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

Supplementary Material

Lucas Maystre¹ Matthias Grossglauser¹

The supplementary material consists of three parts. In Section A, we present formal proofs for the results given in Section 3 of the main text. In Section B we show that in the Poisson model, $\Omega(n)$ comparison outcomes are necessary to discriminate between two neighboring items. In Section C, we present figures that complement the ones presented in the experimental evaluation (Section 4 of the main text).

A Proofs

Section A.1 contains the proofs of Lemmas 2 and 3. Section A.2 presents the proof for our result on the displacement of the output of a single call to Quicksort (Theorem 1), and Section A.3 that of our result on the displacement of the Copeland aggregation of multiple outputs.

A.1 Lemmas 2 and 3

We start by briefly presenting a result from graph theory that will be useful in the proof of Lemma 2. A *tournament* is a directed graph obtained by assigning a direction to every edge of a complete graph. The *score sequence* of a tournament is defined as the nondecreasing sequence of the vertices' outdegrees. The following proposition is due to Landau (1953).

Proposition 1. *Let (s_1, \dots, s_n) with $0 \leq s_1 \leq \dots \leq s_n$ be the score sequence of a tournament on n vertices. Then,*

$$\frac{k-1}{2} \leq s_k \leq \frac{n+k-2}{2} \quad \forall k \in [n].$$

We use a tournament on n vertices to represent the outcome of a comparison between each pair of items. In particular, we represent the outcome $i \prec j$ by an edge (i, j) . In this case, the outdegree of a vertex i corresponds to the number of items which “won” in a comparison against i . Note that the comparison outcomes do not need to be transitive, i.e., the tournament can contain cycles.

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

The proof of Lemma 2 is adapted from standard results on Quicksort, see, e.g., Dubhashi & Panconesi (2009, Section 3.3.3). These results are based on the fact that it is likely that the random choice of pivot leads to a well-balanced partition into subsets L and R . In our setting, the comparison outcomes do not need to be consistent with an ordering of the items, therefore we cannot use the standard argument based on the pivot's *rank*. Instead, we use the tournament representation of the comparison outcomes and analyze the pivot's *out-degree* (using Proposition 1) to ensure that the partition is balanced often enough.

Proof of Lemma 2. We show that the maximum call depth of Quicksort is at most $\lceil 48 \log n \rceil$ with high probability. The statement follows by noticing that at most n comparisons are used at each level of the call tree.

By Lemma 1, Quicksort samples a comparison outcome for each pair of items at most once. Therefore, we can represent these (a priori unobserved) pairwise outcomes as a tournament $T = ([n], A)$. At each step of the recursion, we select a pivot p uniformly at random in the set V (line 3), and compare it to the rest of the items in the set (line 5). Let T_V denote the subgraph of T induced by V . Given that the comparison outcomes follow from the edges of the tournament, L is equal to the set of incoming neighbors of p in T_V . (Correspondingly, R is equal to the set of the outgoing neighbors.) Hence, the outdegree of p in T_V determines how balanced the partition is. The probability that the outdegree of p lies in the middle half of the score sequence is $1/2$, and if it does, Proposition 1 tells us that

$$\frac{|V|-7}{8} \leq \text{outdeg}(p) \leq \frac{7|V|-5}{8}.$$

In this case, at the end of the partition $|L|$ and $|R|$ are of size at most $7|V|/8$, and in at most $\log_{8/7}(n) \leq 8 \log n$ such partitions we get to a subset of size one and reach the terminating case. Even though we do not select the pivot in the middle half every time, it is unlikely that more than $c \cdot 8 \log n$ recursions are needed (for some small constant c) to select the pivot in the middle range at least $8 \log n$ times. Let z_d i.i.d $\sim \text{Bern}(1/2)$ be the indicator variable for the event “the pivot is selected in the middle half at level of

recursion d'' . Using a Chernoff bound, we have

$$\mathbf{P} \left[\sum_{d=1}^{\lceil 48 \log n \rceil} z_d \leq 8 \log n \right] \leq \frac{1}{n^2},$$

i.e., the depth of a leaf in the call tree is at most $\lceil 48 \log n \rceil$ with probability at least $1 - 1/n^2$. As there are at most n leaves in the tree, the *maximum* depth is bounded by the same value with probability at least $1 - 1/n$. \square

In order to prove Lemma 3, we introduce some additional notation. For any $\sigma \in \mathbf{S}_n$ and $V \subseteq [n]$, let $\sigma_V : V \rightarrow \{1, \dots, |V|\}$ be the ordering induced by σ on V . We generalize the definition of displacement as

$$\Delta_V(\sigma, \tau) = \sum_{i \in V} |\sigma_V(i) - \tau_V(i)|.$$

For conciseness, we use the shorthand $\Delta_V(\sigma) \doteq \Delta_V(\sigma, \text{id})$, where id is the identity permutation.

Proof of Lemma 3. Denote by \mathcal{V} the collection of working sets that were used as input to one of the recursive calls to Quicksort. For $V \in \mathcal{V}$, let E_V be the set of pairs sampled by Quicksort to partition V and which results in an error. Note that $E_V \cap E_{V'} = \emptyset$ for $V \neq V'$, and that $\bigcup_V E_V = E$. We will show that for all $V \in \mathcal{V}$,

$$\Delta_V(\sigma) \leq \Delta_L(\sigma) + \Delta_R(\sigma) + 2 \sum_{(i,j) \in E_V} |i - j|, \quad (1)$$

where L and R are the two sets obtained at the end of the partition operation. The lemma follows by taking $V = [n]$ and recursively bounding $\Delta_L(\sigma)$ and $\Delta_R(\sigma)$.

Consider the partition operation on V , with pivot p , resulting in partitions L and R . Let $\tilde{\sigma}$ be the ordering on V that a) ranks L at the bottom, p in the middle and R at the top, and b) matches the identity permutation on L and R , i.e., $\Delta_L(\tilde{\sigma}) = \Delta_R(\tilde{\sigma}) = 0$. In a sense, $\tilde{\sigma}$ is the ordering that would be obtained if there were no further errors in the remaining recursive calls. Using the triangle inequality, we have that

$$\Delta_V(\sigma) \leq \Delta_V(\sigma, \tilde{\sigma}) + \Delta_V(\tilde{\sigma}). \quad (2)$$

By definition of $\tilde{\sigma}$, we have that

$$\begin{aligned} \Delta_V(\sigma, \tilde{\sigma}) &= \Delta_L(\sigma, \tilde{\sigma}) + \Delta_R(\sigma, \tilde{\sigma}) \\ &= \Delta_L(\sigma) + \Delta_R(\sigma), \end{aligned} \quad (3)$$

where the first equality follows from a), and the second follows from b).

Finally, we bound $\Delta_V(\tilde{\sigma})$. Let $E_V^- = \{(p, i) \in E_V : i < p\}$, and similarly $E_V^+ = \{(p, i) \in E_V : i > p\}$. Without

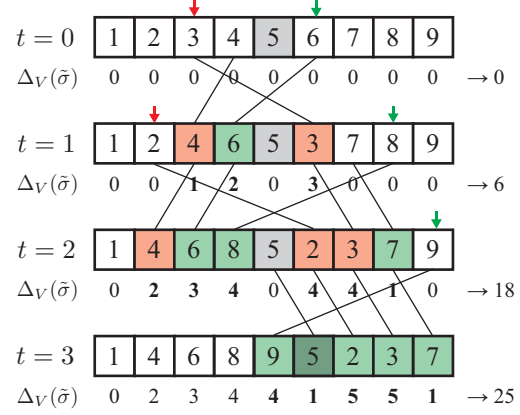


Figure 4. Illustration of the decomposition of $\Delta_V(\tilde{\sigma})$ into contributions of individual errors over a sequence of steps. In this example, $V = \{1, \dots, 9\}$, $p = 5$ and there are five errors. At step $t = 1$, we process the errors $(5, 3)$ and $(5, 6)$; at step $t = 2$, we process the errors $(5, 2)$ and $(5, 8)$, and finally, at step $t = 3$, we process the error $(5, 9)$. The shifts caused by an error are highlighted in red and green. In this example, $\Delta_V(\tilde{\sigma}) = 25 < 2 \sum_{(i,j) \in E_V} |i - j| = 26$.

loss of generality, we can assume that V consists of consecutive integers, and that $\kappa \doteq |E_V^-| \leq |E_V^+|$. We proceed as follows: starting from the ranking id_V , we progressively incorporate errors into the ranking, ending with $\tilde{\sigma}$ once all errors have been treated. To understand the impact of each error on $\Delta_V(\tilde{\sigma})$, we look at errors in the following specific sequence.

1. At steps $t = 1, \dots, \kappa$, we consider the t -th “smallest” errors in E_V^- and E_V^+ . That is, we process $(p, i) \in E_V^-$ and $(p, i') \in E_V^+$ such that $|p - i|$ and $|p - i'|$, respectively, are smallest among errors not yet treated.
2. At steps $t = \kappa + 1, \dots, |E_V^+|$, we process the remaining errors in E_V^+ , once again in increasing order of distance to p .

Figure 4 illustrates the state of the ranking at different steps on a concrete example. We start with the first case, i.e., $t \leq \kappa$. The effect of the errors (p, i) and (p, i') on $\Delta_V(\tilde{\sigma})$ is as follows.

- All items $j < i$ and $j > i'$ are not affected by the two errors: their position remains the same.
- The position of the pivot p remains the same, as the two errors balance out.
- Item i is shifted by $|p - i| + 1$ positions to the right, just right of p . Similarly, item i' is shifted by $|p - i'| + 1$ positions to the left, just left of p .
- The $|p - i| - 1$ items that are between p (excluded) and i are shifted by 1 position to the left. Similarly, the

$|p - i'| - 1$ items that are between p and i' are shifted by 1 position to the right.

Hence, the two errors contribute $2(|p - i| + |p - i'|)$ towards $\Delta_V(\tilde{\sigma})$. Now consider the second case, when $t > \kappa$. The effect of an error (p, i) is as follows.

- All items $j > i$ and all the items on the left of p are not affected by the error: their position remains the same.
- The (at most) $|p - i|$ items that are between p (included) and i are shifted by 1 position to the right.
- Item i is shifted by at most $|p - i|$ positions to the left, just left of p .

As a result, the error contributes at most $2|p - i|$ to the displacement. Adding up the contributions of all the errors, it follows that

$$\Delta_V(\tilde{\sigma}) \leq 2 \sum_{(i,j) \in E_V} |i - j|. \quad (4)$$

Combining (3) and (4) using (2) we obtain (1), which concludes the proof. \square

A.2 Theorem 1

From now on, we focus on parameters drawn from a Poisson process of rate λ , as described in (1) in the main text. We consider a worst-case scenario and assume that Quicksort samples a comparison outcome for every pair of items. Let z_{ij} be the indicator random variable of the event “the comparison between i and j resulted in an error”. By Lemma 3, we have

$$\Delta(\sigma) \leq 2 \sum_{i < j} |i - j| z_{ij} \quad (5)$$

In the following, we will bound some of the statistical properties of the random variables $\{z_{ij}\}$. We start with a lemma that bounds their mean.

Lemma 4. For any $1 \leq i < j \leq n$,

$$\mathbf{E}[z_{ij}] \leq \left(\frac{\lambda}{\lambda + 1} \right)^{j-i}.$$

Proof. Let $d_{ij} = \theta_i - \theta_j$ be the (random) distance between items i and j . This distance is a sum of $k = j - i$ independent exponential random variables, and therefore $d_{ij} \sim \text{Gamma}(k, \lambda)$. The comparison outcome is generated as per the BT model; conditioned on the distance d_{ij} , the random variable z_{ij} is a Bernoulli trial with probability $[1 + \exp(d_{ij})]^{-1}$. Therefore, we have that

$$\mathbf{E}[z_{ij}] \leq \mathbf{E}[\exp(-d_{ij})] = \left(\frac{\lambda}{\lambda + 1} \right)^k$$

\square

Next, we bound their covariance. Note that the random variables $\{z_{ij}\}$ are in general *not* unconditionally independent. They become independent only when conditioned on θ .

Lemma 5. For any $1 \leq i < j \leq n$ and any $1 \leq u < v \leq n$, let $A = \{i \dots j-1\}$ and $B = \{u \dots v-1\}$.

$$\mathbf{Cov}[z_{ij}, z_{uv}] \leq \begin{cases} 0 & \text{if } A \cap B = \emptyset, \\ \left(\frac{\lambda}{\lambda + 1} \right)^{j-i} & \text{if } A = B, \\ \left(\frac{\lambda + 1}{\lambda + 2} \right)^{j-i+v-u} & \text{otherwise.} \end{cases}$$

Proof. If A and B are disjoint, the distances d_{ij} and d_{uv} are independent random variables. Conditioned on the distances, the comparison outcomes are independent Bernoulli trials, and we conclude that z_{ij} and z_{uv} are independent. In the two remaining cases, we bound $\mathbf{E}[z_{ij}z_{uv}] \geq \mathbf{Cov}[z_{ij}, z_{uv}]$. If $A = B$, then $z_{ij} = z_{uv}$ and we have

$$\mathbf{E}[z_{ij}z_{uv}] = \mathbf{E}[z_{ij}^2] = \mathbf{E}[z_{ij}]$$

and we apply Lemma 4. Finally, if A and B are neither equivalent nor disjoint, the two comparison outcomes are independent Bernoulli trials conditioned on the distances d_{ij} and d_{uv} , but the distances are not independent. Consider the case where $i < u < j < v$. Even though d_{ij} and d_{uv} are dependent, the distances d_{iu} , d_{uj} , d_{jv} are independent Gamma random variables of rate λ and shape $u - i$, $j - u$ and $v - j$, respectively, and

$$\begin{aligned} \mathbf{E}[z_{ij}z_{uv}] &\leq \mathbf{E}[\exp\{-(d_{iu} + d_{uj}) - (d_{uj} + d_{jv})\}] \\ &= \left(\frac{\lambda}{\lambda + 1} \right)^{u-i} \left(\frac{\lambda}{\lambda + 2} \right)^{j-u} \left(\frac{\lambda}{\lambda + 1} \right)^{v-j} \\ &\leq \left(\frac{\lambda + 1}{\lambda + 2} \right)^{j-i+v-u} \end{aligned}$$

The other cases are treated analogously. \square

Lemmas 4 and 5 will be useful in proving the first part of Theorem 1. For the second part, we need a result from Ailon (2008), which characterizes the pairwise marginals of the distribution over rankings induced by Quicksort with comparisons sampled from a BT model.

Theorem 3 (Ailon, 2008, Theorem 4.1). *Let σ be the output of Quicksort using comparison outcomes sampled from BT(θ). Then, for any $i, j \in [n]$,*

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

Note that the result is non-trivial as i and j might not have been directly compared to each other: their relative position might have been deduced by transitivity from other comparison outcomes. We are now ready to prove Theorem 1.

Proof of Theorem 1. We begin with the first part of the theorem, which bounds the displacement $\Delta(\sigma)$. For clarity of exposition, we use the notation $z_{i \rightarrow k}$ instead of z_{ij} if $j = i + k$. Using (5) and Lemma 4, we can bound the expected displacement as

$$\begin{aligned} \mathbf{E}[\Delta] &\leq \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} 2k \mathbf{E}[z_{i \rightarrow k}] \\ &\leq n \sum_{k=1}^{\infty} 2k \left(\frac{\lambda}{\lambda+1} \right)^k = 2n\lambda(\lambda+1). \end{aligned}$$

In a similar way, using Lemma 5, we can bound the variance of the displacement as

$$\begin{aligned} \mathbf{Var}[\Delta] &\leq \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} 4k^2 \mathbf{Var}[z_{i \rightarrow k}] \\ &\quad + 2 \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} 2k \sum_{u=i+1}^{i+k} \sum_{\ell=1}^{n-u} 2\ell \mathbf{Cov}[z_{i \rightarrow k}, z_{u \rightarrow \ell}] \\ &\leq n \sum_{k=1}^{\infty} 4k^2 \left(\frac{\lambda}{\lambda+1} \right)^k \\ &\quad + 2n \sum_{k=1}^{\infty} 2k^2 \left(\frac{\lambda+1}{\lambda+2} \right)^k \cdot \sum_{\ell=1}^{\infty} 2\ell \left(\frac{\lambda+1}{\lambda+2} \right)^\ell \\ &\leq 1500n(\lambda^5 + 1). \end{aligned}$$

Combining the bounds for the mean and the variance with Chebyshev's inequality, we have that

$$\mathbf{P}[\Delta(\sigma) \geq 50n(\lambda^2 + 1)] \leq \lambda/n,$$

which concludes the proof of the first part of the claim.

The second part of the theorem bounds the maximum displacement for any single item. We start by showing that with high probability, there is no pair of items separated by at least $O(\lambda \log n)$ positions that is “flipped” in the output of Quicksort. Let i and j be two items such that $i < j$ and let $k = |i - j|$. Then $d_{ij} \sim \text{Gamma}(k, \lambda)$, and using a Chernoff bound we obtain

$$\mathbf{P}[d_{ij} \leq k/(e\lambda)] \leq \exp(-k/e).$$

If $k \geq 3(\lambda + 1)e \log n$, we find that

$$\mathbf{P}[d_{ij} \leq k/(e\lambda)] \leq \mathbf{P}[d_{ij} \leq 3 \log n] \leq n^{-3}. \quad (6)$$

Using the fact that the pairwise marginals of Quicksort match the pairwise comparison outcome probabilities (Theorem 3), we find

$$\begin{aligned} \mathbf{P}[\sigma(j) < \sigma(i)] &= p(j \prec i) \\ &\leq \exp(-3 \log n) = n^{-3}. \end{aligned} \quad (7)$$

Combining (6) and (7), and using a union bound over the $\binom{n}{2}$ pairs, we see that with probability $1 - 1/n$ there is no pairs of items (i, j) separated by at least $3(\lambda + 1)e \log n$ position with $i < j$ but $\sigma(j) < \sigma(i)$. Finally, suppose that there is an i such $|\sigma(i) - i| = k$. Without loss of generality, we can assume that $i < \sigma(i)$. This means that there are k items larger than i that are on the left of i in σ . In particular, there is an item $j > i$ such that $|i - j| \geq k$ and $\sigma(j) < \sigma(i)$. This concludes the proof. \square

A.3 Theorem 2

In order to prove Theorem 2, we first need a basic result on the order statistics of exponential random variables. Let x_1, \dots, x_n , be i.i.d. exponential random variables of rate λ . Let $x_{(1)}, \dots, x_{(n)}$ be their order statistics, i.e., the random variables arranged in increasing order. Then,

$$x_{(i)} = \sum_{j=1}^i \frac{1}{n-j+1} y_j, \quad (8)$$

where y_1, \dots, y_n are i.i.d. exponential random variables of rate λ (see, e.g., Arnold et al., 2008, Section 4.6).

Proof of Theorem 2. We consider the order statistics of the $n - 1$ i.i.d. exponential random variables x_1, \dots, x_{n-1} which define the distances between neighboring items. Let $\hat{n} = \lceil n/\log^2 n \rceil$, and denote by $B \subset [n]$ the set of items at both ends of $x_{(1)}, \dots, x_{(\hat{n}-1)}$. These “bad” items are close to their nearest neighbor, and we simply invoke Theorem 1 to claim that each of these items is shifted by at most $O(\lambda \log n)$ positions with high probability. Consider now the “good” items, i.e., those in $G = [n] \setminus B$. Using (8) and for n large enough,

$$\begin{aligned} \mathbf{P}[x_{(\hat{n})} \leq 1/(e\lambda \log^2 n)] \\ &\leq \mathbf{P}\left[\sum_{j=1}^{\hat{n}} y_j/n \leq 1/(e\lambda \log^2 n)\right] \\ &\leq \exp(-\hat{n}/e) \leq 1/n. \end{aligned}$$

The second-to-last inequality follows from a Chernoff bound similar to that used in the proof of Theorem 1. Therefore, with high probability all items in G are at distance larger than $c/(\lambda \log^2 n)$ from their nearest neighbor.

We will now show that after $m = O(\lambda^2 \log^5 n)$ runs of Quicksort, $\hat{\sigma}(i) = i$ with high probability for all $i \in G$. Let $i \in G$, $j \in [n]$ be a pair of items, and without loss of generality assume that $i < j$. Let t_k be the indicator random variable for the event “ $\sigma(i) < \sigma(j)$ in the k -th run of Quicksort”, and let $p = \mathbf{P}[t_k = 1]$. Then, using

Theorem 3,

$$\begin{aligned} p - \frac{1}{2} &= p(i \prec j) - \frac{1}{2} = \frac{1 - \exp(-d_{ij})}{2[1 + \exp(-d_{ij})]} \\ &\geq \frac{1 - \exp[-1/(e\lambda \log^2 n)]}{4} \\ &\geq \frac{1}{8e\lambda \log^2 n} \end{aligned}$$

with high probability. In the last inequality, we used the fact that $1 - e^{-x} \geq x/2$ for $x \in [0, 1]$. The random variables t_1, \dots, t_m are independent Bernoulli trials, and using a Chernoff bound we obtain

$$\begin{aligned} \mathbf{P}[\hat{\sigma}(j) < \hat{\sigma}(i)] &= \mathbf{P}\left[\sum_{k=1}^m t_k \leq n/2\right] \\ &\leq \exp[-2m(p - 1/2)^2] \leq \exp\left[-\frac{m}{32e^2\lambda^2 \log^4 n}\right]. \end{aligned}$$

With $m = 96e^2\lambda^2 \log^5 n$, we have $\mathbf{P}[\hat{\sigma}(j) < \hat{\sigma}(i)] \leq n^{-3}$, and using a union bound we see that with probability $1 - 1/n$ we have $\hat{\sigma}(i) = i$ for all $i \in G$. Therefore, the total displacement is

$$\begin{aligned} \Delta(\hat{\sigma}) &= \sum_{i \in B} |\hat{\sigma}(i) - i| \leq |B| \cdot 3(\lambda + 1)e \log n \\ &= O(\lambda n / \log n). \end{aligned}$$

This concludes the proof. \square

B Discriminating the Closest Items

The distance between the two closest items is $d_{\min} = \min_i |\theta_{i+i} - \theta_i| = \min_i x_i$, i.e., the minimum of $n - 1$ independent exponential random variables of rate λ . Therefore, $d_{\min} \sim \text{Exp}((n-1)\lambda)$, and for $n \geq 2$ with probability at least $1 - e^{-1/2} \approx 0.39$ we have $d_{\min} \leq (\lambda n)^{-1}$. Suppose that we compare the two closest items m times, and let z_i be the indicator random variable for the event “the outcome of the i -th comparison is incorrect”. Assuming that $d_{\min} \leq (\lambda n)^{-1}$ and that $\lambda n \geq 1/2$,

$$\begin{aligned} \mathbf{P}[z_i = 0] &\leq \frac{1}{1 + \exp[-1/(\lambda n)]} \leq \frac{1}{2 - 1/(\lambda n)} \\ &= \frac{1}{2} \cdot \left(1 + \frac{1}{2\lambda n - 1}\right) \\ &\leq \frac{1}{2} \exp\left[\frac{1}{2\lambda n - 1}\right], \end{aligned}$$

where we used the inequality $e^x \geq 1 + x$ twice. Given the m comparison outcomes, we use a majority vote to decide the relative order of the two items. The probability of making

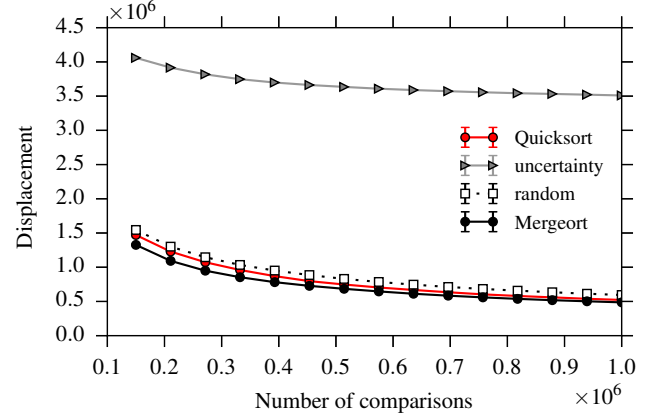


Figure 5. Results on the GIFGIF dataset. The experiment is repeated 10 times, and we report the mean and the standard deviation. The variant of uncertainty sampling performs extremely poorly.

the correct decision is

$$\begin{aligned} \mathbf{P}\left[\sum_{i=1}^m z_i \leq m/2\right] &\leq \sum_{k=1}^{m/2} \binom{m}{k} \mathbf{P}[z_i = 0]^m \\ &\leq \exp\left[\frac{m}{2\lambda n - 1}\right] \cdot 2^{-m} \sum_{k=1}^{m/2} \binom{m}{k} \\ &= \frac{1}{2} \exp\left[\frac{m}{2\lambda n - 1}\right]. \end{aligned}$$

Therefore, if $m = o(\lambda n)$ the probability of making a mistake is bounded from below by a positive constant.

C Additional Figures

In this section, we present a few additional figures that complement the ones presented in Section 4 of the main text.

Figure 5 presents the results on the GIFGIF dataset including a variant of uncertainty sampling. This variant samples, at each iteration, $n - 1$ comparisons consisting of adjacent pairs in the ranking $\hat{\theta}$. This strategy performs surprisingly poorly.

Figure 6 presents results on synthetic datasets with $n = 200$ and $\lambda \in \{1, 2, 5, 10\}$. For the reader’s convenience, we plot every graph on both a linear and a logarithmic scale. Unsurprisingly, the gains of adaptive sampling are greater when the noise is smaller.

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.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. *A First Course in Order Statistics*. SIAM, 2008.

Dubhashi, D. P. and Panconesi, A. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009.

Landau, H. G. On Dominance Relations and the Structure of Animal Societies: III The Condition for a Score Structure. *Bulletin of Mathematical Biophysics*, 15(2): 143–148, 1953.

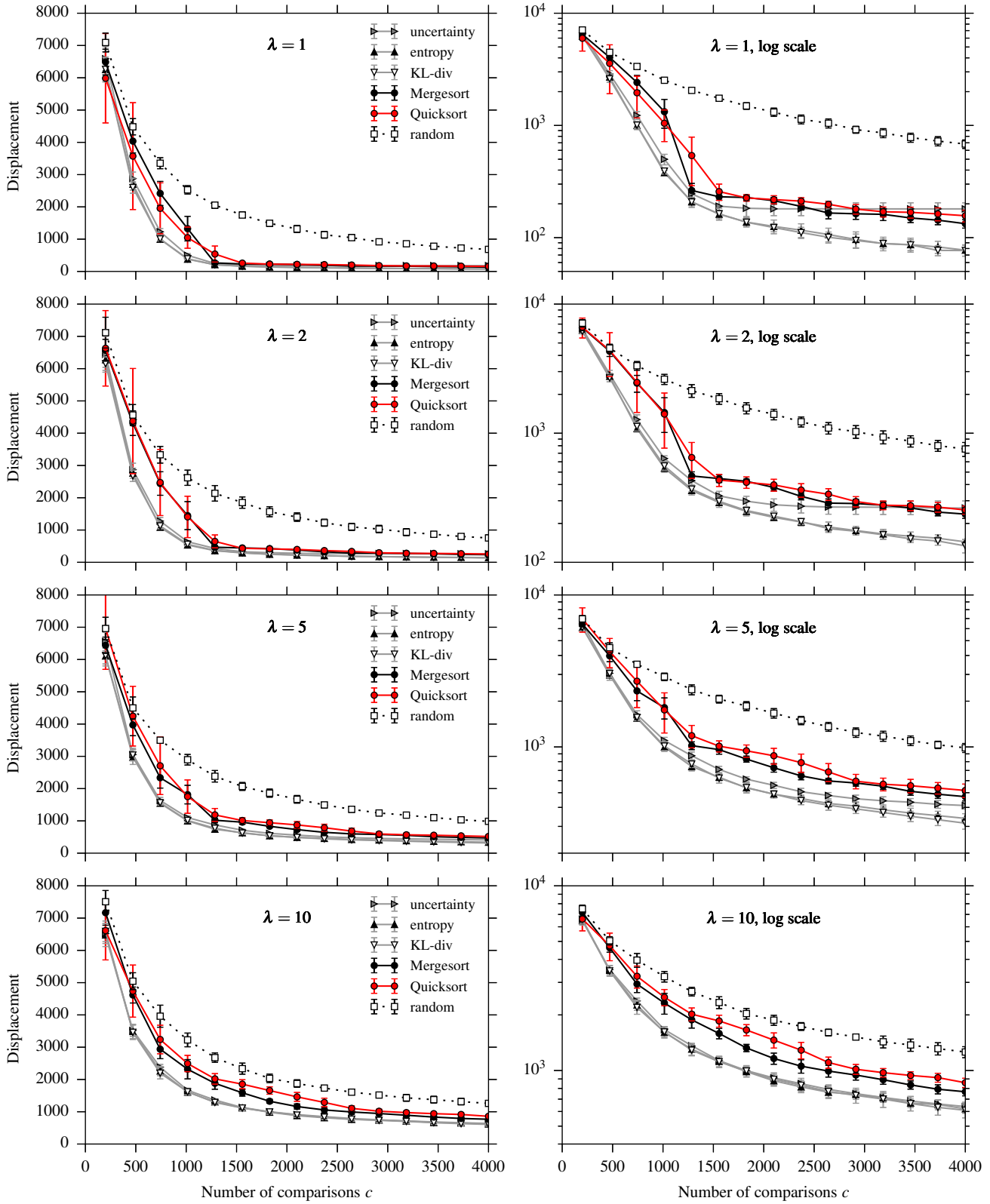


Figure 6. Results on synthetic datasets for $n = 200$ and increasing values of λ . Every experiment is repeated 10 times, and we report the mean and the standard deviation.