Submitted to the Annals of Statistics
arXiv: arXiv:1606.08842

ACTIVE RANKING FROM PAIRWISE COMPARISONS AND WHEN PARAMETRIC ASSUMPTIONS DON'T HELP

By Reinhard Heckel[§,*], Nihar B. Shah[¶,†,‡], Kannan Ramchandran[‖] and Martin J. Wainwright[‖,†]

Rice University[§], Carnegie Mellon University[¶], and University of California, Berkeley[‖]

We consider sequential or active ranking of a set of n items based on noisy pairwise comparisons. Items are ranked according to the probability that a given item beats a randomly chosen item, and ranking refers to partitioning the items into sets of pre-specified sizes according to their scores. This notion of ranking includes as special cases the identification of the top-k items and the total ordering of the items. We first analyze a sequential ranking algorithm that counts the number of comparisons won, and uses these counts to decide whether to stop, or to compare another pair of items, chosen based on confidence intervals specified by the data collected up to that point. We prove that this algorithm succeeds in recovering the ranking using a number of comparisons that is optimal up to logarithmic factors. This guarantee does depend on whether or not the underlying pairwise probability matrix, satisfies a particular structural property, unlike a significant body of past work on pairwise ranking based on parametric models such as the Thurstone or Bradley-Terry-Luce models. It has been a long-standing open question as to whether or not imposing these parametric assumptions allows for improved ranking algorithms. For stochastic comparison models, in which the pairwise probabilities are bounded away from zero, our second contribution is to resolve this issue by proving a lower bound for parametric models. This shows, perhaps surprisingly, that these popular parametric modeling choices offer at most logarithmic gains for stochastic comparisons.

[*]Corresponding author. Supported by the Swiss National Science Foundation under grant P2EZP2_159065.

[†]Supported in part by Office of Naval Research MURI grant DOD-002888, Air Force Office of Scientific Research Grant AFOSR-FA9550-14-1-001, Office of Naval Research grant ONR-N00014, as well as National Science Foundation Grant CIF-31712-23800.

[‡]Supported in part by National Science Foundation grants CIF 1763734 and CIF 1755656.

MSC 2010 subject classifications: Primary 68T99

Keywords and phrases: active learning, online learning, pairwise comparisons, ranking, multi-armed bandits, Bradley-Terry-Luce model

1. Introduction. Given a collection of n items, it is frequently of interest to estimate a ranking based on noisy comparisons between pairs of items. Such rank aggregation problems arise across a wide range of applications. Some traditional examples in sports include identifying the best player in a tournament, selecting the top k teams for playoffs, and finding the full ranking of players. More recently, the internet era has led to a variety of applications involving pairwise comparison data, including recommender systems [28, 1] for rating movies, books, or other consumer items; peer grading [31] for ranking students in massive open online courses; and online sequential survey sampling [29] for assessing the popularity of proposals in a population of voters. In many of these and other such applications, it is possible to make comparisons in an active or adaptive manner—that is, based on the outcomes of comparisons of previously chosen pairs. Motivated by those applications, the focus of this paper is the problem of obtaining statistically sound rankings based on a sequence of actively chosen pairwise comparisons.

We consider a collection of n items, and our data consists of outcomes of comparisons between pairs of items in this collection that are collected in a sequential fashion, also known as the active setting. We assume that the outcomes of comparisons are stochastic—that is, item i beats item j with an unknown probability $M_{ij} \in (0, 1)$. The outcomes of pairwise comparisons are furthermore assumed to be statistically mutually independent. We define the ordering of the items in terms of their (unknown) scores, where the score τ_i of item i is defined as the probability that item i beats an item chosen uniformly at random from all other items:

$$\tau_i := \frac{1}{n-1} \sum_{j \neq i} M_{ij}. \tag{1.1}$$

In the context of social choice theory [12], these sums are also known as the *Borda scores or counts* of the items. Apart from their intuitive appeal, the Borda counts are of particular interest because they provide a natural unification of the assumed orderings in several popular comparison models. Specifically, the parametric Bradley-Terry-Luce (BTL) [5, 24] and Thurstone [36] models, as well as the non-parametric Strong Stochastic Transitivity (SST) model [37], are all based on an assumed ordering of the items; in all of these models, this ordering coincides with that given by the scores $\{\tau_i\}_{i=1}^n$. In this paper, we consider the problem of partitioning the items into sets of pre-specified sizes according to their respective scores. This notion of ranking includes as special cases identification of the top-k items and the total ordering of the items.

We make two primary contributions. We begin by presenting and analyzing a simple active ranking algorithm for estimating a partial or total ranking of the items. At each round, this algorithm first counts the number of comparisons won, then computes confidence bounds from those counts, which it finally uses to select a subset of pairs to be compared at the next time step. We provide performance guarantees showing that with high probability, the algorithm recovers the desired partial or total ranking from a certain number of comparisons. We refer to this sample size as the *sample complexity*, and show that it is a function of the (unknown) scores $\{\tau_i\}_{i=1}^n$, and therefore distribution-dependent. Conversely, we prove distribution-dependent lower bounds that are matching up to logarithmic factors, thereby showing that the algorithm is near-optimal in the number of comparisons. Our analysis leverages the fact that ranking in terms of the scores $\{\tau_i\}_{i=1}^n$ is related to a particular class of multi-armed bandit problems [14, 7, 38]. We note that this connection has been used in past work [40, 19, 38] in the context of finding the top item.

Our second main contribution relates to the popular parametric modeling choices made in the literature. On one hand, the algorithmic analysis of this paper does not impose any assumptions on the pairwise comparison probabilities. On the other hand, much past work (including some of our own) is based on specific parametric assumptions on the pairwise comparisons; for instance, see the papers [35, 16, 26, 15, 9, 34, 32, 25] as well as references therein. Concrete examples of parametric assumptions include the Bradley-Terry-Luce (BTL) and Thurstone parametric models. There is a long standing debate on whether such parametric assumptions are reasonable—that is, in which situations they (approximately) hold, and in which they fail [3]. When such parametric models are suitable, the natural hope is that their structure allows some reduction of the sample complexity. In fact, for essentially deterministic comparison models (meaning that pairwise comparison probabilities may be arbitrarily close to zero or one), there can indeed be significant gains; see the discussion following Theorem 2 for further details. However, as we show in the paper, if one considers *stochastic comparison models* (in which the pairwise probabilities are bounded away from zero and one), then assuming a parametric comparison model versus not making any structural assumption leads to at most a logarithmic gain in the sample complexity. This logarithmic gain needs to be weighed against the potential lack of robustness incurred by using a parametric model—note that parametric modeling assumptions often hold only approximately [3], if at all—which can be significant, as shown in our numerical results section.

Related work: There is a vast literature on ranking and estimation

from pairwise comparison data. Most works assume probabilistic comparison outcomes; we refer to the paper [17] and references therein for ranking problems assuming deterministic comparison outcomes. Several prior works [16, 26, 15, 34, 32, 30, 11] consider settings where pairs to be compared are chosen a priori. In contrast, we consider settings where the pairs may be chosen in an active manner. The recent work [35] assumes the Bradley-Terry-Luce (BTL) parametric model, and considers the problem of finding the top item and the full ranking in an active setup. In the stochastic regime, for certain underlying distributions, the corresponding results [35, Theorem 3 and Theorem 4] are close to what our more general result implies. On the other hand, for several other problem instances, the performance guarantees of Theorem 3 and Theorem 4 in the work [35] lead to a significantly larger sample complexity. Our work thus offers better guarantees for the BTL model in the stochastic regime, despite the additional generality of our setting in that we do not restrict ourselves to the BTL model. However outside the stochastic regime, specifically for models with pairwise comparison probabilities very close to zero and one, [35, Theorem 3 and Theorem 4] offer gains over the results afforded by our more general model; we discuss this regime in more detail later. The paper [25] considers the problem of finding a full ranking of items for a BTL pairwise comparison model, and provides a performance analysis for a probabilistic model on the BTL parameter vector. Eriksson [13] considers the problem of finding the very top items using graph based techniques, whereas Busa-Fekete et al. [8] consider the problem of finding the top-k items. Ailon [2] considers the problem of linearly ordering the items so as to disagree in as few pairwise preference labels as possible. Our work is also related to the literature on multi-armed bandits, and we revisit these relations later in the paper.

Organization: The remainder of this paper is organized as follows. We begin with background and problem formulation in Section 2. We then present a description and a sharp analysis of our ranking algorithm in Section 3. In Section 4, we show that parametric assumptions do not reduce the sample complexity in the stochastic regime. In Section 5, we study numerically whether algorithms designed for parametric models can yield some improvement outside the stochastic regime, and study some additional aspects of our proposed algorithm. We provide proofs of all our results in Section 6, and conclude with a discussion in Section 7.

2. Problem formulation and background. In this section, we formally state the ranking problem considered in this paper and formalize the notion of an active ranking algorithm. We also formally introduce the class

of parametric models in this section.

2.1. *Pairwise probabilities, scores, and rankings.* Given a collection of items $[n] \coloneqq \{1, \dots, n\}$, let us denote by $M_{ij} \in (0,1)$ the (unknown) probability that item i wins a comparison with item j. For all items i and j, we require that each comparison results in a winner (meaning that $M_{ij}+M_{ji} = 1$), and we set $M_{ii} = 1/2$ for concreteness. For each item $i \in [n]$, consider the score (1.1) given as $\tau_i \coloneqq \frac{1}{n-1}\sum_{j\in[n]\setminus\{i\}} M_{ij}$. Note that the unknown score $\tau_i \in (0,1)$ corresponds to the probability that item i wins a comparison with an item j chosen uniformly at random from $[n] \setminus \{i\}$.

Assuming that the scores are all distinct, they define a unique ranking of the n items; more specifically, this unknown ranking is defined by the permutation $\pi : [n] \to [n]$ such that $\tau_{\pi(1)} > \tau_{\pi(2)} > \ldots > \tau_{\pi(n)}$. In words, $\pi(i)$ denotes the i^{th} ranked item according to the scores. A number of ranking problems can be defined in terms of π: at one extreme, finding the best item corresponds to determining the item $\pi(1)$, whereas at the other extreme, finding a complete ranking is equivalent to estimating $\pi(j)$ for all $j \in [n]$. We introduce a general formalism that allows us to handle these and many other ranking problems. In particular, given an integer $L \geq 2$, we let $\{k_\ell\}_{\ell=1}^L$ be a collection of positive integers such that $1 \leq k_1 < k_2 < \ldots < k_{L-1} < k_L = n$. Any such collection of positive integers defines a partition of $[n]$ into L disjoint sets of the form

$$\begin{aligned}(2.1)\qquad \mathcal{S}_1 \coloneqq \{\pi(1), \dots, \pi(k_1)\},\ \ \mathcal{S}_2 \coloneqq \{\pi(k_1+1), \dots, \pi(k_2)\},\\ \dots,\ \ \mathcal{S}_L \coloneqq \{\pi(k_{L-1}+1), \dots, \pi(n)\}.\end{aligned}$$

For instance, if we set $L = 2$ and $k_1 = k$, then the set partition $(\mathcal{S}_1, \mathcal{S}_2)$ corresponds to splitting $[n]$ into the top k items and its complement. At the other extreme, if we set $L = n$ and $(k_1, k_2, \dots, k_n) = (1, 2, \dots, n)$, then the partition $\{\mathcal{S}_\ell\}_{\ell=1}^L$ allows us to recover the full ranking of the items, as specified by the permutation π.

For future reference, we define the set

$$\begin{aligned}(2.2)\qquad \mathcal{C}_{M_{\min}} \coloneqq \{M \in (0,1)^{n\times n} \mid M_{ij} = 1 - M_{ji}, M_{ij} \geq M_{\min},\\ \text{and } \tau_i \neq \tau_j \text{ for all } (i,j)\},\end{aligned}$$

corresponding to the set of pairwise comparison matrices with pairwise comparison probabilities lower bounded by $M_{\min}$, and for which a unique ranking exists. We note that our results actually do not require the entire underlying ordering of the scores to be strict; rather, we require strict inequalities only at the boundaries of the sets $\mathcal{S}_1, \dots, \mathcal{S}_L$.

2.2. *The active ranking problem.* An active ranking algorithm acts on a pairwise comparison model $M \in \mathcal{C}_0$. Consider any specified values of L and $\{k_\ell\}_{\ell=1}^L$, which define a partition of the form (2.1) in terms of their latent scores (1.1). The goal is to obtain a partition of the items $[n]$ into L disjoint sets of the form (2.1) from active comparisons. At each time instant, the algorithm can compare two arbitrary items, and the choice of which items to compare may be based on the outcomes of previous comparisons. As a result of comparing two items i and j, the algorithm receives an independent draw of a binary random variable with success probability M_{ij} in response. After termination dictated by an associated stopping rule, the algorithm returns a ranking $\widehat{\mathcal{S}}_1, \ldots, \widehat{\mathcal{S}}_L$.

For a given tolerance parameter $\delta \in (0,1)$, we say that a ranking algorithm $\mathcal{A}$ is *δ-accurate for a comparison matrix M* if the ranking it outputs obeys

$$\mathbb{P}_M \left[\widehat{\mathcal{S}}_\ell = \mathcal{S}_\ell, \text{ for all } \ell = 1, \ldots, L \right] \geq 1 - \delta. \tag{2.3}$$

For any set of comparison matrices $\mathcal{C}$, we say that the algorithm $\mathcal{A}$ is *uniformly δ-accurate* over $\mathcal{C}$ if it is δ-accurate for each matrix $M \in \mathcal{C}$. The performance of any algorithm is measured by means of its *sample complexity*, by which we mean the number of comparisons required to obtain the desired partition.

2.3. *Active ranking and multi-armed bandits.* It is worthwhile noting that the ranking problem studied here is related to multi-armed bandits [22, 6]. More precisely, a (stochastic) multi-armed bandit model consists of a collection of n "arms", each associated with an unknown and stochastic reward function. When an arm is "pulled", a reward is drawn i.i.d. from a corresponding distribution, and the goal is to maximize the reward obtained via a sequential choice of arms. In past work, various researchers (e.g., [39, 40, 38, 19]) have drawn links between pairwise comparison ranking and such bandit problems. In particular, by definition of the score τ_i, comparing item i to a distinct item chosen from the $n-1$ alternatives can be modeled as drawing a Bernoulli random variable with mean τ_i. Our subsequent analysis in Section 3 relies on this relation. When cast in the multi-armed bandit setting, the setting of pairwise comparisons is often referred to as that of "dueling bandits". Prior works in this setting [39, 40, 19] address the problem of finding the single "best arm"— meaning the item with the highest score—or the set of top-k arms [38], based on noisy comparisons. By contrast, this paper treats the more general problem of finding a partial or total ordering of the items.

Despite these similarities, there is an important distinction between the two settings. If we view our problem as a multi-armed bandit problem with Bernoulli random variables with means $\{\tau_i\}_{i=1}^n$, these means are actually coupled together, in the sense that information about any particular mean imposes constraints on all the other means. In particular, any set of scores $\{\tau_i\}_{i=1}^n$ must be realized by some valid set of pairwise comparison probabilities $\{M_{ij}\}_{i,j\in[n]}$. Since these pairwise comparison probabilities must obey the constraint $M_{ij} = 1 - M_{ji}$, the induced scores must satisfy certain constraints, not all of which are obvious. One obvious constraint, which follows immediately from the definition (1.1), is that $\sum_{i=1}^n \tau_i = n/2$. Another less obvious constraint is the collection of inequalities $\sum_{i=1}^j \tau_{\pi(i)} \geq \frac{1}{n-1}\frac{j(j-1)}{2}$ for $j = 2, \ldots, n-1$; see the papers [23, 20] for discussion. These conditions, while necessary, are certainly not sufficient, as can be seen by studying some simple cases.[1] Our algorithm, presented in the next section, does not take the coupling of the scores explicitly into account. Nevertheless, our algorithm is shown to be optimal up to a logarithmic factor in the stochastic regime.

2.4. *Parametric models.* In this section, we introduce a family of parametric models that form a basis of several prior works [35, 16, 26, 15, 32]. To be clear, we make no modeling assumptions for our algorithm and its analysis in Section 3. Rather, we focus on these parametric models in Section 4, where we show that, perhaps surprisingly, outside of the deterministic regime, none of these parametric assumptions provide more than a logarithmic gain in sample complexity.

Any member of this family is defined by a strictly increasing and continuous function $\Phi\colon \mathbb{R} \to [0,1]$ such that $\Phi(t) = 1 - \Phi(-t)$, for all $t \in \mathbb{R}$. The function Φ is assumed to be known. A pairwise comparison matrix in this family is associated to an unknown vector $w \in \mathbb{R}^n$, where each entry of w represents some quality or strength of the corresponding item. The parametric model $\mathcal{C}_{\mathrm{PAR}}(\Phi)$ associated with the function Φ is defined as:

$$\mathcal{C}_{\mathrm{PAR}}(\Phi) = \{M_{ij} = \Phi(w_i - w_j) \text{ for all } i,j \in [n], \text{ for some } w \in \mathbb{R}^n\}. \tag{2.4}$$

Popular examples of models in this family are the Bradley-Terry-Luce (BTL) model, obtained by setting Φ equal to the sigmoid function ($\Phi(t) = \frac{1}{1+e^{-t}}$), and the Thurstone model, obtained by setting Φ equal to the Gaussian CDF. Note that $\tau_1 > \tau_2 > \ldots > \tau_n$ is equivalent to $w_1 > w_2 > \ldots > w_n$,

[1] For instance, there is no set of pairwise comparison probabilities with scores $[1,1,0,0]$, even though those scores satisfy the aforementioned constraints. In order to verify this fact, note that $\tau_1 = 1$ implies $M_{12} = M_{13} = M_{14} = 1$. Thus, we have $M_{21} = 0$, which implies $\tau_2 \leq 2/3$ and therefore contradicts $\tau_2 = 1$.

meaning that the ranking induced by the scores $\{\tau_i\}_{i=1}^n$ is equivalent to that induced by w.

It is worthwhile noting that a common assumption in the setting of parametric models [26, 32, 9] is that $\|w\|_\infty \leq B$ for some finite constant B. This boundedness assumption implies that the pairwise comparison probabilities $\{M_{ij}\}_{i,j=1}^n$ are all uniformly bounded away from 0 and 1, thereby guaranteeing a stochastic comparison model.

3. Active ranking from pairwise comparisons. In this section, we present our algorithm for obtaining the desired partition of the items as described earlier in Section 2, and a sharp analysis of this algorithm proving its optimality up to logarithmic factors.

3.1. *Active ranking (AR) algorithm.* Our active ranking algorithm is based on the following two ingredients:

- Successive estimation of the scores $\{\tau_i\}_{i=1}^n$, where τ_i is estimated by comparing item i with items chosen uniformly at random from $[n]\backslash\{i\}$.
- Assigning an item i to an estimate $\widehat{\mathcal{S}}_\ell$ of the set $\mathcal{S}_\ell$ once a certain confidence level of i belonging to $\mathcal{S}_\ell$ is attained.

This strategy is essentially an adaption of the successive elimination approach from the bandit literature, proposed in the classic paper [27], and studied in a long line of subsequent work (see, for example, the papers [14, 7, 38, 19]). While we focus on an elimination strategy in this paper, we note that an algorithm based on carefully using lower and upper confidence bound (LUCB) information, in a manner similar to the top-k arm identification LUCB algorithm [21, 10], would result in the same sample complexity.

The first input to the algorithm is a collection of positive integers $\{k_\ell\}_{\ell=0}^L$ such that $k_0 = 0 < k_1 < k_2 < \ldots < k_{L-1} < k_L = n$, which define a desired ranking. The second input is a tolerance parameter $\delta \in (0,1)$, which defines the probability with which the algorithm is allowed to fail. Finally, our algorithm uses a confidence bound based on an non-asymptotic version of the law of the iterated algorithm [22, 18], that takes on the form $\alpha_t \propto \sqrt{\frac{\log(\log(t)n/\delta)}{t}}$, for time $t \in \{1, 2, \ldots\}$. We explicitly choose the constants involved by setting

$$\alpha_t = \sqrt{\frac{\beta(t, \delta/n)}{2t}}, \tag{3.1}$$

with $\beta(t, \delta') = \log(1/\delta') + 0.75 \log\log(1/\delta') + 1.5 \log(1 + \log(t/2))$.

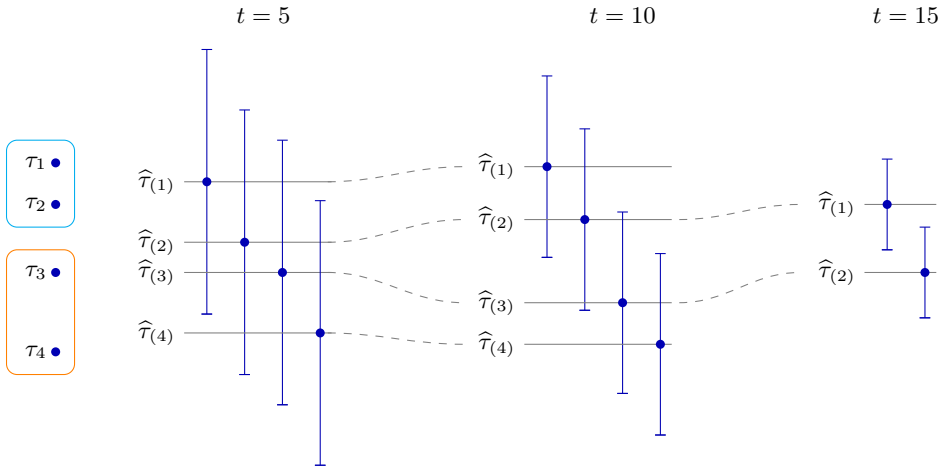


Fig 1: Illustration of the AR algorithm applied to the problem of finding the top 2 items out of $n = 4$ items total, corresponding to $\mathcal{S}_1 = \{1, 2\}, \mathcal{S}_2 = \{3, 4\}$. The figure depicts the estimates $\widehat{\tau}_i(t)$, along with the corresponding confidence intervals $[\widehat{\tau}_i(t) - 4\alpha_t, \widehat{\tau}_i(t) + 4\alpha_t]$, at different time steps t. At time $t = 5$, the algorithm is not confident about the position of any of the items, and hence it continues to sample further. At time $t = 10$, the confidence interval of item (1) indicates that (1) is either the best or the second best item, therefore the AR algorithm assigns (1) to $\widehat{\mathcal{S}}_1$. Likewise, it assigns item (4) to $\widehat{\mathcal{S}}_2$. At time step $t = 15$, the AR algorithm assigns items (1) and (2) to $\widehat{\mathcal{S}}_1$ and $\widehat{\mathcal{S}}_2$, respectively, and terminates.

ALGORITHM 1 (Active Ranking (AR)). *At time $t = 0$, define and initialize the following quantities:*

- $\mathcal{S} = [n]$ *(set of items not ranked yet);*
- $\widehat{\mathcal{S}}_\ell = \emptyset$ *for all* $\ell \in [L]$ *(estimates of the partition);*
- $\widehat{k}_\ell = k_\ell$ *for all* $\ell \in \{0, \dots, L\}$ *(borders of the sets);*
- $\widehat{\tau}_i(0) = 0$ *for all* $i \in [n]$ *(estimates of the scores).*

At any time $t \geq 1$:

1. *For every $i \in \mathcal{S}$: Compare item i to an item chosen uniformly at random from $[n] \setminus \{i\}$, and set*

$$\widehat{\tau}_i(t) = \begin{cases} \frac{t-1}{t}\widehat{\tau}_i(t-1) + \frac{1}{t} & \text{if } i \text{ wins} \\ \frac{t-1}{t}\widehat{\tau}_i(t-1) & \text{otherwise.} \end{cases} \tag{3.2}$$

2. *Sort the items in set $\mathcal{S}$ by their current estimates of the scores: For any $k \in [|\mathcal{S}|]$, let (k) denote the item with the k-th largest estimate of the score (if the k-th largest estimate is not unique, break ties arbitrarily).*
3. *With parameter α_t as defined in equation (3.1), do the following for every $j \in \mathcal{S}$:*
   If the following pair of conditions (3.3a) and (3.3b) hold simultaneously for some $\ell \in [L]$,

$$\widehat{k}_{\ell-1} = 0 \quad \text{or} \quad \widehat{\tau}_j(t) < \widehat{\tau}_{(\widehat{k}_{\ell-1})}(t) - 4\alpha_t \tag{3.3a}$$
$$(j \text{ likely is one of the lower } n - k_{\ell-1} - 1 \text{ items})$$
$$\widehat{k}_{\ell} = |\mathcal{S}| \quad \text{or} \quad \widehat{\tau}_j(t) > \widehat{\tau}_{(\widehat{k}_{\ell}+1)}(t) + 4\alpha_t \tag{3.3b}$$
$$(j \text{ likely is one of the top } k_\ell \text{ items}),$$

   then add j to $\widehat{\mathcal{S}}_\ell$, remove j from $\mathcal{S}$, and set $\widehat{k}_{\ell'} \leftarrow \widehat{k}_{\ell'} - 1$ for all $\ell' \geq \ell$.
4. *If $\mathcal{S} = \emptyset$, terminate.*

See Figure 1 for an illustration of the progress of this algorithm on a particular instance.

3.2. *Guarantees and optimality of the AR algorithm.* In this section, we establish guarantees on the number of samples for the AR algorithm to succeed. As we show below, the sample complexity is a function of the gaps between the scores, defined as

$$\bar{\Delta}_{\ell,i} := \tau_{\pi(k_{\ell-1})} - \tau_i, \quad \text{and} \quad \underline{\Delta}_{\ell,i} := \tau_i - \tau_{\pi(k_\ell+1)}. \tag{3.4}$$

The dependence on these gaps is controlled via the functions

$$f_0(x) := \frac{1}{x^2}, \quad \text{and} \quad f_{\mathrm{AR}}(x) := \frac{\log(2\log(2/x))}{x^2}. \tag{3.5}$$

In part (a) of the theorem to follow, we prove an upper bound involving f_{AR} on the AR algorithm, and in part (b), we prove a lower bound involving f_0 that applies to *any* uniformly δ-accurate algorithm. As one might intuitively expect, the number of comparisons required is lower when the gaps between the underlying scores are larger. See Figure 2 for an illustration of the gaps for the particular problem of finding a partitioning of the items $\{1, 2, \ldots, 6\}$ into three sets of cardinality two each.

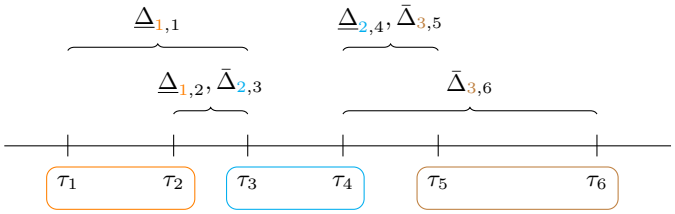


Fig 2: Illustration of the gaps $\bar{\Delta}_{\ell,i}$ and $\underline{\Delta}_{\ell,i}$ relevant for finding a partitioning of the items $\{1, 2, \ldots, 6\}$ into the sets $\mathcal{S}_1 = \{1, 2\}$, $\mathcal{S}_2 = \{3, 4\}$, and $\mathcal{S}_3 = \{5, 6\}$.

THEOREM 1.
There are positive universal constants $(c_{\mathrm{up}}, c_{\mathrm{low}})$ *such that*[2]*:*

(a) For any $M \in \mathcal{C}_0$*, and any* $\delta \in (0, 0.14]$ *the AR algorithm is* δ*-accurate for* M *using a number of comparisons at most*

$$c_{\mathrm{up}} \log\left(\frac{n}{\delta}\right) \Big\{ \sum_{i \in \mathcal{S}_1} f_{AR}(\underline{\Delta}_{1,i}) + \sum_{\ell=2}^{L-1} \sum_{i \in \mathcal{S}_\ell} \max\Big\{ f_{AR}(\underline{\Delta}_{\ell,i}), f_{AR}(\bar{\Delta}_{\ell,i}) \Big\} + \sum_{i \in \mathcal{S}_L} f_{AR}(\bar{\Delta}_{L,i}) \Big\}. \tag{3.6a}$$

(b) For any $\delta \in (0, 0.14]$*, consider a ranking algorithm that is uniformly* δ*-accurate over* $\mathcal{C}_{1/8}$*. Then when applied to a given pairwise comparison model* $M \in \mathcal{C}_{3/8}$*, it must make at least*

$$c_{\mathrm{low}} \log\left(\frac{1}{2\delta}\right) \Big\{ \sum_{i \in \mathcal{S}_1} f_0(\underline{\Delta}_{1,i}) + \sum_{\ell=2}^{L-1} \sum_{i \in \mathcal{S}_\ell} \max\Big\{ f_0(\underline{\Delta}_{\ell,i}), f_0(\bar{\Delta}_{\ell,i}) \Big\} + \sum_{i \in \mathcal{S}_L} f_0(\bar{\Delta}_{L,i}) \Big\} \tag{3.6b}$$

comparisons on average.

Part (a) of Theorem 1 proves that the AR algorithm is δ-accurate, and characterizes the number of comparisons required to find a ranking as a function of the gaps between scores. In contrast, part (b) shows that, up to logarithmic factors, the AR algorithm is optimal, not only in a minimax sense, but in fact when acting on any given problem instance. The proof of

[2]Without having optimized for the constants, our proof yields $c_{\mathrm{up}} = 2003$ and $c_{\mathrm{low}} = 1/16$.

part (b) involves constructing pairs of comparison matrices that are especially hard to distinguish, and makes use of a change of measure lemma [22, Lem. 1] from the bandit literature. For the special case of top-1 identification (corresponding to $L = 2$ and $k_1 = 1$), Jamieson et al. [19] and Urvoy et al. [38] observe that by using the relation to multi-armed bandits discussed in Section 2.3, a standard multi-armed bandit algorithm can be applied which in turn is known to achieve the sample complexity (3.6a). Again for the special case of top-1 identification, part (b) of Theorem 1 recovers Theorem 1 in [19]. For the special case of top-k identification (corresponding to $L = 2$ and $k_1 = k$), part (a) of Theorem 1 reduces to Theorem 1 in [38], which applies to a generic elimination algorithm (SAVAGE) particularized to the dueling bandit problem (see [38, Sec. 4.2]). However, the dependency of Theorem 1 on the gaps is by a logarithmic factor better than that in Theorem 1 in [38]. Note that our negative result in part (b) applies to the stochastic regime, where the pairwise comparison probabilities are bounded away from zero, and does therefore not rule out the possibility that in the regime where the pairwise comparison probabilities are very close to one, improvements in sample complexity are possible.

As Theorem 1 shows, while the AR algorithm is optional, the number of comparisons is at least linear in n, and can be (significantly) larger, depending on the gaps between the scores. In order to gain intuition on this result, in particular the dependence on the squared gaps, it is useful to specialize to the toy case $n = 2$. In this special case with $n = 2$, we have $\tau_1 = M_{12}$ and $\tau_2 = M_{21} = 1 - M_{12}$. Thus, the ranking problem reduces to testing the hypothesis $\{\tau_1 > \tau_2\}$. One can verify that the hypothesis $\{\tau_1 > \tau_2\}$ is equivalent to $\{M_{12} > \frac{1}{2}\}$. Let $X_i, i = 1, \dots, Q$ be the outcomes of Q independent comparisons of items 1 and 2, that is, $\mathbb{P}[X_i = 1] = M_{12}$ and $\mathbb{P}[X_i = 0] = 1 - M_{12}$. A natural test for $\{M_{12} > \frac{1}{2}\}$ is to test whether $\bar{X} > 1/2$, where $\bar{X} \coloneqq \frac{1}{Q}\sum_{i=1}^{Q} X_i$. Supposing without loss of generality that $M_{12} > \frac{1}{2}$, by Hoeffding's inequality, we can upper bound the corresponding error probability as

$$\mathbb{P}\left[\bar{X} \le 1/2\right] = \mathbb{P}\left[\bar{X} - M_{12} \le 1/2 - M_{12}\right] \le e^{-2Q(1/2-M_{12})^2} = e^{-8Q(\tau_1-\tau_2)^2}.$$

Thus, for $Q \ge \frac{\log(1/\delta)}{8(\tau_1-\tau_2)^2}$ the error probability is less than δ. The bound (3.6a) in Theorem 1(a) yields an identical result up to a logarithmic factor.

More generally, testing for the inclusion $i \in \mathcal{S}_\ell$ amounts to testing for $\bar{\Delta}_{\ell,i} > 0$ and $\underline{\Delta}_{\ell,i} > 0$, where $\bar{\Delta}_{\ell,i} = \tau_{\pi(k_{\ell-1})} - \tau_i$ and $\underline{\Delta}_{\ell,i} = \tau_i - \tau_{\pi(k_\ell+1)}$. These requirements provide some intuition regarding the dependence of our bounds on the inverses of the squared gaps.

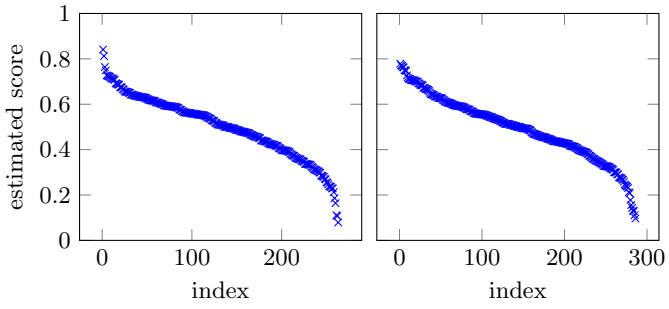
3.3. *Gains due to active estimation.* In order to understand the benefits of an active strategy, it is worthwhile to compare the performance of our active method to the (minimax optimal) guarantees obtainable by passive comparison strategies. We hasten to add that these gains should not be seen as surprising in of themselves, since it is well-known that active estimators can often yield significant improvements over passive schemes.

Recent work by a subset of the current authors [30] considers the problem of ranking items from pairwise comparisons in a passive random design setup. On one hand, it is shown (Theorem 1) that a simple passive scheme—namely, one that ranks items according to the total number of comparisons won—recovers the top k items with high probability using $\frac{n \log n}{(\tau_k - \tau_{k+1})^2}$ comparisons in total (assuming without loss of generality that $\tau_1 > \tau_2 > \ldots > \tau_n$); the same paper also establishes a matching lower bound, meaning that no passive scheme can do better up to constant factors. In contrast, Theorem 1 of the present paper shows that in the active setting, the number of comparisons necessary and sufficient for finding the top k items is of the order

$$\sum_{i=1}^{k} \frac{1}{(\tau_i - \tau_{k+1})^2} + \sum_{i=k+1}^{n} \frac{1}{(\tau_k - \tau_i)^2},$$

up to a logarithmic factor. Readers familiar with the bandit literature, will notice that this expression is equivalent to the complexity of top-k arm identification with τ_i being the expected reward of arm i, see e.g., [21]. By comparing this guarantee to the passive sample complexity $\frac{n \log n}{(\tau_k - \tau_{k+1})^2}$, we can understand when active strategies do or do not lead to substantial gains. First, note that the complexity of the non-active estimator is always higher, except for scores satisfying the linear constraints $\tau_1 = \ldots = \tau_k$ and $\tau_{k+1} = \ldots = \tau_n$, in which case the two estimators would have similar performance. Second, the difference in sample complexity can be as large as a factor of n, up to logarithmic factors. In particular, suppose that the score difference $\tau_i - \tau_{i+1}$ is on the order of $1/n$: in this case, up to logarithmic factors, the sample complexity of the active and passive schemes scale as n^2 and n^3 respectively. A similar conclusion holds if we compare the results of the paper [30] with those of the present paper for the problem of recovering the full ranking.

Having seen that the gains from active estimation depend on the distribution of the scores $\{\tau_i\}_{i=1}^n$, it is natural to wonder how these scores behave in real-world settings. As one illustration, Figure 3 shows some real-world examples of this distribution for data collected by Salganik and Levy [29]; the left panel shows the scores estimated in the paper [29] of a collection of



(a) PlaNYC (b) OECD

Fig 3: Estimated scores from comparisons of the proposals in the PlaNYC (a) and OECD (b) surveys, as reported in the paper [29] (only scores of items (proposals) that were rated at least 50 times are depicted). Estimation of the top k proposals or another ranking with an active scheme would require a significantly smaller number of queries compared to a non-active estimator. For example, for the PlaNYC survey and top-k, $k = 10$, identification the ratio of the sample complexity of the passive estimator over the sample complexity of the active estimator is about $n(\tau_k - \tau_{k+1})^{-2} / \left(\sum_{i=1}^{k}(\tau_i - \tau_{k+1})^{-2} + \sum_{i=k+1}^{n}(\tau_k - \tau_i)^{-2}\right) \approx 30$.

environmental proposals for New York City, whereas the right panel shows a collection of educational proposals for the Organisation for Economic Co-operation and Development (OECD). These data were collected by asking interviewees in corresponding online surveys for preferences between two options. The goal of such online surveys is, for example, to identify the top proposals or a total ranking of the proposals. Our results show that estimation of the top k proposals or another ranking with an active scheme would require a significantly smaller number of queries compared to an non-active estimator.

4. When parametric assumptions don't help. The active ranking algorithm described and analyzed in the previous section applies to *any* comparison matrix M—that is, it neither assumes nor exploits any particular structure in M, such as that imposed by the parametric models described in Section 2.4. Given that the AR algorithm imposes no conditions on the model, one might suspect that when ranking data is actually drawn from a parametric model—for example, of BTL or Thurstone type—it could be possible to come up with another algorithm with a lower sample complexity. Surprisingly, as we show in this section, this intuition turns out to be false in

the following sense: for stochastic comparison models—in which the comparison probabilities are bounded strictly away from zero and one—imposing parametric assumptions can lead to at most a logarithmic reduction in sample complexity.

Recall that a parametric model is described by a continuous and strictly increasing CDF Φ; in this section, we prove a lower bound that applies even to algorithms that are given *a priori* knowledge of the function Φ. For any pair of constants $0 < \phi_{\min} \leq \phi_{\max} < \infty$, we say that a CDF Φ is $(\phi_{\min}, \phi_{\max}, M_{\min})$-bounded if it is differentiable, and its derivative Φ' satisfies the bounds

$$\phi_{\min} \leq \Phi'(t) \leq \phi_{\max}, \quad \text{for all } t \in [\Phi^{-1}(M_{\min}), \Phi^{-1}(1 - M_{\min})]. \tag{4.1}$$

Note that these conditions hold for standard parametric models, such as the BTL and Thurstone models.

The following result applies to any parametric model $\mathcal{C}_{\mathrm{PAR}}(\Phi)$ described by a CDF of this type. It also involves the complexity parameter

$$F(\tau(M)) \coloneqq \sum_{i \in \mathcal{S}_1} f_0(\underline{\Delta}_{1,i}) + \sum_{\ell=2}^{L-1} \sum_{i \in \mathcal{S}_\ell} \max\left\{ f_0(\underline{\Delta}_{\ell,i}), f_0(\bar{\Delta}_{\ell,i}) \right\} + \sum_{i \in \mathcal{S}_L} f_0(\bar{\Delta}_{L,i}),$$

which appeared previously in the lower bound from Theorem 1(b).

THEOREM 2.

(a) Given a tolerance $\delta \in (0, 0.15]$, and a continuous and strictly increasing CDF Φ whose derivative is $(\phi_{\min}, \phi_{\max}, M_{\min})$-bounded, consider any algorithm that is uniformly δ-accurate over $\mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$. Then, when applied to a given pairwise comparison matrix $M \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$, it must make at least

$$c_{\mathrm{par}} \log\left(\frac{1}{2\delta}\right) F(\tau(M)) \tag{4.2}$$

comparisons on average, where $c_{\mathrm{par}} \coloneqq \frac{M_{\min}\phi_{\min}^2}{2.004\phi_{\max}^2}$.

(b) Let $\tau \in (0,1)^n$ be any set of scores that is realizable by some pairwise comparison matrix $M' \in \mathcal{C}_{M_{\min}}$, $M_{\min} > 0$. Then for any continuous and strictly increasing Φ, there exists a pairwise comparison matrix in $M \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$ with scores τ, and in particular with $F(\tau(M)) = F(\tau(M'))$.

First, let us provide some concrete settings of the constant c_{par}: for $M_{\min} = \frac{3}{8}$, we have $c_{\mathrm{par}} = 0.164$ and $c_{\mathrm{par}} = 0.169$ for the BTL and Thurstone models, respectively; whereas for $M_{\min} = \frac{1}{4}$, we have $c_{\mathrm{par}} = 0.07$ and $c_{\mathrm{par}} = 0.079$ for the BTL and Thurstone models, respectively.

Second, let us turn to the implications of Theorem 2. To start, it should be noted that the lower bound (4.2) is, at least in a certain sense, stronger than the lower bound from Theorem 1, because it applies to a broader class of algorithms—namely, those that are δ-accurate *only* over the smaller class of parametric models. On the flip side, it is possible that the lower bound (4.2) might be weaker in some sense. That is, could there be some "difficult" matrix $M' \in \mathcal{C}_{M_{\min}}$ such that the supremum of $F(\tau(M))$ over $M \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$ is much smaller than $F(\tau(M'))$? Part (b) of the theorem rules out this possibility: it guarantees that for any pairwise comparison matrix M'—which need not be generated by a parametric model—there exists a parametric model M for which the ranking problem is equally hard. This result is surprising because one might think that imposing parametric assumptions would simplify the ranking problem. In fact, the full set $\mathcal{C}_{M_{\min}}$ is substantially larger than the parametric subclass $\mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$; in particular, one can demonstrate matrices in $\mathcal{C}_{M_{\min}}$ that cannot be well-approximated by any parametric model; for example, see the paper [33] for inapproximability results of this type.

A consequence of Theorem 2 is that up to logarithmic factors, the AR algorithm is again optimal, even if we restrict ourselves to algorithms that are uniformly δ-accurate *only* over a parametric subclass (provided the comparison probabilities are bounded away from zero and one). Thus, for stochastic comparison models, imposing parametric assumptions only limits the flexibility while failing to provide any significant reductions in sample complexity for ranking. It is worth commenting that for deterministic or near-deterministic comparison models—in which the pairwise probabilities can be arbitrarily close to zero or one— the constant c_{par} in the lower bound (4.2) can become small. For this reason, our lower bound does not contradict the fact that parametric assumptions might help for (near)-deterministic comparison models. As one example, recalling that the BTL model described in Section 2.4 is based on a parameter vector $w \in \mathbb{R}^n$, suppose that we set $w_i = \xi(n-i)$ for all $i \in [n]$, and then let ξ tend to infinity. Since $M_{ij} = \frac{e^{w_i}}{e^{w_i}+e^{w_j}}$ under the model, taking the limit $\xi \to \infty$ leads to a fully deterministic comparison model in which items i beats j with probability one if and only if $w_i > w_j$. In this limit, pairwise ranking reduces to a deterministic sorting problem, and sorting-based algorithms (e.g., [35]) can be used to achieve top item identification with $O(n \log n)$ comparisons. In con-

trast, in this deterministic setting, the AR algorithm requires $O(n^2 \log n)$ comparisons, which can be guaranteed by applying Theorem 1(a) with the associated score vector $\tau_i = 1 - \frac{i-1}{n-1}$. To be very clear, this example does *not* violate any of our claimed results since the lower bound of Theorem 1(b), and hence the associated claim of optimality, applies only to the case when the pairwise comparison probabilities are bounded away from 0 and 1 by some constant $M_{\min}$.

5. Numerical results. We now turn to some numerical comparisons of our active ranking (AR) algorithm with algorithms designed for parametric models. One finding—consistent with our theory—is that the AR algorithm is on par or outperforms these algorithms, unless the pairwise comparison probabilities are close to zero or one. Moreover, we find that algorithms designed for parametric models start to break down even if the parametric modeling assumption is only slightly violated. Finally, in the supplement, we experiment with the choice of constants setting confidence intervals α_t for the AR algorithm, and find that the choice given by our theory is conservative.

5.1. *Comparison to algorithms tailored to parametric models.* Our results in Section 4 show that for stochastic comparison models, algorithms that exploit parametric structure can have sample complexities lower by at most a logarithmic factor. On the other hand, for (near)-deterministic comparison models, we gave an example showing that parametric structure can allow for significant gains. In this section, we perform some numerical experiments to quantify and understand these two different regimes.

To this end, we consider the problem of top-item recovery, that is, the problem specified by $L = 2$ and $k_1 = 1$. We study this top-item recovery problem (also known as the dueling bandit problem) because of availability of previous algorithms for this special case of the more general ranking problem considered in our paper. We compare the AR algorithm to the Plackett-Luce PAC (PLPAC) [35] and Beat the Mean Bandit (BTMB) [39] algorithms. Both algorithms yield an δ-accurate ranking provided the BTL modeling assumptions hold. We choose the PLPAC algorithm for comparison as it is based on sorting: a BTL problem with pairwise comparison probabilities close to one and zero is in essence a noisy sorting problem, thus we expect sorting based procedures to work well here. The BTMB algorithm is guaranteed to succeed if Strong Stochastic Transitivity (SST) (or a relaxed version thereof) and a certain stochastic transitivity triangle inequality hold[3]; both assumptions are satisfied for the BTL model.

[3]A necessary and sufficient condition for a matrix to satisfy the SST condition is the

Regarding the algorithms parameters; for the AR algorithm we set $\alpha_t = \sqrt{\frac{\log(n/\delta)+0.75\log\log(n/\delta)+1.5\log(1+\log(t/2))}{2t}}$, as for this choice the AR algorithm provably succeeds according to our main result (in practice, the constants in α_t may be chosen smaller, see Section 8.3 for a discussion for the choice of α_t, which leads to even better performance). We set $n = 10$ and consider two different BTL models parameterized by $\eta > 0$ and $\xi > 0$, respectively, and denoted by $M^{(\eta)}$ and $M^{(\xi)}$. The parameters η and ξ determine how close the minimal and maximal pairwise comparison probabilities are to 0 and 1; the larger, the closer. Specifically, the parameters of the BTL model $M^{(\eta)}$ are given by $w_i = \log(1/\eta + n - i), i = 1, \ldots, n$. This results in pairwise comparison probabilities $M^{(\eta)}_{ij} = \frac{1/\eta+n-i}{2(1/\eta+n)-i-j}$. The parameters of the second BTL model, $M^{(\xi)}$, are $w_i = \xi(n-i)$ which implies that the probability that item i beats the next best item $i+1$ is $M^{(\xi)}_{i,i+1} = \frac{1}{1+e^{-\xi}}$. Thus, each item beats all lower ranked ones with probability at least $\frac{1}{1+e^{-\xi}}$, which results in *all* the pairwise comparison probabilities being skewed away from 1/2; the larger ξ the "closer" those probabilities are to 0 and 1.

In Figure 4 we depict the empirical sample complexity for both models as a function of $M_{\max} := \max_{i,j} M_{ij}$, along with the corresponding complexity parameters $F(\tau(M^{(\eta)}))$ and $F(\tau(M^{(\xi)}))$. Here, we choose the model parameters η and ξ such that $M_{\max}$ varies between 0.65 and 0.99. The results show, as predicted by our theory, that the sample complexity of the AR algorithm is essentially a constant times the complexity parameter F. In contrast, the sample complexity of the PLPAC and the BTMB algorithms improves in $M_{\max}$ relative to the complexity parameter F. Note that the AR algorithm performs better than PLPAC and BTMB if $M_{\max}$ is not too large, while both PLPAC and BTMB have lower sample complexity than the AR algorithm in the regime where $M_{\max}$ is very close to one. We remark that the relative improvement is not determined solely by $M_{\max}$, as shown by the curves for the two differently parameterized BTL models differing.

Our next simulation shows that, however, even if the pairwise comparison matrix only deviates slightly from the BTL model, both the sample complexity and more pertinently the failure probability (that is, $\mathbb{P}_M\left[\widehat{\mathcal{S}}_\ell \neq \mathcal{S}_\ell, \text{ for one or more } \ell = 1, \ldots, L\right]$) can become very large. Specifically, as before, we generate a BTL model M with $n = 10$ and parameters $w_i = \log(1+n-i), i = 1, \ldots, n$. We then substitute a fraction of λ of the off-

existence of a permutation of the items, such that the permuted pairwise comparison matrix M is non-decreasing across rows and non-increasing across columns. The stochastic transitivity inequality demands that for each triplet with $\tau_1 > \tau_j > \tau_k$, we have that $M_{1j} - 1/2 + M_{jk} - 1/2 \geq M_{1k} - 1/2$.

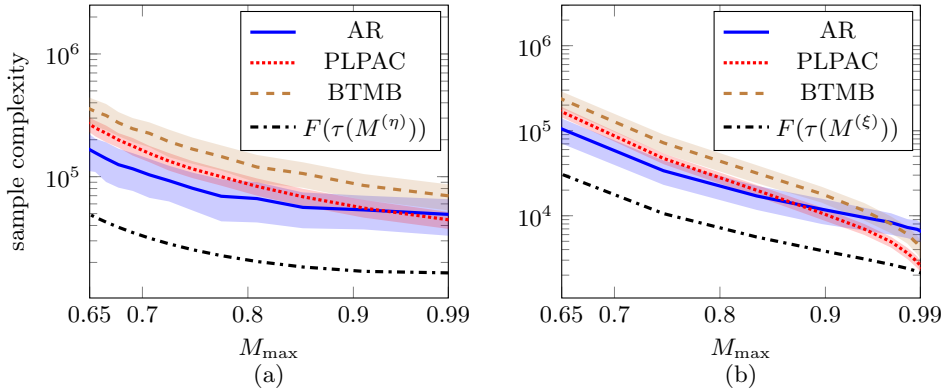


Fig 4: (a) Empirical sample complexity of the AR, PLPAC, and BTMB algorithms applied to the BTL model $M^{(\eta)}$ with parameters $w_i = \log(\eta + n-i), i = 1, \ldots, 10$, and (b) applied to the BTL model $M^{(\xi)}$ with parameters $w_i = \xi(n-i), i = 1, \ldots, 10$, as a function of $M_{\max} := \max_{i,j} M_{ij}$. For panel (a) and (b) we varied η and ξ such that $M_{\max} \in [0.65, 0.99]$. The error bars correspond to one standard deviation from the mean. While the AR algorithm has even lower sample complexity than the PLPAC and BTMB algorithms in the regime where $M_{\max}$ is not to close to 1; the PLPAC and BTMB perform better when $M_{\max}$ is close to one.

diagonal elements of M with a number drawn uniformly from $[0, 1]$. Thus, the model M transitions from a BTL model to a random pairwise comparison matrix in λ; for small λ, the model M is close to the original BTL model. The results, depicted in Figure 5, show that, while the AR algorithm succeeds for all values of λ as expected, the sample complexity and more importantly the failure probability of the PLPAC and BTMB algorithms become very large. We hasten to add that both the PLPAC and BTMB algorithm are not designed for this scenario; therefore it might not be surprising that they fail. The results show that these algorithms are, however, not robust to violations of their assumed models.

6. Proofs. In this section, we provide the proofs of our two main theorems. In order to simplify notation, we take the underlying permutation π equal to the identity, so that $\tau_1 > \tau_2 > \ldots > \tau_n$. This assumption entails no loss of generality, since it can also be satisfied by re-indexing the items if necessary.

6.1. *Proof of Theorem 1(a).* In this section, we provide a proof of the achievable result stated in part (a) of Theorem 1. Our proof consists of three

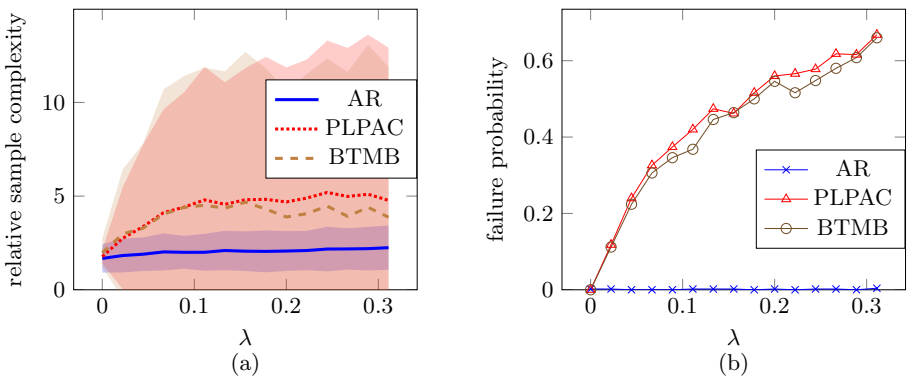


Fig 5: (a) Relative sample complexity defined as the number of comparisons until termination, Q, divided by the complexity parameter $F(\tau(M))$, and (b) failure probability on a BTL model M with $n = 10$ and with a fraction of λ of the off-diagonals of M substituted by a random pairwise comparison probability. The model transitions from a BTL model to a random pairwise comparison matrix in λ; the closer λ to zero the closer M to the original BTL model. The results show that, while the AR algorithm yields an δ-accurate ranking after $O(F(\tau(M)))$ comparisons, irrespectively of λ, the sample complexity and more importantly the failure probability of the PLPAC and BTMB algorithms become very large in λ.

main steps. We begin by showing that the estimate $\widehat{\tau}_i(t)$ is guaranteed to be α_t-close to τ_i, for all $i \in \mathcal{S}$, with high probability. We then use this result to show that the AR algorithm never misclassifies any item, and that it stops with the number of comparisons satisfying the claimed upper bound.

Throughout the paper, we use $\mathcal{S}$ to denote the set of items that have not been ranked yet; to be clear, since items are eliminated from $\mathcal{S}$ at certain time steps t, the set $\mathcal{S}$ changes with t, but we suppress this dependence for notational simplicity.

The following lemma ensures that the estimated score $\widehat{\tau}_i$ is close to the latent score τ_i. As shown below, the lemma follows by noting that $\widehat{\tau}_i(t)$ is a sum of t independent Bernoulli random variables, each of which has mean τ_i/t, and application of a version of the law of the iterated logarithm.

LEMMA 1. *Under the theorem's assumptions, the event*

$$\mathcal{E}_\alpha := \{|\widehat{\tau}_i(t) - \tau_i| \le \alpha_t, \quad \textit{for all } i \in \mathcal{S} \textit{ and for all } t \ge 1\} \tag{6.1}$$

occurs with probability at least $1 - \delta$.

Our next step is to show that provided that the event $\mathcal{E}_\alpha$ occurs, the AR algorithm never misclassifies any item, that is, $\widehat{\mathcal{S}}_\ell \subseteq \mathcal{S}_\ell$ for all ℓ and for all $t \geq 1$. First suppose that, at a given time step t, the AR algorithm did not misclassify any item at a previous time step. We show that, at time t, conditioned on the event $\mathcal{E}_\alpha$, any item $j \in \mathcal{S}$ is added to $\widehat{\mathcal{S}}_\ell$ only if $j \in \mathcal{S}_\ell$, which implies that the AR algorithm does not misclassify any item at time t. This fact is a consequence of our second auxiliary result.

In order to state this second lemma, we require some additional notation. Let $\tau_{\{k\}}$ denote the k-th largest score among the *latent* scores $\tau_i,\ i \in \mathcal{S}$. Note that we use the notation $\{\cdot\}$ to emphasize that the index $\{k\}$ is not necessarily equal to the index (k), since the latter corresponds to the k-th largest score amongst the *estimated* scores $\widehat{\tau}_i(t),\ i \in \mathcal{S}$.

LEMMA 2. *Suppose that the event $\mathcal{E}_\alpha$ occurs. Then both of the implications*

(6.2a)
- *for any $j \in \mathcal{S}$, $\quad \widehat{\tau}_j(t) < \widehat{\tau}_{(\widehat{k}_{\ell-1})}(t) - 4\alpha_t \quad$ implies $\quad \tau_j < \tau_{\{\widehat{k}_{\ell-1}\}}$, and*

(6.2b)
- *for any $j \in \mathcal{S}$, $\quad \widehat{\tau}_j(t) > \widehat{\tau}_{(\widehat{k}_\ell+1)}(t) + 4\alpha_t \quad$ implies $\quad \tau_j > \tau_{\{\widehat{k}_\ell+1\}}$,*

hold for all $t \geq 1$.

Provided that the AR algorithm did not misclassify any item at a previous time step, some consequences of implications (6.2a) and (6.2b) are the following:

- first, for any index ℓ, an item is added to $\widehat{\mathcal{S}}_\ell$ at time t only if $j \in \mathcal{S}_\ell$.
- therefore, we are guaranteed that $\widehat{\mathcal{S}}_\ell \subseteq \mathcal{S}_\ell$ at time $t+1$.

These consequences allow us to apply an inductive argument to conclude that the AR algorithm never misclassifies any item.

Our next step is to show that, conditioned on the event $\mathcal{E}_\alpha$ on which the AR algorithm does not misclassify any item, all items are eliminated after the number of comparisons given in equation (3.6a) have been carried out. Since, by Lemma 1, the event $\mathcal{E}_\alpha$ holds with probability at least $1-\delta$, this concludes the proof of Theorem 1(a).

In order to establish the former claim, we use the following lemma, in which we made the dependence of the set of candidates $\mathcal{S}$ on t explicit by writing $\mathcal{S}(t)$.

LEMMA 3. *Suppose that the event $\mathcal{E}_\alpha$ occurs. For any index $\ell \in \{2, \dots, L\}$ and any item $i \in \mathcal{S}_\ell \cap \mathcal{S}(\bar{t}_i)$, we have, with $c_1 := 654$,*

$$\text{(6.3a)} \quad \widehat{\tau}_i(\bar{t}_i) < \widehat{\tau}_{(\widehat{k}_{\ell-1})}(\bar{t}_i) - 4\alpha_{\bar{t}_i}, \quad \text{where } \bar{t}_i := \frac{c_1}{\bar{\Delta}_{\ell,i}^2} \log\left(\frac{n}{\delta} \log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right),$$

$\bar{\Delta}_{\ell,i} = \tau_{k_{\ell-1}} - \tau_i$, and for $\ell \in \{1, \dots, L-1\}$ and any item $i \in \mathcal{S}_\ell \cap \mathcal{S}(\underline{t}_i)$, we have

$$\text{(6.3b)} \quad \widehat{\tau}_i(\underline{t}_i) > \widehat{\tau}_{(\widehat{k}_\ell+1)}(\underline{t}_i) - 4\alpha_{\underline{t}_i}, \quad \text{where } \underline{t}_i := \frac{c_1}{\underline{\Delta}_{\ell,i}^2} \log\left(\frac{n}{\delta} \log\left(\frac{2}{\underline{\Delta}_{\ell,i}}\right)\right),$$

$\underline{\Delta}_{\ell,i} = \tau_i - \tau_{k_\ell+1}$.

Consequently, the index $i \in \mathcal{S}_\ell$ is eliminated from the set of candidates $\mathcal{S}$ after no more than the following number of many time steps (and hence comparisons):

$$\begin{cases} \underline{t}_i, & \text{if } \ell = 1 \\ \max(\underline{t}_i, \bar{t}_i), & \text{if } \ell \in \{2, \dots, L-1\} \\ \bar{t}_i, & \text{if } \ell = L \end{cases}.$$

Using the relations

$$\bar{t}_i \le c_{\text{up}} \frac{\log(2\log(2/\bar{\Delta}_{\ell,i}))}{\bar{\Delta}_{\ell,i}^2} \log(n/\delta), \text{ and } \underline{t}_i \le c_{\text{up}} \frac{\log(2\log(2/\underline{\Delta}_{\ell,i}))}{\underline{\Delta}_{\ell,i}^2} \log(n/\delta),$$

where the inequalities hold for some constant c_{up} (in particular, we can set $c_{\text{up}} = 2003$), it follows that the AR algorithm terminates after the number of comparisons stated in equation (3.6a) has been carried out.

It remains to prove Lemmas 1, 2, and 3, and we do so in the supplement.

6.2. *Proof of Theorem 1(b).* We now turn to the proof of the lower bound from Theorem 1. We first introduce some notation required to state a useful lemma [22, Lem. 1] from the bandit literature. Let $\nu = \{\nu_j\}_{j=1}^m$ be a collection of m probability distributions, each supported on the real line $\mathbb{R}$. Consider an algorithm $\mathcal{A}$, that, at times $t = 1, 2, \dots$, selects the index $i_t \in [m]$ and receives an independent draw X_t from the distribution ν_{i_t} in response. Algorithm $\mathcal{A}$ may select i_t only based on past observations, that is, i_t is $\mathcal{F}_{t-1}$ measurable, where $\mathcal{F}_t$ is the σ-algebra generated by $i_1, X_{i_1}, \dots, i_t, X_{i_t}$. Algorithm $\mathcal{A}$ has a stopping rule χ that determines the termination of $\mathcal{A}$. We assume that χ is a stopping time measurable with respect to $\mathcal{F}_t$ and obeying $\mathbb{P}[\chi < \infty] = 1$.

Let $Q_i(\chi)$ denote the total number of times index i has been selected by the algorithm $\mathcal{A}$ (until termination). For any pair of distributions ν and ν', we let $\mathrm{KL}(\nu, \nu')$ denote their Kullback-Leibler divergence, and for any $p, q \in [0,1]$, let $d(p,q) := p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$ denote the Kullback-Leiber divergence between two binary random variables with success probabilities p, q.

With this notation, the following lemma relates the cumulative number of comparisons to the uncertainty between the actual distribution ν and an alternative distribution ν'.

LEMMA 4 ([22, Lem. 1]). *Let ν, ν' be two collections of m probability distributions on $\mathbb{R}$. Then for any event $\mathcal{E} \in \mathcal{F}_\chi$ with $\mathbb{P}_\nu[\mathcal{E}] \in (0,1)$, we have*

$$\sum_{i=1}^{m} \mathbb{E}_\nu [Q_i(\chi)] \, \mathrm{KL}(\nu_i, \nu_i') \geq d(\mathbb{P}_\nu[\mathcal{E}], \mathbb{P}_{\nu'}[\mathcal{E}]). \tag{6.4}$$

Let us now use Lemma 4 to prove Theorem 1(b). In particular, we apply it using the event

$$\mathcal{E} := \left\{ \widehat{\mathcal{S}}_\ell = \mathcal{S}_\ell, \quad \text{for all } \ell = 1, \ldots, L \right\}, \tag{6.5}$$

which corresponds to success of the algorithm $\mathcal{A}$. Note that here $\{\widehat{\mathcal{S}}_\ell\}_{\ell=1}^L$ are the estimated sets at termination of the algorithm. Recalling that χ is the stopping rule of algorithm $\mathcal{A}$, we are guaranteed that $\mathcal{E} \in \mathcal{F}_\chi$. Given the linear relations $M_{ij} = 1 - M_{ji}$, the pairwise comparison matrix M is determined by the entries $\{M_{ij}, i = 1, \ldots, n, \ j = i+1, \ldots, n\}$. Let $Q_{ij}(\chi)$ be the total number of comparisons between items i and j made by $\mathcal{A}$. For any other pairwise comparison matrix $M' \in \mathcal{C}_0$, Lemma 4 ensures that

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{E}_M [Q_{ij}] \, d(M_{ij}, M_{ij}') \geq d(\mathbb{P}_M[\mathcal{E}], \mathbb{P}_{M'}[\mathcal{E}]). \tag{6.6}$$

In order to aid in subsequent exposition, we augment the notation $\mathcal{S}_\ell$ defined in (2.1) earlier to explicitly depict the underlying pairwise-probability matrix which $\mathcal{S}_\ell$ depends on. Specifically, for any matrix $M \in [0,1]^{n \times n}$ such that $M_{ij} + M_{ji} = 1$ for every entry (i,j), we let $\mathcal{S}_\ell(M)$ denote the value of $\mathcal{S}_\ell$ when the probabilities of the outcomes of the pairwise-comparisons are governed by the matrix M (we will drop this additional dependence on the matrix whenever it is clear from the context).

For some $\ell > 1$ and item $m \in \mathcal{S}_\ell(M)$, our next step is to construct a matrix $M' \in \mathcal{C}_{1/8}$ such that $m \notin \mathcal{S}_\ell(M')$ under the distribution M'. Since the algorithm $\mathcal{A}$ is uniformly δ-accurate over $\mathcal{C}_{1/8}$ by assumption, we are guaranteed that

$$\mathbb{P}_M[\mathcal{E}] \geq 1 - \delta \quad \text{and} \quad \mathbb{P}_{M'}[\mathcal{E}] \leq \delta,$$

from which it follows that

$$d(\mathbb{P}_M[\mathcal{E}], \mathbb{P}_{M'}[\mathcal{E}]) \geq d(\delta, 1-\delta) = (1-2\delta)\log\frac{1-\delta}{\delta} \geq \log\frac{1}{2\delta}, \tag{6.7}$$

where the last inequality holds for $\delta \leq 0.15$.

It remains to specify the alternative matrix $M' \in \mathcal{C}_0$ for use in inequality (6.7): it is defined with entries

$$M'_{ij} := \begin{cases} M_{mj} + (\tau_{k_{\ell-1}} - \tau_m), & \text{if } i = m, j \in [n] \setminus \{m\} \\ M_{im} - (\tau_{k_{\ell-1}} - \tau_m), & \text{if } j = m, i \in [n] \setminus \{m\} \\ M_{ij} & \text{otherwise.} \end{cases} \tag{6.8}$$

From this definition, it follows that

$$\tau'_m = \frac{1}{n-1}\sum_{j \in [n]\setminus\{m\}} M'_{mj} = \frac{1}{n-1}\sum_{j \in [n]\setminus\{m\}} \big(M_{mj} + (\tau_{k_{\ell-1}} - \tau_m)\big) \;=\; \tau_{k_{\ell-1}}.$$

Similarly, all other scores $\tau'_i, i \in [n] \setminus \{m\}$, are smaller than τ_i by a common constant, that is, for $i \in [n] \setminus \{m\}$ $\tau'_i = \tau_i - \frac{1}{n-1}(\tau_{k_{\ell-1}} - \tau_m)$. See Figure 8 in the supplement for an illustration. It follows that, under the distribution M', the score of item m is among the $k_{\ell-1}$ highest scoring items, which ensures $m \notin \mathcal{S}_\ell(M')$. Moreover, we claim that $M' \in \mathcal{C}_{1/8}$. This inclusion follows from the assumption $M \in \mathcal{C}_{3/8}$, which implies that $M'_{mj} \leq \frac{5}{8} + \left(\frac{5}{8} - \frac{3}{8}\right) \leq \frac{7}{8}$. An analogous argument shows that $M'_{mj} \geq \frac{1}{8}$.

Next, consider the total number of comparisons item m is involved in that is, $Q_m = \sum_{j \in [n]\setminus\{m\}} Q_{mj}$. Recall that Q_{ij} is the total number of comparisons between items i and j made by the algorithm $\mathcal{A}$. By the linearity of expectation, we have the following bound on the expectation of Q_m:

$$\begin{aligned} \max_{j \in [n]\setminus\{m\}} d(M_{mj}, M'_{mj}) \mathbb{E}_M[Q_m] &= \max_{j \in [n]\setminus\{m\}} d(M_{mj}, M'_{mj}) \sum_{j' \in [n]\setminus\{m\}} \mathbb{E}_M\left[Q_{mj'}\right] \\ &\geq \sum_{j \in [n]\setminus\{m\}} \mathbb{E}_M[Q_{mj}]\, d(M_{mj}, M'_{mj}). \end{aligned}$$

Now observe that by the definition of M' in equation (6.8), we have $d(M_{ij}, M'_{ij}) = 0$ for all (i,j) outside of the sets $\{(m,j) \mid j \in [n] \setminus \{m\}\}$ and $\{(i,m) \mid i \in [n] \setminus \{m\}\}$. Removing these terms from the sum yields

$$\begin{aligned} \max_{j \in [n]\setminus\{m\}} d(M_{mj}, M'_{mj}) \mathbb{E}_M [Q_m] &\geq \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{E}_M [Q_{ij}]\, d(M_{ij}, M'_{ij}) \\ &\overset{\text{(i)}}{\geq} d(\mathbb{P}_M [\mathcal{E}], \mathbb{P}_{M'} [\mathcal{E}]) \overset{\text{(ii)}}{\geq} \log \frac{1}{2\delta}, \end{aligned} \tag{6.9}$$

where step (i) follows from inequality (6.6) in Lemma 4; and step (ii) follows from inequality (6.7).

We next upper bound the KL divergence on the left hand side of inequality (6.9). Using the inequality $\log x \leq x - 1$ valid for $x > 0$, we have

$$d(M_{mj}, M'_{mj}) \leq \frac{(M_{mj} - M'_{mj})^2}{M'_{mj}(1 - M'_{mj})} \leq 10(\tau_{k_{\ell-1}} - \tau_m)^2, \tag{6.10}$$

where the last step uses the definition of M' in equation (6.8), as well as the inclusion $\frac{1}{8} \leq M'_{mj} \leq \frac{7}{8}$, which implies that $\frac{1}{M'_{mj}(1-M'_{mj})} \leq 64/7 < 10$.

Applying inequality (6.10) to the left hand side of inequality (6.9) yields

$$\mathbb{E}_M [Q_m] \geq \frac{\log(1/(2\delta))}{10(\tau_{k_{\ell-1}} - \tau_m)^2}, \qquad \text{valid for each } m \in \mathcal{S}_\ell(M) \text{ and } \ell > 1. \tag{6.11}$$

Now consider an index $m \in \mathcal{S}_\ell(M)$ for some $\ell < L$. In this case, again construct an alternative pairwise comparison matrix M' under which $m \notin \mathcal{S}_\ell(M')$. Specifically, for notational convenience, we set

$$M'_{ij} = \begin{cases} M_{mj} - (\tau_m - \tau_{k_\ell+1}), & i = m, j \in [n] \setminus \{m\} \\ M_{im} + (\tau_m - \tau_{k_\ell+1}), & j = m, i \in [n] \setminus \{m\} \\ M_{ij} & \text{otherwise.} \end{cases}$$

In a similar manner to our earlier argument, we have $\tau'_i = \tau_i + \frac{1}{n-1}(\tau_m - \tau_{k_\ell+1})$ for $i \in [n] \setminus \{m\}$ and $\tau'_m = \tau_{k_\ell+1}$ (relative to the scores τ_i, the score of m is smaller and all others are larger by the same factor). Under M', item m is not amongst the k_ℓ items with the largest scores, and therefore $m \notin \mathcal{S}_\ell(M')$. Carrying out the same computations as above yields:

$$\mathbb{E}_M [Q_m] \geq \frac{\log(1/(2\delta))}{10(\tau_m - \tau_{k_\ell+1})^2}. \tag{6.12}$$

Combining inequalities (6.11) and (6.12) across all items m yields the bound

$$\mathbb{E}_M[Q] = \sum_{i=1}^{n} \mathbb{E}[Q_i]$$
$$\geq c_{\text{low}} \log(1/(2\delta)) \left[\sum_{i \in \mathcal{S}_1} \underline{\Delta}_{1,i}^{-2} + \sum_{\ell=2}^{L-1} \sum_{i \in \mathcal{S}_\ell} \max\left\{ \underline{\Delta}_{\ell,i}^{-2}, \bar{\Delta}_{\ell,i}^{-2} \right\} + \sum_{i \in \mathcal{S}_L} \bar{\Delta}_{L,i}^{-2} \right],$$

with $c_{\text{low}} = 1/10$, thereby yielding the claimed result.

6.3. *Proof of Theorem 2(b).* Let $\tau \in (0,1)^n$ be any set of scores that is realizable by some pairwise comparison matrix in $\mathcal{C}_{M_{\min}}$ that is not necessarily in $\mathcal{C}_{\text{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$. Theorem 2(b) is proven by showing that for any continuous and strictly increasing Φ, there exists a pairwise comparison matrix in $\mathcal{C}_{\text{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$ with scores τ. As mentioned before, the proof of Theorem 2(b) relies on results established by Joe [20] on majorization orderings of pairwise probability matrices. For convenience, we define the set of pairwise probability matrices with scores $\tau = (\tau_1, \ldots, \tau_n)$ as

$$\mathcal{C}(\tau) = \left\{ M \in \mathcal{C}_0 \mid \frac{1}{n-1} \sum_{j \neq i} M_{ij} = \tau_i, \text{ for all } i \right\}.$$

Minimality for pairwise comparison matrices. Our proof requires some background on majorization and a certain notion of minimality for pairwise comparison matrices. We say that a vector $y \in \mathbb{R}^m$ is non-increasing if its entries satisfy $y_1 \geq y_2 \geq \ldots \geq y_m$. Given two non-increasing vectors $y, z \in \mathbb{R}^m$ such that $\sum_{i=1}^m y_i = \sum_{i=1}^m z_i$, we say y majorizes z, written $y \succ z$, if $\sum_{i=1}^k y_i \geq \sum_{i=1}^k z_i$, for all $k = 1, \ldots, m-1$.

Given pairwise comparison matrices $M, M' \in \mathcal{C}(\tau)$, we let $v(M), v(M') \in (0,1)^{n(n-1)}$ be vectors with entries corresponding to the off-diagonal elements of M and M', respectively, in non-increasing order. We say that M *majorizes* M' if $v(M) \succ v(M')$, and we use the shorthand $M \succ M'$ to denote this relation. Finally, a matrix $M \in \mathcal{C}(\tau)$ is *minimal* if any other $M' \in \mathcal{C}(\tau)$ obeying $M \succ M'$ satisfies the relation $v(M') = v(M)$.

In order to prove Theorem 2(b), we show that there is a minimal $M \in \mathcal{C}(\tau) \cap \mathcal{C}_{M_{\min}}$, and this minimal M takes a parametric form and thus $M \in \mathcal{C}_{\text{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$. We first note that Joe [20, Thm. 2.7] observed that the argument minimizing any *Schur convex*[4] function over the set $\mathcal{C}(\tau)$ is a minimal M. Let us now construct a function that is Schur convex. In particular,

[4]In our context, a function $f\colon (0,1)^{n \times n} \to \mathbb{R}$ is Schur convex (or order-preserving) if for all $M, M' \in \mathcal{C}(\tau)$ such that M is majorized by M', we have $f(M) \leq f(M')$.

we first define a scalar function $\psi : [0,1] \to [0,\infty]$ as

$$\psi(u) = \begin{cases} \frac{1}{2}\int_{1/2}^{u} \Phi^{-1}(x)dx, & u \in \left[\frac{1}{2}, 1\right], \\ -\frac{1}{2}\int_{u}^{1/2} \Phi^{-1}(x)dx, & u \in \left[0, \frac{1}{2}\right). \end{cases} \tag{6.13}$$

The function ψ is well defined since the inverse Φ^{-1} exists due to our assumption that Φ is strictly increasing and continuous. Since Φ is strictly increasing, so is Φ^{-1}. It follows that ψ is strictly convex. From the property that all symmetric and strictly convex functions are also strictly Schur convex, it follows that the function $\sum_{i,j\neq i} \psi(M_{ij})$ is strictly Schur convex over $\mathcal{C}(\tau)$. As a result, we are guaranteed that the argument minimizing the following convex program corresponds to a minimal matrix:

$$\begin{aligned} \text{minimize} \quad & \sum_{i,j>i} (\psi(M_{ij}) + \psi(1 - M_{ij})) \\ \text{subject to } & 0 \leq M_{ij} \leq 1, \quad \text{for all } i \in [n], \quad j = i+1, \ldots, n, \text{ and} \\ & \frac{1}{n-1}\sum_{j=1}^{i-1}(1 - M_{ji}) + \frac{1}{n-1}\sum_{j=i+1}^{n} M_{ij} = \tau_i, \quad \text{for all } i \in [n] \end{aligned} \tag{6.14}$$

Here the minimization is performed over the variables M_{ij} for $i = 1, \ldots, n$ and $j = i+1, \ldots, n$.

First note that any M that is feasible for the problem (6.14) obeys $M \in \mathcal{C}(\tau)$. We next show that any optimal solution M^* to the problem (6.14) has entries satisfying the interval inclusion $M^*_{ij} \in [M_{\min}, 1 - M_{\min}]$ for all pairs (i,j), and therefore $M^* \in \mathcal{C}_{M_{\min}}$, as desired. Indeed, suppose that there were an optimal solution M^* that violated this inclusion. By assumption, there exists a matrix $M' \in \mathcal{C}(\tau) \cap \mathcal{C}_{M_{\min}}$. Thus, if the inclusion were violated, then there would be some index pair (i,j) such that $M^*_{ij} > 1 - M_{\min}$. This would imply that M^* is strictly larger than M' in the majorization ordering. But since the objective function (6.14) is Schur convex, this contradicts the optimality of M^*.

We have established that $M^* \in \mathcal{C}(\tau) \cap \mathcal{C}_{M_{\min}}$. We next show that M^* takes a parametric form, which establishes $M^* \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$, which concludes the proof. Since there exists a solution to the convex optimization problem (6.14) that satisfies the inequality constraints strictly (due to $M_{\min} > 0$, by assumption), Slater's conditions hold, and the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for optimality (see, for instance, [4, Sec. 5.5]). Thus, the primal and dual optimal solutions M^*_{ij}

and $\{\lambda^*_{ij}, \kappa^*_{ij}, \nu^*_i\}$ must satisfy the KKT conditions

$$\lambda^*_{ij}, \kappa^*_{ij} \geq 0, \tag{6.15a}$$

$$\lambda^*_{ij}(M^*_{ij} - 1) = 0, \quad \kappa^*_{ij} M^*_{ij} = 0, \quad \text{and} \tag{6.15b}$$

$$\psi'(M^*_{ij}) - \psi'(1 - M^*_{ij}) + \lambda^*_{ij} - \kappa^*_{ij} + \nu^*_i - \nu^*_j = 0. \tag{6.15c}$$

Since $M^*_{ij} \in (0,1)$ for all pairs (i,j), the KKT conditions imply that $\lambda^*_{ij} = 0$ and $\kappa^*_{ij} = 0$. Consequently, equation (6.15c) takes the simpler form

$$\begin{aligned} \nu^*_j - \nu^*_i \;&=\; \psi'(M^*_{ij}) - \psi'(1 - M^*_{ij}) = \frac{1}{2}\Phi^{-1}(M^*_{ij}) - \frac{1}{2}\Phi^{-1}(1 - M^*_{ij}) \\ &\overset{\text{(i)}}{=} \Phi^{-1}(M^*_{ij}), \end{aligned} \tag{6.16}$$

where step (i) follows because $\Phi(t) = 1 - \Phi(-t)$ for all $t \in \mathbb{R}$ by assumption. It follows that $M^*_{ij} = \Phi(\nu^*_j - \nu^*_i)$ for all pairs (i,j), meaning that M^* takes a parametric form, as claimed.

7. Discussion. In this paper, we considered the problem of finding a partial or complete ranking from active pairwise comparisons. We proved that a simple and computationally efficient algorithm succeeds in recovering the ranking with a sample complexity that is optimal up to logarithmic factors. We furthermore proved that this algorithm remains optimal when imposing common parametric assumptions such as the popular BTL or Thurstone models—provided the pairwise comparison probabilities are bounded away from 0 and 1. This show that, perhaps surprisingly, imposing common parametric assumptions cannot reduce the sample complexity of ranking by more than a log-factor in the stochastic regime. That being said, it should be noted that in practice, the possibility of gaining (at most) a log factor from assuming the parametric model may be overshadowed by the significant additional robustness afforded by our more general model class. For instance, see Ballinger et al. [3] for some empirical evidence that parametric models do not provide good fit in many applications, and, as our numerical results demonstrated, algorithms relying on parametric models can be quite sensitive to violations of those modeling assumptions.

There are a number of open and practically relevant questions suggested by our work. From a theoretical perspective, it would be interesting to provide an algorithm and corresponding guarantees for parametric models that matches our lower bound in the regime where the comparison probabilities are bounded away from zero and one, and at the same time is optimal in the regime where the pairwise comparison probabilities are very close to zero and one. A final interesting topic is related to approximate rankings, as in

practice, one might only be interested in finding an approximate ranking, or might only be able to find an approximate ranking due to a limited budget.

References.

[1] Aggarwal, C. C. (2016). *Recommender systems: The textbook*. Springer.

[2] Ailon, N. (2011). Active learning ranking from pairwise preferences with almost optimal query complexity. In *Advances in Neural Information Processing Systems* 810–818.

[3] Ballinger, T. P. and Wilcox, N. T. (1997). Decisions, error and heterogeneity. *The Economic Journal* **107** 1090–1105.

[4] Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

[5] Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika* **39** 324–345.

[6] Bubeck, S. and Cesa-Bianchi, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning* **5**.

[7] Bubeck, S., Wang, T. and Viswanathan, N. (2013). Multiple identifications in multi-armed bandits. In *International Conference on Machine Learning* 258–265.

[8] Busa-Fekete, R., Szorenyi, B., Cheng, W., Weng, P. and Hüllermeier, E. (2013). Top-k selection based on adaptive sampling of noisy preferences. In *International Conference on Machine Learning* 1094–1102.

[9] Chen, Y. and Suh, C. (2015). Spectral MLE: Top-k rank aggregation from pairwise comparisons. In *International Conference on Machine Learning* 371-380.

[10] Chen, S., Lin, T., King, I., Lyu, M. R. and Chen, W. (2014). Combinatorial Pure Exploration of Multi-Armed Bandits. In *Advances in Neural Information Processing Systems* 379–387.

[11] Chen, X., Gopi, S., Mao, J. and Schneider, J. (2017). Competitive analysis of the top-K ranking problem. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms* 1245–1264.

[12] De Borda, J. C. (1781). Mémoire sur les élections au scrutin. *Histoire de l'Académie Royale des Sciences*.

[13] Eriksson, B. (2013). Learning to top-k search using pairwise comparisons. In *International Conference on Machine Learning* 265–273.

[14] Even-Dar, E., Mannor, S. and Mansour, Y. (2006). Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal on Machine Learning Research* **7** 1079–1105.

[15] Hajek, B., Oh, S. and Xu, J. (2014). Minimax-optimal inference from partial rankings. In *Advances in Neural Information Processing Systems* 1475–1483.

[16] Hunter, D. (2004). MM algorithms for generalized Bradley-Terry models. *Annals of Statistics* 384–406.

[17] Jamieson, K. and Nowak, R. (2011). Active ranking using pairwise comparisons. In *Advances in Neural Information Processing Systems* 2240–2248.

[18] Jamieson, K., Malloy, M., Nowak, R. and Bubeck, S. (2014). lil' UCB: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory* 423–439.

[19] Jamieson, K., Katariya, S., Deshpande, A. and Nowak, R. (2015). Sparse dueling bandits. In *International Conference on Artificial Intelligence and Statistics* 416–424.

[20] Joe, H. (1988). Majorization, entropy and paired comparisons. *The Annals of Statistics* **16** 915–925.

[21] Kalyanakrishnan, S., Tewari, A., Auer, P. and Stone, P. (2012). PAC Subset Selection in Stochastic Multi-armed Bandits. In *International Conference on Machine Learning* 655–662.
[22] Kaufmann, E., Cappé, O. and Garivier, A. (2016). On the complexity of best arm identification in multi-armed bandit models. *Journal on Machine Learning Research* **17** 1–42.
[23] Landau, H. G. (1953). On dominance relations and the structure of animal societies: III The condition for a score structure. *The Bulletin of Mathematical Biophysics* **15** 143–148.
[24] Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.
[25] Maystre, L. and Grossglauser, M. (2015). Robust active ranking from sparse noisy comparisons. *arXiv:1502.05556*.
[26] Negahban, S. N., Oh, S. and Shah, D. (2012). Iterative ranking from pair-wise comparisons. In *Advances in Neural Information Processing Systems* 2474–2482.
[27] Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. *The Annals of Mathematical Statistics* **35** 174–180.
[28] Piech, C., Huang, J., Chen, Z., Do, C., Ng, A. and Koller, D. (2013). Tuned models of peer assessment in MOOCs. In *International Conference on Educational Data Mining*.
[29] Salganik, M. J. and Levy, K. E. C. (2015). Wiki surveys: Open and quantifiable social data collection. *PLOS ONE* **10** e0123483.
[30] Shah, N. B. and Wainwright, M. J. (2015). Simple, robust and optimal ranking from pairwise comparisons. *arXiv:1512.08949*.
[31] Shah, N. B., Bradley, J. K., Parekh, A., Wainwright, M. J. and Ramchandran, K. (2013). A case for ordinal peer-evaluation in MOOCs. In *NIPS Workshop on Data Driven Education*.
[32] Shah, N. B., Balakrishnan, S., Bradley, J., Parekh, A., Ramchandran, K. and Wainwright, M. J. (2016a). Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. *Journal on Machine Learning Research* 11–20.
[33] Shah, N. B., Balakrishnan, S., Guntuboyina, A. and Wainwright, M. J. (2016b). Stochastically transitive models for pairwise comparisons: Statistical and computational issues. In *International Conference on Machine Learning*.
[34] Soufiani, H., Parkes, D. and Xia, L. (2014). Computing parametric ranking models via rank-breaking. In *International Conference on Machine Learning* 360–368.
[35] Szörényi, B., Busa-Fekete, R., Paul, A. and Hüllermeier, E. (2015). Online rank elicitation for Plackett-Luce: A dueling bandits approach. In *Advances in Neural Information Processing Systems* 604–612.
[36] Thurstone, L. (1927). A law of comparative judgment. *Psychological Review* **34** 273–286.
[37] Tversky, A. and Edward Russo, J. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology* **6** 1–12.
[38] Urvoy, T., Clerot, F., Féraud, R. and Naamane, S. (2013). Generic exploration and K-armed voting bandits. In *International Conference on Machine Learning* 91–99.
[39] Yue, Y. and Joachims, T. (2011). Beat the mean bandit. In *International Conference on Machine Learning* 241–248.
[40] Yue, Y., Broder, J., Kleinberg, R. and Joachims, T. (2012). The K-armed dueling bandits problem. *Journal of Computer and System Sciences* **78** 1538–1556.

SUPPLEMENTARY MATERIAL

Additional numerical results and proofs

(doi: COMPLETED BY THE TYPESETTER; .pdf). We provide additional numerical results as well as the proofs of some of the results in our paper.

8. Additional numerical results. In this section we present additional numerical results.

8.1. *Comparison to algorithms tailored to parametric models.* In the experiment reported in Figure 6, we fixed $n = 10$, and saw that the PLPAC and BTMB algorithm improve slightly over the AR algorithm if $M_{\max}$ is very close to one. In our next experiment, we show that this relative improvement does only mildly depend on n. Specifically, we consider the model $M^{(\xi)}$, with parameters $w_i = \xi(n-i), i = 1, \dots, n$, where for each n, we choose ξ such that $M_{\max} \approx 0.9$; thus we are in the regime where PLPAC outperforms the AR algorithm. As the results in Figure 6(a) show, even in that regime, the AR algorithm requires at most 30% more samples across all n.

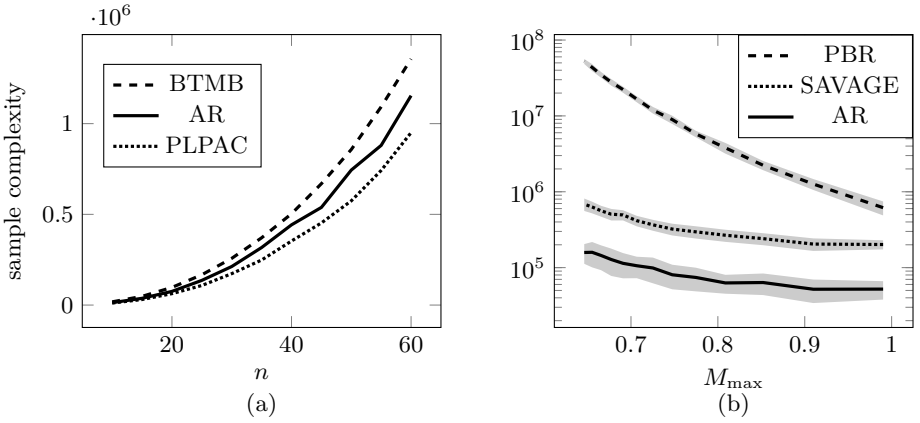


Fig 6: (a) Empirical sample complexity of the AR, PLPAC, and BTMB algorithms applied to the BTL model $M^{(\xi)}$ with parameters $w_i = \xi(n - i), i = 1, \dots, 10$ chosen such that $M_{\max} = 0.9$. Even in this regime where the PLPAC algorithm has a lower sample complexity than the AR algorithm, the difference is minor (30% improvement at the most). (b) Empirical sample complexity of the PBR, SAVAGE, and AR algorithms applied to the BTL model $M^{(\eta)}$ with parameters $w_i = \log(\eta + n - i), i = 1, \dots, 10$, as a function of $M_{\max} := \max_{i,j} M_{ij}$. We varied η such that $M_{\max} \in [0.65, 0.99]$. The error bars correspond to one standard deviation from the mean.

8.2. *Comparison to top item identification algorithms.* We next compare the AR algorithm to the SAVAGE algorithm introduced in [38] and the PBR algorithm from [8]. Both algorithms provably succeed in recovering the top item. We perform the same experiment as in Section 5.1. The results, depicted in Figure 6(b), show that the AR algorithm has a significantly lower sample complexity. This observation may be explained as follows. The SAVAGE is a generic elimination algorithm, which, when applied to top-1 recovery, is similar to our approach, however the corresponding confidence bound α_t is larger which results in a larger sample complexity. The PBR algorithm proposed in [8] estimates the *pairwise comparison probabilities* instead of the scores directly, and therefore is not competitive in terms of sample complexity.

8.3. *Selection of confidence interval.* Recall that the AR algorithm eliminates an item if the confidence that it belongs to one of the sets $\mathcal{S}_1, \dots, \mathcal{S}_L$, is sufficiently large. Our main results show that the AR algorithm succeeds at recovering the ranking with probability at least $1-\delta$, provided that the length of the confidence interval is chosen as $\alpha_t = \sqrt{\frac{\log(n/\delta)+0.75\log\log(n/\delta)+1.5\log(1+\log(t/2))}{2t}}$. While this result is optimal up to log-factors, the particular choice of the constants might be overly conservative, and improvements in the (empirical) sample complexity might be obtained by choosing the constants in α_t smaller, as we show next. To investigate this claim, we set $\alpha_t = \frac{1}{4}\sqrt{\frac{\log(n/3(\log(t)+1)/\delta)}{t}}$. We generate a pairwise comparison model with $n = 5$, scores $\tau = (0.9, 0.7, 0.5, 0.3, 0.1)$, and use the AR algorithm to find the top 2 items, for different values of the desired accuracy δ. The results, depicted in Figure 7, show that, even with those significantly smaller constants, the AR algorithm is δ-accurate.

9. Proof of Lemmas 1, 2, and 3. In this section we provide the proofs of Lemmas 1, 2, and 3, which are part of the proof of Theorem 1(a).

9.1. *Proof of Lemma 1.* In order to show that the event $\mathcal{E}_\alpha$ occurs with probability at least $1-\delta$, first recall that comparing item i to an item chosen uniformly at random from $[n]\backslash\{i\}$ is equivalent to taking an independent draw from a Bernoulli random variable with mean τ_i. One can verify from the recursion (3.2) that $\widehat{\tau}_i(t)$ is a sum of t independent Bernoulli random variables, each of which has mean τ_i/t. In order to control the fluctuations of $\widehat{\tau}_i(t)$, we make use of a non-asymptotic version of the law of the iterated logarithm from Kaufmann, Cappé and Garivier [22] and Jamieson et al. [18].

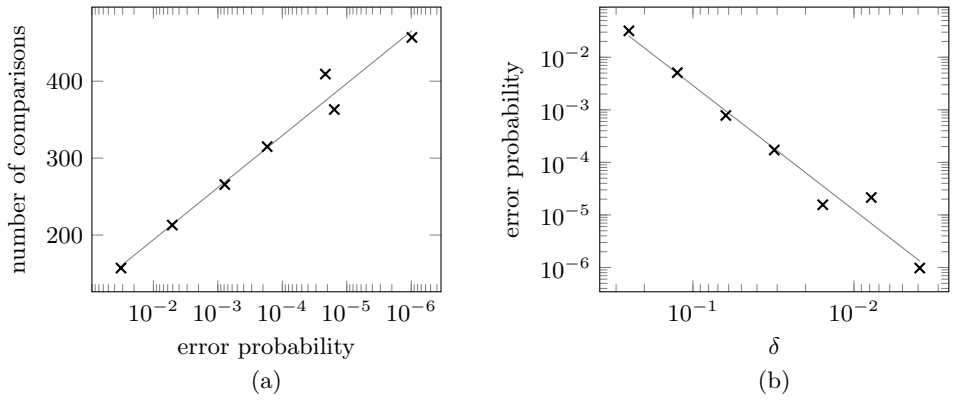


Fig 7: (a) Number of comparisons required to find the top-2 items out of 5 items, for $\alpha_t = \frac{1}{4}\sqrt{\log(n/3(\log(t)+1)/\delta)/t}$. (b) Empirical error probability required to find the top-2 items out of 5 items. The particular choice of the constants in $\alpha_t = \sqrt{\log(125n\log(1.12t)/\delta)/t}$ in our theoretical results is very conservative, in the sense that for obtaining a δ-accurate ranking, the constants in α_t can be chosen smaller, which in turn results in fewer comparisons.

LEMMA 5 ([22, Lem. 19]). *Given an i.i.d. sequence $\{X_s\}_{s=1}^{\infty}$ of Bernoulli variables with mean μ, then for any $\delta' \in (0, 0.0005)$, we have*

$$\left|\frac{1}{t}\sum_{s=1}^{t}(X_s - \mu)\right| \leq \sqrt{\frac{\beta(t,\delta')}{2t}}, \qquad \text{for all } t \geq 1$$

with probability at least $1-\delta'$. Here, $\beta(t,\delta') := \log(1/\delta') + 0.75\log\log(1/\delta') + 1.5\log(1+\log(t/2))$. The statements continues to hold for any $\delta \in (0,1)$ with $\beta(t,\delta') = 2\log(125\log(1.12t)/\delta')$.

In the current context, applying Lemma 5 with $\delta' = n/\delta$ and $\alpha_t = \sqrt{\frac{\beta(t,\delta')}{2t}}$ yields

$$\mathbb{P}\Big[\,|\widehat{\tau}_i(t) - \tau_i| \geq \alpha_t \quad \text{for some } t \geq 1\Big] \leq \frac{\delta}{n}.$$

Taking the union bound over all indices $i \in \mathcal{S} \subseteq [n]$ yields that $\mathbb{P}[\mathcal{E}_\alpha] \geq 1-\delta$, as claimed.

9.2. *Proof of Lemma 2.* We show that implications (6.2a) and (6.2b) follow from the inequality in event $\mathcal{E}_\alpha$. We start by showing that the inequality

in event $\mathcal{E}_\alpha$ implies that

$$|\tau_{\{k\}} - \tau_{(k)}| \leq 2\alpha_t. \tag{9.1}$$

We claim that $\tau_{\{k\}} - \tau_{(k)} \geq -2\alpha_t$. By definition of (k), there are k indices $\{i_1, \ldots, i_k\}$ such that $\widehat{\tau}_{i_\ell}(t) \geq \widehat{\tau}_{(k)}(t)$ for every $\ell \in [k]$. In conjunction with the inequality in event $\mathcal{E}_\alpha$, we obtain

$$\tau_{i_\ell} + \alpha_t \geq \tau_{(k)} - \alpha_t.$$

Since this inequality holds for k many indices $\{i_1, \ldots, i_k\}$, it must also hold for $\{k\}$, due to $\tau_{\{1\}} \geq \tau_{\{2\}} \geq \ldots \geq \tau_{\{|\mathcal{S}|\}}$. It follows that $\tau_{\{k\}} - \tau_{(k)} \geq -2\alpha_t$. It remains to establish that $\tau_{\{k\}} - \tau_{(k)} \leq 2\alpha_t$. By definition of (k), there are $|\mathcal{S}| - k + 1$ many indices j in $\mathcal{S}$ obeying

$$\widehat{\tau}_j(t) \leq \widehat{\tau}_{(k)}(t).$$

By the inequality in event $\mathcal{E}_\alpha$, this yields $\tau_j - \alpha_t \leq \tau_{(k)} + \alpha_t$. Since this inequality holds for $|\mathcal{S}| - k + 1$ indices j, it must hold for $j = \{k\}$, which implies $\tau_{\{k\}} \leq \tau_{(k)} + 2\alpha_t$. Thus, we have established that inequality (9.1) must hold under event $\mathcal{E}_\alpha$.

We are now ready to establish the claim (6.2a). As long as $\widehat{\tau}_j(t) < \widehat{\tau}_{(\widehat{k}_{\ell-1})} - 4\alpha_t$, we have

$$\begin{aligned} 4\alpha_t \;<\; \widehat{\tau}_{(\widehat{k}_{\ell-1})}(t) - \widehat{\tau}_j(t) &\overset{\text{(i)}}{\leq} \tau_{(\widehat{k}_{\ell-1})} + \alpha_t - \tau_j + \alpha_t \\ &\leq \tau_{\{\widehat{k}_{\ell-1}\}} - \tau_j + 2\alpha_t + 2\alpha_t. \end{aligned} \tag{9.2}$$

Here step (i) follows by the inequality in event $\mathcal{E}_\alpha$, whereas inequality (9.2) follows by inequality (9.1). Noting that inequality (9.2) is equivalent to $\tau_j < \tau_{\{\widehat{k}_{\ell-1}\}}$, we have established the claim (6.2a). The proof of claim (6.2b) is analogous, so we omit the details.

9.3. *Proof of Lemma 3.* We first prove that, if the event $\mathcal{E}_\alpha$ occurs, then for any given index $i \in \mathcal{S}_\ell$, and for all $\ell > 1$, inequality (6.3a) holds. Recall that $\tau_{\{k\}}$ is the k-th largest score out of the latent scores $\tau_i, i \in \mathcal{S}$. On the event $\mathcal{E}_\alpha$, we have, for any $\bar{t}_i$, that

$$\begin{aligned} \widehat{\tau}_{(\widehat{k}_{\ell-1})}(\bar{t}_i) \;\geq\; \tau_{(\widehat{k}_{\ell-1})} - \alpha_{\bar{t}_i} &\overset{\text{(i)}}{\geq} \tau_{\{\widehat{k}_{\ell-1}\}} - 3\alpha_{\bar{t}_i} \\ &\overset{\text{(ii)}}{\geq} \tau_{k_{\ell-1}} - 3\alpha_{\bar{t}_i} \\ &= \bar{\Delta}_{\ell,i} - 3\alpha_{\bar{t}_i} + \tau_i \\ &\overset{\text{(iii)}}{\geq} \bar{\Delta}_{\ell,i} - 4\alpha_{\bar{t}_i} + \widehat{\tau}_i(\bar{t}_i). \end{aligned} \tag{9.3}$$

Here step (i) follows by inequality (9.1), which holds on the event $\mathcal{E}_\alpha$, and inequality (ii) follows from $\tau_{\{\widehat{k}_{\ell-1}\}} \geq \tau_{k_{\ell-1}}$, which is seen as follows. As shown above, on the event $\mathcal{E}_\alpha$, the AR algorithm never misclassifies any item, meaning that an item j is only assigned to the set $\widehat{\mathcal{S}}_\ell$ if $j \in \mathcal{S}_\ell$. Therefore the $\widehat{k}_{\ell-1}$-th largest score among the items in the set $\mathcal{S}$ must be larger or equal to the $k_{\ell-1}$-th largest score among all scores. Finally, inequality (iii) follows from $\widehat{\tau}_i(\bar{t}_i) - \tau_i \leq \alpha_{\bar{t}_i}$, which holds on the event $\mathcal{E}_\alpha$.

From the definition of $\alpha_{\bar{t}_i}$, some algebra leads to the lower bound

$$\bar{\Delta}_{\ell,i} > 8\alpha_{\bar{t}_i}, \tag{9.4}$$

with $\bar{t}_i$ as defined in Lemma 3. See the end of this subsection for details of this calculation. Application of inequality (9.4) on the RHS of inequality (9.3) yields

$$\widehat{\tau}_{(\widehat{k}_{\ell-1})}(\bar{t}_i) > 4\alpha_{\bar{t}_i} + \widehat{\tau}_i(\bar{t}_i),$$

which concludes the proof of inequality (6.3a). Analogously, it follows that inequality (6.3b) holds for a given item i if the event $\mathcal{E}_\alpha$ occurs.

Proof of the lower bound (9.4). We prove the bound for the more conservative choice of α_t given by $\alpha_t = \sqrt{\frac{\log(125n\log(1.12t)/\delta)}{t}}$. By definition of α_t and $\bar{t}_i$, we have that

$$\begin{aligned}
\alpha^2_{\bar{t}_i} &= \frac{\log\left(\frac{125n}{\delta}\log\left[1.12\frac{c_1}{\bar{\Delta}^2_{\ell,i}}\log\left(\frac{n}{\delta}\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)\right]\right)}{\frac{c_1}{\bar{\Delta}^2_{\ell,i}}\log\left(\frac{n}{\delta}\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)} \\
&\leq \frac{\bar{\Delta}^2_{\ell,i}}{c_1}\frac{\log\left(\frac{125n}{\delta}\log\left[\frac{1.12c_1}{4}\frac{n}{\delta}\frac{2^3}{\bar{\Delta}^3_{\ell,i}}\right]\right)}{\log\left(\frac{n}{\delta}\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)} \\
&\leq \frac{\bar{\Delta}^2_{\ell,i}}{c_1}\frac{\log\left(\frac{125n}{\delta}\left(\frac{1.12c_1}{4}\frac{n}{\delta}\right)^{1/3}3\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)}{\log\left(\frac{n}{\delta}\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)} \\
&\leq \frac{\bar{\Delta}^2_{\ell,i}}{c_1}\frac{\frac{4}{3}\log\left(375\left(\frac{1.12c_1}{4}\right)^{1/3}\frac{n}{\delta}\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)}{\log\left(\frac{n}{\delta}\log\left(\frac{2}{\bar{\Delta}_{\ell,i}}\right)\right)}, \\
&< \frac{\bar{\Delta}^2_{\ell,i}}{8^2},
\end{aligned}$$

where the last inequality holds for $c_1 = 654$.

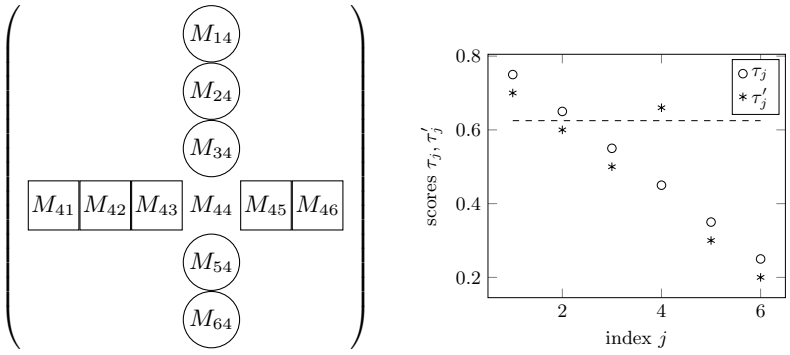


Fig 8: Illustration of the distributions M, M' and the corresponding scores τ_j, τ'_j: Suppose that $m = 4$ and $k_{\ell-1} = 2$. The probabilities M'_{ij} are obtained from the probabilities M_{ij} by increasing the probabilities surrounded by a rectangle, and decreasing the probabilities surrounded by a circle, all others remain unchanged.

10. Proof of Theorem 2(a). As stated before in the original paper, our goal is to prove that any algorithm $\mathcal{A}$ that is uniformly δ-accurate over $\mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$, when applied to a given pairwise comparison model $M \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$, must make at least

$$\mathbb{E}_M[Q] \geq \frac{M_{\min}\phi_{\min}^2}{2.004\phi_{\max}^2} \log\left(\frac{1}{2\delta}\right) F(\tau(M))$$

comparisons on average. Here $F(\tau(M))$ is the complexity parameter defined in Section 4.

The proof is similar to that of Theorem 1(b), with the primary difference being that the alternative matrix M' must now be constructed such that it lies in the parametric class. In what follows, we show how to modify the proof of Theorem 1 at appropriate positions in order to accommodate this difference.

Consider any parametric pairwise comparison matrix $M \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$. Then there exists a parameter vector $w \in \mathbb{R}^n$ such that $M_{ij} = \Phi(w_i - w_j)$. By the assumption $\tau_1 > \ldots > \tau_n$, this parameter vector obeys $w_1 > \ldots > w_n$. Consider an item $m \in \mathcal{S}_\ell(M), \ell > 1$, and set $k := k_{\ell-1}$, for notational convenience. We construct an alternative matrix $M' \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$ as follows. Consider some scalar value ρ that lies in the interval $0 <$

$\rho < w_k - w_{k-1}$. Define a set of alternative parameters as

$$w_i' := \begin{cases} w_k & \text{if } i = m, \\ w_k - \rho & \text{if } i = k, \\ w_i & \text{otherwise.} \end{cases}$$

Now let M' be the matrix with pairwise comparison probabilities $M'_{ij} = \Phi(w'_i - w'_j)$. By definition, we have $w_1 \geq w'_i \geq w_n$ for all $i \in [n]$, which ensures that $M' \in \mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$. Moreover, by definition, item m is among the top k items, so that $m \notin \mathcal{S}_\ell(M')$. Since (by assumption) algorithm $\mathcal{A}$ is uniformly δ-accurate over $\mathcal{C}_{\mathrm{PAR}}(\Phi) \cap \mathcal{C}_{M_{\min}}$, we have both $\mathbb{P}_M[\mathcal{E}] \geq 1 - \delta$ and $\mathbb{P}_{M'}[\mathcal{E}] \leq \delta$, which ensures that inequality (6.7) holds. Here $\mathcal{E}$ denotes the previously defined event (6.5) that the algorithm $\mathcal{A}$ correctly recovers the set structure.

Next consider the total number of comparisons of item m with all others items, denoted by Q_m. As in inequality (6.9), we are guaranteed that

$$\begin{aligned} \max_{j \in [n] \setminus \{m\}} d(M_{mj}, M'_{mj}) \mathbb{E}_M[Q_m] &\geq \sum_{j \in [n] \setminus \{m\}} \mathbb{E}_M[Q_{mj}]\, d(M_{mj}, M'_{mj}) \\ &\overset{\text{(i)}}{=} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{E}_M[Q_{ij}]\, d(M_{ij}, M'_{ij}) - \sum_{j \in [n] \setminus \{k,m\}} \mathbb{E}_M[Q_{jk}]\, d(M_{jk}, M'_{jk}) \\ &\overset{\text{(ii)}}{\geq} d(\mathbb{P}_M[\mathcal{E}], \mathbb{P}_{M'}[\mathcal{E}]) - \sum_{j \in [n] \setminus \{k,m\}} \mathbb{E}_M[Q_{jk}]\, d(M_{jk}, M'_{jk}) \\ &\overset{\text{(iii)}}{\geq} d(\mathbb{P}_M[\mathcal{E}], \mathbb{P}_{M'}[\mathcal{E}]) - 0.001 d(\mathbb{P}_M[\mathcal{E}], \mathbb{P}_{M'}[\mathcal{E}]) \\ &\geq 0.999 \log \frac{1}{2\delta}. \end{aligned} \tag{10.1}$$

Here inequality (i) follows from the fact that $d(M_{ij}, M'_{ij}) = 0$ for all (i, j) with $i, j \in [n] \setminus \{k, m\}$, by definition of M'. Inequality (ii) follows from inequality (6.6) (that is, from Lemma 4). Inequality (iii) is a result of the fact that $\lim_{\rho \to 0} d(M_{ik}, M'_{ik}) = \lim_{\rho \to 0} d(\Phi(w_i - w_k), \Phi(w_i - w_k + \rho)) = 0$ for every $i \in [n] \setminus \{k, m\}$, where we have also employed the continuous mapping theorem: Due to this relation we can choose ρ sufficiently close to 0 to obtain the bound of (iii). Finally, inequality (10.1) is a consequence of inequality (6.7).

Our next step is to upper bound the KL divergence $d(M_{mj}, M'_{mj})$ For

each $j \in [n] \setminus \{k, m\}$, we have

$$\begin{aligned} d(M_{mj}, M'_{mj}) &\leq \frac{(M'_{mj} - M_{mj})^2}{M'_{mj}(1 - M'_{mj})} \overset{\text{(i)}}{=} \frac{(M_{kj} - M_{mj})^2}{M_{kj}(1 - M_{kj})} \\ &\overset{\text{(ii)}}{\leq} \frac{2}{M_{\min}} (\Phi(w_k - w_j) - \Phi(w_m - w_j))^2 \\ &\overset{\text{(iii)}}{\leq} \frac{2}{M_{\min}} (\phi_{\max}(w_k - w_m))^2 \leq \frac{2\phi_{\max}^2}{M_{\min}\phi_{\min}^2} (\tau_k - \tau_m)^2. \end{aligned} \tag{10.2}$$

Here step (i) follows by definition of the parameters M'_{ij}; step (ii) follows because M_{ij} belongs to the interval $[M_{\min}, 1 - M_{\min}]$; and step (iii) is a consequence of assumption (4.1). Finally, the last inequality (10.2) follows from the relations

$$\begin{aligned} \tau_k - \tau_m &= \frac{1}{n-1} \Big(\Phi(w_k - w_m) - \Phi(w_m - w_k) \\ &\quad + \sum_{j \in [n] \setminus \{k,m\}} (\Phi(w_k - w_j) - \Phi(w_m - w_j)) \Big) \\ &\overset{\text{(i)}}{\geq} \frac{1}{n-1} \Big(\phi_{\min}(w_k - w_m - (w_m - w_k)) + \sum_{j \in [n] \setminus \{k,m\}} \phi_{\min}(w_k - w_m) \Big) \\ &= \frac{n}{n-1} \phi_{\min}(w_k - w_m) \geq \phi_{\min}(w_k - w_m). \end{aligned} \tag{10.3}$$

Here inequality (i) follows from assumption (4.1); in particular, recall that $w_k > w_m$, so the difference $w_k - w_m$ above is positive.

Similarly, we have

$$\begin{aligned} d(M_{mk}, M'_{mk}) \;&\leq\; \frac{2}{M_{\min}} (\phi_{\max}(\rho + w_k - w_m))^2 \overset{\text{(i)}}{\leq} \frac{2.001}{M_{\min}} (\phi_{\max}(w_k - w_m))^2 \\ &\overset{\text{(ii)}}{\leq} \frac{2.001\phi_{\max}^2}{M_{\min}\phi_{\min}^2} (\tau_k - \tau_m)^2, \end{aligned} \tag{10.4}$$

where inequality (i) follows from choosing ρ sufficiently close to 0, whereas inequality (ii) follows from the relation (10.3).

Given an index m in a set $\mathcal{S}_\ell(M)$ with $\ell > 1$, combining inequalities (10.2) and (10.4) with inequality (10.1) yields

$$\mathbb{E}_M [Q_m] \geq \frac{M_{\min}\phi_{\min}^2}{2.004\phi_{\max}^2} \frac{\log(1/(2\delta))}{(\tau_{k_{\ell-1}} - \tau_m)^2}. \tag{10.5a}$$

Similarly, for an index $m \in \mathcal{S}_\ell(M)$ with $\ell < L$, we define an alternative matrix M' by defining corresponding parameters as $w'_m = w_{k_\ell+1}$, $w'_{k_\ell+1} = w_{k_\ell+1} + \rho$ for $\rho \in (0, w_{k_\ell} - w_{k_\ell+1})$, and $w'_i = w_i$, for all $i \notin \{m, k_\ell + 1\}$. Under the model specified by M', item m is not amongst the k_ℓ items with the largest scores, and therefore $m \notin \mathcal{S}_\ell(M')$. The same line of arguments as above yields

$$\mathbb{E}_M [Q_m] \geq \frac{M_{\min}\phi^2_{\min}}{2.004\phi^2_{\max}} \frac{\log(1/(2\delta))}{(\tau_m - \tau_{k_\ell+1})^2}. \tag{10.5b}$$

Combining the lower bounds (10.5a) and (10.5b) concludes the proof.

Reinhard Heckel
Dept. of Electrical and Computer Engineering
Rice University
6100 Main St
Houston, TX 77005
E-mail: rh43@rice.edu

Nihar B. Shah
Machine Learning Department
and Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
E-mail: nihars@cs.cmu.edu

Kannan Ramchandran
Dept. of Electrical Engineering
and Computer Sciences
Cory Hall
Berkeley, CA 94720
E-mail: kannanr@eecs.berkeley.edu

Martin J. Wainwright
Dept. of Electrical Engineering
and Computer Sciences
and Department of Statistics
Cory Hall
Berkeley, CA 94720
E-mail: wainwrig@berkeley.edu