Aggregating Incomplete and Noisy Rankings: Supplementary Material

Appendices

A Proof of Theorem 1

The proof of Theorem 1 follows the same steps as the proof provided by Caragiannis et al. [2013] for the upper bound of the sample complexity of finding the central ranking given complete Mallows samples.

Assume that $\mathcal{S}$ is $p-$frequent and $\Pi \sim \mathcal{M}^{\mathcal{S}}_{\pi_0,\beta}$. Also, without loss of generality, let π_0 be the identity permutation id. For the following, we denote with $n_{i\succ j} = n_{i\succ j}[\Pi]$ the number of samples where $i \succ j$, for any $i, j \in [n]$ and with $W_{ij} = W_{ij}(\mathcal{S})$ the total number of samples where they both appear. We, then, have that:

$$\Pr[\hat{\pi}[\Pi] \neq \pi_0] \leq \Pr[\exists i < j : n_{i\succ j} < n_{j\succ i}] \leq \sum_{i<j} \Pr[n_{i\succ j} < n_{j\succ i}]$$

For any fixed pair $i < j$, the probability that $n_{i\succ j} < n_{j\succ i}$ is maximized in the case when they are adjacent in π_0. In this case, $n_{i\succ j} = \sum_{\ell\in W_{ij}} X_\ell$, where $X_\ell \sim \text{Be}(e^{-\beta}/(1+e^{-\beta})), \forall \ell \in [W_{ij}]$. Let $Y_\ell = 1 - X_\ell, \forall \ell \in [W_{ij}]$. Then, we have:

$$\Pr[n_{i\succ j} < n_{j\succ i}] \leq \Pr\left[\sum_{\ell\in[W_{ij}]} X_\ell - Y_\ell \geq 0\right] \leq \exp\left(-2W_{ij}\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2\right) \leq \exp\left(-2p|\mathcal{S}|\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2\right),$$

where the second inequality follows from the Hoeffding bound and the third from the fact that $\mathcal{S}$ is $p-$frequent.

Demanding that the last term is less than ϵ and solving for $|\mathcal{S}|$ concludes the proof of Theorem 1.

B Proof of Theorem 2

Let $\mathcal{S}$ contain $p|\mathcal{S}|$ full sets and $(1-p)|\mathcal{S}|$ sets of size at most $n\sqrt{p/(1-p)}$. For any $i, j \in [n]$, let $W_{ij}(\mathcal{S})$ be the number of sets of $\mathcal{S}$ containing both i and j, that is, the number of the appearances of pair (i, j).

Clearly, $\mathcal{S}$ is $p-$frequent and:

$$\sum_{i<j} W_{ij}(\mathcal{S}) \leq pn^2|\mathcal{S}|, \tag{1}$$

since each full set has no more than $n^2/2$ pairs of alternatives and each of the remaining sets have no more than $n^2p/(2(1-p))$ pairs.

Assume that $|\mathcal{S}| < \frac{1}{8p\beta}\log(\frac{n(1-\epsilon)}{4\epsilon})$. From Eq. (1) we get that:

$$\sum_{i<j} W_{ij}(\mathcal{S}) < \frac{n^2}{8\beta}\log\left(\frac{n(1-\epsilon)}{4\epsilon}\right), \tag{2}$$

We will show that there exists a set of $n/2$ disjoint pairs of alternatives which we observe only a few times in the samples. For simplicity, assume that $n/2 \in \mathbb{N}$. Consider the following family $\{P_t\}_{t\in[n/2]}$ of perfect matchings on the set of alternatives, that is, $n/2$ sets of $n/2$ disjoint pairs:

$$P_1 = \{(1,2), (3,4), \dots, (n-1,n)\}$$
$$P_2 = \{(1,4), (3,6), \dots, (n-1,2)\}$$
$$\dots$$
$$P_t = \{(1, (2t) \bmod n), \dots, (n-1, (2t+n-2) \bmod n)\}$$
$$\dots$$
$$P_{n/2} = \{(1,n), (3,2), \dots, (n-1,n-2)\}$$

Observe that no pair of alternatives appears in more than one perfect matching of the above family. Therefore:

$$\sum_{t\in[n/2]} \sum_{(i,j)\in P_t} W_{ij}(\mathcal{S}) \le \sum_{i<j} W_{ij}(\mathcal{S}) \tag{3}$$

Therefore, combining Eq. (2) and Eq. (3), we get that:

$$\exists t \in [n/2] : \sum_{(i,j)\in P_t} W_{ij}(\mathcal{S}) < \frac{n}{4\beta} \log\left(\frac{n(1-\epsilon)}{4\epsilon}\right)$$

Therefore, since $|P_t| = [n/2]$, there exist at least $n/4$ pairs $(i,j) \in P_t$ with:

$$W_{ij}(\mathcal{S}) < \frac{1}{\beta} \log\left(\frac{n(1-\epsilon)}{4\epsilon}\right) \tag{4}$$

We proceed with an information-theoretic argument which is based on the observation that if the pairs of $P = P_t$, for which Eq. (4) holds, are adjacent in the central ranking, then the probability of swap is maximized for each pair and also knowledge of the pairwise orders of any fraction of the pairs does not give information about the pairwise order of any remaining pair. For simplicity and without loss of generality, assume that $P = P_1$,

For the selection vector $\mathcal{S} = (S_1, \dots, S_{|\mathcal{S}|})$ we denote with $\mathrm{Sym}(\mathcal{S})$ the support of $\mathcal{M}^{\mathcal{S}}_{\pi_0,\beta}$, namely the set of vectors containing in any position $\ell \in [|\mathcal{S}|]$ a permutation of S_ℓ. For the following, $\hat{\Pi}$ will be used to denote elements of $\mathrm{Sym}(\mathcal{S})$, while $\Pi \sim \mathcal{M}^{\mathcal{S}}_{\pi_0,\beta}$ will be a random variable.

Let $\tilde{\pi}$ be any (randomized) algorithm for estimating the central ranking. With the notation $\Pr[A]$ we refer to the probability of the event A, taking into consideration any randomness involved in A (for example the randomness of Π and $\tilde{\pi}$). Also, let $\mathcal{F} \subseteq \mathfrak{S}_n$ be as follows:

$$\mathcal{F} = \{\pi \in \mathfrak{S}_n : \{\pi(2i-1), \pi(2i)\} = \{2i-1, 2i\}\} \tag{5}$$

For example, if $n = 4$ then $\mathcal{F} = \{1 \succ 2 \succ 3 \succ 4, 1 \succ 2 \succ 4 \succ 3, 2 \succ 1 \succ 3 \succ 4, 2 \succ 1 \succ 4 \succ 3\}$.

Fix $\pi \in \mathcal{F}$. For any $\pi' \in \mathcal{F}$, let $D(\pi')$ denote the number of pairwise disagreements between π and π' on the elements of P:

$$D(\pi') = \{(i,j) \in P : (\pi(i) - \pi(j))(\pi'(i) - \pi'(j)) < 0\}$$

Fix $\hat{\Pi} \in \mathrm{Sym}(\mathcal{S})$. Then, assuming the following notation:

$$\Pr[\hat{\Pi}|\pi] := \Pr_{\Pi\sim\mathcal{M}^{\mathcal{S}}_{\pi,\beta}}[\Pi = \hat{\Pi}],$$

we apply the triangle inequality property of Kendall tau distance and get that:

$$\Pr[\hat{\Pi}|\pi] \ge e^{-\beta \sum_{(i,j)\in D(\pi')} W_{ij}(\mathcal{S})} \Pr[\hat{\Pi}|\pi'] \tag{6}$$

Also, since the estimator $\tilde{\pi}$ must have a single output:

$$\sum_{\pi' \in \mathcal{F}} \Pr[\tilde{\pi}[\hat{\Pi}] = \pi'] \leq 1 \tag{7}$$

We multiply the terms of Ineq. (7) with $\Pr[\hat{\Pi}|\pi]$, apply Ineq. (6) and sum over all $\hat{\Pi} \in \mathrm{Sym}(\mathcal{S})$ to get:

$$\sum_{\pi' \in \mathcal{F}} e^{-\beta \sum_{(i,j)\in D(\pi')} W_{ij}(\mathcal{S})} \sum_{\hat{\Pi} \in \mathrm{Sym}(\mathcal{S})} \Pr[\tilde{\pi}[\hat{\Pi}] = \pi'] \Pr[\hat{\Pi}|\pi'] \leq \sum_{\hat{\Pi} \in \mathcal{F}} \Pr[\hat{\Pi}|\pi] \leq 1$$

Assume, for contradiction, that for every $\pi' \in \mathcal{F}$ it holds:

$$\Pr_{\Pi \sim \mathcal{M}^{\mathcal{S}}_{\pi,\beta}}[\tilde{\pi}[\Pi] = \pi'] \geq 1 - \epsilon$$

Then, since it holds that $\Pr_{\Pi \sim \mathcal{M}^{\mathcal{S}}_{\pi,\beta}}[\tilde{\pi}[\Pi] = \pi'] = \sum_{\hat{\Pi} \in \mathrm{Sym}(\mathcal{S})} \Pr[\tilde{\pi}[\hat{\Pi}] = \pi'] \Pr[\hat{\Pi}|\pi']$, we get:

$$(1-\epsilon) \sum_{\pi' \in \mathcal{F}} e^{-\beta \sum_{(i,j)\in D(\pi')} W_{ij}(\mathcal{S})} \leq 1$$

However, from Ineq. (4), we get that:

$$\sum_{\pi' \in \mathcal{F}} e^{-\beta \sum_{(i,j)\in D(\pi')} W_{ij}(\mathcal{S})} > 1 + \frac{n}{4} \frac{4\epsilon}{n(1-\epsilon)}$$

We conclude that: $1 - \epsilon + \epsilon < 1$, contradiction.

C Proof of Theorem 3

For the following, for any ranking $\pi \in \mathfrak{S}_n$ and $S \subseteq [n]$, let $\pi|_S$ denote the reduced central ranking of π to S, that is, the permutation[1] of the elements of S that agrees with their order in π.

Assume that $\Pi \sim \mathcal{M}^{\mathcal{S}}_{\pi_0,\beta}$, where $\mathcal{S}$ is $p-$frequent. The proof of Theorem 3 is based on the definition of a notion of neighborhood for each alternative $i \in [n]$. In particular, for every $i \in [n]$, $L > 0$ and $\lambda > 0$ we define $\mathcal{N}_i(L,\lambda) = \mathcal{N}_i(L,\lambda)[\Pi]$ to be the subset of $[n]$ containing all the alternatives j for which there exist at least r/λ sets S of $\mathcal{S}$ for each of which it holds that $|\pi_0|_S(i) - \pi_0|_S(j)| \leq L$.

Observe that the neighborhoods are formed according to the input data but they are unknown to the algorithm, since π_0 is unknown.

Furthermore, the following Lemma holds, that controls the neighborhoods' size.

Lemma C.1. *For every $i \in [n]$, $L > 0$, $\lambda > 0$, it holds that:*

$$|\mathcal{N}_i(L,\lambda)| \leq 2\lambda L$$

Proof. There are at most $2|\mathcal{S}|L$ total positions for the neighbors of i and each neighbor takes at least $|\mathcal{S}|/\lambda$ of them. □

We now prove that we can pick L and λ so that, with high probability, for any $i \in [n]$, every element outside its neighborhood $\mathcal{N}_i(L,\lambda)$ is ranked correctly relatively to i in the majority of samples where they both appear. More specifically, the following Lemma holds:

[1] Specifically, for a permutation $\pi : [n] \to [n]$ and $S \subseteq [n]$, the mapping $\pi|_S$ is a bijection from S to $[|S|]$.

Lemma C.2. *Assume that* $\Pi = (\pi_1, \ldots, \pi_r) \sim \mathcal{M}^{\mathcal{S}}_{\pi_0,\beta}$, $c \in (0, 1/2]$, *c constant and* $\epsilon \in (0, 1)$. *Then, there exists some constant* $C = C(c)$ *such that, by considering:*

$$L = \frac{C}{\beta} + \frac{2}{\beta cpr} \log(2n^2/\epsilon) \text{ and } \lambda = \frac{2}{(1-c)p},$$

it holds that:

$$n_{i \succ j} \geq (1-c) W_{ij}$$

for every $i \in [n]$ *and for every* $j \in [n] \setminus (\mathcal{N}_i(L, \lambda) \cup \{i\})$ *with probability at least* $1 - \epsilon$.

Proof. Fix $i \in [n]$ and $j \in [n] \setminus (\mathcal{N}_i(L) \cup \{i\})$. It is sufficient to show that the probability of the event that $n_{j \succ i} > cW_{ij}$ is less than ϵ/n^2, since, in that case, the union of the corresponding events over all pairs i', j' such that $j' \in [n] \setminus (\mathcal{N}_{i'}(L) \cup \{i'\})$ would hold with probability less than ϵ.

We have that $|\mathcal{S}| \geq W_{ij} \geq p|\mathcal{S}|$. From the selection of j we have that the number of elements of the set $\mathcal{I}$ of indices $\ell \in [|\mathcal{S}|]$ such that: $|\pi_0|_{S_\ell}(i) - \pi_0|_{S_\ell}(j)| > L$ (i and j are initially distant and hence they do not appear swapped in the corresponding sample, with high probability) is at least $W_{ij} - (1/\lambda)|\mathcal{S}|$.

For some index $\lambda \in \mathcal{I}$ we have that the probability that i and j appear swapped in π_ℓ is upper bounded as follows, according to Bhatnagar and Peled [2015]:

$$\rho = \Pr[(\pi_\ell(i) - \pi_\ell(j))(i - j) < 0] \leq 2e^{-\beta L/2}$$

Assume that $i < j$. We will show that in most of the W_{ij} samples where both i and j appear, it holds that $i \succ_{\pi_\ell} j$. That is, following the notation introduced in Section A of the supplement:

$$\Pr[n_{i \succ j} > cW_{ij}] < \epsilon/(2n^2)$$

Let $n'_{i \succ j}$ be the random variable that corresponds to the number of indices $\ell \in \mathcal{I}$ for which $i \succ_{\pi_\ell} j$. Clearly:

$$\Pr[n_{i \succ j} > cW_{ij}] \leq \Pr[n'_{i \succ j} > cW_{ij}]$$

Also, let $W'_{ij} = |\mathcal{I}| (\geq W_{ij} - (1/\lambda)|\mathcal{S}|)$. The random variable $n'_{i \succ j}$ follows the Binomial distribution $\text{Bin}(W'_{ij}, \rho)$. We want to find some constant $c' = c'(c) \in (0, 1)$ such that $c' W'_{ij} \geq cW_{ij}$, because, in this case, it suffices to bound the following probability:

$$\Pr[n'_{i \succ j} > c' W'_{ij}]$$

We pick $c' = c\lambda p/(\lambda p - 1)$. For the selected λ, c' is indeed a constant (since c is considered a constant) taking some value within the interval $(0, 1)$ (since $c \in (0, 1)$) and from the Chernoff bound we get:

$$\Pr[n'_{i \succ j} > c' W'_{ij}] \leq \exp(-W'_{ij} \mathrm{D_{KL}}(c' \parallel \rho)) \leq \exp(-(p - 1/\lambda) r \mathrm{D_{KL}}(c' \parallel \rho)),$$

where $\mathrm{D_{KL}}(c' \parallel \rho) = c' \log(c'/\rho) + (1 - c') \log((1 - c')/(1 - \rho)) \geq c' \beta L/2 - C'$, where $C' = C'(c') > 0$ is some positive constant.[2] For the selected L, if $C = 2C'/c'$, we get that $\Pr[n'_{i \succ j} > c' W'_{ij}] \leq \epsilon/(2n^2)$, concluding the proof of Lemma C.2. □

We conclude the proof of Theorem 3 by combining Lemmata C.1 and C.2 to get that with probability at least $1 - \epsilon$, for every alternative $i \in [n]$, the number of other alternatives with which i is ordered reversely in the majority of samples where they both appear is upper bounded by N:

$$N = \frac{2C}{(1-c)p\beta} + \frac{4}{\beta c(1-c)pr} \log(2n^2/\epsilon)$$

Hence, after tie braking, the resulting permutation ranks each alternative no more than $2N$ places away from its position in π_0. More specifically, if before breaking ties, for $i, j \in [n]$ we have $\hat{\pi}(j) \leq \hat{\pi}(i)$ then: $\hat{\pi}(j) + N \leq \hat{\pi}(i) + N$. However: $j \leq \hat{\pi}(j) + N$ and $\hat{\pi}(i) + N \leq i + 2N$ therefore: $j \leq i + 2N \Rightarrow j - i \leq 2N$. Therefore, after tie braking: $\hat{\pi}(i) \leq i + 2N$. Symmetrically: $\hat{\pi}(i) \geq i - 2N$.

[2]Let μ, ν be two discrete probability measures on Ω. The Kullback–Leibler divergence between μ is defined as $\mathrm{D_{KL}}(\mu \parallel \nu) = \sum_{x \in \Omega} \mu(x) \log(\frac{\mu(x)}{\nu(x)})$

D Proof of Theorem 4

Theorem 4 consists of two parts. The first one considers the runtime of an algorithm finding a likelier than nature estimation of the central ranking given $p-$frequent selective Mallows samples while the second one refers to solving the maximum likelihood estimation problem. Both parts are based on Lemma 2 (which originates to the work of Braverman and Mossel [2009]).

Part 1. A careful examination of the proof of Lemma 2 (which can be found in [Braverman and Mossel, 2009]) reveals that if $\pi_0 \in \mathcal{A} = \{\pi \in \mathfrak{S}_n : |\pi(i) - \hat{\pi}(i)| \leq N, \forall i \in [n]\}$, then, we can find a likelier than nature estimation π° in time $O(n \cdot N^2 \cdot 2^{6N})$. Picking N according to Theorem 3, π_0 is indeed an element of $\mathcal{A}$ and therefore, we get the desired result.

Part 2. In this case, we want to show that $\pi^* \in \mathcal{A}' = \{\pi \in \mathfrak{S}_n : |\pi(i) - \hat{\pi}(i)| \leq N', \forall i \in [n]\}$. We claim that with probability at least $1 - \epsilon/2$: $\{|\pi^*(i) - \pi_0(i)| \leq K, \forall i \in [n]\}$ (π^* and π_0 are pointwise close) for some $K = O(\frac{1}{\beta p^3} + \frac{1}{\beta p^4 r}\log(n/\epsilon))$. Therefore, picking $N' = K + N = O(K)$, we have that $\pi^* \in \mathcal{A}'$, which gives the desired result.

To prove our claim, we generalize the proof that the maximum likelihood estimation of the central ranking from complete Mallows samples is pointwise close to the central ranking.

Assume, without loss of generality, that $\pi_0 = \text{id}$. Let $h > 0$ and $c \in (0, 1/2)$, which will be defined later. For any $i, j \in [n]$, $n_{i \succ j}$ is the number of samples where $i \succ j$ and W_{ij} is the number of samples where both i and j appear. Clearly, it holds that $n_{i\succ j} + n_{j \succ i} = W_{ij}$.

From Lemmata C.2 and C.1, with probability at least $1 - \epsilon$, there exist λ and L such that for every alternative $i \in [n]$ and any constant $c \in (0, 1/2]$ there exists some constant $C = C(c)$ such that :

1. $|\mathcal{N}_i(L, \lambda)| \leq N = \frac{2C}{\beta(1-c)p} + \frac{4}{\beta(1-c)cp^2 r}\log(2n^2/\epsilon)$
2. $j \in \{i+1, i+2, \dots, n\} \setminus \mathcal{N}_i(L, \lambda) \Rightarrow n_{j\succ i} \leq cW_{ij}$ (and $W_{ij} \geq pr$).
3. Symmetrically, for $j \in \{1, \dots, i-1\} \setminus \mathcal{N}_i(L, \lambda)$: $n_{i \succ j} \leq cW_{ij}$.

Fix $i \in [n]$ such that $|\pi^*(i) - i| = K$, where $K \geq hN$. Without loss of generality, assume $\pi^*(i) = i + K$. It suffices to find values of c and h that contradict the assumption $\pi^*(i) = i + K$.

Let $\mathcal{D} = \{j \in [n] : i \leq \pi^*(j) < i + K\}$ and: $\mathcal{D}_1 = \{j \in \mathcal{D} : j < i\}$, $\mathcal{D}_2 = \{j \in \mathcal{D} : j \in \mathcal{N}_i(L, \lambda)\}$, $\mathcal{D}_3 = \{j \in \mathcal{D} : j > i \text{ and } j \notin \mathcal{N}_i(L, \lambda)\}$. Apparently: $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$.

Observe that since π^* maximizes the following score function:

$$s: \mathfrak{S}_n \to \mathbb{N}$$
$$\pi \to s(\pi) = \sum_{i_1 \succ_\pi i_2} n_{i_1 \succ i_2}$$

It must hold that:

$$0 \leq \sum_{j \in \mathcal{D}} (n_{j \succ i} - n_{i \succ j})$$

For any $j \in \mathcal{D}_3$, from item (1) it holds that:

$$n_{j\succ i} \leq cW_{ij} \Rightarrow n_{j \succ i} - n_{i \succ j} \leq -(1 - 2c)pr$$

Let $|\mathcal{D}_1| = T$. Furthermore: $|\mathcal{D}_2| \leq N$ and $|\mathcal{D}_3| \geq K - N - T \geq (h-1)N - T$. Therefore, we have that:

$$0 \leq rT + rN - (1 - 2c)pr((h-1)N - T) \Rightarrow$$

$$T \geq \frac{(1-2c)p(h-1) - 1}{1 + (1-2c)p} N \tag{8}$$

Observe that, since there are at least T alternatives $j < i$ such that $\pi^*(j) \geq i$, say $T_1 \subset [n]$, there must be at least T alternatives $j \geq i$ such that $\pi^*(j) < i$, say $T_2 \subset [n]$. Let $H_1 = \{1, ..., i-1\}$ and $H_2 = \{i, ..., n\}$. We construct $\sigma_0 \in \mathfrak{S}_n$ by concatenating $\pi^*|_{H_1}$ and $\pi^*|_{H_2}$. It remains to select appropriate values for h (h does not need to be constant) and c (must be constant) for which $s(\sigma_0) > s(\pi^*)$, which is a contradiction.

Create the following sets:

1. P_1: The pairs of elements $i_1, i_2 \in [n], i_1 < i_2$ for which $i_2 \in \mathcal{N}_{i_1}(L, \lambda)$ and σ_0, π^* disagree on their relative ranking. Note that: $|P_1| \leq 2TN$.

2. P_2: The pairs of elements $i_1, i_2 \in [n], i_1 < i_2$ for which σ_0, π^* disagree, but $i_2 \notin \mathcal{N}_{i_1}(L, \lambda)$ (and $i_1 \notin \mathcal{N}_{i_2}(L, \lambda)$). Note that σ_0 has the right answer for this pair and $q(i_1 \succ i_2) - q(i_2 \succ i_1) \geq (1-2c)pr$. Also: $|P_2| \geq T(T-N)$ (select an element of T_1 and an element of T_2 which is not in the first element's neighborhood).

Then: $s(\sigma_0) - s(\pi^*) = \sum_{(i_1,i_2) \in P_1} (n_{i_1 \succ i_2} - n_{i_2 \succ i_1}) + \sum_{(i_1,i_2) \in P_2} (n_{i_1 \succ i_2} - n_{i_2 \succ i_1}) \geq -2rTN + (1-2c)prT(T-N) = rT((1-2c)pT - ((1-2c)p+2)N)$

Using Ineq. (8), we get that:

$$s(\sigma_0) - s(\pi^*) \geq rTN \left[\frac{(1-2c)p((1-2c)p(h-1)-1)}{1+(1-2c)p} - (2 + (1-2c)p) \right]$$

We search for values of c and h such that the quantity inside the brackets is positive. After some algebra, we choose $c < 1/4$ (constant) and $h = 2 + 8/p + 8/p^2 = O(\frac{1}{p^2})$.

E Proof of Theorem 5

The proof we provide almost coincides with the sketch we provided in the main part. However, here we have established the appropriate notation that enables us to be more formal.

Let $\Pi \sim \mathcal{M}^{\mathcal{S}}_{\pi_0,\beta}$ be our sample profile. We will make use of the PosEst $\hat{\pi} = \hat{\pi}[\Pi]$ and, without loss of generality, assume that π_0 is the identity permutation id. We will bound the probability that there exists some $i \in [k]$ such that $\hat{\pi}(i) \neq i$.

For any $i \in [k]$, we can partition the remaining alternatives into $A_1(i) = \mathcal{N}_i(L, \lambda)$ and $A_2(i) = [n] \setminus (A_1(i) \cup \{i\})$.

From Lemma C.2, it holds that for some L, λ such that $|A_1(i)| = O(\frac{1}{p\beta} + \frac{1}{p^2\beta|\mathcal{S}|}\log(n/\epsilon))$, with probability at least $1 - \epsilon/2$, for every $i \in [n]$, for every alternative $j \in A_2(i)$ it holds: $(n_{i \succ j} - n_{j \succ i})(j - i) > 0$.

Picking L, λ so that the above result holds, there exists some $r_1 = O(\frac{1}{p^2\beta k}\log(n/\epsilon))$ such that, if $|\mathcal{S}| \geq r_1$, then $|A_1(i)| = O(k)$.

Furthermore, following the same technique used to prove Theorem 1, we get that, if $|\mathcal{S}| \geq r_2$ for some $r_2 = O(\frac{1}{p(1-e^{-b})^2}\log(k/\epsilon))$, then, with probability at least $1 - \epsilon/2$, for every pair of alternatives (i, j) such that $i \in [k]$ and $j \in A_1(i)$ it holds that: $(n_{i \succ j} - n_{j \succ i})(j - i) > 0$, since the total number of such pairs is $O(k^2)$.

Therefore, with probability at least $1 - \epsilon$, both events hold and for any fixed $i \in [k]$, $\hat{\pi}(i) = i$, because for all j: $(n_{i \succ j} - n_{j \succ i})(j - i) > 0$ and also because for every other alternative $j > k$, we have that $\hat{\pi}(j) > k$, since $n_{i \succ j} > n_{j \succ i}, \forall i \in [k]$.

F Randomized $p-$frequent assumption

In this section, we consider an alternative assumption to the $p-$frequent one, namely the randomized $p-$frequent assumption and we present empirical data that indicate that at least some of our results under the $p-$frequent assumption continue to hold under the randomized one.

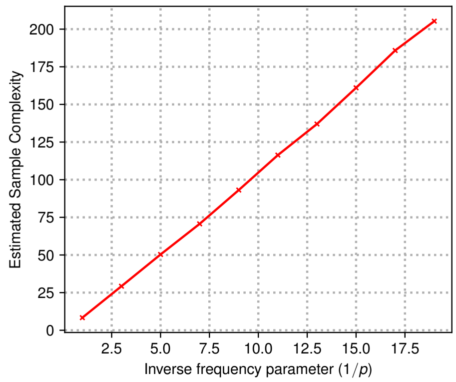

Figure 1: Estimated sample complexity of retrieving, with probability at least 0.95 and using PosEst, the central ranking from selective Mallows samples, with $n = 20$, $\beta = 2$, over the randomized frequency parameter's inverse.

Definition. We say that a distribution $\mathcal{D}$ supported on $2^{[n]}$ is $p-$frequent if for any $i, j \in [n]$ it holds:

$$\Pr_{S \sim \mathcal{D}}[S \ni i, j] \geq p$$

If the selection sets are independently drawn according to some $p-$frequent distribution, then we say that the selective Mallows model is randomly $p-$frequent. Note that the randomized $p-$frequent assumption, contrary to the simple $p-$frequent assumption, permits the event that some pair of alternatives never appear together in the samples, yet with small probability.

Empirical evaluation. We estimate the sample complexity of retrieving the central ranking from randomly selective Mallows samples where $n = 20$ and $\beta = 2$, with probability at least 0.95, using PosEst by performing binary search over the size of the sample profile. During a binary search, for every value, say r, of the sample profile size we examine, we estimate the probability that PosEst outputs the central ranking by drawing 100 independent randomly $p-$frequent selective Mallows profiles of size r, computing PosEst for each one of them and counting successes. We then compare the empirical success rate to 0.95 and proceed with our binary search accordingly. For a specific value of p, we estimate the corresponding sample complexity, by performing 50 independent binary searches and computing the average value. The results, which are shown in Figure 1, indicate that the dependence of sample complexity on the frequency parameter p is indeed $\Theta(1/p)$.

References

Ioannis Caragiannis, Ariel D Procaccia, and Nisarg Shah. When Do Noisy Votes Reveal the Truth? In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 143–160, 2013. URL `https://dl.acm.org/doi/10.1145/2892565`.

Nayantara Bhatnagar and Ron Peled. Lengths of Monotone Subsequences in a Mallows Permutation. *Probability Theory and Related Fields*, 161(3-4):719–780, 2015.

Mark Braverman and Elchanan Mossel. Sorting from Noisy Information. *arXiv preprint arXiv:0910.1191*, 2009. URL `https://arxiv.org/pdf/0910.1191.pdf`.