PAC Rank Elicitation through Adaptive
Sampling of Stochastic Pairwise Preferences

Róbert Busa-Fekete[*]
University of Marburg, Germany
busarobi@inf.u-szeged.hu

Balázs Szörényi[*]
INRIA Lille - Team SequeL, France
szorenyi@inf.u-szeged.hu

Eyke Hüllermeier
University of Paderborn, Germany
eyke@upb.de

Abstract

We introduce the problem of PAC rank elicitation, which consists of sorting a given set of options based on adaptive sampling of stochastic pairwise preferences. More specifically, we assume the existence of a *ranking procedure*, such as Copeland's method, that determines an underlying *target order* of the options. The goal is to predict a ranking that is sufficiently close to this target order with high probability, where closeness is measured in terms of a suitable distance measure. We instantiate this setting with combinations of two different distance measures and ranking procedures. For these instantiations, we devise efficient strategies for sampling pairwise preferences and analyze the corresponding sample complexity. We also present first experiments to illustrate the practical performance of our methods.

Introduction

Exploiting revealed (pairwise) preferences to learn a ranking (total order) over a set of options is a challenging problem with many practical applications. For example, think of crowd-sourcing services like the Amazon Mechanical Turk, where simple questions such as pairwise comparisons between decision alternatives are asked to a group of annotators. The task is to approximate an underlying target ranking on the basis of these pairwise comparisons, which are possibly noisy and partially inconsistent (Chen et al. 2013). Another application worth mentioning is the ranking of XBox gamers based on their pairwise online duels; the ranking system of XBox is called TrueSkill™(Guo et al. 2012).

In this paper, we focus on a problem that we call *PAC rank elicitation*. In the setting of this problem, we consider a finite set of options $\mathcal{O} = \{o_1, \dots, o_K\}$, on which a weighted relation $\mathbf{Y} = (y_{i,j})_{1 \le i,j \le K}$ is defined. As will be explained in more detail later on, this relation specifies the probability of observing preferences $o_j \prec o_i$, suggesting that, in a single comparison of two options o_i and o_j, the former was liked more than the latter. Furthermore, we assume the existence

[1]Róbert Busa-Fekete and Balázs Szörényi are also affiliated with the MTA-SZTE Research Group on Artificial Intelligence

of a *ranking procedure* $\mathcal{R}$ that determines an underlying *target (strict) order* $\prec^*$ of the options $\mathcal{O}$ based on $\mathbf{Y}$.

In rank elicitation, we assume that $\mathcal{R}$ is given whereas $\mathbf{Y}$ is not known. Instead, information about $\mathbf{Y}$ can only be obtained through (adaptive) sampling of pairwise preferences. The goal, then, is to quickly gather enough information so as to enable the prediction of a ranking that is sufficiently close to the target order $\prec^*$ with high probability. We shall describe this rank elicitation setting more formally and, moreover, instantiate it with combinations of two different distance measures and two ranking procedures for determining the target order. For these instantiations, we devise efficient sampling strategies and analyze them in terms of expected sample complexity. Finally, we also present an experimental study, prior to concluding the paper.

Related work

Ranking based on sampling pairwise relations has a long history in the literature (Braverman and Mossel 2008; 2009; Eriksson 2013; Feige et al. 1994). Existing algorithms for *noisy sorting* typically solve this problem with sample complexity $O(K \log K)$. However, these algorithms make strong assumptions: the target relation is a total order, and the comparisons are representative of that order (if o_i precedes o_j, then $\mathbb{P}(o_i \prec o_j) > 1/2$).

Pure exploration algorithms for the stochastic multi-armed bandit problem sample the arms a certain number of times (not necessarily known in advance), and then output a recommendation, such as the best arm or the m best arms (Bubeck, Munos, and Stoltz 2009; Even-Dar, Mannor, and Mansour 2002; Bubeck, Wang, and Viswanathan 2013; Gabillon et al. 2011; Cappé et al. 2012). While our algorithm can be viewed as a pure exploration strategy, too, we do not assume that *numerical* feedback can be generated for *individual* options; instead, our feedback is *qualitative* and refers to *pairs* of options.

Seen from this point of view, our approach is closer to the dueling bandits problem introduced by (Yue et al. 2012), where feedback is provided in the form of noisy comparisons between option. However, apart from making strong structural assumptions (namely strong stochastic transitivity and stochastic triangle inequality), their problem of cumulative regret minimization is of an exploration-exploitation nature.

The kind of feedback assumed in our rank elicitation setup is in fact the one considered by (Busa-Fekete et al. 2013) and (Urvoy et al. 2013), who both solve the top-k subset selection (or EXPLORE-k) problem: Find the k best options with respect to a target ranking based on sampling pairwise preferences. Interestingly, rank elicitation can be seen as solving the top-k problem for all $k \in [K]$ simultaneously, and indeed, our approach builds on this connection. Our starting point is the recent paper (Kalyanakrishnan et al. 2012), which introduces a PAC-bandit algorithm for the top-k problem in the stochastic multi-armed bandit environment (i.e., based on numerical feedback, not pairwise preferences).

In the formulation of (Kalyanakrishnan et al. 2012), an algorithm is an (ϵ, m, δ)-PAC bandit algorithm if it selects the m best options (those with the highest expected value) under the PAC-bandit conditions (Even-Dar, Mannor, and Mansour 2002). The concrete algorithm they propose is based on the widely-known UCB index-based multi-armed bandit method (Auer, Cesa-Bianchi, and Fischer 2002). Our theoretical analysis partly relies on their results, using an expected sample complexity and a high probability bound for the worst case sample complexity. In fact, although our setup is based on preferences, we aim at a similar kind of sample complexity result.

Problem setting and terminology

PAC rank elicitation setup

Our point of departure are pairwise preferences over the set of options $\mathcal{O} = \{o_1, \dots, o_K\}$. More specifically, we allow three possible outcomes of a single pairwise comparison between o_i and o_j, namely (strict) preference for o_i, (strict) preference for o_j, and incomparability/indifference. These outcomes are denoted by $o_i \succ o_j$, $o_i \prec o_j$, and $o_i \perp o_j$, respectively. In our setting, we consider the outcome of a comparison between o_i and o_j as a random variable $Y_{i,j}$ which assumes the value 1 if $o_j \prec o_i$, 0 if $o_i \prec o_j$, and $1/2$ otherwise. Thus, the case $o_i \perp o_j$ is handled by giving half a point to both options. Essentially, this means that these outcomes are treated in a neutral way by the ranking procedures.

The expected values $y_{i,j} = \mathbb{E}[Y_{i,j}]$ can be summarized in the relation $\mathbf{Y} = [y_{i,j}] \in [0,1]^{K \times K}$. A natural idea to define a pairwise preference relation $\prec$ on $\mathcal{O}$ is to "binarize" $\mathbf{Y}$: $o_i \prec o_j$ if and only if $y_{i,j} < y_{j,i}$. This relation, however, may contain preferential cycles and, therefore, may not define a proper order relation. In decision making, this problem is commonly avoided by using a ranking procedure $\mathcal{R}$ (concrete choices of $\mathcal{R}$ will be discussed in the next section) that turns $\mathbf{Y}$ into a strict order relation $\prec^{\mathcal{R}}$ of the options $\mathcal{O}$. Formally, a ranking procedure $\mathcal{R}$ is a map $[0,1]^{K \times K} \to \mathcal{S}_{\mathcal{O}}$, where $\mathcal{S}_{\mathcal{O}}$ denotes the set of strict orders on $\mathcal{O}$. We denote the strict order produced by the ranking procedure $\mathcal{R}$ on the basis of $\mathbf{Y}$ by $\prec_{\mathbf{Y}}^{\mathcal{R}}$, or simply by $\prec^{\mathcal{R}}$ if $\mathbf{Y}$ is clear from the context.

The task in PAC rank elicitation is to approximate $\prec^{\mathcal{R}}$ without knowing the $y_{i,j}$. Instead, relevant information can only be obtained *through sampling pairwise comparisons from the underlying distribution*. Thus, we assume that options can be compared in a pairwise manner, and that a single sample essentially informs about a pairwise preference between two options o_i and o_j. The goal is to devise a *sampling strategy* that keeps the size of the sample (the sample complexity) as small as possible while producing an estimation $\prec$ that is "good" in a PAC sense: $\prec$ is supposed to be sufficiently "close" to $\prec^{\mathcal{R}}$ with high probability. Actually, our algorithms even produce a total order as a prediction, i.e., $\prec$ is a ranking that can be represented by a permutation τ of order K, where τ_i denotes the rank of option o_i in the order (with smaller ranks indicating higher preference, i.e., $o_i \prec o_j$ if $\tau_i > \tau_j$).

To formalize the notion of "closeness", we make use of appropriate distance measures that compare a (predicted) permutation τ with a (target) strict order $\prec$. In particular, we adopt the following two measures: The *number of discordant pairs* (NDP), which is closely connected to Kendall's rank correlation (Kendall 1955), and can be expressed in terms of the indicator function $\mathbb{I}\{\cdot\}$ as follows:

$$d_{\mathcal{K}}(\tau, \prec) = \sum_{i=1}^{K} \sum_{j \neq i} \mathbb{I}\{\tau_j > \tau_i\} \mathbb{I}\{o_i \prec o_j\}.$$

The *maximum rank difference* (MRD) is defined as the maximum difference between the rank of an object o_i according to τ and $\prec$, respectively. More specifically, since $\prec$ is a partial but not necessarily total order, we compare τ to the set $\mathcal{L}^{\prec}$ of its linear extensions[2]:

$$d_{\mathcal{M}}(\tau, \prec) = \min_{\tau' \in \mathcal{L}^{\prec}} \max_{1 \leq i \leq K} |\tau_i - \tau_i'|.$$

Our setup allows for small approximation errors, formalized by a tolerance parameter $\rho \in \mathbb{N}^+$.[3] We call an algorithm $\mathscr{A}$ a (ρ, δ)-*PAC rank elicitation algorithm* with respect to a ranking procedure $\mathcal{R}$ and rank distance d, if it returns a ranking τ for which $d(\tau, \prec^{\mathcal{R}}) < \rho$ with probability at least $1 - \delta$.

Ranking procedures

In the following, we introduce two instantiations of the ranking procedure $\mathcal{R}$, namely Copeland's ranking (binary voting) and the sum of expectations (weighted voting). To define the former, let $d_i = \#\{k \in [K] \mid 1/2 < y_{i,k}\}$ denote the number of options that are "beaten" by o_i. Copeland's ranking (CO) is then defined as follows (Moulin 1988): $o_i \prec^{\mathrm{CO}} o_j$ if and only if $d_i < d_j$. The sum of expectations (SE) ranking is a "soft" version of CO: $o_i \prec^{\mathrm{SE}} o_j$ if and only if

$$y_i = \frac{1}{K-1} \sum_{k \neq i} y_{i,k} < \frac{1}{K-1} \sum_{k \neq j} y_{j,k} = y_j \; . \quad (1)$$

Since $\mathcal{R}$ is mapping the continuous space $[0,1]^{K \times K}$ to the discrete space $\mathcal{S}_{\mathcal{O}}$, ranking is a "non-smooth" operation.

---

[2] $\tau \in \mathcal{L}^{\prec}$ iff $\forall \, i, j \in [K]: (o_i \prec o_j) \Rightarrow (\tau_j < \tau_i)$

[3] Note that our distance measures assume values in $\mathbb{N}_0$ and are not normalized. Although a normalization to $[0,1]$ could easily be done, it would unnecessarily complicate the description of the algorithms and their analysis.

In the case of the Copeland order $\prec^{\mathrm{CO}}$, for example, a minimal change of a value $y_{i,j} \approx \frac{1}{2}$ may strongly influence $\prec^{\mathrm{CO}}$. Consequently, the number of samples needed to assure (with high probability) a certain approximation quality may become arbitrarily large. A similar problem arises for $\prec^{\mathrm{SE}}$ as a target order if some of the individual scores y_i are very close or equal to each other.

As a practical (yet meaningful) solution to this problem, we propose to make the relations $\prec^{\mathrm{CO}}$ and $\prec^{\mathrm{SE}}$ a bit more "partial" by imposing stronger requirements on the strict order. To this end, let $d_i^* = \#\{k \,|\, 1/2 + \epsilon < y_{i,k}, i \neq k\}$ denote the number of options that are beaten by o_i with a margin $\epsilon > 0$, and let $s_i^* = \#\{k : |1/2 - y_{i,k}| \leq \epsilon, i \neq k\}$. Then, we define the ϵ-insensitive Copeland relation as follows: $o_i \prec^{\mathrm{CO}_\epsilon} o_j$ if and only if $d_i^* + s_i^* < d_j^*$. Likewise, in the case of $\prec^{\mathrm{SE}}$, we neglect small differences of the y_i and define the ϵ-insensitive sum of expectations relation as follows: $o_i \prec^{\mathrm{SE}_\epsilon} o_j$ if and only if $y_i + \epsilon < y_j$.

These ϵ-insensitive extensions are interval (and hence strict) orders, that is, they are obtained by characterizing each option o_i by the interval $[d_i^*, d_i^* + s_i^*]$ and sorting intervals according to $[a,b] \prec [a',b']$ iff $b < a'$. It is readily shown that $\prec^{\mathrm{CO}_\epsilon} \subseteq \prec^{\mathrm{CO}_{\epsilon'}} \subseteq \prec^{\mathrm{CO}}$ for $\epsilon > \epsilon'$, with equality $\prec^{\mathrm{CO}_0} \equiv \prec^{\mathrm{CO}}$ if $y_{i,j} \neq 1/2$ for all $i \neq j \in [K]$ (and similarly for SE). Subsequently, ϵ will be taken as a parameter that controls the strictness of the order relations, and thereby the difficulty of the (ρ, δ)-rank elicitation task.

A general rank elicitation algorithm

In this section, we introduce a general rank elicitation framework (RANKEL) that provides the basic statistics needed to solve the PAC rank elicitation problem, notably estimates of the pairwise probabilities $y_{i,j}$ and the number of samples drawn from $Y_{i,j}$ so far. It contains a subroutine that implements sampling strategies for the different distance measures and ϵ-insensitive ranking models.

Our general framework is shown in Algorithm 1. The set A contains all pairs of options that still need to be sampled; it is initialized with all $K^2 - K$ pairs of indices (line 3). In each iteration, the algorithm samples those $Y_{i,j}$ with $(i,j) \in A$ (lines 7) and maintains the estimates $\bar{\mathbf{Y}} = [\bar{y}_{i,j}]_{K \times K}$, where $\bar{y}_{i,j} = \frac{1}{n_{i,j}} \sum_{\ell=1}^{n_{i,j}} y_{i,j}^\ell$ is the mean of the $n_{i,j}$ samples drawn from $Y_{i,j}$ so far. These numbers are maintained by the algorithm, too, and are stored in the matrix $\mathbf{N} = [n_{i,j}]_{K \times K}$. The sampling strategy subroutine returns the indices of option pairs to be sampled. If A is empty, then RANKEL stops and returns a ranking τ over $\mathcal{O}$, which is calculated based on $\bar{\mathbf{Y}}$ (line 15). The sampling strategy depends on the ranking procedure and the distance measure used. We shall describe its concrete implementations in subsequent sections.

We refer to our algorithm as $\mathrm{RANKEL}_d^{\mathcal{R}}$, depending on which ranking procedure $\mathcal{R}$ (ϵ-insensitive Copeland (CO_ϵ) or sum of expectations (SE_ϵ)) and which distance measure d ($d_{\mathcal{K}}$ or $d_{\mathcal{M}}$) are used. For example, $\mathrm{RANKEL}_{d_{\mathcal{K}}}^{\mathrm{CO}_\epsilon}$ denotes the instance of our algorithm that seeks to find a ranking close to the ϵ-insensitive Copeland order in terms of $d_{\mathcal{K}}$.

Algorithm 1 RANKEL $(Y_{1,1}, \ldots, Y_{K,K}, \rho, \delta, \epsilon)$

1: **for** $i, j = 1 \to K$ **do**   ▷ Initialization
2:     $\bar{y}_{i,j} = 0, n_{i,j} = 0$
3: $A = \{(i,j) | i \neq j, 1 \leq i, j \leq K\}$
4: $t = 0$
5: **repeat**
6:     **for** $(i,j) \in A$ **do**
7:         $y \sim Y_{i,j}$ ▷ Draw a random sample
8:         $n_{i,j} = n_{i,j} + 1$
9: ▷ Keep track the number of samples drawn for each $Y_{i,j}$
10:        Update $\bar{y}_{i,j}$ with y
11:                          ▷ $\bar{\mathbf{Y}} = [\bar{y}_{i,j}]_{K \times K} \approx \mathbf{Y} = [y_{i,j}]_{K \times K}$
12:    $t = t + 1$
13:    $A = \mathrm{SAMPLINGSTRATEGY}(\bar{\mathbf{Y}}, \mathbf{N}, \delta, \epsilon, t, \rho)$
14: **until** $0 < |A|$
15: $\tau = \mathrm{GETESTIMATEDRANKING}(\bar{\mathbf{Y}}, \mathbf{N}, \delta, \epsilon, t)$ ▷
    Calculate a ranking based on $\bar{\mathbf{Y}}$ by using $\mathcal{R}$
16: **return** τ

Sampling strategies

The case of ϵ-insensitive Copeland

In the following, we denote the estimate of $y_{i,j} = \mathbb{E}(Y_{i,j})$ at time step t by $\bar{y}_{i,j}^t$, and the number of samples taken from $Y_{i,j}$ up to that time step by $n_{i,j}^t$ (omitting the time index if not needed). We start the description of our sampling strategy by determining reasonable confidence intervals for the $\bar{y}_{i,j}^t$ values.[4]

Lemma 1. *For any sampling strategy in line 13 of Algorithm 1,* $\sum_{i=1}^{K} \sum_{j \neq i} \sum_{t=1}^{\infty} \mathbb{P}(A_{i,j}^t) \leq \delta$, *where* $A_{i,j}^t = \{y_{i,j} \notin [\bar{y}_{i,j}^t - c(n_{i,j}^t, t, \delta), \bar{y}_{i,j}^t + c(n_{i,j}^t, t, \delta)]\}$ *with* $c(n, t, \delta) = \sqrt{\frac{1}{2n} \ln\left(\frac{5K^2 t^4}{4\delta}\right)}$.

From now on, we will concisely write $c_{i,j}^t$ for $c(n_{i,j}^t, t, \delta)$ and $C_{i,j}^t$ for the confidence interval $[\bar{y}_{i,j}^t - c_{i,j}^t, \bar{y}_{i,j}^t + c_{i,j}^t]$. Now, one can calculate a lower bound of d_i^* based on $\bar{\mathbf{Y}}^t$ and $\mathbf{N}^t$. First, let us define $d_i^t = \#D_i^t$, where

$$D_i^t = \left\{ j \,|\, 1/2 - \epsilon < \bar{y}_{i,j}^t - c_{i,j}^t, j \neq i \right\} \ .$$

Put in words, d_i^t denotes the number of options that are already known to be beaten by o_i. Similarly, we define the number of "undecided" pairwise preferences for an option o_i as $u_i^t = \#U_i^t$, where

$$U_i^t = \left\{ j \,|\, [1/2 - \epsilon, 1/2 + \epsilon] \subseteq C_{i,j}^t, j \neq i \right\} \ .$$

Based on d_i^t and u_i^t, we define a ranking τ^t over $\mathcal{O}$ by sorting the options o_i in increasing order according to d_i^t, and in case of a tie ($d_i^t = d_j^t$) according to the sum $d_i^t + u_i^t$. The following corollary upper-bounds the NDP and MRP distances between τ^t and the underlying order $\prec^{\mathrm{CO}_\epsilon}$ based on only empirical estimates.

Corollary 2. *Using the notation introduced above, let*

$$\mathbb{I}_{i,j}^t = \mathbb{I}\left\{ (d_i^t < d_j^t + u_j^t) \wedge (d_j^t < d_i^t + u_i^t) \right\}$$

---

[4] Due to space limitations, all proofs are omitted.

for all $1 \leq i \neq j \leq K$. Then for any time step t, and for any sampling strategy, $d_{\mathcal{K}}(\tau^t, \prec^{\text{CO}_\epsilon}) \leq \frac{1}{2}\sum_{i=1}^{K}\sum_{j\neq i}\mathbb{I}_{i,j}^t$ holds with probability at least $1 - \delta$, and $d_{\mathcal{M}}(\tau^t, \prec^{\text{CO}_\epsilon}) \leq \max_{1\leq i\leq K}\sum_{j\neq i}\mathbb{I}_{i,j}^t$ holds again with probability at least $1 - \delta$.

Corollary 2 implies that sampling can be stopped as soon as

$$\sum_{i=1}^{K}\sum_{j\neq i}\mathbb{I}_{i,j}^t < \rho \quad \text{and} \quad \max_{1\leq i\leq K}\sum_{j\neq i}\mathbb{I}_{i,j}^t < \rho \qquad (2)$$

in the case of NDP and MRD, respectively. Moreover, it suggests a simple greedy strategy for sampling, namely to sample those pairwise preferences that promise a maximal decrease of the respective upper bound in (2). For NDP, this comes down to sampling all undecided pairs of objects ($\cup_i U_i^t$), although this strategy can still be improved: If the rank of an object o_i can be determined based on the samples seen so far ($\mathbb{I}_{i,j}^t = 0$ for all $j \in [K]$), then there is no need to sample any more pairwise preference involving o_i. Formally, the set of object pairs to be sampled can thus be written

$$\widetilde{A}_{\mathcal{K}}^t = \left\{ (i,j) \,|\, (j \in U_i^t) \wedge \exists j' : (\mathbb{I}_{i,j'}^t = 1) \right\}.$$

Further considering the stopping rule in (2), the pairwise preferences to be sampled by $\text{RANKEL}_{d_{\mathcal{K}}}^{\text{CO}_\epsilon}$ in iteration t is given by

$$A_{\mathcal{K}}^t = \begin{cases} \widetilde{A}_{\mathcal{K}}^t & \text{if } \rho \leq \sum_{i=1}^{K}\sum_{j\neq i}\mathbb{I}_{i,j}^t \\ \emptyset & \text{otherwise} \end{cases}. \qquad (3)$$

In the case of the MRD distance, the goal is to decrease the upper bound on $d_{\mathcal{M}}(\tau^t, \prec^{\text{CO}})$. Correspondingly, the greedy strategy samples the set of pairs

$$\widetilde{A}_{\mathcal{M}}^t = \left\{ (i,j) \,|\, (j \in U_i^t) \wedge \rho \leq \sum_{j'\neq i}\mathbb{I}_{i,j'}^t \right\}.$$

Thus, again considering the stopping rule in (2), we can formally write the set of pairs to be sampled by $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$ in iteration t as follows:

$$A_{\mathcal{M}}^t = \begin{cases} \widetilde{A}_{\mathcal{M}}^t & \text{if } \rho \leq \max_{1\leq i\leq K}\sum_{j\neq i}\mathbb{I}_{i,j}^t \\ \emptyset & \text{otherwise} \end{cases} \qquad (4)$$

As a last step, the RANKEL algorithm calls a subroutine to calculate the estimated ranking. According to Corollary 2, τ^t is a suitable choice, because its distance to $\prec^{\text{CO}_\epsilon}$ is smaller than ρ with probability at least $1 - \delta$.

The case of ϵ-insensitive sum of expectations

The SE ranking procedure assigns a real number $y_i = \frac{1}{K-1}\sum_{k\neq i} y_{i,k}$ to every option o_i. Based on the pairwise estimates $\bar{y}_{i,1}^t, \ldots, \bar{y}_{i,K}^t$, an estimate for y_i can simply be obtained as $\bar{y}_i^t = \frac{1}{K-1}\sum_{k\neq i}\bar{y}_{i,k}^t$. Similarly to Lemma 1, one can determine a reasonable confidence interval for the $\bar{y}_i^t$ values.

Lemma 3. *Let $c(n, t, \delta)$ be the function defined in Lemma 1. Then, for any sampling strategy in line 13 of Algorithm 1 that ensures $n_{i,1}^t = n_{i,2}^t = \cdots = n_{i,K}^t$ for any $1 \leq i \leq K$, it holds that $\sum_{i=1}^{K}\sum_{t=1}^{\infty}\mathbb{P}(B_i^t) \leq \delta$, where $B_i^t = \left\{ y_i \notin [\bar{y}_i^t - c(n_i^t, t, \delta), \bar{y}_i^t + c(n_i^t, t, \delta)] \right\}$ and $n_i^t = \sum_{k\neq i} n_{i,k}^t$.*

From now on, we will concisely write c_i^t for $c(n_i^t, t, \delta)$ and C_i^t for the confidence interval $[\bar{y}_i^t - c_i^t, \bar{y}_i^t + c_i^t]$. Given the above estimates, the most natural way to define a ranking σ^t on $\mathcal{O}$ is to sort the options o_i in increasing order according to their scores $\bar{y}_i^t$ (again breaking ties at random). The following corollary upper-bounds the rank distances between σ^t thus defined and $\prec^{\text{SE}_\epsilon}$ in terms of the overlapping confidence intervals of $\bar{y}_1^t, \ldots, \bar{y}_K^t$.

Corollary 4. *Under the condition of Lemma 3, $d_{\mathcal{K}}(\sigma^t, \prec^{\text{SE}_\epsilon}) \leq \frac{1}{2}\sum_{i=1}^{K}\sum_{j\neq i}\mathbb{O}_{i,j}^t$ holds with probability at least $1 - \delta$ for any time step t, where $\mathbb{O}_{i,j}^t = \mathbb{I}\left\{ |C_i^t \cap C_j^t| > \epsilon \right\}$ indicates that the confidence intervals of $\bar{y}_i^t$ and $\bar{y}_j^t$ are overlapping by more than ϵ. Moreover, $d_{\mathcal{M}}\left(\sigma^t, \prec^{\text{SE}_\epsilon}\right) \leq \max_{1\leq i\leq K}\sum_{j\neq i}\mathbb{O}_{i,j}^t$ is again valid with probability at least $1 - \delta$.*

Based on Corollary 4, one can devise greedy sampling strategies that gradually decrease the upper bound of the distances between the current ranking and $\prec^{\text{SE}_\epsilon}$ with respect to $d_{\mathcal{K}}$ or $d_{\mathcal{M}}$, similar to the one described in the previous section for ϵ-sensitive Copeland procedure.

The ranking eventually returned by RANKEL (Algorithm 1, line 15) is simply the one introduced above, namely the permutation that sorts the options o_i according to their scores $\bar{y}_i$.

Complexity analysis

From Propositions 2 and 4, it is immediate that all instantiations of our RANKEL algorithm ($\text{RANKEL}_{d_{\mathcal{K}}}^{\text{CO}_\epsilon}$, $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$, $\text{RANKEL}_{d_{\mathcal{K}}}^{\text{SE}_\epsilon}$, $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{SE}_\epsilon}$) are correct, and hence they are all (ρ, δ)-PAC rank elicitation algorithms. In this section, we analyze $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$ and calculate an upper bound for its expected sample complexity. In our preference-based setup, the sample complexity of an algorithm is the expected number of pairwise comparisons drawn for a given instance of the rank elicitation problem.

The technique we shall use for analyzing $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$ can be applied for $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{SE}_\epsilon}$, too. It cannot be used, however, to characterize the complexity of the rank elicitation task in the case of the $d_{\mathcal{K}}$ distance (see Lemma 6), whence we leave the analysis of $\text{RANKEL}_{d_{\mathcal{K}}}^{\text{CO}_\epsilon}$ and $\text{RANKEL}_{d_{\mathcal{K}}}^{\text{SE}_\epsilon}$ as an open problem.

Expected sample complexity of $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$

Step 1: The following lemma upper-bounds the probability of an estimate $\bar{y}_{i,j}^t$ being significantly bigger than $1/2$ while $y_{i,j} < 1/2$ and vice versa. More specifically, it shows that the error probability decreases with the number of iterations

t as fast as $O(1/t^3)$, a fact that will be useful in our sample complexity analysis later on.

Lemma 5. *Let $\mathcal{E}_{i,j}^t$ denote the event that either $\bar{y}_{i,j}^t - c_{i,j}^t > 1/2 - \epsilon$ and $y_{i,j} < 1/2 - \epsilon$ or $\bar{y}_{i,j}^t + c_{i,j}^t < 1/2 + \epsilon$ and $y_{i,j} > 1/2 + \epsilon$. Then $\textsc{RankEl}_d^{\mathrm{CO}_\epsilon}$ satisfies $\sum_{i=1}^K \sum_{j \neq i} \mathbb{P}\left[\mathcal{E}_{i,j}^t\right] < \frac{4\delta}{5t^3}$.*

Step 2: An interesting property of our problem setting, which distinguishes it from related ones such as top-k and best arm identification, is that it does not only incorporate an ϵ-tolerance on the level of pairwise probability estimates ($y_{i,j}$ values), but also relaxes the required accuracy of the solution along another dimension, namely the proximity of the predicted ranking and the target order. More precisely, the algorithm receives a parameter ρ, and has to guarantee with high confidence that the ranking τ it outputs is at most of distance ρ from some ranking in $\mathcal{L}^{\prec_{\mathbf{Y}}^{\mathrm{CO}_\epsilon}}$.

Unfortunately, one cannot directly determine the smallest distance between a given τ and $\mathcal{L}^{\prec_{\mathbf{Y}}^{\mathrm{CO}_\epsilon}}$ without knowing the entries of $\mathbf{Y}$ with high accuracy. Instead, an indirect method has to be used in order to bound the sample complexity. To this end, denote by $(\mathbf{Y})_r$ the set of matrices that are obtained from $\mathbf{Y}$ as follows

$$(\mathbf{Y})_r = \{\tilde{\mathbf{Y}} \mid \tilde{y}_{i,j} < 1/2 \text{ if } y_{i,j} < 1/2 - \epsilon \text{ and}$$
$$\tilde{y}_{i,j} > 1/2 \text{ if } y_{i,j} > 1/2 + \epsilon \text{ where}$$
$$(i,j) \in A' \subset A, |A \setminus A'| = r\}$$

where $A = \{(i,j) \mid i \neq j, 1 \leq i, j \leq K\}$ is the set of all off-diagonal index pairs.

Now, if all but at most r entries in $\bar{\mathbf{Y}}^t$ are known to be either bigger than $1/2 + \epsilon$ or smaller than $1/2 - \epsilon$ with sufficiently high confidence (i.e., if all but at most r pairs (i,j) satisfy $j \notin U_i^t$), then $\bar{\mathbf{Y}}^t \in (\mathbf{Y})_r$ with high probability. Moreover, note that no algorithm can safely terminate as long as no ranking τ exists that satisfies both that it is consistent with the current information (i.e., $\tau \in \mathcal{L}^{\prec_{\bar{Y}}^{\mathrm{CO}_\epsilon}}$), and that it is of distance at most ρ from any possible strict order—that is formally

$$\max_{\mathbf{Y}': \bar{\mathbf{Y}}^t \in (\mathbf{Y}')_r} d_{\mathcal{M}}(\tau, \prec_{\mathbf{Y}'}^{\mathrm{CO}_\epsilon}) \leq \rho .$$

Accordingly, one should define the variation of distance $d_{\mathcal{M}}$ around $\mathbf{Y}$ at radius r as

$$v_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}(r, \mathbf{Y}) = \max_{\tilde{\mathbf{Y}} \in (\mathbf{Y})_r} \min_{\tau \in \mathcal{L}^{\prec_{\tilde{Y}}^{\mathrm{CO}_\epsilon}}} \max_{\mathbf{Y}': \tilde{\mathbf{Y}} \in (\mathbf{Y}')_r} d_{\mathcal{M}}\left(\tau, \prec_{\mathbf{Y}'}^{\mathrm{CO}_\epsilon}\right)$$

The next result shows that the ranking output by $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ is always within this distance ($v_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}(r, \mathbf{Y})$) and thus, it is indeed a reasonable definition.

Lemma 6. *Assume that $A^t = \cap_{i=1}^K \cap_{j \neq i} A_{i,j}^t$ holds, where $A_{i,j}^t$ denotes the event defined in Lemma 1. Let τ denote some ranking that satisfies $\tau_i > \tau_j$ whenever $(d_i^t < d_j^t)$ or $(d_i^t = d_j^t) \wedge (d_i^t + u_i^t < d_j^t + u_j^t)$ holds for some $t > 0$. Then $d_{\mathcal{M}}(\tau, \prec^{\mathrm{CO}_\epsilon}) \leq \max_i I_i^t$, where $I_i^t = \sum_{j \neq i} \mathbb{I}_{i,j}^t = \#\{j : (d_i^t < d_j^t + u_j^t) \wedge (d_j^t < d_i^t + u_i^t)\}$. Moreover,*

$\max_i I_i^t \leq 2v_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}(r^t, \mathbf{Y})$, where $r^t = \sum_{i=1}^k |U_i^t|$ is the number of pairwise preferences which cannot yet be decided with high probability.

Remark 7. *Lemma 6 establishes the existence of a fast and easy method for computing the largest MRD distance possible, given some $\bar{\mathbf{Y}}$ and r. Needless to say, having an approximation with similar properties (at least for an approximation of the largest distance) for the NDP measure would be quite desirable. However, as it is not clear how such a result can be obtained (if at all), determining the complexity of this task is left as an open problem.*

Remark 8. *Lemma 6 assumes A^t to hold for a particular $t > 0$. This lemma can be restated so that it holds for any $t > 0$ with probability at least $1 - \delta$, since, according to Lemma 1, $\sum_{i=1}^K \sum_{j \neq i} \sum_{t=1}^\infty \mathbb{P}(A_{i,j}^t) \leq \delta$.*

Step 3: We will use $\Delta_{i,j} = |1/2 - y_{i,j}|$ as a complexity measure of the rank elicitation task. Furthermore, let $\Delta_{(r)}$ denote the r-th smallest value among $\Delta_{i,j}$ for all distinct $i, j \in [K]$. The next lemma upper-bounds (building on Lemma 6) the probability that $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ does not terminate at iteration t.

Lemma 9. *With $A_{\mathcal{M}}^t$ the set of pairs $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ samples in round t, it holds that*

$$\mathbb{P}\left\{A_{\mathcal{M}}^t \neq \emptyset \wedge \forall (i,j) : (\Delta_{i,j} \geq \Delta_{(r_1)}) \Rightarrow (n_{i,j}^t > 2b_{i,j}^t)\right\}$$
$$\leq \frac{3\delta}{10K^2 t^4} \sum_{r=1}^{K^2 - r_1} \frac{1}{\left(\Delta_{(r)} + \epsilon\right)^2} ,$$

where $b_{i,j}^t = \left\lceil \frac{1}{2(\Delta_{i,j} + \epsilon)^2} \ln\left(\frac{5K^2 t^4}{4\delta}\right) \right\rceil$ and $r_1 = 2 \arg\max\left\{r \in [K^2] | v_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}(r, \mathbf{Y}) < \rho\right\}$.

Step 4: Using Lemmas 5 and 9, one can calculate an upper bound for the expected sample complexity of $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$.

Theorem 10. *Using the notation introduced in Lemma 9, the expected sample complexity for $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ is $O\left(R_1 \log\left(\frac{R_1}{\delta}\right)\right)$, where $R_1 = \sum_{r=1}^{K^2 - r_1} \left(\Delta_{(r)} + \epsilon\right)^{-2}$.*

Proof sketch: First, it can be shown that $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ terminates before iteration $T \in O\left(R_1 \log\left(\frac{R_1}{\delta}\right)\right)$ if enough samples are drawn from each $Y_{i,j}$ ($n_{i,j}^t > 2b_{i,j}^t$ according to Lemma 9) and no error occurs for any of the $\bar{y}_{i,j}^t$ (Lemma 5). Consequently, after iteration T, the probability of an error along with the probability of the non-termination of the algorithm (if enough samples are drawn) upper-bounds the number of iterations taken by $\textsc{RankEl}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ after T. This probability can be upper-bounded by $4/3\pi^2 \delta$ for iterations $> T$ based on Lemmas 5 and 9. $\qquad \square$

The expected sample complexity bound given in Theorem 10 is similar in spirit to the one given for LUCB1 in the framework of stochastic multi-armed bandits (Kalyanakrishnan et al. 2012), but the complexity measure of the rank elicitation task is essentially of different nature.

Expected sample complexity of $\text{RANKEL}_{d_\mathcal{M}}^{\text{SE}_\epsilon}$

The sample complexity analysis of $\text{RANKEL}_{d_\mathcal{M}}^{\text{SE}_\epsilon}$ is very similar to the one we carried out for the ϵ-insensitive Copeland ranking, although the complexity measure of the rank elicitation task in this case can be given as follows: let $\lambda_{i,j} = |y_i - y_j|$, and furthermore, let $\lambda_{(r)}$ denote the r-th smallest value among $\lambda_{i,j}$ for all distinct $i, j \in [K]$. Now, the expected sample complexity of $\text{RANKEL}_{d_\mathcal{M}}^{\text{SE}_\epsilon}$ can be upper-bounded in terms of $\Lambda_1 = \sum_{r=1}^{K^2 - \ell_1} \left(\lambda_{(r)} + \epsilon \right)^{-2}$ (similarly to Theorem 10) where $\ell_1 = 2 \arg\max \left\{ r \in [K^2] | v_{d_\mathcal{M}}^{\text{SE}_\epsilon}(r, \mathbf{Y}) < \rho \right\}$. We omit the technical details, since the analysis is straightforward based on the previous section and (Kalyanakrishnan et al. 2012).

Experiments

To illustrate our PAC rank elicitation method, we applied it to sports data, namely the soccer matches of the last ten seasons of the German Bundesliga. Our goal was to learn the corresponding Copeland or SE ranking. We restricted to the 8 teams that participated in each Bundesliga season between 2002 to 2012. Each pair of teams o_i and o_j met 20 times; we denote the outcomes of these matches by $y_{i,j}^1, \ldots, y_{i,j}^{20}$ and take the corresponding frequency distribution as the (ground-truth) probability distribution of $Y_{i,j}$. The matrix $\mathbf{Y}$ thus obtained is shown in Figure 1(a).

As a baseline, we run the RANKEL algorithm with uniform sampling, meaning that all pairwise comparisons are sampled in each iteration. The accuracy of a run is 1 if $d(\tau, \prec^\mathcal{R}) \leq \rho$ for the ranking τ that was produced, and 0 otherwise. The relative empirical sample complexity achieved by RANKEL with respect to the uniform sampling is shown in Table 1(b) for various parameter settings. Our results confirm that RANKEL has a significantly smaller empirical sample complexity than uniform sampling (while providing the same guarantees in terms of approximation quality).

Conclusion and future work

We introduced a PAC rank elicitation problem and proposed an algorithm for solving this task, that is, for eliciting a ranking that is close to the underlying target order with high probability. Our algorithm consistently outperforms the uniform sampling strategy that was taken as a baseline. Moreover, it scales gracefully with the parameters ϵ and ρ that specify, respectively, the strictness of the target order and the sought quality of approximation to that order.

There is still a number of theoretical questions to be addressed in future work, as well as interesting variants of our setting. First, as mentioned in Remark 7, the sample complexity for $\text{RANKEL}_{d_\mathcal{K}}^{\text{SE}_\epsilon}$ and $\text{RANKEL}_{d_\mathcal{K}}^{\text{SE}_\epsilon}$ is still an open question. Second, noting that the $Y_{i,j}$ are trinomial random variables for which a Clopper-Pearson-type high probability confidence bound exists (Chafaï and Concordet 2009), there is hope to significantly improve our bound on expected sample complexity. Third, based on (Kalyanakrishnan et al. 2012), a high probability bound for the sample complexity

B. Munchen	0.5	0.7	0.55	0.575	0.75	0.55	0.775	0.7
B. Dortmund	0.3	0.5	0.55	0.475	0.425	0.525	0.6	0.675
B. Leverkusen	0.45	0.45	0.5	0.425	0.55	0.55	0.65	0.6
VfB Stuttgart	0.425	0.525	0.575	0.5	0.4	0.6	0.5	0.65
Schalke 04	0.25	0.575	0.45	0.6	0.5	0.45	0.65	0.675
W. Bremen	0.45	0.475	0.45	0.4	0.55	0.5	0.55	0.65
VfL Wolfsburg	0.225	0.4	0.35	0.5	0.35	0.45	0.5	0.675
Hannover 96	0.3	0.325	0.4	0.35	0.325	0.35	0.325	0.5

[7,7]	[4,7]	[0.73,0.75]	[0.73,0.83]
[4,4]	[1,6]	[0.58,0.60]	[0.58,0.68]
[4,4]	[1,7]	[0.60,0.62]	[0.60,0.70]
[4,5]	[1,7]	[0.60,0.62]	[0.60,0.70]
[4,4]	[2,6]	[0.59,0.61]	[0.59,0.69]
[3,3]	[1,7]	[0.58,0.60]	[0.58,0.68]
[1,2]	[1,4]	[0.49,0.51]	[0.49,0.59]
[0,0]	[0,1]	[0.41,0.43]	[0.41,0.51]

(a) Matrix $\mathbf{Y}$ for Bundesliga data, and the intervals for the interval orders $\prec^{\text{CO}_{0.02}}$, $\prec^{\text{CO}_{0.1}}$, $\prec^{\text{SE}_{0.02}}$ and $\prec^{\text{SE}_{0.1}}$, respectively

$\prec^*$	$d(.,.)$	ρ	ϵ	Improvement (%)
$\prec^{\text{CO}}$	$d_\mathcal{K}$	3	0.02	25.3 ± 0.4
$\prec^{\text{CO}}$	$d_\mathcal{M}$	3	0.02	24.0 ± 0.4
$\prec^{\text{SE}}$	$d_\mathcal{K}$	3	0.02	21.9 ± 0.2
$\prec^{\text{SE}}$	$d_\mathcal{M}$	3	0.02	23.1 ± 0.2
$\prec^{\text{CO}}$	$d_\mathcal{K}$	3	0.1	43.6 ± 0.7
$\prec^{\text{CO}}$	$d_\mathcal{M}$	3	0.1	43.9 ± 0.7
$\prec^{\text{SE}}$	$d_\mathcal{K}$	3	0.1	24.7 ± 0.1
$\prec^{\text{SE}}$	$d_\mathcal{M}$	3	0.1	23.5 ± 0.2
$\prec^{\text{CO}}$	$d_\mathcal{K}$	5	0.1	49.1 ± 0.6
$\prec^{\text{CO}}$	$d_\mathcal{M}$	5	0.1	64.3 ± 0.8
$\prec^{\text{SE}}$	$d_\mathcal{K}$	5	0.1	25.4 ± 0.2
$\prec^{\text{SE}}$	$d_\mathcal{M}$	5	0.1	31.8 ± 0.4

(b) Improvement in empirical sample complexity

Figure 1: The top panel (1(a)) shows the matrix $\mathbf{Y}$ for the Bundesliga data, and the $[d_i^*, d_i^* + s_i^*]$ intervals for $\prec^{\text{CO}_{0.02}}$ and $\prec^{\text{CO}_{0.1}}$, and the $[y_i, y_i + \epsilon]$ intervals for $\prec^{\text{SE}_{0.02}}$ and $\prec^{\text{SE}_{0.1}}$, respectively. The bottom panel (1(b)) shows the reduction of the empirical sample complexity achieved by RANKEL for various parameter settings, taking the complexity of uniform sampling as 100%. Mean and standard deviation of the improvement were obtained by averaging over 100 repetitions. The confidence parameter δ was set to 0.1 for each run; accordingly, the average accuracy was significantly above $1 - \delta = 0.9$ in each case.

might be devised instead of the expected complexity bound. Last but not least, there are other interesting ranking procedures $\mathcal{R}$ and distance measures that can be used to instantiate our setting.

Acknowledgments

References

Auer, P.; Cesa-Bianchi, N.; and Fischer, P. 2002. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* 47:235–256.

Braverman, M., and Mossel, E. 2008. Noisy sorting without resampling. In *Proceedings of the nineteenth annual ACM-SIAM Symposium on Discrete algorithms*, 268–276.

Braverman, M., and Mossel, E. 2009. Sorting from noisy information. *CoRR* abs/0910.1191.

Bubeck, S.; Munos, R.; and Stoltz, G. 2009. Pure exploration in multi-armed bandits problems. In *Proceedings*

of the 20th international conference on Algorithmic learning theory, ALT'09, 23–37. Berlin, Heidelberg: Springer-Verlag.

Bubeck, S.; Wang, T.; and Viswanathan, N. 2013. Multiple identifications in multi-armed bandits. In *Proceedings of The 30th International Conference on Machine Learning*, 258–265.

Busa-Fekete, R.; Szörényi, B.; Weng, P.; Cheng, W.; and Hüllermeier, E. 2013. Top-k selection based on adaptive sampling of noisy preferences. In *Proceedings of the 30th International Conference on Machine Learning, JMLR W&CP*, volume 28.

Cappé, O.; Garivier, A.; Maillard, O.-A.; Munos, R.; and Stoltz, G. 2012. Kullback-Leibler upper confidence bounds for optimal sequential allocation. *Submitted to the Annals of Statistics*.

Chafaï, D., and Concordet, D. 2009. Confidence regions for the multinomial parameter with small sample size. *Journal of the American Statistical Association* 104(487):1071–1079.

Chen, X.; Bennett, P. N.; Collins-Thompson, K.; and Horvitz, E. 2013. Pairwise ranking aggregation in a crowdsourced setting. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 193–202.

Eriksson, B. 2013. Learning to Top-K search using pairwise comparisons. *Journal of Machine Learning Research - Proceedings Track* 31:265–273.

Even-Dar, E.; Mannor, S.; and Mansour, Y. 2002. PAC bounds for multi-armed bandit and markov decision processes. In *Proceedings of the 15th Annual Conference on Computational Learning Theory*, 255–270.

Feige, U.; Raghavan, P.; Peleg, D.; and Upfal, E. 1994. Computing with noisy information. *SIAM J. Comput.* 23(5):1001–1018.

Gabillon, V.; Ghavamzadeh, M.; Lazaric, A.; and Bubeck, S. 2011. Multi-bandit best arm identification. In Shawe-Taylor, J.; Zemel, R.; Bartlett, P.; Pereira, F.; and Weinberger, K., eds., *Advances in Neural Information Processing Systems 24*. MIT. 2222–2230.

Guo, S.; Sanner, S.; Graepel, T.; and Buntine, W. 2012. Score-based bayesian skill learning. In *European Conference on Machine Learning*, 1–16.

Hoeffding, W. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58:13–30.

Kalyanakrishnan, S.; Tewari, A.; Auer, P.; and Stone, P. 2012. Pac subset selection in stochastic multi-armed bandits. In *Proceedings of the Twenty-ninth International Conference on Machine Learning (ICML 2012)*, 655–662.

Kalyanakrishnan, S. 2011. *Learning Methods for Sequential Decision Making with Imperfect Representations*. Ph.D. Dissertation, University of Texas at Austin.

Kendall, M. 1955. *Rank correlation methods*. London: Charles Griffin.

Moulin, H. 1988. *Axioms of cooperative decision making*. Cambridge University Press.

Urvoy, T.; Clerot, F.; Féraud, R.; and Naamane, S. 2013. Generic exploration and K-armed voting bandits. In *Proceedings of the 30th International Conference on Machine Learning, JMLR W&CP*, volume 28, 91–99.

Yue, Y.; Broder, J.; Kleinberg, R.; and Joachims, T. 2012. The k-armed dueling bandits problem. *Journal of Computer and System Sciences* 78(5):1538–1556.

Supplementary material for "PAC Rank Elicitation through Adaptive Sampling of Stochastic Pairwise Preferences"

Proof of Lemma 1

Proof. Regardless of the sampling strategy used, it obviously holds that $1 \leq n_{i,j}^t \leq t$ for any $t > 0$. Then by applying the Hoeffding bound (Hoeffding 1963), we have

$$\sum_{t=1}^{\infty} \mathbb{P}(A_{i,j}^t) \leq \sum_{t=1}^{\infty} \sum_{n=1}^{t} 2\exp(-2nc(n,t,\delta)^2) \leq \sum_{t=1}^{\infty} \frac{8\delta}{5K^2 t^3} < \frac{\delta}{K^2}$$

Summing up $\sum_{t=1}^{\infty} \mathbb{P}(A_{i,j}^t)$ for all pairs of objects, we obtain $\sum_{i=1}^{K} \sum_{j \neq i} \sum_{t=1}^{\infty} \mathbb{P}(A_{i,j}^t) \leq \delta$. □

Pseudo-codes for ϵ-Copeland's ranking procedure ($\text{RANKEL}_{d_{\mathcal{K}}}^{\text{CO}_\epsilon}$ and $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$)

The pseudo-code of the sampling strategy for ϵ-Copeland's ranking procedure

---

Procedure 2 $\text{SAMPLINGSTRATEGY}(\bar{\mathbf{Y}}, \mathbf{N}, \delta, \epsilon, t, \rho)$

1: ▷ Compute the confidence bounds
2: **for** $i, j = 1 \to K$ **do**
3: **if** $(i \neq j)$ **then**
4:   $c_{i,j} = c(n_{i,j}, t, \delta) = \sqrt{\frac{1}{2n_{i,j}} \ln\left(\frac{5K^2 t^4}{4\delta}\right)}$
5: ▷ Compute the d_i and u_i values for determining the order which estimates the target order with high probability
6: **for** $i = 1 \to K$ **do**
7: $d_i = 0, u_i = 0$
8: **for** $j = 1 \to K$ **do**
9:   **if** $(i \neq j)$ **then**   ▷ $C_{i,j} = [\bar{y}_{i,j} - c_{i,j}, \bar{y}_{i,j} + c_{i,j}]$
10:    $u_{i,j} = \mathbb{I}\{[1/2 - \epsilon, 1/2 + \epsilon] \subseteq C_{i,j}\}$ ▷ If $u_{i,j} = 0$ then $j \notin U_i$
11:    $u_i = u_i + u_{i,j}$
12:    **if** $1/2 - \epsilon < \bar{y}_{i,j} - c_{i,j}$ **then**
13:     $d_i = d_i + 1$
14: ▷ Compute upper bounds for the NDP, resp. the MDP distances
15: **for** $i = 1 \to K$ **do**
16: $I_i = 0$
17: **for** $j = 1 \to K$ **do**
18:   **if** $(i \neq j)$ **then**
19:    **if** $(d_i < d_j + u_j) \wedge (d_j < d_i + u_i)$ **then**
20:     $I_i = I_i + 1$
21: ▷ Determine the pairs that need to be sampled in the current round:
22: $\rho' = \begin{cases} \frac{1}{2} \sum_{i=1}^{K} I_i & \text{For NDP distance } d_{\mathcal{K}}(\cdot, \cdot) \\ \max_{1 \leq i \leq K} I_i & \text{For MRD distance } d_{\mathcal{M}}(\cdot, \cdot) \end{cases}$
23: $A = \emptyset$
24: **if** $\rho \leq \rho'$ **then**
25: **for** $i = 1 \to K$ **do**
26:   **if** $I_i > \begin{cases} 0 & \text{For NDP distance } d_{\mathcal{K}}(\cdot, \cdot) \\ \rho & \text{For MRD distance } d_{\mathcal{M}}(\cdot, \cdot) \end{cases}$ **then**
27:    **for** $j = 1 \to K$ **do**
28:     **if** $(i \neq j) \wedge (u_{i,j} = 1)$ **then**
29:      $A = A \cup (i, j)$
30: **return** A

Procedure for estimating ϵ-Copeland's ranking

Procedure 3 GETESTIMATEDRANKING ($\bar{\mathbf{Y}}, \mathbf{N}, \delta, \epsilon, t$)

1: ▷ Compute the confidence bounds
2: **for** $i, j = 1 \rightarrow K$ **do**
3:     **if** $(i \neq j)$ **then**
4:         $c_{i,j} = c(n_{i,j}, t, \delta) = \sqrt{\frac{1}{2n_{i,j}} \ln\left(\frac{5K^2 t^4}{4\delta}\right)}$
5: ▷ Compute the d_i and u_i values for determining the order which estimates the target order with high probability
6: **for** $i = 1 \rightarrow K$ **do**
7:     $d_i = 0, u_i = 0$
8:     **for** $j = 1 \rightarrow K$ **do**
9:         **if** $(i \neq j)$ **then**
10:             $u_{i,j} = \mathbb{I}\{[1/2 - \epsilon, 1/2 + \epsilon] \subseteq C_{i,j}\}$ ▷ If $u_{i,j} = 0$ then $j \notin U_i$
11:             $u_i = u_i + u_{i,j}$
12:             **if** $1/2 - \epsilon < \bar{y}_{i,j} - c_{i,j}$ **then**
13:                 $d_i = d_i + 1$
14: ▷ Sort the options based on the following relation:
15:                               $\tau_i > \tau_j \Leftrightarrow (d_i < d_j)$ or $(d_i = d_j) \wedge (d_i + u_i < d_j + u_j)$
16: $\tau = \text{SORT}((d_1, d_1 + u_1), \cdots, (d_K, d_K + u_K))$
17: **return** τ

Proof of Corollary 2

Proof. First, note that, for any t, $d_i^* \leq d_i^t + u_i^t$, because d_i^t denotes the number of options for which $1/2 - \epsilon < \bar{y}_{i,j}^t - c_{i,j}^t$, and u_i^t denotes the number of "undecided" options (not known yet that either $1/2 - \epsilon < \bar{y}_{i,j}^t - c_{i,j}^t$ or $1/2 + \epsilon > \bar{y}_{i,j}^t + c_{i,j}^t$). Second, it can be seen that $d_i^t \leq d_i^* + s_i^*$, since $1/2 - \epsilon < \bar{y}_{i,j}^t - c_{i,j}^t$ implies that $1/2 - \epsilon < y_{i,j}$. Now, assume that for a pair of options o_i and o_j, we have

$$d_j^* \leq d_j^t + u_j^t < d_i^t \leq d_i^* + s_i^*$$

which implies that $o_i \not\prec^{\text{CO}_\epsilon} o_j$. Consequently, if $\mathbb{I}\left\{(d_j^t < d_i^t + u_i^t) \wedge (d_i^t < d_j^t + u_j^t)\right\} = 0$, then we can decide whether $o_i \not\prec^{\text{CO}_\epsilon} o_j$ or $o_j \not\prec^{\text{CO}_\epsilon} o_i$. In addition, $d_j^t + u_j^t < d_i^t$ readily implies that $d_j^t < d_i^t$, therefore the ordering τ^t defined in the claim is valid in a sense that if $d_j^t < d_i^t$ (which implies that $o_i \not\prec^{\text{CO}_\epsilon} o_j$) then by definition of τ^t, $\tau_j^t > \tau_i^t$. Then simple calculation yields the bounds given in the corollary. The high probability bound follows from Lemma 1 that upper bounds the probability of a $\bar{y}_{i,j}^t$ violates the confidence interval at any time t and for any pairs of objects. $\qquad\square$

Proof of Lemma 3

Proof. Recall that

$$B_i^t = \left\{ y_i \notin \left[\bar{y}_i^t - c\left(n_i^t, t, \delta\right), \bar{y}_i^t - c\left(n_i^t, t, \delta\right)\right] \right\}$$

where $n_i^t = \sum_{j \neq i} n_{i,j}^t$ and $c(n, t, \delta)$ defined in Lemma 1. First, let us note that if $n_{i,1}^t = n_{i,2}^t = \cdots = n_{i,K}^t$ for all $1 \leq i \leq K$ and any $t > 0$, then the empirical estimate $\frac{1}{\sum_{j \neq i} n_{i,j}^t} \sum_{j \neq i} \sum_{t'=1}^{n_{i,j}^t} y_{i,j}^{t'}$ is unbiased. Formally, we have

$$\mathbb{E}\left[\frac{1}{\sum_{j \neq i} n_{i,j}^t} \sum_{j \neq i} \sum_{t'=1}^{n_{i,j}^t} y_{i,j}^{t'}\right] = \mathbb{E}\left[\frac{1}{K-1} \sum_{j \neq i} \frac{1}{n_{i,j}^t} \sum_{t'=1}^{n_j} y_{i,j}^{t'}\right] = \frac{1}{K-1} \sum_{j \neq i} y_{i,j} = y_i$$

Second, it obviously holds that $1 \leq n_{i,j}^t \leq t$ for any $t > 0$, and thus, $\bar{y}_i^t$ is sum of at most tK terms. More precisely, it can only consists of $t'K$ terms where $1 \leq t' \leq K$. Now, by applying the Hoeffding bound (Hoeffding 1963), we have

$$\sum_{t=1}^{\infty} \mathbb{P}(B_i^t) \leq \sum_{t'=1}^{\infty} \sum_{t'=1}^{t} \sum_{\ell=1}^{K} 2 \exp(-2t'Kc(t'K, t, \delta)^2) \leq \sum_{\ell=1}^{K} \sum_{t=1}^{\infty} \frac{8\delta}{5K^2 t^3} < \frac{\delta}{K}$$

Summing up $\sum_{t=1}^{\infty} \mathbb{P}(B_i^t)$ for all objects, we obtain $\sum_{i=1}^{K} \sum_{t=1}^{\infty} \mathbb{P}(B_i^t) \leq \delta$. $\qquad\square$

Pseudo-codes for ϵ-Sum of expectations ranking procedure ($\text{RANKEL}_{d_\mathcal{K}}^{\text{SE}_\epsilon}$ and $\text{RANKEL}_{d_\mathcal{M}}^{\text{SE}_\epsilon}$)

The pseudo-code of the sampling strategy for ϵ-Sum of expectations ranking procedure

Figure 4 shows the pseudo-code of the sampling strategy for ϵ-Sum of expectations ranking procedure. First, the estimate for y_i's and their confidence intervals are calculated in line 1 - 4. Then, the indicator values $\mathbb{O}_{i,j}^t$ defined in Proposition 4 for each pair of object o_i and o_j are calculated that indicates if the confidence intervals of $\bar{y}_i^t$ and $\bar{y}_j^t$ are overlapping. Next, based on Proposition 4, we decide to terminate or not (line 5-13), and finally, following the greedy sampling strategy we select the pairs of objects should be compared next (line 14 - 20). Note that this sampling strategy selects all $Y_{i,1}, \ldots, Y_{i,K}$ to be sampled in an iteration for an option o_i or neither of them (see line 18-20), therefore it satisfies the condition of Lemma 3.

Procedure 4 $\text{SAMPLINGSTRATEGY}(\bar{\mathbf{Y}}, \mathbf{N}, \delta, \epsilon, t, \rho)$

1: $\triangleright$ Compute the confidence bounds
2: **for** $i = 1 \to K$ **do**
3:     $\bar{y}_i = \frac{1}{K-1} \sum_{j \neq i} \bar{y}_{i,j}$
4:     $c_i = c(n_i, t, \delta) = \sqrt{\frac{1}{2 \sum_{j \neq i} n_j} \ln\left(\frac{5K^2 t^4}{4\delta}\right)}$
5: $\triangleright$ Compute upper bounds for the NDP, resp. the MDP distances
6: **for** $i = 1 \to K$ **do** $\triangleright$ Calculate $\mathbb{O}_{i,j}^t$
7:     $o_i = 0$
8:     **for** $j = 1 \to K$ **do**
9:        **if** $(i \neq j)$ **then**
10:           **if** $[\bar{y}_i + \epsilon/2 - c_i, \bar{y}_i - \epsilon/2 + c_i] \cap [\bar{y}_j + \epsilon/2 - c_j, \bar{y}_j - \epsilon/2 + c_j] \neq \emptyset$ **then**
11:             $o_i = o_i + 1$
12: $\rho' = \begin{cases} \frac{1}{2} \sum_{i=1}^{K} o_i & \text{For NDP distance } d_\mathcal{K}(\cdot, \cdot) \\ \max_{1 \leq i \leq K} o_i & \text{For MRD distance } d_\mathcal{M}(\cdot, \cdot) \end{cases}$
13: $\triangleright$ Determine the pairs that need to be sampled in the current round:
14: $A = \emptyset$
15: **if** $\rho \leq \rho'$ **then**
16:     **for** $i = 1 \to K$ **do**
17:        **if** $o_i > \begin{cases} 0 & \text{For NDP distance } d_\mathcal{K}(\cdot, \cdot) \\ \rho & \text{For MRD distance } d_\mathcal{M}(\cdot, \cdot) \end{cases}$ **then**
18:           **for** $j = 1 \to K$ **do**
19:             **if** $i \neq j$ **then**
20:                $A = A \cup (i, j)$
    return A

Procedure for estimating the ϵ-Sum of expectations ranking

Procedure 5 GETESTIMATEDRANKING $(\bar{\mathbf{Y}}, \mathbf{N}, \delta, \epsilon, t)$

1: $\bar{y}_i = \sum_{j \neq i} \bar{y}_{i,j}$ $\triangleright$ Initialization
2: $\tau = $ SORT$(\bar{y}_1, \cdots, \bar{y}_K)$ $\triangleright$ Sort the options based on $\bar{y}_i$
3: **return** τ

Proof of Corollary 4

Proof. First, note that if $\bar{y}_i^t + c_i^t - \epsilon < \bar{y}_j^t - c_j^t$, then $o_i \prec_\epsilon^{\text{SE}} o_j$ with high probability, and if $|C_i^t \cap C_j^t| \leq \epsilon$ then the order of o_i and o_j cannot be determined yet with respect $\prec^{\text{SE}_\epsilon}$. Then simple calculation yields the bounds given in the proposition. The high probability bound follows from Lemma 3 that upper bounds the probability of any $\bar{y}_i^t$ violates the confidence interval at any time t and for any pairs of objects. $\square$

Proofs for the expected sample complexity analysis

For the reading convenience we restate all lemmas and theorem.

Proofs of Lemma 5

Lemma 4. *Let $\mathcal{E}_{i,j}^t$ denote the event that either $\bar{y}_{i,j}^t - c_{i,j}^t > 1/2 - \epsilon$ and $y_{i,j} < 1/2 - \epsilon$ or $\bar{y}_{i,j}^t + c_{i,j}^t < 1/2 + \epsilon$ and $y_{i,j} > 1/2 + \epsilon$. Then RANKEL$_d^{\text{CO}_\epsilon}$ satisfies $\sum_{i=1}^K \sum_{j \neq i} \mathbb{P}\left[\mathcal{E}_{i,j}^t\right] < \frac{4\delta}{5t^3}$.*

Proof. Let us fix some i and j first. Assume that $y_{i,j} > 1/2$ (the other case can be proved similarly).

$$
\begin{aligned}
\mathbb{P}\left[\mathcal{E}_{i,j}^t\right] &\leq \mathbb{P}\left[\bar{y}_{i,j}^t - c_{i,j}^t < y_{i,j}\right] \\
&\leq \sum_{n=1}^t \mathbb{P}\left[(\bar{y}_{i,j}^t - c_{i,j}^t < y_{i,j}) \wedge (n_{i,j}^t = n)\right] \\
&= \sum_{n=1}^t \exp\left(-2nc(n,t,\delta)^2\right) = \frac{4\delta}{5K^2t^3} \ .
\end{aligned}
$$

The lemma follows by summing up for all distinct i and j. $\square$

Example for calculating $v_{d_\mathcal{M}}^{\text{CO}_\epsilon}$

Example 11. *Let us consider a rank elicitation task with $\mathbf{Y}$ for which $d_i^* = i$ for every $i \in [K]$. Thus option o_i has rank $K - i + 1$ according to $\prec_\mathbf{Y}^{\text{CO}_0}$. Let us now set $r = 2$, and thereby allow two entries in $\mathbf{Y}$ to alter their order with $1/2$. By the nature of Copeland's ranking procedure, all but two d_i^* can remain the same or one d_i^* can be changed by at most 2. Then it follows easily that $v_{d_\mathcal{M}}^{\text{CO}_0}(2, \mathbf{Y}) = 2$, since $\min_{\tau \in \mathcal{L}^{\prec_{\bar{\mathbf{Y}}}^{\text{CO}_0}}} \max_{\mathbf{Y}':\bar{\mathbf{Y}} \in (\mathbf{Y}')_2} d_\mathcal{M}(\tau, \prec_{\mathbf{Y}'}^{\text{CO}_0})$ is at most 2 for any $\bar{\mathbf{Y}} \in (\mathbf{Y})_2$.*

We can obtain a more difficult rank elicitation task using some $\mathbf{Y}$ with $d_i^* = K - 2$ for $i = 1, \ldots, K - 1$, and $d_K^* = K - 3$. Let us use again $r = 2$, and note that $\tau_K = 1$ for any $\tau \in \mathcal{L}^{\prec_\mathbf{Y}^{\text{CO}_\epsilon}}$, whereas $\bar{\tau}_K = K$ for any $\bar{\tau} \in \mathcal{L}^{\prec_{\bar{\mathbf{Y}}}^{\text{CO}_0}}$ where $\bar{\mathbf{Y}}$ is obtained from $\mathbf{Y}$ by changing it so that o_K is now beaten by every other option. This immediately implies that $v_{d_\mathcal{M}}^{\text{CO}_0}(2, \mathbf{Y}) = K - 1$.

Modifying the last example slightly, and using $\mathbf{Y}$ with $d_i^* = K - 1$ for $i = 2, \ldots, K - 1$, $d_1^* = K - 2$ and $d_K^* = K$, one even obtains $v_{d_\mathcal{K}}^{\text{CO}_0}(2, \mathbf{Y}) = 2(K - 2)$. Note, however that in this case $v_{d_\mathcal{K}}^{\text{CO}_0}(1, \mathbf{Y}) = K - 2$.

Proof of Lemma 6

Lemma 5. *Assume that $A^t = \cap_{i=1}^K \cap_{j \neq i} A_{i,j}^t$ holds, where $A_{i,j}^t$ denotes the event defined in Lemma 1. Let τ denote some ranking that satisfies $\tau_i > \tau_j$ whenever $(d_i^t < d_j^t)$ or $(d_i^t = d_j^t) \wedge (d_i^t + u_i^t < d_j^t + u_j^t)$ holds for some $t > 0$. Then $d_\mathcal{M}(\tau, \prec^{\text{CO}_\epsilon}) \leq \max_i I_i^t$, where $I_i^t = \sum_{j \neq i} \mathbb{I}_{i,j}^t = \#\{j : (d_i^t < d_j^t + u_j^t) \wedge (d_j^t < d_i^t + u_i^t)\}$. It also holds that $\max_i I_i^t \leq 2v_{d_\mathcal{M}}^{\text{CO}_\epsilon}(r^t, \mathbf{Y})$ where $r^t = \sum_{i=1}^k |U_i^t|$ is the number of pairwise preferences which cannot yet be decided with high probability.*

Proof. For convenience, we drop the t indices throughout the proof.

First of all note that, because of our assumption, for each i

$$
d_i \leq d_i^* \leq d_i + u_i \ . \tag{5}
$$

Let τ^* denote the ranking that satisfies $\tau_i^* > \tau_j^*$ whenever $(d_i^* < d_j^*)$ or $(d_i^* = d_j^*) \wedge (\tau_i > \tau_j)$. That is,

$$
\begin{aligned}
K - \tau_i^* &= \#\{j : d_j^* < d_i^*\} + \#\{j : (d_j^* = d_i^*) \wedge (\tau_j > \tau_i)\} \\
&= \#\{j : (d_j^* < d_i^*) \wedge (\tau_j > \tau_i)\} + \#\{j : (d_j^* < d_i^*) \wedge (\tau_j < \tau_i)\} \\
&\quad + \#\{j : (d_j^* = d_i^*) \wedge (\tau_j > \tau_i)\} \ .
\end{aligned}
\tag{6}
$$

By construction, τ^* is a possible target ranking, as it is one of the possible linear extensions of $\prec_{\mathbf{Y}}^{\mathrm{CO}_\epsilon}$. Contrasting (6) with

$$
\begin{aligned}
K - \tau_i &= \#\{j : \tau_j > \tau_i\} \\
&= \#\{j : (d_j^* < d_i^*) \wedge (\tau_j > \tau_i)\} + \#\{j : (d_j^* = d_i^*) \wedge (\tau_j > \tau_i)\} \\
&\quad + \#\{j : (d_j^* > d_i^*) \wedge (\tau_j > \tau_i)\}
\end{aligned}
$$

it is clear that

$$
\begin{aligned}
|\tau_i^* - \tau_i| &= |\#\{j : (d_j^* < d_i^*) \wedge (\tau_j < \tau_i)\} - \#\{j : (d_j^* > d_i^*) \wedge (\tau_j > \tau_i)\}| \\
&\leq \max\left(\#\{j : (d_j^* < d_i^*) \wedge (\tau_j < \tau_i)\} \,,\, \#\{j : (d_j^* > d_i^*) \wedge (\tau_j > \tau_i)\}\right)
\end{aligned}
$$

for any $i = 1, \ldots, K$. Consequently,

$$
d_{\mathcal{M}}(\tau, \prec_{\mathbf{Y}}^{\mathrm{CO}_\epsilon}) \leq d_{\mathcal{M}}(\tau, \tau^*) \leq \max_{i=1,\ldots,K} \max\left(\#\{j : (d_j^* < d_i^*) \wedge (\tau_j < \tau_i)\} \,,\, \#\{j : (d_j^* > d_i^*) \wedge (\tau_j > \tau_i)\}\right)
\tag{7}
$$

To prove the first claim, it thus suffice to show that $\max_i I_i^t$ upper bounds (7). In specific, we show below that $\#\{j : (\tau_i < \tau_j) \wedge (d_i^* < d_j^*)\} + \#\{j : (\tau_i > \tau_j) \wedge (d_i^* > d_j^*)\}$ is upper bounded by I_i^t for any $i = 1, \ldots, K$.

Let us fix some (i, j) satisfying $(\tau_i < \tau_j) \wedge (d_i^* < d_j^*)$ (if no such pair exists, then $\max_i I_i^t$ trivially upper bounds (7)). Below we show that $(d_i < d_j + u_j) \wedge (d_j < d_i + u_i)$ must hold, which, in turn, implies that $\max_i I_i^t$ upper bounds (7). To this end note that, by construction, $\tau_i < \tau_j$ can only hold if either $(d_i > d_j)$ or $(d_i = d_j) \wedge (d_i + u_i \geq d_j + u_j)$. If $d_i > d_j$, then $d_j < d_i + u_i$ automatically, and, because of (5), $d_i \leq d_i^* < d_j^* \leq d_j + u_j$. If, on the other hand, $(d_i = d_j) \wedge (d_i + u_i \geq d_j + u_j)$ holds, then by (5), $d_i \leq d_i^* < d_j^* \leq d_j + u_j$ and $d_j = d_i \leq d_i^* < d_j^* \leq d_j + u_j \leq d_i + u_i$. This completes the proof of the first claim.

To prove the second claim fix some index i satisfying $I_i = \max_j I_j$. Let $\mathbf{Y}'$ (resp. $\mathbf{Y}''$) be the matrix obtained from $\bar{\mathbf{Y}}^t$ by setting all $\bar{y}_{i,j}^t$ values in it with $j \in U_i^t$ to 1 (resp. to 0), and all $\bar{y}_{k,j}^t$ values in $\bar{\mathbf{Y}}$ with $j \in U_k^t$ and $k \neq i$ to 0 (resp. to 1). Note that $\mathbf{Y}', \mathbf{Y}'' \in (\bar{\mathbf{Y}}^t)_{r^t}$ (recall our assumption from the beginning of the proof). Additionally, $\tau'' \in \mathcal{L}^{\prec_{\mathbf{Y}''}^{\mathrm{CO}_\epsilon}}$ it holds that

$$
K - \tau_i' = \#\{j : \tau_i' < \tau_j'\} \geq \#\{j : d_i + u_i > d_j\} \qquad \text{for any } \tau' \in \mathcal{L}^{\prec_{\mathbf{Y}'}^{\mathrm{CO}_\epsilon}}
$$

and

$$
K - \tau_j'' = \#\{j : \tau_i'' < \tau_j''\} \leq \#\{j : d_i \geq d_j + u_j\} \qquad \text{for any } \tau'' \in \mathcal{L}^{\prec_{\mathbf{Y}''}^{\mathrm{CO}_\epsilon}},
$$

implying that

$$
\tau_i' - \tau_i'' \geq \#\{j : d_i + u_i > d_j\} - \#\{j : d_i \geq d_j + u_j\} = \#\{j : (d_i + u_i > d_j) \wedge (d_i \leq d_j + u_j)\} = I_i^t,
$$

and thus also that if $K - \tau_i \leq (\#\{j : d_i + u_i > d_j\} + \#\{j : d_i \geq d_j + u_j\}/2)$ for some $\tau \in \mathcal{L}^{\prec_{\mathbf{Y}}^{\mathrm{CO}_\epsilon}}$ then $d_{\mathcal{M}}(\tau, \prec_{\mathbf{Y}'}^{\mathrm{CO}_\epsilon}) \geq I_i^t/2$, otherwise $d_{\mathcal{M}}(\tau, \prec_{\mathbf{Y}''}^{\mathrm{CO}_\epsilon}) \geq I_i^t/2$. $\qquad\square$

Proof of Lemma 9

Lemma 7. *Denoting by $A_{\mathcal{M}}^t$ the set of pairs* $\mathrm{RANKEL}_{d_{\mathcal{M}}}^{\mathrm{CO}_\epsilon}$ *samples in round t it holds that*

$$
\mathbb{P}\left\{(A_{\mathcal{M}}^t \neq \emptyset) \wedge \left(n_{i,j}^t > 2b_{i,j}^t \text{ for each } (i,j) \text{ with } \Delta_{i,j} \geq \Delta_{(r_1)}\right)\right\} \leq \frac{3\delta}{10K^2 t^4} \sum_{r=1}^{K^2 - r_1} \frac{1}{(\Delta_{(r)} + \epsilon)^2} \ ,
$$

where $b_{i,j}^t = \left\lceil \frac{1}{2(\Delta_{i,j} + \epsilon)^2} \ln\left(\frac{5K^2 t^4}{4\delta}\right) \right\rceil$ and $r_1 = 2 \operatorname{argmax}\left\{r \in [K^2] | v_{d_{\mathcal{M}}}^{\mathrm{CO}}(r) < \rho\right\}$.

Proof. Recall that $(i,j) \in \mathcal{A}_{\mathcal{M}}^t$ iff $j \in U_i^t$ iff $[0.5 - \epsilon, 0.5 + \epsilon] \subseteq C_{i,j}^t$. We shall first prove an upper bound on the probabilities $\mathbb{P}\left\{(n_{i,j}^t > 2b_{i,j}^t) \wedge (j \in U_i^t)\right\}$ for each $i \neq j$. To this end, fix some arbitrary indices $i \neq j$. Assume without loss of generality

that $y_{i,j} > 1/2$. (The other case is handled the same way.) Then

$$\mathbb{P}\left\{\left(n_{i,j}^t > 2b_{i,j}^t\right) \wedge \left(j \in U_i^t\right)\right\} \leq \mathbb{P}\left\{\left(n_{i,j}^t > 2b_{i,j}^t\right) \wedge \left(0.5 - \epsilon > \bar{y}_{i,j}^t - c_{i,j}^t\right)\right\}$$

$$= \mathbb{P}\left\{\left(n_{i,j}^t > 2b_{i,j}^t\right) \wedge \left(y_{i,j} - \bar{y}_{i,j}^t > y_{i,j} - 0.5 + \epsilon - c_{i,j}^t\right)\right\}$$

$$\leq \mathbb{P}\left\{\left(n_{i,j}^t > 2b_{i,j}^t\right) \wedge \left(y_{i,j} - \bar{y}_{i,j}^t > \Delta_{i,j} + \epsilon - c_{i,j}^t\right)\right\}$$

$$\leq \sum_{n=2b_{i,j}^t+1}^{\infty} \mathbb{P}\left\{\left(n_{i,j}^t = n > 2b_{i,j}^t\right) \wedge \left(y_{i,j} - \bar{y}_{i,j}^t > \Delta_{i,j} + \epsilon - c_{i,j}^t\right)\right\}$$

$$\leq \sum_{n=2b_{i,j}^t+1}^{\infty} \exp\left(-2n\left(\Delta_{i,j} + \epsilon - c_{i,j}^t\right)^2\right) \tag{8}$$

$$= \sum_{n=2b_{i,j}^t+1}^{\infty} \exp\left(-2n\left(\Delta_{i,j} + \epsilon - \sqrt{\frac{1}{2n}\ln\left(\frac{5K^2t^4}{4\delta}\right)}\right)^2\right)$$

$$= \sum_{n=2b_{i,j}^t+1}^{\infty} \exp\left(-2(\Delta_{i,j} + \epsilon)^2\left(\sqrt{n} - \sqrt{\frac{1}{2(\Delta_{i,j} + \epsilon)^2}\ln\left(\frac{5K^2t^4}{4\delta}\right)}\right)^2\right)$$

$$\leq \sum_{n=2b_{i,j}^t+1}^{\infty} \exp\left(-2(\Delta_{i,j} + \epsilon)^2\left(\sqrt{n} - \sqrt{b_{i,j}^t}\right)^2\right)$$

$$\leq \cdots \leq \frac{3\delta}{10K^2t^4(\Delta_{i,j} + \epsilon)^2} \tag{9}$$

Here (8) follows from the Hoeffding bound, and the derivation of (9) is similar to Appendix B.2 from (Kalyanakrishnan 2011). (For $y_{i,j} < 1/2$, one can obtain the same bound.)

According to Lemma 6 and the definition of r_1, $\sum_{i=1}^K \sum_{j \neq i} \mathbb{I}_{i,j}^t \geq \rho$ whenever all but at most r_1 entries are recovered correctly. Recalling furthermore that $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}}$ terminates by setting $\mathcal{A}_{\mathcal{M}}^t = \emptyset$ whenever $\sum_{i=1}^K \sum_{j \neq i} \mathbb{I}_{i,j}^t \leq \rho$ one obtains, based on the derivations above, that

$$\mathbb{P}\left\{\left(A_{\mathcal{M}}^t \neq \emptyset\right) \wedge \left(n_{i,j}^t > 2b_{i,j}^t \text{ for each } (i,j) \text{ with } \Delta_{i,j} \geq \Delta_{(r)}\right)\right\}$$

$$\leq \mathbb{P}\left\{\exists(i,j) : \left((i,j) \in A_{\mathcal{M}}^t\right) \wedge \left(\Delta_{i,j} \geq \Delta_{(r)}\right) \wedge \left(n_{i,j}^t > 2b_{i,j}^t\right)\right\} \leq \frac{3\delta}{10K^2t^4} \sum_{r=1}^{K^2-r_1} \frac{1}{(\Delta_{i,j} + \epsilon)^2}$$

where the last inequality follows using the union bound. $\qquad\square$

Proof of Theorem 10

Theorem 8. *Using the notation introduced in Lemma 9, the expected sample complexity for $\text{RANKEL}_{d_{\mathcal{M}}}^{\text{CO}_\epsilon}$ is $O\left(R_1 \log\left(\frac{R_1}{\delta}\right)\right)$ where $R_1 = \sum_{r=1}^{K^2-r_1}\left(\Delta_{(r)} + \epsilon\right)^{-2}$.*

Proof. The proof follows closely the one given for LUCB1(Kalyanakrishnan et al. 2012). Let be $T^* = \lceil 146R_1 \log\left(\frac{R_1}{\delta}\right)\rceil$. Let us consider an iteration T which $T^* < T$, and let be $\bar{T} = \lceil T^*/2 \rceil$. Consider the following events

$$E_1 := \exists t \in \left[\bar{T}, \cdots, T-1\right] \text{ and } \exists i, j : \mathcal{E}_{i,j}^t$$

where $\mathcal{E}_{i,j}^t$ is defined in Lemma 5, and

$$E_2 := \exists t \in \left[\bar{T}, \cdots, T-1\right] \text{ and } \exists(i,j) : \left\{\left((i,j) \in A_{\mathcal{M}}^t\right) \wedge \left(n_{i,j}^t > 2b_{i,j}^t \text{ where } \Delta_{i,j} \geq \Delta_{(r_1)}\right)\right\}$$

Assuming that neither E_1 nor E_2 occur, we can upper bound the rounds $\text{RANKEL}^{\text{CO}}_{d_{\mathcal{M}}}$ takes between iterations $\bar{T}$ and T as

$$\#rounds = \sum_{t=\bar{T}}^{T} \mathbb{I}\left\{\exists (i,j) : \left((i,j) \in A^t_{\mathcal{M}}\right) \wedge (n^t_{i,j} \leq b^t_{i,j}) \wedge (\Delta_{i,j} \geq \Delta_{(r_1)})\right\}$$

$$= \sum_{t=\bar{T}}^{T} \sum_{i=1}^{K} \sum_{j\neq i} \mathbb{I}\left\{\left((i,j) \in A^t_{\mathcal{M}}\right) \wedge (n^t_{i,j} \leq b^t_{i,j}) \wedge (\Delta_{i,j} \geq \Delta_{(r_1)})\right\}$$

$$\leq \sum_{i=1}^{K} \sum_{j\neq i} \sum_{t=\bar{T}}^{T} \mathbb{I}\left\{\left((i,j) \in A^t_{\mathcal{M}}\right) \wedge (n^t_{i,j} \leq b^t_{i,j}) \wedge (\Delta_{i,j} \geq \Delta_{(r_1)})\right\}$$

$$\leq \sum_{i=1}^{K^2-r_1} \left\lceil \frac{1}{2(\Delta_{(i)} + \epsilon)^2} \ln\left(\frac{5K^2 T^4}{4\delta}\right) \right\rceil \tag{10}$$

The last equation follows form $\sum_{t=\bar{T}}^{T} \mathbb{I}\left\{(n^t_{i,j} \leq b^t_{i,j})\right\} < b^T_{i,j}$ and the definition of r_1.

According to p.272 from (Kalyanakrishnan 2011) (with a slight modification), first let us assume that $T = CR_1 \log\left(\frac{R_1}{\delta}\right)$ where $C \geq 146$ and $R_1 = \sum_{r=1}^{K^2-r_1} \frac{1}{\left(\Delta_{(r)} + \epsilon\right)^2}$, then

$$2 + 8\sum_{i=1}^{K^2-r_1} \left\lceil \frac{1}{2(\Delta_{(i)} + \epsilon)^2} \ln\left(\frac{5K^2 T^4}{4\delta}\right) \right\rceil \leq 2 + 8(K^2 - r_1) + 4R_1 \ln\left(\frac{5K^2 T^4}{4\delta}\right)$$

$$\leq (10 + 4\ln(5))R_1 + 4R_1 \ln\left(\frac{n}{\delta}\right) + 16R_1 \ln(T)$$

$$\leq \cdots \leq$$

$$< CR_1 \log\left(\frac{R_1}{\delta}\right)$$

$$= T \ .$$

If neither E_1 nor $E_2 = 0$ occur, then $\text{RANKEL}^{\text{CO}_\epsilon}_{\mathcal{M}}$ terminates after at most iteration T. Consequently $\mathbb{P}[E_1 + E_2]$ upper bounds the probability that $\text{RANKEL}^{\text{CO}_\epsilon}_{\mathcal{M}}$ does not terminate after T round. Applying Lemma 5 and 9, we have

$$\mathbb{P}[E_1 + E_2] = \sum_{t=\bar{T}}^{T} \left(\frac{4\delta}{5t^3} + \frac{3\delta R_1}{10K^2 t^4}\right) = \frac{4\delta}{5T^2}\left(1 + \frac{3R_1}{8K^2 T}\right) \leq \frac{8\delta}{T^2}$$

Summarizing the things above, the expected sample complexity can be then obtained as

$$\left\lceil 146R_1 \log\left(\frac{R_1}{\delta}\right) \right\rceil + 8\delta \sum_{t=T^*}^{\infty} \frac{1}{t^2} \leq \left\lceil 146R_1 \log\left(\frac{R_1}{\delta}\right) \right\rceil + 4/3\pi^2 \delta$$

$\square$

The dependency of empirical sample complexity on parameters ρ and ϵ

In this experiment, we investigate the dependency of the empirical sample complexity achieved by RANKEL on the parameters ρ and ϵ. We used artificial data where the pairwise comparisons obey Bernoulli distribution i.e. for options o_i and o_j, if the realisation of $\text{Bern}(p_{i,j})$ with a parameter $p_{i,j}$ is 1, then o_i is preferred to o_j. The parameters of the Bernoulli distributions are generated as $p_{i,j} = (1 - b_{i,j})(1 - b'_{i,j}) + b_{i,j}b'_{i,j}$ where $b_{i,j} \sim \text{Bern}(1/2)$ and $b'_{i,j} \sim \text{Beta}(2,6)$. The probability density function of the distribution of $p_{i,j}$ values is plotted in Figure 2. For each run, we generated a new set of $p_{i,j}$ values. Note that $p_{i,j}$ coincides with $y_{i,j}$ for all $1 \leq i, j \leq K$ since here there is no incomparable or indifferent case.

Figure 3 and 4 show the relative empirical sample complexity achieved by RANKEL with respect to the uniform sampling versus parameters ρ and ϵ. As we seen from the plots, the RANKEL sampling strategy lessen significantly the pairwise comparisons are needed to solve the rank elicitation task given.
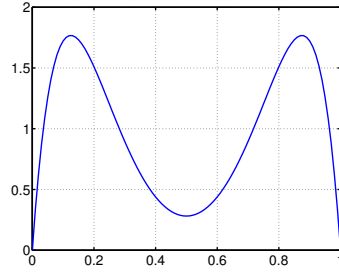
Figure 2: The probability density function of the distribution of the $p_{i,j}$ values.
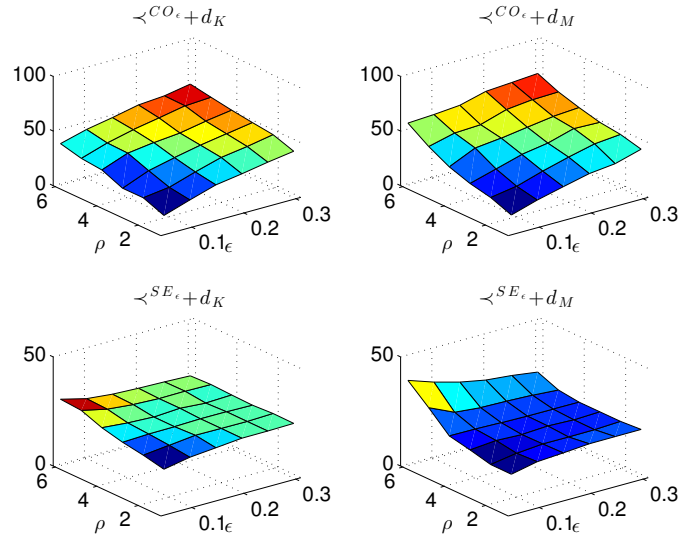

Figure 3: The relative empirical sample complexity achieved by $\text{RANKEL}_{d_\mathcal{K}}^{\text{CO}_\epsilon}$, $\text{RANKEL}_{d_\mathcal{M}}^{\text{CO}_\epsilon}$, $\text{RANKEL}_{d_\mathcal{K}}^{\text{SE}_\epsilon}$ and $\text{RANKEL}_{d_\mathcal{M}}^{\text{SE}_\epsilon}$, respectively, with respect to uniform sampling are plotted versus parameter ρ and ϵ. The number of options K is set to 8. Each result is the average of 100 repetitions. In each run, the number of pairwise comparisons drawn by uniform sampling was considered 100%. The confidence parameter δ was set to 0.1 for each run accordingly, the average accuracy were significantly above $1 - \delta = 0.9$ in each case.
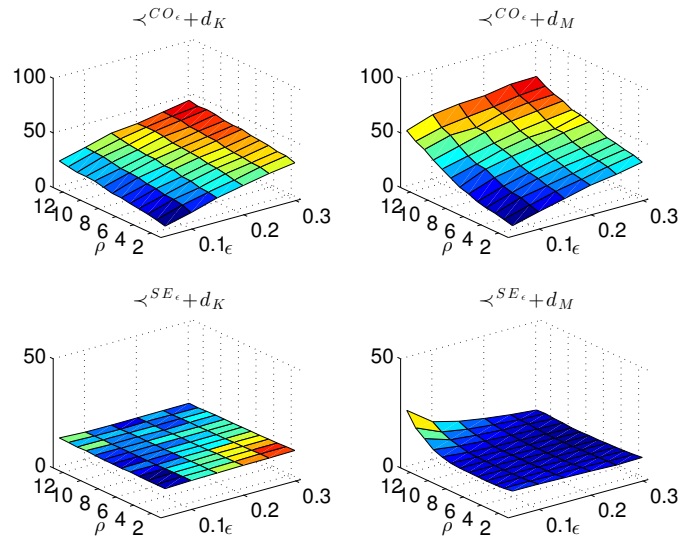

Figure 4: The same experiments as in Figure 3, but with $K = 15$