Supplement to "Optimal Full Ranking from Pairwise Comparisons"

Pinhan Chen[1], Chao Gao[1], and Anderson Y. Zhang[2]

[1] *University of Chicago*
[2] *University of Pennsylvania*

January 18, 2022

This supplement includes all the technical proofs. In Appendix A, we first give the proof of Theorem 3.1. In Appendix B, we give the proof of Theorem 4.1. After that, we prove Lemma 4.1, Lemma 4.2, and Lemma 4.3 in Appendix C. We then present the proofs of Lemma 8.1 and Lemma 8.2 in Appendix D.

A Proof of Theorem 3.1

We prove Theorem 3.1 in this section. We first state and prove a few lemmas.

Lemma A.1. *Assume $p \geq \frac{c_0 \log n}{n}$ for some sufficiently large $c_0 > 0$. Then, we have*

$$\|A - \mathbb{E}(A)\|_{\mathrm{op}} \leq C\sqrt{np}, \tag{S1}$$

$$\|D - \mathbb{E}(D)\|_{\mathrm{op}} \leq C\sqrt{np \log n} \tag{S2}$$

for some constant $C > 0$ with probability at least $1 - O(n^{-10})$.

Proof. Bound (S1) is a direct consequence of Theorem 5.2 in [3] and Bound (S2) is from standard concentration of sums of i.i.d. Bernoulli random variables. $\square$

Lemma A.2. *Assume $p \geq \frac{c_0 \log n}{n}$ for some sufficiently large $c_0 > 0$. Then, we have*

$$np - 2C\sqrt{np \log n} \leq \lambda_{\min,\perp}(\mathcal{L}_A) = \min_{u \neq 0, \mathbb{1}_n^T u = 0} \frac{u^T \mathcal{L}_A u}{\|u\|},$$

$$np + 2C\sqrt{np \log n} \geq \lambda_{\max,\perp}(\mathcal{L}_A) = \max_{u \neq 0, \mathbb{1}_n^T u = 0} \frac{u^T \mathcal{L}_A u}{\|u\|}$$

for some constant $C > 0$ with probability at least $1 - O(n^{-10})$.

Proof. Note the decomposition

$$\mathcal{L}_A = \mathbb{E}\mathcal{L}_A + D - \mathbb{E}D - (A - \mathbb{E}A)$$

and $\lambda_{\min,\perp}(\mathbb{E}\mathcal{L}_A) = \lambda_{\max,\perp}(\mathbb{E}\mathcal{L}_A) = np$. By Lemma A.1, we have

$$\|D - \mathbb{E}D - (A - \mathbb{E}A)\|_{\mathrm{op}} \leq 2C\sqrt{np\log n}$$

with probability at least $1 - O(n^{-10})$ for some $C > 0$. The Lemma can be seen immediately by Weyl's inequality. $\square$

Recall the notation of $r^{*(i,j)} \in \mathfrak{S}_n$ defined in (36), which is a permutation by swapping the i, jth position in $r^* \in \mathfrak{S}_n$ while keeping other positions fixed.

Lemma A.3. *Assume $\frac{np}{\log n} \to \infty$. There exists $\delta = o(1)$, such that for any $\theta^* \in \Theta_n(\beta, C_0)$, any $r^* \in \mathfrak{S}_n$, any $i, j \in [n], i \neq j$, we have*

$$\begin{aligned}
&\inf_{\widehat{r}} \frac{\mathbb{P}_{(\theta^*,\sigma^2,r^*)}\left(\widehat{r} \neq r^*\right) + \mathbb{P}_{(\theta^*,\sigma^2,r^{*(i,j)})}\left(\widehat{r} \neq r^{*(i,j)}\right)}{2} \\
&\gtrsim \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta^*_{r^*_i} - \theta^*_{r^*_j})^2}} \exp\left(-\frac{(1+\delta)np(\theta^*_{r^*_i} - \theta^*_{r^*_j})^2}{4\sigma^2}\right)\right\}
\end{aligned}$$

Proof. Assume $r^*_i = a < r^*_j = b$ and thus $\theta^*_a \geq \theta^*_b$. Let $\mathcal{F}$ be the event about A on which Lemma A.1 holds. We have $\mathbb{P}(\mathcal{F}) > 1/2$. To simplify notation, let $\mathbb{P}_A(\cdot) = \mathbb{P}_{(\theta^*,\sigma^2,r^*)}(\cdot|A)$ be the conditional probability. For any A, by Neyman-Pearson Lemma, the optimal procedure is given by the likelihood ratio test. Then

$$\begin{aligned}
&\inf_{\widehat{r}} \frac{\mathbb{P}_{(\theta^*,\sigma^2,r^*)}\left(\widehat{r} \neq r^*\right) + \mathbb{P}_{(\theta^*,\sigma^2,r^{*(i,j)})}\left(\widehat{r} \neq r^{*(i,j)}\right)}{2} \\
&\geq \mathbb{P}(\mathcal{F}) \inf_{A \in \mathcal{F}} \mathbb{P}_A\left(\frac{d\mathbb{P}_{(\theta^*,\sigma^2,r^{*(i,j)})}}{d\mathbb{P}_{(\theta^*,\sigma^2,r^*)}} \geq 1\right) \\
&\gtrsim \inf_{A \in \mathcal{F}} \mathbb{P}_A\left(-4A_{ij}(\theta^*_a - \theta^*_b + w_{ij}) + \sum_{k \neq i,j} -A_{ik}(\theta^*_a - \theta^*_b + 2w_{ik}) + \sum_{k \neq i,j} A_{jk}(\theta^*_b - \theta^*_a + 2w_{jk}) \geq 0\right) \\
&= \inf_{A \in \mathcal{F}} \mathbb{P}_A\left(\mathcal{N}(0, \frac{\sigma^2}{D_{ii} + D_{jj} + 2A_{ij}}) \geq \frac{|\theta_a - \theta_b|}{2}\right) \\
&\gtrsim \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta^*_a - \theta^*_b)^2}} \exp\left(-\frac{(1+\delta)np(\theta^*_a - \theta^*_b)^2}{4\sigma^2}\right)\right\} \qquad \text{(S3)}
\end{aligned}$$

for some $\delta = o(1)$, where (S3) comes from standard Gaussian tail bound and Lemma A.1. $\square$

Now we are ready to state the proof of Theorem 3.1.

Proof of Theorem 3.1. We prove the theorem for any $\theta^* \in \Theta_n(\beta, C_0)$. Note that conditional on A, the solution of the least squares problem (11) can be written as

$$\widehat{\theta} = c\mathbb{1}_n + \theta^*_{r^*} + Z,$$

where $\theta^*_{r^*} = (\theta^*_{r_1^*}, ..., \theta^*_{r_n^*})^T$, $Z \sim \mathcal{N}(0, \sigma^2 \mathcal{L}_A^\dagger)$ and $c\mathbb{1}_n$ is a global shift of the skill parameters. Let $x_{ij} = e_i - e_j$ where $\{e_1, ..., e_n\}$ are the standard basis of $\mathbb{R}^n$. Let $\mathcal{F}$ be the event about A when Lemma A.2 holds. Then

$$\begin{aligned}
&\mathbb{E}_{(\theta^*,\sigma^2,r^*)}\left[\mathsf{K}(\widehat{r}, r^*)\right] = \frac{1}{n}\sum_{1\le i<j\le n} \mathbb{P}_{(\theta^*,\sigma^2,r^*)}\left(\text{sign}(\widehat{r}_i - \widehat{r}_j)\text{sign}(r_i^* - r_j^*) < 0\right) \\
&= \frac{1}{n}\sum_{1\le i<j\le n} \mathbb{P}_{(\theta^*,\sigma^2,r^*)}\left(\text{sign}(\widehat{\theta}_i - \widehat{\theta}_j)\text{sign}(r_i^* - r_j^*) > 0\right) \\
&\le \frac{1}{n}\sum_{1\le i<j\le n} \sup_{A\in\mathcal{F}} \mathbb{P}\left(\mathcal{N}(0, \sigma^2 x_{ij}^T \mathcal{L}_A^\dagger x_{ij}) > \left|\theta^*_{r_i^*} - \theta^*_{r_j^*}\right| |A\right) + O(n^{-9}) \\
&\le \frac{1}{n}\sum_{1\le i<j\le n} \sup_{A\in\mathcal{F}} \min\left\{1, \sqrt{\frac{\sigma^2 x_{ij}^T \mathcal{L}_A^\dagger x_{ij}}{2\pi(\theta^*_{r_i^*} - \theta^*_{r_j^*})^2}} \exp\left(-\frac{(\theta^*_{r_i^*} - \theta^*_{r_j^*})^2}{2\sigma^2 x_{ij}^T \mathcal{L}_A^\dagger x_{ij}}\right)\right\} + O(n^{-9}) \\
&\le \frac{1}{n}\sum_{1\le i<j\le n} \min\left\{1, \sqrt{\frac{\sigma^2(np - 2C\sqrt{np\log n})^{-1}}{\pi(\theta^*_{r_i^*} - \theta^*_{r_j^*})^2}} \exp\left(-\frac{(np - 2C\sqrt{np\log n})(\theta^*_{r_i^*} - \theta^*_{r_j^*})^2}{4\sigma^2}\right)\right\} \qquad \text{(S4)} \\
&\quad + O(n^{-9}) \\
&\lesssim \frac{1}{n}\sum_{1\le i<j\le n} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i^* - \theta_j^*)^2}} \exp\left(-\frac{(1-\delta_1')np(\theta_i^* - \theta_j^*)^2}{4\sigma^2}\right)\right\} + n^{-9}
\end{aligned}$$

for some $\delta_1' = o(1)$ independent of θ^*, σ^2 and r^*, where (S4) is due to Lemma A.2.

We first consider the high signal-to-noise ratio regime, where $\frac{np\beta^2}{\sigma^2} > 1$. In this scenario,

$$\begin{aligned}
&\sum_{1\le i<j\le n} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i^* - \theta_j^*)^2}} \exp\left(-\frac{(1-\delta_1')np(\theta_i^* - \theta_j^*)^2}{4\sigma^2}\right)\right\} \\
&\le \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \exp\left(-\frac{(1-\delta_1')np(\theta_i^* - \theta_j^*)^2}{4\sigma^2}\right) \\
&\le \sum_{i=1}^{n-1} \exp\left(-\frac{(1-\delta_1')np(\theta_i^* - \theta_{i+1}^*)^2}{4\sigma^2}\right)\sum_{j=i+1}^{n} \exp\left(-\frac{(1-\delta_1')np[(\theta_i^* - \theta_j^*)^2 - (\theta_i^* - \theta_{i+1}^*)^2]}{4\sigma^2}\right) \\
&\le \sum_{i=1}^{n-1} \exp\left(-\frac{(1-\delta_1')np(\theta_i^* - \theta_{i+1}^*)^2}{4\sigma^2}\right)\sum_{j=i+1}^{n} \exp\left(-\frac{(1-\delta_1')np(j-i-1)\beta^2}{4\sigma^2}\right) \\
&\lesssim \sum_{i=1}^{n-1} \exp\left(-\frac{(1-\delta_1')np(\theta_i^* - \theta_{i+1}^*)^2}{4\sigma^2}\right)
\end{aligned}$$

where the last inequality is due to summation of an exponentially decaying series. This gives the exponential rate in high signal-to-noise ratio regime.

Now, when $\frac{np\beta^2}{\sigma^2} \leq 1$,

$$
\begin{aligned}
&\sum_{1\leq i<j\leq n} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i^*-\theta_j^*)^2}} \exp\left(-\frac{(1-\delta_1')np(\theta_i^*-\theta_j^*)^2}{4\sigma^2}\right)\right\} \\
&\leq \sum_{i=1}^{n-1}\sum_{k\geq 1} \sum_{\substack{j>i \\ (k-1)\sqrt{\frac{\sigma^2}{np\beta^2}}<j-i\leq k\sqrt{\frac{\sigma^2}{np\beta^2}}}} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i^*-\theta_j^*)^2}} \exp\left(-\frac{(1-\delta')np(\theta_i^*-\theta_j^*)^2}{4\sigma^2}\right)\right\} \\
&\lesssim \sqrt{\frac{\sigma^2}{np\beta^2}} \sum_{i=1}^{n-1}\left(\sum_{k\geq 0}\exp\left(-\frac{(1-\delta')k^2}{4}\right)\right) \lesssim n\sqrt{\frac{\sigma^2}{np\beta^2}} \wedge n^2
\end{aligned}
$$

where the last inequality also comes from summing an exponentially decaying series and n^2 is a trivial upper bound. This finishes the proof of the upper bound.

Now we look at the lower bound. For any $r^* \in \mathfrak{S}_n$, we have $r^{*(i,j)} \in \mathfrak{S}_n$ defined as in (36). Then for any $\theta^* \in \Theta_n(\beta, C_0)$,

$$
\begin{aligned}
&\inf_{\widehat{r}} \sup_{r^*\in\mathfrak{S}_n} \mathbb{E}_{(\theta^*,\sigma^2,r^*)}\left[\mathsf{K}(\widehat{r}, r^*)\right] \\
&\geq \inf_{\widehat{r}} \frac{1}{n}\sum_{1\leq i<j\leq n}\frac{1}{n!}\sum_{r^*\in\mathfrak{S}_n}\mathbb{P}_{(\theta^*,\sigma^2,r^*)}\left(\text{sign}(\widehat{r}_i-\widehat{r}_j)\text{sign}(r_i^*-r_j^*)<0\right) \\
&= \inf_{\widehat{r}} \frac{1}{n}\sum_{1\leq i<j\leq n}\frac{1}{n!}\sum_{1\leq a<b\leq n}\sum_{r^*:\{r_i^*,r_j^*\}=\{a,b\}}\mathbb{P}_{(\theta^*,\sigma^2,r^*)}\left(\text{sign}(\widehat{r}_i-\widehat{r}_j)\text{sign}(r_i^*-r_j^*)<0\right) \\
&\geq \frac{1}{n}\sum_{1\leq i<j\leq n}\frac{2}{n(n-1)}\sum_{1\leq a<b\leq n}\frac{1}{(n-2)!}\sum_{r^*:r_i^*=a,r_j^*=b}\inf_{\widehat{r}}\frac{\mathbb{P}_{(\theta^*,\sigma^2,r^*)}(\widehat{r}_i\neq a)+\mathbb{P}_{(\theta^*,\sigma^2,r^{*(i,j)})}(\widehat{r}_i\neq b)}{2} \\
&\gtrsim \frac{1}{n}\sum_{1\leq i<j\leq n}\frac{2}{n(n-1)}\sum_{1\leq a<b\leq n} \\
&\qquad\qquad \frac{1}{(n-2)!}\sum_{r^*:r_i^*=a,r_j^*=b}\min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_a^*-\theta_b^*)^2}}\exp\left(-\frac{(1+\delta')np(\theta_a^*-\theta_b^*)^2}{4\sigma^2}\right)\right\} \qquad \text{(S5)} \\
&= \frac{1}{n}\sum_{1\leq a<b\leq n}\min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_a^*-\theta_b^*)^2}}\exp\left(-\frac{(1+\delta')np(\theta_a^*-\theta_b^*)^2}{4\sigma^2}\right)\right\}
\end{aligned}
$$

for some $\delta' = o(1)$, where (S5) comes from Lemma A.3.

We still consider the high signal-to-noise ratio case first.

$$
\begin{aligned}
&\sum_{1\le i<j\le n} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i^*-\theta_j^*)^2}} \exp\left(-\frac{(1+\delta')np(\theta_i^*-\theta_j^*)^2}{4\sigma^2}\right)\right\} \\
&\ge \sum_{i=1}^{n-1} \sqrt{\frac{\sigma^2}{np(\theta_i^*-\theta_{i+1}^*)^2}} \exp\left(-\frac{(1+\delta')np(\theta_i^*-\theta_{i+1}^*)^2}{4\sigma^2}\right) \\
&\gtrsim \sum_{i=1}^{n-1} \exp\left(-\frac{(1+\delta)np(\theta_i^*-\theta_{i+1}^*)^2}{4\sigma^2}\right) \qquad \text{(S6)}
\end{aligned}
$$

where δ in (S6) can be chosen arbitrarily small when $np\beta^2/\sigma^2 > 1$, which concludes the exponential lower bound.

For the polynomial lower bound when signal-to-noise ratio is small,

$$
\begin{aligned}
&\sum_{1\le i<j\le n} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i^*-\theta_j^*)^2}} \exp\left(-\frac{(1+\delta')np(\theta_i^*-\theta_j^*)^2}{4\sigma^2}\right)\right\} \\
&\gtrsim \sum_{i=1}^{n} \sum_{\substack{j\ne i \\ |j-i|\le\sqrt{\frac{\sigma^2}{np\beta^2}}}} \min\left\{1, \sqrt{\frac{\sigma^2}{np(\theta_i-\theta_j)^2}} \exp\left(-\frac{(1+\delta')np(\theta_i-\theta_j)^2}{4\sigma^2}\right)\right\} \\
&\gtrsim \sum_{i=1}^{n} n \wedge \left(\sqrt{\frac{\sigma^2}{np\beta^2}}\right)
\end{aligned}
$$

which concludes the proof. □

B Proof of Theorem 4.1

The following two lemmas are needed for the proof of Theorem 4.1. Recall that $\bar{y}_{ij}^{(1)} = \frac{1}{L_1}\sum_{l=1}^{L_1} y_{ijl}, i \ne j \in [n]$.

Lemma B.1. *There exists a constant $C_1 > 0$ such that for any $\theta^* \in \Theta_n(\beta, C_0)$ and $r^* \in \mathfrak{S}_n$,*

$$
\max_{i\in[n],j\in[n],i\ne j} \left|\bar{y}_{ij}^{(1)} - \psi(\theta_{r_i^*}^* - \theta_{r_j^*}^*)\right| \le C_1\sqrt{\frac{\log n}{L_1}}
$$

holds with probability at least $1 - O(n^{-10})$.

Proof. This can be seen directly by standard Hoeffding's inequality and union bound argument. □

Lemma B.2. *For L_1 such that $\frac{L_1}{\log n} \to \infty$ and constant $M \ge 1$, there exists $0 < \delta_0 = o(1)$ and $0 < \delta_1 = o(1)$ such that for any $\theta^* \in \Theta_n(\beta, C_0)$, any $r^* \in \mathfrak{S}_n$,*

$$
\max_{i\in[n],j\in[n],i\ne j} \left|\bar{y}_{ij}^{(1)} - \psi(\theta_{r_i^*}^* - \theta_{r_j^*}^*)\right| \le \delta_0
$$

and

$$\bigcap_{i=1}^{n}\left\{\underline{\mathcal{E}_{1,i}}\subset\mathcal{E}_{1,i}\subset\overline{\mathcal{E}_{1,i}}\right\}$$

hold with probability at least $1-O(n^{-10})$*, where*

$$\begin{aligned}
\mathcal{E}_{1,i} &= \left\{j\in[n]:\bar{y}_{ij}^{(1)}\leq\psi(-2M)\right\},\\
\underline{\mathcal{E}_{1,i}} &= \left\{j\in[n]:\theta^*_{r_j^*}\geq\theta^*_{r_i^*}+2M+\delta_1\right\},\\
\overline{\mathcal{E}_{1,i}} &= \left\{j\in[n]:\theta^*_{r_j^*}\geq\theta^*_{r_i^*}+2M-\delta_1\right\}.
\end{aligned}$$

Proof. This is a direct consequence of Lemma B.1 and $M=O(1)$. □

Now we are ready to prove Theorem 4.1.

Proof of Theorem 4.1. Let $\mathcal{F}^{(0)}$ be the event on which Lemma B.2 holds. We will always work on this high probability event throughout the proof. Also, we will assume the regime $n\beta\to\infty$. The case $\beta\lesssim 1/n$ is trivial since we only have one league $S_1=[n]$ if M is chosen to be a large enough constant.

To start the exposition, we define a series of quantities iteratively for all $k\in[K-1]$, with the base case $\underline{S_0}=\overline{S_0}=S_0'=\widetilde{S}_0=\emptyset$, $\underline{u^{(0)}}=\overline{u^{(0)}}=0$. Let

$$\underline{t_i^{(k)}}=\left|\left\{j\in[n]\backslash\widetilde{S}_{k-1}:j\in\underline{\mathcal{E}_{1,i}}\right\}\right|,$$

$$\overline{t_i^{(k)}}=\left|\left\{j\in[n]\backslash\widetilde{S}_{k-1}:j\in\overline{\mathcal{E}_{1,i}}\right\}\right|,$$

$$\underline{S_k}=\left\{i\in[n]\backslash\widetilde{S}_{k-1}:\left(1+\frac{0.11}{C_0^2}\right)p\overline{t_i^{(k)}}\leq h\right\},\tag{S7}$$

$$\overline{S_k}=\left\{i\in[n]\backslash\widetilde{S}_{k-1}:\left(1-\frac{0.11}{C_0^2}\right)p\underline{t_i^{(k)}}\leq h\right\},\tag{S8}$$

$$\overline{u^{(k)}}=\max\left\{r_i^*:i\in[n]\backslash\widetilde{S}_{k-1},\overline{t_i^{(k)}}\leq\frac{M}{\left(1-\frac{0.12}{C_0^2}\right)\beta}\right\},$$

$$\underline{u^{(k)}}=\max\left\{r_i^*:i\in[n]\backslash\widetilde{S}_{k-1},\overline{t_i^{(k)}}\leq\frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta}\right\},$$

$$w_i^{(k)\prime}=\sum_{j\in\underline{S_k}\cap\overline{\mathcal{E}_{1,i}}}A_{ij}\mathbb{I}\left\{j\in\mathcal{E}_{1,i}\right\},$$

$$S_k'=\left\{i\in[n]\backslash\widetilde{S}_{k-1}:w_i^{(k)\prime}\leq h\right\},$$

$$\widetilde{S}_k=\uplus_{m=1}^{k}S_m'.$$

We make several remarks about these definition. The above definitions have essentially constructed another partition $S_1', S_2', \ldots$ using $w_i^{(k)\prime}$ comparing to Algorithm 1 using $w_i^{(k)}$. The relationship between S_k and S_k' will be made clear during the exposition. In fact, they will be equal with high probability. We should keep in mind that the partition using $w_i^{(k)\prime}$ is not a bona fide one since the definition uses $\overline{\mathcal{E}_{1,i}}$ and $\underline{S_k}$ which involve the knowledge of θ^*. However, this can be used in theoretical exploration. Our strategy is to show certain properties hold for partitions S_k', then $S_k = S_k'$ with high probability and thus inherits those properties.

We start with some simple but crucial facts which will act as building blocks in the proof.

- $\overline{t_i^{(k)}}$ and $\underline{t_i^{(k)}}$ has the following monotonicity property: for any $i, j \in [n]\backslash \widetilde{S}_{k-1}$ such that $r_i^* \leq r_j^*$,
$$\overline{t_i^{(k)}} \leq \overline{t_j^{(k)}}, \underline{t_i^{(k)}} \leq \underline{t_j^{(k)}}. \tag{S9}$$
This is direct from the definition.

- For any $i \in [n]\backslash \widetilde{S}_{k-1}$,
$$0 \leq \overline{t_i^{(k)}} - \underline{t_i^{(k)}} \leq \frac{2\delta_1}{\beta} + 1, \tag{S10}$$
which comes from $\overline{t_i^{(k)}} - \underline{t_i^{(k)}} = \left|\left\{j \in [n]\backslash \widetilde{S}_{k-1} : 2M - \delta_1 \leq \theta^*_{r_j^*} - \theta^*_{r_i^*} < 2M + \delta_1\right\}\right|$.

- We have
$$\left\{i \in [n]\backslash \widetilde{S}_{k-1} : r_i^* \leq \underline{u^{(k)}}\right\} = \underline{S_k} \subset \overline{S_k} \subset \left\{i \in [n]\backslash \widetilde{S}_{k-1} : r_i^* \leq \overline{u^{(k)}}\right\}. \tag{S11}$$
Here $\underline{S_k} \subset \overline{S_k}$ is due to monotonicity (S9) and $\underline{t_i^{(k)}} \leq \overline{t_i^{(k)}}$ by definition. Recall $h = pM/\beta$. Using (S10), for any $i \in \overline{S_k}$, we have $\overline{t_i^{(k)}} \leq \underline{t_i^{(k)}} + \frac{2\delta_1}{\beta} + 1 \leq \frac{h}{\left(1-\frac{0.11}{C_0^2}\right)p} + \frac{2\delta_1}{\beta} + 1 \leq \frac{M}{\left(1-\frac{0.12}{C_0^2}\right)\beta}$. Hence, we have $\overline{S_k} \subset \left\{i \in [n]\backslash \widetilde{S}_{k-1} : r_i^* \leq \overline{u^{(k)}}\right\}$.

- $\underline{t_i^{(k)}}$, $\overline{t_i^{(k)}}$, $\underline{S_k}$, $\overline{S_k}$, $\underline{u^{(k)}}$, $\overline{u^{(k)}}$ are measurable with respect to the σ-algebra generated by $\widetilde{S}_{k-1}$. This is direct from the definition.

Now, we will prove the following statements by induction on k:

- With probability at least $1 - O(kn^{-10})$,
$$\underline{S_{k'}} \subset S_{k'} \subset \overline{S_{k'}} \tag{S12}$$
for all $0 \leq k' \leq k$.

- With probability at least $1 - O(kn^{-10})$,
$$\left|\underline{S_{k'}}\right| \geq \left(\frac{1.7}{C_0} + \frac{1}{1 + \frac{0.11}{C_0^2}}\right)\frac{M}{\beta} \tag{S13}$$
for all $1 \leq k' \leq k$ and $|S_0| = 0$.

- With probability at least $1 - O(kn^{-10})$,
$$\left|\overline{S_{k'}} \backslash \underline{S_{k'}}\right| \leq \overline{u^{(k')}} - \underline{u^{(k')}} \leq \frac{0.29M}{C_0\beta} \tag{S14}$$
for all $0 \leq k' \leq k$.
- With probability at least $1 - O(kn^{-10})$,
$$\left|\overline{S_{k'}}\right| \leq \left(2 + \frac{0.29}{C_0} + \frac{1}{1 - \frac{0.12}{C_0^2}}\right)\frac{M}{\beta} \tag{S15}$$
for all $0 \leq k' \leq k$.
- With probability at least $1 - O(kn^{-10})$,
$$S_{k'} = S'_{k'} \tag{S16}$$
for all $0 \leq k' \leq k$.

Now, suppose (S12) - (S16) hold until $k-1$, which is the case for $k=1$. In the following, we are going to establish (S12) - (S16) for k one by one.

(Establishment of (S12)). Recall that we assume $\mathcal{F}^{(0)}$ holds. On the intersection of all high probability events before k, we have $\widetilde{S}_{k-1} = S_1 \cup \ldots \cup S_{k-1}$. We sandwich $w_i^{(k)}$ by
$$\underline{w_i^{(k)}} = \sum_{j \in [n] \backslash \widetilde{S}_{k-1}} A_{ij} \mathbb{I}\left\{j \in \underline{\mathcal{E}_{1,i}}\right\} \leq w_i^{(k)} \leq \sum_{j \in [n] \backslash \widetilde{S}_{k-1}} A_{ij} \mathbb{I}\left\{j \in \overline{\mathcal{E}_{1,i}}\right\} = \overline{w_i^{(k)}}.$$
Recall the definition of S_k in Algorithm 1. Then we have $S_k = \left\{i \in [n]\backslash \widetilde{S}_{k-1} : w_i^{(k)} \leq h\right\}$. Hence, $\left\{i \in [n]\backslash \widetilde{S}_{k-1} : \overline{w_i^{(k)}} \leq h\right\} \subset S_k \subset \left\{i \in [n]\backslash \widetilde{S}_{k-1} : \underline{w_i^{(k)}} \leq h\right\}$. To prove (S12), by the definitions in (S7) and (S8), we only need to show
$$\left\{i \in [n]\backslash \widetilde{S}_{k-1} : \underline{w_i^{(k)}} \leq h\right\} \subset \left\{i \in [n]\backslash \widetilde{S}_{k-1} : \left(1 - \frac{0.11}{C_0^2}\right) p\underline{t_i^{(k)}} \leq h\right\},$$
$$\left\{i \in [n]\backslash \widetilde{S}_{k-1} : \left(1 + \frac{0.11}{C_0^2}\right) p\overline{t_i^{(k)}} \leq h\right\} \subset \left\{i \in [n]\backslash \widetilde{S}_{k-1} : \overline{w_i^{(k)}} \leq h\right\},$$
a sufficient condition of which is the following event:
$$\begin{aligned}\mathcal{F}^{(k)} = &\left\{\forall i \in [n]\backslash \widetilde{S}_{k-1} \text{ such that } p\underline{t_i^{(k)}} \leq \frac{h}{2} : \underline{w_i^{(k)}} \leq h\right\} \\ &\bigcap \left\{\forall i \in [n]\backslash \widetilde{S}_{k-1} \text{ such that } p\underline{t_i^{(k)}} > \frac{h}{2} : \left(1 - \frac{0.11}{C_0^2}\right) p\underline{t_i^{(k)}} \leq \underline{w_i^{(k)}}\right\} \\ &\bigcap \left\{\forall i \in [n]\backslash \widetilde{S}_{k-1} \text{ such that } p\overline{t_i^{(k)}} \leq \frac{h}{2} : \overline{w_i^{(k)}} \leq h\right\} \\ &\bigcap \left\{\forall i \in [n]\backslash \widetilde{S}_{k-1} \text{ such that } p\overline{t_i^{(k)}} > \frac{h}{2} : \overline{w_i^{(k)}} \leq \left(1 + \frac{0.11}{C_0^2}\right) p\overline{t_i^{(k)}}\right\}.\end{aligned}$$

Hence to prove (S12), we only need to analyze $\mathbb{P}\left(\mathcal{F}^{(k)}\right)$.

Note that for any $j \in [n]\backslash\widetilde{S}_{k-1}$ we have $r_j^* > \underline{u^{(k-1)}}$ according to the definition of $\underline{S_{k-1}}$ in (S11). Thus

$$\underline{w_i^{(k)}} = \sum_{\substack{j\in[n]\backslash\widetilde{S}_{k-1}\\ r_j^*>\underline{u^{(k-1)}}}} A_{ij}\mathbb{I}\left\{\theta_j^* \geq \theta_i^* + 2M + \delta_1\right\},$$

$$\overline{w_i^{(k)}} = \sum_{\substack{j\in[n]\backslash\widetilde{S}_{k-1}\\ r_j^*>\underline{u^{(k-1)}}}} A_{ij}\mathbb{I}\left\{\theta_j^* \geq \theta_i^* + 2M - \delta_1\right\}.$$

On the other hand, recall that $w_i^{(k-1)\prime} = \sum_{j\in\underline{S_{k-1}}\cap\overline{\mathcal{E}_{1,i}}} A_{ij}\mathbb{I}\left\{j \in \mathcal{E}_{1,i}\right\}$ which only involves A_{ij} such that $r_j^* \leq \underline{u^{(k-1)}}$ due to (S11). By (S11) and induction hypothesis of (S12) we further know $\underline{u^{(1)}} \leq ... \leq \underline{u^{(k-1)}}$. As a result, $w_i^{(1)\prime}, ..., w_i^{(k-1)\prime}$ are independent of $\underline{w_i^{(k)}}, \overline{w_i^{(k)}}$. Since $\widetilde{S}_{k-1}$ is determined by $w_i^{(1)\prime}, ..., w_i^{(k-1)\prime}$, it is also independent of $\underline{w_i^{(k)}}, \overline{w_i^{(k)}}$.

Therefore, conditional on $\widetilde{S}_{k-1}$, we have

$$\underline{w_i^{(k)}}|\widetilde{S}_{k-1} \sim \text{Binomial}(\underline{t_i^{(k)}}, p),$$

$$\overline{w_i^{(k)}}|\widetilde{S}_{k-1} \sim \text{Binomial}(\overline{t_i^{(k)}}, p).$$

Recall that $C_0 \geq 1$ is a constant and $h = pM/\beta \gg \log n$ since $p/(\beta \log n) \to \infty$ by assumption. By Bernstein inequality for the Binomial distributions together with a union bound argument, we have $\mathbb{P}\left(\mathcal{F}^{(k)}|\widetilde{S}_{k-1}\right) \geq 1 - O(n^{-10})$. Since this holds for all $\widetilde{S}_{k-1}$, we have

$$\mathbb{P}\left(\mathcal{F}^{(k)}\right) \geq 1 - O(n^{-10}).$$

Therefore, we have proved (S12).

(Establishment of (S13)). We first present a simple fact from induction hypothesis:

$$\left\{i \in [n], r_i^* \leq \underline{u^{(k-1)}}\right\} \subset \widetilde{S}_{k-1} \subset \left\{i \in [n], r_i^* \leq \overline{u^{(k-1)}}\right\}. \tag{S17}$$

The first containment is because (S11) and (S12) hold up to $k-1$. To prove the second containment, we only need to show $\overline{u^{(1)}} \leq \ldots \leq \overline{u^{(k-1)}}$. Notice that from (S13) and (S14) for $k-1$, we have $\left|\underline{S_{k-1}}\right| \geq \overline{u^{(k-2)}} - \underline{u^{(k-2)}}$. On the other hand, from (S12) for $k-1$, we have $\left|\underline{S_{k-1}}\right| \leq \left|\left\{i \in [n] : r_i^* > \underline{u^{(k-2)}}, r_i^* \leq \overline{u^{(k-1)}}\right\}\right| \leq \overline{u^{(k-1)}} - \underline{u^{(k-2)}}$. Hence, we have $\overline{u^{(k-1)}} \geq \overline{u^{(k-2)}}$ and similarly we can show $\overline{u^{(l+1)}} \geq \overline{u^{(l)}}$ for any $l \leq k-2$, which proves $\overline{u^{(1)}} \leq \ldots \leq \overline{u^{(k-1)}}$.

Using (S17), we have

$$\left|\underline{S_k}\right| = \left|\left\{i \in [n]\backslash\widetilde{S}_{k-1} : \overline{t_i^{(k)}} \le \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta}\right\}\right|$$
$$\ge \left|\left\{i \in [n] : r_i^* > \overline{u^{(k-1)}}, \left|\left\{j \in [n] : r_j^* > \underline{u^{(k-1)}}, \theta^*_{r_j^*} \ge \theta^*_{r_i^*} + 2M - \delta_1\right\}\right| \le \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta}\right\}\right|.$$

For any $i \in [n]$, since $\theta^* \in \Theta_n(\beta, C_0)$, we have

$$\left|\left\{j \in [n] : r_j^* > \underline{u^{(k-1)}}, \theta^*_{r_j^*} \ge \theta^*_{r_i^*} + 2M - \delta_1\right\}\right| \le r_i^* - \left\lfloor \frac{2M-\delta_1}{C_0\beta}\right\rfloor - \underline{u^{(k-1)}}.$$

Hence,

$$\begin{aligned}\left|\underline{S_k}\right| &\ge \left|\left\{i \in [n] : r_i^* > \overline{u^{(k-1)}}, r_i^* \le \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta} + \left\lfloor \frac{2M-\delta_1}{C_0\beta}\right\rfloor + \underline{u^{(k-1)}}\right\}\right| \\ &\ge \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta} + \left\lfloor \frac{2M-\delta_1}{C_0\beta}\right\rfloor + \underline{u^{(k-1)}} - \overline{u^{(k-1)}} \\ &\ge \left(\frac{1.7}{C_0} + \frac{1}{1+\frac{0.11}{C_0^2}}\right)\frac{M}{\beta}.\end{aligned}$$

(Establishment of (S14)). From (S12), we have $\left|\overline{S_k}\backslash\underline{S_k}\right| \le \overline{u^{(k)}} - \underline{u^{(k)}}$. Hence, we only need to show $\overline{u^{(k)}} - \underline{u^{(k)}} \le \frac{0.29M}{C_0\beta}$.

We are going to prove

$$\theta^*_{\overline{u^{(k-1)}}} \ge \theta^*_{\underline{u^{(k)}}} + 2M - \delta_1. \tag{S18}$$

First, by (S14) for $k-1$, (S13), and (S11), we have $\left|\left\{i \in [n] : \underline{u^{(k-1)}} \le r_i^* \le \underline{u^{(k)}}\right\}\right| \ge |\underline{S_k}|$ which leads to $\underline{u^{(k)}} \ge \overline{u^{(k-1)}}$. Let $b \in [n]$ be the index such that $r_b^* = \underline{u^{(k)}} + 1$. Then it means $b \in [n]\backslash\widetilde{S}_{k-1}$ and $\overline{t_b^{(k)}} > \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta}$. By the definition of $\overline{t_i^{(k)}}$, for any $i \in [n]\backslash\widetilde{S}_{k-1}$, we have

$$\left|\left\{j \in [n] : r_j^* \ge \underline{u^{(k-1)}}, \theta^*_{r_j^*} > \theta^*_{r_i^*} + 2M - \delta_1\right\}\right| \ge \left|\left\{j \in [n]\backslash\widetilde{S}_{k-1} : j \in \overline{\mathcal{E}_{1,i}}\right\}\right| = \overline{t_i^{(k)}},$$

which implies $\theta^*_{\underline{u^{(k-1)}}+\overline{t_i^{(k)}}} > \theta^*_{r_i^*} + 2M - \delta_1$. This means

$$\theta^*_{\underline{u^{(k-1)}}} > \theta^*_{r_i^*} + 2M - \delta_1 + \overline{t_i^{(k)}}\beta.$$

Considering the b index here, we have

$$\theta^*_{\underline{u^{(k-1)}}} \ge \theta^*_{\underline{u^{(k)}}+1} + 2M - \delta_1 + \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)}. \tag{S19}$$

Then using (S14) for $k-1$, we have

$$\theta^*_{\overline{u^{(k-1)}}} \geq \theta^*_{\underline{u^{(k)}}+1} + 2M - \delta_1 + \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)} - 0.29M \geq \theta^*_{\underline{u^{(k)}}} + 2M - \delta_1, \tag{S20}$$

which proves (S18). Then for any $i, j \in [n]\backslash\widetilde{S}_{k-1}$ such that $\underline{u^{(k)}} \leq r_i^* < r_j^*$, we have

$$\begin{aligned}
\overline{t_j^{(k)}} - \overline{t_i^{(k)}} &= \left|\left\{l \in [n]\backslash\widetilde{S}_{k-1} : \theta^*_{r_l^*} \geq \theta^*_{r_j^*} + 2M - \delta_1\right\}\right| - \left|\left\{l \in [n]\backslash\widetilde{S}_{k-1} : \theta^*_{r_l^*} \geq \theta^*_{r_i^*} + 2M - \delta_1\right\}\right| \\
&= \left|\left\{l \in [n]\backslash\widetilde{S}_{k-1} : \theta^*_{r_i^*} + 2M - \delta_1 > \theta^*_{r_l^*} \geq \theta^*_{r_j^*} + 2M - \delta_1\right\}\right| \\
&\geq \left|\left\{r_l^* \geq \overline{u^{(k-1)}} : \theta^*_{r_i^*} + 2M - \delta_1 > \theta^*_{r_l^*} \geq \theta^*_{r_j^*} + 2M - \delta_1\right\}\right| \\
&\geq \left|\left\{l \in [n] : \theta^*_{r_i^*} + 2M - \delta_1 > \theta^*_{r_l^*} \geq \theta^*_{r_j^*} + 2M - \delta_1\right\}\right| \\
&\geq \frac{\theta^*_{r_i^*} - \theta^*_{r_j^*}}{C_0\beta} \\
&\geq \frac{r_j^* - r_i^*}{C_0},
\end{aligned}$$

where in the first inequality we use (S17) and in the second inequality we use (S18). The last two inequalities are due to $\theta^* \in \Theta_n(\beta, C_0)$. As a result,

$$\overline{u^{(k)}} - \underline{u^{(k)}} \leq \frac{\frac{M}{\left(1-\frac{0.12}{C_0^2}\right)\beta} - \frac{M}{\left(1+\frac{0.11}{C_0^2}\right)\beta}}{C_0} \leq \frac{0.29M}{C_0\beta}. \tag{S21}$$

(Establishment of (S15)). We first have

$$\left|\overline{S_k}\right| \leq \overline{u^{(k)}} - \underline{u^{(k-1)}} \leq \overline{u^{(k)}} - \left(\overline{u^{(k-1)}} - \frac{0.29M}{C_0\beta}\right)$$

due to induction hypothesis on (S14) for $k-1$ and $\left\{i \in [n] : r_i^* \leq \underline{u^{(k-1)}}\right\} \subset \widetilde{S}_{k-1}$. By the definition of $\overline{u^{(k)}}$, similar to the proof of (S19), we can show

$$\theta^*_{\overline{u^{(k-1)}}} \leq \theta^*_{\overline{u^{(k)}}} + 2M - \delta_1 + \frac{M}{\left(1-\frac{0.12}{C_0^2}\right)}$$

which implies

$$\overline{u^{(k)}} - \overline{u^{(k-1)}} \leq \frac{2M}{\beta} + \frac{M}{\left(1-\frac{0.12}{C_0^2}\right)\beta}.$$

Therefore,

$$\left|\overline{S_k}\right| \leq \left(2 + \frac{0.29}{C_0} + \frac{1}{1-\frac{0.12}{C_0^2}}\right)\frac{M}{\beta}.$$

(Establishment of (S16)). Define

$$\mathcal{F}^{(k)\prime} = \left\{ \min_{i\in[n]:r_i^*>\overline{u^{(k)}}} \sum_{j\in\underline{S_k}} A_{ij}\mathbb{I}\left\{j \in \underline{\mathcal{E}_{1,i}}\right\} > h \right\}.$$

We are going to show the event $\mathcal{F}^{(k)\prime}$ is a sufficient condition for (S16). By definition, since $\underline{S_k} \subset [n]\backslash\widetilde{S}_{k-1}$, we have $w_i^{(k)\prime} \le w_i^{(k)}$ which implies $S_k \subset S_k'$. We only need to show $S_k' \subset S_k$. Note that for any i such that $r_i^* > \overline{u^{(k)}}$, we have

$$w_i^{(k)\prime} = \sum_{j\in\underline{S_k}} A_{ij}\mathbb{I}\left\{j \in \mathcal{E}_{1,i}\right\} \ge \sum_{j\in\underline{S_k}} A_{ij}\mathbb{I}\left\{j \in \underline{\mathcal{E}_{1,i}}\right\} > h,$$

which means $i \notin S_k'$ as $\mathcal{F}^{(k)\prime}$ is assumed to be true. Hence to show $S_k' \subset S_k$, we only need to show $S_k' \cap \{i \in [n] : r_i^* \le \overline{u^{(k)}}\} \subset S_k$. Note that due to (S14), for any $i, j \in [n]$, such that $r_i^* \le \overline{u^{(k)}}$ and $r_j^* > \underline{u^{(k)}}$, we have $r_j^* > \underline{u^{(k)}}$, $\theta^*_{r_i^*} - \theta^*_{r_j^*} \ge \theta^*_{\overline{u^{(k)}}} - \theta^*_{\underline{u^{(k)}}} \ge -0.29M$. Then for any i such that $r_i^* \le \overline{u^{(k)}}$, we have

$$\begin{aligned}
w_i^{(k)\prime} - w_i^{(k)} &= \sum_{j\in\underline{S_k}} A_{ij}\mathbb{I}\left\{j \in \mathcal{E}_{1,i}\right\} - \sum_{j\in[n]\backslash\widetilde{S}_{k-1}} A_{ij}\mathbb{I}\left\{j \in \mathcal{E}_{1,i}\right\} \\
&\ge - \sum_{j\in[n]:r_j^*>\underline{u^{(k)}}} \mathbb{I}\left\{j \in \mathcal{E}_{1,i}\right\} \\
&\ge - \sum_{j\in[n]:r_j^*>\underline{u^{(k)}}} \mathbb{I}\left\{j \in \overline{\mathcal{E}_{1,i}}\right\} \\
&= - \sum_{j\in[n]:r_j^*>\underline{u^{(k)}}} \mathbb{I}\left\{\theta^*_{r_j^*} \ge \theta^*_{r_i^*} + 2M - \delta_1\right\} \\
&= 0,
\end{aligned}$$

where first inequality is due to (S11). Hence we have $S_k' \cap \{i \in [n] : r_i^* \le \overline{u^{(k)}}\} \subset S_k$ which leads to $S_k = S_k'$. As a result, to establish (S16), we only need to analyze $\mathbb{P}\left(\mathcal{F}^{(k)\prime}\right)$.

The analysis of $\mathbb{P}\left(\mathcal{F}^{(k)\prime}\right)$ is similar to that of $\mathbb{P}\left(\mathcal{F}^{(k)}\right)$ in the establishment of (S12). By a similar independence argument, we have

$$\left(\sum_{j\in\underline{S_k}} A_{ij}\mathbb{I}\left\{j \in \underline{\mathcal{E}_{1,i}}\right\}\right) \Bigg| \widetilde{S}_{k-1} \sim \text{Binomial}\left(\left|\underline{S_k} \cap \underline{\mathcal{E}_{1,i}}\right|, p\right)$$

for any $i \in [n]$ such that $r_i^* > \overline{u^{(k)}}$. From (S20), we have

$$\underline{u^{(k)}} - \overline{u^{(k-1)}} \ge \frac{2M - \delta_1}{C_0\beta}. \tag{S22}$$

Together with (S11) and (S17), we have

$$\left|\underline{S_k} \cap \underline{\mathcal{E}_{1,i}}\right| \geq \left|\left\{j \in [n] : \overline{u^{(k-1)}} \leq r_j^* \leq \underline{u^{(k)}}, \theta^*_{r_j^*} \geq \theta^*_{r_i^*} + 2M + \delta_1\right\}\right| \geq \underline{u^{(k)}} - \overline{u^{(k-1)}} \geq \frac{2M - \delta_1}{C_0 \beta}.$$

Recall that $h = pM/\beta$ and $p/(\beta \log n) \to \infty$. By Bernstein inequality, we have

$$\mathbb{P}\left(\mathcal{F}^{(k)\prime} | \widetilde{S}_{k-1}\right) = \mathbb{P}\left(\min_{i \in [n]: r_i^* \geq \overline{u^{(k)}}} \left(\sum_{j \in \underline{S_k}} A_{ij} \mathbb{I}\left\{j \in \underline{\mathcal{E}_{1,i}}\right\}\right) > h \middle| \widetilde{S}_{k-1}\right) \geq 1 - O(n^{-10}).$$

Since this holds for all $\widetilde{S}_{k-1}$, we have $\mathbb{P}\left(\mathcal{F}^{(k)\prime}\right) \geq 1 - O(n^{-10})$.

(Establishment of (S12) - (S16) for K). We have (S12) - (S16) hold for each $k \in [K-1]$ with probability at least $1 - O(n^{-9})$. For the last partition, $S_K = [n] \backslash \widetilde{S}_{K-1}$. Let $S_{K,1}$ be the set obtained by Algorithm 1 before the terminating condition $[n] - |S_1| + ... + |S_{K,1}| \leq |S_{K,1}|/2$ is met. $\underline{S_{K,1}}, \overline{S_{K,1}}$ can be similarly defined and (S12) - (S16) should also be satisfied by $S_{K,1}$. Therefore,

$$|S_K| \leq \frac{3\,|S_{K,1}|}{2} \leq \frac{3}{2}\left|\overline{S_{K,1}}\right| \leq \frac{3}{2}\left(2 + \frac{0.29}{C_0} + \frac{1}{1 - \frac{0.12}{C_0^2}}\right)\frac{M}{\beta},$$

$$|S_K| \geq |S_{K,1}| \geq \left|\underline{S_{K,1}}\right| \geq \left(\frac{1.7}{C_0} + \frac{1}{1 + \frac{0.11}{C_0^2}}\right)\frac{M}{\beta}$$

and

$$\left\{i \in [n] : r_i^* > \overline{u^{(K-1)}}\right\} \subset S_K \subset \left\{i \in [n] : r_i^* > \underline{u^{(K-1)}}\right\}.$$

So far, we have establish (S12) - (S16) for any $k \in [K]$. Now we are ready to use them to prove the conclusions in Theorem 4.1.

1. Conclusion 1 is a consequence of (S12) and (S15).

2. For Conclusion 2, by (S22) we have $\overline{u^{(k-2)}} < \underline{u_{(k-1)}} < \overline{u^{(k-1)}} < \underline{u_{(k)}} < \overline{u^{(k)}} < \underline{u_{(k+1)}}$. Together with (S12) and (S17), we have

   $$\left\{i \in [n] : \overline{u^{(k-2)}} < r_i^* \leq \underline{u^{(k+1)}}\right\} \subset S_{k-1} \cup S_k \cup S_{k+1} \subset \left\{i \in [n] : \underline{u^{(k-2)}} < r_i^* \leq \overline{u^{(k+1)}}\right\}.$$

   Therefore, using (S22), for any i such that $\underline{u^{(k-1)}} < r_i^* \leq \overline{u^{(k)}}$,

   $$\begin{aligned}&\left\{j \in [n] : \left|r_i^* - r_j^*\right| \leq \frac{1.51M}{C_0 \beta}\right\} \\ \subset &\left\{j \in [n] : \underline{u^{(k-1)}} - \frac{1.51M}{C_0\beta} \leq r_j^* \leq \overline{u^{(k)}} + \frac{1.51M}{C_0\beta}\right\} \\ \subset &\left\{j \in [n] : \overline{u^{(k-2)}} < r_j^* \leq \underline{u^{(k+1)}}\right\} \subset S_{k-1} \cup S_k \cup S_{k+1}.\end{aligned}$$

   For $k = 1$ or K, only oneside needs to be considered and the property still holds due to the gap between $\underline{u^{(2)}}$ and $\overline{u^{(1)}}$ as well as the gap between $\underline{u^{(K-1)}}$ and $\overline{u^{(K-2)}}$.

3. For Conclusion 3, by (S12) and (S17), we have

$$\left\{i \in [n] : \overline{u^{(k-1)}} < r_i^* \le \underline{u^{(k)}}\right\} \subset S_k \subset \left\{i \in [n] : \underline{u^{(k-1)}} < r_i^* \le \overline{u^{(k)}}\right\}. \tag{S23}$$

   Using (S22), we have

$$\max\left\{r_i^* : i \in S_k\right\} \le \overline{u^{(k)}} < \underline{u^{(k+1)}} < \min\left\{r_i^* : i \in S_{k+2}\right\}.$$

   Same results can be established for $\max\left\{r_i^* : i \in S_k\right\} < \min\left\{r_i^* : i \in S_l\right\}$ for any $l > k+2$.

4. For Conclusion 4, for any k and any i, the definition of $w_i^{(k)\prime}$ only involves j such that $j \in \overline{\mathcal{E}_{1,i}}$. This implies that the definition of S_k' only involves information of $(A_{ij}, \bar{y}_{ij}^{(1)})$ such that $\theta^*_{r_j^*} - \theta^*_{r_i^*} \ge 2M - \delta_1$. Thus S_k' can be used as the $\check{S}_k$ in Theorem 4.1.

5. For Conclusion 5, note that for any $k \in [K]$ and $i \in S_k$, we have

$$\begin{aligned}\left|\left\{j \in [n] : |\theta^*_{r_i^*} - \theta^*_{r_j^*}| \le \frac{M}{2}\right\} \cap S_k\right| &\ge \left|\left\{j \in [n] : |r_i^* - r_j^*| \le \frac{M}{2C_0\beta}\right\} \cap S_k\right| \\ &\ge \left|\left\{j \in [n] : |r_i^* - r_j^*| \le \frac{M}{2C_0\beta}, \overline{u^{(k-1)}} < r_j^* \le \underline{u^{(k)}}\right\}\right|.\end{aligned}$$

   where the last inequality is by (S23). Again by (S23), we have $\underline{u^{(k-1)}} < r_i^* \le \overline{u^{(k)}}$. From (S22) we know $\underline{u^{(k)}} - \overline{u^{(k-1)}} > M/(2C_0\beta)$. Then we have

$$\begin{aligned}\left|\left\{j \in [n] : |\theta^*_{r_i^*} - \theta^*_{r_j^*}| \le \frac{M}{2}\right\} \cap S_k\right| &\ge \frac{M}{2C_0\beta} - \max\left\{\overline{u^{(k-1)}} - \underline{u^{(k-1)}}, \overline{u^{(k)}} - \underline{u^{(k)}}\right\} \\ &\ge \frac{0.21M}{C_0\beta}\end{aligned}$$

   where the last inequality is by (S14).

The proof is complete. □

C Proofs of Lemma 4.1, Lemma 4.2, and Lemma 4.3

We first prove Lemma 4.1 below.

Proof of Lemma 4.1. Recall that $\widehat{r}$ is obtained by sorting $\left\{\sum_{j\in[n]\setminus\{i\}} R_{ij}\right\}_{i\in[n]}$. Define

$$\widehat{s}_i = \sum_{j\in[n]\setminus\{i\}} R_{ij},$$

$$\widehat{R}_{ij} = \mathbb{I}\left\{\widehat{s}_i > \widehat{s}_j\right\} = \mathbb{I}\left\{\widehat{r}_i < \widehat{r}_j\right\}$$

and

$$s_i^* = \sum_{j\in[n]\setminus\{i\}} R_{ij}^*.$$

Observe that

$$r_i^* = n - s_i^*,$$

we have

$$\widehat{R}_{ij} = \mathbb{I}\left\{\widehat{s}_i > \widehat{s}_j\right\} = \mathbb{I}\left\{\widehat{s}_i - s_i^* + s_i^* > \widehat{s}_j - s_j^* + s_j^*\right\} = \mathbb{I}\left\{\widehat{s}_i - s_i^* - (\widehat{s}_j - s_j^*) > r_i^* - r_j^*\right\}.$$

Thus

$$\begin{aligned}
\mathsf{K}(\widehat{r}, r^*) &= \frac{1}{n}\sum_{1\le i<j\le n} \mathbb{I}\left\{\widehat{R}_{ij} \neq R_{ij}^*\right\} \\
&= \frac{1}{n}\sum_{1\le i<j\le n} \left|\mathbb{I}\left\{\widehat{s}_i - s_i^* - (\widehat{s}_j - s_j^*) > r_i^* - r_j^*\right\} - \mathbb{I}\left\{0 > r_i^* - r_j^*\right\}\right| \\
&\le \frac{1}{n}\sum_{1\le i<j\le n} \mathbb{I}\left\{\left|\widehat{s}_i - s_i^* - (\widehat{s}_j - s_j^*)\right| \ge \left|r_i^* - r_j^*\right|\right\} \\
&\le \frac{1}{n}\sum_{1\le i<j\le n} \mathbb{I}\left\{\left|\left|r_i^* - r_j^*\right| \le \left|\widehat{s}_i - s_i^*\right| + \left|\widehat{s}_j - s_j^*\right|\right|\right\} \\
&= \frac{1}{n}\sum_{k=1}^{n-1}\sum_{\substack{1\le i<j\le n\\ \left|r_i^*-r_j^*\right|=k}} \mathbb{I}\left\{k \le \left|\widehat{s}_i - s_i^*\right| + \left|\widehat{s}_j - s_j^*\right|\right\} \\
&\le \frac{1}{n}\sum_{k=1}^{n-1}\sum_{\substack{1\le i<j\le n\\ \left|r_i^*-r_j^*\right|=k}} \mathbb{I}\left\{\frac{k}{2} \le \left|\widehat{s}_i - s_i^*\right|\right\} + \mathbb{I}\left\{\frac{k}{2} \le \left|\widehat{s}_j - s_j^*\right|\right\} \\
&\le \frac{2}{n}\sum_{i=1}^{n}\sum_{k=1}^{n-1} \mathbb{I}\left\{\frac{k}{2} \le \left|\widehat{s}_i - s_i^*\right|\right\} \le \frac{4}{n}\sum_{i=1}^{n}\sum_{k=1}^{n} \mathbb{I}\left\{k \le \left|\widehat{s}_i - s_i^*\right|\right\} \\
&= \frac{4}{n}\sum_{i=1}^{n}\left|\widehat{s}_i - s_i^*\right| \le \frac{4}{n}\sum_{i=1}^{n}\sum_{j\in[n]\setminus\{i\}}\left|R_{ij} - R_{ij}^*\right| = \frac{4}{n}\sum_{1\le i\neq j\le n}\mathbb{I}\left\{R_{ij} \neq R_{ij}^*\right\}
\end{aligned}$$

which completes the proof. □

Next, we prove Lemma 4.2.

Proof of Lemma 4.2. Recall $\mathcal{E} = \left\{(i,j) : 1 \le i < j \le n, \psi(-M) \le \overline{y}_{ij}^{(1)} \le \psi(M)\right\}$. Then on the event where Lemma B.2 holds, $\mathcal{E}$ can be written as

$$\mathcal{E} = \left\{(i,j) : 1 \le i < j \le n, \left|\theta_{r_i^*} - \theta_{r_j^*}\right| \le M/2\right\} \uplus \left(\mathcal{E} \cap \left\{(i,j) : M/2 < \left|\theta_{r_i^*} - \theta_{r_j^*}\right| < 1.1M\right\}\right)$$

which implies $\check{A}_{ij} = A_{ij}\mathbb{I}\left\{(i,j) \in \mathcal{E}\right\}$. Moreover, on the event where Theorem 4.1 holds, $\check{S}_k = S_k, k \in [K]$ by Conclusion 4. This proves $\ell^{(k)}(\theta) = \check{\ell}^{(k)}(\theta)$ with probability at least $1 - O(n^{-8})$. $\check{\theta}^{(k)}$ and $\widehat{\theta}^{(k)}$ are equivalent up to a common shift since the Hessian in local MLE is well conditioned with probability at least $1 - O(n^{-8})$ due to Lemma C.2. □

Finally, we need to prove Lemma 4.3, which requires us to first establish a few extra lemmas.

Lemma C.1. *For any integer constant $C \geq 1$, define a matrix $M \in \{0,1\}^{n\times n}$ such that $M_{ij} = \mathbb{I}\{|i-j| \leq n/C\}$. Let $\mathcal{L}_M$ be its Laplacian matrix such that*

$$[\mathcal{L}_M]_{ij} = \begin{cases} -M_{ij}, & \text{if } i \neq j \\ \sum_l M_{il}, & \text{if } i = j. \end{cases}$$

Denote $\lambda_{\min,\perp}(\mathcal{L}_M)$ to be the second smallest eigenvalue of $\mathcal{L}_M$, i.e., $\lambda_{\min,\perp}(\mathcal{L}_M) = \min_{u\neq 0, \mathbb{1}_n^T u = 0} \frac{u^T \mathcal{L}_M u}{\|u\|}$. Then there exists another constant $C' > 0$ that only depends on C such that

$$\frac{1}{n}\lambda_{\min,\perp}(\mathcal{L}_M) = \inf_{\substack{x\in\mathbb{R}^n \\ \mathbb{1}_n^T x = 0 \\ \|x\|=1}} \frac{\sum_{|i-j|\leq \frac{n}{C}} (x_i - x_j)^2}{n} \geq C'.$$

Proof. We partition $[n]$ into $4C$ consecutive blocks such that each block contains either $\lceil n/4C \rceil$ or $\lfloor n/4C \rfloor$ consecutive indices. Let these blocks be a sequence of disjoint sets $B_1, ..., B_{4C}$ such that $\max_{i\in B_k} i < \min_{j\in B_l} j$ if $k < l$. The idea is to lower bound the summation over the diagonal region by a sequence of square regions. Thus, for any $x \in \mathbb{R}^n, \mathbb{1}_n^T x = 0, \|x\| = 1$, we have

$$\begin{aligned}
\frac{1}{n} x^T \mathcal{L}_M x &= \frac{\sum_{|i-j|\leq \frac{n}{C}} (x_i - x_j)^2}{n} \\
&\geq \frac{1}{n} \left[\sum_{k,l\in[4C]:|k-l|\leq 1} \sum_{i\in B_k, j\in B_l} (x_i - x_j)^2 \right] \\
&= \sum_{k,l\in[4C]:|k-l|\leq 1} \left(\frac{|B_l|}{n} \sum_{i\in B_k} x_i^2 + \frac{|B_k|}{n} \sum_{i\in B_l} x_i^2 - 2 \left(\sum_{i\in B_k} \frac{x_i}{\sqrt{n}} \right) \left(\sum_{i\in B_l} \frac{x_i}{\sqrt{n}} \right) \right) \\
&= \sum_{k,l\in[4C]:|k-l|\leq 1} \left(p_l z_k + p_k z_l - 2 y_k y_l \right),
\end{aligned}$$

where we denote

$$y_k = \sum_{i\in B_k} \frac{x_i}{\sqrt{n}}, z_k = \sum_{i\in B_k} x_i^2, p_k = \frac{|B_k|}{n}.$$

For any $k \in [4C]$, we define

$$w_{2k-1} = \frac{y_k + \sqrt{p_k z_k - y_k^2}}{2p_k}, \text{ and } w_{2k} = \frac{y_k - \sqrt{p_k z_k - y_k^2}}{2p_k}. \tag{S24}$$

Note that for any $k, l \in [4C]$, we have

$$\begin{aligned}
&p_l p_k \left((w_{2k-1} - w_{2l-1})^2 + (w_{2k-1} - w_{2l})^2 + (w_{2k} - w_{2l-1})^2 + (w_{2k} - w_{2l})^2 \right) \\
&= p_l p_k \left(2 \left(w_{2k-1}^2 + w_{2k}^2 + w_{2l-1}^2 + w_{2k}^2 \right) - 2 \left(w_{2k-1} + w_{2k} \right) \left(w_{2l-1} + w_{2l} \right) \right) \\
&= p_l p_k \left(\frac{z_k}{p_k} + \frac{z_l}{p_l} - 2 \frac{y_k y_l}{p_k p_l} \right) \\
&= p_l z_k + p_k z_l - 2 y_k y_l.
\end{aligned}$$

Then we have

$$\begin{aligned}
&\sum_{k,l \in [4C] : |k-l| \leq 1} \left(p_l z_k + p_k z_l - 2 y_k y_l \right) \\
&= \sum_{k,l \in [4C] : |k-l| \leq 1} p_l p_k \left((w_{2k-1} - w_{2l-1})^2 + (w_{2k-1} - w_{2l})^2 + (w_{2k} - w_{2l-1})^2 + (w_{2k} - w_{2l})^2 \right).
\end{aligned}$$

Note that w is a function of y, z, p which by definition satisfy: $\sum_{k=1}^{4C} y_k = 0$, $\sum_{k=1}^{4C} z_k = 1$, $\min_{k \in [4C]} p_k \geq 1/(5C)$, $\sum_{k=1}^{4C} p_k = 1$, and $y_k^2 \leq p_k z_k$ for all $k \in [4C]$. Define a parameter space T:

$$T = \left\{ (y, z, p) : \sum_{k=1}^{4C} y_k = 0, \sum_{k=1}^{4C} z_k = 1, \min_{k \in [4C]} p_k \geq 1/(5C), \sum_{k=1}^{4C} p_k = 1, \text{ and } y_k^2 \leq p_k z_k, \forall k \in [4C] \right\}.$$

Then we have

$$\begin{aligned}
&\frac{1}{n} \lambda_{\min,\perp}(\mathcal{L}_M) \\
&\geq \inf_{(y,z,p) \in T} \sum_{k,l \in [4C] : |k-l| \leq 1} p_l p_k \left((w_{2k-1} - w_{2l-1})^2 + (w_{2k-1} - w_{2l})^2 + (w_{2k} - w_{2l-1})^2 + (w_{2k} - w_{2l})^2 \right),
\end{aligned} \tag{S25}$$

where w is defined in (S24).

Since T only depends on C, the quantity (S25) also only depends on C. Then, (S25) is equal to some constant $C' \geq 0$ only depending on C. We are going to show $C' > 0$. Otherwise, let the infimum of (S25) be achieved at some w with $(y, z, p) \in T$. Then, we must have $w_{2k-1} = w_{2l-1} = w_{2k} = w_{2l}$ for any $k, l \in [4C]$ such that $|k - l| \leq 1$. This has two immediately implications. First, for any $k \in [4C]$, since $w_{2k-1} = w_{2k}$, we have $y_k^2 = p_k z_k$ and $w_k = y_k/(2p_k)$, Second, since $w_{2k} = w_{2(k+1)}$ for any $k \in [4C - 1]$, there exists some c such that $y_k/p_k = c$ for all $k \in [4C]$. Together with $y_k^2 = p_k z_k$, we obtain $c^2 p_k = z_k$ for all $k \in [C]$. Since $\sum_{k=1}^{4C} z_k = 1$ and $\sum_{k=1}^{4C} p_k = 1$, we conclude $c = \pm 1$. Then using $y_k/p_k = c$, we have $\sum_{k=1}^{4C} y_k = c \sum_{k=1}^{4C} p_k = c \neq 0$, which is a contradiction with $\sum_{k=1}^{4C} y_k = 0$. As a result, we obtain $\frac{1}{n} \lambda_{\min,\perp}(\mathcal{L}_M) \geq C' > 0$. □

Lemma C.2. *Under the assumptions in Lemma 4.3,*

$$\lambda_{\min,\perp}(H(\eta^*)) = \min_{u \neq 0 : \mathbb{1}_m^T u = 0} \frac{u^T H(\eta^*) u}{\|u\|^2} \gtrsim mp$$

with probability at least $1-O(n^{-10})$*, where* $H(\eta^*)$ *is the Hessian matrix of the objective (28), defined by*

$$H_{ij}(\eta^*) = \begin{cases} \sum_{l\in[m]\setminus\{i\}} B_{il}\psi'(\eta_i^* - \eta_l^*), & i = j, \\ -B_{ij}\psi'(\eta_i^* - \eta_j^*), & i \neq j. \end{cases}$$

Proof. We can decompose $H(\eta^*)$ into stochastic part $H(\eta^*) - \mathbb{E}(H(\eta^*))$ ans deterministic part $\mathbb{E}(H(\eta^*))$ and bound them separately. We first look at the stochastic part. Note that

$$H(\eta^*) - \mathbb{E}(H(\eta^*)) = D - \mathbb{E}(D) - \sum_{i<j}(B_{ij} - p_{ij})\psi'(\eta_i^* - \eta_j^*)(E_{ij} + E_{ji})$$

where $D = diag\{D_1, ..., D_m\} = diag\{\sum_{j\neq 1} B_{ij}\psi'(\eta_1^* - \eta_j^*), ..., \sum_{j\neq m} B_{mj}\psi'(\eta_m^* - \eta_j^*)\}$; E_{ij} is an $m \times m$ matrix and has 1 on the entry (i, j) and 0 otherwise. We also have $\|(B_{ij} - p_{ij})\psi'(\eta_i^* - \eta_j^*)(E_{ij} + E_{ji})\|_{\text{op}} \leq 1$ and $\|\sum_{i<j}(B_{ij} - p_{ij})^2\psi'(\eta_i^* - \eta_j^*)^2(E_{ij} + E_{ji})^2\|_{\text{op}} \leq mp$. By matrix Bernstein inequality in [5], we have

$$\mathbb{P}\left(\|\sum_{i<j}(B_{ij} - p_{ij})(E_{ij} + E_{ji})\|_{\text{op}} > t\right) \leq 2m \exp\left(-\frac{t^2/2}{mp + \frac{t}{3}}\right).$$

Taking $t = C_1'\sqrt{mp\log n}$ for some large enough constant $C_1' > 0$, we have

$$\|\sum_{i<j}(B_{ij} - p_{ij})\psi'(\eta_i^* - \eta_j^*)(E_{ij} + E_{ji})\|_{\text{op}} \leq C_1'\sqrt{mp\log n}$$

with probability at least $1-O(n^{-10})$. Standard concentration using Bernstein inequality also yields

$$\|D - \mathbb{E}(D)\|_{\text{op}} \leq C_2'\sqrt{mp\log n}$$

for some constant $C_2' > 0$ with probability at least $1 - O(n^{-10})$. Thus the stochastic part

$$\|H(\eta^*) - \mathbb{E}(H(\eta^*))\|_{\text{op}} \leq (C_1' + C_2')\sqrt{mp\log n} = o(mp) \tag{S26}$$

with probability at least $1 - O(n^{-10})$.

For the deterministic part, we first choose a constant integer $C' > 0$ such that for any $|i - j| \leq \frac{n}{C'}$, $p_{ij} = p$. Thus for any unit vector $x \in \mathbb{R}^m$ such that $\mathbb{1}_m^T x = 0$,

$$\begin{aligned} \frac{x^T\mathbb{E}(H(\eta^*))x}{m} &= \frac{\sum_{i<j} p_{ij}\psi'(\eta_i^* - \eta_j^*)(x_i - x_j)^2}{m} \\ &\geq \frac{\sum_{i<j,|i-j|\leq\frac{m}{C'}} p\psi'(\eta_i^* - \eta_j^*)(x_i - x_j)^2}{m} \\ &\gtrsim p\frac{\sum_{i<j,|i-j|\leq\frac{m}{C'}}(x_i - x_j)^2}{m} && \text{(S27)} \\ &\gtrsim p && \text{(S28)} \end{aligned}$$

where (S27) uses the boundedness of $\eta_1^* - \eta_m^*$; (S28) is a consequence of Lemma C.1 and C' is a constant independent of m and n. Combing (S26) and (S28) concludes the proof. $\square$

The proof of Lemma 4.3 is given below.

Proof of Lemma 4.3. Since $\frac{L(\eta_i^*-\eta_j^*)^2}{2(W_i(\eta^*)+W_j(\eta^*))} \asymp mpL(\eta_i^*-\eta_j^*)^2$, we only need to consider the situation where $mpL(\eta_i^*-\eta_j^*)^2$ is greater than a sufficiently large constant, since otherwise we can use the trivial bound $\mathbb{P}\left(\widehat{\eta}_i < \widehat{\eta}_j\right) \leq 1$. Define

$$\widetilde{\eta}_j = \eta_j^* - \frac{\sum_{l\in[m]\setminus\{j\}} B_{jl}(\bar{y}_{jl} - \psi(\eta_j^* - \eta_l^*))}{\sum_{l\in[m]\setminus\{j\}} B_{jl}\psi'(\eta_j^* - \eta_l^*)}.$$

Following the same argument used in the proof of Theorem 3.2 of [1] (see Equations (63), (64) and (66) of [1]), we have

$$|\widehat{\eta}_i - \widetilde{\eta}_i| \vee |\widehat{\eta}_j - \widetilde{\eta}_j| \leq \delta\Delta, \tag{S29}$$

with probability at least $1 - O(n^{-7}) - \exp(-\Delta^{3/2}Lmp) - \exp\left(-\Delta^2 mpL\frac{mp}{\log(n+m)}\right)$, where $\Delta = \min\left(\eta_i^* - \eta_j^*, \left(\frac{\log(n+m)}{mp}\right)^{1/4}\right)$ and $\delta > 0$ is some sufficiently small constant. In fact, the bound (S29) has only been established in [1] with a random graph that satisfies $p_{ij} = p$ for all $1 \leq i < j \leq m$. To establish (S29) under the more general setting of interest, we first have

$$\lambda_{\min,\perp}(H(\eta^*)) = \min_{u\neq 0:\mathbb{1}_m^T u=0} \frac{u^T H(\eta^*)u}{\|u\|^2} \gtrsim mp, \tag{S30}$$

with high probability, where $H(\eta^*)$ is the Hessian matrix of the objective (28). This is established in Lemma C.2. Note that (S30) is the only difference between the proofs of the current setting and the setting in [1]. With (S29), we have

$$\begin{aligned}\mathbb{P}\left(\widehat{\eta}_i < \widehat{\eta}_j\right) &\leq \mathbb{P}\left(\widetilde{\eta}_j - \eta_j^* - (\widetilde{\eta}_i - \eta_i^*) > (1-\delta)\Delta\right) \\ &\quad + O(n^{-7}) + \exp(-\Delta^{3/2}Lmp) + \exp\left(-\Delta^2 mpL\frac{mp}{\log(n+m)}\right).\end{aligned}$$

Define

$$\mathcal{B} = \left\{B : \left|\frac{\sum_{l\in[m]\setminus\{j\}} p_{jl}\psi'(\eta_j^* - \eta_l^*)}{\sum_{l\in[m]\setminus\{j\}} B_{jl}\psi'(\eta_j^* - \eta_l^*)} - 1\right| \leq \delta, \left|\frac{\sum_{l\in[m]\setminus\{i\}} p_{il}\psi'(\eta_i^* - \eta_l^*)}{\sum_{l\in[m]\setminus\{i\}} B_{il}\psi'(\eta_i^* - \eta_l^*)} - 1\right| \leq \delta'\right\}.$$

By Bernstein's inequality, we have $\mathbb{P}(B \in \mathcal{B}^c) \leq O(n^{-7})$ for some $\delta' = o(1)$. We then have

$$\begin{aligned}&\mathbb{P}\left(\widetilde{\eta}_j - \eta_j^* - (\widetilde{\eta}_i - \eta_i^*) > (1-\delta)\Delta\right) \\ \leq\ & \sup_{B\in\mathcal{B}} \mathbb{P}\left(-\frac{\sum_{l\in[m]\setminus\{j\}} B_{jl}(\bar{y}_{jl} - \psi(\eta_j^* - \eta_l^*))}{\sum_{l\in[m]\setminus\{j\}} B_{jl}\psi'(\eta_j^* - \eta_l^*)}\right. \\ &\left. + \frac{\sum_{l\in[m]\setminus\{i\}} B_{il}(\bar{y}_{il} - \psi(\eta_i^* - \eta_l^*))}{\sum_{l\in[m]\setminus\{i\}} B_{il}\psi'(\eta_i^* - \eta_l^*)} > (1-\delta)\Delta \,\middle|\, B\right) + O(n^{-7}) \\ \leq\ & \exp\left(-\frac{(1-2\delta)L(\eta_i^* - \eta_j^*)^2}{2(W_i(\eta^*) + W_j(\eta^*))}\right) + O(n^{-7}).\end{aligned}$$

Since

$$\exp(-\Delta^{3/2}Lmp) + \exp\left(-\Delta^2 mpL\frac{mp}{\log(n+m)}\right) \lesssim \exp\left(-\frac{(1-2\delta)L(\eta_i^*-\eta_j^*)^2}{2(W_i(\eta^*)+W_j(\eta^*))}\right) + O(n^{-7}),$$

we obtain the desired conclusion.

□

D Proofs of Lemma 8.1 and Lemma 8.2

Proof of Lemma 8.1. Define

$$R_\theta(x, t_1, t_2) = \{i : t_1 \le |\theta_i - x| < t_2\} \tag{S31}$$

It is easy to see that there exist constants $C_1', C_2' > 0$ such that for any $\theta \in \Theta_n(\beta, C_0)$,

$$\frac{C_1'}{\beta \vee 1/n} \le \inf_{x\in[\theta_n,\theta_1]} |R_\theta(x,0,1)| \tag{S32}$$

and

$$\sup_{t\in\mathbb{N}} \sup_{x\in[\theta_n,\theta_1]} |R_\theta(x,t,t+1)| \le \frac{C_2'}{\beta \vee 1/n} \tag{S33}$$

Thus

$$\begin{aligned}
&\inf_{\theta_0\in[\theta_n,\theta_1]} \sum_{i=1}^n \psi'(\theta_0-\theta_i)^\alpha \ge \inf_{\theta_0\in[\theta_n,\theta_1]} \sum_{i\in R_\theta(\theta_0,0,1)} \psi'(\theta_0-\theta_i)^\alpha \\
&= \inf_{\theta_0\in[\theta_n,\theta_1]} \sum_{i\in R_\theta(\theta_0,0,1)} \left[\frac{e^{\theta_0-\theta_i}}{(1+e^{\theta_0-\theta_i})^2}\right]^\alpha \ge \inf_{\theta_0\in[\theta_n,\theta_1]} \sum_{i\in R_\theta(\theta_0,0,1)} \frac{1}{4^\alpha} e^{-\alpha|\theta_0-\theta_i|} \\
&\ge \inf_{\theta_0\in[\theta_n,\theta_1]} |R_\theta(\theta_0,0,1)| \frac{1}{4^\alpha} e^{-\alpha} \ge \frac{C_3'}{\beta\vee 1/n}
\end{aligned}$$

for some constant $C_3' > 0$. On the other hand,

$$\begin{aligned}
&\sup_{\theta_0\in[\theta_n,\theta_1]} \sum_{i=1}^n \psi'(\theta_0-\theta_i)^\alpha = \sup_{\theta_0\in[\theta_n,\theta_1]} \sum_{t\ge 0} \sum_{i\in R_\theta(\theta_0,t,t+1)} \psi'(\theta_0-\theta_i)^\alpha \\
&\le \sup_{\theta_0\in[\theta_n,\theta_1]} \sum_{t\ge 0} \sum_{i\in R_\theta(\theta_0,t,t+1)} e^{-\alpha|\theta_0-\theta_i|} \le \sup_{\theta_0\in[\theta_n,\theta_1]} \sum_{t\ge 0} |R_\theta(\theta_0,t,t+1)| e^{-\alpha t} \\
&\le \frac{C_4'}{\beta\vee 1/n}
\end{aligned}$$

for some constant $C_4' > 0$, which concludes the proof. □

To prove Lemma 8.2, we first establish a few additioinal lemmas.

Lemma D.1 (Central limit theorem, Theorem 2.20 of [4]). *If $Z \sim \mathcal{N}(0,1)$ and $W = \sum_{i=1}^n X_i$ where X_i are independent mean 0 and $\mathsf{Var}(W) = 1$, then*

$$\sup_t |\mathbb{P}(W \le t) - \mathbb{P}(Z \le t)| \le 2\sqrt{3\sum_{i=1}^n \left(\mathbb{E}X_i^4\right)^{3/4}}.$$

Lemma D.2. *Assume $p \ge c_0 \frac{\log n}{n}$ for some sufficiently large constant $c_0 > 0$. For any fixed $\{w_{ijk}\}$, $i,j \in [n], k \in \mathbb{K}$ where $\mathbb{K}$ is a discrete set with cardinality at most n^{c_1} for some constant $c_1 > 0$. Assume $\max_{i,j\in[n],k\in\mathbb{K}} |w_{ijk}| \le c_2$ and*

$$p \min_{i\in[n],k\in\mathbb{K}} \sum_{j\in[n]\setminus\{i\}} w_{ijk}^2 \ge c_3 \log n$$

for some constants $c_2, c_3 > 0$. Then there exists constants $C_1, C_2 > 0$, such that for any $i \in [n]$,

$$\max_{k\in\mathbb{K}} \sum_{j\in[n]\setminus\{i\}} (A_{ij} - p) w_{ijk} \le C_1 \sqrt{p \log n \max_{k\in\mathbb{K}} \sum_{j\in[n]} w_{ijk}^2}$$

with probability at least $1 - C_2 n^{-10}$.

Proof. For any constant $C_1' > 0$, by Bernstein's inequality, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{k\in\mathbb{K}} \sum_{j\in[n]\setminus\{i\}} (A_{ij} - p) w_{ijk} > C_1' \sqrt{p \log n \max_{k\in\mathbb{K}} \sum_{j\in[n]} w_{ijk}^2}\right) \\
&\le |\mathbb{K}| \max_{k\in\mathbb{K}} \exp\left(-\frac{C_1'^2 p \log n \max_{k\in\mathbb{K}} \sum_{j\in[n]} w_{ijk}^2}{2p\sum_{j\in[n]\setminus\{i\}} w_{ijk}^2 + \frac{2}{3}\max_{i,j\in[n],k\in\mathbb{K}} |w_{ijk}|\, C_1' \sqrt{p \log n \max_{k\in\mathbb{K}} \sum_{j\in[n]} w_{ijk}^2}}\right) \\
&\le n^{c_1} \exp\left(-\frac{C_1'^2}{C_2'} \log n\right)
\end{aligned}$$

for some constant $C_2' > 0$. Thus we can set C_1' large enough to make the theorem holds. ☐

Lemma D.3. *Assume $p \ge c_0(\beta \vee \frac{1}{n}) \log n$ for some sufficiently large constant $c_0 > 0$ and $1 \le C_0 = O(1)$. For any constant $\alpha > 0$, there exist constants $C_1, C_2, C_3 > 0$ such that for any $r \in \mathfrak{S}_n, i \ne j \in [n]$, and $\theta \in \Theta_n(\beta, C_0)$,*

$$\inf_{u\in[0,1]} \sum_{k\ne i,j} A_{ik} \psi'(u\theta_{r_i} + (1-u)\theta_{r_j} - \theta_{r_k})^\alpha \ge C_1 \frac{p}{\beta \vee 1/n} \tag{S34}$$

and

$$\sup_{u\in[0,1]} \sum_{k\ne i,j} A_{ik} \psi'(u\theta_{r_i} + (1-u)\theta_{r_j} - \theta_{r_k})^\alpha \le C_2 \frac{p}{\beta \vee 1/n} \tag{S35}$$

with probability at least $1 - O(n^{-10})$ for n large enough.

Proof. We remark that $p \geq c_0(\beta \vee \frac{1}{n}) \log n$ necessarily implies $0 < \beta = o(1)$. We only give the proof of (S35). The inf part (S34) can be proved similarly. For (S35),

$$
\begin{aligned}
&\sup_{u\in[0,1]} \sum_{k\neq i,j} A_{ik}\psi'(u\theta_{r_i} + (1-u)\theta_{r_j} - \theta_{r_k})^\alpha \\
&\leq \frac{C_1' p}{\beta \vee 1/n} + \sup_{u\in[0,1]} \sum_{k\neq i,j} (A_{ik} - p)\psi'(u\theta_{r_i} + (1-u)\theta_{r_j} - \theta_{r_k})^\alpha \qquad \text{(S36)}
\end{aligned}
$$

for some constant $C_1' > 0$, where (S36) uses Lemma 8.1. To bound the second term in (S36), we use standard discretization technique. Let $u_a = \frac{a}{n}, a \in [n]$. Then for any $u \in [0,1]$, let $a(u) = \arg\min_{a\in[n]} |u - u_a|$. We have $\left|u - u_{a(u)}\right| \leq 1/n$. Observe that for any $u \in [0,1]$,

$$
\begin{aligned}
&\left|\sum_{k\neq i,j} (A_{ik} - p)\left(\psi'(u\theta_{r_i} + (1-u)\theta_{r_j} - \theta_{r_k})^\alpha - \psi'(u_{a(u)}\theta_{r_i} + (1-u_{a(u)})\theta_{r_j} - \theta_{r_k})^\alpha\right)\right| \\
&\leq \alpha \sup_{\xi\in[u\wedge u_{a(u)}, u\vee u_{a(u)}]} \sum_{k\neq i,j} \psi'(\xi\theta_{r_i} + (1-\xi)\theta_{r_j} - \theta_{r_k})^\alpha \left|u - u_{a(u)}\right| \left|\theta_{r_i} - \theta_{r_j}\right| \qquad \text{(S37)} \\
&\leq \frac{C_2' n\beta}{n} \frac{1}{\beta \vee 1/n} \leq C_2' \frac{p}{\beta \vee 1/n} \qquad \text{(S38)}
\end{aligned}
$$

for some constant $C_2' > 0$, where (S37) is due to mean value theorem and $|\psi''(x)| \leq \psi'(x)$ while (S38) comes from Lemma 8.1. Therefore,

$$
\begin{aligned}
&\sup_{u\in[0,1]} \sum_{k\neq i,j} A_{ik}\psi'(u\theta_{r_i} + (1-u)\theta_{r_j} - \theta_{r_k})^\alpha \\
&\leq \frac{C_3' p}{\beta \vee 1/n} + \max_{a\in[n]} \sum_{k\neq i,j} (A_{ik} - p)\psi'(u_a\theta_{r_i} + (1-u_a)\theta_{r_j} - \theta_{r_k})^\alpha \qquad \text{(S39)} \\
&\leq \frac{C_3' p}{\beta \vee 1/n} + C_4' \sqrt{p \log n \max_{a\in[n]} \sum_{k\neq i,j} \psi'(u_a\theta_{r_i} + (1-u_a)\theta_{r_j} - \theta_{r_k})^{2\alpha}} \qquad \text{(S40)} \\
&\leq \frac{C_5' p}{\beta \vee 1/n} \qquad \text{(S41)}
\end{aligned}
$$

for some constants $C_3', C_4', C_5' > 0$ with probability at least $1 - O(n^{-10})$, where (S39) is due to (S38) and $\frac{p}{\beta\vee 1/n} \gtrsim \log n \gg 1$. (S40) comes from Lemma 8.1, $|\psi'(x)| \leq 1/4$ and Lemma D.2. (S41) is a consequence of Lemma 8.1 and $\log n \lesssim \frac{p}{\beta\vee 1/n}$, which concludes the proof. □

Recall the definition of $G_{i,j,k,\theta,r}(u)$ in (35). We provide some properties of this term.

Lemma D.4. *Assume $1 \leq C_0 = O(1)$ and $0 < \beta = o(1)$. For any constant $C > 0$, there exist constants $C_1, C_2, C_3 > 0$ such that for any $\theta \in \Theta_n(\beta, C_0)$, any $r \in \mathfrak{S}_n$ and any $i \neq j \in [n]$ such that $\left|\theta_{r_i} - \theta_{r_j}\right| \leq C$, the following hold for n large enough,*

$$
\sup_{u\in[0,1]} \sup_{k\neq i,j} G_{i,j,k,\theta,r}(u) \leq C_1, \qquad \text{(S42)}
$$

$$\sup_{u\in[0,1]}\sum_{k\neq i,j} G_{i,j,k,\theta,r}(u)+G_{i,j,k,\theta,r}(1-u)\leq \sum_{k\neq i,j}\log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}, \tag{S43}$$

$$\sup_{u\in[0,1]}\sum_{k\neq i,j} G_{i,j,k,\theta,r}(u)^2+G_{i,j,k,\theta,r}(1-u)^2\leq C_2\frac{(\theta_{r_i}-\theta_{r_j})^4}{\beta\vee 1/n}, \tag{S44}$$

$$C_3\frac{\left|\theta_{r_i}-\theta_{r_j}\right|^2}{\beta\vee 1/n}\leq \sum_{k\neq i,j}\log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}\leq C_2\frac{\left|\theta_{r_i}-\theta_{r_j}\right|^2}{\beta\vee 1/n}. \tag{S45}$$

Proof. We first look at (S42). Note that

$$\begin{aligned}
G_{i,j,k,\theta,r}(u) &= \log\frac{\psi(u\theta_{r_i}+(1-u)\theta_{r_j}-\theta_{r_k})}{\psi(\theta_{r_i}-\theta_{r_k})^u\psi(\theta_{r_j}-\theta_{r_k})^{1-u}}\\
&\leq \log\frac{\psi(u\theta_{r_i}+(1-u)\theta_{r_j}-\theta_{r_k})}{\psi(\theta_{r_i}-\theta_{r_k})\wedge\psi(\theta_{r_j}-\theta_{r_k})}\\
&= \log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})}{e^{(1-u)(\theta_{r_i}-\theta_{r_j})}+e^{\theta_{r_i}-\theta_{r_k}}}\vee\log\frac{(1+e^{\theta_{r_j}-\theta_{r_k}})}{e^{-u(\theta_{r_i}-\theta_{r_j})}+e^{\theta_{r_j}-\theta_{r_k}}}\leq C
\end{aligned}$$

where the last inequality comes from $\left|\theta_{r_i}-\theta_{r_j}\right|\leq C$.

Now we look at (S43).

$$\begin{aligned}
&\sup_{u\in[0,1]}\sum_{k\neq i,j} G_{i,j,k,\theta,r}(u)+G_{i,j,k,\theta,r}(1-u)\\
&=\sup_{u\in[0,1]}\sum_{k\neq i,j}\log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{1+e^{u\theta_{r_i}+(1-u)\theta_{r_j}-\theta_{r_k}}+e^{(1-u)\theta_{r_i}+u\theta_{r_j}-\theta_{r_k}}+e^{\theta_{r_i}+\theta_{r_j}-2\theta_{r_k}}}\\
&\leq\sup_{u\in[0,1]}\sum_{k\neq i,j}\log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{1+2e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}+e^{\theta_{r_i}+\theta_{r_j}-2\theta_{r_k}}}=\sum_{k\neq i,j}\log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}.
\end{aligned}$$

To see (S44), we first note that

$$\begin{aligned}
&G_{i,j,k,\theta,r}(u)\\
&=u\log(1+e^{\theta_{r_i}-\theta_{r_k}})+(1-u)\log(1+e^{\theta_{r_j}-\theta_{r_k}})-\log(1+e^{u\theta_{r_i}+(1-u)\theta_{r_j}-\theta_{r_k}})\\
&\geq 0
\end{aligned}$$

by Jensen's inequality. Therefore,

$$\begin{aligned}
&\sup_{u\in[0,1]}\sum_{k\neq i,j} G_{i,j,k,\theta,r}(u)^2+G_{i,j,k,\theta,r}(1-u)^2\leq\sup_{u\in[0,1]}\sum_{k\neq i,j}(G_{i,j,k,\theta,r}(u)+G_{i,j,k,\theta,r}(1-u))^2\\
&\leq\sum_{k\neq i,j}\left[\log\frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}\right]^2 \qquad \text{(S46)}
\end{aligned}$$

where (S46) can be derived similarly as in the proof of (S43). To upper bound (S46), recall the definition of $R_\theta(\cdot,\cdot,\cdot)$ in (S31). We have that for any k such that $r_k \in R_\theta(\frac{\theta_{r_i}+\theta_{r_j}}{2}, t, t+1)$,

$$\log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} = \log\left(\frac{\cosh(\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}) + \cosh\frac{\theta_{r_i}-\theta_{r_j}}{2}}{\cosh(\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k})+1}\right)$$

$$\leq \frac{\cosh\frac{\theta_{r_i}-\theta_{r_j}}{2} - 1}{\cosh(\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k})+1} \leq \frac{C_1'(\theta_{r_i}-\theta_{r_j})^2}{e^t} \tag{S47}$$

for some constant $C_1' > 0$. (S47) can be seen from $\cosh x \leq 1 + C_2' x^2$ for some constant $C_2' > 0$ when $|x| \leq C/2$. and the fact that $t \leq |\frac{\theta_{r_i}+\theta_{r_j}}{2} - \theta_{r_k}| \leq t+1$. Therefore, using (S33) and (S47),

$$\sum_{k\neq i,j}\left[\log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}\right]^2$$

$$\leq \sum_{t\geq 0}\sum_{k:r_k\in R_\theta(\frac{\theta_{r_i}+\theta_{r_j}}{2},t,t+1)} \frac{C_1'^2(\theta_{r_i}-\theta_{r_j})^4}{e^{2t}} \leq \frac{C_3'(\theta_{r_i}-\theta_{r_j})^4}{\beta \vee 1/n}$$

for some constant $C_3' > 0$. The upper bound of (S45) can be proved similarly.

Finally, we turn to the lower bound of (S45). Note that we also have $\cosh x \geq 1 + C_4' x^2$ for some constant $C_4' > 0$ when $|x| \leq C/2$. Therefore, when $r_k \in R_\theta(\frac{\theta_{r_i}+\theta_{r_j}}{2}, 0, 1)$,

$$\log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} = \log\left(1 + \frac{\cosh\frac{\theta_{r_i}-\theta_{r_j}}{2} - 1}{\cosh(\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k})+1}\right)$$

$$\geq \frac{\cosh\frac{\theta_{r_i}-\theta_{r_j}}{2} - 1}{\cosh\frac{\theta_{r_i}-\theta_{r_j}}{2} + \cosh(\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k})} \geq C_5' \left|\theta_{r_i}-\theta_{r_j}\right|^2 \tag{S48}$$

for some constant $C_5' > 0$, where the first inequality is due to the fact that $\log(1+x) \geq x/(1+x)$ for any $x > -1$. Thus,

$$\sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} \geq \sum_{\substack{k:r_k\in R_\theta(\frac{\theta_{r_i}+\theta_{r_j}}{2},0,1)\\ k\neq i,j}} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}$$

$$\geq \left(\left|R_\theta(\frac{\theta_{r_i}+\theta_{r_j}}{2}, 0, 1)\right| - 2\right) C_5' \left|\theta_{r_i}-\theta_{r_j}\right|^2 \geq \frac{C_6'\left|\theta_{r_i}-\theta_{r_j}\right|^2}{\beta \vee 1/n} \tag{S49}$$

for some constant $C_6' > 0$, where (S49) is a result of (S32) and (S48). $\square$

Lemma D.5. *Assume* $\frac{p}{\log n(\beta \vee 1/n)} \to \infty$ *and* $1 \leq C_0 = O(1)$. *For any constant* $C_1 > 0$, *there exists* $\delta = o(1)$ *and constant* $C_2 > 0$, *such that for any* $\theta \in \Theta_n(\beta, C_0)$, *any* $r \in \mathfrak{S}_n$ *and any* $i \neq j \in [n]$ *such that* $\left|\theta_{r_i} - \theta_{r_j}\right| \leq C_1$, *the following holds with probability at least* $1 - O(n^{-10})$ *for* n *large enough,*

$$\sup_{u\in[0,1]} \sum_{k\neq i,j} A_{ik} G_{i,j,k,\theta,r}(u) + A_{jk} G_{i,j,k,\theta,r}(1-u) \leq (1+\delta) p \sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2}.$$

Proof. First we have

$$\psi(a-c) \wedge \psi(b-c) \leq \frac{1}{a-b} \log \frac{1+e^{a-c}}{1+e^{b-c}} \leq \psi(a-c) \vee \psi(b-c), \tag{S50}$$

for any $a, b, c \in \mathbb{R}$. To see why (S50) holds, let us study the function $f(\delta) = \log(1+\exp(x+\delta)) - \log(1+\exp(x)) - \delta \exp(x)/(1+\exp(x))$ for any x. Note that $f'(\delta) = \exp(x+\delta)/(1+\exp(x+\delta)) - \exp(x)/(1+\exp(x))$ is positive when $\delta > 0$ and negative when $\delta < 0$. Since $f(0) = 0$, we have $f(\delta) \geq 0$. As a result, we have $\exp(x)/(1+\exp(x)) \leq \delta^{-1} \log((1+\exp(x+\delta))/(1+\exp(x)))$ when $\delta > 0$ and the direction of the inequality is reversed when $\delta < 0$. WLOG, we assume $a - b > 0$. Then the first inequality of (S50) is proved by taking $x = b - c$ and $\delta = a - b$ and second one is proved by taking $x = a - c$ and $\delta = -(a-b)$.

Recall the definition of $G_{i,j,k,\theta,r}$ in (35). Then

$$\begin{aligned} \left|G'_{i,j,k,\theta,r}(u)\right| &= \left|\theta_{r_i} - \theta_{r_j}\right| \left| \frac{1}{\theta_{r_i} - \theta_{r_j}} \log \frac{1+e^{\theta_{r_i}-\theta_{r_k}}}{1+e^{\theta_{r_j}-\theta_{r_k}}} - \frac{e^{u(\theta_{r_i}-\theta_{r_j})+\theta_{r_j}-\theta_{r_k}}}{1+e^{u(\theta_{r_i}-\theta_{r_j})+\theta_{r_j}-\theta_{r_k}}} \right| \\ &\leq \left|\theta_{r_i} - \theta_{r_j}\right| \left|\psi(\theta_{r_i} - \theta_{r_k}) - \psi(\theta_{r_j} - \theta_{r_k})\right| \qquad \text{(S51)} \\ &\leq \left|\theta_{r_i} - \theta_{r_j}\right|^2. \qquad \text{(S52)} \end{aligned}$$

Here (S51) is due to the observation that both terms are in the interval $[\psi(\theta_{r_i} - \theta_{r_k}) \wedge \psi(\theta_{r_j} - \theta_{r_k}), \psi(\theta_{r_i} - \theta_{r_k}) \vee \psi(\theta_{r_j} - \theta_{r_k})]$ for any $u \in [0,1]$, where the first term is due to (S50) and the second term is due to the monotonicity of $\exp(x)/(1+\exp(x))$. Hence the difference between these two terms are bounded by $\left|\psi(\theta_{r_i} - \theta_{r_k}) - \psi(\theta_{r_j} - \theta_{r_k})\right|$ in absolute value. (S52) is due to $\psi'(x) \leq 1/4$. Following the line of discretization, let $u_a = \frac{a}{n}, a = 1, ..., n$. Then for any $u \in [0,1]$, let $a(u) = \arg\min_{a\in[n]} |u - u_a|$. We have $\left|u - u_{a(u)}\right| \leq 1/n$. Thus,

$$\begin{aligned} &\left| \sum_{k\neq i,j} (A_{ik} - p)(G_{i,j,k,\theta,r}(u) - G_{i,j,k,\theta,r}(u_{a(u)})) + (A_{jk} - p)(G_{i,j,k,\theta,r}(1-u) - G_{i,j,k,\theta,r}(1-u_{a(u)})) \right| \\ &\leq 2\left|\theta_{r_i} - \theta_{r_j}\right|^2 (n-2) \left|u - u_{a(u)}\right| \leq 2\left|\theta_{r_i} - \theta_{r_j}\right|^2. \qquad \text{(S53)} \end{aligned}$$

Then

$$
\begin{aligned}
&\sup_{u\in[0,1]} \sum_{k\neq i,j} A_{ik}G_{i,j,k,\theta,r}(u) + A_{jk}G_{i,j,k,\theta,r}(1-u) \\
&\leq p\sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} \qquad \text{(S54)}\\
&\quad + \sup_{u\in[0,1]} \sum_{k\neq i,j} (A_{ik}-p)G_{i,j,k,\theta,r}(u) + (A_{jk}-p)G_{i,j,k,\theta,r}(1-u) \\
&\leq p\sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} \\
&\quad + 2\left|\theta_{r_i}-\theta_{r_j}\right|^2 + \max_{a\in[n]} \sum_{k\neq i,j} (A_{ik}-p)G_{i,j,k,\theta,r}(u_a) + (A_{jk}-p)G_{i,j,k,\theta,r}(1-u_a) \qquad \text{(S55)}\\
&\leq p\sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} + 2\left|\theta_{r_i}-\theta_{r_j}\right|^2 \\
&\quad + C_1'\sqrt{p\log n \max_{a\in[n]} \sum_{k\neq i,j} G_{i,j,k,\theta,r}(u_a)^2 + G_{i,j,k,\theta,r}(1-u_a)^2} \qquad \text{(S56)}\\
&\leq p\sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} + 2\left|\theta_{r_i}-\theta_{r_j}\right|^2 + C_2'\left|\theta_{r_i}-\theta_{r_j}\right|^2\sqrt{\frac{p\log n}{\beta\vee 1/n}} \qquad \text{(S57)}\\
&= (1+\delta)p\sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} \qquad \text{(S58)}
\end{aligned}
$$

with probability at least $1-O(n^{-10})$ for some constants $C_1', C_2' > 0$ and $\delta = o(1)$. (S54) is due to Lemma D.4. (S55) comes from (S53). (S56) and (S57) are a consequence of Lemma D.2 and Lemma D.4. (S58) is because of Lemma D.4 and

$$
p\frac{1}{\left|\theta_{r_i}-\theta_{r_j}\right|^2} \sum_{k\neq i,j} \log \frac{(1+e^{\theta_{r_i}-\theta_{r_k}})(1+e^{\theta_{r_j}-\theta_{r_k}})}{\left(1+e^{\frac{\theta_{r_i}+\theta_{r_j}}{2}-\theta_{r_k}}\right)^2} \gtrsim \frac{p}{\beta\vee 1/n} \gg \sqrt{\frac{p\log n}{\beta\vee 1/n}} \gg 1,
$$

which concludes the proof.

□

Now we are ready to prove the Lemma of 8.2.

Proof of Lemma of 8.2. By Neyman-Pearson Lemma, the optimal procedure is the likelihood ratio test:

$$
\begin{aligned}
&\inf_{\widehat{r}} \frac{\mathbb{P}_{(\theta^*,r^*)}\left(\widehat{r}\neq r^*\right)+\mathbb{P}_{(\theta^*,r^{*(i,j)})}\left(\widehat{r}\neq r^{*(i,j)}\right)}{2}\\
&=\frac{\mathbb{P}_{(\theta^*,r^*)}\left(\ell_n(\theta^*,r^*)\geq \ell_n(\theta^*,r^{*(i,j)})\right)+\mathbb{P}_{(\theta^*,r^{*(i,j)})}\left(\ell_n(\theta^*,r^*)\leq \ell_n(\theta^*,r^{*(i,j)})\right)}{2}
\end{aligned}
$$

We only need to lower bound $\mathbb{P}_{(\theta^*,r^*)}\left(\ell_n(\theta^*,r^*)\geq \ell_n(\theta^*,r^{*(i,j)})\right)$ and the other term can be bounded similarly. WLOG, assume $i<j$ and $r_i^*=a<r_j^*=b$. Let

$$
Z_{kl}=y_{ikl}\log\frac{\psi(\theta_b^*-\theta_{r_k^*}^*)}{\psi(\theta_a^*-\theta_{r_k^*}^*)}+(1-y_{ikl})\log\frac{1-\psi(\theta_b^*-\theta_{r_k^*}^*)}{1-\psi(\theta_a^*-\theta_{r_k^*}^*)}, k\neq i,j,
$$

$$
\bar{Z}_{kl}=y_{jkl}\log\frac{\psi(\theta_a^*-\theta_{r_k^*}^*)}{\psi(\theta_b^*-\theta_{r_k^*}^*)}+(1-y_{jkl})\log\frac{1-\psi(\theta_a^*-\theta_{r_k^*}^*)}{1-\psi(\theta_b^*-\theta_{r_k^*}^*)}, k\neq i,j
$$

and

$$
Z_{0l}=y_{ijl}\log\frac{\psi(\theta_b^*-\theta_a^*)}{\psi(\theta_a^*-\theta_b^*)}+(1-y_{ijl})\log\frac{1-\psi(\theta_b-\theta_a^*)}{1-\psi(\theta_a^*-\theta_b^*)}.
$$

To simplify notation, we use $\mathbb{P}_A(\cdot)$ as $\mathbb{P}_{(\theta^*,r^*)}(\cdot|A)$ and $\mathbb{E}_A[\cdot]$ as $\mathbb{E}_{(\theta^*,r^*)}[\cdot|A]$. Then

$$
\mathbb{P}_A\left(\ell_n(\theta^*,r^*)\geq \ell_n(\theta^*,r^{*(i,j)})\right)=\mathbb{P}_A\left(\sum_{l=1}^L\left(A_{ij}Z_{0l}+\sum_{k\neq i,j}A_{ik}Z_{kl}+A_{jk}\bar{Z}_{kl}\right)\geq 0\right).
$$

Let $\mu_{i'j'}=\psi(\theta_{r_{i'}^*}^*-\theta_{r_{j'}^*}^*)$ for any $i'\neq j'$. Define

$$
\begin{aligned}
\nu_{r^*}(u)&=\log\mathbb{E}_A\left\{\exp\left[u\left(A_{ij}Z_{01}+\sum_{k\neq i,j}A_{ik}Z_{k1}+A_{jk}\bar{Z}_{k1}\right)\right]\right\}\\
&=A_{ij}\nu_{0,r^*}(u)+\sum_{k\neq i,j}A_{ik}\nu_{k,r^*}(u)+\sum_{k\neq i,j}A_{jk}\bar{\nu}_{k,r^*}(u)
\end{aligned}
$$

where

$$
\nu_{0,r^*}(u)=\log\left[\mu_{ij}^u(1-\mu_{ij})^{1-u}+\mu_{ij}^{1-u}(1-\mu_{ij})^u\right]=-\log\frac{1+e^{\theta_a^*-\theta_b^*}}{e^{u(\theta_a^*-\theta_b^*)}+e^{(1-u)(\theta_a^*-\theta_b^*)}}
$$

$$
\nu_{k,r^*}(u)=\log\left[\mu_{jk}^u\mu_{ik}^{1-u}+(1-\mu_{jk})^u(1-\mu_{ik})^{1-u}\right]=-G_{i,j,k,\theta^*,r^*}(1-u)
$$

$$
\bar{\nu}_{k,r^*}(u)=\log\left[\mu_{ik}^u\mu_{jk}^{1-u}+(1-\mu_{ik})^u(1-\mu_{jk})^{1-u}\right]=-G_{i,j,k,\theta^*,r^*}(u).
$$

$\nu_{r^*}(u)$ is the conditional cumulant generating function of $A_{ij}Z_{01}+\sum_{k\neq i,j}A_{ik}Z_{kl}+A_{jk}\bar{Z}_{kl}$. We also have $\nu_{0,r^*}(u),\nu_{k,r^*}(u),\bar{\nu}_{k,r^*}(u)$ as the cumulant generating functions of $Z_{01},Z_{k1},\bar{Z}_{k1}$ respectively. Define

$$
u_{r^*}^*=\arg\min_{u\geq 0}\nu_{r^*}(u).
$$

Since cumulant generating functions are convex and $\nu_{r^*}(0) = \nu_{r^*}(1) = 0$, it can be seen easily that $u^*_{r^*} \in (0,1)$ and depends on A. Following the change-of-measure argument in the proof of Lemma A.4 and Lemma B.3 of [2], we have

$$\begin{aligned}
&\mathbb{P}_A\left(\sum_{l=1}^{L}\left(A_{ij}Z_{0l} + \sum_{k\neq i,j} A_{ik}Z_{kl} + A_{jk}\bar{Z}_{kl}\right) \geq 0\right) \\
&\geq \exp\left(-u^*_{r^*}T + L\nu_{r^*}(u^*_{r^*})\right)\mathbb{Q}_A\left(0 \leq \sum_{l=1}^{L}\left(A_{ij}Z_{0l} + \sum_{k\neq i,j} A_{ik}Z_{kl} + A_{jk}\bar{Z}_{kl}\right) \leq T\right) \quad \text{(S59)}
\end{aligned}$$

for any T in (S59) to be determined later and $\mathbb{Q}_A$ is a measure under which $Z_{0l}, Z_{kl}, \bar{Z}_{kl}, l \in [L], k \neq i,j$ are all independent given A and follow

$$\mathbb{Q}_A(Z_{0l} = s) = e^{u^*_{r^*}s - \nu_{0,r^*}(u^*_{r^*})}\mathbb{P}_A\left(Z_{0l} = s\right),$$

$$\begin{aligned}
\mathbb{Q}_A(Z_{kl} = s) &= e^{u^*_{r^*}s - \nu_{k,r^*}(u^*_{r^*})}\mathbb{P}_A\left(Z_{kl} = s\right), k \neq i,j, k \in [n], \\
\mathbb{Q}_A(\bar{Z}_{kl} = s) &= e^{u^*_{r^*}s - \bar{\nu}_{k,r^*}(u^*_{r^*})}\mathbb{P}_A\left(\bar{Z}_{kl} = s\right), k \neq i,j, k \in [n].
\end{aligned}$$

Furthermore, by definition of $u^*_{r^*}$, the expectation of $A_{ij}Z_{0l} + \sum_{k\neq i,j} A_{ik}Z_{kl} + A_{jk}\bar{Z}_{kl}$ under $\mathbb{Q}_A$ is 0.

We can compute the 2nd and 4th moments under Q_A, denoted as $\mathsf{Var}_{\mathbb{Q}_A}(\cdot)$ and $\kappa_{\mathbb{Q}_A}(\cdot)$ respectively:

$$\begin{aligned}
\mathsf{Var}_{\mathbb{Q}_A}(Z_{0l}) &= \nu''_{0,r^*}(u^*_{r^*}) = 4\mu_{ij}(1-\mu_{ij})\frac{(\theta^*_a - \theta^*_b)^2 e^{2u^*_{r^*}(\theta^*_a - \theta^*_b)}}{((1-\mu_{ij})e^{2u^*_{r^*}(\theta^*_a - \theta^*_b)} + \mu_{ij})^2} \\
&= 4(\theta^*_a - \theta^*_b)^2\psi'\left((1-2u^*_{r^*})(\theta^*_a - \theta^*_b)\right), \quad \text{(S60)}
\end{aligned}$$

$$\begin{aligned}
\mathsf{Var}_{\mathbb{Q}_A}(Z_{kl}) &= \nu''_{k,r^*}(u^*_{r^*}) = \mu_{ik}(1-\mu_{ik})\frac{(\theta^*_a - \theta^*_b)^2 e^{u^*_{r^*}(\theta^*_a - \theta^*_b)}}{((1-\mu_{ik})e^{u^*_{r^*}(\theta^*_a - \theta^*_b)} + \mu_{ik})^2} \\
&= (\theta^*_a - \theta^*_b)^2\psi'\left((1-u^*_{r^*})\theta^*_a + u^*_{r^*}\theta^*_b - \theta^*_{r^*_k}\right), k \neq i,j, k \in [n], \quad \text{(S61)}
\end{aligned}$$

$$\begin{aligned}
\mathsf{Var}_{\mathbb{Q}_A}(\bar{Z}_{kl}) &= \bar{\nu}''_{k,r^*}(u^*_{r^*}) = \mu_{jk}(1-\mu_{jk})\frac{(\theta^*_a - \theta^*_b)^2 e^{-u^*_{r^*}(\theta^*_a - \theta^*_b)}}{((1-\mu_{jk})e^{-u^*_{r^*}(\theta^*_a - \theta^*_b)} + \mu_{jk})^2} \\
&= (\theta^*_a - \theta^*_b)^2\psi'\left(u^*_{r^*}\theta^*_a + (1-u^*_{r^*})\theta^*_b - \theta^*_{r^*_k}\right), k \neq i,j, k \in [n] \quad \text{(S62)}
\end{aligned}$$

and

$$\begin{aligned}
&\kappa_{\mathbb{Q}_A}(Z_{0l}) = \mathbb{Q}_A\left((Z_{0l} - \mathbb{Q}_A(Z_{0l}))^4\right) = \nu''''_{0,r^*}(u^*_{r^*}) + 3\nu''_{0,r^*}(u^*_{r^*})^2 \\
&\leq 16\mu_{ij}(1-\mu_{ij})\frac{(\theta^*_a - \theta^*_b)^4 e^{2u^*_{r^*}(\theta^*_a - \theta^*_b)}}{[(1-\mu_{ij})e^{2u^*_{r^*}(\theta^*_a - \theta^*_b)} + \mu_{ij}]^2} + 3\nu''_{0,r^*}(u^*_{r^*})^2 \\
&= 4(\theta^*_a - \theta^*_b)^2\nu''_{0,r^*}(u^*_{r^*}) + 3\nu''_{0,r^*}(u^*_{r^*})^2 \leq 7(\theta^*_a - \theta^*_b)^4\psi'\left((1-2u^*_{r^*})(\theta^*_a - \theta^*_b)\right), \quad \text{(S63)}
\end{aligned}$$

$$\kappa_{\mathbb{Q}_A}(Z_{kl}) = \mathbb{Q}_A\left((Z_{kl} - \mathbb{Q}_A(Z_{kl}))^4\right) \leq (\theta_a^* - \theta_b^*)^2 \nu''_{k,r^*}(u^*_{r^*}) + 3\nu''_{k,r^*}(u^*_{r^*})^2$$
$$\leq 4(\theta_a^* - \theta_b^*)^4 \psi'\left((1-u^*_{r^*})\theta_a^* + u^*_{r^*}\theta_b^* - \theta^*_{r^*_k}\right), k \neq i,j, k \in [n] \tag{S64}$$

$$\kappa_{\mathbb{Q}_A}(\bar{Z}_{kl}) = \mathbb{Q}_A\left((\bar{Z}_{kl} - \mathbb{Q}_A(\bar{Z}_{kl}))^4\right) \leq (\theta_a^* - \theta_b^*)^2 \bar{\nu}''_{k,r^*}(u^*_{r^*}) + 3\bar{\nu}''_{k,r^*}(u^*_{r^*})^2$$
$$\leq 4(\theta_a^* - \theta_b^*)^4 \psi'\left(u^*_{r^*}\theta_a^* + (1-u^*_{r^*})\theta_b^* - \theta^*_{r^*_k}\right), k \neq i,j, k \in [n]. \tag{S65}$$

Let $\mathcal{F}_1$ be the event on which the following holds:

$$\inf_{u\in[0,1]} \sum_{k\neq i,j} A_{ik}\psi'((1-u)\theta_a^* + u\theta_b^* - \theta^*_{r^*_k}) + A_{jk}\psi'(u\theta_a^* + (1-u)\theta_b^* - \theta^*_{r^*_k}) \geq C_1' \frac{p}{\beta \vee 1/n},$$

$$\sup_{u\in[0,1]} \sum_{k\neq i,j} A_{ik}\psi'((1-u)\theta_a^* + u\theta_b^* - \theta^*_{r^*_k})^{3/4} + A_{jk}\psi'(u\theta_a^* + (1-u)\theta_b^* - \theta^*_{r^*_k})^{3/4} \leq C_2' \frac{p}{\beta \vee 1/n}$$

for some constants $C_1', C_2' > 0$. We shall choose C_1', C_2' to make $\mathcal{F}_1$ happen with probability at least $1 - O(n^{-10})$ by Lemma D.3. Therefore, we shall choose T as

$$T = \sqrt{L\left(A_{ij}\mathsf{Var}_{\mathbb{Q}_A}(Z_{01}) + \sum_{k\neq i,j} A_{ik}\mathsf{Var}_{\mathbb{Q}_A}(Z_{k1}) + A_{jk}\mathsf{Var}_{\mathbb{Q}_A}(\bar{Z}_{k1})\right)}$$
$$\leq \sqrt{C_3' L \frac{p(\theta_a^* - \theta_b^*)^2}{\beta \vee 1/n}}$$

on $\mathcal{F}_1$ for some constant $C_3' > 0$ using (S60)-(S62). With this choice of T, the $\mathbb{Q}_A$ measure can be lower bounded by some constant $C_4' > 0$ on $\mathcal{F}_1$. This can be seen by bounding the 4th moment approximation bound using Lemma D.1 :

$$\sqrt{L \frac{A_{ij}\kappa_{\mathbb{Q}_A}(Z_{01})^{3/4} + \sum_{k\neq i,j} A_{ik}\kappa_{\mathbb{Q}_A}(Z_{kl})^{3/4} + A_{jk}\kappa_{\mathbb{Q}_A}(\bar{Z}_{kl})^{3/4}}{\left(LA_{ij}\mathsf{Var}_{\mathbb{Q}_A}(Z_{01}) + L\sum_{k\neq i,j} A_{ik}\mathsf{Var}_{\mathbb{Q}_A}(Z_{k1}) + A_{jk}\mathsf{Var}_{\mathbb{Q}_A}(\bar{Z}_{k1})\right)^{3/2}}}$$
$$\leq \sqrt{C_5' L \frac{\sum_{k\neq i,j} A_{ik}\psi'\left((1-u^*_{r^*})\theta_a^* + u^*_{r^*}\theta_b^* - \theta^*_{r^*_k}\right)^{3/4} + A_{jk}\psi'\left(u^*_{r^*}\theta_a^* + (1-u^*_{r^*})\theta_b^* - \theta^*_{r^*_k}\right)^{3/4}}{\left(L\sum_{k\neq i,j} A_{ik}\psi'((1-u^*_{r^*})\theta_a^* + u^*_{r^*}\theta_b^* - \theta^*_{r^*_k}) + A_{jk}\psi'(u^*_{r^*}\theta_a^* + (1-u^*_{r^*})\theta_b^* - \theta^*_{r^*_k})\right)^{3/2}}} \tag{S66}$$

$$\leq C_6' \left(L\frac{p}{\beta \vee 1/n}\right)^{-1/4} \tag{S67}$$

on $\mathcal{F}_1$ for some constants $C_5', C_6' > 0$ and this bound tends to 0. (S66) is due to (S60)-(S62) and (S63)-(S65). (S67) is a consequence of Lemma D.3.

Now we turn to $L\nu_{r^*}(u^*_{r^*})$. Let $\mathcal{F}_2$ be the event on which the following holds:

$$\sup_{u\in[0,1]} \sum_{k\neq i,j} (A_{ik}G_{i,j,k,\theta^*,r^*}(1-u) + A_{jk}G_{i,j,k,\theta^*,r^*}(u))$$
$$\leq (1+\delta_1')2p \sum_{k\neq i,j} G_{i,j,k,\theta^*,r^*}(1/2).$$

By Lemma D.5, there exists $\delta_1' = o(1)$ independent of i, j, θ^*, r^* such that $\mathcal{F}_2$ holds with probability at least $1 - O(n^{-10})$. Then, on this event,

$$\nu_{r^*}(u_{r^*}) \geq -\sup_{u\in[0,1]}\left(-A_{ij}\nu_{0,r^*}(u) - \sum_{k\neq i,j}(A_{ik}\nu_{k,r^*}(u) + A_{jk}\bar{\nu}_{k,r^*}(u))\right)$$

$$\geq -A_{ij}\sup_{u\in[0,1]}\log\frac{1+e^{\theta_a^*-\theta_b^*}}{e^{u(\theta_a^*-\theta_b^*)}+e^{(1-u)(\theta_a^*-\theta_b^*)}} - \sup_{u\in[0,1]}\sum_{k\neq i,j}(A_{ik}G_{i,j,k,\theta^*,r^*}(1-u) + A_{jk}G_{i,j,k,\theta^*,r^*}(u))$$

$$\geq -C_7'\left|\theta_a^*-\theta_b^*\right|^2 - (1+\delta_1')2p\sum_{k\neq i,j}G_{i,j,k,\theta^*,r^*}(1/2) \tag{S68}$$

$$\geq -(1+\delta_2')2p\sum_{k\neq i,j}G_{i,j,k,\theta^*,r^*}(1/2) \tag{S69}$$

for some $\delta_2' = o(1)$. (S68) comes from

$$\log\frac{1+e^{\theta_a^*-\theta_b^*}}{e^{u(\theta_a^*-\theta_b^*)}+e^{(1-u)(\theta_a^*-\theta_b^*)}} \leq \log\cosh\frac{\theta_a^*-\theta_b^*}{2} \leq \cosh\frac{\theta_a^*-\theta_b^*}{2} - 1 \leq C_7'\left|\theta_a^*-\theta_b^*\right|^2$$

for some constant $C_7' > 0$ when $|\theta_a^* - \theta_b^*| \leq C$. (S69) is because of Lemma D.4 and $\frac{p}{\beta\vee 1/n} \gg 1$. Note that δ_2' can also be chosen independent of i, j, θ^*, r^*.

Thus, we can further lower bound (S59) on $\mathcal{F}_1 \cap \mathcal{F}_2$:

$$\mathbb{P}_A\left(\sum_{l=1}^{L}\left(A_{ij}Z_{0l} + \sum_{k\neq i,j}A_{ik}Z_{kl} + A_{jk}\bar{Z}_{kl}\right) \geq 0\right)$$

$$\geq C_4'\exp\left(-\sqrt{C_3'Lp\frac{(\theta_a^*-\theta_b^*)^2}{\beta\vee 1/n}} + L\nu_{r^*}(u_{r^*}^*)\right)$$

$$\geq C_4'\exp\left(-\sqrt{C_3'Lp\frac{(\theta_a^*-\theta_b^*)^2}{\beta\vee 1/n}} - L\sup_{u\in[0,1]}\left(-A_{ij}\nu_{0,r^*}(u) - \sum_{k\neq i,j}(A_{ik}\nu_{k,r^*}(u) + A_{jk}\bar{\nu}_{k,r^*}(u))\right)\right)$$

$$\geq C_4'\exp\left(-\sqrt{C_3'Lp\frac{(\theta_a^*-\theta_b^*)^2}{\beta\vee 1/n}} - (1+\delta_2')2Lp\sum_{k\neq i,j}G_{i,j,k,\theta^*,r^*}(1/2)\right).$$

which finishes the proof. □

References

[1] Pinhan Chen, Chao Gao, and Anderson Y Zhang. Partial recovery for top-k ranking: Optimality of mle and sub-optimality of spectral method. *The Annals of Statistics (to appear)*.

[2] Pinhan Chen, Chao Gao, and Anderson Y. Zhang. Supplement to "partial recovery for top-k ranking:
optimality of mle and sub-optimality of spectral method".

[3] Jing Lei and Alessandro Rinaldo. Consistency of spectral clustering in stochastic block models. *The Annals of Statistics*, 43(1):215–237, 2015.

[4] Sheldon M Ross and Erol A Peköz. *A second course in probability*. www. Probability-Bookstore. com, 2007.

[5] Joel A Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015.