

# Approximate Ranking from Pairwise Comparisons

Reinhard Heckel\*   Max Simchowitz\*   Kannan Ramchandran\*   Martin J. Wainwright<sup>†,\*</sup>

Department of ECE\*, Rice University  
Department of Statistics<sup>†</sup>, and EECS\*, University of California, Berkeley

January 4, 2018

## Abstract

A common problem in machine learning is to rank a set of  $n$  items based on pairwise comparisons. Here ranking refers to partitioning the items into sets of pre-specified sizes according to their scores, which includes identification of the top- $k$  items as the most prominent special case. The score of a given item is defined as the probability that it beats a randomly chosen other item. Finding an exact ranking typically requires a prohibitively large number of comparisons, but in practice, approximate rankings are often adequate. Accordingly, we study the problem of finding approximate rankings from pairwise comparisons. We analyze an active ranking algorithm that counts the number of comparisons won, and decides whether to stop or which pair of items to compare next, based on confidence intervals computed from the data collected in previous steps. We show that this algorithm succeeds in recovering approximate rankings using a number of comparisons that is close to optimal up to logarithmic factors. We also present numerical results, showing that in practice, approximation can drastically reduce the number of comparisons required to estimate a ranking.

## 1 Introduction

The problem of ranking a collection of  $n$  items from noisy pairwise comparisons arises in a wide range of applications, including recommender systems for rating movies, books, or other consumer items [Pie+13; Agg16]; peer grading for ranking students in massive open online courses [Sha+13]; ranking players in tournaments; search engines; quantifying people’s perception of cities from pairwise comparison of street views of the cities [Sal+13]; and online sequential survey sampling for assessing the popularity of proposals in a population of voters [SL15].

In each of these applications, the aim is to obtain a statistically sound ranking from as few comparisons as possible. In this work, we investigate the power of adaptively selecting which pairs to compare based on the outcomes of previous comparisons, a setting we call *active* or *adaptive* ranking. In contrast, *passive* or *non-adaptive* ranking approaches fix the comparisons to make before any data is collected. It is well understood that one can typically learn a ranking using fewer adaptively chosen comparisons than one would need when passively choosing comparisons [Hec+16]. However, for moderately large or large collections of items—such as the ones that appear in most of the applications mentioned above—or for collections with many items of “similar quality” (to be made rigorous below), learning the *exact* ground-truth ranking may still require prohibitively many comparisons.

Motivated by these large-scale ranking problems, this work studies the problem of adaptivity obtaining *approximate* rankings. We demonstrate that learning an *approximate* ranking may still be statistically tractable even when recovering the *exact* ranking is not. Formally, we consider a collection of  $n$  items, and make comparison queries between pairs of items  $i, j \in [n] := \{1, 2, \dots, n\}$ .

We assume that the response to those queries are stochastic, where the probability that item  $i$  “beats” item  $j$  is given by  $M_{ij} \in (0, 1)$ . We assume that the outcomes of all queries are statistically independent, and assume that either item  $i$  or item  $j$  “wins” the comparison with probability 1, which means that  $M_{ij} + M_{ji} = 1$  for all  $i \neq j$ . Our aim is to rank the items in terms of their Borda scores [DB81], defined as the probability that item  $i$  defeats an item chosen uniformly at random from  $[n] \setminus \{i\}$ :

$$\tau_i := \frac{1}{n-1} \sum_{j \neq i} M_{ij} . \quad (1)$$

Apart from their intuitive appeal, the Borda scores generalize the orderings considered in several popular comparison models, including the classical, parametric Bradley-Terry-Luce (BTL) [BT52; Luc59] and Thurstone [Thu27] models, as well as the non-parametric Strong Stochastic Transitivity (SST) model [TER69]. In all of these models, the intrinsic model-defined ordering coincides with that given by the scores  $\{\tau_i\}_{i=1}^n$ . Rather than learning the scores  $\{\tau_i\}_{i=1}^n$  exactly, or ranking items according to their exact score, this paper considers the problem of *approximately* partitioning the items into sets of pre-specified sizes according to their respective scores. This includes finding a total ordering that is approximately correct, and the task of finding a set of  $k$  items that is close to the top- $k$  items. For simplicity, we exclusively focus on the latter problem in this paper.

**Contributions:** Our main contribution is to present and analyze a novel active ranking algorithm for estimating an *approximate* ranking of the items. The algorithm is based on adaptively estimating the scores to within sufficient resolution to deduce a ranking. We establish that with high probability, the algorithm returns a ranking which satisfies the desired approximation guarantee, and attains a distribution-dependent sample complexity which can be parameterized in terms of the scores  $\{\tau_i\}_{i=1}^n$ . We then prove distribution-dependent lower bounds that match our upper bound up to logarithmic factors for many problem instances. 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 [ED+06; Bub+13; Urv+13]; this same connection has been observed in the context of finding the top item [Yue+12; Jam+15; Urv+13]. Since to the best of our knowledge, the *approximate* subset selection problem has not been studied in the bandit literature, a version of our algorithm and results are also new when specialized to the multi-armed bandit problem. Finally, we examine pathological distributions for which the complexity of approximate ranking (or approximate subset selection in the multi-armed bandit setup) seems to diverge from what one would expect. In these cases, we show that careful randomized guessing strategies can yield significant improvements in sample complexity.

**Motivation for Approximate Rankings:** In order to understand how approximation can drastically reduce the number of comparisons required, let us consider a motivating example. Suppose that we are interested in identifying the top- $k$  items, and suppose for simplicity that the items are ordered, i.e.,  $\tau_1 > \tau_2 > \dots > \tau_n$  (of course this ordering is not known a-priori). The paper [Hec+16] shows that in the active setting, the number of comparisons necessary and sufficient for finding the top  $k$  items is of the order

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

up to a logarithmic factor. Thus, the sample complexity depends on the distribution of the scores; see Figure 1 how these scores are distributed in some applications. In practice, the differences between the scores often obey the scaling  $\tau_i - \tau_{i+1} \approx 1/n$  on average (see Figure 1). To identify the top- $k$  items exactly, the aforementioned optimal active scheme would require on the order of  $n^2$  comparisons, and a minimax-optimal passive ranking scheme would even require on the order of  $n^3$  comparisons [SW15].

Theorem 1 in this paper shows that if one does not need to extract the exact top- $k$  items, but is instead willing to tolerate a few—say,  $h$  many—mistakes, then the number of comparisons shrinks drastically, specifically by a factor proportional to  $h$ . In particular, if we want to find a set  $\mathcal{S}$  of 10% of the items ( $k = 0.1n$ ) such that all but 10% of the elements of  $\mathcal{S}_1$  are among the *true* top 10% of items ( $h = 0.1k$ ,  $k = 0.1n$ ), then the overall number of comparisons required would be on the order of  $n^2/h = 100n$ . Thus, relaxing to approximate ranking can yield speedups that are *linear* and *quadratic* in the number of items, compared to optimal exact active and exact passive schemes. Moreover, our algorithm (Algorithm 1 below) that obtains this factor-of- $h$  speedup does not require priori information about the spacings of the  $\{\tau_i\}_{i=1}^n$ , but instead learns a near-optimal measurement allocation for these scores adaptively.

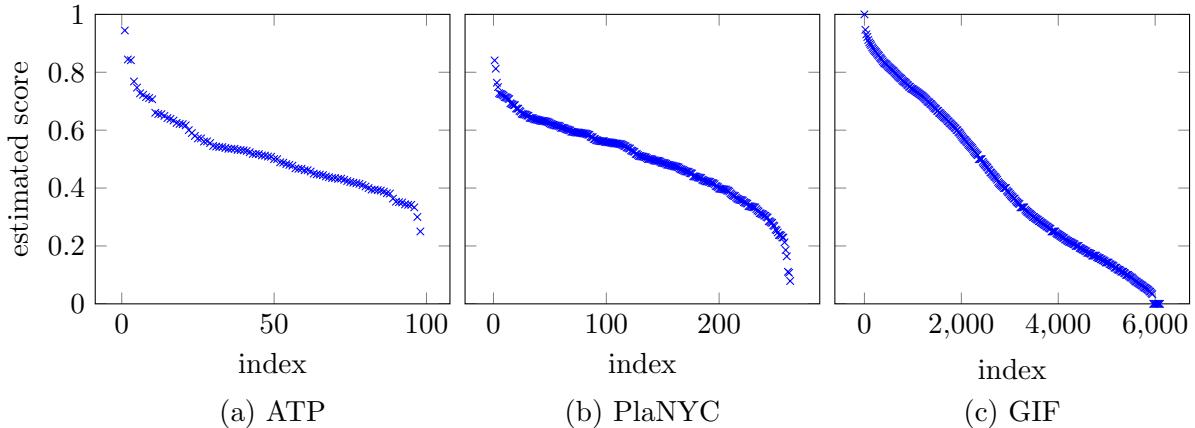
**Related works:** There is a vast literature on ranking and estimation from pairwise comparison data; however, most work focuses on finding *exact* rankings. There are a number of papers [Hum04; Neg+12; Haj+14; Sha+16a; SW15] devoted to settings in which pairs to be compared are chosen a priori, whereas here we assume that the pairs may be chosen in an active manner. Moreover, several works impose restrictions on the pairwise comparison probabilities, e.g., by assuming the Bradley-Terry-Luce (BTL) parametric model (discussed below) [Sz15; Hum04; Neg+12; Haj+14; Sha+16a]. Eriksson [Eri13] considers the problem of finding the very top items using graph-based techniques, whereas Busa-Fekete et al. [BF+13] consider the problem of finding the top- $k$  items. Ailon [Ail11] 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, as discussed later in the paper.

## 2 Problem formulation and background

In this section we formally state the approximate ranking problem considered in this paper.

### 2.1 Pairwise probabilities and scores

Given a collection of items  $[n] := \{1, \dots, n\}$ , let us denote by  $M_{ij} \in (0, 1)$  the (unknown) probability that item  $i$  wins a comparison with item  $j$ . We let  $X_{ij}$  denote a Bernoulli random variable taking a value of 1 if  $i$  beats  $j$  and 0 otherwise, so that  $M_{ij} = \mathbb{E}[X_{ij}]$ . Moreover, we require that any comparison results in a winner, so that  $M_{ij} + M_{ji} = 1$ . For each item  $i \in [n]$ , recall that the score (1) defined by  $\tau_i := \frac{1}{n-1} \sum_{j \in [n] \setminus \{i\}} M_{ij}$  corresponds to the probability that item  $i$  wins a comparison with an item  $j$  chosen uniformly at random from  $[n] \setminus \{i\}$ . We let  $\pi: [n] \rightarrow [n]$  denote any (possibly non-unique) permutation such that  $\tau_{\pi(1)} \geq \tau_{\pi(2)} \geq \dots \geq \tau_{\pi(n)}$ . In words,  $\pi(i)$  denotes the item with the  $i^{\text{th}}$  largest score. Ranking corresponds to partitioning the items into disjoint sets according to its scores. For simplicity, in this paper we focus on the ranking problem of splitting  $[n]$  into the top- $k$  items and its complement  $\mathcal{S}_1 := \{\pi(1), \dots, \pi(k)\}$ ,  $\mathcal{S}_2 := \{\pi(k+1), \dots, \pi(n)\}$ . In this work,



**Figure 1.** Estimated scores from three different domains: (a) Scores of the Association of Tennis Professionals (ATP) world tour, computed from the games played within a 52 week interval as the fraction of games won by the total number of games played. (b) Comparisons of the proposals in the PlaNYC survey, as reported in the paper [SL15] (only scores of items (proposals) that were rated at least 50 times are depicted). (c) Scores from comparisons of Gif’s according to whether they display a certain emotion (see <http://www.gif.gf/>).

our goal is to find an approximation to  $\mathcal{S}_1$  and  $\mathcal{S}_2$  in terms of the Hamming distance between two sets  $\mathcal{S}, \mathcal{S}'$ , defined as  $D_H(\mathcal{S}, \mathcal{S}') := |(\mathcal{S} \cup \mathcal{S}') \setminus (\mathcal{S} \cap \mathcal{S}')|$ . Specifically, we say the ranking  $\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2$  with  $|\widehat{\mathcal{S}}_\ell| = |\mathcal{S}_\ell|$  is  $h$ -Hamming-accurate if

$$D_H(\widehat{\mathcal{S}}_\ell, \mathcal{S}_\ell) \leq 2h, \quad \text{for } \ell \in \{1, 2\}.$$

For future reference, we define

$$\mathcal{C}_{M_{\min}} := \{M \in (0, 1)^{n \times n} \mid M_{ij} = 1 - M_{ji}, M_{ij} \geq M_{\min}\},$$

corresponding to the set of pairwise comparison matrices with pairwise comparison probabilities lower bounded by  $M_{\min}$ .

## 2.2 The active approximate ranking problem

An active ranking algorithm acts on a pairwise comparison model  $M \in \mathcal{C}_0$ . The goal is to obtain an approximate partition of the items into disjoint sets from active comparisons. At each time instant, the algorithm can compare two arbitrary items, which the algorithm may select based on the outcomes of previous comparisons. When comparing  $i$  and  $j$ , the algorithm obtains an independent draw of the random variable  $M_{ij}$  in response. The algorithm terminates based on an associated stopping rule, and returns an approximate ranking  $\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2$ . For a given tolerance parameter  $\delta \in (0, 1)$ , we say a ranking algorithm  $\mathcal{A}$  is  $(h, \delta)$ -accurate for a pairwise comparison matrix  $M$ , if the ranking returned is  $h$ -Hamming accurate with probability at least  $1 - \delta$ . Moreover, we say that  $\mathcal{A}$  is *uniformly*  $(h, \delta)$ -accurate over a given set of pairwise comparison models  $\mathcal{C}$  if it is  $\delta$ -accurate for each  $M \in \mathcal{C}$ .

### 2.3 Relation to multi-armed bandits

The *exact* version of the ranking problem considered in this paper is related to the subset selection problem in the bandit literature [Kal+12]. Specifically, a multi-armed bandit model consists of  $n$  arms, each a random variable with unknown distribution. The subset selection problem is concerned with identifying the top arms (according to the means) by taking independent draws of the random variables. Various works [YJ11; Yue+12; Urv+13; Jam+15] have observed that, by definition of the score  $\tau_i$ , comparing item  $i$  to an item chosen uniformly at random from  $[n] \setminus \{i\}$  can be modeled as drawing a Bernoulli random variable with mean  $\tau_i$ . Our subsequent analysis relies on this relation.

However, when viewing our problem as a multi-armed bandit problem with means  $\{\tau_i\}_{i=1}^n$ , we are ignoring the fact that the means are coupled, as they must be realized by some pairwise comparison matrix  $M$ . Due to  $M_{ij} = 1 - M_{ji}$ , this matrix must satisfy certain constraints, such as  $\sum_{i=1}^n \tau_i = n/2$  and  $\sum_{i=1}^j \tau_{\pi(i)} \geq \frac{1}{n-1} \frac{j(j-1)}{2}$  (e.g., see the papers [Lan53; Joe88]). Our algorithm turns out to be near-optimal, even though it does not take those constraints into account. This seems to corroborate the observation in [Sim+17] that many types of constraints surprisingly do not improve the sample complexity of bandit problems.

Finally, at least to the best of our knowledge, the problem of *approximate* subset selection has not been studied in the bandit literature, meaning that our algorithm and results are also new when specialized to the multi-armed bandit problem. However, it should be noted that other versions of approximation have been considered in the literature; for instance, Zhou et al. [Zho+14] studied the problem of selecting  $k$  arms with low aggregate regret, defined as the gap between the average reward of the optimal solution and the solution given by the algorithm.

### 2.4 Parametric models

In this section, we introduce a family of parametric models that are popular in the pairwise comparison literature [Sz15; Hun04; Neg+12; Haj+14; Sha+16a]. We focus on these parametric models in Section 3.3, where we show that, perhaps surprisingly, if the pairwise comparison probabilities are bounded away from zero, for most constellations of scores, these assumptions can at most provide little gains in sample complexity.

Any member of this family is defined by a strictly increasing and continuous function  $\Phi: \mathbb{R} \rightarrow [0, 1]$  obeying  $\Phi(t) = 1 - \Phi(-t)$ , for all  $t \in \mathbb{R}$ . The function  $\Phi$  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  $\mathcal{C}_{\text{PAR}(\Phi)}$  associated with the function  $\Phi$  is defined as:

$$\mathcal{C}_{\text{PAR}(\Phi)} = \{M_{ij} = \Phi(w_i - w_j) \quad \forall i, j \in [n], w \in \mathbb{R}^n\}.$$

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

## 3 Hamming-LUCB: Algorithm and analysis

In this section, we present our approximate ranking algorithm, and an analysis proving that it is near optimal for many interesting and natural problem instances.

### 3.1 The Hamming-LUCB algorithm

Our algorithm is based on actively identifying sets  $\tilde{\mathcal{S}}_1$  and  $\tilde{\mathcal{S}}_2$  consisting of  $k - h$  items and  $n - k - h$  items, respectively, such that with high confidence the items in the first set have a larger score than the items in the second set. Once we have found such sets, we can arbitrarily distribute the remaining items to the sets  $\tilde{\mathcal{S}}_1$  and  $\tilde{\mathcal{S}}_2$  in order to obtain a Hamming-accurate ranking with high confidence.

Our algorithm identifies those sets based on adaptively estimating the scores  $\{\tau_i\}_{i=1}^n$ . We estimate the score of item  $i$  by comparing item  $i$  with items chosen uniformly at random from  $[n] \setminus \{i\}$ , which yields an unbiased estimate of  $\tau_i$ . The key idea is to only estimate the scores sufficiently well so we can obtain the two sets  $\tilde{\mathcal{S}}_1$  and  $\tilde{\mathcal{S}}_2$  from them. This strategy decides based on the current estimates of the scores and associated confidence intervals which estimate to “update”, by comparing it to a randomly chosen item. Our strategy to update the estimates of the scores is guided by the insight that the “easiest” items to distinguish are the top  $k - h$  items,  $\{\pi(1), \dots, \pi(k - h)\}$ , and the bottom  $n - k - h$  items,  $\{\pi(k + h + 1), \dots, \pi(n)\}$ . Hence, our algorithm focuses on what it “thinks” are those top and bottom items.

We define a confidence bound based on a non-asymptotic version of the law of the iterated algorithm [Kau+16; Jam+14]; it is of the form  $\alpha(u) \propto \sqrt{\frac{\log(\log(u)n/\delta)}{u}}$ , where  $u$  is an integer corresponding to the number of comparisons, and with the constants involved explicitly chosen by setting

$$\alpha(u) = \sqrt{\frac{\beta(u, \delta/n)}{2u}}, \quad \text{with } \beta(u, \delta') = \log(1/\delta') + 0.75 \log \log(1/\delta') + 1.5 \log(1 + \log(u/2)).$$

For each item  $i \in [n]$ , the algorithm stores a counter  $T_i$  of the number of comparisons in which it has been involved, along with an empirical estimate of the associated score  $\hat{\tau}_i(T_i)$ . For notational convenience, we adopt the shorthands  $\hat{\tau}_i = \hat{\tau}_i(T_i)$  and  $\alpha_i = \alpha(T_i)$ . Within each round, we also let  $(\cdot)$  denote a permutation of  $[n]$  such that  $\hat{\tau}_{(1)} \geq \hat{\tau}_{(2)} \geq \dots \geq \hat{\tau}_{(n)}$ . We then define the indices

$$d_1 = \arg \min_{i \in \{(1), \dots, (k-h)\}} \hat{\tau}_i - \alpha_i, \quad d_2 = \arg \max_{i \in \{(k+1+h), \dots, (n)\}} \hat{\tau}_i + \alpha_i. \quad (3)$$

These indices are the analogues of the standard indices of the Lower-Upper Confidence Bound (LUCB) strategy from the bandit literature [Kal+12] for the top  $k - h$  and bottom  $n - k - h$  items. The LUCB strategy for exact top  $k$  recovery would update the scores  $d_1$  and  $d_2$  (for  $h = 0$ ) at each round. As mentioned before, our strategy will go after what it “thinks” are the top  $k - h$  items,  $\tilde{\mathcal{S}}_1 = \{(1), \dots, (k - h)\}$ , and what it “thinks” are the bottom  $n - k - h$  items,  $\tilde{\mathcal{S}}_2 = \{(k + 1 + h), \dots, (n)\}$ . Moreover, the algorithm keeps all the other items in consideration for inclusion in these sets, by keeping their confidence intervals below the confidence intervals of the items in  $\tilde{\mathcal{S}}_1$  and  $\tilde{\mathcal{S}}_2$  (cf. equation (4) in the algorithm below). This is crucial to ensure that the algorithm does not get stuck trying to distinguish the middle items  $\{\pi(k - h + 1), \dots, \pi(k + h)\}$ , which in general requires many comparisons, as their scores are typically closer. In Figure 2 we show an example run of the Hamming-LUCB algorithm, to illustrate the idea.

### 3.2 Guarantees and optimality of the Hamming-LUCB algorithm

We next establish guarantees on the number of comparisons for the Hamming-LUCB algorithm to succeed. As we show below, the number of comparisons depends on the following gaps between the

---

**Algorithm 1:** Hamming-LUCB
 

---

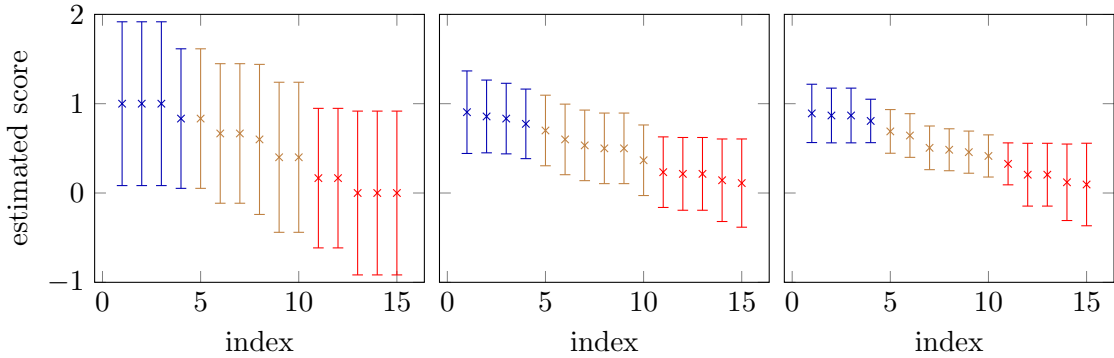
- 1 **Input:** Confidence parameter  $\delta$ .
- 2 **Initialization:** For every item  $i \in [n]$ , compare  $i$  to an item  $j$  chosen uniformly at random from  $[n] \setminus \{i\}$ , and set  $\hat{\tau}_i(1) = \mathbb{1}\{i \text{ wins}\}$ ,  $T_i = 1$ .
- 3 **Do** until termination:
- 4   Let  $(\cdot)$  denote a permutation of  $[n]$  such that  $\hat{\tau}_{(1)} \geq \hat{\tau}_{(2)} \geq \dots \hat{\tau}_{(n)}$ .
- 5   For  $d_1$  and  $d_2$  defined by equation (3), define the indices

$$b_1 = \arg \max_{i \in \{d_1, (k-h+1), \dots, (k)\}} \alpha_i, \quad \text{and} \quad b_2 = \arg \max_{i \in \{d_2, (k+1), \dots, (k+h)\}} \alpha_i. \quad (4)$$

- 6   **For**  $i \in \{b_1, b_2\}$ , increment  $T_i \leftarrow T_i + 1$ , compare  $i$  to an item  $j$  chosen uniformly at random from  $[n] \setminus \{i\}$ , and update  $\hat{\tau}_i \leftarrow \frac{T_i-1}{T_i} \hat{\tau}_i + \frac{1}{T_i} \mathbb{1}\{i \text{ wins}\}$ .
- 7   **End Loop** once the termination condition holds:

$$\hat{\tau}_{d_1} - \alpha_{d_1} \geq \hat{\tau}_{d_2} + \alpha_{d_2}. \quad (5)$$

- 8 **Return** the estimates of the partitions  $\hat{\mathcal{S}}_1 = \{(1), \dots, (k)\}$  and  $\hat{\mathcal{S}}_2 = \{(k+1), \dots, (n)\}$ .
- 



(a) after 30 comparisons    (b) after 200 comparisons    (c) at termination

**Figure 2.** Visualization of a run of the Hamming-LUCB algorithm on a problem instance with scores evenly spaced in the interval  $[0.1, 0.9]$ , and parameters  $k = 7, h = 3$ . The estimates of the scores  $\hat{\tau}_i$  of the top items  $\{(1), \dots, (k-h)\}$ , the middle items  $\{\pi(k-h+1), \dots, \pi(k+h)\}$ , and the bottom items  $\{(k+1+h), \dots, (n)\}$ , along with the confidence intervals  $[\hat{\tau}_i - \alpha_i, \hat{\tau}_i + \alpha_i]$  are depicted in blue, brown, and red, respectively, after 30 and 200 comparisons, and at termination. Note that once the confidence intervals of the top and bottom items are separated, the algorithm terminates.



scores

$$\Delta_{i,k+1+h} \quad \text{and} \quad \Delta_{k-h,i}, \quad \text{where} \quad \Delta_{i,j} := \tau_i - \tau_j.$$

Thus, as one might intuitively expect, the number of comparisons is typically smaller when  $h$  is larger, as the corresponding gaps typically become larger.

**Theorem 1.** *For any  $M \in \mathcal{C}_0$ , the Hamming-LUCB algorithm run with confidence parameter  $\delta$  is  $(h, \delta)$ -Hamming-accurate, and with probability at least  $1 - \delta$ , makes at most  $N_h^{\text{up}}(M)$  comparisons, where*

$$N_h^{\text{up}}(M) = \tilde{O} \left( \sum_{i=1}^{k-h} \Delta_{i,k+1+h}^{-2} + \sum_{i=k+1+h}^n \Delta_{k-h,i}^{-2} + 2h \Delta_{k-h,k+1+h}^{-2} \right). \quad (6)$$

The notation  $\tilde{O}$  absorbs factors logarithmic in  $n$ , and doubly logarithmic in the gaps.

Theorem 1 proves that the Hamming-LUCB algorithm is  $(h, \delta)$ -accurate, and characterizes the number of comparisons that it requires as a function of the gaps between the scores.

Comparing  $N_h^{\text{up}}(M)$  to the number of comparisons necessary and sufficient for finding the top- $k$  items, we see that the Hamming-LUCB algorithm depends on the gaps  $\Delta_{i,k+1+h}$  and  $\Delta_{k-h,i}$  instead of the gaps  $\Delta_{i,k+1}$  and  $\Delta_{k,i}$  which appear in the sample complexity for finding the top  $k$  items (cf. equation (2)). These gaps are typically significantly larger, resulting in a lower sample complexity. For example, in practice, the scores are often increasing in that  $\tau_i - \tau_{i+1}$  is on average on the order of  $1/n$ . Thus, for sufficiently large  $h$ , several real world models belong to the class (see Figure 1 for plausible members of this class):

$$\mathcal{C}_{\beta,h} := \{M \in (0,1)^{n \times n} \mid M_{ij} = 1 - M_{ji}, \text{ and } \tau_i - \tau_{i+h} \geq \beta h/n, \text{ for all } i\}. \quad (7)$$

For this class, the complexity of finding the top- $k$  items with the Hamming-LUCB algorithm is on the order of  $\tilde{O}(n^2/(\beta^2 h))$ , which is by a factor of  $h$  smaller than the complexity for finding the exact top- $k$  items.

Moreover, Hamming LUCB provides a strict improvement over the optimal sample complexity in the passive setup, for which Shah and Wainwright [SW15] establish upper bounds and minimax lower bounds which state that  $O(n \log n / \Delta_{k-h,k+1+h}^2)$  comparisons are necessary and sufficient to identify the top  $k$  items up to a Hamming error  $h$  with high probability.

As  $h$  increases, the upper bound depends on gaps between items with increasingly disparate position in the ranking, and thus, the upper bound on the sample complexity decreases. The following lower bound shows that, up to logarithmic factors in  $n$ , doubly logarithmic factors in the gaps, and a multiplicative scaling of  $h$ , the Hamming-LUCB algorithm is optimal.

**Theorem 2.** *For any  $\delta \in (0, 0.14]$ , let  $\mathcal{A}$  denote an algorithm which is uniformly  $(h, \delta)$ -accurate over  $\mathcal{C}_{1/8}$ . Then, when  $\mathcal{A}$  is run on any comparison instance  $M \in \mathcal{C}_{3/8}$ ,  $\mathcal{A}$  must make at least  $N_h^{\text{low}}(M)$  comparisons in expectation, where*

$$N_h^{\text{low}}(M) := c_{\text{low}} \log \left( \frac{1}{2\delta} \right) \left( \sum_{i=1}^{k-2h} \Delta_{i,k+1+2h}^{-2} + \sum_{i=k+1+2h}^n \Delta_{k-2h,i}^{-2} \right),$$

for some universal constant  $c_{\text{low}}$ .



Note that the above lower bound *does not* depend on the gaps involving the items  $k - 2h + 1, \dots, k + 2h$ . However, we can still relate the lower bound to the upper bound by (see Section A for the simple proof)

$$N_{3h}^{\text{up}}(M) \leq \tilde{O}(N_h^{\text{low}}(M)), \quad (8)$$

so that we see that, up to rescaling our Hamming error tolerance  $h$ , our upper and lower bounds ( $N_h^{\text{up}}(M)$  and  $N_h^{\text{low}}(M)$ , respectively) match up to logarithmic factors. For many problem instances of interest—such as models in the class  $\mathcal{C}_{\beta,h}$  in equation (7)—the sample complexity bounds  $N_{3h}^{\text{up}}(M)$  and  $N_h^{\text{low}}(M)$  degrade gracefully with the Hamming tolerance  $h$ , so that typically we have  $N_h^{\text{up}}(M) \leq \tilde{O}(N_h^{\text{low}}(M))$ .

Observe that if  $h = 0$ , we recover the exact top- $k$  recovery upper bound in equation (2), which is related to similar results for multi armed bandits [Kal+12]. We believe that by modifying the confidence intervals in Hamming LUCB as in the LUCB++ algorithm of Simchowitz et al. [Sim+17], one can sharpen the upper bound  $N_h^{\text{up}}(M)$  on the sample complexity by replacing  $\log n$  with  $\log k$  on the terms  $\Delta_{k-h,i}^{-2}$  corresponding to items  $i \in \{k + h + 1, \dots, n\}$ , thereby matching known lower bounds for top- $k$  subset selection problem in the bandit literature [Sim+17; Che+17; Kal+12]. In the interest of simplicity, we defer refining these logarithmic factors to later work.

### 3.3 Parametric models

Even though the lower bound of  $N_h^{\text{low}}(M)$  qualitatively matches the upper bound  $N_h^{\text{up}}(M)$ , it gives the misleading impression that an  $h$ -approximate algorithm can get away without querying the items in  $\{k - 2h, \dots, k + 2h + 1\}$ . In the proof section, we use techniques from [Sim+17] and [Che+17] to establish a more refined technical lower bound showing that all items, including those with ranks close to  $k$  must be compared an “adequate” number of times. For simplicity, we state a consequence of this lower bound applied to the parametric models described in Section 2.4. In addition to showing that each item has to be compared a certain number of times, this bound also establishes that *even knowledge of the exact parametric form of the pairwise comparison probabilities*  $M$  cannot drastically improve the performance of an active ranking algorithm.

In more detail, we say that a model is parametric, if there exists a strictly increasing CDF  $\Phi: \mathbb{R} \rightarrow [0, 1]$  such that  $M_{ij} = \Phi(w_i - w_j)$  for some weights  $\{w_j\}$ . For any pair of constants  $0 < \phi_{\min} \leq \phi_{\max} < \infty$ , we say that a CDF  $\Phi$  is  $(\phi_{\min}, \phi_{\max}, M_{\min})$ -bounded, if it is differentiable, and if its derivative  $\Phi'$  satisfies the bounds

$$\phi_{\min} \leq \Phi'(t) \leq \phi_{\max}, \quad \text{for all } t \in [\Phi^{-1}(M_{\min}), \Phi^{-1}(1 - M_{\min})]. \quad (9)$$

Note that for the popular BTL and Thurstone models, equation (9) holds with  $\phi_{\min}/\phi_{\max}$  close to one, provided that  $M_{\min}$  is not too small. We say that an algorithm is symmetric if its distribution of comparisons commutes with permutations of the items. For any such algorithm, our main lower bound is as follows:

**Theorem 3.** *For a given  $\delta \leq \frac{1}{2} \min(\frac{1}{k}, \frac{1}{n-k})$ , let  $\mathcal{A}$  be any symmetric algorithm that is uniformly  $(h, \delta)$ -Hamming accurate over  $M \in \mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$ . Then, when  $\mathcal{A}$  is run on the instance  $M \in \mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$ , for any integer  $q \geq 1$  and any item  $a \in [n]$ , it must make at least*

$$\frac{M_{\min} \phi_{\min}^2}{3\phi_{\max}^2} \left( \frac{2q-1}{2h+q} \right)^2 \max_{b \in \{k-2(h+q), k+1+2(h+q)\}} \Delta_{a,b}^{-2}$$

comparisons involving item  $a$  on average.

In particular, by choosing  $q = h$ , we see that the total sample complexity is lower bounded by

$$\sum_{i=1}^{k-3h} \Delta_{i,k+1+3h}^{-2} + \sum_{i=k+1+3h}^n \Delta_{k-3h,i}^{-2} + 6h\Delta_{k-3h,k+1+3h}^{-2}, \quad (10)$$

which is equivalent to the upper bound  $N_{3h}^{\text{up}}(M)$  achieved by the Hamming-LUCB algorithm up to logarithmic factors. The lower bound from Theorem 3 is stronger than the lower bound from Theorem 2, in that it applies to the larger class of algorithms that are only  $(h, \delta)$ -accurate over the smaller class of parametric models. In fact, the parametric subclass  $\mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$  is significantly smaller than the full set of pairwise comparison models  $\mathcal{C}_{M_{\min}}$ , in the sense that one can find matrices in  $\mathcal{C}_{M_{\min}}$  that cannot be well-approximated by any parametric model [Sha+16b]. Therefore, theorem 3 shows that, up to rescaling the Hamming error tolerance  $h$  and logarithmic factors, the Hamming-LUCB algorithm is optimal, even if we restrict ourself to algorithms that are uniformly  $(h, \delta)$ -accurate *only* over a parametric subclass. Thus, in the regime where the pairwise comparison probabilities are bounded away from zero, parametric assumptions cannot substantially reduce the sample complexity of finding an *approximate* ranking; an observation that has been made previously in the paper [Hec+16] for exact rankings.

A second and equally important consequence of Theorem 3 is that each item has to be sampled a certain number of times, an intuition not captured by Theorem 2. This conclusion continues to hold for general pairwise comparison matrices, please see Theorem 4 in Section 5.3 for a formal statement.

### 3.4 Random guessing

Even though our the upper and lower bounds essentially match whenever  $N_{3h}^{\text{up}}(M) \approx N_h^{\text{up}}(M)$ , there are there are pathological instances where  $N_{3h}^{\text{up}}(M) \ll N_h^{\text{up}}(M)$ , and where the Hamming-LUCB algorithm will make considerably more comparisons than a careful random guessing strategy.

As an example, consider a problem instance parameterized by  $\kappa, \epsilon$ , with scores given by

$$\tau_i = \begin{cases} 1/2 + \kappa, & i \in \{1, \dots, k - h - 2\} \\ 1/2 + 2\epsilon, & i = k - h - 1 \\ 1/2 + \epsilon, & i \in \{k - h, \dots, k\} \\ 1/2 - \epsilon, & i \in \{k + 1, \dots, k + 1 + h\} \\ 1/2 - 2\epsilon, & i = k + 2 + h \\ 1/2 - \kappa, & i \in \{k + h + 2, \dots, n\} \end{cases},$$

for some  $\kappa$  and  $\epsilon$ . The upper bound (6) for the Hamming-LUCB strategy is at least on the order of  $h/\epsilon^2$ , since the gap between the  $(k - h)$ -th and the  $(k + 1 + h)$ -th largest score is  $4\epsilon$ . However, the lower bound provided by Theorem 2 is  $(n - 2(h + 2))/\kappa^2$ , which is independent of  $\epsilon$ . Thus, by making  $\epsilon$  small, the ratio of upper and lower bounds becomes arbitrarily large. Intuitively, Hamming-LUCB is wasteful because it is attempting to identify the *exact* top  $k - h$  arms with too much precision. However, for this particular problem instance, the following random guessing strategy will attain our lower bound. First, we obtain estimates  $\hat{\tau}_i$  of each score  $\tau_i$  by comparing item  $i$  to  $T_i = c \log(n/\delta)/\kappa^2$  randomly chosen items. For each score, test whether there

are  $k - h - 2$  items obeying  $\widehat{\tau}_i \geq 1/2 + c\sqrt{\log(n)/T_i}$  and whether there are  $n - (k + h + 2)$  items obeying  $\widehat{\tau}_i \leq 1/2 - c\sqrt{\log(n)/T_i}$ . If yes, assign these items the estimates  $\widehat{\mathcal{S}}_1$  and  $\widehat{\mathcal{S}}_2$ , respectively, and assign all remaining items uniformly at random to the sets  $\widehat{\mathcal{S}}_1$  and  $\widehat{\mathcal{S}}_2$ , and terminate.

## 4 Experimental results

In this section, we provide experimental evidence that corroborates our theoretical claims that the Hamming-LUCB algorithm allows to significantly reduce the number of comparisons if one is content with an approximate ranking. We show that these gains are attained on a real-world data set. Specifically, we generate a pairwise comparison model by choosing  $M$  such that the Borda scores  $\tau_i$  coincide with those found empirically in the PlaNYC survey [SL15]; see panel (b) of Figure 1. We emphasize that, since Hamming LUCB depends only on the Borda scores  $\tau_i$  and not on the comparison probabilities  $M_{ij}$ , these simulations provide a faithful representation of how Hamming LUCB performs on real-world data. In Figure 3, we plot the results of running the Hamming-LUCB algorithm on the PlanNYC-pairwise comparison model in order to determine the top  $k = 35$  items, for different values of  $h$ . We observed that the results for other values of  $k$  are very similar. As suggested by our theory, the number of comparisons to find an approximate ranking decays in a manner inversely proportional in  $h$ . We compare the Hamming-LUCB algorithm to another sensible active ranking strategy for obtaining an Hamming-accurate ranking. Specifically, we consider a version of the successive elimination strategy proposed in [Hec+16, Sec. 3.1] for finding an exact ranking. This strategy can be adapted to yield an Hamming-accurate ranking by changing its stopping criterium. Instead of stopping once all items have been eliminated, we stop when either  $k - h$  items have been assigned to the top, or  $n - k - h$  items have been assigned to the bottom. While this strategy yields an Hamming accurate ranking, its sample complexity is, up to logarithmic factors, equal to  $\sum_{i=1}^{k-h} \frac{1}{(\tau_i - \tau_{k+1+h})^2} + \sum_{i=k+1+h}^n \frac{1}{(\tau_k - \tau_i)^2} + 2h(\tau_{k-h} - \tau_{k+1+h})$ , which is strictly smaller than that of the Hamming-LUCB algorithm. As Figure 3 shows, this strategy requires significantly more comparisons for finding an approximate ranking, thereby validating the benefits of our approach.

## 5 Proofs

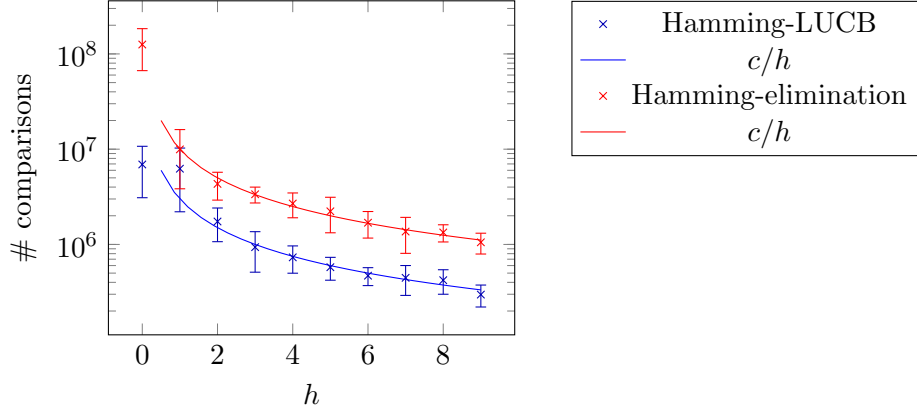
In this section, we provide the proofs of our theorems. In order to simplify notation, we assume without loss of generality (re-indexing as needed) that the underlying permutation  $\pi$  equal to the identity, so that  $\tau_1 > \tau_2 > \dots > \tau_n$ .

### 5.1 Proof of Theorem 1

Our analysis uses an argument inspired by the proof of the performance guarantee of the original LUCB algorithm from the bandit literature, presented in [Kal+12]. We begin by showing that the estimate  $\widehat{\tau}_i(T_i)$  is guaranteed to be  $\alpha_i$ -close to  $\tau_i$ , for all  $i$ , with high probability.

**Lemma 1** ([Kau+16, Lem. 19]). *For any  $\delta \in (0, 0.0005)$ , with probability at least  $1 - \delta$ , the event*

$$\mathcal{E}_\alpha := \{|\widehat{\tau}_i(t) - \tau_i| \leq \alpha_i, \quad \text{for all } i \in [n] \text{ and for all } t \geq 1\} \quad (11)$$



**Figure 3.** Sample complexity of the Hamming-LUCB algorithm and an elimination strategy run on a pairwise comparison model resembling the PlaNYC online sequential survey. Both algorithms find the top 35 proposals out of 263 proposals, up to Hamming error  $h$ . The error bars correspond to one standard deviation from the mean. The results show that the sample complexity of the Hamming-LUCB algorithm for finding an  $h$ -accurate ranking drops by a factor of about  $h$ . Moreover, the Hamming-LUCB algorithm requires significantly fewer samples than the elimination strategy.

occurs. The statement continues to hold for any  $\delta \in (0, 1)$  with  $\alpha_i = \alpha(T_i) = \sqrt{\frac{\beta(T_i, \delta')}{2T_i}}$ ,  $\beta(t, \delta') = 2 \log(125 \log(1.12t)/\delta')$ .

Lemma 1 is a non-asymptotic version of the law of the iterated logarithm from Kaufmann et al. [Kau+16] and Jamieson et al. [Jam+14].

We first show that, on the event  $\mathcal{E}_\alpha$  defined in equation (11), the Hamming-LUCB algorithm returns sets  $\widehat{\mathcal{S}}_1$  and  $\widehat{\mathcal{S}}_2$  obeying  $D(\widehat{\mathcal{S}}_\ell, \mathcal{S}_\ell) \leq 2h$  for  $\ell = 1, 2$ , as desired. Indeed, suppose that  $\{(1), \dots, (k-h)\} \subseteq \mathcal{S}_1$ . This implies that  $\mathcal{S}_1$  and  $\widehat{\mathcal{S}}_1$  differ in at most  $h$  values, which in turn implies that  $\mathcal{S}_2$  and  $\widehat{\mathcal{S}}_2$  differ by at most  $h$  values. Therefore,  $D(\widehat{\mathcal{S}}_\ell, \mathcal{S}_\ell) \leq 2h$  for  $\ell = 1, 2$ . Next, suppose that  $\{(1), \dots, (k-h)\} \not\subseteq \mathcal{S}_1$ . Then, at least one item in  $\{(1), \dots, (k-h)\}$  is in  $\mathcal{S}_2$ . Thus, on  $\mathcal{E}_\alpha$ , the termination condition (5) implies that  $\{(k+1+h), \dots, (n)\} \subset \{k+1, \dots, n\} = \mathcal{S}_2$ . Similarly as above, this in turn implies that  $D(\widehat{\mathcal{S}}_\ell, \mathcal{S}_\ell) \leq 2h$  for  $\ell = 1, 2$ .

We next show that on the event  $\mathcal{E}_\alpha$ , Hamming-LUCB terminates after the desired number of comparisons. Let  $\gamma := \frac{\tau_{k-h} + \tau_{k+1+h}}{2}$ , and define the event that item  $i$  is bad as

$$\mathcal{E}_{\text{bad}}(i) = \begin{cases} \widehat{\tau}_i < \gamma + 3\alpha_i, & i \in \{1, \dots, k-h\} \\ \widehat{\tau}_i > \gamma - 3\alpha_i, & i \in \{k+1+h, \dots, n\} \\ \alpha_i > \frac{\tau_{k-h} - \tau_{k+1+h}}{4}, & \text{otherwise.} \end{cases}$$

**Lemma 2.** *If  $\mathcal{E}_\alpha$  occurs and the termination condition (5) is false, then either  $\mathcal{E}_{\text{bad}}(b_1)$  or  $\mathcal{E}_{\text{bad}}(b_2)$  occurs.*

Given Lemma 2, we can complete the proof in the following way. For an item  $i$ , define

$$\Delta_i = \begin{cases} \tau_i - \tau_{k+1+h}, & i \in \{1, \dots, k-h\} \\ \tau_{k-h} - \tau_i & i \in \{k+1+h, \dots, n\} \\ \tau_{k-h} - \tau_{k+1+h}, & \text{otherwise,} \end{cases}$$

and let  $\tilde{T}_i$  be the largest integer  $u$  satisfying the bound  $\alpha(u) \leq \Delta_i/4$ . A simple calculation (see Section 5.1.1 for the details) yields that

$$\text{On the event } \mathcal{E}_\alpha, \text{ if } T_i \geq \tilde{T}_i \text{ holds, then } \mathcal{E}_{\text{bad}}(i) \text{ is false.} \quad (12)$$

Let  $t \geq 1$  be the  $t$ -th iteration of the steps in the LUCB algorithm, and let  $b_1$  and  $b_2$  be the two items selected in Step ?? of the algorithm. Note that in each iteration only those two items are compared to other items. By Lemma 2, we can therefore bound the total number comparisons by

$$\begin{aligned} 2 \sum_{t=1}^{\infty} \mathbb{1}\{\mathcal{E}_{\text{bad}}(b_1) \cup \mathcal{E}_{\text{bad}}(b_2)\} &\leq 2 \sum_{t=1}^{\infty} \sum_{i=1}^n \mathbb{1}\{(i = b_1 \cup i = b_2) \cap \mathcal{E}_{\text{bad}}(i)\} \\ &\stackrel{(i)}{\leq} 2 \sum_{t=1}^{\infty} \sum_{i=1}^n \mathbb{1}\{(i = b_1 \cup i = b_2) \cap T_i \leq \tilde{T}_i\} \\ &\stackrel{(ii)}{\leq} 2 \sum_{i=1}^n \tilde{T}_i. \end{aligned} \quad (13)$$

For inequality (i), we used the fact (12), and inequality (ii) follows because  $T_i(t) \leq \tilde{T}_i$  can only be true for  $\tilde{T}_i$  iterations  $t$ .

We conclude the proof by noting that the definition of  $\alpha(\cdot)$  and some algebra yields (see [Hec+16, Eq. (20)]) that for  $c_1$  sufficiently large

$$\tilde{T}_i \leq \frac{c_1}{(\Delta_i/4)^2} \log \left( \frac{n}{\delta} \log \left( \frac{2}{(\Delta_i/4)^2} \right) \right) \leq c_2 \log \left( \frac{n}{\delta} \right) \frac{\log(2 \log(2/\Delta_i))}{\Delta_i^2}.$$

Applying this inequality to the RHS of equation (13) above concludes the proof.

### 5.1.1 Proof of fact (12)

First, consider an item  $i \in \{k+1+h, \dots, n\}$ . We show that if  $T_i \geq \tilde{T}_i$ , then  $\mathcal{E}_{\text{bad}}(i)$  is false. On the event  $\mathcal{E}_\alpha$ ,

$$\hat{\tau}_i(\tilde{T}_i) + \alpha(\tilde{T}_i) \leq \tau_i + 2\alpha(\tilde{T}_i) \stackrel{(i)}{\leq} \tau_i + \frac{\Delta_i}{2} = \gamma + \frac{\Delta_i}{2} - \frac{\tau_{k-h} - \tau_i + \tau_{k+1+h} - \tau_i}{2} \leq \gamma, \quad (14)$$

where inequality (i) follows from  $\alpha(T_i) \leq \Delta_i/4$  for  $T_i \geq \tilde{T}_i$ , by definition of  $\tilde{T}_i$ , and the last inequality follows from  $\Delta_i = \tau_{k-h} - \tau_i$  and  $\tau_{k+1+h} - \tau_i \geq 0$ . Thus,  $\mathcal{E}_{\text{bad}}(i)$  does not occur.

For an item  $i \in \{1, \dots, k-h\}$ ,  $\mathcal{E}_{\text{bad}}(i)$  that is false, the argument is equivalent. For an item in the middle  $i \in \{k-h+1, \dots, k+h\}$ , the event  $\mathcal{E}_{\text{bad}}(i)$  is false by definition. This concludes the proof.

### 5.1.2 Proof of Lemma 2

We prove the lemma by considering all different values the indices  $b_1$  and  $b_2$  selected by the LUCB algorithm can take on, and showing that in each case  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  cannot occur simultaneously. For notational convenience, we define the indices

$$m_1 = \arg \max_{i \in \{(k-h+1), \dots, (k)\}} \alpha_i, \quad m_2 = \arg \max_{i \in \{(k+1), \dots, (k+h)\}} \alpha_i,$$

and note that

$$b_1 = \arg \max_{i \in \{d_1, m_1\}} \alpha_i, \quad \text{and} \quad b_2 = \arg \max_{i \in \{d_2, m_2\}} \alpha_i.$$

1. Suppose that  $b_1 \in \{1, \dots, k-h\}$  and  $b_2 \in \{k+1+h, \dots, n\}$ , and that both  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  do not occur. First note that

$$\widehat{\tau}_{d_1} - \alpha_{d_1} \geq \widehat{\tau}_{b_1} - \alpha_{b_1}. \quad (15)$$

In order to establish this claim, note that the inequality holds trivially with equality if  $b_1 = d_1$ . If  $b_1 = m_1$ , then it follows from  $\widehat{\tau}_{d_1} \geq \widehat{\tau}_{m_1}$  and  $\alpha_{d_1} \leq \alpha_{b_1}$ . Thus, we obtain

$$\widehat{\tau}_{d_1} - \alpha_{d_1} \geq \widehat{\tau}_{b_1} - \alpha_{b_1} > \gamma, \quad (16)$$

where the last inequality holds by the assumption that  $\mathcal{E}_{\text{bad}}(b_1)$  does not occur. An analogous argument yields that

$$\gamma > \widehat{\tau}_{b_2} + \alpha_{b_2} \geq \widehat{\tau}_{d_2} + \alpha_{d_2}. \quad (17)$$

Combining those inequalities yields  $\widehat{\tau}_{d_1} - \alpha_{d_1} > \widehat{\tau}_{d_2} + \alpha_{d_2}$ , which contradicts that the termination condition (5) is false.

2. Next, suppose that  $b_1$  is an index in the middle and  $b_2$  is in the very bottom, i.e.,  $b_1 \in \{k-h+1, \dots, k+h\}$ , and  $b_2 \in \{k+1+h, \dots, n\}$ , and both  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  do not occur.

First note that from  $b_2 \in \{k+1+h, \dots, n\}$  and  $\mathcal{E}_{\text{bad}}(b_2)$  not occurring, we have that

$$\gamma \geq \widehat{\tau}_{b_2} + 3\alpha_{b_2} \stackrel{(i)}{\geq} \widehat{\tau}_{d_2} + \alpha_{d_2} + 2\alpha_{b_2} \stackrel{(ii)}{\geq} \widehat{\tau}_i + \alpha_i + 2\alpha_{b_2}, \quad \text{for } i \in \{(k+1+h), \dots, (n)\}.$$

Here, inequality (i) holds by  $\widehat{\tau}_{b_2} \geq \widehat{\tau}_{d_2}$  and  $\alpha_{b_2} \geq \alpha_{d_2}$ , and inequality (ii) follows by the definition of  $d_2$ . On the event  $\mathcal{E}_\alpha$ , this implies

$$\gamma \geq \tau_i + 2\alpha_{b_2}. \quad (18)$$

Inequality (18) can only be true for all  $i \in \{(k+1+h), \dots, (n)\}$  if  $\gamma - \tau_{k+1+h} \geq 2\alpha_{b_2}$ , which is equivalent to

$$\alpha_{b_2} \leq \frac{\Delta}{4}, \quad \Delta := \tau_{k-h} - \tau_{k+1+h}.$$

Again using that  $b_2 \in \{k+1+h, \dots, n\}$  and  $\mathcal{E}_{\text{bad}}(b_2)$  not occurring, we have that

$$\gamma \geq \widehat{\tau}_{d_2} + \alpha_{d_2} \stackrel{(i)}{\geq} \widehat{\tau}_{d_1} - \alpha_{d_1} \stackrel{(ii)}{\geq} \widehat{\tau}_{d_1} - \frac{\Delta}{4}, \quad (19)$$

where inequality (i) holds since the termination condition (5) is false, and inequality (ii) follows from  $\alpha_{d_1} \leq \alpha_{b_1} \leq \frac{\Delta}{4}$ , where the last inequality holds since  $\mathcal{E}_{\text{bad}}(b_2)$  does not occur, by assumption.

From  $\widehat{\tau}_{d_1} \geq \widehat{\tau}_i$  for all  $i \in \{(k-h+1), \dots, (n)\}$ , it follows that for  $i \in \{d_1\} \cup \{(k-h+1), \dots, (n)\}$ ,

$$\gamma > \widehat{\tau}_i - \frac{\Delta}{4} \geq \tau_i - \alpha_i - \frac{\Delta}{4}. \quad (20)$$

Below, we show that

$$\alpha_i \leq \frac{\Delta}{4}, \quad \text{for all } i \in \{d_1\} \cup \{(k-h+1), \dots, (k+h)\}. \quad (21)$$

It follows that

$$\gamma > \tau_i - \frac{\Delta}{4} - \frac{\Delta}{4} \Leftrightarrow \tau_{k-h} > \tau_i, \quad \text{for all } i \in \{d_1\} \cup \{(k-h+1), \dots, (k+h)\}. \quad (22)$$

Together with equation (18), this yields that  $\tau_{k-h} > \tau_i$  for all  $i \in \{d_1\} \cup \{(k-h+1), \dots, (n)\}$ , which is a contradiction. This concludes the proof.

It remains to establish the claim (21). From the bound  $\alpha_{b_2} \leq \frac{\Delta}{4}$ , as shown above, we have  $\frac{\Delta}{4} \geq \alpha_{b_2} \geq \alpha_{m_2} \geq \alpha_i$  for all  $i \in \{(k+1), \dots, (k+h)\}$ , by definition of  $m_2$ . Moreover, for  $i \in \{(k-h), \dots, (k)\}$ , we have  $\alpha_i \leq \alpha_{d_1} \leq \alpha_{b_1} \leq \frac{\Delta}{4}$ , where the last inequality holds since  $b_1$  is in the middle and is not bad. This concludes the proof of (21).

3. The case where  $b_1$  lies in the very top and  $b_2$  lies in the middle, i.e.,  $b_1 \in \{1, \dots, k-h\}$  and  $b_2 \in \{k-h+1, \dots, k+h\}$ , and both  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  do not occur, can be treated analogously as the previous case.
4. Next, suppose that both  $b_1$  and  $b_2$  lie in the middle, i.e.,  $b_1, b_2 \in \{k-h+1, \dots, k+h\}$  and both  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  do not occur. We show that this leads a contradiction.

Towards this goal, first note that either

$$\gamma < \widehat{\tau}_{d_2} + \alpha_{d_2}. \quad (23)$$

holds true or

$$\gamma > \widehat{\tau}_{d_1} - \alpha_{d_1}, \quad (24)$$

holds true, but not both. In order to see this fact, note that if inequality (23) is violated, then

$$\gamma \geq \widehat{\tau}_{d_2} + \alpha_{d_2} \stackrel{(i)}{<} \widehat{\tau}_{d_1} - \alpha_{d_1}, \quad (25)$$

where step (i) follows from the termination condition (5) being false, by assumption. Likewise, if inequality (24) does not hold, then

$$\gamma \leq \widehat{\tau}_{d_1} - \alpha_{d_1} < \widehat{\tau}_{d_2} + \alpha_{d_2}.$$

We have shown that either condition (23) or (24) holds true, but not both simultaneously; consequently, we may conclude that at least one of these two conditions does not hold true. Next, we show that if either inequality (23) or inequality (24) does not hold true, then this leads to a contradiction, which concludes the proof.



First, suppose that inequality (23) *does not* hold true. Then, by definition of  $d_2$ , on  $\mathcal{E}_\alpha$ ,

$$\gamma \geq \widehat{\tau}_{d_2} + \alpha_{d_2} \geq \widehat{\tau}_i + \alpha_i \geq \tau_i, \quad \text{for all } i \in \{(k+1+h), \dots, (n)\}. \quad (26)$$

Moreover, by inequality (25) together with  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  not occurring, which implies that  $\alpha_{b_1}, \alpha_{b_2} \leq \frac{\Delta}{4}$ , the following inequality follows by the same argument as inequality (22) follow from inequality (19):

$$\tau_{k-h} > \tau_i, \quad \text{for all } i \in \{d_1\} \cup \{(k-h+1), \dots, (k+h)\}. \quad (27)$$

Together with condition (26), this yields a contradiction.

The argument for the case in which claim (23) is true is entirely analogous.

5. Finally, if  $b_1 \in \{k+1+h, \dots, n\}$  or if  $b_2 \in \{1, \dots, k-h\}$ , and both  $\mathcal{E}_{\text{bad}}(b_1)$  and  $\mathcal{E}_{\text{bad}}(b_2)$  do not occur, we reach a contradiction using similar arguments as in the previous cases.

## 5.2 Proof of Theorem 2

We now turn to the proof of the lower bound from Theorem 2.

We first introduce some notation required to state a useful lemma [Kau+16, 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  $\nu_{i_t}$  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  $\sigma$ -algebra generated by  $i_1, X_{i_1}, \dots, i_t, X_{i_t}$ . Algorithm  $\mathcal{A}$  has a stopping rule  $\xi$  that determines the termination of  $\mathcal{A}$ . We assume that  $\xi$  is a stopping time measurable with respect to  $\mathcal{F}_t$  and obeying  $\mathbb{P}[\xi < \infty] = 1$ .

Let  $N_i(\xi)$  denote the total number of times index  $i$  has been selected by the algorithm  $\mathcal{A}$  (until termination). For any pair of distributions  $\nu$  and  $\nu'$ , we let  $\text{KL}(\nu, \nu')$  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-Leibler 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  $\nu$  and an alternative distribution  $\nu'$ .

**Lemma 3** ([Kau+16, Lem. 1]). *Let  $\nu, \nu'$  be two collections of  $m$  probability distributions on  $\mathbb{R}$ . Then for any  $\mathcal{E} \in \mathcal{F}_\xi$  with  $\mathbb{P}_\nu[\mathcal{E}] \in (0, 1)$ , we have*

$$\sum_{i=1}^m \mathbb{E}_\nu[N_i(\xi)] \text{KL}(\nu_i, \nu'_i) \geq d(\mathbb{P}_\nu[\mathcal{E}], \mathbb{P}_{\nu'}[\mathcal{E}]). \quad (28)$$

Let us now use Lemma 3 to prove Theorem 2.

Define the event

$$\mathcal{E} := \left\{ D(\widehat{\mathcal{S}}_\ell, \mathcal{S}_\ell) \leq 2h \text{ for } \ell = 1, 2 \right\},$$

corresponding to success of the algorithm  $\mathcal{A}$ . Recalling that  $\xi$  is the stopping rule of algorithm  $\mathcal{A}$ , we are guaranteed that  $\mathcal{E} \in \mathcal{F}_\xi$ . Given the linear relations  $M_{ij} = 1 - M_{ji}$ , the pairwise comparison matrix  $M$  is determined by the entries  $\{M_{ij}, i = 1, \dots, n, j = i+1, \dots, n\}$ . Let  $N_{ij}(\xi)$  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 3 ensures that

$$\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}_M [N_{ij}] d(M_{ij}, M'_{ij}) \geq d(\mathbb{P}_M [\mathcal{E}], \mathbb{P}_{M'} [\mathcal{E}]). \quad (29)$$

Let  $\mathcal{M} := \{m_1, \dots, m_{2h+1}\}$  be a set of distinct items in  $\mathcal{S}_1$ . We next construct  $M' \in \mathcal{C}_{1/8}$  such that  $m_1, \dots, m_{2h+1} \notin \mathcal{S}_1(M')$  under the distribution  $M'$ . Since we assume algorithm  $\mathcal{A}$  to be uniformly  $(h, \delta)$ -Hamming-accurate over  $\mathcal{C}_{1/8}$ , we have both  $\mathbb{P}_M [\mathcal{E}] \geq 1 - \delta$  and  $\mathbb{P}_{M'} [\mathcal{E}] \leq \delta$ . To see this note that since  $\mathcal{S}_1$  and  $\mathcal{S}_1(M')$  differ in  $2h + 1$  elements, there is no set of cardinality  $k$  that differs from both  $\mathcal{S}_1$  and  $\mathcal{S}_1(M')$  in only  $h$  elements. It follows that

$$d(\mathbb{P}_M [\mathcal{E}], \mathbb{P}_{M'} [\mathcal{E}]) \geq d(\delta, 1 - \delta) = (1 - 2\delta) \log \frac{1 - \delta}{\delta} \geq \log \frac{1}{2\delta}, \quad (30)$$

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

It remains to specify the alternative matrix  $M' \in \mathcal{C}_0$ . The alternative matrix  $M'$  is defined as

$$M'_{ij} = \begin{cases} M_{mj} - \frac{n-1}{n-1-2h}(\tau_m - \tau_{k+1+2h}), & \text{if } i = m \text{ for } m \in \mathcal{M}, j \in [n] \setminus \mathcal{M} \\ M_{im} + \frac{n-1}{n-1-2h}(\tau_m - \tau_{k+1+2h}), & \text{if } j = m \text{ for } m \in \mathcal{M}, i \in [n] \setminus \mathcal{M} \\ M_{ij} & \text{otherwise.} \end{cases} \quad (31)$$

It follows that, for  $m \in \mathcal{M}$ ,

$$\begin{aligned} \tau'_m &= \frac{1}{n-1} \sum_{j \in [n] \setminus \{m\}} M'_{mj} \\ &= \frac{1}{n-1} \sum_{j \in [n] \setminus \{m\}} M_{mj} - \frac{1}{n-1} \sum_{j \in [n] \setminus \mathcal{M}} \frac{n-1}{n-1-2h}(\tau_m - \tau_{k+1+2h}) \\ &= \tau_{k+1+2h}. \end{aligned}$$

Similarly, all other scores  $\tau'_i$  are larger than  $\tau_i$  by a common constant, that is, for  $i \in [n] \setminus \mathcal{M}$ ,

$$\tau'_i = \tau_i + \frac{1}{n-1-2h} \sum_{m \in \mathcal{M}} (\tau_m - \tau_{k+1+2h}).$$

It follows that, under the distribution  $M'$  the items in the set  $\mathcal{M}$  are not among the  $k$  highest scoring items, which ensures that  $\mathcal{M} \cap \mathcal{S}_1(M) = \emptyset$ . Moreover,  $M' \in \mathcal{C}_{1/8}$ . This follows from the assumption  $M \in \mathcal{C}_{3/8}$ , which implies

$$M'_{mj} \leq \frac{5}{8} + \left( \frac{5}{8} - \frac{3}{8} \right) \leq \frac{7}{8},$$

and similarly  $M'_{mj} \geq \frac{1}{8}$ .

Next consider the total number of comparisons of item  $m$  with all others items, that is,  $N_m = \sum_{j \in [n] \setminus \{m\}} N_{mj}$ . By linearity of expectation, we have

$$\begin{aligned}
\sum_{m \in \mathcal{M}} \max_{j \in [n] \setminus \{m\}} d(M_{mj}, M'_{mj}) \mathbb{E}_M [N_m] &= \sum_{m \in \mathcal{M}} \max_{j \in [n] \setminus \{m\}} d(M_{mj}, M'_{mj}) \sum_{j' \in [n] \setminus \{m\}} \mathbb{E}_M [N_{mj'}] \\
&\stackrel{(i)}{\geq} \sum_{m \in \mathcal{M}} \sum_{j \in [n] \setminus \{m\}} \mathbb{E}_M [N_{mj}] d(M_{mj}, M'_{mj}) \\
&\stackrel{(ii)}{=} \sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}_M [N_{ij}] d(M_{ij}, M'_{ij}) \\
&\stackrel{(iii)}{\geq} d(\mathbb{P}_M [\mathcal{E}], \mathbb{P}_{M'} [\mathcal{E}]) \\
&\geq \log \frac{1}{2\delta}. \tag{32}
\end{aligned}$$

Here steps (i) and (ii) follows from the fact that  $d(M_{ij}, M'_{ij}) = 0$  for all  $(i, j)$  not in  $\{(m, j) \mid m \in \mathcal{M}, j \in [n] \setminus \mathcal{M}\}$  and not in  $\{(i, m) \mid m \in \mathcal{M}, i \in [n] \setminus \mathcal{M}\}$ , by definition of the  $M'_{ij}$  (see equation (31)), and step (iii) follows from inequality (29) (that is, from Lemma 3). Finally, inequality (32) follows from inequality (30).

We next upper bound the KL divergence on the left hand side of inequality (32). Using the inequality  $\log x \leq x - 1$  valid for  $x > 0$ , we have that

$$d(M_{mj}, M'_{mj}) \leq \frac{(M_{mj} - M'_{mj})^2}{M'_{mj}(1 - M'_{mj})} \leq d_m, \quad d_m := 16(\tau_m - \tau_{k+1+2h})^2. \tag{33}$$

Here, the last inequality follows from the definition of  $M'$  in equation (31), for  $j \in [n] \setminus \{m\}$ , and from  $\frac{1}{8} \leq M'_{mj} \leq \frac{7}{8}$ , which implies  $\frac{1}{M'_{mj}(1 - M'_{mj})} \leq 16$ . Applying inequality (33) to the left hand side of inequality (32) yields

$$\sum_{m \in \mathcal{M}} d_m \mathbb{E}_M [N_m] \geq \log \frac{1}{2\delta}, \quad \text{valid for each subset } \mathcal{M} \subseteq \mathcal{S}_1 \text{ of cardinality } 2h + 1. \tag{34}$$

We can therefore obtain a lower bound on  $\sum_{i \in \mathcal{S}_1} \mathbb{E}_M [N_m]$  by solving the minimization problem:

$$\begin{aligned}
&\underset{e_m \geq 0}{\text{minimize}} \sum_{m \in \mathcal{S}_1} e_m \\
&\text{subject to} \quad \sum_{m \in \mathcal{M}} d_m e_m \geq \log \frac{1}{2\delta} \quad \text{for each subset } \mathcal{M} \subseteq \mathcal{S}_1 \text{ of cardinality } 2h + 1. \tag{35}
\end{aligned}$$

Since the  $d_m$  are decreasing in  $m$ , the solution to this optimization problem is  $e_{k-2h}, \dots, e_k = 0$  and  $e_m = \log(1/2\delta)/d_m$ .

Using an analogous line of arguments for items in the set  $\mathcal{S}_2$ , we arrive at the following lower bound

$$\log \frac{1}{2\delta} \left( \sum_{i=1}^{k-2h} \frac{1}{8(\tau_i - \tau_{k+1+2h})} + \sum_{i=k+1+2h}^n \frac{1}{8(\tau_{k-2h} - \tau_i)} \right)$$

on the number of comparisons. This concludes the proof.

### 5.3 Alternative lower bound

In this section, we state a second lower bound on the number of comparisons, which shows that to obtain an  $(h, \delta)$ -Hamming accurate ranking, an algorithm has to compare *each* item a certain number of times. The proof of this lower bound also forms the foundation for the proof of Theorem 3.

**Theorem 4.** *Let  $\mathcal{A}$  be a symmetric algorithm, i.e., its distribution of comparisons commutes with permutations of the items, that is uniformly  $(h, \delta)$ -Hamming accurate over  $\mathcal{C}$ , with  $\delta \leq \frac{1}{2} \min(\frac{1}{k}, \frac{1}{n-k})$ . Choose an integer  $q \geq 1$ . Then, for any item  $a \in [n]$ , when applied to a given pairwise comparison model  $M \in \mathcal{C}$ , the algorithm  $\mathcal{A}$  must make at least*

$$\frac{2}{3} \left( \frac{2q-1}{2h+q} \right)^2 / \left( \max_{b \in \{k-2(h+q), \dots, k+1+2(h+q)\}} \max_{j \neq \{a, b\}} \left( \max d(M_{aj}, M_{bj}), d(M_{ab}, 1/2) \right) \right)$$

comparisons on average.

In the remainder of this section, we provide a proof of Theorem 4. For a given item  $a$ , we divide our proof into two cases, corresponding to whether or not  $\mathbb{P}[a \notin \widehat{\mathcal{S}}_1] > c_1 + \eta$ , where we define the scalar  $\eta := \frac{1-c_1-c_2}{2}$ .

**Case 1:** First, suppose that  $\mathbb{P}[a \notin \widehat{\mathcal{S}}_1] > c_1 + \eta$ . Pick some other item  $b$  in  $\{k-h', \dots, k\}$  that obeys  $\mathbb{P}[b \notin \widehat{\mathcal{S}}_1] \leq c_1$ . The following lemma guarantees that such an item exists:

**Lemma 4.** *Let  $\mathcal{A}$  be an algorithm that is  $(h, \delta)$ -Hamming accurate, with  $\delta \leq \frac{1}{2} \min(\frac{1}{k}, \frac{1}{n-k})$ . Let  $\widehat{\mathcal{S}}_1$  and  $\widehat{\mathcal{S}}_2$  be  $\mathcal{A}$ 's estimate of the top  $k$  items  $\mathcal{S}_1$  and the bottom  $n-k$  items  $\mathcal{S}_2$ , respectively. Choose constants  $c_1, c_2$  and  $h'$  such that  $h + \frac{1}{2} \leq c_1 h'$ ,  $h + \frac{1}{2} \leq c_2 h'$ , and  $c_1 + c_2 < 1$ . Then*

- i) *there exists an item  $b \in \{k-h', \dots, k\}$  such that  $\mathbb{P}[b \notin \widehat{\mathcal{S}}_1] \leq c_1$ , and*
- ii) *there exists an item  $b' \in \{k+1, \dots, k+1+h'\}$  such that  $\mathbb{P}[b' \in \widehat{\mathcal{S}}_1] \leq c_2$ .*

We use Lemma 3 from Kaufmann et al., which relates the expected number of comparisons to the uncertainty between the actual distribution  $M$  and an alternative distribution  $M'$  about the events  $\mathcal{E}_a := \{a \notin \widehat{\mathcal{S}}_1\}$  and  $\mathcal{E}_b = \{b \notin \widehat{\mathcal{S}}_1\}$ . Concretely, define the alternative matrix  $M'$  as

$$M'_{ij} = \begin{cases} M_{bj}, & i = a, j \in [n] \setminus \{a, b\} \\ M_{ib}, & j = a, i \in [n] \setminus \{a, b\} \\ 1/2, & i = a \text{ and } j = b, \text{ or } i = b \text{ and } j = a \\ M_{i,j}, & \text{otherwise.} \end{cases} \quad (36)$$

Since the algorithm  $\mathcal{A}$  is invariant to permutations of the labels, by assumption, we have that  $\mathbb{P}_{M'}[\mathcal{E}_a] = \mathbb{P}_{M'}[\mathcal{E}_b]$ , since  $a$  and  $b$  have the same distribution under the distribution specified by  $M'$ , and we assume  $\mathcal{A}$  to be symmetric. Moreover, by construction of  $M'$ , we have

$$\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}_\nu [N_{ij}(\xi)] \text{KL}(M_{ij}, M'_{ij}) = \sum_{j \notin \{a, b\}} \mathbb{E}_M [N_{aj}(\xi)] \text{KL}(M_{aj}, M_{bj}) + \mathbb{E}_M [N_{ab}(\xi)] \text{KL}(M_{aj}, 1/2).$$

Applying Lemma 3 from Kaufmann et al. (see equation (29)) then yields

$$\begin{aligned}
& \sum_{j \notin \{a,b\}} \mathbb{E}_M [N_{aj}(\xi)] \text{KL}(M_{aj}, M_{bj}) + \mathbb{E}_M [N_{ab}(\xi)] \text{KL}(M_{aj}, 1/2) \\
& \geq \max\{d(\mathbb{P}_\nu[\mathcal{E}_a], \mathbb{P}_{\nu'}[\mathcal{E}_a]), d(\mathbb{P}_\nu[\mathcal{E}_b], \mathbb{P}_{\nu'}[\mathcal{E}_b])\} \\
& = \max\{d(\mathbb{P}_\nu[\mathcal{E}_a], \mathbb{P}_{\nu'}[\mathcal{E}_a]), d(\mathbb{P}_\nu[\mathcal{E}_b], \mathbb{P}_{\nu'}[\mathcal{E}_a])\} \\
& \geq \min_{p \in [0,1]} \max\{d(\mathbb{P}_\nu[\mathcal{E}_a], p), d(\mathbb{P}_\nu[\mathcal{E}_b], p)\} \\
& \geq \frac{2}{3}\eta^2,
\end{aligned} \tag{37}$$

where the last inequality follows from Lemma 5 stated below, together with  $\mathbb{P}[a \notin \widehat{\mathcal{S}}_1] - \mathbb{P}[b \notin \widehat{\mathcal{S}}_1] \geq \eta$ , which follows from  $\mathbb{P}[a \notin \widehat{\mathcal{S}}_1] > c_1 + \eta$  and  $\mathbb{P}[b \notin \widehat{\mathcal{S}}_1] \leq c_1$ .

**Lemma 5.** For scalar  $p_a, p_b \in [0, 1]$ , let  $\bar{p}$  denote their average  $\frac{p_a + p_b}{2}$ . Then,

$$d(p_a, \bar{p}) + d(p_b, \bar{p}) \geq \min_{p \in [0,1]} \max\{d(p_a, p), d(p_b, p)\} \geq \frac{1}{2}(d(p_a, \bar{p}) + d(p_b, \bar{p})). \tag{38}$$

Moreover, if  $p_b - p_a \geq \eta$ , then

$$\frac{1}{2}(d(p_a, \bar{p}) + d(p_b, \bar{p})) \geq \frac{2}{3}\eta^2. \tag{39}$$

**Case 2:** Turning to the other case, suppose that  $\mathbb{P}[a \notin \widehat{\mathcal{S}}_1] \leq c_1 + \eta$ . Pick some other item  $b'$  in  $\{k+1, \dots, k+1+h'\}$  obeying  $\mathbb{P}[b' \notin \widehat{\mathcal{S}}_1] > 1 - c_2$ , and note that Lemma 4 ensures that such an item exists. Using a line of argument analogous to that above, we find that

$$\sum_{j \notin \{a,b\}} \mathbb{E}_M [N_{aj}(\xi)] \text{KL}(M_{aj}, M_{b'j}) + \mathbb{E}_M [N_{ab}(\xi)] \text{KL}(M_{aj}, 1/2) \geq \frac{2}{3}\eta^2. \tag{40}$$

Here we used that Lemma 5 together with the lower bound  $\mathbb{P}[a \in \widehat{\mathcal{S}}_1] - \mathbb{P}[b' \in \widehat{\mathcal{S}}_1] \geq (1 - c_1 - \eta) - c_2 = \eta$ , which in turn follows from the relations  $\mathbb{P}[a \notin \widehat{\mathcal{S}}_1] \leq c_1 + \eta$ ,  $\mathbb{P}[b' \notin \widehat{\mathcal{S}}_1] > 1 - c_2$ , and  $1 - c_1 - c_2 = 2\eta$ .

Combining inequalities (37) and (40) yields

$$\max_{b \in \{k-h', \dots, k+1+h'\}} \left\{ \sum_{j \notin \{a,b\}} \mathbb{E}_M [N_{aj}(\xi)] \text{KL}(M_{aj}, M_{bj}) + \mathbb{E}_M [N_{ab}(\xi)] \text{KL}(M_{aj}, 1/2) \right\} \geq \frac{2}{3}\eta^2. \tag{41}$$

Choosing  $h' = 2(h + q)$  and  $c_1 = c_2 = 1/2 - \eta$  concludes the proof.

### 5.3.1 Proof of Lemma 4

Since  $\mathcal{A}$  is  $(h, \delta)$ -accurate, we have

$$\sum_{i \in \mathcal{S}_1} \mathbb{P} \left[ i \notin \widehat{\mathcal{S}} \right] \leq h + \delta k \leq h + \frac{1}{2} \leq c_1 h',$$

where the last inequality holds by assumption. Thus, there are at most  $h'$  many  $i \in \mathcal{S}_1$  with  $\mathbb{P} \left[ i \notin \widehat{\mathcal{S}} \right] \geq c_1$ , which implies that for at least  $k - h'$  many items  $i \in \mathcal{S}_1$ , we have that  $\mathbb{P} \left[ i \notin \widehat{\mathcal{S}} \right] \leq c_1$ .

This in turn implies that there is at least one item  $b \in \{k - h', \dots, k\}$  obeying  $\mathbb{P} \left[ b \notin \widehat{\mathcal{S}} \right] \leq c_1$ .

Likewise, assuming that  $\mathcal{A}$  is  $(h, \delta)$ -accurate, we have

$$\sum_{i \in \mathcal{S}_2} \mathbb{P} \left[ i \in \widehat{\mathcal{S}} \right] \leq h + \delta(n - k) \leq h + \frac{1}{2} \leq c_2 h'.$$

Then there exists at least one arm  $b' \in \{k + 1, \dots, k + 1 + h'\}$  such that  $\mathbb{P} \left[ b' \in \widehat{\mathcal{S}} \right] \leq c_2$ .

### 5.3.2 Proof of Lemma 5

We start with proving inequality (38). Observe that, since  $d(x, y) \geq 0$ , we have

$$\min_{p \in [0, 1]} (d(p_a, p) + d(p_b, p)) \geq \min_{p \in [0, 1]} \max\{d(p_a, p), d(p_b, p)\} \geq \min_{p \in [0, 1]} \frac{1}{2} (d(p_a, p) + d(p_b, p)).$$

Hence, it suffices to show that  $\min_{p \in [0, 1]} \frac{1}{2} (d(p_a, p) + d(p_b, p)) = \frac{1}{2} (d(p_a, \bar{p}) + d(p_b, \bar{p}))$ . To this end, define the binary entropy  $H(q) := -q \log q - (1 - q) \log(1 - q)$ . We then have

$$\begin{aligned} \frac{1}{2} (d(p_a, p) + d(p_b, p)) &= -\frac{1}{2} (H(p_a) + H(p_b)) + \bar{p} \log \frac{1}{p} + (1 - \bar{p}) \log \frac{1}{1 - p} \\ &= -\frac{1}{2} (H(p_a) + H(p_b)) + H(\bar{p}) + \bar{p} \log \frac{\bar{p}}{p} + (1 - \bar{p}) \log \left( \frac{1 - \bar{p}}{1 - p} \right) \\ &= -\frac{1}{2} (H(p_a) + H(p_b)) + H(\bar{p}) + d(\bar{p}, p), \end{aligned}$$

which is minimized by taking  $p = \bar{p}$ , for which  $d(\bar{p}, \bar{p}) = 0$ . We can then expand

$$\begin{aligned} -H(\bar{p}) &= \frac{p_a + p_b}{2} \log(\bar{p}) + \left( 1 - \frac{p_a + p_b}{2} \right) \log(1 - \bar{p}) \\ &= \frac{p_a}{2} \log(\bar{p}) + \frac{1 - p_a}{2} \log(1 - \bar{p}) + \frac{p_b}{2} \log(\bar{p}) + \frac{1 - p_b}{2} \log(1 - \bar{p}). \end{aligned}$$

Thus

$$\begin{aligned} -\frac{1}{2} (H(p_a) + H(p_b)) + H(\bar{p}) &= \frac{1}{2} (-H(p_a) - p_a \log(\bar{p}) - (1 - p_a) \log(1 - \bar{p})) \\ &\quad + \frac{1}{2} (-H(p_b) - p_b \log(\bar{p}) - (1 - p_b) \log(1 - \bar{p})) \\ &= \frac{1}{2} \{d(p_a, \bar{p}) + d(p_b, \bar{p})\}, \end{aligned}$$

as needed.

We next prove inequality (39). We have

$$\begin{aligned}
d(p_a, \bar{p}) &= d\left(p_a, \frac{pb - p_a}{2} + p_a\right) \\
&\geq \min_{p \in [0,1]} d(p, \eta/2 + p) = d(1/2 + \eta/4, 1/2 - \eta/4) = \frac{\eta}{2} \log\left(\frac{1/2 + \eta/4}{1/2 - \eta/4}\right) \\
&\geq \frac{\eta}{2} \left(1 - \frac{1/2 - \eta/4}{1/2 + \eta/4}\right) \geq \frac{2}{3}\eta^2,
\end{aligned}$$

where the second to last, and the last inequality follow from  $\log x \geq 1 - 1/x$  and  $\eta \in [0, 1]$ , respectively. This concludes the proof of inequality (39).

#### 5.4 Proof of Theorem 3

The proof is analogous to that of the proof of Theorem 4 in Section 5.3, and only requires minor changes. Specifically, we only need to show that for a given pairwise comparison matrix  $M \in \mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$ , we can construct an alternative matrix obeying equality (36), that lies in  $\mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$  as well.

Consider any parametric pairwise comparison matrix  $M \in \mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$ . Then there exists a parameter vector  $w \in \mathbb{R}^n$  such that  $M_{ij} = \Phi(w_i - w_j)$ . For the items  $a, b \in [n]$ , in the proof of Theorem 4, define a set of alternative parameters as

$$w'_i := \begin{cases} w_b & \text{if } i = a, \\ w_i & \text{otherwise.} \end{cases}$$

Now let  $M'$  be the matrix with pairwise comparison probabilities  $M'_{ij} = \Phi(w'_i - w'_j)$ . Note that  $M' \in \mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$ , and observe that it obeys equality (36), as desired.

Thus, the proof of Theorem 4 yields that for any item  $a$ , when applied to a given pairwise comparison model  $M \in \mathcal{C}_{\text{PAR}(\Phi)} \cap \mathcal{C}_{M_{\min}}$ , the algorithm  $\mathcal{A}$  must make at least

$$\begin{aligned}
&\frac{2}{3} \left(\frac{2q-1}{2h+q}\right)^2 / \left(\max_{b \in \{k-2(h+q), \dots, k+1+2(h+q)\}} \max_{j \neq \{a,b\}} \left(d(M_{aj}, M_{bj}), d(M_{ab}, 1/2)\right)\right) \\
&\leq \frac{2}{3} \left(\frac{2q-1}{2h+q}\right)^2 / \left(\frac{2\phi_{\max}^2}{M_{\min}\phi_{\min}^2} (\tau_a - \tau_b)^2\right)
\end{aligned}$$

comparisons on average. Here, the last inequality follows from [Hec+16, Eq. (31)], which holds for any  $i \in [n]$ :

$$d(M_{ia}, M'_{ib}) \leq \frac{2\phi_{\max}^2}{M_{\min}\phi_{\min}^2} (\tau_a - \tau_b)^2. \quad (42)$$

Moreover, we used that

$$d(M_{ab}, 1/2) \leq \frac{2\phi_{\max}^2}{M_{\min}\phi_{\min}^2} (\tau_a - \tau_b)^2, \quad (43)$$

which follows along the lines as [Hec+16, Eq. (31)]. This concludes the proof.



## 6 Discussion

In this paper, we considered the problem of finding an Hamming-approximate ranking from pairwise comparisons. We provided an algorithm that allows to significantly reduce the sample complexity if one is content with an approximate ranking. Moreover, we showed that our algorithm is near optimal and remains near optimal when imposing common parametric assumptions. There are a number of open and practically relevant questions suggested by our work. As our work shows, it is non-trivial to adapt to approximate notions of ranking. It would be interesting to further understand how one can optimally adapt to approximate notions of ranking, by closing the gap of our bounds for pathological problem instances, and more importantly by studying other notions of approximate rankings. It would also be interesting to study algorithms that work with a limited budget of queries and quantify their approximation accuracy.

### Acknowledgements

The work of RH was supported by the Swiss National Science Foundation under grant P2EZP2\_159065. The work of MJW was partially supported by grants DOD ONR-N00014 and NSF-DMS-1612948.

## References

- [Agg16] C. C. Aggarwal. *Recommender systems: The textbook*. Springer, 2016.
- [Ail11] N. Ailon. “Active learning ranking from pairwise preferences with almost optimal query complexity”. In: *Advances in Neural Information Processing Systems*. 2011, pp. 810–818.
- [BT52] R. A. Bradley and M. E. Terry. “Rank analysis of incomplete block designs: I. The method of paired comparisons”. In: *Biometrika* 39.3/4 (1952), pp. 324–345.
- [Bub+13] S. Bubeck, T. Wang, and N. Viswanathan. “Multiple identifications in multi-armed bandits”. In: *International Conference on Machine Learning*. 2013, pp. 258–265.
- [BF+13] R. Busa-Fekete, B. Szorenyi, W. Cheng, P. Weng, and E. Hüllermeier. “Top-k selection based on adaptive sampling of noisy preferences”. In: *International Conference on Machine Learning*. 2013, pp. 1094–1102.
- [Che+17] L. Chen, J. Li, and M. Qiao. “Nearly instance optimal sample complexity bounds for top-k arm selection”. In: *arXiv preprint arXiv:1702.03605* (2017).
- [DB81] J. C. De Borda. “Mémoire sur les élections au scrutin”. In: *Histoire de l’Académie Royale des Sciences* (1781).
- [Eri13] B. Eriksson. “Learning to top-k search using pairwise comparisons”. In: *International Conference on Machine Learning*. 2013, pp. 265–273.
- [ED+06] E. Even-Dar, S. Mannor, and Y. Mansour. “Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems”. In: *Journal on Machine Learning Research* 7 (2006), pp. 1079–1105.
- [Haj+14] B. Hajek, S. Oh, and J. Xu. “Minimax-optimal inference from partial rankings”. In: *Advances in Neural Information Processing Systems*. 2014, pp. 1475–1483.

- [Hec+16] R. Heckel, N. B. Shah, K. Ramchandran, and M. J. Wainwright. “Active ranking from pairwise comparisons and when parametric assumptions don’t help”. In: *arXiv:1606.08842* (2016).
- [Hun04] D. Hunter. “MM algorithms for generalized Bradley-Terry models”. In: *Annals of Statistics* (2004), pp. 384–406.
- [Jam+14] K. Jamieson, M. Malloy, R. Nowak, and S. Bubeck. “lil’ UCB: An optimal exploration algorithm for multi-armed bandits”. In: *Conference on Learning Theory*. 2014, pp. 423–439.
- [Jam+15] K. Jamieson, S. Katariya, A. Deshpande, and R. Nowak. “Sparse dueling bandits”. In: *International Conference on Artificial Intelligence and Statistics*. 2015, pp. 416–424.
- [Joe88] H. Joe. “Majorization, entropy and paired comparisons”. In: *The Annals of Statistics* 16.2 (1988), pp. 915–925.
- [Kal+12] S. Kalyanakrishnan, A. Tewari, P. Auer, and P. Stone. “PAC subset selection in stochastic multi-armed Bandits”. In: *International Conference on Machine Learning*. Vol. 655–662. 2012.
- [Kau+16] E. Kaufmann, O. Cappé, and A. Garivier. “On the complexity of best arm identification in multi-armed bandit models”. In: *Journal on Machine Learning Research* 17.1 (2016), pp. 1–42.
- [Lan53] H. G. Landau. “On dominance relations and the structure of animal societies: III The condition for a score structure”. In: *The Bulletin of Mathematical Biophysics* 15.2 (1953), pp. 143–148.
- [Luc59] R. D. Luce. *Individual choice behavior: A theoretical analysis*. Wiley, 1959.
- [Neg+12] S. N. Negahban, S. Oh, and D. Shah. “Iterative ranking from pair-wise comparisons”. In: *Advances in Neural Information Processing Systems*. 2012, pp. 2474–2482.
- [Pie+13] C. Piech, J. Huang, Z. Chen, C. Do, A. Ng, and D. Koller. “Tuned models of peer assessment in MOOCs”. In: *International Conference on Educational Data Mining*. 2013.
- [Sal+13] P. Salesses, K. Schechtner, and C. A. Hidalgo. “The collaborative image of the city: Mapping the inequality of urban perception”. In: *PLOS ONE* 8.7 (2013), e68400.
- [SL15] M. J. Salganik and K. E. C. Levy. “Wiki surveys: Open and quantifiable social data collection”. In: *PLOS ONE* 10.5 (2015), e0123483.
- [SW15] N. B. Shah and M. J. Wainwright. “Simple, robust and optimal ranking from pairwise comparisons”. In: *arXiv:1512.08949* (2015).
- [Sha+13] N. B. Shah, J. K. Bradley, A. Parekh, M. J. Wainwright, and K. Ramchandran. “A case for ordinal peer-evaluation in MOOCs”. In: *NIPS Workshop on Data Driven Education*. 2013.
- [Sha+16a] N. B. Shah, S. Balakrishnan, J. Bradley, A. Parekh, K. Ramchandran, and M. J. Wainwright. “Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence”. In: *Journal on Machine Learning Research* (2016).

- [Sha+16b] N. B. Shah, S. Balakrishnan, A. Guntuboyina, and M. J. Wainwright. “Stochastically transitive models for pairwise comparisons: Statistical and computational issues”. In: *International Conference on Machine Learning*. 2016.
- [Sim+17] M. Simchowitz, K. Jamieson, and B. Recht. “The simulator: Understanding adaptive sampling in the moderate-confidence regime”. In: *arXiv:1702.05186* (2017).
- [Sz15] B. Szörényi, R. Busa-Fekete, A. Paul, and E. Hüllermeier. “Online rank elicitation for Plackett-Luce: A dueling bandits approach”. In: *Advances in Neural Information Processing Systems*. 2015, pp. 604–612.
- [Thu27] L. Thurstone. “A law of comparative judgment”. In: *Psychological Review* 34.4 (1927), pp. 273–286.
- [TER69] A. Tversky and J. Edward Russo. “Substitutability and similarity in binary choices”. In: *Journal of Mathematical Psychology* 6.1 (1969), pp. 1–12.
- [Urv+13] T. Urvoy, F. Clerot, R. Féraud, and S. Naamane. “Generic exploration and K-armed voting bandits”. In: *International Conference on Machine Learning*. 2013, pp. 91–99.
- [YJ11] Y. Yue and T. Joachims. “Beat the mean bandit”. In: *International Conference on Machine Learning*. 2011, pp. 241–248.
- [Yue+12] Y. Yue, J. Broder, R. Kleinberg, and T. Joachims. “The K-armed dueling bandits problem”. In: *Journal of Computer and System Sciences* 78.5 (2012), pp. 1538–1556.
- [Zho+14] Y. Zhou, X. Chen, and J. Li. “Optimal PAC multiple arm identification with applications to crowdsourcing”. In: *International Conference on Machine Learning*. 2014, pp. 217–225.

## A Proof of equation (8)

Equation (8) follows by upper bounding the terms in

$$N_h^{\text{up}}(M) = \tilde{O} \left( \sum_{i=1}^{k-3h} \Delta_{i,k+1+3h}^{-2} + \sum_{i=k+1+3h}^n \Delta_{k-3h,i}^{-2} + 2(3h)\Delta_{k-3h,k+1+3h}^{-2} \right).$$

Specifically, if  $i_1 < i_2$  and  $j_2 > j_1$ , then  $\Delta_{i_1,j_1} \leq \Delta_{i_2,j_2}$ . Therefore, the terms above can be upper bounded by

$$\begin{aligned} \sum_{i=1}^{k-3h} \Delta_{i,k+1+3h}^{-2} &\leq \sum_{i=1}^{k-3h} \Delta_{i,k+1+2h}^{-2}, & \sum_{i=k+1+3h}^n \Delta_{k-3h,i}^{-2} &\leq \sum_{i=k+1+3h}^n \Delta_{k-2h,i}^{-2}, \\ \text{and } 2h\Delta_{k-3h,k+1+3h}^{-2} &\leq \sum_{i=k-3h+1}^{k-2h} \Delta_{i,k-2h+1}^{-2} + \sum_{i=k+2h+1}^{k-3h} \Delta_{k-2h,i}^{-2}. \end{aligned}$$