# Efficient Ranking from Pairwise Comparisons - Supplementary Material 

## 1. SVM has Optimal Sample Complexity

In this section we will show that the SVM, applied to ranking (as described in Section 3) has an $O(n)$ sample complexity. A related claim (without complete proof) has been made in (Radinsky \& Ailon, 2011). We then show that this sample complexity is tight.
Proposition 3.1. There is a constant d, so that for any $0<\eta<1$, if we noiselessly measure dn/ $\eta^{2}$ binary comparisons, chosen uniformly at random with replacement, and $n>n_{0}$ is large enough, the SVM will produce a prediction $\hat{\pi}$, which satisfies

$$
\begin{equation*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) \leq \frac{\eta}{2}\binom{n}{2} \tag{1}
\end{equation*}
$$

Proof. First, note that in the noiseless case the datapoints $x_{i, j}$ with associated labels $\bar{y}_{i, j}$ of Section 3 are linearly separable. The SVM finds such a separator. Since the measured comparisons are chosen uniformly at random with replacement, we can prove the outcome by appealing to learning-theoretic generalization bounds. Using results in (for example) Bousquet et al. (2003) and a VC dimension bound of (Radinsky \& Ailon, 2011), we have the following Lemma:

Lemma. For $\delta>0$, and $0<\eta<1$, if we noiselessly measure $d n / \eta^{2}$ binary comparisons, chosen uniformly at random with replacement, then for some constant $c$, with probability at least $1-\delta$ the $S V M$ produces a permutation $\hat{\pi}$ with

$$
\begin{equation*}
\operatorname{inv}(\hat{\pi}) \leq \eta\left[\frac{c}{\sqrt{d}}+\sqrt{\frac{2 \log \left(\frac{2}{\delta}\right)}{d n}}\right]\binom{n}{2} \tag{2}
\end{equation*}
$$

To use this Lemma, first define

$$
\begin{equation*}
t=\eta\left[\frac{c}{\sqrt{d}}+\sqrt{\frac{2 \log \left(\frac{2}{\delta}\right)}{d n}}\right]\binom{n}{2} . \tag{3}
\end{equation*}
$$

Then by the Lemma $\mathbb{P}(\operatorname{inv}(\hat{\pi})>t) \leq \delta$. Notice that if we plug in $\delta=1$ into $t$, we get a value $t_{1}$ for which $\mathbb{P}\left(\operatorname{inv}(\hat{\pi})>t_{1}\right) \leq \delta=1$

$$
\begin{equation*}
t_{1}=\eta\left[\frac{c}{\sqrt{d}}+\sqrt{\frac{2 \log (2)}{d n}}\right]\binom{n}{2} . \tag{4}
\end{equation*}
$$

We will thus assume that for all $t \leq t_{1}$, we have $\mathbb{P}(\operatorname{inv}(\hat{\pi})>t) \leq 1$. We can use this result to upper bound $\mathbb{E}(\operatorname{inv}(\hat{\pi}))$ as follows. Since $\operatorname{inv}(\hat{\pi}) \geq 0$,

$$
\begin{equation*}
\mathbb{E}(\operatorname{inv}(\hat{\pi}))=\int_{0}^{\infty} \mathbb{P}(\operatorname{inv}(\hat{\pi})>t) d t \leq t_{1}+\int_{t_{1}}^{\infty} \mathbb{P}(\operatorname{inv}(\hat{\pi})>t) d t \tag{5}
\end{equation*}
$$

All that remains is to express $\delta$ in terms of $t$. Rewriting, we find that

$$
\begin{align*}
\delta & =2 \exp \left\{-\frac{1}{2 \sigma_{n}^{2}}\left(t-\mu_{n}\right)^{2}\right\}  \tag{6}\\
\sigma_{n}^{2} & =\frac{\eta^{2}(n(n-1))^{2}}{4 d n} \quad \mu_{n}=\eta \frac{c n(n-1)}{2 \sqrt{d}} \tag{7}
\end{align*}
$$

Returning to our original expectation,

$$
\begin{align*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) & \leq t_{1}+\int_{t_{1}}^{\infty} \delta d t  \tag{8}\\
& =t_{1}+2 \sqrt{2 \pi \sigma_{n}^{2}} \int_{t_{1}}^{\infty} \frac{1}{\sqrt{2 \pi \sigma_{n}^{2}}} \exp \left\{-\frac{1}{2 \sigma_{n}^{2}}\left(t-\mu_{n}\right)^{2}\right\} d t  \tag{9}\\
& =t_{1}+2 \sqrt{2 \pi \sigma_{n}^{2}} \mathbb{P}\left(\sigma_{n} z+\mu_{n}>t_{1}\right)  \tag{10}\\
& \leq t_{1}+2 \sqrt{2 \pi \sigma_{n}^{2}}  \tag{11}\\
& =\eta\left[\frac{c}{\sqrt{d}}+\sqrt{\frac{2 \log (2)}{d n}}\right]\binom{n}{2}+2 \eta \sqrt{2 \pi} \frac{n(n-1)}{2 \sqrt{d n}}  \tag{12}\\
& =\eta\left[\frac{c}{\sqrt{d}}+\frac{\sqrt{2 \log (2)}+\sqrt{8 \pi}}{\sqrt{d n}}\right]\binom{n}{2} \tag{13}
\end{align*}
$$

Suppose we set $d=16 c^{2}$. Then for any $\eta / 4>\epsilon_{0}>0$, there is an $n_{0}$ so that if $n>n_{0}$,

$$
\begin{equation*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) \leq\left[\frac{\eta}{4}+\epsilon_{0}\right]\binom{n}{2} \leq \frac{\eta}{2}\binom{n}{2} \tag{14}
\end{equation*}
$$

Next we prove the following proposition via results of Giesen et al. (Giesen et al., 2009) to show that the sample complexity in Proposition 3.1 is effectively tight.
Proposition 3.2. For $\eta<1$, any randomized, comparison-based algorithm that produces for all $\pi^{*}$ a prediction $\hat{\pi}$ with an expected risk of

$$
\begin{equation*}
\mathbb{E}(i n v(\hat{\pi})) \leq \frac{\eta}{2}\binom{n}{2} \tag{15}
\end{equation*}
$$

must on expectation use at least $\Omega(n)$ comparisons in the worst case.
Proof. Recall that for two permutations $\hat{\pi}, \pi^{*}$, Spearman's footrule distance is defined to be

$$
\begin{equation*}
D(\hat{\pi})=\sum_{j=1}^{n}\left|\hat{\pi}(j)-\pi^{*}(j)\right| \tag{16}
\end{equation*}
$$

Giesen et al. (2009) give the following theorem (we restate slightly):
Theorem (Giesen et al. (2009)). Any randomized, comparison-based algorithm that produces for each input permutation $\pi^{*}$ a prediction $\hat{\pi}$ with expected Spearman's footrule distance $D(\hat{\pi})$ of at most $n^{2} / \nu(n)$, must on expectation use at least $n(\min (\log (\nu(n)), \log (n))-6)$ comparisons in the worst case.

Because Spearman's footrule can be upper bounded by twice Kendall's tau distance (Diaconis \& Graham, 1977) (i.e. $D(\hat{\pi}) \leq 2 \operatorname{inv}(\hat{\pi})$ ) we have that if

$$
\begin{equation*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) \leq \frac{\eta}{2}\binom{n}{2} \tag{17}
\end{equation*}
$$

then

$$
\begin{equation*}
\mathbb{E}(D(\hat{\pi})) \leq \mathbb{E}(2 \operatorname{inv}(\hat{\pi})) \leq \eta\binom{n}{2}<\eta \frac{n^{2}}{2} \tag{18}
\end{equation*}
$$

Combining with the above theorem of Giesen et al. (Giesen et al., 2009), we have the desired result.

## 2. Proofs for BRE and URE

### 2.1. Proof of Theorem 4.1

Balanced Rank Estimation Algorithm (BRE): Measure each binary comparison independently with probability $m(n) / n$. Define the scores

$$
\begin{equation*}
\hat{\Pi}(j)=\frac{\sum_{i \neq j} s_{i, j}\left(2 \bar{c}_{i, j}-1\right)}{2 m(n)} \propto \sum_{i \neq j} s_{i, j}\left(2 \bar{c}_{i, j}-1\right) \tag{19}
\end{equation*}
$$

Predict $\pi^{*}$ by the ordering $\hat{\pi}$ of the estimated scores, breaking ties randomly.
Theorem 4.1. For any fixed $0<\eta<1$ there is a constant $c(p, \eta) \in \Theta\left(1 /(2 p-1)^{2} \eta^{2}\right)$ so that if $m(n) / n \geq$ $c(p, \eta) / n$ and $n>n_{0}$, the Balanced Rank Estimation Algorithm satisfies

$$
\begin{equation*}
\mathbb{E}(i n v(\hat{\pi})) \leq \frac{\eta}{2}\binom{n}{2} \tag{20}
\end{equation*}
$$

Proof. Without loss of generality, we suppose that $\pi^{*}=(1,2, \ldots, n)$, so that $\pi^{*}(j)=j$. The true (rescaled) rank score of object $j$ is denoted by

$$
\begin{equation*}
\Pi^{*}(j)=\frac{1}{n} \sum_{i \neq j} c_{i, j}=\frac{j-1}{n} \tag{21}
\end{equation*}
$$

Knowing the true scores $\Pi^{*}$ clearly suffices to predict the permutation $\pi^{*}$ with 0 Kendall tau distance. Equivalently, it suffices to know the true scores up to global scaling and additive constants. We will show that the Balanced Rank Estimation Algorithm produces an unbiased estimate of the scores

$$
\begin{equation*}
\tilde{\Pi}^{*}(j)=\Pi^{*}(j)(2 p-1)+\left(1-p-\frac{1}{2}\right) \frac{n-1}{n} \tag{22}
\end{equation*}
$$

where $p$ parameterizes the observation model. We will then show that for large enough $m(n)$, the probability that we predict an incorrect binary comparison (i.e. we find that $\hat{\Pi}(j)<\hat{\Pi}(i)$ even though $i<j$ ) will decay fast enough with growing $|j-i|$ to guarantee that the expected Kendall tau distance achieves the target. Our interest lies in computing

$$
\begin{align*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) & =\mathbb{E}\left(\sum_{i<j} \mathbf{1}(\hat{\Pi}(j) \leq \hat{\Pi}(i))\right)  \tag{23}\\
& =\sum_{i<j} \mathbb{P}\left(\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right] \geq\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]\right) \tag{24}
\end{align*}
$$

Bernstein concentration: Introduce binary random variables $b_{i, j}$ that capture if a comparison was flipped or not. Let $b_{i, j}=1$ indicate that the measurement was not flipped. By assumption, the $b_{i, j}$ are i.i.d. and $\mathbb{P}\left(b_{i, j}=1\right)=p$. Hence, we can write $\bar{c}_{i, j}=b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)$. We study the difference between the estimated scores and the target scores:

$$
\begin{align*}
& \hat{\Pi}(j)-\tilde{\Pi}^{*}(j)=\hat{\Pi}(j)-\left(\Pi^{*}(j)(2 p-1)+\left(1-p-\frac{1}{2}\right) \frac{n-1}{n}\right)  \tag{25}\\
= & \frac{1}{n} \sum_{i \neq j}\left[\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right] \tag{26}
\end{align*}
$$

Note that given $\pi^{*}$, the variables in the sum are independent. Some algebra reveals:

$$
\begin{align*}
& \mathbb{E}\left(\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right)=0  \tag{27}\\
& \operatorname{var}\left(\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right) \leq \frac{n}{4 m(n)}  \tag{28}\\
&\left|\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right| \leq \frac{n+m(n)}{2 m(n)} . \tag{29}
\end{align*}
$$

The difference $\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right]$ can be written as a sum of $2 n-3$ independent terms with magnitude no more than twice and variance no more than four times that stated above. Thus, we can derive a Bernstein concentration inequality (Boucheron et al., 2004). Conditioned on $\pi^{*}$, one can show that

$$
\begin{align*}
& \mathbb{P}\left(\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right] \geq\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]\right)  \tag{30}\\
= & \mathbb{P}\left(\frac{2 m(n)}{2(n+m(n))} \frac{n}{2 n-3}\left(\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right]\right) \geq \frac{2 m(n)}{2(n+m(n))} \frac{n}{2 n-3}\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]\right)  \tag{31}\\
\leq & \exp \left\{-\frac{(2 n-3)\left(\frac{n}{2 n-3}\right)^{2}\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]^{2}}{2 \frac{(n+m(n))^{2}}{m(n)^{2}}\left(\frac{4 n}{4 m(n)} \frac{m(n)^{2}}{(n+m(n))^{2}}+\frac{\left(\frac{n}{2 n-3}\right)\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]}{3} \frac{m(n)}{n+m(n)}\right)}\right\}  \tag{32}\\
\leq & \exp \left\{-\left[\frac{j-i}{n}\right]^{2} \frac{3}{32}(2 p-1)^{2} m(n)\right\}, \tag{33}
\end{align*}
$$

where we used the fact that we must eventually have $n>3$ as well as $m(n)<n-3$ if $n>n_{0}$. Returning to the expected Kendall tau distance of Eq. (24), we may bound

$$
\begin{align*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) & \leq \sum_{i<j} \exp \left\{-\left[\frac{j-i}{n}\right]^{2} \frac{3}{32}(2 p-1)^{2} m(n)\right\}  \tag{34}\\
& =\sum_{k=1}^{n-1}(n-k) \exp \left\{-\left[\frac{k}{n}\right]^{2} \frac{3}{32}(2 p-1)^{2} m(n)\right\}  \tag{35}\\
& \leq \int_{0}^{n}(n-k) \exp \left\{-\left[\frac{k}{n}\right]^{2} \frac{3}{32}(2 p-1)^{2} m(n)\right\} d k  \tag{36}\\
& =n^{2} \frac{\sqrt{\pi \frac{3}{32}(2 p-1)^{2} m(n)} \operatorname{erf}\left(\sqrt{\frac{3}{32}(2 p-1)^{2} m(n)}\right)+\exp \left\{-\frac{3}{32}(2 p-1)^{2} m(n)\right\}-1}{2 \frac{3}{32}(2 p-1)^{2} m(n)}  \tag{37}\\
& \leq \frac{n}{n-1} \sqrt{\frac{128}{3}} \frac{2}{2(2 p-1) \sqrt{m(n)}}\binom{n}{2} . \tag{38}
\end{align*}
$$

For this bound to be no larger than $\frac{\eta}{2}\binom{n}{2}$, we need

$$
\begin{equation*}
\eta>\frac{n}{n-1} \sqrt{\frac{128}{3}} \frac{2}{(2 p-1) \sqrt{m(n)}} \tag{39}
\end{equation*}
$$

so that a suitable $m(n)$ exists which satisfies $m(n) \in \Theta\left(1 /\left((2 p-1)^{2} \eta^{2}\right)\right)$.

### 2.2. Proof of Theorem 4.2

Balanced Rank Estimation Algorithm (BRE): Measure each binary comparison independently with probability $m(n) / n$. Define the scores

$$
\begin{equation*}
\hat{\Pi}(j)=\frac{\sum_{i \neq j} s_{i, j}\left(2 \bar{c}_{i, j}-1\right)}{2 m(n)} \propto \sum_{i \neq j} s_{i, j}\left(2 \bar{c}_{i, j}-1\right) \tag{40}
\end{equation*}
$$

Predict $\pi^{*}$ by the ordering $\hat{\pi}$ of the estimated scores, breaking ties randomly.
Theorem 4.2. For $c>0$, if each comparison is measured with probability $m(n) / n=c \log (n) / n$, Balanced Rank Estimation produces with probability at least $1-2 n^{1-a_{n} \frac{3}{8}(2 p-1)^{2} \nu^{2} c}$ a permutation $\hat{\pi}$ with

$$
\begin{equation*}
\max _{j}\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq \nu n \tag{41}
\end{equation*}
$$

where $a_{n}$ is a sequence with $a_{n} \rightarrow 1$.

Proof. Without loss of generality, we suppose that $\pi^{*}=(1,2, \ldots, n)$, so that $\pi^{*}(j)=j$. The true (rescaled) rank score of object $j$ is denoted by

$$
\begin{equation*}
\Pi^{*}(j)=\frac{1}{n} \sum_{i \neq j} c_{i, j}=\frac{j-1}{n} \tag{42}
\end{equation*}
$$

Knowing the true scores $\Pi^{*}$ clearly suffices to predict the permutation $\pi^{*}$ with 0 Kendall tau distance. Equivalently, it suffices to know the true scores up to global scaling and additive constants. We will show that the Balanced Rank Estimation Algorithm produces an unbiased estimate of the scores

$$
\begin{equation*}
\tilde{\Pi}^{*}(j)=\Pi^{*}(j)(2 p-1)+\left(1-p-\frac{1}{2}\right) \frac{n-1}{n} . \tag{43}
\end{equation*}
$$

In particular, we will show that for a large enough $m(n)$, the difference $\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|$ is with high probability small for all $j$ and that we will thus not confuse the relative ordering of two objects that are further than twice this difference apart in $\pi^{*}$.

Bernstein concentration: Introduce binary random variables $b_{i, j}$ that capture if a comparison was flipped or not. Let $b_{i, j}=1$ indicate that the measurement was not flipped. We assume that $\mathbb{P}\left(b_{i, j}=1\right)=p$ and that the $b_{i, j}$ are i.i.d. Hence, we can write $\bar{c}_{i, j}=b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)$. As in Theorem 4.2, we study the difference between the estimated scores and the rescaled and translated target scores:

$$
\begin{align*}
& \hat{\Pi}(j)-\tilde{\Pi}^{*}(j)=\hat{\Pi}(j)-\left(\Pi^{*}(j)(2 p-1)+\left(1-p-\frac{1}{2}\right) \frac{n-1}{n}\right)  \tag{44}\\
= & \frac{1}{n} \sum_{i \neq j}\left[\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right] \tag{45}
\end{align*}
$$

Note that given $\pi^{*}$, the variables in the sum are independent. The following result, which were previously used in Theorem 4.2, can be shown

$$
\begin{align*}
\mathbb{E}\left(\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right)=0  \tag{46}\\
\operatorname{var}\left(\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right) \leq \frac{n}{4 m(n)}  \tag{47}\\
\quad\left|\frac{n}{2 m(n)} s_{i, j}\left(2\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-1\right)-\left(c_{i, j}(2 p-1)+\left(1-p-\frac{1}{2}\right)\right)\right| \leq \frac{n+m(n)}{2 m(n)} . \tag{48}
\end{align*}
$$

Thus, we can derive a Bernstein concentration inequality (Boucheron et al., 2004). Conditioned on $\pi^{*}$,

$$
\begin{align*}
\mathbb{P}\left(\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>t\right) & =\mathbb{P}\left(\left|\frac{n}{n-1} \frac{2 m(n)}{n+m(n)} \hat{\Pi}(j)-\frac{n}{n-1} \frac{2 m(n)}{n+m(n)} \tilde{\Pi}^{*}(j)\right|>\frac{n}{n-1} \frac{2 m(n)}{n+m(n)} t\right)  \tag{49}\\
& \leq 2 \exp \left\{-\frac{(n-1) t^{2}\left(\frac{n}{n-1}\right)^{2}}{2 \frac{(n+m(n))^{2}}{4 m(n)^{2}}\left(\frac{n}{4 m(n)} \frac{4 m(n)^{2}}{(n+m(n))^{2}}+\frac{t}{3} \frac{n}{n-1} \frac{2 m(n)}{n+m(n)}\right)}\right\}  \tag{50}\\
& \leq 2 \exp \left\{-\frac{n}{n+m(n)} \frac{t^{2} 4 m(n)}{2\left(1+\frac{2 t}{3}\right)}\right\} . \tag{51}
\end{align*}
$$

Since there are $n$ items to be sorted, we apply a union bound

$$
\begin{equation*}
\mathbb{P}\left(\exists j:\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>t\right) \leq 2 \exp \left\{-\frac{n}{n+m(n)} \frac{t^{2} 4 m(n)}{2\left(1+\frac{2 t}{3}\right)}+\log (n)\right\} \tag{52}
\end{equation*}
$$

The concentration result tells us that the relative ordering of two objects that are far apart in the true ordering (large $t$ ) is harder to confuse than that of nearby objects (small $t$ ). Thus, as $n$ gets large, the relative ordering of any sufficiently well-separated pair in $\pi^{*}$ should with high probability be predicted correctly in $\hat{\pi}$. Specifically, we have the following Lemma.

Lemma 4.3. For some $a>0$ and $b \in \mathbb{R}$, if $\forall j,\left|\hat{\Pi}(j)-\left(\Pi^{*}(j) a+b\right)\right| \leq t$, then $\forall j,\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq 2$ tn $/ a$.
Proof. We have $\hat{\pi}(j) \neq \pi^{*}(j)$ when one or more elements in $\hat{\Pi}$ are mapped to the wrong side of $\hat{\Pi}(j)$. Equivalently, to bound $\left|\hat{\pi}(j)-\pi^{*}(j)\right|$ we can count how many elements of $\hat{\Pi}$ can at most map to the same value $\hat{\Pi}(j)$ and to assume that the sorting algorithm breaks ties in the least favorable way. Note that

$$
\begin{equation*}
\frac{n}{a}\left|\hat{\Pi}(j)-\left(\Pi^{*}(j) a+b\right)\right|=\left|\frac{n}{a} \hat{\Pi}(j)-\left(j-1+\frac{n b}{a}\right)\right|<\frac{t n}{a} \tag{53}
\end{equation*}
$$

Hence, $\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq 2 t n / a$.
Putting it together: By the definition of the mean score $\tilde{\Pi}^{*}$ in Eq. (43), we see that we need $a=(2 p-1)$ for Lemma 4.3. Then, in order to show that $\forall j,\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq \nu n$ with high probability, we need that with high probability $\forall j,\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right| \leq(2 p-1) \nu / 2$. Looking at Eq. (52), we can achieve this if we let $m(n) \geq c(p, \nu) \log (n)$, for a sufficiently large constant $c(p, \nu)$. Then, as $n \rightarrow \infty$, with high probability $\forall j$, $\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq \nu n$.

Probability of success. The probability that the preconditions to Lemma 4.3 hold depend on the constant $c(p, \nu)$. Specifically,

$$
\begin{align*}
& \mathbb{P}\left(\forall j:\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|<\frac{(2 p-1) \nu}{2}\right)  \tag{54}\\
\geq & 1-2 \exp \left\{\left[1-\frac{n}{n+c(p, \nu) \log (n)} \frac{((2 p-1) \nu)^{2} c(p, \nu)}{2\left(1+\frac{(2 p-1) \nu}{3}\right)}\right] \log (n)\right\}  \tag{55}\\
\geq & 1-2 \exp \left\{\left[1-\frac{n}{n+c(p, \nu) \log (n)} \frac{3}{8}((2 p-1) \nu)^{2} c(p, \nu)\right] \log (n)\right\}  \tag{56}\\
= & 1-2 n^{1-a_{n} \frac{3}{8}(2 p-1)^{2} \nu^{2} c(p, \nu)}, \tag{57}
\end{align*}
$$

where $a_{n}=n /(n+c(p, \nu) \log (n)) \rightarrow 1$.

### 2.3. Proof of Theorem 4.5

Unbalanced Rank Estimation Algorithm (URE): Measure each binary comparison independently with probability $m(n) / n$. Define the scores

$$
\begin{equation*}
\hat{\Pi}(j)=\frac{1}{m(n)} \sum_{i \neq j} s_{i, j} \bar{c}_{i, j} \propto \sum_{i \neq j} s_{i, j} \bar{c}_{i, j} \tag{58}
\end{equation*}
$$

Predict $\pi^{*}$ by the ordering $\hat{\pi}$ of the estimated scores, breaking ties randomly.
Theorem 4.5. For any fixed $0<\eta<1$, there is a constant $c(p, \eta) \in \Theta\left(1 /\left((2 p-1)^{2} \eta^{2}\right)\right)$ so that if each comparison is measured with probability at least $m(n) / n \geq c(p, \eta) / n$, the Unbalanced Rank Estimation Algorithm satisfies

$$
\begin{equation*}
\mathbb{E}(i n v(\hat{\pi})) \leq \frac{\eta}{2}\binom{n}{2} \tag{59}
\end{equation*}
$$

Proof. Without loss of generality, we suppose that $\pi^{*}=(1,2, \ldots, n)$, so that $\pi^{*}(j)=j$. The true (rescaled) rank of object $j$ is denoted by

$$
\begin{equation*}
\Pi^{*}(j)=\frac{1}{n} \sum_{i \neq j} c_{i, j}=\frac{j-1}{n} \tag{60}
\end{equation*}
$$

It suffices to know the true ranking scores up to global scaling and additive constants. One can show that the Unbalanced Rank Estimation Algorithm produces an unbiased estimate of the scores

$$
\begin{equation*}
\tilde{\Pi}^{*}(j)=\Pi^{*}(j)(2 p-1)+(1-p) \frac{n-1}{n} \tag{61}
\end{equation*}
$$

We will show that for a large enough $m(n)$, the probability that we predict an incorrect binary comparison (i.e. we find that $\hat{\Pi}(j)<\hat{\Pi}(i)$ even though $i<j)$ will decay fast enough with growing $|j-i|$ to guarantee that the expected Kendall tau distance achieves the target. Our interest lies in upper bounding

$$
\begin{align*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) & =\mathbb{E}\left(\sum_{i<j} \mathbf{1}(\hat{\Pi}(j) \leq \hat{\Pi}(i))\right)  \tag{62}\\
& =\sum_{i<j} \mathbb{P}\left(\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right] \geq\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]\right) \tag{63}
\end{align*}
$$

Bernstein concentration: We introduce binary random variables $b_{i, j}$ which encode whether or not a comparison was flipped. Let $b_{i, j}=1$ indicate that the measurement was not flipped. By assumption, the $b_{i, j}$ are i.i.d. and $\mathbb{P}\left(b_{i, j}=1\right)=p$. With this, we can write $\bar{c}_{i, j}=b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)$ and so

$$
\begin{align*}
& \hat{\Pi}(j)-\tilde{\Pi}^{*}(j)=\hat{\Pi}(j)-\left(\Pi^{*}(j)(2 p-1)+(1-p) \frac{n-1}{n}\right)  \tag{64}\\
= & \frac{1}{n} \sum_{i \neq j}\left[\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right] . \tag{65}
\end{align*}
$$

Given $\pi^{*}$, the random variables inside the sum are independent and one can show that

$$
\begin{align*}
& \mathbb{E}\left(\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right)=0  \tag{66}\\
& \frac{1}{n-1} \sum_{i \neq j} \operatorname{var}\left(\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right)  \tag{67}\\
\leq & \frac{n}{m(n)} \frac{1}{n-1}[(j-1) p+(n-j)(1-p)]  \tag{68}\\
& \left|\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right| \leq \frac{n}{m(n)} . \tag{69}
\end{align*}
$$

The difference $\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right]$ can be written as a sum of $2 n-3$ independent, zero-mean random variables, with magnitude at most twice and variance at most four times the above. Using the variance and magnitude bound, we can derive the following Bernstein concentration result (Boucheron et al., 2004). Conditioned on $\pi^{*}$,

$$
\begin{align*}
& \mathbb{P}\left(\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right] \geq\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]\right)  \tag{70}\\
= & \mathbb{P}\left(\frac{m(n)}{2 n} \frac{n}{2 n-3}\left(\left[\hat{\Pi}(i)-\tilde{\Pi}^{*}(i)\right]-\left[\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right]\right) \geq \frac{m(n)}{2 n} \frac{n}{2 n-3}\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]\right)  \tag{71}\\
\leq & \exp \left\{-\frac{(2 n-3)\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]^{2}}{2 \frac{(2(2 n-3))^{2}}{m(n)^{2}}\left(4 \frac{n}{m(n)} \frac{1}{n-1}[(j-1) p+(n-j)(1-p)] \frac{m(n)^{2}}{4 n^{2}}+\frac{\left[\tilde{\Pi}^{*}(j)-\tilde{\Pi}^{*}(i)\right]}{3} \frac{m(n)}{2(2 n-3)}\right)}\right\}  \tag{72}\\
\leq & \exp \left\{-\left[\frac{j-i}{n}\right]^{2} \frac{3}{100}(2 p-1)^{2} m(n)\right\} . \tag{73}
\end{align*}
$$

Then plugging this in the expected Kendall tau distance,

$$
\begin{align*}
\mathbb{E}(\operatorname{inv}(\hat{\pi})) & \leq \sum_{i<j} \exp \left\{-\left[\frac{j-i}{n}\right]^{2} \frac{3}{100}(2 p-1)^{2} m(n)\right\}  \tag{74}\\
& =\sum_{k=1}^{n-1}(n-k) \exp \left\{-\left[\frac{k}{n}\right]^{2} \frac{3}{100}(2 p-1)^{2} m(n)\right\}  \tag{75}\\
& \leq \int_{0}^{n}(n-k) \exp \left\{-\left[\frac{k}{n}\right]^{2} \frac{3}{100}(2 p-1)^{2} m(n)\right\} d k  \tag{76}\\
& =n^{2} \frac{\sqrt{\pi \frac{3}{100}(2 p-1)^{2} m(n)} \operatorname{erf}\left(\sqrt{\frac{3}{100}(2 p-1)^{2} m(n)}\right)+\exp \left\{-\frac{3}{100}(2 p-1)^{2} m(n)\right\}-1}{2 \frac{3}{100}(2 p-1)^{2} m(n)}  \tag{77}\\
& \leq \frac{n}{n-1} \sqrt{\frac{400}{3}} \frac{2}{2(2 p-1) \sqrt{m(n)}}\binom{n}{2} . \tag{78}
\end{align*}
$$

For this bound to be no larger than $\frac{\eta}{2}\binom{n}{2}$, we need

$$
\begin{equation*}
\eta>\frac{n}{n-1} \sqrt{\frac{400}{3}} \frac{2}{(2 p-1) \sqrt{m(n)}} \tag{79}
\end{equation*}
$$

so that a suitable $m(n)$ exists which satisfies $m(n) \in \Theta\left(1 /\left((2 p-1)^{2} \eta^{2}\right)\right)$.

### 2.4. Proof of Theorem 4.6

Unbalanced Rank Estimation Algorithm (URE): Measure each binary comparison independently with probability $m(n) / n$. Define the scores

$$
\begin{equation*}
\hat{\Pi}(j)=\frac{1}{m(n)} \sum_{i \neq j} s_{i, j} \bar{c}_{i, j} \propto \sum_{i \neq j} s_{i, j} \bar{c}_{i, j} \tag{80}
\end{equation*}
$$

Predict $\pi^{*}$ by the ordering $\hat{\pi}$ of the estimated scores, breaking ties randomly.
Theorem 4.6. For $c>0$, if each comparison is measured with probability $m(n) / n=c \log (n) / n$, Unbalanced Rank Estimation produces with probability at least

$$
\begin{equation*}
1-2 n^{1-\frac{3}{2}\left[(2 p-1)^{2} \nu^{2} /(3(1-p)+(5 p-1) \nu)\right] c} \tag{81}
\end{equation*}
$$

a permutation $\hat{\pi}$ with

$$
\left|\pi^{*}(j)-\hat{\pi}(j)\right| \leq\left\{\begin{array}{cl}
4 \nu n & \text { if } \pi^{*}(j)<\nu n  \tag{82}\\
4 \sqrt{\nu} \sqrt{\pi^{*}(j) n} & \text { if } \pi^{*}(j) \geq \nu n
\end{array}\right.
$$

Proof. Without loss of generality, we suppose that $\pi^{*}=(1,2, \ldots, n)$, so that $\pi^{*}(j)=j$. To prove this theorem, we need to refine the Bernstein concentration from Theorem 4.5. The true (rescaled) rank score of object $j$ is denoted by

$$
\begin{equation*}
\Pi^{*}(j)=\frac{1}{n} \sum_{i \neq j} c_{i, j}=\frac{j-1}{n} \tag{83}
\end{equation*}
$$

It suffices to know the true ranking scores up to global scaling and additive constants. We will show that the Unbalanced Rank Estimation Algorithm produces an unbiased estimate of the scores

$$
\begin{equation*}
\tilde{\Pi}^{*}(j)=\Pi^{*}(j)(2 p-1)+(1-p) \frac{n-1}{n} \tag{84}
\end{equation*}
$$

Bernstein concentration: We introduce binary random variables $b_{i, j}$ which encode whether or not a comparison was flipped. Let $b_{i, j}=1$ indicate that the measurement was not flipped. By assumption, the $b_{i, j}$ are i.i.d. and $\mathbb{P}\left(b_{i, j}=1\right)=p$. With this, we can write $\bar{c}_{i, j}=b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)$ and so

$$
\begin{align*}
& \hat{\Pi}(j)-\tilde{\Pi}^{*}(j)=\hat{\Pi}(j)-\left(\Pi^{*}(j)(2 p-1)+(1-p) \frac{n-1}{n}\right)  \tag{85}\\
= & \frac{1}{n} \sum_{i \neq j}\left[\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right] . \tag{86}
\end{align*}
$$

Given $\pi^{*}$, the random variables inside the sum are independent. Furthermore, one can show the following results, previously used in Theorem 4.5

$$
\begin{align*}
& \mathbb{E}\left(\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right)=0  \tag{87}\\
& \frac{1}{n-1} \sum_{i \neq j} \operatorname{var}\left(\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right)  \tag{88}\\
\leq & \frac{n}{m(n)} \frac{1}{n-1}[(j-1) p+(n-j)(1-p)]  \tag{89}\\
& \left|\frac{n}{m(n)} s_{i, j}\left[b_{i, j} c_{i, j}+\left(1-b_{i, j}\right)\left(1-c_{i, j}\right)\right]-\left(c_{i, j}(2 p-1)+(1-p)\right)\right| \leq \frac{n}{m(n)} . \tag{90}
\end{align*}
$$

With this we can derive a refined Bernstein concentration result (Boucheron et al., 2004). Conditioned on $\pi^{*}$,

$$
\begin{align*}
& \mathbb{P}\left(\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>t\right)  \tag{91}\\
= & \mathbb{P}\left(\frac{m(n)}{n} \frac{n}{n-1}\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>\frac{m(n)}{n} \frac{n}{n-1} t\right)  \tag{92}\\
\leq & 2 \exp \left\{-\frac{(n-1) t^{2}\left(\frac{n}{n-1}\right)^{2}}{2 \frac{n^{2}}{m(n)^{2}}\left(\frac{n}{m(n)} \frac{1}{n-1}[(j-1) p+(n-j)(1-p)] \frac{m(n)^{2}}{n^{2}}+\frac{t}{3} \frac{m(n)}{n} \frac{n}{n-1}\right)}\right\}  \tag{93}\\
\leq & 2 \exp \left\{-\frac{t^{2} m(n)}{2\left(\frac{j}{n} p+(1-p)+\frac{t}{3}\right)}\right\} . \tag{94}
\end{align*}
$$

Let us now substitute different values of $t$. To begin, if $j<\nu n$, then by rescaling $t$,

$$
\begin{align*}
\mathbb{P}\left(\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>\sqrt{\nu} t\right) & \leq 2 \exp \left\{-\frac{\nu t^{2} m(n)}{2\left(\frac{j}{n} p+(1-p)+\frac{t}{3} \sqrt{\nu}\right)}\right\}  \tag{95}\\
& \leq 2 \exp \left\{-\frac{\sqrt{\nu} t^{2} m(n)}{2\left(\sqrt{\nu} p+\frac{1}{\sqrt{\nu}}(1-p)+\frac{t}{3}\right)}\right\} \tag{96}
\end{align*}
$$

And if $j \geq \nu n$, by rescaling $t$,

$$
\begin{align*}
\mathbb{P}\left(\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>\sqrt{\frac{j}{n}} t\right) & \leq 2 \exp \left\{-\frac{t^{2} m(n)}{2 \frac{n}{j}\left(\frac{j}{n} p+(1-p)+\frac{t}{3} \sqrt{\frac{j}{n}}\right)}\right\}  \tag{97}\\
& \leq 2 \exp \left\{-\frac{\sqrt{\nu} t^{2} m(n)}{2\left(\sqrt{\nu} p+\frac{1}{\sqrt{\nu}}(1-p)+\frac{t}{3}\right)}\right\} \tag{98}
\end{align*}
$$

Notice that the upper bounds do not depend on $j$ and are identical in the two cases. Hence, we see that the concentration result becomes strong for small $j<\nu n$ as $n$ gets large, but remains relatively weak for large $j \approx n$. Define the events

$$
A_{j}= \begin{cases}\left\{\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>\sqrt{\nu} t\right\} & \text { if } \quad j<\nu n  \tag{99}\\ \left.\left|\hat{\Pi}(j)-\tilde{\Pi}^{*}(j)\right|>\sqrt{\frac{j}{n}} t\right\} & \text { if } \quad j \geq \nu n\end{cases}
$$

Applying a union bound, we find

$$
\begin{equation*}
\mathbb{P}\left(\bigcup_{j=1}^{n} A_{j}\right) \leq 2 \exp \left\{-\frac{\sqrt{\nu} t^{2} m(n)}{2\left(\sqrt{\nu} p+\frac{1}{\sqrt{\nu}}(1-p)+\frac{t}{3}\right)}+\log (n)\right\} \tag{100}
\end{equation*}
$$

To use the bound in Eq. (100), we first prove the following Lemma.
Lemma 4.7. For some $a>0,0<\gamma<a^{2}$ and arbitrary $b \in \mathbb{R}$, if

$$
\left|\hat{\Pi}(j)-\left(\Pi^{*}(j) a+b\right)\right| \leq \begin{cases}\gamma / a & \text { if } j<\gamma n / a^{2}  \tag{101}\\ \sqrt{\gamma j / n} & \text { if } j \geq \gamma n / a^{2}\end{cases}
$$

then

$$
\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq \begin{cases}4 \gamma n / a^{2} & \text { if } j<\gamma n / a^{2}  \tag{102}\\ 4 \sqrt{\gamma j n} / a & \text { if } j \geq \gamma n / a^{2}\end{cases}
$$

Proof. Let $\tilde{\Pi}^{*}=\Pi^{*}(j) a+b$. We need to bound the number of elements of $\hat{\Pi}$ that can appear on the wrong side of $\hat{\Pi}(j)$. The prediction $\hat{\pi}(j)$ can deviate from $\pi^{*}(j)$ by at most the number of such misplaced elements. It suffices to bound the number of elements of $\hat{\Pi}$ that can take on the same value $\hat{\Pi}(j)$ and then to assume that the Unbalanced Rank Estimation Algorithm breaks ties in the least favorable way. To maximize the number of such confusions, the estimated score $\tilde{\Pi}(j)$ must deviate $u p$ from its mean value $\tilde{\Pi}^{*}(j)$ as much as possible, since then the most elements $k$ with $\tilde{\Pi}^{*}(k)>\hat{\Pi}(j)$ can map down to $\hat{\Pi}(j)$, and the most elements $k$ with $\hat{\Pi}(j)>\tilde{\Pi}^{*}(k)$ can map up to $\hat{\Pi}(j)$ (The conditions of the lemma ensure that larger deviations are possible for large $k$ than small $k$.) Specifically, if $j \geq \gamma n / a^{2}$ then we should have $\hat{\Pi}(j)=\tilde{\Pi}^{*}(j)+\sqrt{\gamma j / n}$. If $j<\gamma n / a^{2}$, then we should have, $\hat{\Pi}(j)=\tilde{\Pi}^{*}(j)+\gamma / a$. For the following, denote $t_{j}=\sqrt{\gamma j / n} / a$, which is how much the rescaled score $\hat{\Pi}(j) / a$ can differ from $\tilde{\Pi}^{*}(j) / a=\Pi^{*}(j)+b / a$ if $j \geq \gamma n / a^{2}$.
Suppose that $j \geq \gamma n / a^{2}$. The largest element $\bar{k}$ that can overlap with $j$ satisfies

$$
\begin{align*}
\frac{\tilde{\Pi}^{*}(\bar{k}) n}{a}-t_{\bar{k}} n & =\frac{\tilde{\Pi}^{*}(j) n}{a}+t_{j} n  \tag{103}\\
\bar{k}-1+\frac{b n}{a}-t_{\bar{k}} n & =j-1+\frac{b n}{a}+t_{j} n  \tag{104}\\
\bar{k}-t_{\bar{k}} n & =j+t_{j} n  \tag{105}\\
\bar{k}-\frac{\sqrt{n \bar{k}} \sqrt{\gamma}}{a} & =j+\frac{\sqrt{n j} \sqrt{\gamma}}{a} \tag{106}
\end{align*}
$$

Since $\sqrt{\bar{k}} \geq 0$, we use the positive solution given by the quadratic formula:

$$
\begin{align*}
\sqrt{\bar{k}} & =\frac{\sqrt{n \gamma} / a+\sqrt{n \gamma / a^{2}+4 \sqrt{n j} \sqrt{\gamma} / a+4 j}}{2}  \tag{107}\\
& =\frac{\sqrt{n \gamma} / a+\sqrt{(\sqrt{n \gamma} / a+2 \sqrt{j})^{2}}}{2}  \tag{108}\\
& =\frac{\sqrt{n \gamma}}{a}+\sqrt{j} \tag{109}
\end{align*}
$$

With this, we have

$$
\begin{equation*}
t_{\bar{k}}=\sqrt{\frac{\bar{k}}{n}} \frac{\sqrt{\gamma}}{a}=\frac{\sqrt{n \gamma} / a+\sqrt{j}}{\sqrt{n}} \frac{\sqrt{\gamma}}{a}=\frac{\gamma}{a^{2}}+\frac{\sqrt{\gamma}}{a} \sqrt{\frac{j}{n}} \tag{110}
\end{equation*}
$$

By this derivation, at most $t_{\bar{k}} n$ elements $k$ with $\tilde{\Pi}_{\hat{\Pi}}(k)>\hat{\Pi}(j)$ can map down to $\hat{\Pi}(j)$. Furthermore, at most $t_{j} n$ elements $k$ with $\hat{\Pi}(j)>\tilde{\Pi}^{*}(k)$ can map up to $\hat{\Pi}(j)$. Taken together, because $t_{j} \leq t_{\bar{k}}$, at most $\left(t_{j}+t_{\bar{k}}\right) n \leq$ $2 t_{\bar{k}} n=2 \gamma n / a^{2}+2 \sqrt{\gamma} \sqrt{j n} / a$ elements $k$ can map onto $\hat{\Pi}(j)$. This means, even if the sorting algorithm breaks ties in the least favorable way, we have $\forall j \geq \gamma n / a^{2}$

$$
\begin{equation*}
\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq 2 t_{\bar{k}} n=\frac{2 \gamma n}{a^{2}}+\frac{2 \sqrt{\gamma j n}}{a} \tag{111}
\end{equation*}
$$

Note that if $j \geq \gamma n / a^{2}$, the second term in Eq. (111) is at least as large as the first term. So $\forall j \geq \gamma n / a^{2}$

$$
\begin{equation*}
\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq \frac{4 \sqrt{\gamma j n}}{a} \tag{112}
\end{equation*}
$$

On the other hand, if $j \leq \gamma n / a^{2}$, then at most $\bar{k}$ elements can map to $\hat{\Pi}(j)$. The upper limit $\bar{k}$ is largest when $j=\gamma n / a^{2}$. Thus,

$$
\begin{equation*}
\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq \bar{k}=\sqrt{\bar{k}}^{2}=\frac{\gamma n}{a^{2}}+\frac{2 \sqrt{\gamma n j}}{a}+j=\frac{4 \gamma n}{a^{2}} . \tag{113}
\end{equation*}
$$

We can now prove the theorem. In the context of $\tilde{\Pi}^{*}$ in Eq. (84), $a=(2 p-1)$. For any $0<\nu<1$, set $\gamma=\nu a^{2}$ in Lemma 4.7. If we set $t=a \sqrt{\nu}$, then the union bound in Eq. (100) controls the probability that the bounds on the scores required by the lemma will be satisfied so that we can use the lemma to draw the desired conclusion. Specifically, with these settings, if $m(n) \geq c(p, \nu) \log (n)$ with the constant $c(p, \nu)$ large enough, then by the lemma we predict with high probability the first $\nu n$ elements $j$ of $\pi^{*}$ with accuracy $\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq 4 \nu n$ and the remaining elements with accuracy $\left|\hat{\pi}(j)-\pi^{*}(j)\right| \leq 4 \sqrt{\nu} \sqrt{j n}$.

Probability of success. The probability that the preconditions to Lemma 4.7 hold depend on the constant $c(p, \nu)$. Specifically,

$$
\begin{align*}
1-\mathbb{P}\left(\bigcup_{j=1}^{n} A_{j}\right) & \geq 1-2 \exp \left\{-\frac{\sqrt{\nu}((2 p-1) \sqrt{\nu})^{2} c(p, \nu) \log (n)}{2\left(\sqrt{\nu} p+\frac{1}{\sqrt{\nu}}(1-p)+\frac{(2 p-1) \sqrt{\nu}}{3}\right)}+\log (n)\right\}  \tag{114}\\
& =1-2 n^{1-\frac{3}{2} \frac{(2 p-1)^{2} \nu^{2}}{(3(1-p)+(5 p-1) \nu)} c(p, \nu)} \tag{115}
\end{align*}
$$

## References

Boucheron, S., Lugosi, G., and Bousquet, O. Concentration inequalities. In Advanced Lectures in Machine Learning, pp. 208-240. Springer, 2004.

Bousquet, O., Boucheron, S., and Lugosi, G. Introduction to statistical learning theory. In Bousquet, O., von Luxburg, U., and Rätsch, G. (eds.), Advanced Lectures on Machine Learning, ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures, volume 3176 of Lecture Notes in Computer Science, pp. 169-207. Springer, 2003.

Diaconis, P. and Graham, R. L. Spearman's footrule as a measure of disarray. Journal of the Royal Statistical Society. Series B (Methodological), 39(2):262-268, 1977.

Giesen, J., Schuberth, E., and Stojaković, M. Approximate sorting. Fundamenta Informaticae, 90(1-2):67-72, 2009.

Radinsky, K. and Ailon, N. Ranking from pairs and triplets: Information quality, evaluation methods and query complexity. In King, I., Nejdl, W., and Li, H. (eds.), Fourth ACM International Conference on Web Search and Data Mining (WSDM), pp. 105-114. ACM, 2011.

