\end{aligned}$$

where $\|W\|_F \leq \sqrt{n}\|W\| \leq Cn\bar{\theta}$ with probability $1 - o(n^{-3})$.

The analysis for bounding $\mathcal{T}_2$ and $\mathcal{T}_3$ is similar, we provide the details for $\mathcal{T}_2$ only.

$$\begin{aligned}\mathcal{T}_2 = \|\widehat{\Theta}(W \circ (\widehat{\Pi} - \Pi)P\widehat{\Pi}')\widehat{\Theta}\| \leq C\bar{\theta}^2\|W \circ (\widehat{\Pi} - \Pi)P\widehat{\Pi}'\| &\leq C\bar{\theta}^2\|W \circ (\widehat{\Pi} - \Pi)P\widehat{\Pi}'\|_F \\ &\leq C\bar{\theta}^2\sqrt{\sum_{i:\hat{\pi}_i \neq \pi_i}\sum_j W_{ij}^2[(\hat{\pi}_i - \pi_i)'P\hat{\pi}_j]^2} \\ &\leq C\bar{\theta}^2\sqrt{\sum_{i:\hat{\pi}_i \neq \pi_i}\sum_j W_{ij}^2 \cdot 4\|P\|_{\max}^2} \\ &\leq C\bar{\theta}^2\|W\|_{2\to\infty}\|\widehat{\Pi} - \Pi\| \\ &\leq \bar{\theta}^2\sqrt{n\bar{\theta}^2}\|\widehat{\Pi} - \Pi\|\end{aligned}$$

The last step is due to fact that $\|e_{n,i}'W\| \leq \sqrt{n\bar{\theta}^2}$ simultaneously for all $1 \leq i \leq n$ with probability $1 - o(n^{-3})$ by Bernstein inequality. From here to the end of this subsection, with a slight abuse of notation, we will use $\{e_i\}_{i=1}^n$ to denote the standard basis of $\mathbb{R}^n$ for simplicity.

Next, for $\mathcal{T}_4$ and $\mathcal{T}_5$, the analysis is also analogous, and we show the bound for $\mathcal{T}_4$ in detail and omit the proof for $\mathcal{T}_5$.

$$\mathcal{T}_4 = \|(\widehat{\Theta} - \Theta)(W \circ \Pi P\Pi')\widehat{\Theta}\| \leq C\bar{\theta}^2\|W \circ \Pi P\Pi'\| \leq C\bar{\theta}^2\sqrt{n\bar{\theta}^2}$$

where we bound $\|W \circ \Pi P \Pi'\| \leq \sqrt{n\bar{\theta}^2}$ by non-asymptotic bound for random matrices since $\|\Pi P \Pi'\|_{\max} \leq C$ and $W \circ \Pi P \Pi'$ is symmetric random matrix with independent upper triangular entries.

Combining the discussions above, we have

$$\|W \circ (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)\| \leq C\Big(n\bar{\theta}^3\|\widehat{P} - P\|_{\max} + \bar{\theta}^2\sqrt{n\bar{\theta}^2}\|\widehat{\Pi} - \Pi\| + \bar{\theta}^2\sqrt{n\bar{\theta}^2}\Big)$$

This further gives rise to

$$\begin{aligned}\|\widetilde{W}\| &\leq \|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \|W \circ (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)\| + C\sqrt{n\bar{\theta}^2} \\ &\leq \|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + C\Big(n\bar{\theta}^3\|\widehat{P} - P\|_{\max} + \bar{\theta}^2\sqrt{n\bar{\theta}^2}\|\widehat{\Pi} - \Pi\| + \sqrt{n\bar{\theta}^2}\Big) \\ &\ll |\lambda_K(\widetilde{\Omega})|\end{aligned}$$

under the assumptions that

$$\|\widehat{P} - P\|_{\max} \ll \lambda_{\min}(P)/\bar{\theta} \quad \|\widehat{\Pi} - \Pi\| \ll \sqrt{n}\lambda_{\min}(P)/\bar{\theta} \quad \|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| \ll |\lambda_K(\widetilde{\Omega})|$$

and $\sqrt{n\bar{\theta}^2}\lambda_{\min}(P) \geq c_3\log(n)$.

We note that $\lambda_1(\widetilde{\Omega}) \asymp n\bar{\theta}^2$ and $|\lambda_K(\widetilde{\Omega})| \asymp n\bar{\theta}^2|\lambda_{\min}(P)|$. In addition, we have the decomposition $\Omega \oslash \widehat{N} = \widetilde{\Omega} + (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}$. Since $\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| \ll |\lambda_K(\widetilde{\Omega})|$ with high probability, we obtain that

$$\lambda_k(\Omega \oslash \widehat{N}) = \lambda_k(\widetilde{\Omega})(1 + o(1)), \quad \text{for } 1 \leq k \leq K.$$

Consequently, recall the definition that $\tau_n = n\bar{\theta}^3\|\widehat{P} - P\|_{\max} + \bar{\theta}^2\sqrt{n\bar{\theta}^2}\|\widehat{\Pi} - \Pi\|$, we obtain

$$\begin{aligned}\min\{\|\hat{\xi}_1 - \xi_1\|, \|\Xi_1'\hat{\xi}_1\|\} &\leq C\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})} = o(1) \\ \|\widehat{\Xi}_1 O - \Xi\|_F &\leq C\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n + \sqrt{\lambda_1(\widetilde{\Omega})}}{\big|\lambda_K(\widetilde{\Omega})\big|} = o(1) \qquad (8)\end{aligned}$$

with probability $1 - o(n^{-3})$.

To proceed, we need to study the entry-wise error for $\hat{\xi}_1 - \xi_1$. By $(A \oslash \widehat{N})\hat{\xi}_1 = \hat{\lambda}_1\hat{\xi}_1$ and $A \oslash \widehat{N} = \widetilde{\Omega} + \widetilde{W}$, we derive

$$\hat{\xi}_1 - \xi_1 = (\hat{\lambda}_1^{-1}\lambda_1\xi_1'\hat{\xi}_1 - 1)\xi_1 + \hat{\lambda}_1^{-1}\Xi_1\mathrm{diag}(\lambda_2, \cdots, \lambda_K)\Xi_1'\hat{\xi}_1 + \hat{\lambda}_1^{-1}\widetilde{W}\hat{\xi}_1\,.$$

We can bound

$$\begin{aligned}|\hat{\lambda}_1^{-1}\lambda_1\xi_1'\hat{\xi}_1 - 1| &\leq C(|\lambda_1^{-1}(\hat{\lambda}_1 - \lambda_1)| + |\xi_1'\hat{\xi}_1 - 1|) \leq C\Big(\frac{\|\widetilde{W}\|}{\lambda_1(\widetilde{\Omega})} + \frac{\|\widetilde{W}\|^2}{\lambda_1(\widetilde{\Omega})^2}\Big) \\ &\leq C\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})}\end{aligned}$$

and

$$\|\hat{\lambda}_1^{-1}\mathrm{diag}(\lambda_2, \cdots, \lambda_K)\Xi_1'\hat{\xi}_1\| \leq \|\Xi_1'\hat{\xi}_1\| \leq C\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})}$$

These give rise to

$$|\hat{\xi}_1(i) - \xi_1(i)| \leq C\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})} \cdot \frac{1}{\sqrt{n}} + \frac{|e_i'\widetilde{W}\hat{\xi}_1|}{n\bar{\theta}^2} \qquad (9)$$

since $\|\Xi_1(i)\| \leq 1/\sqrt{n}$ and $\xi_1(i) \asymp 1/\sqrt{n}$ following from the assumptions on $\widetilde{\Omega}$ (see (1)).

What remains to bound $|e_i'\widetilde{W}\hat{\xi}_1|/n\bar{\theta}^2$. Using the definition of $\widetilde{W}$, we first have

$$
\begin{aligned}
&|e_i'\widetilde{W}\hat{\xi}_1| \\
&\leq \big|e_i'[(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}]\hat{\xi}_1\big| + \big|(N \circ \widetilde{\Omega} \oslash \widehat{N})_{ii}\hat{\xi}_1(i)\big| + \big|e_i'(W \oslash \widehat{N})\hat{\xi}_1\big| \\
&\leq \|e_i'[(\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta) \circ \widetilde{\Omega} \circ N]\|_1\|\hat{\xi}_1\|_{\max} + \bar{\theta}^2\Big(\frac{1}{\sqrt{n}} + |\hat{\xi}_1(i) - \xi_1(i)|\Big) + \big|e_i(W \oslash \widehat{N})\hat{\xi}_1\big| \\
&\leq n\bar{\theta}^4\Big(\frac{1}{\sqrt{n}} + \|\hat{\xi}_1 - \xi\|_{\max}\Big) + \bar{\theta}^2\Big(\frac{1}{\sqrt{n}} + |\hat{\xi}_1(i) - \xi_1(i)|\Big) + \big|e_i(W \oslash \widehat{N})\hat{\xi}_1\big|
\end{aligned}
$$

Here we crudely bound

$$
\|e_i'[(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}]\|_1 \leq \|e_i'[(\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta) \circ \widetilde{\Omega} \circ N]\|_1 \leq Cn\bar{\theta}^4 .
$$

and $\|N \circ \widetilde{\Omega} \oslash \widehat{N}\|_{\max} \leq C\|\widetilde{\Omega}\|_{\max} \leq C\bar{\theta}^2$ by the fact that all entries in N and $\widehat{N}$ are lowered bound by a positive constant.

Regarding the last term on the RHS, i.e., $\big|e_i(W \oslash \widehat{N})\hat{\xi}_1\big|$, we have

$$
\big|e_i(W \oslash \widehat{N})\hat{\xi}_1\big| \leq \big|e_i(W \oslash N)\hat{\xi}_1\big| + \big|e_i'W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\hat{\xi}_1\big|
$$

We study the second term below. Note that $W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') = W \circ (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)$. We bound

$$
\begin{aligned}
\big|e_i'W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\hat{\xi}_1\big| \leq\ & |e_i'(W \circ \widehat{\Theta}\widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\widehat{\Theta})\hat{\xi}_1| + |e_i'(W \circ \widehat{\Theta}(\widehat{\Pi} - \Pi)P\widehat{\Pi}'\widehat{\Theta})\hat{\xi}_1| \\
&+ |e_i'(W \circ \widehat{\Theta}\Pi P(\widehat{\Pi} - \Pi)'\widehat{\Theta})\hat{\xi}_1| + |e_i'(W \circ (\widehat{\Theta} - \Theta)\Pi P\Pi'\widehat{\Theta})\hat{\xi}_1| \\
&+ |e_i'(W \circ \Theta\Pi P\Pi'(\widehat{\Theta} - \Theta))\hat{\xi}_1|
\end{aligned}
$$

For each term, we further have

$$
\begin{aligned}
|e_i'(W \circ \widehat{\Theta}\widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\widehat{\Theta})\hat{\xi}_1| &\leq |e_i'\widehat{\Theta}(W \circ \widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}')\widehat{\Theta}\hat{\xi}_1| \leq \bar{\theta}^2\|e_i'W \circ \widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\|_1\|\hat{\xi}_1\|_{\max} \\
&\leq \bar{\theta}^2\|e_i'W\|\|e_i'\widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\|\|\hat{\xi}_1\|_{\max} \leq n\bar{\theta}^3\|\widehat{P} - P\|_{\max}\|\hat{\xi}_1\|_{\max} \\
|e_i'(W \circ \widehat{\Theta}(\widehat{\Pi} - \Pi)P\widehat{\Pi}'\widehat{\Theta})\hat{\xi}_1| &\leq C\bar{\theta}\|e_i'W\|\|e_i'(\widehat{\Pi} - \Pi)P\widehat{\Pi}'\widehat{\Theta}) \circ \hat{\xi}_1\| \leq C\bar{\theta}^2\|e_i'W\|\|(\widehat{\Pi} - \Pi)P\widehat{\Pi}'\|_{\max}\|\hat{\xi}_1\| \\
&\leq C\bar{\theta}^2\|e_i'W\| \leq C\bar{\theta}^2\sqrt{n\bar{\theta}^2} \\
|e_i'(W \circ \widehat{\Theta}\Pi P(\widehat{\Pi} - \Pi)'\widehat{\Theta})\hat{\xi}_1| &\leq C\bar{\theta}\|e_iW\|\|e_i'\Pi P(\widehat{\Pi} - \Pi)'\widehat{\Theta}) \circ \hat{\xi}_1\| \leq C\bar{\theta}^2\|e_iW\| \leq C\bar{\theta}^2\sqrt{n\bar{\theta}^2} \\
|e_i'(W \circ (\widehat{\Theta} - \Theta)\Pi P\Pi'\widehat{\Theta})\hat{\xi}_1| &\leq |\hat{\theta}_i - \theta_i||e_i'(W \circ \Pi P\Pi')\widehat{\Theta}\hat{\xi}_1| \leq C\bar{\theta}^2\|e_i'(W \circ \Pi P\Pi')\| \leq C\bar{\theta}^2\sqrt{n\bar{\theta}^2} \\
|e_i'(W \circ \Theta\Pi P\Pi'(\widehat{\Theta} - \Theta))\hat{\xi}_1| &\leq C\theta_i|e_i'(W \circ \Pi P\Pi')(\widehat{\Theta} - \Theta)\hat{\xi}_1| \leq C\bar{\theta}^2\|e_i'(W \circ \Pi P\Pi')\| \leq C\bar{\theta}^2\sqrt{n\bar{\theta}^2}
\end{aligned}
$$

Combining all these inequalities, we arrive at

$$
\begin{aligned}
\frac{|e_i'\widetilde{W}\hat{\xi}_1|}{n\bar{\theta}^2} \leq\ & \frac{C}{n}\Big(\frac{1}{\sqrt{n}} + |\hat{\xi}_1(i) - \xi_1(i)|\Big) + C(\bar{\theta}^2 + \bar{\theta}\|\widehat{P} - P\|_{\max})\Big(\frac{1}{\sqrt{n}} + \|\hat{\xi}_1 - \xi_1\|_{\max}\Big) \\
&+ \frac{C\bar{\theta}}{\sqrt{n}} + \frac{\big|e_i'(W \oslash N)\hat{\xi}_1\big|}{n\bar{\theta}^2} \qquad (10)
\end{aligned}
$$

In the sequel, we analyze $\big|e_i'(W \oslash N)\hat{\xi}_1\big|$ by leave-one-out technique. Let $\hat{\xi}_1^{(i)}$ be the first eigenvector of

$$
A^{(i)} \oslash N = \widetilde{\Omega} - \mathrm{diag}(\widetilde{\Omega}) + W^{(i)} \oslash N
$$

where $W^{(i)}$ is obtained by zeroing out the i-th row and column of W. Without loss of generality, we assume $\mathrm{sgn}(\hat{\xi}_1'\hat{\xi}_1^{(i)}) = 1$. Thus,

$$
\begin{aligned}
\big|e_i'(W \oslash N)\hat{\xi}_1\big| &\leq \big|e_i'(W \oslash N)\hat{\xi}_1^{(i)}\big| + \|e_i(W \oslash N)\|\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\| \\
&\leq C\Big(\bar{\theta}\sqrt{\log(n)} + \|\hat{\xi}_1^{(i)}\|_{\max}\log(n) + \sqrt{n\bar{\theta}^2}\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\|\Big) \\
&\leq C\Big(\bar{\theta}\sqrt{\log(n)} + \|\hat{\xi}_1\|_{\max}\log(n) + \sqrt{n\bar{\theta}^2}\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\|\Big)
\end{aligned}
$$

where we applied Bernstein inequality on $e_i'(W \oslash N)\hat{\xi}_1^{(i)}$ as $e_i'(W \oslash N)$ is independent of $\hat{\xi}_1^{(i)}$. And the last step is due to the derivation

$$\|\hat{\xi}_1^{(i)}\|_{\max} \log(n) \le \big(\|\hat{\xi}_1\|_{\max} + \|\hat{\xi}_1^{(i)} - \hat{\xi}_1\|_{\max}\big) \log(n) \le \|\hat{\xi}_1\|_{\max} \log(n) + \|\hat{\xi}_1^{(i)} - \hat{\xi}_1\| \sqrt{n\bar{\theta}^2}$$

under the condition that $\sqrt{n\bar{\theta}^2}\lambda_{\min}(P) \ge c_3 \log(n)$. Next, by sine-theta theorem, we bound

$$\begin{aligned}
\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\| &\le \frac{\|\big(\widetilde{W} - W^{(i)} \oslash N + \mathrm{diag}(\widetilde{\Omega})\big)\hat{\xi}_1\|}{n\bar{\theta}^2} \\
&\le \frac{\|e_i e_i'(W \oslash N)\hat{\xi}_1 + (W \oslash N)e_i e_i'\hat{\xi}_1\|}{n\bar{\theta}^2} + \frac{\|\big((N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\big)\hat{\xi}_1\|}{n\bar{\theta}^2} \\
&\quad + \frac{\|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\hat{\xi}_1\|}{n\bar{\theta}^2} + \frac{C}{n}
\end{aligned} \tag{11}$$

where we used the decomposition

$$\widetilde{W} = (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega} - \mathrm{diag}(N \circ \widetilde{\Omega} \oslash \widehat{N}) + W \oslash N + W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')$$

and the crude bound

$$\big\|\big[\mathrm{diag}(\widetilde{\Omega}) - \mathrm{diag}(N \circ \widetilde{\Omega} \oslash \widehat{N})\big]\hat{\xi}_1\big\| \le C\bar{\theta}^2 \|\hat{\xi}_1\| \le C\bar{\theta}^2$$

following from $N \oslash \widehat{N} \le C$ with high probability and $\|\widetilde{\Omega}\|_{\max} \le C\bar{\theta}^2$. To proceed, we further bound

$$\frac{\|\big((N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\big)\hat{\xi}_1\|}{n\bar{\theta}^2} \le \frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|}{n\bar{\theta}^2} \le \frac{\|N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n'\|_F}{n} \tag{12}$$

And we analyze the upper bound for $\|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\hat{\xi}_1\|$ below. By definition,

$$\begin{aligned}
\|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\hat{\xi}_1\| &= \sqrt{\sum_i \Big(\sum_j (W \oslash N)_{ij}(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')_{ij}\hat{\xi}_1(j)\Big)^2} \\
&\le \sqrt{\sum_i \sum_j (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')_{ij}^2 \cdot \sum_j (W \oslash N)_{ij}^2 \hat{\xi}_1(j)^2} \\
&\le \sqrt{\sum_{i,j} (N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')_{ij}^2 \cdot \max_i \sum_j (W \oslash N)_{ij}^2 \|\hat{\xi}_1\|_{\max}^2} \\
&\le \|N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n'\|_F \|\hat{\xi}_1\|_{\max} \sqrt{\max_i \sum_j (W \oslash N)_{ij}^2}
\end{aligned}$$

where in the second step we used Cauchy-Schwarz inequality. Regarding the last factor inside the square root, for each fixed i, it is a sum of independent r.v.s, so we can use Bernstein inequality to get its high probability bound. Specifically, fixed an i, as N is deterministic and each entry of N is $\asymp 1$, we can derive the mean of $\sum_j (W \oslash N)_{ij}^2$ is given by

$$\sum_j \mathbb{E}(W \oslash N)_{ij}^2 \asymp \sum_j \theta_i\theta_j \asymp n\bar{\theta}^2\,;$$

And the variance can be estimated by

$$\sum_j \mathrm{var}(W \oslash N)_{ij}^2 \le \sum_j \mathbb{E}(W \oslash N)_{ij}^4 \le Cn\bar{\theta}^2\,.$$

Consequently, by Bernstein inequality, it is not hard to derive

$$\Big|\sum_j (W \oslash N)_{ij}^2 - \mathbb{E}(W \oslash N)_{ij}^2\Big| \le C\sqrt{n\bar{\theta}^2 \log(n)} + C\log(n) \ll Cn\bar{\theta}^2$$

with probability $1 - o(n^{-4})$. Then, combining all i, it gives that

$$\max_i \sum_j (W \oslash N)_{ij}^2 \le Cn\bar{\theta}^2$$

with probability $1 - o(n^{-3})$. We thus obtain that

$$\|W \oslash N \circ (N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n')\hat{\xi}_1\| \le C\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F \cdot \sqrt{n}\bar{\theta}\|\hat{\xi}_1\|_{\max} \tag{13}$$

Next, we study the bound of $\|e_i e_i'(W \oslash N)\hat{\xi}_1 + (W \oslash N)e_i e_i'\hat{\xi}_1\|$ below.

$$\begin{aligned}
&\|e_i e_i'(W \oslash N)\hat{\xi}_1 + (W \oslash N)e_i e_i'\hat{\xi}_1\| \\
&\le |e_i'(W \oslash N)\hat{\xi}_1| + \|e_i'(W \oslash N)\||\hat{\xi}_1(i)| \\
&\le C\Big(\bar{\theta}\sqrt{\log(n)} + \|\hat{\xi}_1\|_{\max}\log(n) + \sqrt{n\bar{\theta}^2}\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\| + \sqrt{n\bar{\theta}^2}|\hat{\xi}_1(i)|\Big).
\end{aligned} \tag{14}$$

Combining (11) - (14), we get

$$\begin{aligned}
\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\| \le C\Bigg(&\frac{\sqrt{\log(n)}}{n\bar{\theta}} + \|\hat{\xi}_1\|_{\max}\frac{\log(n)}{n\bar{\theta}^2} + \frac{\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\|}{\sqrt{n\bar{\theta}^2}} \\
&+ \frac{|\hat{\xi}_1(i)|}{\sqrt{n\bar{\theta}^2}} + \frac{\|\hat{\xi}_1\|_{\max}\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{\sqrt{n\bar{\theta}^2}}\Bigg)
\end{aligned}$$

Rearranging both sides gives that

$$\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\| \le C\Bigg(\frac{\sqrt{\log(n)}}{n\bar{\theta}} + \|\hat{\xi}_1\|_{\max}\frac{1 + \|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{\sqrt{n\bar{\theta}^2}}\Bigg)$$

Consequently,

$$\begin{aligned}
\frac{\big|e_i(W \oslash N)\hat{\xi}_1\big|}{n\bar{\theta}^2} &\le C\Bigg(\frac{\sqrt{\log(n)}}{n\bar{\theta}} + \frac{\|\hat{\xi}_1\|_{\max}\log(n)}{n\bar{\theta}^2} + \frac{\|\hat{\xi}_1 - \hat{\xi}_1^{(i)}\|}{\sqrt{n\bar{\theta}^2}}\Bigg) \\
&\le C\Bigg(\frac{\sqrt{\log(n)}}{n\bar{\theta}} + \|\hat{\xi}_1\|_{\max}\Big[\frac{\log(n)}{n\bar{\theta}^2} + \frac{\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{n\bar{\theta}^2}\Big]\Bigg)
\end{aligned}$$

Plugging this and (10) into (9), we have

$$\begin{aligned}
|\hat{\xi}_1(i) - \xi_1(i)| \le\ & C\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n + \sqrt{\lambda_1(\widetilde{\Omega})}}{\lambda_1(\widetilde{\Omega})} \cdot \frac{1}{\sqrt{n}} + \frac{C\bar{\theta}}{\sqrt{n}} \\
&+ C(\bar{\theta}^2 + \bar{\theta}\|\widehat{P} - P\|_{\max})\Big(\frac{1}{\sqrt{n}} + \|\hat{\xi}_1 - \xi\|_{\max}\Big) + \frac{C}{n}|\hat{\xi}_1(i) - \xi_1(i)| \\
&+ C\Bigg(\frac{\sqrt{\log(n)}}{n\bar{\theta}} + \|\hat{\xi}_1\|_{\max}\Big[\frac{\log(n)}{n\bar{\theta}^2} + \frac{\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{n\bar{\theta}^2}\Big]\Bigg)
\end{aligned}$$

Rearranging both sides, together with $\|\hat{\xi}_1\|_{\max} \le \|\xi_1\|_{\max} + \|\hat{\xi}_1 - \xi_1\|_{\max} \le C/\sqrt{n} + \|\hat{\xi}_1 - \xi_1\|_{\max}$, gives rise to

$$\begin{aligned}
&|\hat{\xi}_1(i) - \xi_1(i)| \\
&\le C\Big(\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n}{n\bar{\theta}^2} + \frac{\sqrt{\log(n)}}{\sqrt{n\bar{\theta}^2}} + \frac{\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{n\bar{\theta}^2} + \bar{\theta} + \bar{\theta}\|\widehat{P} - P\|_{\max}\Big)\frac{1}{\sqrt{n}} \\
&\quad + C\Big(\frac{\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{n\bar{\theta}^2} + \bar{\theta}^2 + \bar{\theta}\|\widehat{P} - P\|_{\max} + \frac{\log(n)}{n\bar{\theta}^2}\Big)\|\hat{\xi}_1 - \xi_1\|_{\max}
\end{aligned}$$

We further take maximum for both sides. Under the conditions that $\|(N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\| \ll n\bar{\theta}^2$, $\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F \ll n\bar{\theta}^2$ and $n\bar{\theta}^2 \ge C(\log(n))^2$, together with $\bar{\theta}\|\widehat{P} - P\|_{\max} \ll \lambda_{\min}(P) \le C$ and $\bar{\theta} = o(1)$, it yields that with probability $1 - o(n^{-3})$,

$$\begin{aligned}
\|\hat{\xi}_1 - \xi_1\|_{\max} &\le \frac{C}{\sqrt{n}}\Big(\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n') \circ \widetilde{\Omega}\| + \tau_n}{n\bar{\theta}^2} + \frac{\sqrt{\log(n)}}{\sqrt{n\bar{\theta}^2}} + \frac{\|N \oslash \widehat{N} - \mathbf{1}_n \mathbf{1}_n'\|_F}{n\bar{\theta}^2} \\
&\qquad + \bar{\theta}^2 + \bar{\theta}\|\widehat{P} - P\|_{\max}\Big) \\
&\ll \frac{1}{\sqrt{n}}
\end{aligned}$$

Further by $|\xi_1(i)| \asymp 1/\sqrt{n}$ for $1 \leq i \leq n$, we deduce that $|\hat{\xi}_1(i)| \asymp 1/\sqrt{n}$ for $1 \leq i \leq n$.

Now, by the definition of $\widehat{R}$, we can derive

$$
\begin{aligned}
\|\widehat{R}O - R\|_F^2 &= \|\mathrm{diag}(\hat{\xi}_1)^{-1}(\widehat{\Xi}_1 O - \Xi) - \big[\mathrm{diag}(\hat{\xi}_1)^{-1} - \mathrm{diag}(\xi_1)^{-1}\big]\Xi_1\|_F^2 \\
&\leq C\Big(\|\mathrm{diag}(\hat{\xi}_1)^{-1}(\widehat{\Xi}_1 O - \Xi)\|_F^2 + \|\big[\mathrm{diag}(\hat{\xi}_1)^{-1} - \mathrm{diag}(\xi_1)^{-1}\big]\Xi_1\|_F^2\Big) \\
&\leq C\Big(n\|\widehat{\Xi}_1 O - \Xi\|_F^2 + \|\hat{\xi}_1 - \xi_1\|^2\|\mathrm{diag}(\hat{\xi}_1)^{-1}\mathrm{diag}(\xi_1)^{-1}\Xi_1\|_{2\to\infty}^2\Big) \\
&\leq C\Big(n\|\widehat{\Xi}_1 O - \Xi\|_F^2 + n\,\|\hat{\xi}_1 - \xi_1\|^2\Big)
\end{aligned}
$$

By (8), we conclude that

$$
\|\widehat{R}O - R\|_F^2 \leq Cn\frac{\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\circ\Omega\|^2 + \tau_n^2 + \lambda_1(\widetilde{\Omega})}{\big|\lambda_K(\widetilde{\Omega})\big|^2}
$$

with probability $1 - o(n^{-3})$. Now, combining the above result with Lemma B.1, we conclude the proof of Lemma 3.1.

C.2 The error rate of θ

In this subsection, we prove the error rate for refitting θ under the assumptions in Theorem 3.2.The results is collected in the following Lemma.

Lemma C.1 *Under the assumptions in Theorem 3.2, it holds that with probability* $1 - o(n^{-3})$,

$$
\begin{aligned}
&\big|\hat{\theta}_i - \theta_i\big| \leq C\Big(\sqrt{\log(n)/n} + r_n/\bar{\theta}\Big), && \text{if} \quad \hat{\pi}_i = \pi_i; \\
&\big|\hat{\theta}_i - P_{kk_0}\theta_i\big| \leq C\Big(\bar{\theta}(\log(n)/n\bar{\theta}^2)^{1/4} + \sqrt{r_n}\Big), && \text{if} \quad \hat{\pi}_i = e_k \neq e_{k_0} = \pi_i.
\end{aligned}
$$

where r_n *is the Hamming error of the* $\widehat{\Pi}$ *by directly SCORE and* $\{e_k\}_{k=1}^K$ *represents the standard basis of* $\mathbb{R}^K$.

We prove Lemma C.1 below.

Recall the refitting formula for θ:

$$
\hat{\theta}_i = \sqrt{\frac{\sum_{j\neq t\in\hat{S}_{k,i}} A_{ij}(1 - A_{jt})A_{ti}}{\sum_{j\neq t\in\hat{S}_{k,i}}(1 - A_{ij})A_{jt}(1 - A_{ti})}} \tag{15}
$$

where $i \in \widehat{\mathcal{C}}_k$ and $\hat{S}_{k,i} = \widehat{\mathcal{C}}_k \setminus \{i\}$. Using the error rate of $\widehat{\Pi}$ from SCORE, i.e., $\|\widehat{\Pi} - \Pi\|_1 \leq nr_n$, we first crudely bound the numerator and denominator in the expression of $\hat{\theta}_i$. Let $\hat{1}_k$ and 1_k denote the k0th column of $\widehat{\Pi}$ and Π, respectively.

$$
\begin{aligned}
&\sum_{j\neq t\in\hat{S}_{k,i}} A_{ij}(1 - A_{jt})A_{ti} \\
&= e_{n,i}'A\mathrm{diag}\big(\hat{1}_k\big)(\mathbf{1}_n\mathbf{1}_n' - I_n - A)\mathrm{diag}\big(\hat{1}_k\big)Ae_{n,i} \\
&= e_{n,i}'A\mathrm{diag}\big(1_k\big)(\mathbf{1}_n\mathbf{1}_n' - I_n - A)\mathrm{diag}\big(1_k\big)Ae_{n,i} + e_{n,i}'A\mathrm{diag}\big(\hat{1}_k - 1_k\big)(\mathbf{1}_n\mathbf{1}_n' - I_n - A)\mathrm{diag}\big(1_k\big)Ae_{n,i} \\
&\quad + e_{n,i}'A\mathrm{diag}\big(\hat{1}_k\big)(\mathbf{1}_n\mathbf{1}_n' - I_n - A)\mathrm{diag}\big(\hat{1}_k - 1_k\big)Ae_{n,i}
\end{aligned}
$$

For the second and third terms on the RHS above, we bound

$$
\begin{aligned}
&\big|e_{n,i}'A\mathrm{diag}\big(\hat{1}_k - 1_k\big)(\mathbf{1}_n\mathbf{1}_n' - I_n - A)\mathrm{diag}\big(1_k\big)Ae_{n,i} + e_{n,i}'A\mathrm{diag}\big(\hat{1}_k\big)(\mathbf{1}_n\mathbf{1}_n' - I_n - A)\mathrm{diag}\big(\hat{1}_k - 1_k\big)Ae_{n,i}\big| \\
&\leq 2\|e_{n,i}'A\|_1 \cdot \|e_{n,i}'A\mathrm{diag}\big(\hat{1}_k - 1_k\big)\|_1 \\
&\leq Cn\bar{\theta}^2\|\hat{1}_k - 1_k\|_1 \leq Cn\bar{\theta}^2\|\widehat{\Pi} - \Pi\|_1 \leq Cn^2 r_n\bar{\theta}^2
\end{aligned}
$$

where to obtain the third line, we used

$$\|e_{n,i}'A\|_1 = \sum_{j\neq i} A_{ij} = \sum_{j\neq i}\Omega_{ij} + \sum_{j\neq i} W_{ij} = n\bar{\theta}^2 + O(\sqrt{n\bar{\theta}^2\log(n)}\,)$$
$$\|e_{n,i}'A\mathrm{diag}\big(\hat{1}_k - 1_k\big)\|_1 \leq \|e_{n,i}'A\|_{\max}\|\hat{1}_k - 1_k\|_1 \leq \|\hat{1}_k - 1_k\|_1$$

simultaneously for all $1 \leq i \leq n$, with probability $1 - o(n^{-3})$. Here Bernstein inequality is employed to derive

$$\Big|\sum_{j\neq i} W_{ij}\Big| \leq C\big(\sqrt{\sum_{j\neq i}\mathrm{var}(W_{ij})\log(n)} + \log(n)\big) \leq C\sqrt{n\bar{\theta}^2\log(n)}$$

by noting that W_{ij} is with mean 0 and variance

$$\mathrm{var}(W_{ij}) = \Omega_{ij}(1-\Omega_{ij}) \leq \theta_i\theta_j$$

and the condition $n\bar{\theta}^2 \geq C\log(n)$. Therefore,

$$\Big|\sum_{j\neq t\in\hat{S}_{k,i}} A_{ij}(1-A_{jt})A_{ti} - \sum_{j\neq t\in S_{k,i}} A_{ij}(1-A_{jt})A_{ti}\Big| \leq Cn^2 r_n\bar{\theta}^2\,. \tag{16}$$

simultaneously for all $1 \leq i \leq n$, with probability $1 - o(n^{-3})$. Similarly, we can show that

$$\Big|\sum_{j\neq t\in\hat{S}_{k,i}} (1-A_{ij})A_{jt}(1-A_{ti}) - \sum_{j\neq t\in S_{k,i}} (1-A_{ij})A_{jt}(1-A_{ti})\Big|$$
$$\leq C\|\hat{1}_k - 1_k\|_1\cdot n\bar{\theta}^2 \leq C\|\hat{\Pi} - \Pi\|_1 n\bar{\theta}^2 \leq n^2 r_n\bar{\theta}^2\,. \tag{17}$$

To proceed, we study $\sum_{j\neq t\in S_{k,i}} A_{ij}(1-A_{jt})A_{ti}$ and $\sum_{j\neq t\in S_{k,i}}(1-A_{ij})A_{jt}(1-A_{ti})$ instead. Recall the decomposition $A = \Omega - \mathrm{diag}(\Omega) - W$. For the numerator, we decompose

$$\begin{aligned}
&\sum_{j\neq t\in S_{k,i}} A_{ij}(1-A_{jt})A_{ti} \\
&= \sum_{j\neq t\in S_{k,i}} \Omega_{ij}(1-\Omega_{jt})\Omega_{ti} + \sum_{j\neq t\in S_{k,i}} W_{ij}(1-\Omega_{jt})\Omega_{ti} + \Omega_{ij}(-W_{jt})\Omega_{ti} + \Omega_{ij}(1-\Omega_{jt})W_{ti} \\
&\quad + \sum_{j\neq t\in S_{k,i}} W_{ij}(-W_{jt})\Omega_{ti} + \Omega_{ij}(-W_{jt})W_{ti} + W_{ij}(1-\Omega_{jt})W_{ti} \\
&\quad + \sum_{j\neq t\in S_{k,i}} W_{ij}(-W_{jt})W_{ti} \\
&=: \sum_{j\neq t\in S_{k,i}} \Omega_{ij}(1-\Omega_{jt})\Omega_{ti} + \sum_{a=1}^{3} T_{1a} + \sum_{a=1}^{3} T_{2a} + T_3
\end{aligned}$$

We analyze each term on the RHS above one by one. Note that W_{ij} is with mean 0 and variance

$$\mathrm{var}(W_{ij}) = \Omega_{ij}(1-\Omega_{ij}) \leq \theta_i\theta_j$$

and the trivial bound $|W_{ij}\sum_{t\neq j}(1-\Omega_{jt})\Omega_{ti}| < n\bar{\theta}^2$ (similarly for each summand in T_{12} and T_{13}). By Bernstein inequality,

$$|T_{11}| \leq C\Big(\sqrt{\sum_j \theta_i\theta_j\big(\sum_{t\neq j}(1-\Omega_{jt})\Omega_{ti}\big)^2\log(n)} + n\bar{\theta}^2\log(n)\Big) \leq Cn\bar{\theta}^2\big(\sqrt{n\bar{\theta}^2\log(n)} + \log(n)\big)$$

$$|T_{12}| \leq 2C\Big(\sqrt{\sum_{j<t}\theta_j\theta_t(\Omega_{ij})^2(\Omega_{ti})^2\log(n)} + \log(n)\Big) \leq C\big(n\bar{\theta}^5\sqrt{\log(n)} + \log(n)\big)$$

$$|T_{13}| \leq C\Big(\sqrt{\sum_t \theta_t\theta_i\big(\sum_{j\neq t}\Omega_{ij}(1-\Omega_{jt})\big)^2\log(n)} + n\bar{\theta}^2\log(n)\Big) \leq Cn\bar{\theta}^2\big(\sqrt{n\bar{\theta}^2\log(n)} + \log(n)\big)$$

simultaneously for all $1 \le i \le n$, with probability $1 - o(n^{-3})$. Consequently,

$$\Big|\sum_{a=1}^{3} T_{1a}\Big| \le Cn\bar{\theta}^2\sqrt{n\bar{\theta}^2\log(n)}\,.$$

following from the condition that $n\bar{\theta}^2 \ge C\log(n)$, which is implied by condition (16) in the manuscript.

Regarding T_{2a} for $a = 1, 2, 3$ and T_3, their large deviation bounds can be tackled by decoupling inequality for U-statistics in de la Pena & Montgomery-Smith (1995). Specifically, implied by this technique, the large deviation bound of T_{21} is dominated by that of

$$\tilde{T}_{21} := \sum_{j\neq t\in S_{k,i}} W_{ij}(-W_{jt}^{(1)})\Omega_{ti}$$

where $W^{(1)}$ is an i.i.d. copy of W. Thanks to this independence, we first condition on $W^{(1)}$ and use Bernstein inequality to get

$$|\tilde{T}_{21}| \le C\Big[\sqrt{\sum_{j\in S_{k,i}}\theta_i\theta_j\big(\sum_{t\neq j\in S_{k,i}}\Omega_{ti}W_{jt}^{(1)}\big)^2\log(n)} + \max_j\Big|\sum_{t\neq j\in S_{k,i}}\Omega_{ti}W_{jt}^{(1)}\Big|\log(n)\Big]$$

Next, by Bernstein inequality again, we obtain

$$\Big|\sum_{t\neq j\in S_{k,i}}\Omega_{ti}W_{jt}^{(1)}\Big| \le C\Big(\sqrt{n\bar{\theta}^6\log(n)} + \bar{\theta}^2\log(n)\Big)$$

Combining the above inequalities, we arrive at

$$|\tilde{T}_{21}| \le C\big(n\bar{\theta}^4\mathrm{log}(n) + \sqrt{n\bar{\theta}^2}\bar{\theta}^2(\mathrm{log}(n))^{3/2}\big)\,.$$

simultaneously for all $1 \le i \le n$, with probability $1 - o(n^{-3})$. Therefore,

$$|T_{21}| \le C\big(n\bar{\theta}^4\mathrm{log}(n) + \sqrt{n\bar{\theta}^2}\bar{\theta}^2(\mathrm{log}(n))^{3/2}\big)\,.$$

In the same manner, we can show that

$$|T_{22}| \le C\big(n\bar{\theta}^4\mathrm{log}(n) + \sqrt{n\bar{\theta}^2}\bar{\theta}^2(\mathrm{log}(n))^{3/2}\big), \qquad |T_{23}| \le Cn\bar{\theta}^2\mathrm{log}(n)$$

As a result,

$$\Big|\sum_{a=1}^{3} T_{2a}\Big| \le Cn\bar{\theta}^2\mathrm{log}(n)\,.$$

under the condition $n\bar{\theta}^2 \ge C\log(n)$ and $\bar{\theta} \le C$.

Lastly, we prove the large deviation bound for T_3. Using decoupling inequality for U-statistics, it suffices to prove a large deviation bound for $\tilde{T}_3$ with

$$\tilde{T}_3 := \sum_{j\neq t\in S_{k,i}} W_{ij}(-W_{jt}^{(1)})W_{ti}^{(2)}$$

Here $W^{(1)}$ and $W^{(2)}$ are two i.i.d. copies of W. Condition on $W^{(1)}, W^{(2)}$, by Bernstein inequality,

$$|\tilde{T}_3| \le C\Big(\sqrt{\sum_{j\in S_{k,i}}\theta_i\theta_j\big(\sum_{t\neq j\in S_{k,i}}W_{jt}^{(1)}W_{ti}^{(2)}\big)^2\log(n)} + \max_j\Big|\sum_{t\neq j\in S_{k,i}}W_{jt}^{(1)}W_{ti}^{(2)}\Big|\log(n)\Big)$$

In addition, for $\sum_{t\neq j\in S_{k,i}}W_{jt}^{(1)}W_{ti}^{(2)}$, each summand is independent of each other. By Bernstein inequality, we can similarly get

$$\Big|\sum_{t\neq j\in S_{k,i}}W_{jt}^{(1)}W_{ti}^{(2)}\Big| \le C\Big(\sqrt{n\bar{\theta}^4\log(n)} + \log(n)\Big)$$

Consequently,

$$|\tilde{T}_3| \le C\Big(n\bar{\theta}^3 \log(n) + \sqrt{n\bar{\theta}^2}(\log(n))^{3/2}\Big)$$

This, by $n\bar{\theta}^2 \ge C\log(n)$ and $\bar{\theta} \le C$, further implies

$$|T_3| \le Cn\bar{\theta}^2 \log(n)$$

simultaneously for all $1 \le i \le n$, with probability $1 - o(n^{-3})$. Based on the large deviation bounds for $\sum_{a=1}^3 T_{1a}, \sum_{a=1}^3 T_{2a}$ and T_3, we therefore conclude that

$$\sum_{j \ne t \in S_{k,i}} A_{ij}(1 - A_{jt})A_{ti} = \sum_{j \ne t \in S_{k,i}} \Omega_{ij}(1 - \Omega_{jt})\Omega_{ti} + O_p\big((n\bar{\theta}^2)^{3/2}\sqrt{\log(n)}\,\big) \tag{18}$$

simultaneously for all i, where the probability is $1 - o(n^{-3})$.

Next, we analyze the denominator $\sum_{j \ne t \in S_{k,i}}(1 - A_{ij})A_{jt}(1 - A_{ti})$. Analogously, we decompose

$$\begin{aligned}
&\sum_{j \ne t \in S_{k,i}} (1 - A_{ij})A_{jt}(1 - A_{ti}) \\
&= \sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti}) \\
&\quad + \sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(-W_{ti}) + (1 - \Omega_{ij})W_{jt}(1 - \Omega_{ti}) + (-W_{ij})\Omega_{jt}(1 - \Omega_{ti}) \\
&\quad + \sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})W_{jt}(-W_{ti}) + (-W_{ij})W_{jt}(1 - \Omega_{ti}) + (-W_{ij})\Omega_{jt}(-W_{ti}) \\
&\quad + \sum_{j \ne t \in S_{k,i}} (-W_{ij})W_{jt}(-W_{ti}) \\
&=: \sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti}) + \sum_{a=1}^{3} \mathcal{T}_{1a} + \sum_{a=1}^{3} \mathcal{T}_{2a} - T_3
\end{aligned}$$

Similarly to $\sum_{a=1}^3 T_{1a}$ and $\sum_{a=1}^3 T_{2a}$, we can derive

$$\begin{aligned}
&|\mathcal{T}_{11}| \le Cn\bar{\theta}^2\sqrt{n\bar{\theta}^2 \log(n)}, \quad |\mathcal{T}_{12}| \le Cn\bar{\theta}\sqrt{\log(n)}, \quad |\mathcal{T}_{13}| \le Cn\bar{\theta}^2\sqrt{n\bar{\theta}^2 \log(n)} \\
&|\mathcal{T}_{21}| \le Cn\bar{\theta}^2 \log(n), \quad |\mathcal{T}_{22}| \le Cn\bar{\theta}^2 \log(n), \quad |\mathcal{T}_{23}| \le Cn\bar{\theta}^4 \log(n)
\end{aligned}$$

by Bernstein inequality and decoupling inequality. Since the details are rather similar, we omit the details. The above estimates, together with $|T_3| \le Cn\bar{\theta}^2\log(n)$, give rise to

$$\sum_{j \ne t \in S_{k,i}} (1 - A_{ij})A_{jt}(1 - A_{ti}) = \sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti}) + O_p\big([(n\bar{\theta}^2)^{3/2} + n\bar{\theta}]\sqrt{\log(n)}\,\big) \tag{19}$$

simultaneously for all $1 \le i \le n$, with probability $1 - o(n^{-3})$.

Combining (16), (17), (18) and (19) into (15), we can further derive that

$$\begin{aligned}
\hat{\theta}_i &= \sqrt{\frac{\sum_{j \ne t \in \hat{S}_{k,i}} A_{ij}(1 - A_{jt})A_{ti}}{\sum_{j \ne t \in \hat{S}_{k,i}} (1 - A_{ij})A_{jt}(1 - A_{ti})}} \\
&= \sqrt{\frac{\sum_{j \ne t \in S_{k,i}} \Omega_{ij}(1 - \Omega_{jt})\Omega_{ti} + O_p\big((n\bar{\theta}^2)^{3/2}\sqrt{\log(n)} + n^2 r_n \bar{\theta}^2\big)}{\sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti}) + O_p\big([(n\bar{\theta}^2)^{3/2} + n\bar{\theta}]\sqrt{\log(n)} + n^2 r_n \bar{\theta}^2\big)}} \\
&= \sqrt{\frac{\sum_{j \ne t \in S_{k,i}} \Omega_{ij}(1 - \Omega_{jt})\Omega_{ti}}{\sum_{j \ne t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti})} + O_p\big(\bar{\theta}\sqrt{\log(n)/n} + r_n\big)}
\end{aligned} \tag{20}$$

simultaneously for all $1 \leq i \leq n$, where the high probability is at least $1 - o(n^{-3})$. Here we used the crude estimate

$$\sum_{j \neq t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti}) \asymp (n\bar{\theta})^2$$

under the assumption that the number of nodes in each community is balanced and the diagonal entries of P are one so that $\Omega_{jt} \propto \theta_j \theta_t$ if $j, t \in \mathcal{C}_k$.

To proceed, we separate the analysis into two cases: (1) $\hat{\pi} = \pi_0 = e_k$; (2) $\hat{\pi}_i = e_k \neq e_{k_0} = \pi_i$. This is because the leading term in $\hat{\theta}_i$ may vary with the two different cases.

For case (1), $i \in \mathcal{C}_k$, it follows that

$$\frac{\sum_{j \neq t \in S_{k,i}} \Omega_{ij}(1 - \Omega_{jt})\Omega_{ti}}{\sum_{j \neq t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti})} = \frac{\sum_{j \neq t \in S_{k,i}} N_{ij} N_{jt} N_{ti} \theta_i \theta_j \theta_t \theta_i}{\sum_{j \neq t \in S_{k,i}} N_{ij} N_{jt} N_{ti} \theta_j \theta_t} = \theta_i^2$$

which is also claimed by Lemma 2.2 . In light of this, further with the condition in Theorem 3.2 that $r_n \ll \bar{\theta}^2$ (note that $r_n \asymp \delta_n$), we conclude that

$$\hat{\theta}_i = \theta_i + O_p\big(\sqrt{\log(n)/n} + r_n/\bar{\theta}\big) . \tag{21}$$

simultaneously for all $1 \leq i \leq n$, where the high probability is at least $1 - o(n^{-3})$.

For case (2), $i \notin \mathcal{C}_k$. Therefore, $\Omega_{ij} = N_{ij}\theta_i\theta_j \cdot P_{k_0 k}$ where $N_{ij} = (1 + \theta_i\theta_j \cdot P_{k_0 k})^{-1}$, for all $j \in \mathcal{C}_k$. As a result,

$$\frac{\sum_{j \neq t \in S_{k,i}} \Omega_{ij}(1 - \Omega_{jt})\Omega_{ti}}{\sum_{j \neq t \in S_{k,i}} (1 - \Omega_{ij})\Omega_{jt}(1 - \Omega_{ti})} = \frac{\sum_{j \neq t \in S_{k,i}} N_{ij} N_{jt} N_{ti} \theta_i \theta_j \theta_t \theta_i \cdot P_{k_0 k}^2}{\sum_{j \neq t \in S_{k,i}} N_{ij} N_{jt} N_{ti} \theta_j \theta_t} = \theta_i^2 P_{k_0 k}^2$$

Note that $P_{k_0 k} = \hat{\pi}_i' P \pi_i$ under this case and $\|P\|_{\max} \leq C$. We therefore conclude that

$$\begin{aligned}
\hat{\theta}_i &= e_{n,i}'(\widehat{\Pi} P \Pi') e_{n,i} \theta_i + O_p\Big(\min\Big\{ \bar{\theta}(\log(n)/n\bar{\theta}^2)^{1/4} + \sqrt{r_n}, \frac{\sqrt{\log(n)/n} + r_n/\bar{\theta}}{e_{n,i}'(\widehat{\Pi} P \Pi') e_{n,i}} \Big\} \Big) \\
&= e_{n,i}'(\widehat{\Pi} P \Pi') e_{n,i} \theta_i + O_p\big(\bar{\theta}(\log(n)/n\bar{\theta}^2)^{1/4} + \sqrt{r_n}\big) . \qquad (22)
\end{aligned}$$

simultaneously for all $1 \leq i \leq n$, where the high probability is at least $1 - o(n^{-3})$. By (21) and (22), we complete the proof.

C.3 THE ERROR RATE OF P

In this subsection, we prove the error rate of P, which is presented in the following lemma.

Lemma C.2 *Under the assumptions in Theorem 3.2, it hold with probability* $1 - o(n^{-3})$ *that*

$$\|\widehat{P} - P\|_{\max} \leq C\Big(\sqrt{\frac{\log(n)}{n\bar{\theta}^2}} + \frac{r_n}{\bar{\theta}^2}\Big)$$

where r_n *denotes the Hamming error by directly applying SCORE.*

Recalling the refitting formula of P,

$$\widehat{P}_{k\ell} = \frac{\sum_{i \in \widehat{\mathcal{C}}_k} \sum_{j \in \widehat{\mathcal{C}}_\ell} A_{ij}}{\sum_{i \in \widehat{\mathcal{C}}_k} \sum_{j \in \widehat{\mathcal{C}}_\ell} \hat{\theta}_i \hat{\theta}_j (1 - A_{ij})} \tag{23}$$

for $k \neq \ell$. We can rewrite it as

$$\widehat{P}_{k\ell} = \frac{\hat{1}_k' A \hat{1}_\ell}{\hat{1}_k' \widehat{\Theta}(\mathbf{1}_n \mathbf{1}_n' - A)\widehat{\Theta}\hat{1}_\ell}$$

For the numerator, we derive

$$\begin{aligned}
\hat{1}_k' A \hat{1}_\ell &= 1_k' \Omega 1_\ell + 1_k' W 1_\ell + (\hat{1}_k - 1_k)' A 1_\ell + 1_k' A(\hat{1}_\ell - 1_\ell) + (\hat{1}_k - 1_k)' A(\hat{1}_\ell - 1_\ell) \\
&= 1_k' \Omega 1_\ell + O_p(n\bar{\theta}\sqrt{\log(n)} + n^2 r_n \bar{\theta}^2) \qquad (24)
\end{aligned}$$

where the high probability is at least $1 - o(n^{-3})$. Here to get the RHS, we used the following estimates which can be obtained by employing Bernstein inequality,

$$|1_k' W 1_\ell| \le C\Big(\sqrt{\sum_{i\in\mathcal{C}_k, j\in\mathcal{C}_\ell} \theta_i\theta_j P_{k\ell}\log(n)} + \log(n)\Big) \le C\big(n\bar{\theta}\sqrt{P_{k\ell}\log(n)} + \log(n)\big),$$

$$|e_{n,i}' A 1_\ell - e_{n,i}'\Omega 1_\ell| \le C\Big(\sqrt{\sum_{j\in\mathcal{C}_\ell} \theta_i\theta_j\log(n)} + \log(n)\Big) \le C\big(\sqrt{n\bar{\theta}^2\log(n)} + \log(n)\big),$$

$$|e_{n,i}'\Omega 1_\ell| \le Cn\bar{\theta}^2\,.$$

simultaneously for all $1 \le i \le n$ and $1 \le k,\ell \le K$, with probability $1 - o(n^{-3})$. Note that the second and third inequalities also imply that $\max_i(e_{n,i}' A\mathbf{1}_n) \le C\bar{\theta}^2$ and furthermore,

$$\begin{aligned}
\big|(\hat{1}_k - 1_k)' A 1_\ell\big| &\le \|\hat{1}_k - 1_k\|_1 \max_i (e_{n,i}' A\mathbf{1}_n) \le Cn^2 r_n\bar{\theta}^2\\
|(\hat{1}_k - 1_k)' A(\hat{1}_\ell - 1_\ell)| &\le \|\hat{1}_k - 1_k\|_1 \max_i (e_{n,i}' A\mathbf{1}_n) \le Cn^2 r_n\bar{\theta}^2
\end{aligned}$$

Next for denominator, we first have

$$\begin{aligned}
&\hat{1}_k'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}\hat{1}_\ell\\
&= 1_k'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}1_\ell + (\hat{1}_k - 1_k)'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}\hat{1}_\ell + 1_k'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}(\hat{1}_\ell - 1_\ell)\\
&= 1_k'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}1_\ell + O_p(n^2 r_n\bar{\theta}^2)
\end{aligned}$$

Write $\Delta = \widehat{\Theta} - \mathrm{diag}(\widehat{\Pi}P\Pi')\Theta$. We further derive

$$\begin{aligned}
1_k'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}1_\ell =& 1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta(\mathbf{1}_n\mathbf{1}_n' - A)\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell\\
&+ 1_k'\Delta(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}1_\ell + 1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta(\mathbf{1}_n\mathbf{1}_n' - A)\Delta 1_\ell\\
&=: J_1 + J_2 + J_3
\end{aligned}$$

We analyze each term on the RHS as follows.

$$\begin{aligned}
J_1 &= 1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta(\mathbf{1}_n\mathbf{1}_n' - A)\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell\\
&= 1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta(\mathbf{1}_n\mathbf{1}_n' - \Omega)\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell - 1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta W\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell
\end{aligned}$$

where with probability $1 - o(n^{-3})$, since $\mathbf{1}_n\mathbf{1}_n' - \Omega = \mathbf{1}_n\mathbf{1}_n' - \widetilde{\Omega}\circ N = N$,

$$\begin{aligned}
&1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta N\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell\\
&= 1_k'\Theta N\Theta 1_\ell + 1_k'(\mathrm{diag}(\widehat{\Pi}P\Pi) - I_n)\Theta N\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell + 1_k'\Theta N\Theta(\mathrm{diag}(\widehat{\Pi}P\Pi) - I_n)1_\ell\\
&= \sum_{i\in\mathcal{C}_k, j\in\mathcal{C}_\ell} \theta_i\theta_j N_{ij} + O_p\big(n^2 r_n\bar{\theta}^2\big) \asymp n^2\bar{\theta}^2
\end{aligned}$$

and

$$\big|1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta W\Theta\mathrm{diag}(\widehat{\Pi}P\Pi)1_\ell\big| \le \|W\|\|1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta\|^2 \le n\bar{\theta}^2\sqrt{n\bar{\theta}^2}\,.$$

The last step of the above inequality is due to the non-asymptotic theory of random matrix which gives $\|W\| \le \sqrt{n\bar{\theta}^2}$ with high probability. As a result,

$$J_1 = \sum_{i\in\mathcal{C}_k, j\in\mathcal{C}_\ell} \theta_i\theta_j N_{ij} + O_p\big(n^2 r_n\bar{\theta}^2\big) + O_p\big((n\bar{\theta}^2)^{3/2}\big)\,.$$

To proceed, we note that

$$\begin{aligned}
\|1_k'\Delta\|_1 &= \sum_{i\in\mathcal{C}_k:\hat{\pi}_i = \pi_i} |\hat{\theta}_i - \theta_i| + \sum_{i\in\mathcal{C}_k:\hat{\pi}_i\neq\pi_i} |\hat{\theta}_i - \hat{\pi}_i' P\pi_i\theta_i|\\
&\le Cn\Big(\sqrt{\frac{\log(n)}{n}} + r_n/\bar{\theta}\Big) + Cnr_n\Big(\bar{\theta}\Big(\frac{\log(n)}{n\bar{\theta}^2}\Big)^{1/4} + \sqrt{r_n}\Big)\\
&\le C\sqrt{n\log(n)} + Cnr_n/\bar{\theta}
\end{aligned}$$

For J_2,

$$|J_2| = |1_k'\Delta(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}1_\ell| \le C\|1_k'\Delta\|_1 n\bar{\theta} \le Cn\sqrt{n\bar{\theta}^2\log(n)} + n^2 r_n\,.$$

For J_3,

$$|J_3| = |1_k'\mathrm{diag}(\widehat{\Pi}P\Pi)\Theta(\mathbf{1}_n\mathbf{1}_n' - A)\Delta 1_\ell| \le C\|1_\ell'\Delta\|_1 n\bar{\theta} \le Cn\sqrt{n\bar{\theta}^2\log(n)} + n^2 r_n\,.$$

Consequently,

$$\hat{1}_k'\widehat{\Theta}(\mathbf{1}_n\mathbf{1}_n' - A)\widehat{\Theta}\hat{1}_\ell = \sum_{i\in\mathcal{C}_k, j\in\mathcal{C}_\ell}\theta_i\theta_j N_{ij} + O_p\Big(n\sqrt{n\bar{\theta}^2\log(n)} + n^2 r_n\Big)$$

This, with (24), gives rise to

$$\begin{aligned}
\widehat{P}_{k\ell} &= \frac{\sum_{i\in\mathcal{C}_k, j\in\mathcal{C}_\ell}\theta_i\theta_j N_{ij}P_{k\ell} + O_p(n\bar{\theta}\sqrt{\log(n)} + n^2 r_n\bar{\theta}^2)}{\sum_{i\in\mathcal{C}_k, j\in\mathcal{C}_\ell}\theta_i\theta_j N_{ij} + O_p\Big(n\sqrt{n\bar{\theta}^2\log(n)} + n^2 r_n\Big)} \\
&= P_{k\ell} + O_p\Big(\sqrt{\log(n)}/(n\bar{\theta}) + r_n + \sqrt{\log(n)/n\bar{\theta}^2} + r_n/\bar{\theta}^2\Big) \\
&= P_{k\ell} + O_p\Big(\sqrt{\log(n)/n\bar{\theta}^2} + r_n/\bar{\theta}^2\Big)
\end{aligned}$$

simultaneously for all $1 \le k \neq \ell \le K$, where the probability is at least $1 - o(n^{-3})$. This completes the proof.

C.4 THE ERROR RATE OF N

In this subsection, we prove bounds for $\|N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n'\|_F$ and $\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|$ under the assumptions in Theorem 3.2. The results are provided as below.

Lemma C.3 *Suppose the assumptions in Theorem 3.2 hold. Then,*

$$\begin{aligned}
&\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')\|_F \le C\Big(\sqrt{n\bar{\theta}^2\log(n)} + nr_n + n\bar{\theta}^2\sqrt{r_n}\Big) \ll \lambda_1(\widetilde{\Omega}) \\
&\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| \le C\Big(\bar{\theta}^2\sqrt{n\bar{\theta}^2\log(n)} + n\bar{\theta}^2 r_n + n\bar{\theta}^4\sqrt{r_n}\Big) \ll |\lambda_K(\widetilde{\Omega})|
\end{aligned}$$

with probability $1 - o(n^{-3})$.

We prove Lemma C.3 below.

By definition,

$$N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n' = (\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta) \circ N$$

It follows that

$$\begin{aligned}
& e_{n,i}'(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n')e_{n,j} \\
\le\ & e_{n,i}'(\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)e_{n,j} \\
\le\ & e_{n,i}'(\widehat{\Theta}\widehat{\Pi}(\widehat{P} - P)\widehat{\Pi}'\widehat{\Theta})e_{n,j} + e_{n,i}'(\widehat{\Theta}(\widehat{\Pi} - \Pi)P\widehat{\Pi}'\widehat{\Theta})e_{n,j} + e_{n,i}'(\widehat{\Theta}\Pi P(\widehat{\Pi} - \Pi)'\widehat{\Theta})e_{n,j} \\
& + e_{n,i}'(\widehat{\Theta}\Pi P\Pi'\widehat{\Theta} - \Theta\Pi P\Pi'\Theta)e_{n,j}
\end{aligned}$$

By the error rates for refitting θ and P, i.e.,

$$\|\widehat{P} - P\|_{\max} \le C\Big(\sqrt{\frac{\log(n)}{n\bar{\theta}^2}} + \frac{r_n}{\bar{\theta}^2}\Big), \qquad |\hat{\theta}_i - \theta_i| \le C\Big(\sqrt{\frac{\log(n)}{n}} + \frac{r_n}{\bar{\theta}}\Big), \quad \text{if } \hat{\pi}_i = \pi_i$$

and

$$|\hat{\theta}_i - P_{kk_0}\theta_i| \le C\Big(\bar{\theta}(\log(n)/n\bar{\theta}^2)^{1/4} + \sqrt{r_n}\Big), \quad \text{if } \hat{\pi}_i = e_k, \pi_i = e_{k_0}, k \neq k_0$$

We can derive that with probability $1-o(n^{-3})$, simultaneously for all $1\le i,j\le n$,

$$
\begin{aligned}
\big|e_{n,i}'\big(\widehat{\Theta}\widehat{\Pi}(\widehat{P}-P)\widehat{\Pi}'\widehat{\Theta}\big)e_{n,j}\big| &= \begin{cases} 0 & \text{if } \hat{\pi}_i=\hat{\pi}_j \\ O_p\Big(\bar{\theta}\sqrt{\log(n)/n}+r_n\Big) & \text{if } \hat{\pi}_i\neq\hat{\pi}_j \end{cases} \\
\big|e_{n,i}'\big(\widehat{\Theta}(\widehat{\Pi}-\Pi)P\widehat{\Pi}'\widehat{\Theta}\big)e_{n,j}\big| &= \begin{cases} 0 & \text{if } \hat{\pi}_i=\pi_i \\ O_p\big(\bar{\theta}^2\big) & \text{if } \hat{\pi}_i\neq\pi_i \end{cases} \\
\big|e_{n,i}'\big(\widehat{\Theta}\Pi P(\widehat{\Pi}-\Pi)'\widehat{\Theta}\big)e_{n,j}\big| &= \begin{cases} 0 & \text{if } \hat{\pi}_j=\pi_j \\ O_p\big(\bar{\theta}^2\big) & \text{if } \hat{\pi}_j\neq\pi_j \end{cases}
\end{aligned}
\tag{25}
$$

and

$$
\big|e_{n,i}'\big(\widehat{\Theta}\Pi P\Pi'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\big)e_{n,j}\big| = \begin{cases} O_p\Big(\bar{\theta}\sqrt{\log(n)/n}+r_n\Big) & \text{if } \hat{\pi}_i=\pi_i, \hat{\pi}_j=\hat{\pi}_j \\ O_p\big(\bar{\theta}^2\big) & \text{if } \hat{\pi}_i\neq\pi_i \text{ or } \hat{\pi}_j\neq\pi_j \end{cases}
\tag{26}
$$

Here we used the fact that $\widehat{P}_{kk}=P_{kk}=1$.

Combining the above estimates together, we obtain that

$$
(N\oslash\widehat{N}-\mathbf{1}_n\mathbf{1}_n')_{ij} = \begin{cases} O_p\Big(\bar{\theta}\sqrt{\log(n)/n}+r_n\Big) & \text{if } \hat{\pi}_i=\pi_i, \hat{\pi}_j=\hat{\pi}_j \\ O_p\big(\bar{\theta}^2\big) & \text{if } \hat{\pi}_i\neq\pi_i \text{ or } \hat{\pi}_j\neq\pi_j \end{cases}
$$

Therefore, in light of $\|\widehat{\Pi}-\Pi\|_1\le nr_n$, it yields that

$$
\begin{aligned}
\|N\oslash\widehat{N}-\mathbf{1}_n\mathbf{1}_n'\|_F &= \sqrt{\sum_{i,j:\hat{\pi}_i=\pi_i,\hat{\pi}_j=\pi_j}((N\oslash\widehat{N})_{ij}-1)^2+\sum_{i,j:\hat{\pi}_i\neq\pi_i \text{ or } \hat{\pi}_j\neq\pi_j}((N\oslash\widehat{N})_{ij}-1)^2} \\
&\le C\sqrt{n^2\big(\bar{\theta}\sqrt{\log(n)/n}+r_n\big)^2+n^2r_n\bar{\theta}^4} \\
&\le C\big(\bar{\theta}\sqrt{n\log(n)}+nr_n+n\bar{\theta}^2\sqrt{r_n}\,\big)
\end{aligned}
$$

with probability $1-o(n^{-3})$. Due to the conditions that

$$
r_n\ll\bar{\theta}^2\to 0,\qquad n\bar{\theta}^2\gg\log(n),
$$

we easily see that

$$
\|N\oslash\widehat{N}-\mathbf{1}_n\mathbf{1}_n'\|_F\le C\big(\bar{\theta}\sqrt{n\log(n)}+nr_n+n\bar{\theta}^2\sqrt{r_n}\,\big)\ll\lambda_1(\widetilde{\Omega})
$$

since $\lambda_1(\widetilde{\Omega})\asymp n\bar{\theta}^2$.

Next, we consider $\|(N\oslash\widehat{N}-\mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\|$. Note that

$$
N\oslash\widehat{N}-\mathbf{1}_n\mathbf{1}_n'=\big(\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\big)\circ N\,.
$$

We consider $\|\big(\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\big)\circ\widetilde{\Omega}\circ N\|$ instead. Since the rank of $\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta$ is at most $2K$, we bound

$$
\begin{aligned}
\|(N\oslash\widehat{N}-\mathbf{1}_n\mathbf{1}_n')\circ\widetilde{\Omega}\| &\le \sqrt{\sum_{i,j}(N\circ\widetilde{\Omega})_{ij}^2\big(\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\big)_{ij}^2} \\
&\le \|N\circ\widetilde{\Omega}\|_{\max}\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|_F \\
&\le \sqrt{2K}\|N\circ\widetilde{\Omega}\|_{\max}\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\| \\
&\le C\bar{\theta}^2\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|
\end{aligned}
$$

To proceed, we study the upper bound of $\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|$. Note that

$$
\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|\le\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|_F\le\sqrt{2K}\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|
$$

It suffices to study the upper bound of $\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|_F$, which by previous arguments (25) and (26), is given by

$$
\|\widehat{\Theta}\widehat{\Pi}\widehat{P}\widehat{\Pi}'\widehat{\Theta}-\Theta\Pi P\Pi'\Theta\|_F\le C\big(\bar{\theta}\sqrt{n\log(n)}+nr_n+n\bar{\theta}^2\sqrt{r_n}\,\big)
$$

We thus conclude that

$$\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| \le C\big(\bar{\theta}^2\sqrt{n\bar{\theta}^2\log(n)} + n\bar{\theta}^2 r_n + n\bar{\theta}^4\sqrt{r_n}\,\big)$$

Our assumptions in Theorem 3.2 says that

$$\sqrt{n\bar{\theta}^2}\lambda_{\min}(P) \ge C\log(n), \quad r_n \ll \lambda_{\min}(P), \quad r_n \ll \lambda^2_{\min}(P)/\bar{\theta}^4$$

Using these conditions, together with $|\lambda_K(\widetilde{\Omega})| \asymp n\bar{\theta}^2\lambda_{\min}(P)$, we easily derive that

$$\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\| \le C\big(\bar{\theta}^2\sqrt{n\bar{\theta}^2\log(n)} + n\bar{\theta}^2 r_n + n\bar{\theta}^4\sqrt{r_n}\,\big) \ll |\lambda_K(\widetilde{\Omega})|$$

This finishes the proof of error rates for $\widehat{N}$.

C.5 Proof of Theorem 3.2

We now prove Theorem 3.2 using the results in Sections C.3 and C.4, and Lemma 3.1.

To begin with, we verify the additional conditions in Lemma 3.1 (i.e., (19) and (20)) under the assumptions in Theorem 3.2. Notice that $r_n \asymp \delta_n$ where $\delta_n = \max\{\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|^2, \lambda_1(\widetilde{\Omega})\}/\lambda_K^2(\widetilde{\Omega})$. Specifically, in Section C.3, we have shown that

$$\|\widehat{P} - P\|_{\max} \le C\Big(\sqrt{\frac{\log(n)}{n\bar{\theta}^2}} + \frac{r_n}{\bar{\theta}^2}\Big)$$

From the assumptions in Theorem 3.2, we have

$$\sqrt{n\bar{\theta}^2}\lambda_{\min}(P) \ge C\log(n), \qquad r_n \ll \min\{|\lambda_{\min}(P)|\bar{\theta}, \bar{\theta}^2\}$$

It follows that $\|\widehat{P} - P\|_{\max} = o(1)$ and

$$\sqrt{\frac{\log(n)}{n\bar{\theta}^2}}|\lambda_{\min}(P)|^{-1}\bar{\theta} = o(1), \qquad \frac{r_n}{\bar{\theta}^2}|\lambda_{\min}(P)|^{-1}\bar{\theta} = o(1)\,.$$

Next, thanks to $r_n \asymp \delta_n \ll \lambda^2_{\min}(P)/\bar{\theta}^2$,

$$\|\widehat{\Pi} - \Pi\|(\sqrt{n}\,|\lambda_{\min}(P)|)^{-1}\bar{\theta} \le C\sqrt{nr_n}(\sqrt{n}\,|\lambda_{\min}(P)|)^{-1}\bar{\theta} \ll 1\,.$$

Therefore, the conditions in (19) are satisfied. It was mentioning that in Section C.4, we have validated (20). Lastly, by Lemma C.1, and the conditions that $n\bar{\theta}^2 \gg \log(n), r_n \ll \bar{\theta}^2$, we easily see that $\hat{\theta}_i < C\bar{\theta}$.

Therefore, we can apply the results in Lemma 3.1, which gives that

$$r_n(\widehat{\Pi}^{rscore}) \le \frac{C[\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|^2 + \tau_n^2 + \lambda_1(\widetilde{\Omega})]}{\left|\lambda_K(\widetilde{\Omega})\right|^2}$$

where $\tau_n = \sqrt{n}\bar{\theta}^3[\sqrt{n}\|\widehat{P} - P\|_{\max} + \|\widehat{\Pi}^{score} - \Pi\|]$.

Furthermore, we plug in the upper bounds of $\|\widehat{P} - P\|_{\max}$ and $\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|$ in Sections C.3 and C.4 and note that $\|\widehat{\Pi}^{score} - \Pi\| \le C\sqrt{nr_n} \le C\sqrt{n\delta_n}$. Elementary computations lead to

$$[\|(N \oslash \widehat{N} - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|^2 + \tau_n^2 \le C\Big(n\bar{\theta}^4\log(n) + n^2\bar{\theta}^2\delta_n^2 + n^2\bar{\theta}^6\delta_n\Big)$$

Thereby, we conclude the proof of Theorem 3.2.

C.6 Proof of Corollary 3.1

Using the assumptions that $\lambda_{\min}(P) \geq C$ for a constant $C > 0$, we see that the condition of δ_n in Theorem 3.2 is reduced to

$$\delta_n \ll \bar{\theta}^2$$

Also, we can derive

$$\|(N - \mathbf{1}_n\mathbf{1}_n') \circ \widetilde{\Omega}\|^2/\lambda_K^2(\widetilde{\Omega}) \leq \|(N - \mathbf{1}_n\mathbf{1}_n')\|_{\max}^2\|\widetilde{\Omega}\|^2/\lambda_K^2(\widetilde{\Omega}) \leq C\bar{\theta}^4 \ll \bar{\theta}^2 \to 0$$

and

$$\lambda_1(\widetilde{\Omega})/\lambda_K^2(\widetilde{\Omega}) \leq \frac{1}{n\bar{\theta}^2} \ll \bar{\theta}^2 .$$

by the assumption that $n\bar{\theta}^4 \to \infty$. Therefore, by Theorem 3.1, we obtain that

$$\delta_n \asymp r_n(\widehat{\Pi}^{score}) \leq C\Big(\frac{1}{n\bar{\theta}^2} + \bar{\theta}^4\Big)$$

and $\delta_n \ll \bar{\theta}^2$. Then, the conditions in Theorem 3.2 are satisfied. Therefore,

$$\begin{aligned} r_n(\widehat{\Pi}^{rscore}) &\leq \frac{C}{\lambda_K^2(\widetilde{\Omega})}\Big(\lambda_1(\widetilde{\Omega}) + n\bar{\theta}^4\log(n) + n^2\bar{\theta}^2\delta_n^2 + n^2\bar{\theta}^6\delta_n\Big) \\ &\leq \frac{C}{n^2\bar{\theta}^4}\Big(n\bar{\theta}^2 + n\bar{\theta}^4\log(n) + n^2\bar{\theta}^2[1/(n\bar{\theta}^2)^2 + \bar{\theta}^8] + n^2\bar{\theta}^6[1/n\bar{\theta}^2 + \bar{\theta}^4]\Big) \\ &\leq C\Big(\frac{1}{n\bar{\theta}^2} + \bar{\theta}^6 + \frac{\log(n)}{n}\Big) \end{aligned}$$

where we used $\bar{\theta} = o(1)$ and $n\bar{\theta}^4 \to \infty$. We thus complete the proof of Corollary 3.1.

References

Afonso S Bandeira and Ramon Van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.*, 44(4):2479–2506, 2016.

Chandler Davis and William Morton Kahan. The rotation of eigenvectors by a perturbation. iii. *SIAM J. Numer. Anal.*, 7(1):1–46, 1970.

Victor H de la Pena and Stephen J Montgomery-Smith. Decoupling inequalities for the tail probabilities of multivariate u-statistics. *The Annals of Probability*, pp. 806–816, 1995.

Roger Horn and Charles Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

Jiashun Jin. Fast community detection by score. *Ann. Statist.*, 43(1):57–89, 2015.

Yi Yu, Tengyao Wang, and Richard J Samworth. A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.