Ranking from Stochastic Pairwise Preferences: Recovering Condorcet Winners and Tournament Solution Sets at the Top

Supplementary Material

Before we give the appendix material, we note the following: In general, all the algorithms described (both previous algorithms and those proposed in this paper) can be run either on a true preference matrix P as input (where $P_{ij}+P_{ji}=1$ for all $i \neq j$) or on an empirical preference matrix $\widehat{\mathbf{P}}$ as input (where there might be some (i,j) pairs for which $\widehat{P}_{ij}=\widehat{P}_{ji}=0$). The notation $\overline{\mathbf{P}}$ used in describing the algorithms allows for both possibilities. All the theorems we give apply to the empirical preference matrix and therefore use the notation $\widehat{\mathbf{P}}$; on the other hand, the counter-examples we give in Examples show that even if the corresponding algorithms are run on the true matrix $\mathbf{P}$, the algorithms can fail to satisfy the desired property, and therefore use the notation $\mathbf{P}$.

A. Supplement to Section 3 (Related Work and Existing Results)

For completeness, here we include descriptions of the Rank Centrality (RC), Matrix Borda (MB) and SVM-RankAggregation (SVM-RA) algorithms that form the backdrop to our work. For the SMV-RA algorithm, we will need the following definition:

Definition 12 ($\overline{\mathbf{P}}$-Induced Dataset (Rajkumar & Agarwal, 2014)**).** *Let $\overline{\mathbf{P}} \in [0,1]^{n\times n}$ satisfy the following conditions: (i) for all $i \neq j$: $\overline{p}_{ij}+\overline{p}_{ji}=1$ or $\overline{p}_{ij}=\overline{p}_{ji}=0$ and (ii) for every i: $\overline{p}_{ii}=0$. Define the $\overline{\mathbf{P}}$-*induced dataset *$S_{\overline{\mathbf{P}}} = \{\mathbf{v}_{ij}, z_{ij}\}_{i<j:\overline{p}_{ij}+\overline{p}_{ji}=1}$ as consisting of the vectors $\mathbf{v}_{ij}=(\overline{\mathbf{p}}_i-\overline{\mathbf{p}}_j)\in\mathbb{R}^n$ for every pair (i,j) such that $i<j$ and $\overline{p}_{ij}+\overline{p}_{ji}=1$, where $\overline{\mathbf{p}}_i$ denotes the i-th column of $\overline{\mathbf{P}}$, together with binary labels $z_{ij}=\mathrm{sign}(\overline{p}_{ji}-\overline{p}_{ij})\in\{\pm1\}$.*

Algorithm 4 Rank Centrality (RC) Algorithm ((Negahban et al., 2012))

Input: Pairwise comparison matrix $\overline{\mathbf{P}} \in [0,1]^{n\times n}$ satisfying the following conditions:
(i) for all $i \neq j$: $\overline{p}_{ij}+\overline{p}_{ji}=1$ or $\overline{p}_{ij}=\overline{p}_{ji}=0$
(ii) for every i: $\overline{p}_{ii}=0$

Construct an empirical Markov chain with transition probability matrix $\breve{\overline{\mathbf{P}}}$ as follows:

$$\breve{\overline{p}}_{ij} = \begin{cases} \frac{1}{n}\overline{p}_{ij} & \text{if } i \neq j \\ 1-\frac{1}{n}\sum_{k\neq i}\overline{p}_{ik} & \text{if } i=j. \end{cases}$$

Compute $\overline{\boldsymbol{\pi}}$, the stationary probability vector of $\breve{\overline{\mathbf{P}}}$

Output: Permutation $\overline{\sigma}_{\text{RC}} \in \operatorname{argsort}(\overline{\boldsymbol{\pi}})$

Algorithm 5 Matrix Borda (MB) Algorithm ((Rajkumar & Agarwal, 2014))

Input: Pairwise comparison matrix $\overline{\mathbf{P}} \in [0,1]^{n\times n}$ satisfying the following conditions:
(i) for all $i \neq j$: $\overline{p}_{ij}+\overline{p}_{ji}=1$ or $\overline{p}_{ij}=\overline{p}_{ji}=0$
(ii) for every i: $\overline{p}_{ii}=0$

For $i=1$ to n: $\overline{f}_i=\frac{1}{n}\sum_{k=1}^{n}\overline{p}_{ki}$

Output: Permutation $\overline{\sigma}_{\text{MB}} \in \operatorname{argsort}(\overline{\mathbf{f}})$

B. Supplement to Section 4 (Performance of Rank Centrality, Matrix Borda and SVM-RankAggregation Algorithms under Preferences with Cycles) and Example 3

Additional details for Examples 1–3. For completeness, we first give here the permutations output by different algorithms for the preference matrices given in Examples 1, 2 and 3. For the PM algorithm, the value of the parameter c is chosen as the maximum of the values prescribed by Theorem 10 and 11.

Algorithm 6 SVM-RankAggregation (SVM-RA) Algorithm ((Rajkumar & Agarwal, 2014))

Input: Pairwise comparison matrix $\overline{\mathbf{P}} \in [0,1]^{n\times n}$ satisfying the following conditions:
(i) for all $i \neq j$: $\overline{p}_{ij} + \overline{p}_{ji} = 1$ or $\overline{p}_{ij} = \overline{p}_{ji} = 0$
(ii) for every i: $\overline{p}_{ii} = 0$

Construct $\overline{\mathbf{P}}$-induced dataset $S_{\overline{\mathbf{P}}}$ (see Definition 12)

If $S_{\overline{\mathbf{P}}}$ is linearly separable by hyperplane through origin, then
train hard-margin linear SVM on $S_{\overline{\mathbf{P}}}$; obtain $\overline{\boldsymbol{\alpha}} \in \mathbb{R}^n$
else
train soft-margin linear SVM (with any suitable value for regularization parameter) on $S_{\overline{\mathbf{P}}}$; obtain $\overline{\boldsymbol{\alpha}} \in \mathbb{R}^n$

For $i = 1$ to n: $\overline{f}_i = \sum_{k=1}^{n} \overline{\alpha}_k \overline{p}_{ki}$
Output: Permutation $\overline{\sigma}_{\text{SVM-RA}} \in \text{argsort}(\overline{\mathbf{f}})$

For the matrix in Example 1, we have the following:

$$\begin{aligned}
\sigma_{\text{RC}} &= (2\ 1\ 6\ 3\ 4\ 7\ 5) \\
\sigma_{\text{MB}} &= (2\ 1\ 6\ 3\ 4\ 7\ 5) \\
\sigma_{\text{SVM-RA}} &= (7\ 6\ 2\ 1\ 5\ 4\ 3) \\
\sigma_{\text{MC}} &= (1\ 2\ 6\ 3\ 4\ 7\ 5) \\
\sigma_{\text{UM}} &= (1\ 2\ 3\ 4\ 5\ 6\ 7) \\
\sigma_{\text{PM}} &= (1\ 2\ 6\ 7\ 3\ 4\ 5)
\end{aligned}$$

For the matrix in Example 2, we have the following:

$$\begin{aligned}
\sigma_{\text{RC}} &= (4\ 3\ 1\ 2\ 6\ 5) \\
\sigma_{\text{MB}} &= (4\ 1\ 2\ 3\ 6\ 5) \\
\sigma_{\text{SVM-RA}} &= (4\ 2\ 6\ 1\ 3\ 5) \\
\sigma_{\text{MC}} &= (1\ 2\ 3\ 4\ 6\ 5) \\
\sigma_{\text{UM}} &= (1\ 2\ 3\ 4\ 5\ 6) \\
\sigma_{\text{PM}} &= (1\ 2\ 3\ 6\ 5\ 4)
\end{aligned}$$

For the matrix in Example 3, we have the following:

$$\begin{aligned}
\sigma_{\text{RC}} &= (4\ 6\ 5\ 7\ 1\ 2\ 3\ 8) \\
\sigma_{\text{MB}} &= (4\ 6\ 5\ 7\ 1\ 2\ 3\ 8) \\
\sigma_{\text{SVM-RA}} &= (4\ 6\ 5\ 2\ 1\ 7\ 3\ 8) \\
\sigma_{\text{MC}} &= (2\ 4\ 6\ 5\ 1\ 7\ 3\ 8) \\
\sigma_{\text{UM}} &= (5\ 2\ 4\ 6\ 7\ 3\ 1\ 8) \\
\sigma_{\text{PM}} &= (5\ 4\ 6\ 2\ 7\ 3\ 1\ 8)
\end{aligned}$$

We next give additional examples showing that the MB algorithm can fail to rank the Condorcet winner/Copeland set/Markov set at the top even in $\mathcal{P}^{\text{DAG}} \setminus \mathcal{P}^{\text{LN}}$, and the RC algorithm can fail to do so even in $\mathcal{P}^{\text{LN}} \setminus \mathcal{P}^{\text{BTL}}$.

Example 6 (MB algorithm can fail to rank the Condorcet winner/Copeland set/Markov set at the top even in $\mathcal{P}^{\text{DAG}} \setminus \mathcal{P}^{\text{LN}}$)**.** *Let $n = 3$, and consider*

$$\mathbf{P} = \begin{bmatrix} 0 & 0.4 & 0.4 \\ 0.6 & 0 & 0.1 \\ 0.6 & 0.9 & 0 \end{bmatrix}$$

Here $\text{CW}(\mathbf{P}) = 1, \text{CO}(\mathbf{P}) = \text{MA}(\mathbf{P}) = \{1\}$*, but the permutation σ_{MB} produced by running* MB *on* $\mathbf{P}$ *is* $\sigma_{\text{MB}} = (2\ 1\ 3)$.

Example 7 (RC algorithm can fail to rank the Condorcet winner/Copeland set/Markov set at the top even in $\mathcal{P}^{\text{LN}} \setminus \mathcal{P}^{\text{BTL}}$)**.** *Let* $n = 8$*, and consider*

$$\mathbf{P} = \begin{bmatrix} 0 & 0.49 & 0.3095 & 0.3 & 0.3 & 0.3 & 0.3 & 0.3 \\ 0.51 & 0 & 0.1 & 0.49 & 0.3 & 0.3 & 0.3 & 0.3 \\ 0.6905 & 0.9 & 0 & 0.2 & 0.1 & 0.3 & 0.3 & 0.3 \\ 0.7 & 0.51 & 0.8 & 0 & 0.1 & 0.3 & 0.3 & 0.3 \\ 0.7 & 0.7 & 0.9 & 0.9 & 0 & 0.3 & 0.3 & 0.3 \\ 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0 & 0.3 & 0.3 \\ 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0 & 0.3 \\ 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 0 \end{bmatrix}$$

Here $\text{CW}(\mathbf{P}) = 1$, $\text{CO}(\mathbf{P}) = \text{MA}(\mathbf{P}) = \{1\}$, *but the permutation* σ_{RC} *produced by running the* RC *algorithm on* $\mathbf{P}$ *is* $\sigma_{\text{RC}} = (2\ 1\ 3\ 4\ 6\ 5\ 7\ 8)$.

C. Proofs of Results in Sections 5–6

The overall strategy followed by our proofs is largely similar to that of (Rajkumar & Agarwal, 2014): namely, that the empirical pairwise comparison matrix $\widehat{\mathbf{P}}$ concentrates around the true pairwise preference matrix $\mathbf{P}$, and that when $\widehat{\mathbf{P}}$ becomes sufficiently close to $\mathbf{P}$, the specific algorithm satisfies the desired property. However the details differ considerably depending on the algorithm and property of interest. The proofs for the Matrix Copeland (MC) and Unweighted Markov (UM) algorithms are relatively straightforward; the proofs for the Parametrized Markov (PM) algorithm, on the other hand, require additional tools. In particular, as noted in the main text, the proof of Theorem 9 involves reasoning about a new property of preference matrices that we term the *restricted low-noise* (RLN) property; the proofs of Theorems 10–11 make use of the Cho-Meyer perturbation bound for Markov chains. Details follow.

C.1. Proof of Theorem 5

Proof. Let m satisfy the given assumption. Then by Lemma 4, we have with probability at least $1 - \delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\gamma_{\min}}{2} \text{ for all } i, j.$$

Under this event, we have for all $i, j : p_{ij} > \frac{1}{2} \iff \widehat{p}_{ij} > \frac{1}{2}$, and therefore $G_{\mathbf{P}} = G_{\widehat{\mathbf{P}}}$, thus giving for all $i \in [n]$:

$$\widehat{f}_i \;=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}(\widehat{p}_{ij} < \tfrac{1}{2}) \;=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}(p_{ij} < \tfrac{1}{2}) \;=\; \frac{d(i)}{n}\,.$$

The first claim follows. To prove the second claim, note that if $\mathbf{P} \in \mathcal{P}^{\text{DAG}}$, we have for any $i, j,\ i \succ_{\mathbf{P}} j \implies d(i) > d(j)$. The claim follows. □

C.2. Proof of Theorem 6

Proof. Let m satisfy the given assumption. Then by Lemma 4, we have with probability at least $1 - \delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\gamma_{\text{TC}}}{2} \text{ for all } i, j.$$

Under this event, we have for all $i \in \text{TC}(\mathbf{P}), j \notin \text{TC}(\mathbf{P})$, $\widehat{p}_{ij} < \frac{1}{2}$ (and therefore the edge (i, j) is present $G_{\widehat{\mathbf{P}}}$), thus giving

$$\text{for all } i \in \text{TC}(\mathbf{P}):\quad \widehat{f}_i \;=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}(\widehat{p}_{ij} < \tfrac{1}{2}) \;\geq\; \tfrac{n - |\text{TC}(\mathbf{P})|}{n} \;\text{ and} \tag{2}$$

$$\text{for all } j \notin \text{TC}(\mathbf{P}):\quad \widehat{f}_j \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\widehat{p}_{ji} < \tfrac{1}{2}) \;\leq\; \tfrac{n - |\text{TC}(\mathbf{P})| - 1}{n} \;(\text{since } \widehat{p}_{jj} = 0) \tag{3}$$

Thus under the above event, we have $\widehat{f}_i > \widehat{f}_j$ for all $i \in \text{TC}(\mathbf{P}), j \notin \text{TC}(\mathbf{P})$. The claim follows. □

C.3. Proof of Theorem 7

Proof. Let m satisfy the given assumption. Then by Lemma 4, we have with probability at least $1-\delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\gamma_{\min}}{2} \text{ for all } i, j. \tag{4}$$

Under this event, we have for all $i, j : p_{ij} > \frac{1}{2} \iff \widehat{p}_{ij} > \frac{1}{2}$, and therefore $G_{\mathbf{P}} = G_{\widehat{\mathbf{P}}}$; it is easy to verify that in this case, the Markov chain $\breve{\widehat{\mathbf{H}}}$ constructed by the UM algorithm is the same as the Markov chain $\widetilde{\mathbf{P}}$ used to define the Markov set $\mathrm{MA}(\mathbf{P})$, and therefore their corresponding stationary distributions $\widehat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$ are also the same. The claim follows. □

C.4. Proof of Theorem 8

Proof. Let m satisfy the given assumption. Then by Lemma 4, we have with probability at least $1-\delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\gamma_{\min}}{2} \text{ for all } i, j. \tag{5}$$

Under this event, we have for all $i, j : \widehat{p}_{ij} > \frac{1}{2} \iff p_{ij} > \frac{1}{2}$. In particular, we have for all $i \in \mathrm{TC}(\mathbf{P}), j \notin \mathrm{TC}(\mathbf{P})$, $\widehat{p}_{ij} < \frac{1}{2}$, and thus the Markov chain $\breve{\widehat{\mathbf{H}}}$ constructed by the UM algorithm satisfies

$$\breve{\widehat{h}}_{ij} = \begin{cases} 0 & \text{if } i \in \mathrm{TC}(\mathbf{P}), j \notin \mathrm{TC}(\mathbf{P}) \\ \frac{1}{n} & \text{if } i \notin \mathrm{TC}(\mathbf{P}), j \in \mathrm{TC}(\mathbf{P})\,. \end{cases} \tag{6}$$

This implies in particular that the stationary distribution $\widehat{\boldsymbol{\pi}}$ of $\breve{\widehat{\mathbf{H}}}$ satisfies

$$\widehat{\pi}_i = 0 \quad \text{for all } i \notin \mathrm{TC}(\mathbf{P})\,.$$

Moreover, for all $i, j \in \mathrm{TC}(\mathbf{P})$, we have $\widehat{p}_{ij} > \frac{1}{2} \iff p_{ij} > \frac{1}{2}$ as above, and therefore the items in $\mathrm{TC}(\mathbf{P})$ form a recurrent class in the above Markov chain; this implies that

$$\widehat{\pi}_i > 0 \quad \text{for all } i \in \mathrm{TC}(\mathbf{P})\,.$$

Thus under the above event, we have $\widehat{\pi}_i > \widehat{\pi}_j$ for all $i \in \mathrm{TC}(\mathbf{P}), j \notin \mathrm{TC}(\mathbf{P})$. The claim follows. □

C.5. Auxiliary Results Needed for Proof of Theorem 9

Theorem 9 claims that for $\mathbf{P} \in \mathcal{P}^{\mathrm{BTL}}$, the PM algorithm using any $1 \le c < \infty$ (for sufficiently large sample size, with high probability) recovers an optimal permutation w.r.t. PD error. As noted earlier, we will prove the result of Theorem 9 for a slightly larger set of preference matrices than $\mathcal{P}^{\mathrm{BTL}}$, namely for all $\mathbf{P}$ satisfying the *restricted low-noise* (RLN) property, defined as follows:

$$\mathcal{P}_n^{\mathrm{RLN}} = \Big\{ \mathbf{P} \in \mathcal{P}^n : \forall\, i \neq j \neq k : i \succ_{\mathbf{P}} j \implies p_{kj} < p_{ki} \Big\}\,.$$

We will show below the following three results:

(i) $\mathcal{P}^{\mathrm{BTL}} \subseteq \mathcal{P}^{\mathrm{RLN}}$ (Proposition 13).

(ii) For any $1 \le c < \infty$, $\mathcal{P}^{\mathrm{RLN}}$ is closed under g_c, i.e. $\mathbf{P} \in \mathcal{P}^{\mathrm{RLN}} \implies \mathbf{H}^c = g_c(\mathbf{P}) \in \mathcal{P}^{\mathrm{RLN}}$ (Proposition 14).

(iii) For any $\mathbf{P} \in \mathcal{P}^{\mathrm{RLN}}$, the permutation σ_{RC} produced by running the RC algorithm on $\mathbf{P}$ satisfies $\sigma_{\mathrm{RC}} \in \mathrm{argmin}_{\sigma \in \mathcal{S}_n} \mathrm{er}^{\mathrm{PD}}_{\mathbf{P}}[\sigma]$ (Theorem 15).

These results will imply that for any $\mathbf{P} \in \mathcal{P}^{\mathrm{RLN}}$ (and therefore in particular for any $\mathbf{P} \in \mathcal{P}^{\mathrm{BTL}}$) and any $1 \le c < \infty$, the permutation σ_{PM} produced by running the PM algorithm on $\mathbf{P}$ (i.e. the RC algorithm on $\mathbf{H}^c = g_c(\mathbf{P})$) satisfies

$$\sigma_{\mathrm{PM}} \in \mathrm{argmin}_{\sigma \in \mathcal{S}_n} \mathrm{er}^{\mathrm{PD}}_{\mathbf{H}^c}[\sigma]\,.$$

Moreover, since $p_{ij} > \frac{1}{2} \Longleftrightarrow g_c(p_{ij}) > \frac{1}{2}$, we have $\text{argmin}_{\sigma \in S_n} \text{er}^{\text{PD}}_{\mathbf{H}^c}[\sigma] = \text{argmin}_{\sigma \in S_n} \text{er}^{\text{PD}}_{\mathbf{P}}[\sigma]$, which further gives that for $\mathbf{P}$ and c as above,

$$\sigma_{\text{PM}} \in \text{argmin}_{\sigma \in S_n} \text{er}^{\text{PD}}_{\mathbf{P}}[\sigma]\,.$$

We will then be able to argue that for sufficiently large sample size m, when $\widehat{\mathbf{P}}$ is sufficiently close to $\mathbf{P}$, the permutation $\widehat{\sigma}_{\text{PM}}$ returned by running the PM algorithm on $\widehat{\mathbf{P}}$ will also minimize the PD error w.r.t. $\mathbf{P}$, which will prove Theorem 9.

Proposition 13. $\mathcal{P}_n^{\text{BTL}} \subseteq \mathcal{P}_n^{\text{RLN}}$.

Proof. We have,

$$\begin{aligned}
\mathbf{P} \in \mathcal{P}_n^{\text{BTL}} &\implies \text{there exists } \mathbf{w} \in \mathbb{R}_+^n : \text{ for all } i \neq j\,,\;\; p_{ij} = \frac{w_j}{w_i + w_j} \\
&\implies \text{for any } i,j, \text{ if } p_{ij} < p_{ji}, \text{ then } w_i > w_j \\
&\implies \text{for any } i,j, \text{ for any } k \neq \{i,j\}, p_{ik} = \frac{w_k}{w_i + w_k} > p_{jk} = \frac{w_k}{w_j + w_k} \\
&\implies \mathbf{P} \in \mathcal{P}_n^{\text{RLN}}\,.
\end{aligned}$$

This proves the claim. ☐

Proposition 14. *Let* $\mathbf{P} \in \mathcal{P}_n^{\text{RLN}}$*. Let* $1 \le c < \infty$ *and define* $\mathbf{H}^c$ *as* $h^c_{ij} = g_c(p_{ij})$ *for all* $i \neq j$ *and* $h^c_{ii} = 0$ *for all* i*. Then* $\mathbf{H}^c \in \mathcal{P}_n^{\text{RLN}}$*.*

Proof. Let $i, j \in [n]$. We have,

$$\begin{aligned}
i \succ_{\mathbf{H}^c} j &\implies h^c_{ij} < h^c_{ji} \\
&\implies \frac{p^c_{ij}}{p^c_{ij} + p^c_{ji}} < \frac{p^c_{ji}}{p^c_{ji} + p^c_{ij}} \\
&\implies p_{ij} < p_{ji} \\
&\implies p_{ki} > p_{kj} \text{ for all } k \neq \{i,j\} \quad (\text{since } \mathbf{P} \in \mathcal{P}^{\text{RLN}}) \\
&\implies p_{ki} + 1 - p_{jk}p_{ik} > p_{kj} + 1 - p_{jk}p_{ik} \text{ for all } k \neq \{i,j\} \\
&\implies p_{ki}p_{jk} > p_{ik}p_{kj} \text{ for all } k \neq \{i,j\} \\
&\implies \frac{p_{ik}}{p_{ki}} < \frac{p_{kj}}{p_{jk}} \text{ for all } k \neq \{i,j\} \\
&\implies \left(\frac{p_{ik}}{p_{ki}}\right)^c < \left(\frac{p_{kj}}{p_{jk}}\right)^c \text{ for all } c \ge 1, \text{ for all } k \neq \{i,j\} \\
&\implies g_c(p_{ki}) > g_c(p_{kj}) \text{ for all } c \ge 1, \text{ for all } k \neq \{i,j\} \\
&\implies h^c_{ki} > h^c_{jk} \text{ for all } c \ge 1, \text{ for all } k \neq \{i,j\}
\end{aligned}$$

Thus we have $\mathbf{H}^c \in \mathcal{P}_n^{\text{RLN}}$. ☐

Theorem 15. *Let* $\mathbf{P} \in \mathcal{P}_n^{\text{RLN}}$*. Then the permutation* σ_{RC} *produced by running the Rank Centrality algorithm on* $\mathbf{P}$ *satisfies*

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

Proof. Let $\breve{\mathbf{P}}$ be the Markov chain constructed from $\mathbf{P}$ by the Rank Centrality algorithm and let $\boldsymbol{\pi}$ be its stationary distribution. Fix any i, j such that $i \succ_{\mathbf{P}} j$. Assume for the sake of contradiction that $\pi_i \le \pi_j$. We have from the stationary distribution equations,

$$\begin{aligned}\pi_i &= \sum_{k=1}^{n} \pi_k \breve{p}_{ki} \\ &= \sum_{k\neq\{i,j\}} \pi_k \breve{p}_{ki} + \pi_i \breve{p}_{ii} + \pi_j \breve{p}_{ji} \\ &= \sum_{k\neq\{i,j\}} \pi_k \frac{p_{ki}}{n} + \pi_i (1 - \sum_{k\neq i} \breve{p}_{ik}) + \pi_j \frac{p_{ji}}{n} \\ &= \sum_{k\neq\{i,j\}} \pi_k \frac{p_{ki}}{n} + \pi_i \Big(\frac{1}{n} + \sum_{k\neq i} \frac{p_{ki}}{n}\Big) + \pi_j \frac{p_{ji}}{n}\end{aligned}$$

Similarly we have

$$\pi_j = \sum_{k\neq\{i,j\}} \pi_k \frac{p_{kj}}{n} + \pi_j \Big(\frac{1}{n} + \sum_{k\neq j} \frac{p_{kj}}{n}\Big) + \pi_i \frac{p_{ij}}{n}$$

Notice that the equations for π_i and π_j are of the form

$$\begin{aligned}\pi_i &= a_1 + b_1 \pi_i + c_1 \pi_j \\ \pi_j &= a_2 + b_2 \pi_j + c_2 \pi_i\end{aligned}$$

where

$$\begin{aligned}a_1 &= \sum_{k\neq\{i,j\}} \pi_k \frac{p_{ki}}{n}\,, \quad b_1 = \frac{1}{n} + \sum_{k\neq i} \frac{p_{ki}}{n}\,, \quad c_1 = \frac{p_{ji}}{n} \\ a_2 &= \sum_{k\neq\{i,j\}} \pi_k \frac{p_{kj}}{n}\,, \quad b_2 = \frac{1}{n} + \sum_{k\neq j} \frac{p_{kj}}{n}\,, \quad c_2 = \frac{p_{ij}}{n}\,.\end{aligned}$$

Since $\mathbf{P} \in \mathcal{P}^{\text{RLN}}$ and $i \succ_{\mathbf{P}} j$, we have that $a_1 \geq a_2, b_1 > b_2$ and $c_1 > c_2$. Observe that $0 < b_1, b_2 \leq 1$. We claim that neither b_1 nor b_2 can be exactly equal to 1. To see this, first consider b_2. We have $b_2 = 1$ only if $p_{kj} = 1$ for all $k \neq j$. But as $i \succ_{\mathbf{P}} j$, we have $p_{ij} < p_{ji}$ and so this cannot happen. Thus $b_2 < 1$. Next consider b_1. We have $b_1 = 1$ only if $p_{ki} = 1$ for all $k \neq i$. When this happens, it is easy to see that $\text{CW}(\mathbf{P}) = i$ and so $\pi_i = 1$ and $\pi_j = 0$ for all $j \neq i$. This contradicts our assumption that $\pi_i \leq \pi_j$. So in what follows, assume $0 < b_1, b_2 < 1$. In that case, we have

$$\begin{aligned}\pi_i &= \frac{a_1}{1-b_1} + \frac{c_1}{1-b_1}\pi_j \\ &> \frac{a_2}{1-b_2} + \frac{c_2}{1-b_2}\pi_i \quad (\text{as } \mathbf{P} \in \mathcal{P}^{\text{RLN}} \text{ and by our assumption that } \pi_i \leq \pi_j) \\ &= \pi_j\end{aligned}$$

which contradicts the assumption that $\pi_i \leq \pi_j$. Thus, we must have that

$$\text{for any } i, j, \quad i \succ_{\mathbf{P}} j \implies \pi_i > \pi_j\,.$$

The claim follows. □

C.6. Proof of Theorem 9

Proof. Let m satisfy the given assumption. Then by Lemma 4, we have with probability at least $1 - \delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\min(\gamma_{\min}, r_{\min})}{2} \text{ for all } i, j\,. \tag{7}$$

Under this event, we have

$$\begin{aligned}\text{for all } i,j\text{:}\quad \widehat{p}_{ij} > \tfrac{1}{2} &\implies p_{ij} > \tfrac{1}{2} \quad (\text{since } |\widehat{p}_{ij} - p_{ij}| < \tfrac{\gamma_{\min}}{2} \text{ for all } i,j) \\ &\implies p_{ki} > p_{kj} \text{ for all } k \neq \{i,j\} \quad (\text{since } \mathbf{P} \in \mathcal{P}^{\text{BTL}} \subseteq \mathcal{P}^{\text{RLN}} \text{ from Proposition 13}) \\ &\implies \widehat{p}_{ki} > \widehat{p}_{kj} \text{ for all } k \neq \{i,j\} \quad (\text{since } |\widehat{p}_{km} - p_{km}| < \frac{r_{\min}}{2} \text{ for all } k,m)\,.\end{aligned}$$

Thus under the above event, we have $\widehat{\mathbf{P}} \in \mathcal{P}^{\text{RLN}}$, and therefore by Proposition 14 we have that the matrix $\widehat{\mathbf{H}}^c$ constructed by the PM algorithm satisfies $\widehat{\mathbf{H}}^c \in \mathcal{P}^{\text{RLN}}$, and by Theorem 15 that

$$\widehat{\sigma}_{\text{PM}} \in \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\widehat{\mathbf{H}}^c}[\sigma]\,. \tag{8}$$

Since under the above event we also have

$$\text{for all } i,j: \quad p_{ij} > \tfrac{1}{2} \iff \widehat{p}_{ij} > \tfrac{1}{2} \iff \widehat{h}^c_{ij} > \tfrac{1}{2}\,,$$

it follows that under this event, $\text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\widehat{\mathbf{H}}^c}[\sigma] = \text{argmin}_{\sigma \in \mathcal{S}_n} \text{er}^{\text{PD}}_{\mathbf{P}}[\sigma]$ and therefore

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

This proves the result. □

C.7. Auxiliary Results Needed for Proofs of Theorems 10–11

The first result we will need for proving Theorems 10–11 is the following perturbation bound for Markov chains:

Theorem 16 ((Cho & Meyer, 2001; Seneta, 1988))**.** *Let $\mathbf{R}$ and $\mathbf{R}'$ be the transition probability matrices of finite Markov chains having unique stationary distributions $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ respectively. Then*

$$\|\boldsymbol{\pi} - \boldsymbol{\pi}'\|_1 \leq \frac{\|\mathbf{R} - \mathbf{R}'\|_\infty}{1 - \tau(\mathbf{R})}\,,$$

where $1 - \tau(\mathbf{R}) = \min_{i \neq j} \sum_{k=1}^n \min\{r_{ik}, r_{jk}\}$.

We will use the above theorem to show the following two results:

(i) For any $\mathbf{P}$ and sufficiently large c, the permutation σ_{PM} produced by running the PM algorithm on $\mathbf{P}$ satisfies $\sigma_{\text{PM}}(i) < \sigma_{\text{PM}}(j)$ for all $i \in \text{MA}(\mathbf{P}), j \notin \text{MA}(\mathbf{P})$ (Theorem 19).

(ii) For any $\mathbf{P}$ and sufficiently large c, the permutation σ_{PM} produced by running the PM algorithm on $\mathbf{P}$ satisfies $\sigma_{\text{PM}}(i) < \sigma_{\text{PM}}(j)$ for all $i \in \text{TC}(\mathbf{P}), j \notin \text{TC}(\mathbf{P})$ (Theorem 20).

In proving these results, we will make use of Theorem 16 to bound the difference between the stationary probability vectors $\boldsymbol{\pi}$ and $\boldsymbol{\pi}^c$ of the Markov chains associated with the matrices $\mathbf{H}^c = g_c(\mathbf{P})$ and $\mathbf{H} = g_\infty(\mathbf{P})$. For this, we will find the following two lemmas useful:

Lemma 17. *Let $\mathbf{P} \in \mathcal{P}_n$. Let $\mathbf{H} = g_\infty(\mathbf{P})$, with $h_{ij} = \mathbf{1}(p_{ij} > 1/2)$, and let $\breve{\mathbf{H}}$ be the Markov chain transition matrix corresponding to $\mathbf{H}$ as constructed by the Rank Centrality algorithm applied to $\mathbf{H}$. Then*

$$1 - \tau(\breve{\mathbf{H}}) \geq \frac{1}{n}\,.$$

Proof. We have,

$$
\begin{aligned}
1-\tau(\breve{\mathbf{H}}) &= \min_{i\neq j}\sum_{k=1}^{n}\min\{\breve{h}_{ik},\breve{h}_{jk}\} \\
&= \min_{i\neq j}\Big(\sum_{k\neq\{i,j\}}\min\{\breve{h}_{ik},\breve{h}_{jk}\}+\min\{\breve{h}_{ii},\breve{h}_{ji}\}+\min\{\breve{h}_{ij},\breve{h}_{jj}\}\Big) \\
&= \min_{i\neq j}\Big(\sum_{k\neq\{i,j\}}\min\{\breve{h}_{ik},\breve{h}_{jk}\}+\breve{h}_{ji}+\breve{h}_{ij}\Big) \text{ (since } \breve{h}_{ii}\geq\breve{h}_{ij},\breve{h}_{ji} \text{ for all } i\neq j) \\
&= \min_{i\neq j}\Big(\sum_{k\neq\{i,j\}}\min\{\breve{h}_{ik},\breve{h}_{jk}\}+\breve{h}_{ji}+\breve{h}_{ij}\Big) \\
&= \min_{i\neq j}\Big(\sum_{k\neq\{i,j\}}\min\{\breve{h}_{ik},\breve{h}_{jk}\}+\frac{1}{n}\Big) \text{ (since } \breve{h}_{ji}+\breve{h}_{ij}=\frac{1}{n} \text{ for all } i\neq j) \\
&\geq \frac{1}{n}\,.
\end{aligned}
$$

□

Lemma 18. *Let* $\mathbf{P}\in\mathcal{P}_n$ *and* $1\leq c<\infty$. *Let* $\mathbf{H}^c=g_c(\mathbf{P})$ *and* $\mathbf{H}=g_\infty(\mathbf{P})$, *and let* $\breve{\mathbf{H}}^c$, $\breve{\mathbf{H}}$ *be the Markov chain transition matrices corresponding to* $\mathbf{H}^c$ *and* $\mathbf{H}$ *as constructed by the Rank Centrality algorithm applied to* $\mathbf{H}^c$ *and* $\mathbf{H}$. *Let* $\alpha_{\min}=\min_{i\neq j:i\succ_{\mathbf{P}}j}\frac{p_{ji}}{p_{ij}}$. *Let* $\epsilon>0$. *If* $c\geq\frac{\ln(2/\epsilon)}{\ln(\alpha_{\min})}$, *then*

$$\|\breve{\mathbf{H}}^c-\breve{\mathbf{H}}\|_\infty<\epsilon$$

Proof. Consider any $i\neq j$. If $p_{ij}<\frac{1}{2}$, we have

$$|\breve{h}^c_{ij}-\breve{h}_{ij}|=\frac{|g_c(p_{ij})-\mathbf{1}(p_{ij}>\frac{1}{2}))|}{n}=\frac{g_c(p_{ij})}{n}=\frac{1}{n(1+(\frac{p_{ji}}{p_{ij}})^c)}\,.$$

Similarly if $p_{ij}>\frac{1}{2}$, we have

$$|\breve{h}^c_{ij}-\breve{h}_{ij}|=\frac{|g_c(p_{ij})-\mathbf{1}(p_{ij}>\frac{1}{2})|}{n}=\frac{|g_c(p_{ij})-1|}{n}=\frac{1}{n(1+(\frac{p_{ij}}{p_{ji}})^c)}\,.$$

In both cases, we have

$$|\breve{h}^c_{ij}-\breve{h}_{ij}|\leq\frac{1}{n(1+\alpha^c_{\min})}\,.$$

Let c satisfy the given assumption. Then we have for all $i\neq j$,

$$|\breve{h}^c_{ij}-\breve{h}_{ij}|\leq\frac{\epsilon}{2n}\,.$$

Moreover, for all i, we have

$$|\breve{h}^c_{ii}-\breve{h}_{ii}|=\Big|\sum_{k\neq i}(\breve{h}^c_{ik}-\breve{h}_{ik})\Big|\leq\sum_{k\neq i}|\breve{h}^c_{ik}-\breve{h}_{ik}|\leq\frac{\epsilon}{2}\,.$$

Thus

$$\|\breve{\mathbf{H}}^c-\breve{\mathbf{H}}\|_\infty=\max_i\sum_k|\breve{h}^c_{ik}-\breve{h}_{ik}|\leq\max_i\left(|\breve{h}^c_{ii}-\breve{h}_{ii}|+\sum_{k\neq i}|\breve{h}^c_{ik}-\breve{h}_{ik}|\right)\leq\frac{\epsilon}{2}+\frac{(n-1)\epsilon}{2n}<\epsilon\,.$$

□

Theorem 19. *Let* $\mathbf{P} \in \mathcal{P}_n$. *Let* $\mathbf{H} = g_\infty(\mathbf{P})$, *and let* $\breve{\mathbf{H}}$ *be the Markov chain transition matrix corresponding to* $\mathbf{H}$ *as constructed by the Rank Centrality algorithm applied to* $\mathbf{H}$. *Let* $\boldsymbol{\pi}$ *be the stationary probability vector of* $\breve{\mathbf{H}}$. *Let* $\beta_2(\mathbf{P}) = \max_{i\in[n]} \pi_i - \max_{k\notin \mathrm{MA}(\mathbf{P})} \pi_k$ *and* $\alpha_{\min} = \min_{i,j:i\succ_{\mathbf{P}} j} \frac{p_{ji}}{p_{ij}}$. *If* $c \geq \frac{\ln(4n/\beta_2(\mathbf{P}))}{\ln(\alpha_{\min})}$, *then the permutation* σ_{PM} *produced by running the PM algorithm on* $\mathbf{P}$ *satisfies*

$$\sigma_{\mathrm{PM}}(i) < \sigma_{\mathrm{PM}}(j) \text{ for all } i \in \mathrm{MA}(\mathbf{P}), j \notin \mathrm{MA}(\mathbf{P})\,.$$

Proof. Let c satisfy the given assumption. Let $\breve{\mathbf{H}}^c$ be the Markov chain transition matrix corresponding to $\mathbf{H}^c = g_c(\mathbf{P})$ as constructed by the Rank Centrality algorithm applied to $\mathbf{H}^c$. Let $\boldsymbol{\pi}^c$ be the stationary probability vector of $\breve{\mathbf{H}}^c$. By Lemma 18, we have

$$\|\breve{\mathbf{H}}^c - \breve{\mathbf{H}}\|_\infty \leq \frac{\beta_2(\mathbf{P})}{2n}\,.$$

By Theorem 16 and Lemma 17, it then follows that

$$\|\boldsymbol{\pi}^c - \boldsymbol{\pi}\|_\infty \leq \|\boldsymbol{\pi}^c - \boldsymbol{\pi}\|_1 \leq \frac{\beta_2(\mathbf{P})}{2}\,.$$

By definition of $\beta_2(\mathbf{P})$ and $\mathrm{MA}(\mathbf{P})$, this implies in particular that

$$\pi^c_i > \pi^c_j \text{ for all } i \in \mathrm{MA}(\mathbf{P}), j \notin \mathrm{MA}(\mathbf{P})\,.$$

The claim follows. □

Theorem 20. *Let* $\mathbf{P} \in \mathcal{P}_n$. *Let* $\alpha_{\min} = \min_{i,j:i\succ_{\mathbf{P}} j} \frac{p_{ji}}{p_{ij}}$. *If* $c \geq \frac{(n+1)\ln(4n)}{\ln(\alpha_{\min})}$, *then the permutation* σ_{PM} *produced by the running the PM algorithm on* $\mathbf{P}$ *satisfies*

$$\sigma_{\mathrm{PM}}(i) < \sigma_{\mathrm{PM}}(j) \text{ for all } i \in \mathrm{TC}(\mathbf{P}), j \notin \mathrm{TC}(\mathbf{P})\,.$$

Proof. Let c satisfy the given assumption. Let $\breve{\mathbf{H}}^c$ and $\breve{\mathbf{H}}$ be the Markov chain transition matrices corresponding to $\mathbf{H}^c = g_c(\mathbf{P})$ and $\mathbf{H} = g_\infty(\mathbf{P})$ as constructed by the Rank Centrality algorithm applied to $\mathbf{H}^c$ and $\mathbf{H}$. Let $\boldsymbol{\pi}^c, \boldsymbol{\pi}$ be the stationary probability vectors of $\breve{\mathbf{H}}^c, \breve{\mathbf{H}}$. By Lemma 18, we have

$$\|\breve{\mathbf{H}}^c - \breve{\mathbf{H}}\|_\infty \leq \frac{1}{2n^{n+1}}\,.$$

By Theorem 16 and Lemma 17, it then follows that

$$\|\boldsymbol{\pi}^c - \boldsymbol{\pi}\|_\infty \leq \|\boldsymbol{\pi}^c - \boldsymbol{\pi}\|_1 \leq \frac{1}{2n^n}\,.$$

By definition of $\mathbf{H}$, we have that $\pi_i > 0$ if and only if $i \in \mathrm{TC}(\mathbf{P})$. Now, define

$$\pi^{\mathrm{TC}}_{\max} = \max_{i\in\mathrm{TC}(\mathbf{P})} \pi_i$$

$$\pi^{\mathrm{TC}}_{\min} = \min_{i\in\mathrm{TC}(\mathbf{P})} \pi_i\,,$$

and let

$$m_0 \in \mathrm{argmin}_{i\in\mathrm{TC}(\mathbf{P})} \pi_i\,.$$

If $|\,\mathrm{TC}(\mathbf{P})| = 1$, then clearly m_0 is the unique Condorcet winner and $\pi_{m_0} = 1$, with $\pi_j = 0$ for $j \neq m_0$; in this case

$$\pi^{\mathrm{TC}}_{\min} = 1 \geq \frac{1}{n^n}\,.$$

If $|\,\mathrm{TC}(\mathbf{P})| > 1$, then it must be the case that $m_0 \succ_{\mathbf{P}} m_1$ for some $m_1 \in \mathrm{TC}(\mathbf{P})$ (otherwise $m_0 \notin \mathrm{TC}(\mathbf{P})$). From the stationary equations, we have that

$$\pi^{\mathrm{TC}}_{\min} = \pi_{m_0} = \sum_{k=1}^{n} \pi_k \breve{h}_{km_0} \geq \pi_{m_1}\breve{h}_{m_1m_0} \geq \frac{\pi_{m_1}}{n} \quad \text{(from definition of } \breve{\mathbf{H}} \text{ and since } m_0 \succ_{\mathbf{P}} m_1).$$

If $|\,\mathrm{TC}(\mathbf{P})| = 2$, then $\pi_{m_1} = \pi_{\max} \geq \frac{1}{n}$, which again gives

$$\pi^{\mathrm{TC}}_{\min} \geq \frac{1}{n^2} \geq \frac{1}{n^n}\,.$$

One can similarly show that regardless of the size of $\mathrm{TC}(\mathbf{P})$, we have

$$\pi^{\mathrm{TC}}_{\min} \geq \frac{1}{n^n}\,.$$

Thus the above result gives us

$$\|\boldsymbol{\pi}^c - \boldsymbol{\pi}\|_\infty \leq \frac{\pi^{\mathrm{TC}}_{\min}}{2}\,.$$

By definition of $\pi^{\mathrm{TC}}_{\min}$, this implies in particular that

$$\pi^c_i > \pi^c_j \;\text{ for all } i \in \mathrm{TC}(\mathbf{P}), j \notin \mathrm{TC}(\mathbf{P})\,.$$

The claim follows. □

C.8. Proof of Theorem 10

Proof. Let c and m satisfy the given assumptions. Applying Theorem 19 to $\widehat{\mathbf{P}}$, we have that $\widehat{\sigma}_{\mathrm{PM}}$ satisfies

$$\widehat{\sigma}_{\mathrm{PM}}(i) < \widehat{\sigma}_{\mathrm{PM}}(j) \;\text{ for all } i \in \mathrm{MA}(\widehat{\mathbf{P}}), j \notin \mathrm{MA}(\widehat{\mathbf{P}})\,.$$

Now by Lemma 4, we have with probability at least $1 - \delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\gamma_{\min}}{2} \text{ for all } i, j. \tag{9}$$

Under this event, we have

$$\text{for all } i, j: \quad p_{ij} > \tfrac{1}{2} \Longleftrightarrow \widehat{p}_{ij} > \tfrac{1}{2}\,. \tag{10}$$

Thus under the above event, we have $\mathrm{MA}(\widehat{\mathbf{P}}) = \mathrm{MA}(\mathbf{P})$. The claim follows. □

C.9. Proof of Theorem 11

Proof. Let c and m satisfy the given assumptions. Applying Theorem 20 to $\widehat{\mathbf{P}}$, we have that $\widehat{\sigma}_{\mathrm{PM}}$ satisfies

$$\widehat{\sigma}_{\mathrm{PM}}(i) < \widehat{\sigma}_{\mathrm{PM}}(j) \;\text{ for all } i \in \mathrm{TC}(\widehat{\mathbf{P}}), j \notin \mathrm{TC}(\widehat{\mathbf{P}})\,.$$

Now by Lemma 4, we have with probability at least $1 - \delta$, the following event holds:

$$|p_{ij} - \widehat{p}_{ij}| < \frac{\gamma_{\min}}{2} \text{ for all } i, j. \tag{11}$$

Under this event, we have

$$\text{for all } i, j: \quad p_{ij} > \tfrac{1}{2} \Longleftrightarrow \widehat{p}_{ij} > \tfrac{1}{2}\,. \tag{12}$$

Thus under the above event, we have $\mathrm{TC}(\widehat{\mathbf{P}}) = \mathrm{TC}(\mathbf{P})$. The claim follows. □