Supplementary Document for "Efficient Rank Aggregation with Lehmer Codes"

I Proof of Lemma 4.2

Before proceeding with the proof, we remark that some ideas in our derivatione have been motivated by Lemma 10.7 of [1].

Let $i \triangleq \sigma_0(u)$. Suppose that $n > j \geq i$ and that we want to prove statement 1) (the second case when $0 < j \leq i$ may be handled similarly). When $i = 1$, the underlying ratio is exactly equal to ϕ. Hence, we only consider the case when $i > 1$. Let $E = \{\sigma : \sigma(u) = j\}$ and $T = \{\sigma : \sigma(u) = j+1\}$. In this case, $\mathbb{P}[\sigma(u) = j] = \mathbb{P}[E]$ and $\mathbb{P}[\sigma(u) = j+1] = \mathbb{P}[T]$. Define the sets:

$$
\begin{aligned}
E_1 &= \{\sigma : \sigma(u) = j, \sigma_0(\sigma^{-1}(j+1)) > i\},\\
E_2 &= \{\sigma : \sigma(u) = j, \sigma_0(\sigma^{-1}(j+1)) < i\},\\
T_1 &= \{\sigma : \sigma(u) = j+1, \sigma_0(\sigma^{-1}(j)) > i\},\\
T_2 &= \{\sigma : \sigma(u) = j+1, \sigma_0(\sigma^{-1}(j)) < i\}.
\end{aligned}
$$

Clearly, $\mathbb{P}[E] = \mathbb{P}[E_1] + \mathbb{P}[E_2]$ and $\mathbb{P}[T] = \mathbb{P}[T_1] + \mathbb{P}[T_2]$. By swapping u and $\sigma^{-1}(j+1)$, we can construct two bijections $E_1 \leftrightarrow T_1$ and $E_2 \leftrightarrow T_2$. Statement 1) can then be easily proved by using the following three claims:

$$
\begin{aligned}
\mathbb{P}[T_1] &= \phi\mathbb{P}[E_1],\\
\mathbb{P}[T_2] &= \frac{1}{\phi}\mathbb{P}[S_2],\\
0 < \mathbb{P}[T_2] &\overset{a)}{\leq} \phi_{1:n-1}\mathbb{P}[T_1].
\end{aligned} \tag{1}
$$

Observe that inequality is achieved in a) when $j = n-1$. The first two claims are straightforward to check, and hence we only prove the third claim.

Consider a mapping from T_2 to T_1 based on circular swapping of elements, and let $\sigma \in T_2$. Since $\sigma(u) - 1 = j \geq i$ and $\sigma_0(\sigma^{-1}(j)) < i$, there must exist an element x such that $\sigma_0(x) > \sigma_0(u)$ and $\sigma(x) < j$. Choose the element x with the largest corresponding value of $\sigma(x)$ and construct a new ranking σ' such that

$$
\sigma'(y) = \begin{cases} \sigma(y), & \text{if } \sigma(y) < \sigma(x) \text{ or } \sigma(y) \geq \sigma(u),\\ \sigma(y) - 1, & \text{if } \sigma(x) < \sigma(y) \leq \sigma(u),\\ j, & \text{if } \sigma(y) = \sigma(x). \end{cases}
$$

It is easy to see that $\sigma' \in T_1$. Given that all elements ranked between x and u in σ have rank higher than $\sigma_0(x)$, we have $\mathbb{P}[\sigma] = \phi^{\sigma(u)-\sigma(x)-1}\mathbb{P}[\sigma'] = \phi^{j-\sigma(x)}\mathbb{P}[\sigma']$. Note that the above mapping is neither a bijection nor an injection. Denote the mapping by $\mathcal{M} : T_{j,2} \to T_2$. For each $\sigma' \in T_1$, define $T_{2,\sigma'} \subset T_2$, so that for all $\sigma \in T_{2,\sigma'}$, $\mathcal{M}(\sigma) = \sigma'$. Then, $\cup_{\sigma' \in T_1} T_{2,\sigma'} = T_2$ forms a partition of the set T_2. Next, consider two distinct rankings $\sigma_1, \sigma_2 \in T_{2,\sigma'}$. These rankings must rank the element x differently, i.e., one must have $\sigma_1(x) \neq \sigma_2(x)$. Therefore, $\mathbb{P}[T_{2,\pi'}] \leq \mathbb{P}[\pi']\phi_{1:j-1} = \mathbb{P}[\pi']\phi_{1:j-1}$. As a result, $\mathbb{P}[T_2] \leq \mathbb{P}[T_1]\phi_{1:n-1}$, which proves the third claim. We conclude by observing that the condition under which equality is achieved in the bound stated in the lemma is exactly the same condition under which equality is achieved in the bound stated in the third claim.

II Proof of Lemma 4.3

Let $i \triangleq \sigma_{0,A}(u)$. Suppose that $n > j \geq i$ and that we want to prove statement 1) (the case when $0 < j \leq i$ may be handled similarly). Let $E = \{\pi : \pi_A(u) = j\}$ and $T = \{\pi : \pi_A(u) = j+1\}$. The left-hand-side in the statement of 1) equals the ratio $\frac{\mathbb{P}[T]}{\mathbb{P}[E]}$. Note that removing a fixed number of elements in S of lowest (or highest) rank in the centroid ranking does not change the probability of the ranking involving the remaining elements (see Lemma 5 in Section VI of the supplementary document for the proof). We can hence assume that $\sigma_{0,A}^{-1}(1)$ is the element with highest rank in σ_0.

When $i = 1$, for any ranking σ in T, we can swap the element u with the element $x \in A$ for which $\sigma_A(x) = \sigma_A(u) - 1$ to obtain another ranking $\sigma' \in E$. Moreover, it is easy to check that $\mathbb{P}[\sigma']\phi \geq \mathbb{P}[\sigma]$, so that the ratio in the statement 1) does not exceed ϕ. Note that we have inequality "$\geq$" instead of equality "$=$" in $\mathbb{P}[\sigma']\phi \geq \mathbb{P}[\sigma]$, since there may potentially exists other elements in S/A ranked between x and u in σ.

Next, consider the case when $i > 1$. Define the sets

$$\begin{aligned}
E_1 &= \{\sigma : \sigma_A(u) = j, \sigma_{0,A}(\sigma_A^{-1}(j+1)) > i\},\\
E_2 &= \{\sigma : \sigma_A(u) = j, \sigma_{0,A}(\sigma_A^{-1}(j+1)) < i\},\\
T_1 &= \{\sigma : \sigma_A(u) = j+1, \sigma_{0,A}(\sigma_A^{-1}(j+1)) > i\},\\
T_2 &= \{\sigma : \sigma_A(u) = j+1, \sigma_{0,A}(\sigma_A^{-1}(j+1)) < i\}.
\end{aligned}$$

Then, $\mathbb{P}[E] = \mathbb{P}[E_1] + \mathbb{P}[E_2]$ and $\mathbb{P}[T] = \mathbb{P}[T_1] + \mathbb{P}[T_2]$. By swapping u and $\sigma_A^{-1}(j+1)$, we can construct two bijections $E_1 \leftrightarrow T_1$ and $E_2 \leftrightarrow T_2$ as follows.

Let us consider a finer partition of T_2 in terms of permutations with four labels. More precisely, associate each ranking $\sigma \in T_2$ with a label vector $(x_1, x_2, \ell_1, \ell_2)$, where:

$x_1 = \pi^{-1}(j)$. Note that $\sigma_{0,A}(x_1) < i$ due to the definition of T_2.

$x_2 = \arg\max_{x:\sigma_{0,A}(x)>i, \sigma_A(x)<\sigma_A(u)} \sigma_A(x)$; the label x_2 is well-defined due to the pigeon-hole principle.

$\ell_1 =$ the cardinality of the set F_1 defined as

$$F_1 = \{x \in [n] : \sigma_0(x_1) < \sigma_0(x) < \sigma_0(u), \sigma(\sigma_A^{-1}(j-1)) < \sigma(x) < \sigma(x_1)\}.$$

$\ell_2 =$ the cardinality of the set F_2 defined as

$$F_2 = \{x \in [n] : \sigma_0(x_1) < \sigma_0(x) < \sigma_0(u), \sigma(x_1) < \sigma(x) < \sigma(u)\}.$$

We summarize those labels in a vector $L = (x_1, x_2, \ell_1, \ell_2)$ and thus partition T_2 according to different label vectors L, i.e.,

$$T_2 = \cup_L T_{2,L}. \tag{2}$$

A ranking in T_2 is in $T_{2,L}$ if its corresponding label vector equals L.

We further construct a mapping $\mathcal{M}$ from T_2 to T_1 by swapping elements ranked between x_1 and x_2, so that $\sigma' = \mathcal{M}(\sigma)$ equals

$$\sigma'_A(x) = \begin{cases} j, & \sigma_A(x) = \sigma_A(x_2),\\ \sigma_A(x) - 1, & \sigma_A(x_2) < \sigma_A(x) < j,\\ \sigma_A(x) - 1, & \sigma_A(x) = \sigma_A(x_1),\\ \sigma_A(x), & \text{for other } x \in A. \end{cases}$$

The above mapping basically performs circular swapping by moving x_2 to the position one rank higher and adjacent to u and by moving each element in A between x_2 and x_1, including x_1, to a higher position adjacent to the original one. Based on $\mathcal{M}$, one can also form a partition of T_1 as

$$T_1 = (\cup_L T_{1,L}) \cup T_{1,L^c} \tag{3}$$

where $T_{1,L}$ contains the rankings mapped from $T_{2,L}$ via $\mathcal{M}$. Note that T_{1,L^c} denote the "remainder set" of permutations that do not have a preimage in T_2. In this remainder set, a ranking σ has the property that the elements $\sigma_A^{-1}(j)$ and $\sigma_A^{-1}(j-1)$ are both ranked lower than u in the centroid ranking. Since the swapping operations establish a bijection between $E_1 \leftrightarrow T_1$ and $E_2 \leftrightarrow T_2$, one can also partition E_1, E_2 as

$$E_1 = (\cup_L E_{1,L}) \cup E_{1,L^c}, \tag{4}$$
$$E_2 = \cup_L E_{2,L}. \tag{5}$$

Let $R(\mathcal{L})$ denote $\frac{\mathbb{P}[\cup_{L\in\mathcal{L}}(T_{1,L}\cup T_{2,L})]}{\mathbb{P}[\cup_{L\in\mathcal{L}}(E_{1,L}\cup E_{2,L})]}$ and let $R(L)$ denote the same type of ratio but for a specific choice of L, i.e., $\frac{\mathbb{P}[T_{1,L}\cup T_{2,L}]}{\mathbb{P}[E_{1,L}\cup E_{2,L}]}$. Also, let $\mathcal{L}_0$ denote the set of all possible values of L. To prove the upper bound on $\frac{\mathbb{P}[T]}{\mathbb{P}[E]}$, we proceed through four steps.

1. Partition T and E and verify the validity of (2), (3), (4) and (5).
2. Prove that $\frac{\mathbb{P}[T_{1,L^c}]}{\mathbb{P}[E_{1,L^c}]} \leq \phi$.
3. Prove the upper bound for $R(L)$ when $\ell_1 = \ell_2$.
4. Prove the upper bound for $R(L \cup L')$, where $L = (k_1, k_2, \ell_1, \ell_2)$ and $L' = (k_1, k_2, \ell_2, \ell_1)$, for the case that $\ell_1 \neq \ell_2$.

The second step is easy to prove by directly swapping $\sigma_A^{-1}(j)$ and u in any given ranking $\sigma \in T_{1,L^c}$. We hence only need to establish the validity of the results in Steps 3 and 4.

For any $L = (k_1, k_2, \ell_1, \ell_2)$, the following claims hold:

$$\mathbb{P}[E_{1,L}] \geq \phi^{-1}\mathbb{P}[T_{1,L}],$$
$$\mathbb{P}[E_{2,L}] = \phi^{1+2\ell_2}\mathbb{P}[T_{2,L}],$$
$$\mathbb{P}[T_{2,L}] \leq \phi^{2\ell_1} f_L \mathbb{P}[T_{1,L}], \text{ where } f_L \leq \phi_{1:|A|-2-\ell_1-\ell_2}, \tag{6}$$

where the first two claims are easy to prove, while the equation (6) may be verified similarly as (1) in the proof of Lemma 5.2 (See Section 1 of this supplement).

For any $\sigma \in T_{2,L}$, $\sigma' = \mathcal{M}(\sigma) \in T_{1,L}$. Given that all the elements in A ranked between x_2 and u in σ are ranked lower than u, x_2 in the centroid, and due to swapping, we have

$$\frac{\mathbb{P}[\sigma]}{\mathbb{P}[\sigma']} \leq \phi^{\sigma_A(u)-\sigma_A(x_2)-1+2\ell_1} = \phi^{j-\sigma_A(x_2)+2\ell_1}.$$

Consider two distinct rankings $\sigma_1, \sigma_2 \in T_{2,L}$. If $\mathcal{M}(\sigma_1) = \mathcal{M}(\sigma_2)$, both rankings rank the element x_2 differently over A, i.e., $\sigma_{1,A}(x_2) \neq \sigma_{2,A}(x_2)$. Therefore, we must have

$$\sum_{\sigma:\mathcal{M}(\sigma)=\sigma'} \frac{\mathbb{P}[\sigma]}{\mathbb{P}[\sigma']} \leq \phi^{2(\ell_1+\ell_2)}\phi_{1:|A|-2-\ell_1-\ell_2}.$$

By examining all mappings from $\mathbb{P}[T_{2,L}]$ to $\mathbb{P}[T_{1,L}]$, we conclude that $f_L(\phi) = \phi^{-2(\ell_1+\ell_2)}\frac{\mathbb{P}[T_{2,L}]}{\mathbb{P}[T_{1,L}]} \leq \phi_{1:|A|-2-s}$, which establishes the third claim in (6).

Substituting the above expressions into $R(L)$, we have

$$R(L) \leq \frac{1+\phi^{2\ell_1}\phi_{1:|A|-2-\ell_1-\ell_2}}{\phi^{-1}+\phi^{1+2(\ell_1+\ell_2)}\phi_{1:|A|-2-\ell_1-\ell_2}}. \tag{7}$$

Suppose next that $\ell_1 = \ell_3 = \ell$. Then,

$$R(L) \leq \frac{1+\phi^{2\ell}\phi_{1:|A|-2-2\ell}}{\phi^{-1}+\phi^{1+4\ell}\phi_{1:|A|-2-2\ell}}. \tag{8}$$

which completes the proof of the bound in Step 3.

Let us now consider the bound in Step 4. When $\ell_1 \neq \ell_2$, direct optimization over ℓ_1, ℓ_2 cannot yield the required upper bound as $\ell_2 \to \infty$ may increase the right-hand-side of (9). Hence, in addition to $L = (k_1, k_2, \ell_1, \ell_2)$, let us also simultaneously consider $L' = (k_1, k_2, \ell_2, \ell_1)$, as a larger ℓ_2 will yield a smaller $R(L')$, .

Without loss of generality, suppose that $\ell_2 > \ell_1$. First, we prove the following Lemma.

Lemma 1. *For a pair (L, L') defined as above with $\ell_2 > \ell_1$, one has*

$$\frac{\mathbb{P}[T_{1,L}]}{\mathbb{P}[T_{1,L'}]} \leq \phi^{(\ell_2-\ell_1)}.$$

Proof. Recall the definition of the set F_2 and the fact that for a ranking $\sigma \in T_{1,L}$, σ is obtained via $\mathcal{M}(\pi)$ for some $\pi \in T_{2,L}$. Hence, in σ, the elements in F_1 are now ranked higher than x_2 and lower than x_1, while the elements in F_2 are now ranked higher than u and lower than x_2. Based on this structure of σ, for each ranking $\sigma \in T_{1,L}$, one may perform a swapping operation to obtain another ranking σ' in $T_{1,L'}$. The swapping constitutes a bijection. To see this, consider the set of $\ell_2 - \ell_1$ elements with highest rank in σ_{F_2} (Note that we assumed $\ell_2 > \ell_1$ but could have otherwise considered the $\ell_1 - \ell_2$ elements with lowest rank in σ_{F_1}.). Swapping the element x_2 and the selected elements in F_2 ranked from high to low yields a ranking $\sigma' \in T_{1,L'}$. Since for any element $x \in F_2$, $\sigma_0(x) < \sigma_0(u) < \sigma_0(x_2)$, we have $\mathbb{P}([\sigma]) \leq \mathbb{P}([\sigma'])\phi^{\ell_2-\ell_1}$. This completes the proof. ∎

Let $P \triangleq \frac{\mathbb{P}[T_{1,L}]}{\mathbb{P}[T_{1,L'}]}$. Substituting the results of all claims (6) into $R(L \cup L')$, we obtain

$$\begin{aligned} R(L \cup L') &= \frac{\mathbb{P}[T_{1,L} \cup T_{2,L}] + \mathbb{P}[T_{1,L'} \cup T_{2,L'}]}{\mathbb{P}[E_{1,L} \cup E_{2,L}] + \mathbb{P}[E_{1,L'} \cup E_{2,L'}]} &(9)\\ &\leq \frac{P(1+\phi^{2\ell_1}\phi_{1:|A|-2-\ell_1-\ell_2}) + (1+\phi^{2\ell_2}\phi_{1:|A|-2-\ell_1-\ell_2})}{(P+1)(\phi^{-1}+\phi^{1+2(\ell_1+\ell_2)}\phi_{1:|A|-2-\ell_1-\ell_2})} &(10)\\ &\overset{b)}{\leq} \frac{1+\phi^{\ell_1+\ell_2}\phi_{1:|A|-2-\ell_1-\ell_2}}{(\phi^{-1}+\phi^{2(\ell_1+\ell_2)}\phi_{2:|A|-1-\ell_1-\ell_2})}. &(11)\end{aligned}$$

Here, the inequality b) follows from Lemma 1. By using $|A| \leq n$ and letting $\ell = \ell_1 + \ell_2$, we obtain

$$R(L \cup L') \leq \frac{1+\phi^{\ell}\phi_{1:n-\ell-2}}{\phi^{-1}+\phi^{2\ell}\phi_{2:n-\ell-1}},$$

which completes the proof of the result in Step 4.

Lemma 2. *If $\phi + \phi^2 < 1 + \phi^n$, for all $\ell \in \mathbb{N}$, one has*

$$\frac{1+\phi^{\ell}\phi_{1:n-\ell-2}}{\phi^{-1}+\phi^{2\ell}\phi_{2:n-\ell-1}} \leq \frac{\phi_{1:n-1}}{1+\phi_{3:n}}.$$

Proof. First, for $\ell \geq 1$,

$$\frac{\phi^{\ell+1}}{\phi_{2\ell+1:2\ell+2} - \phi^{n+\ell}} \geq \frac{1}{\phi+\phi^2-\phi^{n-1}} > 1.$$

Thus,

$$\frac{\phi_{2:\ell+1}}{\phi_{3:2\ell+2} - \phi_{n+1:n+\ell}} = \frac{\sum_{t=1}^{\ell}\phi^{t+1}}{\sum_{t=1}^{\ell}\phi_{2t+1:2t+2} - \sum_{t=1}^{\ell}\phi^{n+t}} > 1.$$

Since we also have $\frac{\phi_{1:n-1}}{1+\phi_{3:n}} < 1$, it follows that

$$\frac{\phi_{1:n-1}}{1+\phi_{3:n}} > \frac{\phi_{1:n-1} - \phi_{2:\ell+1}}{1+\phi_{3:n} - (\phi_{3:2\ell+2} - \phi_{n+1:n+\ell})} = \frac{\phi + \phi^{\ell}\phi_{2:n-\ell-1}}{1+\phi^{2\ell}\phi_{3:n-\ell}}.$$

This proves the claimed result and completes the proof of the lemma. ∎

III Proof of Lemma 4.5

Lemma 4.5 is a corollary of the following lemma.

Lemma 3. *Let $\sigma \sim$ MM(σ_0, ϕ). If two subsets of elements A, A' satisfy $A' = A \cup \{x\}$, where $x \notin A$, and if $u \in A$, then for all $t \in [|A|]$ one has*

1) $\mathbb{P}[\sigma_A(u) \geq t] \leq \mathbb{P}[\sigma_{A'}(u) \geq t]$.

2) $\mathbb{P}[\sigma_A(u) \leq t] \leq \mathbb{P}[\sigma_{A'}(u) \leq t+1]$.

Proof. Because of symmetry, it suffices to prove the first claim only. The left-hand-side of the first inequality equals the probability of the event that the element u is ranked in the t-th position or lower within the set of elements in A. The right-hand-side of the inequality equals the probability of the event that the element u is ranked in the t-th position or lower within the set of elements in A'. Since A' is the union of A and another element $x \notin A$, inserting x into a ranking may only increase the rank of already present elements. □

Now, consider Lemma 4.5 in the main text. Choose an element $x \in S$ if there is such and element that satisfies $\sigma_0(x) > \sigma(u)$. Let $A' = A \cup \{x\}$. Then, the statement of the above result implies that the probability that element u is ranked lower than or equal to its correct rank $\sigma_{0,A}(u)$ will increase if we add an element x to A that is ranked lower than u in σ_0. Therefore, by removing all elements from A that are ranked lower than u in σ_0, we obtain a new subset A'' and consequently have $\mathbb{P}[\sigma_A(u) \geq \sigma_{0,A}(u)] \geq \mathbb{P}[\sigma_{A''}(u) \geq \sigma_{0,A''}(u)]$. Note that u is the element with the lowest rank in $\sigma_{0,A''}$. Therefore, it is easy to check that $\mathbb{P}[\sigma_{A''}(u) = \sigma_{0,A''}(u)] \geq \frac{1}{\phi_{0:|A''|-1}}$. Then, one has the inequality $\mathbb{P}[\sigma_A(u) \geq \sigma_{0,A}(u)] \geq \frac{1}{\phi_{0:|A''|-1}}$.

IV Proof of Lemma 4.7

For convenience, we restate Lemma 4.7 first.

Lemma 4. *Let $\sigma \sim GMM(\sigma_0, \phi)$ and let A be a subset of elements containing an element u. Let $A' = A - \{x \in A : x \neq u, \sigma_{0,A}(x) \leq \sigma_{0,A}(u)\}$. Define $W = \{r_{\sigma_A(u)} \leq r_{\sigma_{0,A}(u)}\}$, $Q_j = \{r_{\sigma_A(u)} = j + r_{\sigma_{0,A}(u)},\ l_{\sigma_A(u)} \leq r_{\sigma_{0,A}(u)}\}$, $W' = \{r_{\sigma_{A'}(u)} \leq r_{\sigma_{0,A'}(u)}\}$ and $Q'_j = \{r_{\sigma_{A'}(u)} = j + r_{\sigma_{0,A'}(u)},\ l_{\sigma'_A(u)} \leq r_{\sigma_{0,A'}(u)}\}$. Then, the following two claims hold.*

$$
\begin{aligned}
\mathbb{P}[W] + & \sum_{j=1}^{|A|-r_{\sigma_{0,A}}(u)} V_j \mathbb{P}[Q_j] \\
\geq & \mathbb{P}[W'] + \sum_{j=1}^{|A'|-r_{\sigma_{0,A'}}(u)} \frac{1}{j+1} V_j \mathbb{P}[Q'_j] \qquad (12) \\
\geq & 1 - \frac{1}{2}\phi^{1/2} - \frac{1}{2}\phi. \qquad (13)
\end{aligned}
$$

Proof. The idea behind the proof is similar to that of the proof of Lemma 3 in this supplement. Our first goal is to show that removing the element x from A that is ranked highest in $\sigma_{0,A}$ decreases the left-hand-side of (12). Then, by induction, we may prove (12).

For simplicity, let $A'' = A - \{x\}$. Note that because of the choice of the rank of the element x in $\sigma_{0,A}$, we have $r_{\sigma_{0,A''}(u)} = r_{\sigma_{0,A}(u)} - 1$ and $l_{\sigma_{0,A''}(u)} = l_{\sigma_{0,A}(u)} - 1$. For a ranking $\sigma \in \{\sigma : l_{\sigma_A}(u) \leq r_{\sigma_{0,A}}(u)\}$, the removal of x produces another ranking σ''. When $r_{\sigma_A(x)} < r_{\sigma_A(u)}$, σ and σ'' will contribute the same "vote" to the left-hand-side of (12). When $r_{\sigma_A(x)} = r_{\sigma_A(u)}$, σ'' contributes the same vote when $\sigma \in W$, or smaller vote when $\sigma \in Q_j$ for some j. When $r_{\sigma_A(x)} > r_{\sigma_A(u)}$, σ'' will always contribute a smaller vote. For a ranking $\sigma \notin \{\sigma : l_{\sigma_A}(u) \leq r_{\sigma_{0,A}}(u)\}$, both σ and σ'' contribute a zero vote. Therefore, removing x from A strictly decreases the left-hand-side of (12).

We now prove inequality (13). Note that due to the definition of A', we have $r_{\sigma_{0,A'}}(y) = 1$. Also, due to Lemma 5 of this document, we can further assume that $r_{\sigma_0}(y) = 1$, i.e., that y is the only element in the first bucket of σ_0. Because of its definition, W' includes the rankings σ such that y is the only element in the first bucket of σ'_A while Q'_j includes the rankings σ such that there are, in addition to y, some j other elements in the first bucket of σ'_A.

Partition W' according to the size of the second bucket of σ'_A, i.e., let $W' = \cup_{|\mathcal{B}_2(\sigma'_A)|=j} W'_j$. Then, we can construct a bijection from W'_j to Q'_j by putting y into the second bucket. It is easy to check that

$$\mathbb{P}[Q'_j] \leq \mathbb{P}[W'_j]\phi^{j/2}.$$

Denote the set of partial rankings σ for which the first bucket of σ_A does not contain y but some j other elements in U'_j. We can also construct a mapping from Q'_j to U'_j by moving y from the first bucket to any other possible position higher than the first bucket. Hence, we have

$$\mathbb{P}[U'_j] \leq \mathbb{P}[Q'_j]\phi^{j/2}(1+\phi^{1/2}+\phi+\cdots) \leq \mathbb{P}[Q'_j]\frac{\phi^{j/2}}{1-\phi^{1/2}}.$$

Let $Z'_j = W'_j \cup Q'_j \cup U'_j$, so that $\cup_j Z'_j$ covers all possible partial rankings. Hence,

$$\frac{\mathbb{P}(W'_j)+\frac{1}{j+1}\mathbb{P}(Q'_j)}{\mathbb{P}(Z'_j)} = \frac{\mathbb{P}(W'_j)+\frac{1}{j+1}\mathbb{P}(Q'_j)}{\mathbb{P}(W'_j)+\mathbb{P}(Q'_j)+\mathbb{P}(U'_j)} \geq \frac{1+\frac{1}{j+1}\phi^{j/2}}{1+\phi^{j/2}+\frac{\phi^j}{1-\phi^{1/2}}} \geq \frac{1+\frac{1}{2}\phi^{1/2}}{1+\phi^{1/2}+\frac{\phi}{1-\phi^{1/2}}} = 1-\frac{1}{2}\phi^{1/2}-\frac{1}{2}\phi,$$

where the second inequality follows from $\phi^{1/2}+\phi<1$. This completes the proof. $\square$

V Proof of the Main Results

I Proof for permutation aggregation

Let the Lehmer code of the output permutation σ be denoted by $\hat{\boldsymbol{c}}_\sigma$. We say that the LCA algorithm succeeds if $\sigma = \sigma_0$, or equivalently, if $\hat{\boldsymbol{c}}_\sigma = \boldsymbol{c}_{\sigma_0}$. Given that $\hat{\boldsymbol{c}}_\sigma(1) = 0 = \boldsymbol{c}_{\sigma_0}(1)$, by using the union bound, we arrive at

$$\mathbb{P}[\sigma=\sigma_0] = \mathbb{P}[\hat{\boldsymbol{c}}_\sigma = \boldsymbol{c}_{\sigma_0}] \geq 1-\sum_{t=2}^{n-1}\mathbb{P}[\hat{\boldsymbol{c}}_\sigma(t) \neq \boldsymbol{c}_{\sigma_0}(t)].$$

In Section 4, we explained that the algorithm based on the Lehmer code $\hat{\boldsymbol{c}}_\sigma$ may be viewed as a form of InsertionSort, in which during the t-th iteration one places the element t at the $(t-\hat{\boldsymbol{c}}_\sigma(t))$-th position over the subset of elements $S_t = [t]$. With this specific choice of subset S_t, for any permutation π, we have $\pi_{S_t}(t) = t - \boldsymbol{c}_\pi(t)$. Hence, the event $\{\hat{\boldsymbol{c}}_\sigma(t) \neq \boldsymbol{c}_{\sigma_0}(t)\}$ is equivalent to the event $\{\sigma_{S_t}(t) \neq \sigma_{0,S_t}(t)\}$, which we denote by D_t. For convenience, we let $i \triangleq \sigma_{0,S_t}(t)$.

Given that the ranking $\sigma_k \in \Sigma$ is sampled from a $\text{MM}(\sigma_0,\phi)$ distribution, we also define a random variable $X_k(j) = 1_{\{\sigma_{k,S_t}(t)=j\}}$ to indicate whether the element t is ranked at the j-th position in σ_{k,S_t}. Therefore, $\sum_{k\in[m]} X_k(j)$ equals the number of rankings in Σ in which element t is ranked at the j-th position.

I.I Proof of Theorem 4.1 (Mode)

Given that we aggregate using the mode function, we have $\sigma_{S_t}(t) = \arg\max_{j\in[t]}\sum_{k\in[m]} X_k(j)$. In what follows, we aim to prove an upper bound on $\mathbb{P}[\sigma_{S_t}(t) \neq \sigma_{0,S_t}(t)] = \mathbb{P}[D(t)]$.

To this end, let $q = \frac{\phi_{1:n-1}}{1+\phi_{3:n}}$, so that when $\phi+\phi^2 < 1+\phi^n$, we have $q<1$. Based on Lemma 4.3 in the main text, we have

$$\mathbb{P}[X_k(i)=1] = \frac{\mathbb{P}[X_k(i)=1]}{\sum_{j=1}^t \mathbb{P}[X_k(j)=1]} = \frac{\mathbb{P}[\sigma_{k,S_t}(t)=i]}{\sum_{j=1}^t \mathbb{P}[\sigma_{k,S_t}(t)=j]} \geq \frac{1}{1+2q/(1-q)} = \frac{1-q}{1+q}.$$

Moreover, if $\mathcal{E} = \mathbb{E}[X_k(i)-X_k(j)]$, then

$$\mathcal{E} \geq \mathbb{P}[X_k(i)=1](1-q^{|i-j|}) \geq \frac{(1-q)^2}{1+q}.$$

Therefore, since the σ_k, $k \in [m]$, are i.i.d, Hoeffding's inequality establishes

$$\mathbb{P}[\sum_{k\in[m]} X_k(i) < \sum_{k\in[m]} X_k(j)] = \mathbb{P}[\sum_{k\in[m]} (X_k(i) - X_k(j)) \leq 0] \leq \exp(-\frac{m\mathcal{E}^2}{2}) \leq \exp(-\frac{m(1-q)^4}{2(1+q)^2}).$$

Hence,

$$\mathbb{P}[D_t] \leq \sum_{j\in[t],j\neq i} \mathbb{P}\left[\sum_{k\in[m]} X_k(i) < \sum_{k\in[m]} X_k(j)\right] < (t-1)\exp(-\frac{m(1-q)^4}{2(1+q)^2}).$$

As a result, for $m \geq c\log\frac{n^2}{2\delta}$ with $c = \frac{2(1+q)^2}{(1-q)^4}$ and $q = \frac{\phi_{1:n-1}}{1+\phi_{3:n}}$, we have $\mathbb{P}[\sigma = \sigma_0] > 1-\delta$.

I.II Proof of Theorem 4.3 (Median)

Let $Y_k(j_0) = \sum_{j=1}^{j_0} X_k(j)$. Since we use the median to form the aggregate, we need to establish that $\sigma_{S_t}(t) = \min\{j : \frac{1}{m}\sum_{k\in[m]} Y_k(j) \geq 0.5\}$. According to Lemma 4.5 of the main text, we have $\mathbb{P}[Y_k(i) = 1] = 1 - \mathbb{P}[\sigma_{k,A}(x) > i] \geq 1-\phi$ while $\mathbb{P}[Y_k(i-1) = 1] = \mathbb{P}[\sigma_{k,A}(x) < i] \leq \phi$. Therefore, if $\phi < 0.5$, using Hoeffding's inequality, we have

$$\mathbb{P}[D_t] \leq \mathbb{P}\left[\frac{1}{m}\sum_{k\in[m]} Y_k(i) < 0.5\right] + \mathbb{P}\left[\frac{1}{m}\sum_{k\in[m]} Y_k(i-1) > 0.5\right] \leq 2e^{-2m(\frac{1}{2}-\phi)^2}.$$

As a result, for $m \geq c\log\frac{2n}{\delta}$ with $c = \frac{2}{(1-2\phi)^2}$, we have $\mathbb{P}[\sigma = \sigma_0] > 1 - 2ne^{-2m(\frac{1}{2}-\phi)^2} \geq 1-\delta$.

II Proof of the Performance Guarantees for Partial Ranking Aggregation

Denote the Lehmer code of the output permutation σ of Algorithm 2 of the main text by $\hat{\boldsymbol{c}}_\sigma$. We say that the LCA algorithm succeeds if σ is in Σ_0, which is equivalent to saying that $\hat{\boldsymbol{c}}_\sigma(t) \in [\boldsymbol{c}_\sigma(t), \boldsymbol{c}'_\sigma(t)]$. Given that $\hat{\boldsymbol{c}}_\sigma(1) = 0 = \boldsymbol{c}_{\sigma_0}(1) = \boldsymbol{c}'_\sigma(1)$, from the union bound, we have

$$\mathbb{P}[\sigma \in \Sigma_0] = \mathbb{P}[\hat{\boldsymbol{c}}_\sigma(t) \in [\boldsymbol{c}_\sigma(t), \boldsymbol{c}'_\sigma(t)], \forall t] \geq 1 - \sum_{t=2}^{n-1} \mathbb{P}[\hat{\boldsymbol{c}}_\sigma(t) \notin [\boldsymbol{c}_\sigma(t), \boldsymbol{c}'_\sigma(t)]].$$

In Section 4 of the main text, we explained how the Lehmer code transform $\hat{\boldsymbol{c}}_\sigma$ may be viewed as a form of InsertionSort, which in the t-th iteration places the element t at the $(t - \hat{\boldsymbol{c}}_\sigma(t))$th position within the subset of elements $S_t = [t]$. With this choice of subset S_t, for any π, we have that $\pi_{S_t}(t) = t - \boldsymbol{c}_\pi(t)$. Hence, the event $\{\hat{\boldsymbol{c}}_\sigma(t) \notin [\boldsymbol{c}_\sigma(t), \boldsymbol{c}'_\sigma(t)]\}$ is equivalent to the event $\{\sigma_{S_t}(t) < l_{\sigma_0,S_t(t)}$ or $\sigma_{S_t}(t) > r_{\sigma_0,S_t(t)}\}$, which we denote by D_t. The proof reduces to finding a lower bound on $\mathbb{P}[D_t]$.

For convenience, we let $l \triangleq l_{\sigma_0,S_t(t)}$ and $r \triangleq r_{\sigma_0,S_t(t)}$. Given that the ranking $\sigma_k \in \Sigma$ is sampled from a GMM(σ_0, ϕ), we define the random variable $X_k(j)$ as the vote that σ_k cast for t to be at position j in S_t. Then, $V(j) = \sum_{k\in[m]} X_k(j)$ is the total vote cast by all partial rankings in Σ to rank t at the j-th position.

II.I Proof of Theorem 4.6 (Median)

Let $Y_k(j_0) = \sum_{j=1}^{j_0} X_k(j)$. Since we use the median to form the aggregate, we have $\sigma_{S_t}(t) = \min\{j : \frac{1}{m}\sum_{k\in[m]} Y_k(j) \geq 0.5\}$. Define the event $W = \{r_{\sigma_{k,S_t}(t)} \leq r\}$. When W occurs, σ_k contributes 1 to $Y_k(r)$. Let $Q = \cup_{j=1}^{n-r} Q_j$, where $Q_j = \{r_{\sigma_{k,S_t}(t)} = j + r,\ l_{\sigma_{k,S_t}(u)} \leq r\}$. When Q_j occurs, σ_k contributes a fractional vote V_j to $Y_k(r)$, where $V_j = \frac{r - l_{\sigma_{k,S_t}(t)} + 1}{r_{\sigma_{k,S_t}(t)} - l_{\sigma_{k,S_t}(t)} + 1} \geq V'_j = \frac{1}{j+1}$. In fact, $V_j = V'_j$ when $l_{\sigma_{k,S_t}(t)} = r$. Therefore, based on the Lemma 4.7 of the main text, we have

$$\mathbb{E}[Y_k(r)] \geq \mathbb{P}[W] + \sum_{j=1}^{t-r} \frac{1}{j+1}\mathbb{P}[Q_j] \tag{14}$$

$$\geq 1 - \frac{1}{2}\phi^{1/2} - \frac{1}{2}\phi. \tag{15}$$

Let $q' = 1 - \frac{1}{2}\phi^{1/2} - \frac{1}{2}\phi$. When $\phi^{1/2} + \phi < 1$, it follows that $q' > 0.5$. By using Hoeffding's inequality, we obtain

$$\mathbb{P}\left[\frac{1}{m}\sum_{k\in[m]} Y_k(r) < 0.5\right] \leq \exp(-2m(1/2 - q')^2).$$

Let $Z_k(j_0) = \sum_{j=j_0}^{t} X_k(j)$. In an analogous manner, we can prove that

$$\mathbb{P}\left[\frac{1}{m}\sum_{k\in[m]} Z_k(l) < 0.5\right] \leq \exp(-2m(1/2 - q')^2).$$

Therefore, the probability of success of iteration t may be bounded as

$$\mathbb{P}[D_t] \leq \mathbb{P}\left[\frac{1}{m}\sum_{k=1}^{m} Y_k(r) < 0.5\right] + \mathbb{P}\left[\frac{1}{m}\sum_{k=1}^{m} Z_k(l) < 0.5\right] \leq 2e^{-2m(1/2-q')^2}.$$

As a result, when $m \geq c \log \frac{2n}{\delta}$ with $c = \frac{2}{(1-2q')^2}$, where $q' = 1 - \frac{1}{2}\phi^{1/2} - \frac{1}{2}\phi$, we have $\mathbb{P}[\sigma \in \Sigma_0] > 1 - \delta$.

VI Other Lemmas and Proofs

Lemma 5. *Let σ_0 be a ranking over S and let $A \subseteq S$ be such that A contains the elements ranked highest in σ_0. Consider a ranking $\sigma \sim MM(\sigma_0, \phi)$. Then, the marginal distribution of σ over S/A is the distribution $MM(\sigma_{0,S/A}, \phi)$.*

Proof. It suffices to prove the result for $A = \{x\}$, where x is the element ranked highest in σ_0, as this result may be applied inductively. Consider all permutations σ such that for $\sigma_{S/\{x\}} = \pi$ and some $j \in [|S|]$, one has

$$\sigma^{-1}(t) = \begin{cases} \pi^{-1}(t), & 1 \leq t < j, \\ \pi^{-1}(t-1), & j < t \leq |S|, \\ x, & t = j. \end{cases}$$

For simplicity of notation, we use $\sigma^{(j)}$ to denote a permutation with the above property. Then,

$$\begin{aligned} \mathbb{P}[\sigma_{S/\{x\}} = \sigma'] &= \sum_{j=1}^{|S|} \mathbb{P}[\sigma^{(j)}] = \frac{1}{Z_{|S|}} \sum_{j=1}^{|S|} \phi^{d_\tau(\sigma^{(j)},\sigma_0)} = \frac{1}{Z_{|S|}} \sum_{j=1}^{|S|} \phi^{j-1+d_\tau(\sigma^{(j)}_{S/\{x\}},\sigma_{0,S/\{x\}})} \\ &= \frac{\sum_{j=1}^{|S|} \phi^{j-1}}{Z_{|S|}} \phi^{d_\tau(\pi,\sigma_{0,S/\{x\}})} = \frac{1}{Z_{|S/\{x\}|}} \phi^{d_\tau(\pi,\sigma_{0,S/\{x\}})}, \end{aligned}$$

where $Z_n = \prod_{i=1}^{n-1} \sum_{j=0}^{i} \phi^j$ denotes the normalization constant in the Mallows distribution of permutations with n elements. $\square$

Observe that the same result holds when A is assumed to contain the lowest ranked element in σ_0.

VII Supplementary Algorithms

I Efficient Algorithms for Computing the Mode/Median for Partial Ranking Aggregation

In Section 3 of the main text which discusses partial ranking aggregation, we pointed out that one can efficiently compute the voting function $V_x(y)$, and hence the mode/median $\hat{c}$ as well. Algorithm VII.1 of this text explains how to efficiently compute $V_x(y)$, provided that for fixed k, x, $v_{k\to x}(y)$ is positive over a continuous interval, or more precisely, when $[x - c_{\sigma_k}(x), x - c_{\sigma'_k}(x)]$. Algorithm VII.1 has complexity $\mathcal{O}(m + x)$ and the computation of the mode/median of the component $\hat{c}(x)$ requires $\mathcal{O}(x)$ time. Therefore, the total complexity of Algorithm 2 of the main text for partial rankings equals $\mathcal{O}(mn + n^2)$.

Algorithm VII.1:
Computing $\{V_x(y)\}_{y\in[x]}$ **when** $v_{k\to x}(y)$ **is positive over** $[x-c_{\sigma_k}(x),\, x-c_{\sigma'_k}(x)]$
Input: $\boldsymbol{c}_{\sigma_k}, \boldsymbol{c}'_{\sigma_k}$, votes $v_{k\to x}(y) = v_{k\to x}1_{x-\boldsymbol{c}'_{\sigma_k}\le y\le x-\boldsymbol{c}_{\sigma_k}}$, $\forall\, k\in[m]$;

```
Initialize V_x(y) = 0, for all y ∈ [x+1];
For k from 1 to m do
    V_x(x − c'_{σ_k}(x)) = V_x(x − c'_{σ_k}(x)) + v_{k→x};
    V_x(x + 1 − c_{σ_k}(x)) = V_x(x + 1 − c_{σ_k}(x)) − v_{k→x};
For k from 2 to x+1 do
    V_x(y) = V_x(y − 1) + V_x(y);
Output: V_x(y);
```

II A Kemeny-Distance Optimal Algorithm for Transforming Permutations into Partial Rankings

In Section 3 of the main text pertaining to partial ranking aggregation, we pointed out that one can optimally transform the permutation output of Algorithm 2 into a partial ranking. Algorithm VII.2 of this text explains how to perform this transform. In the description of the algorithm, we used $a:b=(a,a+1,...,b)$ where $a,b\in\mathbb{Z}$, $b\geq a$. For a vector V, we used $V(a:b)$ to denote $(V(a),V(a+1),...,V(b))$. Algorithm VII.2 has complexity $\mathcal{O}(mn^2+n^3)$.

References

[1] Pranjal Awasthi, Avrim Blum, Or Sheffet, and Aravindan Vijayaraghavan, "Learning mixtures of ranking models," in *Advances in Neural Information Processing Systems*, 2014, pp. 2609–2617.

Algorithm VII.2:
Transforms a Permutation into a Partial Ranking that is Kemeny-Distance Optimal
Input: Permutation σ, Set of partial rankings Σ;
1: **Initialize** BucketSize$= (1, 1, ..., 1) \in \mathbb{N}^n$;
2: **Initialize** $W = \{w_{ij}\}_{i,j\in[n]}$ where $w_{ij} = \frac{1}{m}\sum_{k\in[m]} 1_{\sigma_k(\sigma^{-1}(i))<\sigma_k(\sigma^{-1}(j))}$;
3: [Val, BucketSize]=Dynamic-Programming(W, BucketSize);
4: Construct a partial ranking σ' by putting the lowest BucketSize(1) many elements of σ into $\mathcal{B}_1(\sigma')$; Proceed by taking BucketSize(2) many elements of σ and placing them into $\mathcal{B}_2(\sigma')$ and so on;
5: **Output:** σ'.

Dynamic-Programming(W, BucketSize)
1: n' =length(BucketSize);
2: **If** $n' = 1$
  return [0, BucketSize];
3: $s = \lfloor n'/2 \rfloor$;
When $\sigma^{-1}(s)$ and $\sigma^{-1}(s+1)$ are in different buckets (4-6)
4: [Val1, BucketSize1]=Dynamic-Programming($W(1:s, 1:s)$, BucketSize$(1:s)$);
5: [Val2, BucketSize2]=Dynamic-Programming($W(s+1:n', s+1:n')$, BucketSize$(s+1:n)$);
6: Val-div=Val1+Val2+$\sum_{i=s+1}^{n'}\sum_{j=1}^{s} w_{ij} + \frac{1}{2}\sum_{i=s+1}^{n'}\sum_{j=1}^{s}$(BucketSize$(i) *$ BucketSize$(j) - w_{ji} - w_{ij}$).
When $\sigma^{-1}(s)$ and $\sigma^{-1}(s+1)$ are in the same bucket (7-13)
7: $w_{si} = w_{si} + w_{(s+1)i}$, for all $i \in [n']$;
8: $w_{is} = w_{is} + w_{i(s+1)}$, for all $i \in [n']$;
9: Val3=$1/2 * w_{ss}$;
10: Construct $W' \in \mathbb{R}^{n'-1\times n'-1}$ by deleting the $s+1$th row and $s+1$th column of W;
11: Construct newBucketSize:
  newBucketSize(i)=BucketSize(i) for $1 \le i \le s$;
  newBucketSize(i)=BucketSize($i+1$) for $s+1 \le i \le n'-1$;
  newBucketSize(s)=BucketSize(s)+BucketSize($s+1$);
12: [Val4, BucketSize3]=Dynamic-Programming(W', newBucketSize);
13: Val-con= Val3+Val4;
14: **if** Val-con>Val-div,
  Construct BucketSize4 via concatenation of BucketSize1 and BucketSize2;
  return [Val-div, BucketSize4];
15: **else** return [Val-con, BucketSize3];