

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:

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

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 ω, ω' 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

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

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

Part 4. Let m satisfy the given condition. Then

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

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

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

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 π ; 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 $\mathbf{w} = \pi \in \mathbb{R}_+^n$. 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 π given by $\pi_i = \frac{w_i}{\sum_{k=1}^n w_k}$ satisfies

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

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

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 $\|\mathbf{Q}\|_F = (\sum_{i=1}^n \sum_{j=1}^n Q_{ij}^2)^{1/2}$ the Frobenius norm of \mathbf{Q} , by $\|\mathbf{Q}\|_2 = \max_{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0} \frac{\|\mathbf{Q}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$ 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 (μ, \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*

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

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

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

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

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

Now,

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

This proves the result. \square

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*

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

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:

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

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

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

□

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

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

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 $\hat{\mathbf{Q}}$ constructed by the rank centrality algorithm satisfies

$$\|\hat{\mathbf{Q}} - \mathbf{Q}\|_2 \leq \frac{\epsilon}{2} \left(\frac{\pi_{\min}}{\pi_{\max}} \right)^{3/2} P_{\min}. \quad (11)$$

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

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

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

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

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

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

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 $1 - \delta$,

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

But

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

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

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

This proves the result. \square

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 π_i is an optimal permutation w.r.t. the pairwise disagreement error, i.e. we have

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

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

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

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. $\text{argsort}(\mathbf{f}^*) = \text{argsort}(\boldsymbol{\pi})$. \square

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*

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

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

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

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 $|\ln(u) - \ln(u')| \leq \frac{\epsilon}{u - \epsilon}$, thus proving the result. □

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

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

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

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

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

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

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

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

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

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

This proves the result. \square

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

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

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

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

This proves the result. \square

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

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

Thus \mathbf{P} satisfies the LN condition. \square

Proof of Theorem 14

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

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

This proves the result. \square

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 $1 - \delta$,

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

But

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

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

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

This proves the result. \square

Proof of Theorem 16

Proof. We have

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

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

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

where we denote

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

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

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

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

$$\left| \frac{1}{n} \sum_{k=1}^n (P_{kj} - P_{kj}^{\widehat{\boldsymbol{\theta}}}) \right| \leq \frac{r_{\min}}{3} \quad \forall i. \quad (14)$$

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

$$\sum_{k=1}^n P_{kj}^{\hat{\theta}} > \sum_{k=1}^n P_{ki}^{\hat{\theta}} \iff \sum_{k=1}^n P_{kj} > \sum_{k=1}^n P_{ki} \iff f_j^* > f_i^*. \quad (15)$$

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

$$\hat{w}_j > \hat{w}_i \iff \hat{\theta}_j > \hat{\theta}_i \iff \sum_{k=1}^n P_{kj}^{\hat{\theta}} > \sum_{k=1}^n P_{ki}^{\hat{\theta}} \quad (16)$$

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

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

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

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

This proves the result. \square

Proof of Proposition 19

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

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

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

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

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

Thus \mathbf{P} satisfies the GLN condition. \square

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 $\text{sign}(\hat{P}_{ji} - \hat{P}_{ij})$ in $S_{\hat{\mathbf{P}}}$ is the same as the corresponding label $\text{sign}(P_{ji} - P_{ij})$ in $S_{\mathbf{P}}$. We have,

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

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 γ , this yields that with probability at least $1 - \frac{\delta}{2}$,

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

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 α as the corresponding point $(\mathbf{P}_i - \mathbf{P}_j)$ in $S_{\mathbf{P}}$. We have,

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

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

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

By definition, r_{\min}^α is the smallest Euclidean distance of any point $(\mathbf{P}_i - \mathbf{P}_j)$ to the hyperplane defined by α ; 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 α 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 α ; in this case, the SVM-RankAggregation algorithm produces a vector $\widehat{\alpha}$ that correctly classifies all examples in $S_{\widehat{\mathbf{P}}}$, i.e. satisfies $z_{ij} \widehat{\alpha}^\top (\widehat{\mathbf{P}}_i - \widehat{\mathbf{P}}_j) > 0 \quad \forall i < j$ (where $z_{ij} = \text{sign}(P_{ji} - P_{ij})$), and it can be verified that $\widehat{\alpha}$ must then also satisfy $z_{ij} \widehat{\alpha}^\top (\mathbf{P}_i - \mathbf{P}_j) > 0 \quad \forall i < j$, so that $\text{argsort}(\widehat{\alpha}) \subseteq \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}_{\mu, \mathbf{P}}^{\text{PD}}[\sigma]$. This yields that with probability at least $1 - \delta$,

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

□