# Supplementary Material

### Proof of Lemma 3

*Proof.*

*Part 1.* Let $m \geq \frac{4}{\mu_{\min}}$. Recall the definitions of $\widehat{P}_{ij}$, $N_{ij}$, $N_{ij}^{(1)}$ from Eq. (3). In the following, we will make the dependence of these quantities on the training sample explicit; specifically, for any $\omega \in (\mathcal{X} \times \{0,1\})^m$, we will write the corresponding quantities as $\widehat{P}_{ij}(\omega)$, $N_{ij}(\omega)$, and $N_{ij}^{(1)}(\omega)$, respectively.

Clearly, for any $\omega, \omega' \in (\mathcal{X} \times \{0,1\})^m$, since $\widehat{P}_{ij}(\omega), \widehat{P}_{ij}(\omega') \in [0,1]$, we have

$$|\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| \leq 1\,.$$

We will prove the result for the case $i < j$; the case $i > j$ can be proved similarly. Assume $i < j$, and let $B_{ij}$ be the following 'bad' event:

$$B_{ij} = \left\{\omega \in (\mathcal{X} \times \{0,1\})^m : N_{ij}(\omega) \leq \tfrac{m\mu_{ij}}{2}\right\}.$$

Then by a straightforward application of Hoeffding's inequality, we have

$$\mathbf{P}(S \in B_{ij}) \;\leq\; \exp(-m\mu_{ij}^2/2) \;\leq\; \exp(-m\mu_{\min}^2/2)\,.$$

Now consider $\omega, \omega' \in (\mathcal{X} \times \{0,1\})^m$ such that $\omega \notin B_{ij}$, and $\omega, \omega'$ differ only in one element. We can have the following cases:

(1) $N_{ij}(\omega') = N_{ij}(\omega)$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega)$

(2) $N_{ij}(\omega') = N_{ij}(\omega)$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) + 1$

(3) $N_{ij}(\omega') = N_{ij}(\omega)$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) - 1$

(4) $N_{ij}(\omega') = N_{ij}(\omega) + 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) + 1$

(5) $N_{ij}(\omega') = N_{ij}(\omega) + 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega)$

(6) $N_{ij}(\omega') = N_{ij}(\omega) - 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) - 1$

(7) $N_{ij}(\omega') = N_{ij}(\omega) - 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega)$

We will consider each of these cases separately, and will show that in each case, the difference $|\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')|$ is upper bounded by $\frac{2}{m\mu_{\min}}$.

- Case (1): $N_{ij}(\omega') = N_{ij}(\omega)$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega)$

  In this case nothing changes with respect to the pair $(i,j)$ and hence

$$|\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| = 0$$

- Case (2): $N_{ij}(\omega') = N_{ij}(\omega)$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) + 1$

  In this case we have

$$\begin{aligned}
|\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| &= \left|\frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)} - \frac{N_{ij}^{(1)}(\omega)+1}{N_{ij}(\omega)}\right| \\
&= \frac{1}{N_{ij}(\omega)} \\
&\leq \frac{2}{m\mu_{ij}} \leq \frac{2}{m\mu_{\min}}
\end{aligned}$$

- Case (3): $N_{ij}(\omega') = N_{ij}(\omega)$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) - 1$

  In this case we have

$$\begin{aligned} |\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| &= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)} - \frac{N_{ij}^{(1)}(\omega) - 1}{N_{ij}(\omega)} \right| \\ &= \frac{1}{N_{ij}(\omega)} \leq \frac{2}{m\mu_{ij}} \leq \frac{2}{m\mu_{\min}} \end{aligned}$$

- Case (4): $N_{ij}(\omega') = N_{ij}(\omega) + 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{1}(\omega) + 1$

  In this case we have

$$\begin{aligned} |\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| &= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)} - \frac{N_{ij}^{(1)}(\omega) + 1}{N_{ij}(\omega) + 1} \right| \\ &= \left| \frac{N_{ij}(\omega) - N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)(N_{ij}(\omega) + 1)} \right| \leq \left| \frac{N_{ij}(\omega)}{N_{ij}(\omega)(N_{ij}(\omega) + 1)} \right| \\ &\leq \frac{1}{N_{ij}(\omega)} \leq \frac{2}{m\mu_{ij}} \leq \frac{2}{m\mu_{\min}} \end{aligned}$$

- Case (5): $N_{ij}(\omega') = N_{ij}(\omega) + 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega)$

  In this case we have

$$\begin{aligned} |\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| &= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)} - \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega) + 1} \right| \\ &= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)(N_{ij}(\omega) + 1)} \right| \leq \left| \frac{N_{ij}(\omega)}{N_{ij}(\omega)(N_{ij}(\omega) + 1)} \right| \\ &\leq \frac{1}{N_{ij}(\omega)} \leq \frac{2}{m\mu_{ij}} \leq \frac{2}{m\mu_{\min}} \end{aligned}$$

- Case (6): $N_{ij}(\omega') = N_{ij}(\omega) - 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega) - 1$

  In this case we have

$$\begin{aligned} |\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| &= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)} - \frac{N_{ij}^{(1)}(\omega) - 1}{N_{ij}(\omega) - 1} \right| \\ &= \left| \frac{N_{ij}(\omega) - N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)(N_{ij}(\omega) - 1)} \right| \\ &\leq \frac{1}{N_{ij}(\omega)} \leq \frac{2}{m\mu_{ij}} \leq \frac{2}{m\mu_{\min}} \end{aligned}$$

  Note that in this case $N_{ij}(\omega) - 1$ cannot equal 0 because $m \geq \frac{4}{\mu_{\min}}$ which guarantees that for $\omega \notin B_{ij}$, $N_{ij}(\omega) \geq 2$. Also the final step follows because this case can happen only when $N_{ij}^{(1)}(\omega) \geq 1$ and so we can upper bound $\frac{(N_{ij}(\omega) - N_{ij}^{(1)}(\omega))}{(N_{ij}(\omega) - 1)}$ by 1

- Case (7): $N_{ij}(\omega') = N_{ij}(\omega) - 1$ and $N_{ij}^{(1)}(\omega') = N_{ij}^{(1)}(\omega)$

In this case we have

$$
\begin{aligned}
|\widehat{P}_{ij}(\omega) - \widehat{P}_{ij}(\omega')| &= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)} - \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)-1} \right| \\
&= \left| \frac{N_{ij}^{(1)}(\omega)}{N_{ij}(\omega)(N_{ij}(\omega)-1)} \right| \\
&\leq \frac{1}{N_{ij}(\omega)} \leq \frac{2}{m\mu_{ij}} \leq \frac{2}{m\mu_{\min}}
\end{aligned}
$$

Again, $N_{ij}(\omega) - 1$ cannot equal 0 because $m \geq \frac{4}{\mu_{\min}}$ which guarantees that for $\omega \notin B_{ij}$, $N_{ij}(\omega) \geq 2$. Also note that this case can occur only when $N_{ij}^{(1)}(\omega) \leq N_{ij}(\omega) - 1$ which is used to upper bound $\frac{N_{ij}^{(1)}}{N_{ij}(\omega)-1}$ by 1.

Thus we have the required bound in all possible cases.

*Part 2.* This follows directly from Part 1 and Theorem 2.

*Part 3.* Let $m \geq \frac{1}{\mu_{\min}} \ln\left(\frac{1}{\epsilon}\right)$. We have,

$$
\mathbf{E}[\widehat{P}_{ij}] \;=\; P_{ij}(1-(1-\mu_{ij})^m)\,.
$$

This gives

$$
\left|\mathbf{E}[\widehat{P}_{ij}] - P_{ij}\right| \;=\; P_{ij}(1-\mu_{ij})^m \;\leq\; (1-\mu_{\min})^m \;\leq\; e^{-m\mu_{\min}} \;\leq\; \epsilon\,,
$$

where the last inequality follows from the given condition on $m$.

*Part 4.* Let $m$ satisfy the given condition. Then

$$
\begin{aligned}
\mathbf{P}\Big(\big|\widehat{P}_{ij} - P_{ij}\big| \geq \epsilon\Big) \;&\leq\; \mathbf{P}\Big(\big|\widehat{P}_{ij} - \mathbf{E}[\widehat{P}_{ij}]\big| + \big|\mathbf{E}[\widehat{P}_{ij}] - P_{ij}\big| \geq \epsilon\Big)\,, \quad \text{by triangle inequality} \\
&\leq\; \mathbf{P}\Big(\big|\widehat{P}_{ij} - \mathbf{E}[\widehat{P}_{ij}]\big| \geq \frac{\epsilon}{2}\Big)\,, \quad \text{by Part 3, since } m \geq \tfrac{1}{\mu_{\min}} \ln\left(\tfrac{2}{\epsilon}\right) \\
&\leq\; 4\exp\left(\frac{-m\epsilon^2\mu_{\min}^2}{128}\right)\,, \quad \text{by Part 2.}
\end{aligned}
$$

*Part 5.* Let $m \geq \frac{1}{\mu_{\min}P_{\min}} \ln\left(\frac{n(n-1)}{\delta}\right)$. Then

$$
\begin{aligned}
\mathbf{P}\Big(\exists (i \neq j) : \widehat{P}_{ij} = 0\Big) \;&\leq\; \sum_{i=1}^{n}\sum_{j\neq i} \mathbf{P}\big(\widehat{P}_{ij} = 0\big)\,, \quad \text{by union bound} \\
&=\; \sum_{i=1}^{n}\sum_{j\neq i} (1-\mu_{ij}P_{ij})^m \\
&\leq\; n(n-1)\big(1-\mu_{\min}P_{\min}\big)^m \\
&\leq\; n(n-1)\,e^{-m\mu_{\min}P_{\min}} \\
&\leq\; \delta\,,
\end{aligned}
$$

where the last inequality follows from the given condition on $m$.

This completes the proof of the lemma. □

**Proof of Lemma 6**

*Proof.* We will first show the forward direction. Assume that the preference matrix $\mathbf{P}$ satisfies the time-reversibility condition. Let $\mathbf{Q}$ be the time-reversible Markov chain corresponding to $\mathbf{P}$, with stationary distribution $\boldsymbol{\pi}$; since $\mathbf{Q}$ is irreducible and aperiodic, we have $\pi_i > 0\ \forall i$. Now let $i \neq j$. By time reversibility and definition of $Q_{ij}$,

$$\pi_i P_{ij} = \pi_j P_{ji}\,.$$

We also have

$$P_{ji} = 1 - P_{ij}\,.$$

Solving for $P_{ij}$, this gives

$$P_{ij} = \frac{\pi_j}{\pi_i + \pi_j}\,.$$

Thus $\mathbf{P}$ satisfies the BTL condition with vector $\mathbf{w} = \boldsymbol{\pi} \in \mathbb{R}^n_+$. This proves the forward direction.

To show the reverse direction, assume that the preference matrix $\mathbf{P}$ satisfies the BTL condition with vector $\mathbf{w} \in \mathbb{R}^n_+$, so that $w_i > 0\ \forall i$ and $P_{ij} = \frac{w_j}{w_i + w_j}\ \forall i \neq j$. Let $\mathbf{Q}$ be the Markov chain constructed from $\mathbf{P}$ as in Eq. (6). Then it is easy to see that the vector $\boldsymbol{\pi}$ given by $\pi_i = \frac{w_i}{\sum_{k=1}^n w_k}$ satisfies

$$\pi_i Q_{ij} = \pi_j Q_{ji}\ \ \forall i, j \in [n]\,,$$

from which it follows that $\boldsymbol{\pi}$ is also the stationary probability vector of $\mathbf{Q}$. Therefore $\mathbf{P}$ satisfies the time-reversibility condition, thus proving the reverse direction. □

**Proof of Theorem 7**

The proof of Theorem 7 builds on techniques of (Negahban et al., 2012). We first state below four lemmas that are used in the proof: two of these are due to Negahban et al. (Negahban et al., 2012); proofs for the remaining two are included below. The statements of the lemmas and corresponding proofs require some additional notation as summarized below:

**Additional notation.** In what follows, for a matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$, we will denote by $\|\mathbf{Q}\|_F = \big(\sum_{i=1}^n \sum_{j=1}^n Q_{ij}^2\big)^{1/2}$ the Frobenius norm of $\mathbf{Q}$, by $\|\mathbf{Q}\|_2 = \max_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Q}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$ the spectral norm of $\mathbf{Q}$, and by $\lambda_{(2)}(\mathbf{Q})$ the second-largest eigenvalue of $\mathbf{Q}$ in absolute value.

**Lemma 21.** *Let $(\mu, \mathbf{P})$ be such that $\mu_{\min} > 0$. Let $\mathbf{Q}$ be defined as in Eq. (6). Let $0 < \epsilon \leq 8$ and $\delta \in (0, 1]$. If*

$$m \ \geq\ \max\left(\frac{256\,n}{\epsilon^2\mu_{\min}^2}\ln\Big(\frac{8n^2}{\delta}\Big)\,,\ B(\mu_{\min})\right)\,,$$

*then with probability at least $1-\delta$ (over the random draw of $S \sim (\mu, \mathbf{P})^m$ from which $\widehat{\mathbf{P}}$ is constructed), the empirical Markov chain $\widehat{\mathbf{Q}}$ constructed by the rank centrality algorithm satisfies*

$$\|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \leq \epsilon\,.$$

*Proof of Lemma 21.* Let $m$ satisfy the given condition. We have,

$$\begin{aligned}
\big\|\mathbf{E}[\widehat{\mathbf{Q}}] - \mathbf{Q}\big\|_F^2 \ &=\ \sum_{i=1}^n \sum_{j\neq i} \big(\mathbf{E}[\widehat{Q}_{ij}] - Q_{ij}\big)^2 + \sum_{i=1}^n \big(\mathbf{E}[\widehat{Q}_{ii}] - Q_{ii}\big)^2 \\
&=\ \sum_{i=1}^n \sum_{j\neq i} \Big(\frac{1}{n}\big(\mathbf{E}[\widehat{P}_{ij}] - P_{ij}\big)\Big)^2 + \sum_{i=1}^n \bigg(\frac{1}{n}\sum_{k\neq i}\big(\mathbf{E}[\widehat{P}_{ik}] - P_{ik}\big)\bigg)^2 \\
&\leq\ \frac{(n-1)}{n}\Big(\frac{\epsilon}{2\sqrt{n-1}}\Big)^2 + \frac{(n-1)^2}{n}\Big(\frac{\epsilon}{2\sqrt{n-1}}\Big)^2\,, \\
&\qquad \text{by Lemma 3 (part 3), since } m \geq \tfrac{256n}{\epsilon^2\mu_{\min}^2}\ln\big(\tfrac{8n^2}{\delta}\big) \geq \tfrac{1}{\mu_{\min}}\ln\big(\tfrac{2\sqrt{n-1}}{\epsilon}\big) \\
&=\ (n-1)\Big(\frac{\epsilon}{2\sqrt{n-1}}\Big)^2 \\
&=\ \frac{\epsilon^2}{4}\,.
\end{aligned} \tag{10}$$

Now,

$$
\begin{aligned}
\mathbf{P}\big(\|\widehat{\mathbf{Q}}-\mathbf{Q}\|_2 \geq \epsilon\big) &\leq \mathbf{P}\big(\|\widehat{\mathbf{Q}}-\mathbf{Q}\|_F \geq \epsilon\big)\,, \quad \text{since Frobenius norm upper bounds spectral norm} \\
&\leq \mathbf{P}\big(\big\|\widehat{\mathbf{Q}}-\mathbf{E}[\widehat{\mathbf{Q}}]\big\|_F + \big\|\mathbf{E}[\widehat{\mathbf{Q}}]-\mathbf{Q}\big\|_F \geq \epsilon\big)\,, \quad \text{by triangle inequality} \\
&\leq \mathbf{P}\Big(\big\|\widehat{\mathbf{Q}}-\mathbf{E}[\widehat{\mathbf{Q}}]\big\|_F \geq \frac{\epsilon}{2}\Big)\,, \quad \text{by Eq. (10)} \\
&= \mathbf{P}\Big(\big\|\widehat{\mathbf{Q}}-\mathbf{E}[\widehat{\mathbf{Q}}]\big\|_F^2 \geq \frac{\epsilon^2}{4}\Big) \\
&= \mathbf{P}\bigg(\sum_{i=1}^n \sum_{j\neq i} \big(\widehat{Q}_{ij}-\mathbf{E}[\widehat{Q}_{ij}]\big)^2 + \sum_{i=1}^n \big(\widehat{Q}_{ii}-\mathbf{E}[\widehat{Q}_{ii}]\big)^2 \geq \frac{\epsilon^2}{4}\bigg) \\
&\leq \mathbf{P}\bigg(\sum_{i=1}^n \sum_{j\neq i} \big(\widehat{Q}_{ij}-\mathbf{E}[\widehat{Q}_{ij}]\big)^2 \geq \frac{\epsilon^2}{8}\bigg) + \mathbf{P}\bigg(\sum_{i=1}^n \big(\widehat{Q}_{ii}-\mathbf{E}[\widehat{Q}_{ii}]\big)^2 \geq \frac{\epsilon^2}{8}\bigg) \\
&\leq \sum_{i=1}^n \sum_{j\neq i} \mathbf{P}\bigg(\big|\widehat{Q}_{ij}-\mathbf{E}[\widehat{Q}_{ij}]\big| \geq \frac{\epsilon}{(\sqrt{8})\,n}\bigg) + \sum_{i=1}^n \mathbf{P}\bigg(\big|\widehat{Q}_{ii}-\mathbf{E}[\widehat{Q}_{ii}]\big| \geq \frac{\epsilon}{\sqrt{8n}}\bigg) \\
&= \sum_{i=1}^n \sum_{j\neq i} \mathbf{P}\bigg(\frac{1}{n}\big|\widehat{P}_{ij}-\mathbf{E}[\widehat{P}_{ij}]\big| \geq \frac{\epsilon}{(\sqrt{8})\,n}\bigg) + \sum_{i=1}^n \mathbf{P}\bigg(\frac{1}{n}\bigg|\sum_{k\neq i}\big(\widehat{P}_{ik}-\mathbf{E}[\widehat{P}_{ik}]\big)\bigg| \geq \frac{\epsilon}{\sqrt{8n}}\bigg) \\
&\leq \sum_{i=1}^n \sum_{j\neq i} \mathbf{P}\bigg(\big|\widehat{P}_{ij}-\mathbf{E}[\widehat{P}_{ij}]\big| \geq \frac{\epsilon}{\sqrt{8}}\bigg) + \sum_{i=1}^n \mathbf{P}\bigg(\frac{1}{n}\sum_{k\neq i}\big|\widehat{P}_{ik}-\mathbf{E}[\widehat{P}_{ik}]\big| \geq \frac{\epsilon}{\sqrt{8n}}\bigg) \\
&\leq \sum_{i=1}^n \sum_{j\neq i} \mathbf{P}\bigg(\big|\widehat{P}_{ij}-\mathbf{E}[\widehat{P}_{ij}]\big| \geq \frac{\epsilon}{\sqrt{8}}\bigg) + \sum_{i=1}^n \sum_{k\neq i} \mathbf{P}\bigg(\big|\widehat{P}_{ik}-\mathbf{E}[\widehat{P}_{ik}]\big| \geq \frac{\epsilon}{\sqrt{8n}}\bigg) \\
&\leq 4n^2 \exp\bigg(\frac{-m\epsilon^2\mu_{\min}^2}{256}\bigg) + 4n^2 \exp\bigg(\frac{-m\epsilon^2\mu_{\min}^2}{256n}\bigg)\,, \quad \text{by Lemma 3 (part 2)} \\
&\leq \frac{\delta}{2}+\frac{\delta}{2} \;=\; \delta\,, \quad \text{since } m \geq \tfrac{256n}{\epsilon^2\mu_{\min}^2}\ln\big(\tfrac{8n^2}{\delta}\big).
\end{aligned}
$$

This proves the result. □

**Lemma 22** ((Negahban et al., 2012)). *Let $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ be time-reversible Markov chains defined on the same transition probability graph $G = ([n], E)$, with stationary probability vectors $\boldsymbol{\pi}$ and $\widetilde{\boldsymbol{\pi}}$, respectively. Let $\alpha = \min_{(i,j)\in E} \frac{\pi_i Q_{ij}}{\widetilde{\pi}_i \widetilde{Q}_{ij}}$ and $\beta = \max_i \frac{\pi_i}{\widetilde{\pi}_i}$. Then*

$$1-\lambda_{(2)}(\mathbf{Q}) \;\geq\; \frac{\alpha}{\beta}\big(1-\lambda_{(2)}(\widetilde{\mathbf{Q}})\big)\,.$$

**Lemma 23.** *Let $\mathbf{Q}\in[0,1]^{n\times n}$ be the transition probability matrix of a time-reversible Markov chain with $Q_{ij} > 0\ \forall i,j$, and let $\boldsymbol{\pi}$ be the stationary probability vector of $\mathbf{Q}$. Let $Q_{\min} = \min_{i,j} Q_{ij}$, $\pi_{\max} = \max_i \pi_i$, and $\pi_{\min} = \min_i \pi_i$. Then the spectral gap of $\mathbf{Q}$ satisfies*

$$1-\lambda_{(2)}(\mathbf{Q}) \;\geq\; n\left(\frac{\pi_{\min}}{\pi_{\max}}\right)Q_{\min}\,.$$

*Proof of Lemma 23.* Since $Q_{ij} > 0\ \forall i,j$, the chain $\mathbf{Q}$ is defined on the complete directed graph $G = ([n], [n]\times[n])$. Define a time-reversible Markov chain $\widetilde{\mathbf{Q}}$ on the same graph as follows:

$$\widetilde{Q}_{ij} = \frac{1}{n}\ \ \forall i,j\in[n]\,.$$

This has stationary probability vector $\widetilde{\boldsymbol{\pi}}$ given by $\widetilde{\pi}_i = \frac{1}{n}\ \forall i$. Now, using the notation of Lemma 22, we have

$$
\begin{aligned}
\alpha &= n^2 \pi_{\min} Q_{\min} \\
\beta &= n\,\pi_{\max}\,.
\end{aligned}
$$

Moreover, $\lambda_{(2)}(\widetilde{\mathbf{Q}}) = 0$. By Lemma 22, we therefore have

$$1 - \lambda_{(2)}(\mathbf{Q}) \;\geq\; \frac{n^2 \pi_{\min} Q_{\min}}{n \pi_{\max}} \;=\; n \Big(\frac{\pi_{\min}}{\pi_{\max}}\Big) Q_{\min} \,.$$

□

**Lemma 24** ((Negahban et al., 2012)). *Let $\mathbf{Q}$ be a time-reversible Markov chain with stationary probability vector $\boldsymbol{\pi}$. Let $\widehat{\mathbf{Q}}$ be any other Markov chain, and let $\mathbf{q}_t$ denote the state distribution of $\widehat{\mathbf{Q}}$ at time $t$ when started with initial distribution $\mathbf{q}_0$. Let $\pi_{\max} = \max_i \pi_i$, $\pi_{\min} = \min_i \pi_i$, and $\rho = \lambda_{(2)}(\mathbf{Q}) + \|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \sqrt{\frac{\pi_{\max}}{\pi_{\min}}}$. Then*

$$\frac{\|\mathbf{q}_t - \boldsymbol{\pi}\|_2}{\|\boldsymbol{\pi}\|_2} \;\leq\; \rho^t \frac{\|\mathbf{q}_0 - \boldsymbol{\pi}\|_2}{\|\boldsymbol{\pi}\|_2} \sqrt{\frac{\pi_{\max}}{\pi_{\min}}} + \frac{1}{1-\rho} \|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \sqrt{\frac{\pi_{\max}}{\pi_{\min}}} \,.$$

We are now ready to prove Theorem 7:

*Proof of Theorem 7.* Let $m$ satisfy the given condition. Then by Lemma 21, we have with probability at least $1 - \frac{\delta}{2}$, the empirical Markov chain $\widehat{\mathbf{Q}}$ constructed by the rank centrality algorithm satisfies

$$\|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \;\leq\; \frac{\epsilon}{2} \Big(\frac{\pi_{\min}}{\pi_{\max}}\Big)^{3/2} P_{\min} \,. \tag{11}$$

In this case, since $\mathbf{Q}$ is time-reversible with $Q_{ij} > 0 \; \forall i, j$ and $Q_{\min} = \min_{i,j} Q_{ij} = \frac{P_{\min}}{n}$, by Lemma 23 and Eq. (11), we have

$$\begin{aligned}
\rho \;=\; \lambda_{(2)}(\mathbf{Q}) + \|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \sqrt{\frac{\pi_{\max}}{\pi_{\min}}} \quad &\leq \quad 1 - \Big(\frac{\pi_{\min}}{\pi_{\max}}\Big) P_{\min} + \frac{\epsilon}{2} \Big(\frac{\pi_{\min}}{\pi_{\max}}\Big) P_{\min} \\
&\leq \quad 1 - \frac{1}{2} \Big(\frac{\pi_{\min}}{\pi_{\max}}\Big) P_{\min} \,. \qquad (12)
\end{aligned}$$

Next, since $m \geq \frac{1024\, n}{\epsilon^2 P_{\min}^2 \mu_{\min}^2} \big(\frac{\pi_{\max}}{\pi_{\min}}\big)^3 \ln\big(\frac{16 n^2}{\delta}\big) \geq \frac{1}{\mu_{\min} P_{\min}} \ln\big(\frac{2n(n-1)}{\delta}\big)$, by Lemma 3 (part 5), we have that with probability at least $1 - \frac{\delta}{2}$, $\widehat{P}_{ij} > 0 \; \forall i \neq j$, and therefore $\widehat{\mathbf{Q}}$ is an irreducible and aperiodic Markov chain.

Putting the above two statements together, with probability at least $1-\delta$, we have that $\widehat{\mathbf{Q}}$ is an irreducible, aperiodic Markov chain satisfying Eqs. (11-12), and the score vector $\widehat{\boldsymbol{\pi}}$ output by the rank centrality algorithm is the stationary probability vector of $\widehat{\mathbf{Q}}$. In this case, by Lemma 24, we have that for any initial distribution $\mathbf{q}_0$ of $\widehat{\mathbf{Q}}$,

$$\begin{aligned}
\frac{\|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2}{\|\boldsymbol{\pi}\|_2} \;=\; \lim_{t\to\infty} \frac{\|\mathbf{q}_t - \boldsymbol{\pi}\|_2}{\|\boldsymbol{\pi}\|_2} \quad &\leq \quad \lim_{t\to\infty} \rho^t \frac{\|\mathbf{q}_0 - \boldsymbol{\pi}\|_2}{\|\boldsymbol{\pi}\|_2} \sqrt{\frac{\pi_{\max}}{\pi_{\min}}} + \frac{1}{1-\rho} \|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \sqrt{\frac{\pi_{\max}}{\pi_{\min}}} \\
&= \quad \frac{1}{1-\rho} \|\widehat{\mathbf{Q}} - \mathbf{Q}\|_2 \sqrt{\frac{\pi_{\max}}{\pi_{\min}}} \,, \quad \text{since } \rho < 1, \text{ by Eq. (12)} \\
&\leq \quad \epsilon \,, \quad \text{by Eqs. (11-12).}
\end{aligned}$$

The result follows since $\|\boldsymbol{\pi}\|_2 \leq 1$, which gives $\|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2 \leq \frac{\|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2}{\|\boldsymbol{\pi}\|_2}$. □

**Proof of Corollary 8**

*Proof.* Let $\epsilon = \frac{r_{\min}}{3}$. By definition of $r_{\min}$, we have $r_{\min} \leq 1$, and therefore $\epsilon \leq \frac{1}{3} < 1$. Therefore if $m$ satisfies the given condition, then by Theorem 7, we have with probability at least $1 - \delta$,

$$\|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2 \leq \frac{r_{\min}}{3} \,.$$

But

$$\begin{aligned}
\|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2 \leq \frac{r_{\min}}{3} \quad & \implies \quad \|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_\infty \leq \frac{r_{\min}}{3}\,, \quad \text{since } L_2 \text{ norm upper bounds } L_\infty \text{ norm} \\
& \implies \quad |\widehat{\pi}_i - \pi_i| \leq \frac{r_{\min}}{3} \;\; \forall i \\
& \implies \quad \Big\{ \forall i,j:\; \pi_j > \pi_i \implies \widehat{\pi}_j > \widehat{\pi}_i \Big\}\,, \quad \text{by definition of } r_{\min} \\
& \implies \quad \Big\{ \forall i,j:\; P_{ij} > P_{ji} \implies \widehat{\pi}_j > \widehat{\pi}_i \Big\}\,, \quad \text{by time-reversibility condition on } \mathbf{P} \\
& \implies \quad \text{argsort}(\widehat{\boldsymbol{\pi}}) \subseteq \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma] \\
& \implies \quad \widehat{\sigma} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.
\end{aligned}$$

Thus we have that with probability at least $1-\delta$,

$$\widehat{\sigma} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.$$

This proves the result. □

**Proof of Lemma 9**

*Proof.* Let $\mathbf{Q}$ be as defined in Eq. (6), and let $\boldsymbol{\pi}$ be the stationary probability vector of $\mathbf{Q}$. From Section 6, we know that $\mathbf{P}$ satisfies the time-reversibility condition, and therefore any permutation that ranks items according to decreasing order of scores $\pi_i$ is an optimal permutation w.r.t. the pairwise disagreement error, i.e. we have

$$\text{argsort}(\boldsymbol{\pi}) \subseteq \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.$$

We will show that argsort$(\mathbf{f}^*) =$ argsort$(\boldsymbol{\pi})$, which will imply the result. We have,

$$\begin{aligned}
f_i^* \;=\; -\frac{1}{n}\sum_{k=1}^n Y_{ik} \;=\; \frac{1}{n}\sum_{k=1}^n \ln\Big(\frac{P_{ki}}{P_{ik}}\Big) \quad &= \quad \frac{1}{n}\ln\left(\prod_{k=1}^n \frac{P_{ki}}{P_{ik}}\right) \\
&= \quad \frac{1}{n}\ln\left(\prod_{k=1}^n \frac{\pi_i}{\pi_k}\right)\,, \quad \text{by time-reversibility} \\
&= \quad \ln \pi_i - \frac{1}{n}\ln(\pi_1 \cdot \ldots \cdot \pi_n)\,.
\end{aligned}$$

The second term on the right-hand side is a constant, and $\ln(\cdot)$ is a strictly monotonically increasing function; therefore $\mathbf{f}^*$ induces the same orderings as $\boldsymbol{\pi}$, i.e. argsort$(\mathbf{f}^*) =$ argsort$(\boldsymbol{\pi})$. □

**Proof of Theorem 10**

The proof makes use of the following technical lemma:

**Lemma 25.** *Let* $0 < u, u' < 1$. *Let* $0 < \epsilon < u$. *Then*

$$|u - u'| \leq \epsilon \implies \big|\ln(u) - \ln(u')\big| \leq \frac{\epsilon}{u - \epsilon}\,.$$

*Proof.* Let $|u-u'| \leq \epsilon$. Thus $u' \in (u-\epsilon, u+\epsilon)$. Now, since $\ln(\cdot)$ is a concave function, we have

$$\ln(y) \leq \ln(x) + \frac{1}{x}(y-x) \;\; \forall x,y > 0\,.$$

Taking $x = u$ and $y = u+\epsilon$ gives

$$\ln(u+\epsilon) \;\leq\; \ln(u) + \frac{\epsilon}{u}\,;$$

taking $x = u-\epsilon$ and $y = u$ gives

$$\ln(u) \;\leq\; \ln(u-\epsilon) + \frac{\epsilon}{u-\epsilon}\,.$$

Combining both, and using the fact that $\ln(\cdot)$ is a monotonically increasing function, we get

$$\ln(u) - \frac{\epsilon}{u-\epsilon} \;\leq\; \ln(u-\epsilon) \;\leq\; \ln(u) \;\leq\; \ln(u+\epsilon) \;\leq\; \ln(u) + \frac{\epsilon}{u}\,.$$

For $\epsilon < u$, we have $\frac{\epsilon}{u-\epsilon} > \frac{\epsilon}{u}$. Thus, since $u' \in (u-\epsilon, u+\epsilon)$, we have either

$$\ln(u) - \frac{\epsilon}{u-\epsilon} \;\leq\; \ln(u-\epsilon) \;\leq\; \ln(u') \;\leq\; \ln(u)$$

or

$$\ln(u) \;\leq\; \ln(u') \;\leq\; \ln(u+\epsilon) \;\leq\; \ln(u) + \frac{\epsilon}{u} \;<\; \ln(u) + \frac{\epsilon}{u-\epsilon}\,;$$

in both cases, we get $\big|\ln(u) - \ln(u')\big| \leq \frac{\epsilon}{u-\epsilon}$, thus proving the result. □

*Proof of Theorem 10.* Let $m$ satisfy the given condition. Since $m \geq \frac{128}{P_{\min}^2 \mu_{\min}^2}\big(1+\frac{2}{\epsilon}\big)^2 \ln\big(\frac{16n^2}{\delta}\big) \geq \frac{1}{\mu_{\min} P_{\min}} \ln\big(\frac{2n(n-1)}{\delta}\big)$, by Lemma 3 (part 5), we have with probability at least $1-\frac{\delta}{2}$, $\widehat{P}_{ij} > 0\ \forall i \neq j$. In this case, we have

$$\widehat{Y}_{ij} = \begin{cases} \ln\Big(\frac{\widehat{P}_{ij}}{\widehat{P}_{ji}}\Big) & \text{if } i \neq j \\ 0 & \text{otherwise,} \end{cases}$$

and $\widehat{E} = \mathcal{X}$. As discussed in (Jiang et al., 2011), the score vector $\widehat{\mathbf{f}}$ output by the least squares algorithm in this case is given by

$$\widehat{f}_i \;=\; -\frac{1}{n}\sum_{k=1}^{n} \widehat{Y}_{ik} \;=\; \frac{1}{n}\sum_{k\neq i} \ln\left(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}}\right).$$

Moreover, since $P_{ij} \in (0,1)\ \forall i \neq j$, we also have

$$f_i^* \;=\; =\; -\frac{1}{n}\sum_{k=1}^{n} Y_{ik} \;=\; \frac{1}{n}\sum_{k\neq i} \ln\left(\frac{P_{ki}}{P_{ik}}\right).$$

Next, we have

$$\begin{aligned}
&\mathbf{P}\bigg(\exists i : \bigg|\frac{1}{n}\sum_{k\neq i}\ln\Big(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}}\Big) - \frac{1}{n}\sum_{k\neq i}\ln\Big(\frac{P_{ki}}{P_{ik}}\Big)\bigg| \geq \epsilon\bigg)\\
&\leq \sum_{i=1}^{n}\mathbf{P}\bigg(\frac{1}{n}\sum_{k\neq i}\Big|\ln\Big(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}}\Big) - \ln\Big(\frac{P_{ki}}{P_{ik}}\Big)\Big| \geq \epsilon\bigg), \quad \text{by union bound and triangle inequality}\\
&\leq \sum_{i=1}^{n}\mathbf{P}\bigg(\exists k \neq i : \Big|\ln\Big(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}} - \ln\Big(\frac{P_{ki}}{P_{ik}}\Big)\Big)\Big| \geq \epsilon\bigg),\\
&\leq \sum_{i=1}^{n}\sum_{k\neq i}\mathbf{P}\bigg(\Big|\ln\Big(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}}\Big) - \ln\Big(\frac{P_{ki}}{P_{ik}}\Big)\Big| \geq \epsilon\bigg), \quad \text{by union bound}\\
&\leq \sum_{i=1}^{n}\sum_{k\neq i}\mathbf{P}\bigg(\big|\ln\widehat{P}_{ki} - \ln P_{ki}\big| + \big|\ln\widehat{P}_{ik} - \ln P_{ik}\big| \geq \epsilon\bigg)\\
&\leq 2\sum_{i=1}^{n}\sum_{k\neq i}\mathbf{P}\bigg(\big|\ln\widehat{P}_{ki} - \ln P_{ki}\big| \geq \frac{\epsilon}{2}\bigg)\\
&\leq 2\sum_{i=1}^{n}\sum_{k\neq i}\mathbf{P}\bigg(\big|\widehat{P}_{ki} - P_{ki}\big| \geq \frac{\epsilon P_{\min}}{2+\epsilon}\bigg), \quad \text{by Lemma 25}\\
&\leq 8n^2\exp\Big(\frac{-m\epsilon^2 P_{\min}^2\mu_{\min}^2}{128(2+\epsilon)^2}\Big), \quad \text{by Lemma 3 (part 4)}\\
&\qquad\qquad \text{(since } m \geq \tfrac{128}{P_{\min}^2\mu_{\min}^2}\big(1+\tfrac{2}{\epsilon}\big)^2\ln\big(\tfrac{16n^2}{\delta}\big) \geq \tfrac{1}{\mu_{\min}}\ln\big(\tfrac{2(2+\epsilon)}{\epsilon P_{\min}}\big)\text{)}\\
&\leq \frac{\delta}{2}, \quad \text{since } m \geq \tfrac{128}{P_{\min}^2\mu_{\min}^2}\big(1+\tfrac{2}{\epsilon}\big)^2\ln\big(\tfrac{16n^2}{\delta}\big).
\end{aligned}$$

In other words, with probability at least $1-\frac{\delta}{2}$, we have

$$\max_i\bigg|\frac{1}{n}\sum_{k\neq i}\ln\Big(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}}\Big) - \frac{1}{n}\sum_{k\neq i}\ln\Big(\frac{P_{ki}}{P_{ik}}\Big)\bigg| \leq \epsilon\,.$$

Putting the above statements together, we have that with probability at least $1-\delta$,

$$\|\widehat{\mathbf{f}} - \mathbf{f}^*\|_\infty = \max_i\big|\widehat{f}_i - f_i^*\big| = \max_i\bigg|\frac{1}{n}\sum_{k\neq i}\ln\Big(\frac{\widehat{P}_{ki}}{\widehat{P}_{ik}}\Big) - \frac{1}{n}\sum_{k\neq i}\ln\Big(\frac{P_{ki}}{P_{ik}}\Big)\bigg| \leq \epsilon\,.$$

This proves the result. □

**Proof of Corollary 11**

*Proof.* Let $\epsilon = \frac{r_{\min}}{3}$. By definition of $r_{\min}$, we have $r_{\min} \leq n$, and therefore $\epsilon \leq \frac{n}{3} \leq (4\sqrt{2})n$. Therefore if $m$ satisfies the given condition, then by Theorem 10, we have with probability at least $1-\delta$,

$$\|\widehat{\mathbf{f}} - \mathbf{f}^*\|_\infty \leq \frac{r_{\min}}{3}\,.$$

But

$$\begin{aligned}
\|\widehat{\mathbf{f}} - \mathbf{f}^*\|_\infty \leq \frac{r_{\min}}{3} \quad &\Longrightarrow \quad |\widehat{f}_i - f_i^*| \leq \frac{r_{\min}}{3} \ \ \forall i\\
&\Longrightarrow \quad \Big\{\forall i,j:\ f_j^* > f_i^* \implies \widehat{f}_j > \widehat{f}_i\Big\}, \quad \text{by definition of } r_{\min}\\
&\Longrightarrow \quad \Big\{\forall i,j:\ P_{ij} > P_{ji} \implies \widehat{f}_j > \widehat{f}_i\Big\}, \quad \text{by Lemma 9}\\
&\Longrightarrow \quad \text{argsort}(\widehat{\mathbf{f}}) \subseteq \text{argmin}_{\sigma\in\mathcal{S}_n}\text{er}_{\mu,\mathbf{P}}^{\text{PD}}[\sigma]\\
&\Longrightarrow \quad \widehat{\sigma} \in \text{argmin}_{\sigma\in\mathcal{S}_n}\text{er}_{\mu,\mathbf{P}}^{\text{PD}}[\sigma]\,.
\end{aligned}$$

Thus we have that with probability at least $1-\delta$,

$$\widehat{\sigma} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.$$

This proves the result. □

**Proof of Lemma 13**

*Proof.* Let $\mathbf{P} \in [0,1]^{n\times n}$ satisfy the BTL condition with vector $\mathbf{w} \in \mathbb{R}^n_+$, so that $w_i > 0\ \forall i$ and $P_{ij} = \frac{w_j}{w_i+w_j}\ \forall i \neq j$. Then we have

$$\begin{aligned}
P_{ij} > P_{ji} \quad &\implies \quad w_j > w_i \\
&\implies \quad \frac{w_j}{w_j+w_k} > \frac{w_i}{w_i+w_k}\ \ \forall k \\
&\implies \quad \sum_{k=1}^n \frac{w_j}{w_j+w_k} > \sum_{k=1}^n \frac{w_i}{w_i+w_k} \\
&\implies \quad \sum_{k=1}^n P_{kj} > \sum_{k=1}^n P_{ki}\,.
\end{aligned}$$

Thus $\mathbf{P}$ satisfies the LN condition. □

**Proof of Theorem 14**

*Proof.* Let $m$ satisfy the given condition. We have,

$$\begin{aligned}
\mathbf{P}\big(\|\widehat{\mathbf{f}}-\mathbf{f}^*\|_\infty \geq \epsilon\big) \quad &= \quad \mathbf{P}\big(\exists i : |\widehat{f}_i - f^*_i| \geq \epsilon\big) \\
&\leq \quad \sum_{i=1}^n \mathbf{P}\big(|\widehat{f}_i - f^*_i| \geq \epsilon\big)\,, \quad \text{by union bound} \\
&= \quad \sum_{i=1}^n \mathbf{P}\left(\left|\frac{1}{n}\sum_{k=1}^n \big(\widehat{P}_{ki} - P_{ki}\big)\right| \geq \epsilon\right), \quad \text{by definition of } \widehat{f}_i \text{ and } f^*_i \\
&\leq \quad \sum_{i=1}^n \mathbf{P}\left(\frac{1}{n}\sum_{k=1}^n \big|\widehat{P}_{ki} - P_{ki}\big| \geq \epsilon\right) \\
&\leq \quad \sum_{i=1}^n \mathbf{P}\Big(\exists k : \big|\widehat{P}_{ki} - P_{ki}\big| \geq \epsilon\Big) \\
&\leq \quad \sum_{i=1}^n \sum_{k=1}^n \mathbf{P}\Big(\big|\widehat{P}_{ki} - P_{ki}\big| \geq \epsilon\Big)\,, \quad \text{by union bound} \\
&\leq \quad 4n^2 \exp\left(\frac{-m\epsilon^2\mu^2_{\min}}{128}\right), \quad \text{by Lemma 3 (part 4)} \\
&\qquad\qquad\qquad \left(\text{since } m \geq \tfrac{128}{\epsilon^2\mu^2_{\min}} \ln\left(\tfrac{4n^2}{\delta}\right) \geq \tfrac{1}{\mu_{\min}} \ln\left(\tfrac{2n}{\epsilon}\right)\right) \\
&\leq \quad \delta\,, \quad \text{since } m \geq \tfrac{128}{\epsilon^2\mu^2_{\min}} \ln\left(\tfrac{4n^2}{\delta}\right).
\end{aligned}$$

This proves the result. □

**Proof of Corollary 15**

*Proof.* Let $\epsilon = \frac{r_{\min}}{3}$. By definition of $r_{\min}$, we have $r_{\min} \leq n$, and therefore $\epsilon \leq \frac{n}{3} \leq (4\sqrt{2})n$. Therefore if $m$ satisfies the given condition, then by Theorem 14, we have with probability at least $1-\delta$,

$$\|\widehat{\mathbf{f}} - \mathbf{f}^*\|_\infty \leq \frac{r_{\min}}{3}\,.$$

But

$$\begin{aligned}
\|\widehat{\mathbf{f}} - \mathbf{f}^*\|_\infty \leq \frac{r_{\min}}{3} \quad &\implies \quad |\widehat{f}_i - f_i^*| \leq \frac{r_{\min}}{3} \;\; \forall i \\
&\implies \quad \Big\{\forall i,j: \; f_j^* > f_i^* \implies \widehat{f}_j > \widehat{f}_i\Big\}, \quad \text{by definition of } r_{\min} \\
&\implies \quad \Big\{\forall i,j: \; P_{ij} > P_{ji} \implies \widehat{f}_j > \widehat{f}_i\Big\}, \quad \text{by extended low-noise condition on } \mathbf{P} \\
&\implies \quad \text{argsort}(\widehat{\mathbf{f}}) \subseteq \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma] \\
&\implies \quad \widehat{\sigma} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.
\end{aligned}$$

Thus we have that with probability at least $1-\delta$,

$$\widehat{\sigma} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.$$

This proves the result. □

**Proof of Theorem 16**

*Proof.* We have

$$\widehat{\boldsymbol{\theta}} \in \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^n} \sum_{i<j} \Big( \ln(1 + \exp(\theta_j - \theta_i)) - \widehat{P}_{ij}(\theta_j - \theta_i) \Big)\,.$$

Setting the gradient of the above objective to $\mathbf{0}$ gives:

$$\forall i: \quad \sum_{k=1}^n \widehat{P}_{ki} \;=\; \sum_{k \neq i} \frac{\exp(\widehat{\theta}_i)}{\exp(\widehat{\theta}_k) + \exp(\widehat{\theta}_i)} \;=\; \sum_{k=1}^n P^{\widehat{\boldsymbol{\theta}}}_{ki}\,, \tag{13}$$

where we denote

$$P^{\widehat{\boldsymbol{\theta}}}_{ij} \;=\; \begin{cases} \frac{\exp(\widehat{\theta}_j)}{\exp(\widehat{\theta}_i)+\exp(\widehat{\theta}_j)} & \text{if } i<j \\ 1 - P^{\widehat{\boldsymbol{\theta}}}_{ji} & \text{if } i>j \\ 0 & \text{if } i=j. \end{cases}$$

Now, we have for any $0 < \epsilon < 4\sqrt{2}$, if $m \geq \max\big(B(\mu_{\min}),\; \frac{1}{\mu_{\min}} \ln(\frac{2}{\epsilon})\big)$, then

$$\begin{aligned}
\mathbf{P}\Big(\exists i: \Big|\frac{1}{n}\sum_{k=1}^n \big(P_{ki} - P^{\widehat{\boldsymbol{\theta}}}_{ki}\big)\Big| \geq \epsilon\Big) \;&=\; \mathbf{P}\Big(\exists i: \Big|\frac{1}{n}\sum_{k=1}^n \big(P_{ki} - \widehat{P}_{ki}\big)\Big| \geq \epsilon\Big), \quad \text{by Eq. (13)} \\
&\leq\; \sum_{i=1}^n \mathbf{P}\Big(\Big|\frac{1}{n}\sum_{k=1}^n \big(P_{ki} - \widehat{P}_{ki}\big)\Big| \geq \epsilon\Big) \\
&\leq\; \sum_{i=1}^n \mathbf{P}\Big(\frac{1}{n}\sum_{k=1}^n \big|P_{ki} - \widehat{P}_{ki}\big| \geq \epsilon\Big) \\
&\leq\; \sum_{i=1}^n \mathbf{P}\Big(\exists k: \big|P_{ki} - \widehat{P}_{ki}\big| \geq \epsilon\Big) \\
&\leq\; \sum_{i=1}^n \sum_{k=1}^n \mathbf{P}\big(\big|P_{ki} - \widehat{P}_{ki}\big| \geq \epsilon\big) \\
&\leq\; 4n^2 \exp\Big(\frac{-m\epsilon^2\mu_{\min}^2}{128}\Big), \quad \text{by Lemma 3 (part 4).}
\end{aligned}$$

Setting $\epsilon = \frac{r_{\min}}{3}$, we get that if $m$ satisfies the given condition, then with probability at least $1-\delta$,

$$\Big|\frac{1}{n}\sum_{k=1}^n \big(P_{kj} - P^{\widehat{\boldsymbol{\theta}}}_{kj}\big)\Big| \;\leq\; \frac{r_{\min}}{3} \quad \forall i\,. \tag{14}$$

By definition of $r_{\min}$, this means that with probability at least $1-\delta$, we have for all $i \neq j$,

$$\sum_{k=1}^{n} P^{\widehat{\boldsymbol{\theta}}}_{kj} > \sum_{k=1}^{n} P^{\widehat{\boldsymbol{\theta}}}_{ki} \iff \sum_{k=1}^{n} P_{kj} > \sum_{k=1}^{n} P_{ki} \iff f^*_j > f^*_i \,. \tag{15}$$

Also, it is easy to verify that for all $i \neq j$,

$$\widehat{w}_j > \widehat{w}_i \iff \widehat{\theta}_j > \widehat{\theta}_i \iff \sum_{k=1}^{n} P^{\widehat{\boldsymbol{\theta}}}_{kj} \; > \; \sum_{k=1}^{n} P^{\widehat{\boldsymbol{\theta}}}_{ki} \tag{16}$$

Combining Eqs. (15-16), we have that with probability at least $1-\delta$,

$$\text{argsort}(\widehat{\mathbf{w}}) = \text{argsort}(\mathbf{f}^*) \,.$$

Since $\text{argsort}(\mathbf{f}^*) \subseteq \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]$, we thus have with probability at least $1-\delta$,

$$\widehat{\sigma} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma] \,.$$

This proves the result. □

**Proof of Proposition 19**

*Proof.* Suppose $\mathbf{P}$ satisfies the GLN condition with vector $\boldsymbol{\alpha} \in \mathbb{R}^n$. Then clearly, since $P_{ij} \neq \frac{1}{2} \; \forall i \neq j$, we have $\forall i < j$:

$$z_{ij} = 1 \implies P_{ji} > P_{ij} \implies \sum_{k=1}^{n} \alpha_k P_{ki} > \sum_{k=1}^{n} \alpha_k P_{kj} \implies \boldsymbol{\alpha}^\top (\mathbf{P}_i - \mathbf{P}_j) > 0$$
$$z_{ij} = -1 \implies P_{ij} > P_{ji} \implies \sum_{k=1}^{n} \alpha_k P_{kj} > \sum_{k=1}^{n} \alpha_k P_{ki} \implies \boldsymbol{\alpha}^\top (\mathbf{P}_i - \mathbf{P}_j) < 0 \,.$$

Thus $S_{\mathbf{P}}$ is linearly separable by the hyperplane $\boldsymbol{\alpha}$ passing through the origin.

Conversely, suppose that $S_{\mathbf{P}}$ is linearly separable by a hyperplane passing through the origin. Then $\exists \boldsymbol{\alpha} \in \mathbb{R}^n$ s.t. $z_{ij} \boldsymbol{\alpha}^\top (\mathbf{P}_i - \mathbf{P}_j) > 0 \; \forall i < j$. Thus we have $\forall i < j$:

$$P_{ij} > P_{ji} \implies z_{ij} = -1 \implies \boldsymbol{\alpha}^\top (\mathbf{P}_i - \mathbf{P}_j) < 0 \implies \sum_{k=1}^{n} \alpha_k P_{kj} > \sum_{k=1}^{n} \alpha_k P_{ki}$$
$$P_{ji} > P_{ij} \implies z_{ij} = 1 \implies \boldsymbol{\alpha}^\top (\mathbf{P}_i - \mathbf{P}_j) > 0 \implies \sum_{k=1}^{n} \alpha_k P_{ki} > \sum_{k=1}^{n} \alpha_k P_{kj} \,.$$

Thus $\mathbf{P}$ satisfies the GLN condition. □

**Proof of Theorem 20**

*Proof.* Let $m$ satisfy the given conditions. We first show that with probability at least $1-\frac{\delta}{2}$, every label $\text{sign}(\widehat{P}_{ji} - \widehat{P}_{ij})$ in $S_{\widehat{\mathbf{P}}}$ is the same as the corresponding label $\text{sign}(P_{ji} - P_{ij})$ in $S_{\mathbf{P}}$. We have,

$$\begin{aligned}
\mathbf{P}\Big(\exists i \neq j : \big|\widehat{P}_{ij} - P_{ij}\big| \geq \gamma\Big) \quad &\leq \quad \sum_{i \neq j} \mathbf{P}\Big(\big|\widehat{P}_{ij} - P_{ij}\big| \geq \gamma\Big) \,, \quad \text{by union bound} \\
&\leq \quad 4n^2 \exp\left(\frac{-m\gamma^2 \mu_{\min}^2}{128}\right) , \quad \text{by Lemma 3 (part 4)} \\
&\qquad\quad (\text{since } m \geq \tfrac{128}{\gamma^2 \mu_{\min}^2} \log(\tfrac{8n^2}{\delta}) \geq \tfrac{1}{\mu_{\min}} \ln(\tfrac{2}{\gamma})) \\
&\leq \quad \frac{\delta}{2} \,, \quad \text{since } m \geq \tfrac{128}{\gamma^2 \mu_{\min}^2} \log(\tfrac{8n^2}{\delta}).
\end{aligned}$$

Thus we have that with probability at least $1-\frac{\delta}{2}$,

$$|\widehat{P}_{ij}-P_{ij}|\leq\gamma\quad\forall i\neq j\,.$$

By definition of $\gamma$, this yields that with probability at least $1-\frac{\delta}{2}$,

$$\widehat{P}_{ij}>\widehat{P}_{ji}\Longleftrightarrow P_{ij}>P_{ji}\quad\forall i\neq j\,,$$

i.e. with probability at least $1-\frac{\delta}{2}$,

$$\text{sign}(\widehat{P}_{ji}-\widehat{P}_{ij})=\text{sign}(P_{ji}-P_{ij})\quad\forall i<j\,.$$

Next, we show that with probability at least $1-\frac{\delta}{2}$, every point $(\widehat{\mathbf{P}}_i-\widehat{\mathbf{P}}_j)$ in $S_{\widehat{\mathbf{P}}}$ falls on the same side of the hyperplane given by $\boldsymbol{\alpha}$ as the corresponding point $(\mathbf{P}_i-\mathbf{P}_j)$ in $S_{\mathbf{P}}$. We have,

$$\begin{aligned}
&\mathbf{P}\Big(\exists(i<j):\|(\widehat{\mathbf{P}}_i-\widehat{\mathbf{P}}_j)-(\mathbf{P}_i-\mathbf{P}_j)\|_2\ \geq\ \frac{r^{\boldsymbol{\alpha}}_{\min}}{2}\Big)\\
&\quad=\ \mathbf{P}\Big(\exists(i<j):\|(\widehat{\mathbf{P}}_i-\mathbf{P}_i)-(\widehat{\mathbf{P}}_j-\mathbf{P}_j)\|_2\ \geq\ \frac{r^{\boldsymbol{\alpha}}_{\min}}{2}\Big)\\
&\quad\leq\ \sum_{i<j}\mathbf{P}\Big(\|(\widehat{\mathbf{P}}_i-\mathbf{P}_i)-(\widehat{\mathbf{P}}_j-\mathbf{P}_j)\|_2\geq\frac{r^{\boldsymbol{\alpha}}_{\min}}{2}\Big)\,,\quad\text{by union bound}\\
&\quad\leq\ \sum_{i<j}\bigg(\mathbf{P}\Big(\|\widehat{\mathbf{P}}_i-\mathbf{P}_i\|_2\geq\frac{r^{\boldsymbol{\alpha}}_{\min}}{4}\Big)+\mathbf{P}\Big(\|\widehat{\mathbf{P}}_j-\mathbf{P}_j\|_2\geq\frac{r^{\boldsymbol{\alpha}}_{\min}}{4}\Big)\bigg)\\
&\quad\leq\ \sum_{i<j}\bigg(\mathbf{P}\Big(\exists k:|\widehat{P}_{ki}-P_{ki}|\geq\frac{r^{\boldsymbol{\alpha}}_{\min}}{4\sqrt{n}}\Big)+\mathbf{P}\Big(\exists k:|\widehat{P}_{kj}-P_{kj}|\geq\frac{r^{\boldsymbol{\alpha}}_{\min}}{4\sqrt{n}}\Big)\bigg)\\
&\quad\leq\ \sum_{i<j}\bigg(\sum_k\mathbf{P}\Big(|\widehat{P}_{ki}-P_{ki}|\geq\frac{r^{\alpha}_{\min}}{4\sqrt{n}}\Big)+\sum_k\mathbf{P}\Big(|\widehat{P}_{kj}-P_{kj}|\geq\frac{r^{\alpha}_{\min}}{4\sqrt{n}}\Big)\bigg)\\
&\quad\leq\ 8n^3\exp\Big(\frac{-m(r^{\boldsymbol{\alpha}}_{\min})^2\mu^2_{\min}}{2048n}\Big)\,,\quad\text{by Lemma 3 (part 4)}\\
&\qquad\qquad(\text{since } m\geq\tfrac{2048\,n}{(r^{\boldsymbol{\alpha}}_{\min})^2\mu^2_{\min}}\log(\tfrac{16n^3}{\delta})\geq\tfrac{1}{\mu_{\min}}\ln(\tfrac{8\sqrt{n}}{r^{\boldsymbol{\alpha}}_{\min}}))\\
&\quad\leq\ \frac{\delta}{2}\,,\quad\text{since } m\geq\tfrac{2048\,n}{(r^{\boldsymbol{\alpha}}_{\min})^2\mu^2_{\min}}\log(\tfrac{16n^3}{\delta}).
\end{aligned}$$

Thus with probability at least $1-\frac{\delta}{2}$,

$$\|(\widehat{\mathbf{P}}_i-\widehat{\mathbf{P}}_j)-(\mathbf{P}_i-\mathbf{P}_j)\|_2\ \leq\ \frac{r^{\boldsymbol{\alpha}}_{\min}}{2}\quad\forall i<j\,.$$

By definition, $r^{\boldsymbol{\alpha}}_{\min}$ is the smallest Euclidean distance of any point $(\mathbf{P}_i-\mathbf{P}_j)$ to the hyperplane defined by $\boldsymbol{\alpha}$; therefore we get that with probability at least $1-\frac{\delta}{2}$, all points $(\widehat{\mathbf{P}}_i-\widehat{\mathbf{P}}_j)$ fall on the same side of the hyperplane $\boldsymbol{\alpha}$ as the corresponding points $(\mathbf{P}_i-\mathbf{P}_j)$.

Combining the above statements yields that with probability at least $1-\delta$, the dataset $S_{\widehat{\mathbf{P}}}$ is also linearly separable by $\boldsymbol{\alpha}$; in this case, the SVM-RankAggregation algorithm produces a vector $\widehat{\boldsymbol{\alpha}}$ that correctly classifies all examples in $S_{\widehat{\mathbf{P}}}$, i.e. satisfies $z_{ij}\widehat{\boldsymbol{\alpha}}^\top(\widehat{\mathbf{P}}_i-\widehat{\mathbf{P}}_j)>0\;\forall i<j$ (where $z_{ij}=\text{sign}(P_{ji}-P_{ij})$), and it can be verified that $\widehat{\boldsymbol{\alpha}}$ must then also satisfy $z_{ij}\widehat{\boldsymbol{\alpha}}^\top(\mathbf{P}_i-\mathbf{P}_j)>0\;\forall i<j$, so that $\text{argsort}(\widehat{\boldsymbol{\alpha}})\subseteq\text{argmin}_{\sigma\in\mathcal{S}_n}\text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]$. This yields that with probability at least $1-\delta$,

$$\widehat{\sigma}\in\text{argmin}_{\sigma\in\mathcal{S}_n}\text{er}^{\text{PD}}_{\mu,\mathbf{P}}[\sigma]\,.$$

□