
Just Sort It! A Simple and Effective Approach to Active Preference Learning

Lucas Maystre¹ Matthias Grossglauser¹

Abstract

We address the problem of learning a ranking by using adaptively chosen pairwise comparisons. Our goal is to recover the ranking accurately but to sample the comparisons sparingly. If all comparison outcomes are consistent with the ranking, the optimal solution is to use an efficient sorting algorithm, such as Quicksort. But how do sorting algorithms behave if some comparison outcomes are inconsistent with the ranking? We give favorable guarantees for Quicksort for the popular Bradley–Terry model, under natural assumptions on the parameters. Furthermore, we empirically demonstrate that sorting algorithms lead to a very simple and effective active learning strategy: repeatedly sort the items. This strategy performs as well as state-of-the-art methods (and much better than random sampling) at a minuscule fraction of the computational cost.

1 Introduction

The problem of recovering a ranking over n items from noisy outcomes of pairwise comparisons has attracted, in the last century, much research interest, driven by applications in sports (Elo, 1978), social sciences (Thurstone, 1927; Sarganik & Levy, 2015) and—more recently—recommender systems (Houlsby et al., 2012). Whereas pairwise comparison models and related inference algorithms have been extensively studied, the issue of *which pairwise comparisons to sample*, also known as active learning, has received significantly less attention. To understand the potential benefits of adaptively selecting samples, consider the case where comparison outcomes are noiseless, i.e., consistent with a linear order on a set of n items. If pairs of items are selected at random, it is necessary to collect $\Omega(n^2)$ comparisons to recover the ranking (Alon et al., 1994). In contrast, by using an efficient sorting algorithm, $O(n \log n)$ adaptively

chosen comparisons are sufficient. In this work, we demonstrate that sorting algorithms can also be helpful in the *noisy* setting, where some comparison outcomes are inconsistent with the ranking: despite errors, sorting algorithms tend to select informative samples. We focus on the Bradley–Terry (BT) model, a widely-used probabilistic model of comparison outcomes. In this model, each item is associated with a parameter on the real line, and the probability of observing an incorrect outcome decreases as the distance between the items’ parameters increases.

First, we study the output of a single execution of Quicksort when comparison outcomes are generated from a BT model, under the assumption that the distance between adjacent parameters is (stochastically) uniform across the ranking. We measure the quality of a ranking estimate by its displacement with respect to the ground truth, i.e., the sum of rank differences. We show that Quicksort’s output is a good approximation to the ground-truth ranking: no method comparing every pair of items at most once can do better (up to constant factors). Furthermore, we show that by aggregating $O(\log^5 n)$ independent runs of Quicksort, it is possible to recover the exact rank for all but a vanishing fraction of the items. These theoretical results suggest that adaptive sampling is able to bring a substantial acceleration to the learning process.

Second, we propose a practical active-learning (AL) strategy that consists of repeatedly sorting the items. We evaluate our sorting-based method on three datasets and compare it to existing AL methods. We observe that *all* the strategies that we consider lead to better ranking estimates noticeably faster than random sampling. However, most strategies are challenging to operate and computationally expensive, thus hindering wider adoption (Schein & Ungar, 2007). In this regard, sorting-based AL stands out, as *a*) it is computationally-speaking as inexpensive as random sampling, *b*) it is trivial to implement, and *c*) it requires no tuning of hyperparameters.

1.1 Preliminaries and Notation

We consider n items that are represented by consecutive integers $[n] = \{1, \dots, n\}$. Without loss of generality, we

¹School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland. Correspondence to: Lucas Maystre <lucas.maystre@epfl.ch>.

assume that the items are ranked by increasing preference¹, i.e., $i < j$ means that j is (in expectation) preferred to i . When j is preferred to i as a result of a pairwise comparison, we denote the observation by $i \prec j$. If $i < j$, we say that $i \prec j$ is a *consistent* outcome and $j \prec i$ an *inconsistent* (incorrect) outcome. In most of the paper, pairwise comparison outcomes follow a Bradley–Terry model with parameters $\theta = [\theta_1 \ \cdots \ \theta_n] \in \mathbf{R}^n$, denoted $\text{BT}(\theta)$. The parameters $\theta_1 < \cdots < \theta_n$ represent the utilities of items $1, \dots, n$, and the probability of observing the outcome $i \prec j$ is

$$p(i \prec j \mid \theta) = \frac{1}{1 + \exp[-(\theta_j - \theta_i)]}.$$

The probability of observing an inconsistent comparison decreases with the distance between the items. This captures the intuitive notion that some pairs of items are easy to compare and some are more difficult (Zermelo, 1928; Bradley & Terry, 1952).

A ranking σ is a function that maps an item to its rank, i.e., $\sigma(i) = \text{rank of item } i$. The (ground-truth) identity ranking is denoted by id , i.e. $\text{id}(i) = i$. To measure the quality of a ranking σ with respect to the ground-truth, we consider the *displacement*

$$\Delta(\sigma) = \sum_{i=1}^n |\sigma(i) - i|,$$

also known as Spearman’s footrule distance. Another metric widely used in practice is the Kendall–Tau distance, defined as $K(\sigma) = \sum_{i < j} \mathbf{1}\{\sigma(i) > \sigma(j)\}$. Both metrics are equivalent up to a factor of two², such that bounds on $\Delta(\sigma)$ also hold for $K(\sigma)$ up to constant factors.

Finally, we say that an event A holds *with high probability* if $\mathbf{P}[A] \rightarrow 1$ as $n \rightarrow \infty$. For a random variable X and a sequence of numbers a_n , we say that $X = O(a_n)$ with high probability if $\mathbf{P}[|X| \leq ca_n] \rightarrow 1$ as $n \rightarrow \infty$ for some constant c that does not depend on n .

Outline of the paper. We begin by briefly reviewing related literature in Section 2. Next, in Section 3, we study the displacement of Quicksort’s output under noisy comparisons. In Section 4, we empirically evaluate several AL strategies on three datasets. Finally, we conclude in Section 5.

2 Related Work

Passive setting. Recently, there have been a number of results on the sample complexity of the BT model, based on

the assumption that all pairs of items are chosen *before* any comparison outcome is revealed (Negahban et al., 2012; Hajek et al., 2014; Rajkumar & Agarwal, 2014; Vojnovic & Yun, 2016). In general, these results reveal that choosing pairs of items uniformly at random is essentially optimal. Furthermore, they suggest that the ranking induced by the BT model cannot be recovered with less than $\Omega(n^2)$ comparisons. Our work shows that by *adaptively* selecting pairs based on observed outcomes, we observe substantial gains.

Active preference learning. AL approaches for learning a ranking based on noisy comparison outcomes have been studied under various assumptions. Braverman & Mossel (2008) examine a model where outcomes of pairwise comparisons are flipped with a small, constant probability. Ailon (2012) considers an adversarial setting (comparison outcomes can be arbitrary) and investigates AL in the context of finding a ranking that minimizes the number of inconsistent outcomes, also known as the minimum feedback-arc set problem on tournaments (MFAST). These theoretical studies imply, in their respective settings, that $O(n \log^k n)$ comparison outcomes are enough to recover a near-optimal ranking. Jamieson & Nowak (2011) propose an efficient active-ranking algorithm that is applicable if items can be embedded in \mathbf{R}^d (e.g., using d features) and assuming that admissible rankings satisfy some geometric constraints. Wang et al. (2014) study a collaborative preference-learning problem and show that a variant of uncertainty sampling (a well-known AL strategy) works well for their problem. In this work, we assume that we do not have access to item features and that comparison outcomes follow a single BT model.

Bayesian methods. From a practical standpoint, Bayesian methods provide an effective way to select informative samples (MacKay, 1992). However, they can be difficult to scale if the number of items is large. Work on Bayesian active preference learning includes Chu & Ghahramani (2005), Houlby et al. (2012), Salimans et al. (2012) and Chen et al. (2013). We compare our AL strategy to these methods in Section 4.

Multi-armed bandit. The *dueling bandit* problem (Yue et al., 2009) is somewhat related to our work. In this problem, the goal is to identify the best item based on noisy comparison outcomes, using as few adaptively chosen samples as possible. Two recent papers also extend the problem to that of recovering the entire ranking (instead of only the top element). The work of Szörényi et al. (2015) is the closest to ours, as it also uses the BT model. One of their results is similar to our Theorem 2: They show that a quasi-linear number of comparisons is sufficient to recover the true ranking, under some conditions on θ . Heckel et al. (2016) investigate a non-parametric model and develop some theoretical guarantees. In contrast to these works, our paper studies practical

¹ This convention greatly simplifies the notation throughout the paper, but differs from that used in most of the preference learning literature. In our paper, the item with rank 1 is the *worst*.

² $\Delta(\sigma)/2 \leq K(\sigma) \leq \Delta(\sigma)$ (Diaconis & Graham, 1977).

Algorithm 1 Quicksort

Require: set of items V

- 1: **if** $|V| < 2$ **then return** $\text{list}(V)$ \triangleright Terminating case.
- 2: $L \leftarrow \emptyset, R \leftarrow \emptyset$
- 3: $p \leftarrow$ element of V selected uniformly at random
- 4: **for** $i \in V \setminus \{p\}$ **do**
- 5: **if** $i < p$ **then** \triangleright Pairwise comparison.
- 6: $L \leftarrow L \cup \{i\}$
- 7: **else**
- 8: $R \leftarrow R \cup \{i\}$
- 9: **return** $\text{Quicksort}(L) \cdot p \cdot \text{Quicksort}(R)$

comparison budgets: we give theoretical guarantees for the output obtained from a single call to Quicksort, and in our experiments we never exceed ≈ 10 calls.

Quicksort. The Quicksort algorithm (Hoare, 1962) is one of the most widely studied sorting procedures. Quicksort has been shown to produce useful rankings beyond classic sorting problems. For example, Ailon et al. (2008) show that Quicksort produces (in expectation) a 3-approximation to the MFAST problem. Quicksort combined with BT comparison outcomes has also been proposed as a probabilistic ranking model (Ailon, 2008). We take advantage of some of the properties of this ranking model in order to derive the theoretical results of Section 3.

3 Theoretical Results

In this section, we begin by studying the behavior and output of Quicksort under inconsistent comparison outcomes, without any assumptions on the noise generating process. Then, starting in Section 3.1, we focus on comparison outcomes generated by the BT model. Due to limited space, most full proofs are deferred to the supplementary material (Section A).

Quicksort (Algorithm 1) is best described as a recursive procedure. At each step of the recursion, a *pivot* item p is chosen uniformly at random (line 3). Then, during the *partition* operation (lines 4–8), every other item is compared to p and added to the set L or R , depending on the outcome. If all comparison outcomes are consistent, it is well-known that Quicksort terminates after sampling $O(n \log n)$ comparisons with high probability. What happens if we drop the consistency assumption? The following two lemmas state that these key properties remain valid, no matter which (and how many) comparison outcomes are inconsistent.

Lemma 1. *Quicksort always terminates and samples each of the $n(n-1)/2$ possible comparisons at most once.*

Proof. The proof is identical to the consistent setting. Consider the state of L and R at the end of a partition operation.

Because $|L| + |R| = |V| - 1$, the recursive calls are made on sets of items of strictly decreasing cardinality, and the algorithm terminates after a finite number of steps. Furthermore, suppose that Quicksort samples an outcome for the pair (i, j) . Then either i or j is the pivot in a partition operation. In either case, the pivot is not included in the recursive calls, which ensures that (i, j) cannot be compared again. \square

Lemma 2. *Quicksort samples $O(n \log n)$ comparisons w.h.p.*

Proof (sketch). We follow a standard analysis of Quicksort (see, e.g., Dubhashi & Panconesi, 2009, Section 3.3.3). With high probability, we choose a “good” pivot (i.e., one that results in a balanced partition) a constant fraction of the time. In this case, the depth of the call tree is $O(\log n)$. As there are at most n comparisons at each level of the call tree, we conclude that Quicksort uses $O(n \log n)$ comparisons in total. With respect to the standard proof, we need some additional work to formalize the notion of “good” pivot to the setting where comparison outcomes are not consistent with a linear order. \square

Lemma 2 complements Theorem 3 in Ailon & Mohri (2010), which states that Quicksort samples $O(n \log n)$ in expectation. These results might suggest that *all* properties of Quicksort carry over to the noisy setting. This is not the case. For example, although Quicksort uses approximately $2n \ln n$ comparisons on average in the noiseless setting (Sedgewick & Wayne, 2011), this number can be distinctly different with inconsistent comparison outcomes³.

Quicksort (and efficient sorting algorithms in general) infer most pairs of items’ relative position by transitivity and thus rely heavily on the consistency of comparison outcomes. In the noisy case, it is therefore important to precisely understand the effect of an inconsistent outcome on the output of the algorithm; this effect extends beyond the pair of items whose comparison outcome was inconsistent. For this purpose, the next Lemma bounds the displacement of Quicksort’s output as a function of the inconsistent outcomes.

Lemma 3. *Let E be the set of pairs sampled by Quicksort and whose outcome is inconsistent with id . Let σ be the output. Then,*

$$\Delta(\sigma) \leq 2 \sum_{(i,j) \in E} |i - j|$$

Proof (sketch). Consider the first partition operation, with pivot p , resulting in partitions L and R . Denote the errors

³E.g., if comparison outcomes are uniformly random, all items are “good” pivots w.h.p., and the average number of comparisons will be closer to $n \log_2 n$ on average, for large n .

made during this partition operation by E_1 . We can show that the displacement is bounded by

$$\Delta(\sigma) \leq \Delta_L(\sigma) + \Delta_R(\sigma) + 2 \sum_{(i,j) \in E_1} |i - j|,$$

where $\Delta_L(\sigma)$ and $\Delta_R(\sigma)$ represent the displacement of the ordering induced by σ on L and R , respectively. In other words, the total displacement can be decomposed into a term that represents the “local” displacement due to the partition operation and into two terms that account for errors in the recursive calls. We obtain the desired result by recursively bounding $\Delta_L(\sigma)$ and $\Delta_R(\sigma)$. \square

Informally, Lemma 3 states that the displacement can be bounded by a sum of “local shifts” due to the inconsistent outcomes and that the price to pay for any information inferred by transitivity is bounded by a factor two. Lemma 3 is a crucial component of our subsequent analysis of BT noise, and we believe that it can be useful in order to investigate Quicksort under a wide variety of other noise generating processes.

3.1 Displacement in the Poisson Model

From here on, we assume that comparison outcomes are generated from $\text{BT}(\theta)$. Clearly, any results on the displacement of a ranking estimated from samples of a BT model will depend on θ ; it is easy to construct a model instance for which it is arbitrarily hard to recover the ranking, by choosing parameters sufficiently close to each other. Our approach is as follows. We postulate a family of distributions over θ , and we give bounds on the displacement that hold with high probability.

We suppose that comparison outcomes are (in expectation) *uniformly noisy across the ranking*: i.e., comparing two elements at the bottom is (a priori) as difficult as comparing two elements at the top or in the middle. This means that the probability distribution over parameters $\theta_1, \dots, \theta_n$ results in (random) distances $|\theta_{i+k} - \theta_i|$ that depend only on k . One such distribution arises if the parameters are drawn from a Poisson point process of rate λ . That is,

$$\text{i.i.d. } x_1, \dots, x_{n-1} \sim \text{Exp}(\lambda), \quad \theta_i = \sum_{k=1}^{i-1} x_k. \quad (1)$$

The average distance between two items separated by k positions in the ordering is $\mathbf{E}[\theta_{i+k} - \theta_i] = k/\lambda$. Although the distance between adjacent items is constant in expectation, we allow some parameters to be arbitrarily close⁴. The parameter λ controls the expected level of noise; a large λ is

⁴ In particular, the expected minimum distance between two items (i.e., the min of n exponential r.v.s) decreases as $(n\lambda)^{-1}$ as n increases.

likely to result in a larger number of inconsistent outcomes. Although the precise choice of this Poisson model is driven by tractability concerns, in Section 3.2 we argue that it is essentially equivalent to choosing the parameters independently and uniformly at random in the interval $[0, (n+1)/\lambda]$, when λ is fixed and n is large. We are now ready to state our main result.

Theorem 1. *Let θ be sampled from a Poisson point process of rate λ . Let σ be the output of Quicksort using comparison outcomes sampled from $\text{BT}(\theta)$. Then, w.h.p.,*

$$\Delta(\sigma) = O(\lambda^2 n), \quad (2)$$

$$\max_i |\sigma(i) - i| = O(\lambda \log n). \quad (3)$$

Proof (sketch). Let z_{ij} be the indicator random variable of the event “the comparison between i and j results in an error”, and let $d_{ij} = |\theta_i - \theta_j|$. The distance d_{ij} is a sum of $|i - j|$ exponential random variables, i.e., $d_{ij} \sim \text{Gamma}(|i - j|, \lambda)$, and we can show that

$$\begin{aligned} \mathbf{E}[z_{ij}] &= \mathbf{E}\left[\frac{1}{1 + \exp(d_{ij})}\right] \\ &\leq \mathbf{E}[\exp(-d_{ij})] = (1 + 1/\lambda)^{-|i-j|}. \end{aligned}$$

Using Lemma 3 and the fact that every pair of items is compared at most once, we find

$$\begin{aligned} \mathbf{E}[\Delta] &\leq 2 \sum_{i < j} |i - j| \mathbf{E}[z_{ij}] \\ &\leq 2n \sum_{k=0}^{\infty} k(1 + 1/\lambda)^{-k} = 2n\lambda(\lambda + 1). \end{aligned}$$

The random variables $\{z_{ij}\}$ are not unconditionally independent (they are independent when conditioned on θ) but, with some more work, we can show that $\text{Var}[\Delta] = O(n)$. By using a Chebyshev bound, (2) follows.

In order to prove (3), we take advantage of a theorem due to Ailon (2008) which states that

$$\mathbf{P}[\sigma(i) < \sigma(j) \mid \theta] = p(i \prec j \mid \theta),$$

even if i and j were not directly compared with each other. We use a Chernoff bound on d_{ij} to show that the relative order between any two items separated by at least $O(\lambda \log n)$ positions is correct with high probability. The second part of the claim follows easily. \square

Note that any method that compares each pair of items at most once results in a ranking estimate τ with displacement $\Delta(\tau) = \Omega(n)$ with high probability: As there is only a single (possibly inconsistent) comparison outcome between each pair of adjacent items, it is likely that a constant fraction of the items will be ranked incorrectly, resulting in a

Algorithm 2 Multisort

Require: set of items V , number of iterations m

- 1: $S \leftarrow \emptyset$
 - 2: **for** $k = 1, \dots, m$ **do**
 - 3: $\sigma \leftarrow \text{Quicksort}(V)$
 - 4: $S \leftarrow S \cup \{\sigma\}$
 - 5: **return** Copeland aggregation of S
-

displacement that grows linearly in n . Hence, our bound on $\Delta(\sigma)$ shows that Quicksort is order-optimal (in n).

In light of Theorem 1, a natural question to ask is as follows. How many comparisons are needed in order to find the correct ranking? Clearly, finding the exact ranking is difficult: in fact, $\Omega(n)$ comparison outcomes are necessary to discriminate the closest pair of items reliably (see supplementary material, Section B). As such, we will focus on finding a ranking that matches the ground truth everywhere, except at a vanishing fraction of the items.

Multiple runs of Quicksort likely produce different outputs, because of the noisy comparison outcomes and because the algorithm itself is randomized (the pivot selection is random). By aggregating m independent outputs of Quicksort, is it possible to produce a better ranking estimate? Similarly to Szörényi et al. (2015), we combine the m outputs $\sigma_1, \dots, \sigma_m$ into an aggregate ranking $\hat{\sigma}$ using Copeland’s method. The method assigns, to each item, a score that corresponds to the number of items that it beats in a majority of the rankings, and it then ranks the items by increasing score (Copeland, 1951). We call the procedure Multisort and describe it in Algorithm 2.

Theorem 2. *Let θ be sampled from a Poisson point process of rate λ . Let $\hat{\sigma}$ be the output of Multisort using $m = O(\lambda^2 \log^5 n)$ and comparison outcomes sampled from $\text{BT}(\theta)$. Then, w.h.p.,*

$$\Delta(\hat{\sigma}) = o(\lambda n).$$

Proof (sketch). We use results on the order statistics of the distances x_1, \dots, x_{n-1} between successive items, as defined in (1), to partition the items into two disjoint subsets B and G . The set B contains a vanishing $(1/\log^2 n)$ -fraction of “bad” items that are difficult to order. The set G is such that the smallest distance d_{ij} from any item $i \in G$ to any other item $j \in [n]$ is bounded from below by $c/(\lambda \log^2 n)$. We can show that with $m = O(\lambda^2 \log^5 n)$, for any $i \in G$ and $j \in [n]$ we have $i < j \iff \sigma(i) < \sigma(j)$ in a majority of the Quicksort outputs (with high probability). This implies that $\hat{\sigma}(i) = i$ for all $i \in G$ with high probability. Using (3) for items in B , we have

$$\Delta(\hat{\sigma}) = |B| \cdot O(\lambda \log n) = O(\lambda n / \log n)$$

with high probability. \square

Theorem 2 states that all but a vanishing fraction of items are correctly ranked using $O(\lambda^2 n \log^6 n)$ comparisons. This result should be compared to the $\Omega(n^2)$ comparisons needed if samples are selected uniformly at random.

Empirical validation. In Figure 1, we illustrate the results of Theorems 1 and 2 by running simulations for increasing n and different values of λ . The bound on $\Delta(\sigma)$ is tight in n , but the dependence on λ appears to be linear rather than quadratic. The bound on $\max_i |\sigma(i) - i|$ appears to be tight in n and λ . Finally, we compare the Copeland aggregation of m outputs of Quicksort with the ranking induced by the maximum-likelihood (ML) estimate, inferred from the outcomes of all the pairwise comparisons sampled by the m runs. Although the ranking induced by the ML estimate does not benefit from the guarantees of Theorem 2, it performs better in practice. We will make use of this observation in Section 4.

3.2 Independent Uniformly-Distributed Parameters

A different (perhaps more natural) assumption on the parameters θ is to consider that they are drawn independently and uniformly at random over some interval. That is,

$$\text{i.i.d. } \bar{\theta}_1, \dots, \bar{\theta}_n \sim \text{U}(0, (n+1)/\lambda),$$

with $\theta_1, \dots, \theta_n$ the order statistics of $\bar{\theta}$, i.e., the random variables arranged in increasing order. From some elementary results on the joint distribution of order statistics (see, e.g., Arnold et al., 2008), we see that

$$|\theta_{i+k} - \theta_i| \sim (n+1)/\lambda \cdot \text{Beta}(k, n-k+1),$$

i.e., a Beta random variable rescaled between 0 and $(n+1)/\lambda$. Letting $f_{k,n}(x)$ be the probability density of $|\theta_{i+k} - \theta_i|$, we have, for any fixed k and λ ,

$$f_{k,n}(x) \propto x^{k-1} \left[1 - \frac{\lambda x}{n+1} \right]^{n-k} \xrightarrow{n \rightarrow \infty} x^{k-1} e^{-\lambda x}.$$

We recognize the functional form of the density of a Gamma(k, λ) distribution. Hence, the Poisson model and the i.i.d. uniform model are essentially equivalent for fixed λ and large n , and we can expect the results developed in Section 3.1 to hold under this distribution as well.

4 Experimental Results

In practice, the comparison budget for estimating a ranking from noisy data might typically be larger than that for a single call to Quicksort, and it might not exactly match the number of comparisons required to run a given number of calls to Quicksort to completion. Building upon the observations made at the end of Section 3.1, we suggest the following practical active-learning strategy: for a budget of

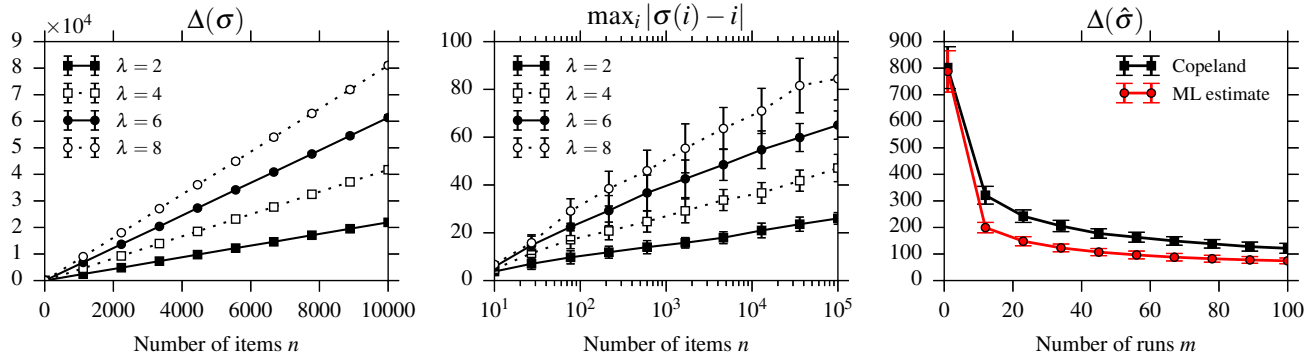


Figure 1. Empirical validation of Theorem 1 and illustration of Theorem 2. Every simulation is repeated 50 times, and we report the mean and the standard deviation. Left and middle: total and maximum displacement (respectively) for increasing n and different values of λ . Right: displacement of the aggregate ranking $\hat{\sigma}$ for increasing m , fixing $n = 200$ and $\lambda = 4$ and using two different aggregation rules.

c pairwise comparisons, run the sorting procedure repeatedly until the budget is depleted (the last call might have to be truncated). Then, retain only the set of c comparison pairs and their outcomes and discard the rankings produced by the sorting procedure. The final ranking estimate is then induced from the ML estimate over the set of c comparison outcomes.

In this section, we demonstrate the effectiveness of this sampling strategy on synthetic and real-world data. In particular, we show that it is comparable to existing AL strategies at a minuscule fraction of the computational cost.

4.1 Competing Sampling Strategies

To assess the relative merits of our sorting-based strategy, we consider three strategies that we believe are representative of the state of the art in active preference learning.

Uncertainty sampling. Developed in the context of classification tasks, this popular active-learning heuristic suggests to greedily sample the point that lies closest to the decision boundary (Settles, 2012). In the context of a ranking task, this corresponds to sampling the pair of items whose relative order is most uncertain. After t observations, given an estimate of model parameters θ^t , the strategy selects the $(t+1)$ -st pair uniformly at random in

$$\arg \min_{i \neq j} |\theta_i^t - \theta_j^t|.$$

This set can be computed in time $O(n \log n)$ by sorting the parameters. The parameters themselves need to be estimated, e.g., using (penalized) ML inference that in practice can be the dominating cost.

Bayesian methods. If we have access to a full posterior distribution $q^t(\theta)$ instead of a point estimate θ^t , we can take advantage of the extra information on the uncertainty of the parameters to improve the selection strategy. A principled approach to AL consists of sampling the point that

maximizes the expected information gain (MacKay, 1992). That is, the pair of items at iteration $t + 1$ is selected in

$$\arg \max_{i \neq j} H(q^t) - \mathbf{E} [H(q^{t+1})], \quad (4)$$

where $H(\cdot)$ denotes the entropy function. A conceptually similar but slightly different selection strategy is given by Chen et al. (2013). Letting q_{ij} be the marginal distribution of (θ_i, θ_j) , the pair is selected in

$$\arg \max_{i \neq j} \mathbf{E} [\text{KL}(q_{ij}^{t+1} \| q_{ij}^t)], \quad (5)$$

where $\text{KL}(\cdot)$ denotes the Kullback–Leibler divergence. Computing the exact posterior is not analytically tractable for the BT model, but a Gaussian approximation can be found in time $O(n^3)$. Criteria (4) and (5) can be computed in constant time for each pair of items. The dominating cost is again that of estimating θ (or, in this case, $q(\theta)$).

In addition to these existing AL strategies, we also include in our experiments a variation of our sorting-based strategy that uses Mergesort instead of Quicksort. In the noiseless setting, Mergesort is known to use on average $\approx 39\%$ fewer comparisons than Quicksort per run (Knuth, 1998), but it does not benefit from the theoretical guarantees developed in Section 3.

4.2 Running Time

In this section, we briefly discuss the running time of the methods. We implement ML and Bayesian approximate inference algorithms for the BT model as a Python library⁵. For ML inference, we find that the fastest running time is achieved by a truncated Newton algorithm (even for large n). For approximate Bayesian inference, we use a variant of the expectation-propagation algorithm outlined by Chu & Ghahramani (2005). All experiments are performed on

⁵See: <http://lucas.maystre.ch/choix>.

Table 1. Time (in seconds) to select the $(n + 1)$ -st pair. Values indicated by ε are below 10^{-5} . See text for details.

Strategy	T [s]		
	$n = 10^2$	$n = 10^3$	$n = 10^4$
uncertainty	0.05	0.5	11
entropy	0.3	40	—
KL-divergence	0.9	71	—
Mergesort	ε	ε	ε
Quicksort	ε	ε	ε
random	ε	ε	ε

a server with a 12-core Xeon X5670 processor running at 2.93 GHz. Numerical computations take advantage of the Intel Math Kernel Library.

We illustrate the running time of AL strategies as follows. For $n \in \{10^2, 10^3, 10^4\}$, we generate outcomes for n comparisons pairs chosen uniformly at random among n items. For each strategy, we then measure the time it takes to select the $(n + 1)$ -st pair of items adaptively. The results are presented in Table 1. Note that these numbers are intended to be considered as orders of magnitude, rather than exact values, as they depend on the particular combination of software and hardware that we use. The running time of the Bayesian AL strategies exceed 10 hours for $n = 10^4$ and the calls were stopped ahead of completion. Our sorting-based methods, like random sampling, are the only AL strategies whose running time is constant for increasing n (and for increasing c). In fact, their running time is negligible in comparison to the other strategies, including uncertainty sampling.

4.3 Empirical Evaluation

We now investigate three datasets and measure the displacement of rankings estimated from adaptively-chosen samples, as a function of the budget c . Note that in order to use uncertainty sampling and Bayesian methods, it is necessary to choose a regularization strength or prior variance in the inference step. Different values can result in drastically different outcomes (in particular for uncertainty sampling) and, in practice, choosing a good value can be a significant challenge⁶. In the following, we report results for the values that worked best *a posteriori*.

Synthetic dataset. We generate n i.i.d. parameters $\theta_1, \dots, \theta_n$ uniformly in $[0, (n + 1)/\lambda]$ and draw samples from $\text{BT}(\theta)$. The ground-truth ranking is the one induced by the parameters. Figure 2 presents results for $n = 200$ and $\lambda = 5$ (plots for different values of λ are presented in the supplementary material, Section C, and are qualitatively

⁶Observe that our sorting-based approach is entirely parameter-free and is therefore not affected by this issue.

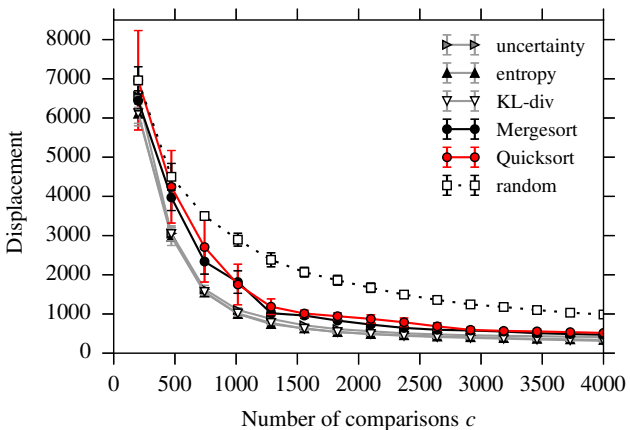


Figure 2. Synthetic dataset with $\lambda = 5$ and $n = 200$. The experiment is repeated 10 times, and we report the mean and the standard deviation. Compared to random sampling, AL results in significantly better rankings for a given budget c .

similar). In comparison to random sampling, AL is very effective and results in significantly better ranking estimates for any given number of comparisons. The two Bayesian methods, though being the most computationally expensive, perform the best for all values of c , but are nearly indistinguishable from uncertainty sampling. The two sorting-based strategies perform similarly (with a small edge for Mergesort). They are slightly worse than the Bayesian methods but are still able to reap most of the benefits of active learning.

Sushi dataset. Next, we consider a dataset of Sushi preferences (Kamishima & Akaho, 2009). In this dataset, 5000 respondents give a strict ordering over 10 different types of sushi. These 10 sushi are chosen among a larger set of $n = 100$ items. To suit our purposes, we decompose each 10-way partial ranking into pairwise comparisons, resulting in 225 000 comparison outcomes. We use all comparisons to fit a BT model that induces a ground-truth ranking⁷.

The comparisons are dense, and there is at least one comparison outcome for almost all pairs. When an outcome for pair (i, j) is requested, we sample uniformly at random over all outcomes observed for this pair. In the rare case where no outcome is available, we return $i \prec j$ with probability $1/2$. This enables us to compare sampling strategies in a realistic setting, where the assumptions of the BT model do not necessarily hold anymore.

Results are shown in Figure 3 (left). Once again, active learning performs noticeably better than random sampling. On this real-world dataset, the performance of our sorting-based strategies is indistinguishable from that of the Bayesian

⁷The BT-induced ranking is almost the same as that obtained using the Copeland score. The results are very similar if the Copeland aggregation is used as ground truth.

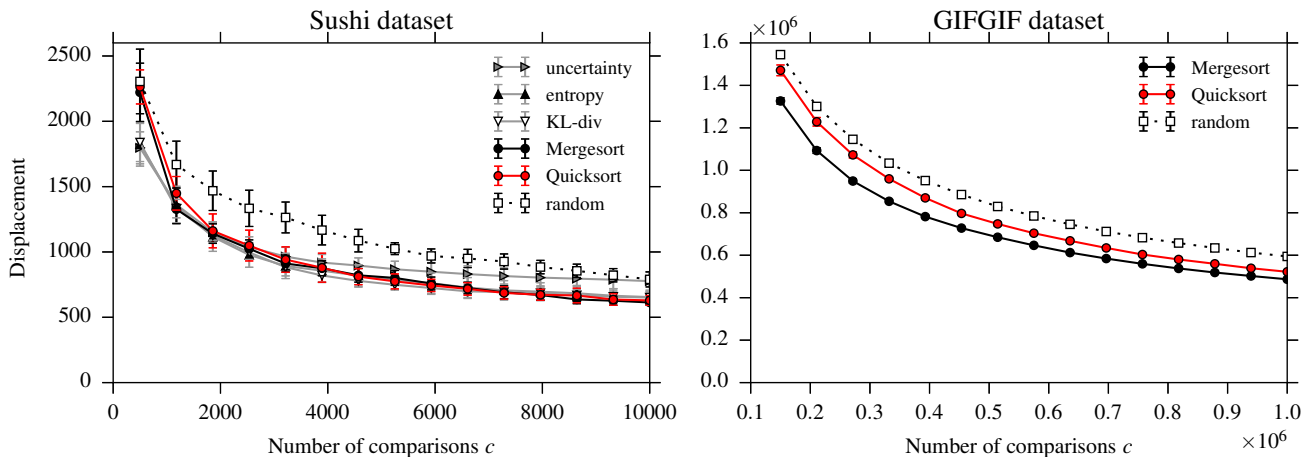


Figure 3. Results on two real-world datasets. Every experiment is repeated 10 times, and we report the mean and the standard deviation. Left: on the sushi dataset, sorting-based and Bayesian AL strategies have near-identical performance starting from $c \approx 1000$. Right: on the GIFGIF dataset, most AL strategies are computationally too expensive—except for sorting-based methods.

methods, after completing one entire call to the sorting procedure (slightly less than 1000 comparisons). This result should be interpreted in light of the time needed to select all 10^4 pairs: a fraction of a second for sorting-based strategies, and several hours for the Bayesian methods. Finally, we observe that the performance of uncertainty sampling progressively degrades as c increases. A detailed analysis reveals that uncertainty sampling increasingly focuses on a small set of hard-to-discriminate pairs, symptomatic of a well-known issue (Settles, 2012).

GIFGIF dataset. GIFGIF⁸ is a project of the MIT Media Lab that aims at explaining the emotions communicated by a collection of animated GIF images. Users of the website are shown a prompt with two images and a question, “Which better expresses x ?” where x is one of 17 emotions. The users can click on either image, or use a third option, *neither*. To date, over three million comparison outcomes have been collected. For the purpose of our experiment, we restrict ourselves to a single emotion, *happiness*; and we ignore outcomes that resulted in *neither*. We consider 106 887 comparison outcomes over $n = 6120$ items—a significant increase in scale compared to the Sushi dataset.

As the data, despite a relatively large number of comparisons, remains sparse (less than 20 comparisons per item on average), we proceed as follows. We fit a BT model by using all the available comparisons and use the induced ranking as ground truth. We then generate new, synthetic comparison outcomes from the BT model. In this sense, the experiment enables us to compare sampling strategies by using a large BT model with realistic parameters. The large number of items makes uncertainty sampling and the two Bayesian

⁸See <http://www.gif.gf/>. Data available at <http://lucas.maystre.ch/gifgif-data>.

methods prohibitively expensive. We try a simplified, computationally less expensive version of uncertainty sampling where, at every iteration, each item is compared to its two closest neighbors, but this heuristic fails spectacularly: The resulting displacement is over $5 \times$ larger than random sampling for $c = 10^6$, and is therefore not reported here (see supplementary material, Section C).

Figure 3 (right) compares the displacement of random sampling to that of the two sorting-based sampling strategies for increasing c . The adaptive sampling approaches perform systematically better. After 10^6 comparisons, the displacement of random sampling is 14 % and 23 % larger than that of Quicksort and Mergesort, respectively. Conversely, in order to reach any target displacement, Mergesort requires approximately $2 \times$ fewer comparisons than random sampling.

5 Conclusion

In this work, we demonstrate that active learning can substantially speed up the task of learning a ranking from noisy comparisons gains—both in theory and in practice. With the advent of large-scale crowdsourced ranking surveys, exemplified by GIFGIF and wiki surveys (Salganik & Levy, 2015), there is a clear need for practical AL strategies. However, existing methods are complex and computationally expensive to operate even for a reasonable number of items (a few thousands). We show that a deceptively simple idea—repeatedly sorting the items—is able to bring in all the benefits of active learning, is trivial to implement, and is computationally no more expensive than random sampling. Therefore, we believe that our method can be broadly useful for machine-learning practitioners interested in ranking problems.

Acknowledgments

We thank Holly Cogliati-Bauereis, Ksenia Konyushkova, Brunella Spinelli and the anonymous reviewers for careful proofreading and helpful comments.

References

- Ailon, N. Reconciling Real Scores with Binary Comparisons: A Unified Logistic Model for Ranking. In *Advances in Neural Information Processing Systems 21*, Vancouver, BC, Canada, 2008.
- Ailon, N. An Active Learning Algorithm for Ranking from Pairwise Preferences with an Almost Optimal Query Complexity. *Journal of Machine Learning Research*, 13 (Jan):137–164, 2012.
- Ailon, N. and Mohri, M. Preference-based learning to rank. *Machine Learning*, 80(2):189–211, 2010.
- Ailon, N., Charikar, M., and Newman, A. Aggregating Inconsistent Information: Ranking and Clustering. *Journal of the ACM*, 55(5):23, 2008.
- Alon, N., Bollobás, B., Brightwell, G., and Janson, S. Linear Extensions of a Random Partial Order. *The Annals of Applied Probability*, 4(1):108–123, 1994.
- Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. *A First Course in Order Statistics*. SIAM, 2008.
- Bradley, R. A. and Terry, M. E. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika*, 39(3/4):324–345, 1952.
- Braverman, M. and Mossel, E. Noisy sorting without resampling. In *Proceedings of SODA’08*, San Francisco, CA, 2008.
- Chen, X., Bennett, P. N., Collins-Thompson, K., and Horvitz, E. Pairwise Ranking Aggregation in a Crowdsourced Setting. In *Proceedings of WSDM’13*, Rome, Italy, 2013.
- Chu, W. and Ghahramani, Z. Extensions of Gaussian Processes for Ranking: Semi-supervised and Active Learning. In *Proceedings of the NIPS 2005 Workshop on Learning to Rank*, Whistler, BC, Canada, 2005.
- Copeland, A. H. A ‘reasonable’ social welfare function. 1951.
- Diaconis, P. and Graham, R. L. Spearman’s Footrule as a Measure of Disarray. *Journal of the Royal Statistical Society, Series B*, 39(2):262–268, 1977.
- Dubhashi, D. P. and Panconesi, A. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009.
- Elo, A. *The Rating Of Chess Players, Past & Present*. Arco, 1978.
- Hajek, B., Oh, S., and Xu, J. Minimax-optimal Inference from Partial Rankings. In *Advances in Neural Information Processing Systems 27*, Montreal, QC, Canada, 2014.
- Heckel, R., Shah, N. B., Ramchandran, K., and Wainwright, M. J. Active Ranking from Pairwise Comparisons and when Parametric Assumptions Don’t Help. preprint, arXiv:1606.08842 [cs.LG], September 2016.
- Hoare, C. A. R. Quicksort. *The Computer Journal*, 5(1): 10–16, 1962.
- Houlsby, N., Huszár, F., Ghahramani, Z., and Hernández-lobato, J. M. Collaborative Gaussian Processes for Preference Learning. In *Advances in Neural Information Processing Systems 25*, Lake Tahoe, CA, 2012.
- Jamieson, K. and Nowak, R. Active Ranking using Pairwise Comparisons. In *Advances in Neural Information Processing Systems 24*, Granada, Spain, 2011.
- Kamishima, T. and Akaho, S. Efficient Clustering for Orders. In *Mining Complex Data*, pp. 261–279. Springer, 2009.
- Knuth, D. E. *The art of computer programming: sorting and searching*, volume 3. Addison-Wesley, 2nd edition, 1998.
- MacKay, D. J. C. *Bayesian Methods for Adaptive Models*. PhD thesis, California Institute of Technology, 1992.
- Negahban, S., Oh, S., and Shah, D. Iterative Ranking from Pair-wise Comparisons. In *Advances in Neural Information Processing Systems 25*, Lake Tahoe, CA, 2012.
- Rajkumar, A. and Agarwal, S. A Statistical Convergence Perspective of Algorithms for Rank Aggregation from Pairwise Data. In *Proceedings of ICML 2014*, Beijing, China, 2014.
- Salganik, M. J. and Levy, K. E. C. Wiki Surveys: Open and Quantifiable Social Data Collection. *PLOS ONE*, 10(5): 1–17, 2015.
- Salimans, T., Paquet, U., and Graepel, T. Collaborative Learning of Preference Rankings. In *Proceedings of RecSys’12*, Dublin, Ireland, 2012.
- Schein, A. I. and Ungar, L. H. Active learning for logistic regression: an evaluation. *Machine Learning*, 68(3):235–265, 2007.
- Sedgewick, R. and Wayne, K. *Algorithms*. Addison-Wesley, 4th edition, 2011.
- Settles, B. *Active Learning*. Morgan & Claypool Publishers, 2012.

- Szörényi, B., Busa-Fekete, R., Paul, A., and Hüllermeier, E. Online Rank Elicitation for Plackett–Luce: A Dueling Bandits Approach. In *Advances in Neural Information Processing Systems 28*, Montreal, QC, Canada, 2015.
- Thurstone, L.L. A Law of Comparative Judgment. *Psychological Review*, 34(4):273–286, 1927.
- Vojnovic, M. and Yun, S. Parameter Estimation for Generalized Thurstone Choice Models. In *Proceedings of ICML 2016*, New York, NY, 2016.
- Wang, J., Srebro, N., and Evans, J. Active Collaborative Permutation Learning. In *Proceedings of KDD'14*, New York, NY, 2014.
- Yue, Y., Broder, J., Kleinberg, R., and Joachims, T. The k -armed dueling bandits problem. In *Proceedings of COLT 2009*, Montreal, QC, Canada, 2009.
- Zermelo, E. Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 29(1):436–460, 1928.