Optimal Differentially Private Ranking from Pairwise Comparisons*

T. Tony Cai, Abhinav Chakraborty, and Yichen Wang

November 4, 2025

Abstract

Data privacy is a central concern in many applications involving ranking from incomplete and noisy pairwise comparisons, such as recommendation systems, educational assessments, and opinion surveys on sensitive topics. In this work, we propose differentially private algorithms for ranking based on pairwise comparisons. Specifically, we develop and analyze ranking methods under two privacy notions: edge differential privacy, which protects the confidentiality of individual comparison outcomes, and individual differential privacy, which safeguards potentially many comparisons contributed by a single individual. Our algorithms–including a perturbed maximum likelihood estimator and a noisy count-based method–are shown to achieve minimax optimal rates of convergence under the respective privacy constraints. We further demonstrate the practical effectiveness of our methods through experiments on both simulated and real-world data.

Keywords: differential privacy, ranking, Bradley-Terry-Luce model, minimax optimality

1 Introduction

As personal data are more extensively collected and analyzed than ever, the importance of privacy protection in data analysis is also increasingly recognized. In this paper, we consider privacy-preserving methods for ranking from pairwise comparisons. In this ranking problem, the data analyst observes random and incomplete pairwise comparisons among items that follow some unknown ranking, with higher ranked items more likely, but not guaranteed, to prevail over lower ranked ones. The analyst then tries to infer the underlying ranking from the noisy comparison results. The extensive applied research on this topic attests to its practical relevance.

*T. T. Cai is with the Department of Statistics and Data Science, The Wharton School, University of Pennsylvania and A. Chakraborty is with Department of Statistics, Columbia University. The research of T.T. Cai was supported in part by NSF Grant DMS-2413106 and NIH grants R01-GM129781 and R01-GM123056. Y. Wang is an independent researcher. This research was conducted outside of Y. Wang's current employment with Amazon.com Services LLC.

- **Pairwise comparison in sensitive survey data:** Pairwise comparisons in surveys provide a structured approach for respondents to express their relative preferences across a wide range of options. For instance, a survey by [53] was conducted to gauge public sentiments on immigration. The study involved 98 student participants, each of whom responded to at least one pairwise comparison drawn from four extreme statements regarding immigrants.

- **Pairwise comparison in recommendation systems.** Pairwise comparisons are widely used in recommendation systems that capture users' preferences between pairs of items such as movies, books, or other consumer products. For example, [5] proposed a method in which customers are presented with a series of paired preference questions (e.g. "Do you prefer item A over item B?"). The preferences are then used to infer personalized rankings.

- **Pairwise comparison in education.** Pairwise comparison can serve as an effective tool for educational assessment. For instance, [30] describes a study in which teachers used a pairwise comparison procedure to grade student works and construct a performance scale. The study found that teacher judgments were internally consistent and showed strong correlation with scores from a large-scale standardized test administered to the same group of students.

Privacy is a critical concern in many applications involving ranking from pairwise comparisons. For example, in the educational assessment study by [30], teachers' preferences between pairs of student assignments should remain confidential. Similarly, in the survey conducted by [53], respondents were asked to express preferences between pairs of political positions–data considered sensitive due to the controversial nature of the statements and the potential ramifications for individuals if their views were made public.

Motivated by the importance of protecting privacy in such settings, we develop *statistically optimal* algorithms for ranking from pairwise comparisons under *differential privacy (DP) constraints*. Differential privacy [24, 23] is the most widely adopted framework for privacy-preserving data analysis, offering strong formal guarantees that the output of an algorithm reveals provably little about any individual data point. At the same time, it supports the design and implementation of efficient algorithms. In this paper, we propose and analyze DP algorithms for ranking from pairwise comparisons and show that they achieve statistical optimality under the DP constraint: no other differentially private algorithm can attain a faster rate of convergence to the true ranking.

1.1 Main Results and Our Contribution

Edge DP and individual DP. A differentially private algorithm guarantees that its outputs on two *adjacent* datasets are statistically similar. Two datasets are *adjacent* if they differ by exactly one unit of data. In essence, a DP algorithm ensures that its output does not depend significantly on any single unit of data, thereby preventing inference about the presence or absence of any individual in the input. However, as discussed in [4], the definition of adjacency, and thus the unit of data, depends on the context. For instance, in a census, a unit might be a person or a household, while in business analytics, it could be a single transaction. In the context of ranking from pairwise comparisons, two natural notions of adjacency arise, leading to two corresponding definitions of differential privacy:

- Edge DP: Two datasets are adjacent if they differ by a single pairwise comparison outcome.
- Individual DP: Two datasets are adjacent if they differ in the comparisons contributed by a single individual.

To the best of our knowledge, this paper is the first to simultaneously consider both definitions of DP and rigorously analyze them in the context of ranking from pairwise comparisons.

Optimal parametric estimation with differential privacy. Under the parametric Bradley-Terry-Luce (BTL) model for pairwise comparisons [9, 36], we introduce in Section 2.1 a perturbed maximum likelihood estimator of the following form:

$$\widetilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\mathbb{R}^n}{\arg\min}\, \mathcal{L}(\boldsymbol{\theta}; y) + \frac{\gamma}{2}\|\boldsymbol{\theta}\|_2^2 + \boldsymbol{w}^\top\boldsymbol{\theta}, \quad \boldsymbol{w} = (w_1, w_2, \cdots, w_n) \overset{\text{i.i.d.}}{\sim} \text{Laplace}(\lambda),$$

where $\mathcal{L}(\boldsymbol{\theta}, y)$ is the likelihood function, $\boldsymbol{\theta}$ is the vector of latent parameters which determine the items' ranks, and y is the data set of pairwise comparisons. We then show that the estimator respectively satisfies edge DP and individual DP for two different choices of (γ, λ), in Sections 2.1 and 3.1. It is further shown by Theorems 2.2 and 3.2 that the estimator's respective rates of convergence under edge DP and individual DP are optimal.

Optimal nonparametric ranking with differential privacy. In the absence of parametric assumptions, we show that ranking items based on noisy counts of wins – with appropriately calibrated noise distributions – satisfies either edge differential privacy or individual differential privacy. Moreover, this simple approach is statistically optimal: it accurately ranks items across the broadest possible regime of sample sizes and privacy levels. Specifically, in Theorems 2.3 and 2.4, we show that ranking by noisy counts succeeds at identifying the top k items whenever the strengths of the k-th and $(k+1)$-th items are separated by a certain threshold, and no other edge DP algorithm will succeed below this threshold. Theorems 3.3 and 3.4 contain similar results for individual DP.

Entrywise analysis of DP algorithms. The study of differentially private ranking also yields insights of broader methodological interest. In particular, the perturbed maximum likelihood estimator (MLE) achieves differential privacy by injecting a single dose of noise into the objective function. This design facilitates a leave-one-out analysis [19], enabling precise control of entrywise errors in optimization problems. While most prior work on differentially private optimization has focused on bounding ℓ_2 errors, our framework extends the analysis to the entrywise setting. Furthermore, our adaptation of the score attack technique [14]–originally developed for ℓ_2 risk–to entrywise error analysis may have broader applicability in establishing entrywise lower bounds for differentially private algorithms more generally.

1.2 Related Work

Some of the most historically significant contributions to the study of pairwise comparisons and ranking include [50], which pioneered the use of pairwise comparisons for measuring psychological values, and the seminal works of [9] and [36], which introduced the Bradley-Terry-Luce (BTL) model. Additionally, [28] was the first to formulate the ranking problem as a maximum likelihood estimation task.

In recent years, there has been growing interest in understanding the minimax rates of convergence for ranking from pairwise comparisons. Several works – including [52, 41, 37, 42, 20, 19] – adopt parametric assumptions, often relying on the BTL model, to study the minimax ℓ_2 or ℓ_∞ risk in estimating underlying parameters. Parallel to this, a nonparametric line of research focuses on identifying the top-ranked items [45, 18] or estimating pairwise comparison probabilities under the assumption of stochastic transitivity [43, 44, 39].

On the other hand, the trade-off between differential privacy and statistical utility has also attracted substantial attention. A wide array of differentially private statistical methods have been proposed and analyzed–ranging from Gaussian mean estimation and linear regression [13], to nonparametric density/regression estimation [51, 11], M-estimators [35], and PCA [27]. These methods are often grounded in foundational design paradigms such as the Laplace and Gaussian mechanisms [24, 25], and private convex optimization techniques [16, 17, 8, 34, 7].

Specifically on differentially private ranking, existing works such as [46, 29, 54, 47] focus on the related but distinct problem of rank aggregation, where the goal is to combine multiple full rankings into a single consensus ranking. [40] proposes a one-shot algorithm for ranking from pairwise comparisons as an application of their private top-k selection method; we compare this with our algorithms in Section 4. [14] studies a special case of our edge DP setting under ℓ_2 loss, but does not directly address the ranking problem considered here.

Understanding the privacy-utility trade-off requires identifying the minimum possible loss in accuracy among all differentially private methods. To this end, several powerful lower-bounding techniques have been developed, including tracing attacks [10, 26, 48, 49, 31, 13, 11] and the score attack [12, 14], as well as differentially private versions of Le Cam's, Fano's, and Assouad's inequalities [6, 32, 15, 2, 3].

1.3 Organization

The rest of the paper is organized as follows. Section 2 focuses on ranking under edge differential privacy (DP). Section 2.1 presents the privacy and optimality analysis of the parametric estimator, while Section 2.2 addresses the privacy guarantees and optimality of ranking by noisy counts under edge DP. Section 3 turns to ranking under individual DP, with Sections 3.1 and 3.2 dedicated to the parametric and nonparametric analyses, respectively. Section 4 provides a numerical evaluation of our algorithms using both simulated and real datasets. Section 5 concludes the paper with a discussion of open questions and directions for future work. Additional technical details and omitted proofs are provided in the supplementary material [1].

1.4 Notation

For real-valued sequences $\{a_n\}, \{b_n\}$, we write $a_n \lesssim b_n$ if $a_n \leq cb_n$ for some universal constant $c \in (0, \infty)$, and $a_n \gtrsim b_n$ if $a_n \geq c'b_n$ for some universal constant $c' \in (0, \infty)$. We say $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $a_n \gtrsim b_n$. $c, C, c_0, c_1, c_2, \cdots$, and so on refer to universal constants in the paper, with their specific values possibly varying from place to place. For a positive integer n, let $[n] = \{1, 2, 3, \cdots, n\}$.

2 Optimal Ranking under Edge Differential Privacy

The ranking problem. There are n distinct items, indexed by $[n] = \{1, 2, 3, \cdots, n\}$. Pairwise comparisons between items are observed randomly and independently, where each pair (i, j), $1 \leq i < j \leq n$, is compared with a known probability $p \in (0, 1]$. This results in an Erdős–Rényi random graph $\mathcal{G}$ with n nodes and the observed comparisons constituting the edges. Every observed pair (i, j) determines a unique winner, symbolized by the outcome $Y_{ij} \in \{0, 1\}$, satisfying $Y_{ij} + Y_{ji} = 1$. Consequently, for $i < j$, the random variable Y_{ij} follows an independent Bernoulli distribution with parameter $\rho_{ij} \in [0, 1]$, and the requirement $Y_{ij} + Y_{ji} = 1$ implies $\rho_{ij} + \rho_{ji} = 1$. We assume $\rho_{ii} = 1/2$ for clarity.

Our objective is to rank the set of n items based on the average true winning probabilities when compared to randomly selected counterparts. This average winning probability for each item $i \in [n]$ is formally denoted by $\tau_i = \frac{1}{n}\sum_{j\in[n]} \rho_{ij}$. We are interested in estimating the index set $\mathcal{S}_k$, where $\mathcal{S}_k = \{i \in [n] : \tau_i \text{ is among the top-}k \text{ largest of } \tau_1, \tau_2, \ldots, \tau_n\}$ for a predetermined $k \in [n]$.

Parametric and nonparametric models. The ranking problem is studied under two models. The first one is a parametric model where $\rho_{ij} = F(\theta_i^* - \theta_j^*)$. Each item $i \in [n]$ is assigned a latent parameter θ_i^*, and $F : \mathbb{R} \to [0, 1]$ is a predetermined link function. This model generalizes well-known the Bradley-Terry-Luce (BTL) model [9, 36] for pairwise comparison, and recovers the BTL model when F is the standard logistic CDF. With this parametric assumption, the ranking of τ_i is equivalent to the ranking of θ_i^*, which further reduces to estimating the parameters $\{\theta_i^*\}_{i\in[n]}$.

The second model is nonparametric, in which we do not assume any parametric form for the ρ_{ij} values, and instead aim to estimate the ranks directly. This nonparametric ranking problem is the focus of a more recent line of work [45, 18, 45, 44]. We define and study these two settings in Sections 2.1 and 2.2 respectively.

Edge differential privacy. Under these models, we study ranking algorithms satisfying (ε, δ) differential privacy. The formal definition of (ε, δ)-DP requires that, for an algorithm M taking values in some domain $\mathcal{R}$ and every measurable subset $A \subseteq \mathcal{R}$,

$$\mathbb{P}(M(X) \in A) \leq e^{\varepsilon} \cdot \mathbb{P}(M(X') \in A) + \delta$$

for any pair of data sets X and X' which differ by one unit of data. A pair of data sets which differ by one unit of data is called *adjacent* data sets.

When specialized to pairwise comparison, one natural interpretation of *adjacency* is that two sets of comparison outcomes differ by a single comparison. Concretely, two sets of comparison outcomes $\boldsymbol{Y} = \{Y_{ij}\}_{(i,j)\in\mathcal{G}}$ and $\boldsymbol{Y}' = \{Y'_{ij}\}_{(i,j)\in\mathcal{G}'}$ are adjacent if they satisfy one of following scenarios.

- The comparison graphs are identical, $\mathcal{G} = \mathcal{G}'$, and there exists exactly one edge $(i^*, j^*) \in \mathcal{G}$ on which the comparison outcomes differ, $Y_{i^*j^*} \neq Y'_{i^*j^*}$. All other comparison outcomes are identical: $Y_{ij} = Y'_{ij}$ for $(i,j) \neq (i^*, j^*)$.

- The comparison graphs $\mathcal{G}$ and $\mathcal{G}'$ differ by exactly one edge: there exist $a^*, b^*, c^*, d^* \in [n]$ and $(a^*, b^*) \neq (c^*, d^*)$, such that

$$\mathcal{G} = \mathcal{G} \cap \mathcal{G}' + \{(a^*, b^*)\}, \mathcal{G}' = \mathcal{G} \cap \mathcal{G}' + \{(c^*, d^*)\}.$$

  The comparison outcomes $\boldsymbol{Y} = \{Y_{ij}\}_{(i,j)\in\mathcal{G}}$ and $\boldsymbol{Y}' = \{Y'_{ij}\}_{(i,j)\in\mathcal{G}'}$ satisfy $Y_{ij} = Y'_{ij}$ for all $(i,j) \in \mathcal{G} \cap \mathcal{G}'$.

This notion of adjacency and the corresponding definition of differential privacy is akin to "edge differential privacy" for graphs [38, 33], and we adopt the same term for our case.

2.1 Parametric Estimation under Edge DP

We first study ranking from pairwise comparisons under parametric assumptions: each item $i \in [n]$ is associated with a latent parameter θ_i^*, and the pairwise probability ρ_{ij} is related to the latent parameters of items i, j by a known increasing function $F : \mathbb{R} \to [0,1]$, specifically $\rho_{ij} = F(\theta_i^* - \theta_j^*)$. These assumptions conveniently reduce the problem of ranking n items by their average winning probability against peers, $\tau_i = n^{-1} \sum_{j\in[n]} \rho_{ij}$, to the problem of estimating $\boldsymbol{\theta}^* = (\theta_i^*)_{i\in[n]}$.

2.1.1 The Edge DP Algorithm for Parametric Estimation

For constructing a differentially private estimator of $\boldsymbol{\theta}^*$, our approach is to minimize a randomly perturbed and ℓ_2-penalized version of the negative log-likelihood function. For a vector $\boldsymbol{v} \in \mathbb{R}^n$, indices $i, j \in [n]$ and a given link function F, let $F_{ij}(\boldsymbol{v}) = F(v_i - v_j)$ and $F'_{ij}(\boldsymbol{v}) = F'(v_i - v_j)$. The negative log-likelihood function is given by

$$\mathcal{L}(\boldsymbol{\theta}; y) = \sum_{(i,j)\in\mathcal{G}} -y_{ij} \log F_{ij}(\boldsymbol{\theta}) - y_{ji} \log(1 - F_{ij}(\boldsymbol{\theta})). \tag{2.1}$$

The estimator is defined by Algorithm 1.

Algorithm 1 Differentially Private Ranking for parametric models

Input: Comparison data $(y_{ij})_{(i,j)\in\mathcal{G}}$, comparison graph $\mathcal{G}$, privacy parameter ε, regularity constants κ_1, κ_2 defined in (2.3) and (2.4).

1: Set $\lambda \geq \frac{8\kappa_1}{\varepsilon}$ and $\gamma \geq \frac{4\kappa_2}{\varepsilon}$ and generate $\boldsymbol{w} = (w_1, w_2, \cdots, w_n) \overset{\text{i.i.d.}}{\sim} \text{Laplace}(\lambda)$.

2: Solve for

$$\widetilde{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\mathbb{R}^n} \mathcal{L}(\boldsymbol{\theta}; y) + \frac{\gamma}{2}\|\boldsymbol{\theta}\|_2^2 + \boldsymbol{w}^\top\boldsymbol{\theta}, \quad \boldsymbol{w} = (w_1, w_2, \cdots, w_n) \overset{\text{i.i.d.}}{\sim} \text{Laplace}(\lambda). \tag{2.2}$$

Output: $\widetilde{\boldsymbol{\theta}}$.

Some regularity conditions on the function F will be helpful throughout our analysis of $\widetilde{\boldsymbol{\theta}}$. We collect them here for convenience.

(A0) $F : \mathbb{R} \to [0, 1]$ is strictly increasing and satisfies $F(x) = 1 - F(-x)$ for every $x \in \mathbb{R}$.

(A1) There is an absolute constant $\kappa_1 > 0$ such that

$$\sup_{x\in\mathbb{R}} \left|\frac{F'(x)}{F(x)(1 - F(x))}\right| = \sup_{x\in\mathbb{R}} \frac{F'(x)}{F(x)(1 - F(x))} < \kappa_1. \tag{2.3}$$

(A2) $\frac{\partial^2}{\partial x^2}(-\log F(x)) > 0$ for every $x \in \mathbb{R}$, and there exists an absolute constant $\kappa_2 > 0$ such that

$$\frac{\partial^2}{\partial x^2}(-\log F(x)) < \kappa_2, \quad \min_{|x|\leq 4} \frac{\partial^2}{\partial x^2}(-\log F(x)) > \frac{1}{\kappa_2}. \tag{2.4}$$

In particular, choosing F to be the standard logistic CDF satisfies these conditions and recovers the BTL model.

Returning to the estimator (2.2), the random perturbation $\boldsymbol{w}^\top\boldsymbol{\theta}$ is an instance of objective perturbation methods in differentially private optimization [17, 34]. Let $\mathcal{R}(\boldsymbol{\theta}; y)$ denote the regularized log-likelihood part, $\mathcal{R}(\boldsymbol{\theta}; y) = \mathcal{L}(\boldsymbol{\theta}; y) + \frac{\gamma}{2}\|\boldsymbol{\theta}\|_2^2$, then $\widetilde{\boldsymbol{\theta}}$ amounts to the solution of a noisy stationary condition $\nabla\mathcal{R}(\widetilde{\boldsymbol{\theta}}; y) = -\boldsymbol{w}$. The solution $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}(y)$ is differentially private when (1) the scale parameter λ of noise vector $\boldsymbol{w}$ is sufficiently large to obfuscate the change in $\nabla\mathcal{R}(\widetilde{\boldsymbol{\theta}})$ over adjacent data sets, and (2) the regularization coefficient γ ensures strong convexity of the objective $\mathcal{R}(\boldsymbol{\theta})$, so that perturbation of the gradient is translated to perturbation of the solution

$\widetilde{\boldsymbol{\theta}}$. The privacy guarantee is formalized by Proposition 2.1.

Proposition 2.1. *Suppose conditions (A0), (A1) and (A2) hold. If* $\lambda \geq 8\kappa_1/\varepsilon$ *and* $\gamma \geq 4\kappa_2/\varepsilon$, $\widetilde{\boldsymbol{\theta}}$ *as defined in Algorithm 1 is* $(\varepsilon, 0)$ *differentially private.*

Proposition 2.1 is proved in the supplement [1].

We have so far not considered the convergence of $\widetilde{\boldsymbol{\theta}}$ to the truth $\boldsymbol{\theta}^*$, and in particular choosing large values of λ and γ for differential privacy compromises the accuracy of the estimator $\widetilde{\boldsymbol{\theta}}$. The optimal choice of λ and γ, which balances privacy and utility, depends on the loss function.

When $\gamma \asymp \sqrt{np\log n}$ and $F(x) = (1+e^{-x})^{-1}$, the ℓ_2-penalized MLE $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\mathbb{R}^n} \mathcal{L}(\boldsymbol{\theta}; y) + \frac{\gamma}{2}\|\boldsymbol{\theta}\|_2^2$ is shown to be a minimax optimal estimator of $\boldsymbol{\theta}^*$ by [19]. By following a similar path as the leave-one-out analysis in [19], we can then characterize the entry-wise convergence of $\widetilde{\boldsymbol{\theta}}$ in terms of the noise scale λ. As the parametric model $\rho_{ij} = F(\theta_i^* - \theta_j^*)$ is invariant to translations of $\boldsymbol{\theta}^*$, we assume without the loss of generality that $\boldsymbol{\theta}^*$ is centered: $\mathbf{1}^\top\boldsymbol{\theta}^* = 0$.

Proposition 2.2. *If* $\gamma = c_0\sqrt{np\log n}$ *for some absolute constant* c_0, $p \geq c_1\lambda\log n/n$ *for some sufficiently large constant* $c_1 > 0$, *and* $c_2 < \lambda < c_2\sqrt{\log n}$ *for some sufficiently large constant* $c_2 > 0$, *it holds with probability at least* $1 - O(n^{-5})$ *that*

$$\|\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\|_\infty \lesssim \sqrt{\frac{\log n}{np}} + \frac{\lambda\log n}{np}. \tag{2.5}$$

The proof is given in [1]. Combining the privacy guarantee, Proposition 2.1, with the rate of convergence, Proposition 2.2 leads to the rate of convergence of our estimator $\widetilde{\boldsymbol{\theta}}$.

Theorem 2.1. *If* $\gamma = c_0\sqrt{np\log n}$ *for some absolute constant* $c_0 > 0$, $p \geq c_1\log n/n\varepsilon$ *for some absolute constant* $c_1 > 0$, $\lambda = 8\kappa_1/\varepsilon$, *and* $c_2(\log n)^{-1/2} < \varepsilon < 1$ *for some absolute constant* $c_2 > 0$, *then the estimator* $\widetilde{\boldsymbol{\theta}}$ *defined in* (2.2) *is* $(\varepsilon, 0)$ *edge differentially private, and it holds with probability at least* $1 - O(n^{-5})$ *that*

$$\|\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\|_\infty \lesssim \sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon}. \tag{2.6}$$

In Theorem 2.1, the assumed conditions ensure Propositions 2.1 and 2.2 are applicable. The upper bound (2.6) follows from (2.5) in Proposition 2.2 by plugging in $\lambda \asymp 1/\varepsilon$. The entry-wise error bound implies that the latent parameters $(\theta_i^*)_{i\in[n]}$ can be ranked correctly as long as the true kth and $(k+1)$th ranked items are sufficiently separated in their θ values for all $k \in [n-1]$,

$$|\theta^*_{(k)} - \theta^*_{(k+1)}| \gtrsim \sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon}. \tag{2.7}$$

More formally, if $\widetilde{\mathcal{S}}_k$ is the index of the top k values of the vector $\tilde{\boldsymbol{\theta}}$ then we have the following result for the recovery of the true top-k set $\mathcal{S}_k$.

Corollary 2.1. *Under conditions of Theorem 2.1, if* (2.7) *holds for a fixed k, then*

$$\mathbb{P}(\widetilde{\mathcal{S}}_k \neq \mathcal{S}_k) = O(n^{-5}).$$

In the separation condition (2.7), the $O\left(\frac{\log n}{np\varepsilon}\right)$ due to the differential privacy constraint can dominate the $O(\sqrt{\frac{\log n}{np}})$ term, which is optimal in the non-private case, when for example $\varepsilon \asymp (\log n)^{-1/2}$ and $p \ll \frac{\log^2 n}{n}$. The potentially severe cost of requiring differential privacy motivates the next section which studies the inevitable cost of privacy for entry-wise estimation of $\boldsymbol{\theta}^*$.

2.1.2 The Edge DP Lower Bound for Parametric Estimation

For an arbitrary (ε, δ)-DP estimator $M(\boldsymbol{Y})$ of $\boldsymbol{\theta}$, we would like to find a lower bound for the maximum ℓ_∞ risk $\sup_{\theta\in\Theta} \mathbb{E}\|M(\boldsymbol{Y}) - \boldsymbol{\theta}\|_\infty$ over the parameter space $\Theta = \{\boldsymbol{\theta} \in \mathbb{R}^n : \|\boldsymbol{\theta}\|_\infty \leq 1\}$, which captures the inevitable cost of differential privacy for estimating $\boldsymbol{\theta}$.

To this end, we consider an entry-wise version of the score attack method [14]:

$$\mathcal{A}^{(k)}(M(\boldsymbol{Y}), Y_{ij}) = \begin{cases} 0 & (i,j) \notin \mathcal{G} \text{ or } i, j \neq k, \\ (M(\boldsymbol{Y})_k - \theta_k)(y_{kj} - F_{kj}(\boldsymbol{\theta}))\frac{F'_{kj}(\boldsymbol{\theta})}{F_{kj}(\boldsymbol{\theta})(1-F_{kj}(\boldsymbol{\theta}))} & (i,j) \in \mathcal{G} \text{ and } i = k, \\ (M(\boldsymbol{Y})_k - \theta_k)(y_{ik} - F_{ik}(\boldsymbol{\theta}))\frac{F'_{ik}(\boldsymbol{\theta})}{F_{ik}(\boldsymbol{\theta})(1-F_{ik}(\boldsymbol{\theta}))} & (i,j) \in \mathcal{G} \text{ and } j = k. \end{cases}$$

It is an entry-wise version of the score attack in the sense that summing $\mathcal{A}^{(k)}(M(\boldsymbol{Y}), Y_{ij})$ over $k \in [n]$ is exactly equal to the score attack for lower bounding the ℓ_2 minimax risk. When the reference to $\boldsymbol{Y}$ and M is clear, we denote $\mathcal{A}^{(k)}(M(\boldsymbol{Y}), Y_{ij})$ by $A_{ij}^{(k)}$.

Our plan for lower bounding the ℓ_∞ risk consists of upper bounding $\sum_{1 \leq i < j \leq n} \mathbb{E} A_{ij}^{(k)}$ by the ℓ_∞ risk and lower bounding the same quantity by a non-negative amount. The results of these steps are condensed in Propositions 2.3 and 2.4 respectively.

Proposition 2.3. *If M is an (ε, δ)-DP estimator with $0 < \varepsilon < 1$ and $p > 6 \log n / n$, then for sufficiently large n, every $\boldsymbol{\theta} \in \Theta$ and every $k \in [n]$, it holds that*

$$\sum_{1 \leq i < j \leq n} \mathbb{E}_{\boldsymbol{Y}|\boldsymbol{\theta}} A_{ij}^{(k)} \leq 4\kappa_1 np\varepsilon \cdot \mathbb{E}_{\boldsymbol{Y}|\boldsymbol{\theta}} |M(\boldsymbol{Y})_k - \theta_k| + 4\kappa_1 (n-1)\delta + 2\kappa_1 n^{-1}. \tag{2.8}$$

Proposition 2.3 is proved by considering $\widetilde{\boldsymbol{Y}}_{ij}$, an adjacent data set of $\boldsymbol{Y}$ obtained by replacing Y_{ij} with an independent copy. By differential privacy of algorithm M, $\mathbb{E} A_{ij}^{(k)}$ should be close to $\mathbb{E} A^{(k)}(M(\widetilde{\boldsymbol{Y}}_{ij}), Y_{ij})$, which is seen to be exactly 0 by the statistical independence of $M(\widetilde{\boldsymbol{Y}}_{ij})$ and Y_{ij}. The full details can be found in the supplement.

In the opposing direction, instead of a pointwise lower bound of $\sum_{1 \leq i < j \leq n} \mathbb{E}_{\boldsymbol{Y}|\boldsymbol{\theta}} A_{ij}^{(k)}$ at every $\boldsymbol{\theta} \in \Theta$, we lower bound the sum over a particular prior distribution of $\boldsymbol{\theta}$ over Θ.

Proposition 2.4. *Suppose M is an estimator of $\boldsymbol{\theta}$ such that $\sup_{\boldsymbol{\theta} \in \Theta} \mathbb{E} \|M(\boldsymbol{Y}) - \boldsymbol{\theta}\|_\infty < c$ for a sufficiently small constant $c > 0$. If each coordinate of $\boldsymbol{\theta}$ has density $\pi(t) = \mathbb{1}(|t| < 1)(15/16)(1 - t^2)^2$, then for every $k \in [n]$ there is some constant $C > 0$ such that*

$$\sum_{1 \leq i < j \leq n} \mathbb{E}_{\boldsymbol{\theta}} \mathbb{E}_{Y|\boldsymbol{\theta}} A_{ij}^{(k)} > C. \tag{2.9}$$

We defer the proof of Proposition 2.4 to the supplement, and combine Propositions 2.3 and 2.4 to arrive at a minimax risk lower bound in ℓ_∞ norm for estimating $\boldsymbol{\theta}$ with differential privacy.

Theorem 2.2. *If $p > 6\log n/n$, $\varepsilon \gtrsim (\log n)^{-1}$, $0 < \varepsilon < 1$ and $\delta \lesssim n^{-1}$, then*

$$\inf_{M \in M_{\varepsilon,\delta}} \sup_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_{\boldsymbol{Y}|\boldsymbol{\theta}} \|M(\boldsymbol{Y}) - \boldsymbol{\theta}\|_\infty \gtrsim \sqrt{\frac{\log n}{np}} + \frac{1}{np\varepsilon}. \tag{2.10}$$

The first term in the lower bound (2.10) is exactly the non-private minimax rate proved in [45, 19]. The second term is attributable to differential privacy.

The lower bound result given in Theorem 2.2 suggests that the perturbed MLE $\widetilde{\boldsymbol{\theta}}$ is essentially optimal except possibly a room of improvement by $O(\log n)$, but there is no implication about differentially private ranking algorithms not based on estimating the latent parameters $\boldsymbol{\theta}$. The next section considers differentially private ranking without relying on the parametric assumptions.

2.2 Nonparametric Estimation under Edge DP

If we drop the parametric assumption $\rho_{ij} = F(\theta_i^* - \theta_j^*)$, the estimand of interest shifts from $\boldsymbol{\theta}^*$ to the index set of top-k items $\mathcal{S}_k$ for $k \in [n-1]$ in terms of the average winning probability, $\tau_i = \frac{1}{n}\sum_{j \in [n]} \rho_{ij}$. In Section 2.2.1, we exhibit a differentially private estimator of $\mathcal{S}_k$ which exactly recovers $\mathcal{S}_k$ when $\tau_{(k)}$ and $\tau_{(k+1)}$ are sufficiently far apart,

$$|\tau_{(k)} - \tau_{(k+1)}| \gtrsim \sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon}.$$

It is not a coincidence that the requisite separation is identical to its parametric counterpart (2.7). We prove in Section 2.2.2 that this separation is the exact threshold for differentially private ranking in either parametric or nonparametric case.

Formally, consider the space of pairwise probability matrices

$$\Theta(k, m, c) = \left\{ \boldsymbol{\rho} \in [0,1]^{n \times n} : \boldsymbol{\rho} + \boldsymbol{\rho}^\top = 11^\top, \tau_{(k-m)} - \tau_{(k+m+1)} \geq c \left(\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon} \right) \right\},$$

for $k \in [n-1]$ and $0 \leq m \leq \min(k-1, n-k-1)$. Let $d_H(\cdot, \cdot)$ denote the Hamming distance between sets, and an estimator $\widehat{\mathcal{S}_k}$ succeeds at recovering $\mathcal{S}_k$ within tolerance m if

$$\sup_{\rho \in \Theta(k,m,c)} \mathbb{P}\left(d_H(\widehat{\mathcal{S}_k}, \mathcal{S}_k) > 2m\right) = o(1).$$

Exact recovery of $\mathcal{S}_k$ corresponds to $m = 0$. By adopting a similar framework to that of [45], we can directly compare the requisite threshold for top-k ranking with or without differential privacy.

2.2.1 The Edge DP Algorithm for Non-Parametric Estimation

[45] shows that the Copeland counting algorithm, which simply ranks the n items by their number of wins, exactly recovers the top-k items when the τ values of the true kth and $(k+1)$th items are separated by at least $O\left(\sqrt{\frac{\log n}{np}}\right)$. Algorithm 2 considers a differentially private version where the items are ranked by noisy numbers of wins.

Algorithm 2 Differentially Private Ranking for nonparametric models

Input: Comparison data $(y_{ij})_{(i,j)\in\mathcal{G}}$, comparison graph $\mathcal{G}$, privacy parameter ε.

1: Set $N_i = \sum_{j\neq i,(i,j)\in\mathcal{G}} \mathbb{1}(Y_{ij} = 1)$ denote the number of comparisons won by item i.

2: Generate

$$\boldsymbol{W} = (W_1, W_2, \cdots, W_n) \overset{\text{i.i.d.}}{\sim} \text{Laplace}\left(\frac{2}{\varepsilon}\right).$$

3: Compute the top-k set

$$\widetilde{\mathcal{S}_k} = \{i \in [n] : N_i + W_i \text{ is among the top } k \text{ largest of } \{N_j + W_j\}_{j\in[n]}\}.$$

Output: $\widetilde{\mathcal{S}_k}$.

The estimator $\widetilde{\mathcal{S}_k}$ defined in Algorithm 2 is $(\varepsilon, 0)$-DP by the Laplace mechanism [24]: the vector $(N_1, N_2, \cdots, N_n)$ has ℓ_1-sensitivity bounded by 2 over edge adjacent data sets, and the set $\widetilde{\mathcal{S}_k}$ is differentially private because it post-processes $\{N_j + W_j\}_{j\in[n]}$.

$\widetilde{\mathcal{S}_k}$ recovers $\mathcal{S}_k$ within tolerance m as long as $\tau_{(k-m)}, \tau_{(k+m+1)}$ are sufficiently separated.

Theorem 2.3. *For every $k \in [n-1]$ and any sufficiently large constant $C > 0$,*

$$\sup_{\boldsymbol{\rho} \in \Theta(k,m,C)} \mathbb{P}\left(d_H(\widetilde{\mathcal{S}_k}, \mathcal{S}_k) > 2m\right) < O(n^{-5}). \tag{2.11}$$

Specializing the theorem to $m = 0$ leads to the threshold for exact recovery.

Corollary 2.2. *For every $k \in [n-1]$, if the matrix of pairwise probabilities $\boldsymbol{\rho}$ is such that*

$$|\tau_{(k)} - \tau_{(k+1)}| \geq C \left(\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon} \right)$$

for a sufficiently large constant $C > 0$, we have $\mathbb{P}_{\boldsymbol{\rho}} \left(\widetilde{\mathcal{S}}_k \neq \mathcal{S}_k \right) < O(n^{-5})$.

As a further consequence, if $|\tau_{(k)} - \tau_{(k+1)}| \geq C \left(\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon} \right)$ for every k, then the union bound implies all n items can be correctly ranked with probability at least $1 - O(n^{-4})$. The next section shows this threshold is optimal in the sense that no differentially private algorithm can succeed at recovering $\mathcal{S}_k$ when $|\tau_{(k)} - \tau_{(k+1)}| < c \left(\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon} \right)$ for a sufficiently small c.

2.2.2 The Edge DP Lower Bound for Non-parametric Estimation

To establish the tightness of the threshold $\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon}$ for differentially private ranking, we shall prove that the supremum of $\mathbb{P}_{\boldsymbol{\rho}} \left(d_H(\widetilde{\mathcal{S}}_k, \mathcal{S}_k) > 2m \right)$ over the set of matrices

$$\Theta(k, m, c) = \left\{ \boldsymbol{\rho} \in [0,1]^{n \times n} : \boldsymbol{\rho} + \boldsymbol{\rho}^\top = 11^\top, \tau_{(k-m)} - \tau_{(k+m+1)} \geq c \left(\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon} \right) \right\},$$

is bounded away from 0 for sufficiently small c. In view of the lower bound, Theorem 2(b), in [45] where the supremum is taken over

$$\Theta^0(k, m, c) := \left\{ \boldsymbol{\rho} \in [0,1]^{n \times n} : \boldsymbol{\rho} + \boldsymbol{\rho}^\top = 11^\top, \tau_{(k-m)} - \tau_{(k+m+1)} \geq c\sqrt{\frac{\log n}{np}} \right\},$$

it suffices to show that the supremum of $\mathbb{P}_{\boldsymbol{\rho}} \left(d_H(\widetilde{\mathcal{S}}_k, \mathcal{S}_k) > 2m \right)$ over the set

$$\widetilde{\Theta}(k, m, 2c) = \left\{ \boldsymbol{\rho} \in [0,1]^{n \times n} : \boldsymbol{\rho} + \boldsymbol{\rho}^\top = 11^\top, \tau_{(k-m)} - \tau_{(k+m+1)} \geq 2c \frac{\log n}{np\varepsilon} \right\}$$

is bounded away from 0, because $\Theta(k, m, c) \subseteq \Theta^0(k, m, 2c) \cup \widetilde{\Theta}(k, m, 2c)$ for $c > 0$.

For proving the lower bound over $\widetilde{\Theta}(k, m, 2c)$, the differentially private Fano's inequality [6, 3] reduces the argument to choosing a number of different $\boldsymbol{\rho}$'s in $\widetilde{\Theta}(k, m, 2c)$ such that the distance among the distributions induced by the chosen $\boldsymbol{\rho}$'s is sufficiently small. We defer the details to the supplement [1] and state the lower bound result below.

Theorem 2.4. *Suppose the tolerance m is bounded by $2m \leq (1+\nu_2)^{-1} \min\{n^{1-\nu_1}, k, n-k\}$, $\frac{\log n}{np\varepsilon} < c_0$, and $\delta < c_0 \left(m \log n \cdot n^{10m}/\varepsilon\right)^{-1}$ for a sufficiently small constant c_0. There is a small constant $c(\nu_1, \nu_2)$ such that every (ε, δ)-DP estimator $\widehat{\mathcal{S}}_k$ satisfies*

$$\sup_{\boldsymbol{\rho} \in \widetilde{\Theta}(k,m,c)} \mathbb{P}_{\boldsymbol{\rho}} \left(d_H(\widehat{\mathcal{S}}_k, \mathcal{S}_k) > 2m \right) \geq \frac{1}{10} \tag{2.12}$$

whenever $c < c(\nu_1, \nu_2)$ and n is sufficiently large. The inequality remains true if $\boldsymbol{\rho} = (\rho_{ij})_{i,j \in [n]}$ is additionally restricted to the parametric model $\rho_{ij} = F(\theta_i^* - \theta_j^*)$, as long as F satisfies regularity condition (A0) in Section 2.1.

In conjunction with Theorem 2.3, Theorem 2.4 yields that $\widetilde{\mathcal{S}}_k$ is an optimal (ε, δ)-DP estimator. Setting $m = 0$ in Theorem 2.4 gives the lower bound for exactly recovering the top k items $\mathcal{S}_k$. In the exact recovery case, the threshold for full ranking of n items is when $|\tau_{(k)} - \tau_{(k+1)}| \geq C\left(\sqrt{\frac{\log n}{np}} + \frac{\log n}{np\varepsilon}\right)$ for every k.

Remark 1. Because the lower bound continues to hold when restricted to the parametric model, it in fact settles the $O(\log n)$ gap between the parametric upper bound Theorem 2.1 and the parametric lower bound Theorem 2.2. If $\omega_{ij} = F(\theta_i^* - \theta_j^*)$ for some F satisfying regularity conditions (A0) and (A1) in Section 2.1 and $\boldsymbol{\theta}^* \in \Theta$, we have $|\tau_{(k)} - \tau_{(k+1)}| \asymp |\theta^*_{(k)} - \theta^*_{(k+1)}|$. The existence of an (ε, δ)-DP estimator with a faster rate of convergence than $\widetilde{\boldsymbol{\theta}}$ would contradict the lower bound above for recovering $\mathcal{S}_k$. Under the parametric assumptions, the perturbed MLE $\widetilde{\boldsymbol{\theta}}$ is minimax optimal for estimating the parameters $\boldsymbol{\theta}^*$.

3 Optimal Ranking under Individual Differential Privacy

The ranking problem. In Section 2, we studied *edge* differential privacy, where each observed comparison between a fixed pair of items was protected.

We now consider a more realistic setting in which each of m users performs multiple item comparisons. Specifically, each user $k \in [m]$ selects L unordered pairs $\{i, j\} \subset [n]$ and reports the outcome of each comparison. We continue to assume that all users share the same underlying preference structure: there is a matrix $\rho = (\rho_{ij})_{1 \le i < j \le n}$ with $\rho_{ij} + \rho_{ji} = 1$, $\rho_{ii} = \frac{1}{2}$, so that whenever any user compares items i and j, item i is chosen (i.e., "preferred") with probability ρ_{ij} and item j with probability ρ_{ji}. By making ρ depend only on the item indices i, j, we assume all users draw comparisons from the same pairwise preference distribution.

Concretely, for each user $k = 1, \ldots, m$ and each draw $l = 1, \ldots, L$:

1. The user k selects an unordered pair $\{i, j\} \subset [n]$ uniformly at random.
2. They record $Y_{ij}^{(k,l)} = +1$ if i is preferred, or $Y_{ij}^{(k,l)} = -1$ if j is perferred.

All comparisons are independent across users and draws. The entire dataset is therefore

$$\mathcal{D} \;=\; \big\{\, Y_{ij}^{(k,l)} : 1 \le k \le m,\; 1 \le l \le L,\; 1 \le i < j \le n \big\}, \tag{3.1}$$

where each nonzero $Y_{ij}^{(k,l)}$ follows $\mathbb{P}\big(Y_{ij}^{(k,l)} = +1\big) \;=\; \rho_{ij}$, and $\mathbb{P}\big(Y_{ij}^{(k,l)} = -1\big) \;=\; \rho_{ji}$.

Our objective remains to rank the n items by their "average preference score" τ_i as defined in Section 2 and, for a given $k \in [n]$, recover the top k set.

Individual Differential Privacy. With the pairwise comparisons made by m users, it is natural to consider *all* L comparisons made by a single user, as opposed to a single pairwise comparison, as one unit of data. As the result, in this section we require the ranking algorithm M to satisfy that, for every event A,

$$\mathbb{P}(M(X) \in A) \le e^{\varepsilon} \cdot \mathbb{P}(M(X') \in A) + \delta$$

for any pair of data sets X and X' which differ by pairwise comparisons made by any single user. To differentiate this definition from the edge DP definition in Section 2, we say that such a ranking algorithm M satisfies (ε, δ) individual differential privacy.

Relationship between Edge DP and Individual DP. The critical, and only, difference between our two definitions of differential privacy is their respective notion of *adjacency*. In the edge DP formulation, any pair of data sets differing by a single comparison is adjacent, and an edge DP ranking algorithm is therefore required to behave similarly over them. In contrast, under individual DP, a pair of data sets is adjacent if their difference consists of up to L comparisons made by a single user, and an individual DP algorithm is required to behave similarly over any such pair. From an adversarial perspective, an edge DP algorithm prevents an adversary from inferring the presence of any single comparison in the input data, and an individual DP prevents an adversary from inferring the presence of a user.

Next, we build on the edge DP algorithms of Section 2 so that the entire bundle of L comparisons from each user is protected. Because each user's L draws can collectively influence counts or likelihoods, our noise scales must be adjusted accordingly, by roughly a multiplier of L compared to the edge DP case. In the next subsection, we show how to privatize the maximum likelihood estimator for parametric scores, and then we turn to a counting-based approach for nonparametric top-k recovery under this stronger, individual-level privacy constraint.

3.1 Parametric Estimation under Individual DP

We assume the same parametric model as in Section 2: each item $i \in [n]$ carries a latent score $\theta_i^* \in \mathbb{R}$, and for any pair (i, j), the probability that i is preferred over j is $\rho_{ij} = F\big(\theta_i^* - \theta_j^*\big)$, where $F : \mathbb{R} \to (0, 1)$ is a known, strictly increasing function satisfying $F(x) = 1 - F(-x)$. Across all m users, each user k draws L item-pairs (with replacement), and records the outcome $Y_{ij}^{(k,l)} \in \{+1, -1\}$ with probability ρ_{ij} and ρ_{ji} respectively, whenever they compare (i, j). All comparisons are independent across users and draws.

Define $M_{ij} \;=\; \big|\big\{\, k \in [m],\, l \in [L] : \text{user } k \text{ compares } (i,j) \big\}\big|$ and

$$\bar{Y}_{ij} \;=\; \frac{1}{M_{ij}} \sum_{\substack{k,l:\\ (i,j)\text{ compared}}} \frac{1 + Y_{ij}^{(k,l)}}{2} \;\in [0,1].$$

In words, M_{ij} counts how many times (i,j) was compared over all users, and $\bar{Y}_{ij}$ is the empirical fraction of times i was chosen over j. Then, up to additive constants independent of θ, the negative log-likelihood over all mL observations is

$$\mathcal{L}(\theta; Y) \;=\; \sum_{1 \le i < j \le n} M_{ij} \Big[-\bar{Y}_{ij}\, \log F_{ij}(\theta) \;-\; (1 - \bar{Y}_{ij})\, \log\big(1 - F_{ij}(\theta)\big) \Big],$$

where $F_{ij}(\theta) = F(\theta_i - \theta_j)$.

Individual DP estimator. Consider the perturbed MLE

$$\widetilde{\boldsymbol{\theta}} = \mathop{\arg\min}_{\boldsymbol{\theta} \in \mathbb{R}^n} \mathcal{L}(\boldsymbol{\theta}; y) + \frac{\gamma}{2} \|\boldsymbol{\theta}\|_2^2 + \boldsymbol{w}^\top \boldsymbol{\theta}, \quad \boldsymbol{w} = (w_1, w_2, \cdots, w_n) \overset{\text{i.i.d.}}{\sim} \text{Laplace}(\lambda). \tag{3.2}$$

Since each user contributes at most L comparisons per pair, our noise scales must scale with L. We have the following individual DP guarantee.

Proposition 3.1. *Suppose conditions (A0), (A1) and (A2) hold. If $\lambda \ge 8L \cdot \kappa_1/\varepsilon$ and $\gamma \ge 8L\kappa_2/\varepsilon$, $\widetilde{\boldsymbol{\theta}}$ as defined in Algorithm 1 is $(\varepsilon, 0)$ differentially private.*

With the noise scales set as above and under the usual sampling regime (m large enough so that $n \log n/(m\,\varepsilon) \;=\; O(1)$), a similar leave-one-out analysis as the edge DP case leads to the following accuracy result.

Theorem 3.1. *Assume (A0)–(A2) and fix $\varepsilon \in \big(c_2 (\log n)^{-1/2},\, 1\big)$, with L a constant. Choose $\gamma = 8\,L\,\kappa_2/\varepsilon$ and $\lambda = 8\,L\,\kappa_1/\varepsilon$. If m and n satisfy $n \log n/(m\,\varepsilon) = O(1)$, then $\widetilde{\boldsymbol{\theta}}$ is $(\varepsilon, 0)$-DP, and it holds with probability at least $1 - O(n^{-5})$ that*

$$\|\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\|_\infty \;\lesssim\; \sqrt{\frac{n\, \log n}{m}} \;+\; \frac{n\, \log n}{m\, \varepsilon}.$$

In particular the first term matches the non-private minimax rate $\sqrt{n \log n/m}$, and the second term quantifies the cost of individual DP.

Minimax lower bound. We can further show that no (ε, δ)-DP estimator can achieve a smaller worst-case ℓ_∞ error. By the same reduction argument from parametric estimation lower bound to top k ranking lower bound made in Remark 1 in Section 2, we shall prove a ranking lower bound in the nonparametric case, which then implies the following estimation lower bound.

Theorem 3.2. *If $\sqrt{(n \log n)/m} + (n \log n)/(m\,\varepsilon) < c_0$ and $\delta \lesssim n^{-1-\omega}$ for some $\omega > 0$, then any (ε, δ)-DP estimator $\widehat{\theta}$ must satisfy*

$$\sup_{\|\boldsymbol{\theta}\|_\infty \le 1} \mathbb{E}\big\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\big\|_\infty \gtrsim \sqrt{\frac{n \log n}{m}} + \frac{n \log n}{m\,\varepsilon}.$$

Hence $\widetilde{\boldsymbol{\theta}}$ of Theorem 3.1 is minimax optimal up to constants.

3.2 Nonparametric Ranking under Individual DP

Next, we drop any parametric assumption on ρ_{ij} and directly recover the top-k set based on $\tau_i = \frac{1}{n}\sum_{j=1}^{n} \rho_{ij}$.

Count-based estimator. Let N_i be the total wins of item i. Replacing one user's L comparisons can change each N_i by at most L, and we therefore add i.i.d. Laplace(L/ε) noise to each coordinate.

By the Laplace mechanism [24], the resulting noisy counts $\widetilde{N}_i$ in Algorithm 3 are $(\varepsilon, 0)$-DP, and ranking items by $\{\widetilde{N}_i\}$ preserves individual-level differential privacy.

Algorithm 3 Individual-DP Top-k via Noisy Counts

Input: All comparisons $\{Y^{(k,l)}\}$ and privacy parameter $\varepsilon > 0$.

1: For each $i \in [n]$, compute $N_i = \sum_{j \neq i} \sum_{k=1}^{m} \sum_{l=1}^{L} \mathbb{1}\big(Y_{ij}^{(k,l)} = +1\big)$.
2: Draw $W_1, \dots, W_n \overset{\text{i.i.d.}}{\sim} \text{Laplace}\big(L/\varepsilon\big)$.
3: Form $\widetilde{N}_i = N_i + W_i$ for $i = 1, \dots, n$.
4: Output $\widetilde{\mathcal{S}}_k =$ the indices of the top k values among $\{\widetilde{N}_i\}$.

Output: $\widetilde{\mathcal{S}}_k$.

Fix a tolerance $u \in \{0, 1, \ldots, \min(k-1, n-k-1)\}$. Define

$$\Theta(k,u,c) = \Big\{\rho : \ [0,1]^{n\times n}, \ \rho_{ij} + \rho_{ji} = 1, \ \tau_{(k-u)} - \tau_{(k+u+1)} \ \geq \ c\Big(\sqrt{\tfrac{n\log n}{m}} + \tfrac{n\log n}{m\,\varepsilon}\Big)\Big\},$$

where $\tau_{(1)} \geq \tau_{(2)} \geq \cdots \geq \tau_{(n)}$. We show that the noisy counts succeed at recovering the top items with overwhelming probability.

Theorem 3.3. *There exists $C > 0$ such that, for any k and u, if $\rho \in \Theta(k, u, C)$, we have*

$$\mathbb{P}\big(d_H(\widetilde{\mathcal{S}}_k, \mathcal{S}_k) > 2u\big) \ = \ O(n^{-5}).$$

In particular, when $u = 0$, exact recovery holds with probability $1 - O(n^{-5})$ provided that

$$\tau_{(k)} - \tau_{(k+1)} \ \geq \ C\Big(\sqrt{\tfrac{n\log n}{m}} + \tfrac{n\log n}{m\,\varepsilon}\Big).$$

As a further consequence, if $|\tau_{(k)} - \tau_{(k+1)}| \geq C\left(\sqrt{\frac{n\log n}{m}} + \frac{n\log n}{m\varepsilon}\right)$ for every k, then the union bound implies all n items can be correctly ranked with probability at least $1 - O(n^{-4})$. The next theorem shows this threshold is optimal in the sense that no differentially private algorithm can succeed at recovering $\mathcal{S}_k$ when $|\tau_{(k)} - \tau_{(k+1)}| < c\left(\sqrt{\frac{n\log n}{m}} + \frac{n\log n}{m\varepsilon}\right)$ for a sufficiently small constant c.

Nonparametric ranking lower bound. By a DP Fano argument analogous to [6, 3], no (ε, δ)-DP procedure can reliably recover $\mathcal{S}_k$ when the gap $\tau_{(k)} - \tau_{(k+1)}$ is smaller.

Theorem 3.4. *Assume $\sqrt{(n\log n)/m} + (n\log n)/(m\,\varepsilon) < c_0$ and $\delta < c_0\big(u\log n\, n^{10u}/\varepsilon\big)^{-1}$. Then for sufficiently small c and large n,*

$$\inf_{M\in\mathcal{M}_{\varepsilon,\delta}} \sup_{\rho\in\Theta(k,u,c)} \mathbb{P}\big(d_H(M(Y), \mathcal{S}_k) > 2u\big) \ \geq \ \tfrac{1}{10}.$$

In particular, exact recovery ($u = 0$) fails if for some k, $\tau_{(k)} - \tau_{(k+1)} < c\big(\sqrt{(n\log n)/m} + (n\log n)/(m\,\varepsilon)\big)$.

Together, Theorems 3.3 and 3.4 show that the requirement $\sqrt{(n \log n)/m} + (n \log n)/(m\,\varepsilon)$ is both necessary and sufficient for privately recovering the top-k set under individual DP.

4 Numerical Experiments

We conduct three sets of experiments to evaluate the numerical performance of our algorithms and to guide their practical application. [1]

In Sections 4.1 and 4.2, we use simulated data to study the accuracy of edge DP and individual DP algorithms respectively. The numerical results qualitatively confirm the dependence of estimation and ranking errors on all parameters: the number of items n, the number of individuals m, sampling probability p. Notably, the results support the most striking difference between the rates of convergence under edge DP and their counterparts in individual DP, namely that the former is a decreasing function of the number of items n but the latter is increasing in n.

To complement these theoretical and simulation studies, Section 4.3 applies the individual DP ranking algorithms to two real-world datasets: a student preference dataset from [22] and an immigration attitude survey from [53]. Both datasets naturally align with the individual DP framework, as each involves individuals comparing multiple pairs of items. In such contexts, participants' opinions–on topics such as university preferences or immigration policy–are sensitive and merit privacy protection.

For practical applications of our ranking algorithms, we highlight three findings from the experiment results.

- For recovering ranks, the non-parametric, counting based algorithm generally outperform the parametric estimation algorithm, in both edge and individual DP and across all settings of n, p, m, and ε. Both the parametric and nonparametric algorithms significantly outperform the one-shot algorithm in [40] in our numerical experiments.

- The individual DP algorithms are accurate when the number of items is small but the number

[1] The code for reproducing the numerical experiments is at https://github.com/abhinavc3/DP-Ranking.

of comparisons is large. The edge DP algorithms, in contrast, perform well when both the number of items and the number of comparisons are large. With the caveat that the privacy protections offered by edge DP and individual DP are quite different, this observation may nevertheless be relevant when there is flexibility in choosing which DP framework to apply to the data analysis task.

- The estimation errors at $\varepsilon = 2.5$ or greater tend to be much closer to the non-private estimation errors than to the estimation errors at, say $\varepsilon = 1$, possibly suggesting that choosing $\varepsilon = 2.5$ is a promising choice for balancing the privacy and utility of our algorithms. While the ε values in different contexts and applications are not inherently comparable [4], a privacy parameter of $\varepsilon = 2.5$ is smaller than or comparable with the ε values adopted by prominent adopters of DP, such as Apple, LinkedIn, or the US Census Bureau [21].

4.1 Simulated Data Experiments under Edge DP

4.1.1 Setup

Data Generation The pairwise comparison outcomes are sampled from the BTL model, with various values of the number of items n and sampling probability p as specified in the experiments below. We fix the value of $k = n/4$ and generate our θ_i, up to centering, by

$$e^{\theta_i} \sim \begin{cases} 1 & \text{if} \quad i < k, \\ \text{Unif}(0.2, 0.7) & \text{if} \quad i \geq k. \end{cases}$$

Evaluation Metrics for Parameter Estimation: For evaluating the parametric estimation algorithms, we consider the ℓ_∞ and ℓ_2 relative errors on logarithmic scale, $\log\left(\frac{\|\hat{\theta}-\theta^*\|_\infty}{\|\theta^*\|_\infty}\right)$ and $\log\left(\frac{\|\hat{\theta}-\theta^*\|_2}{\|\theta^*\|_2}\right)$, where $\hat{\theta}$ is the estimator and θ^* is the true parameter.

Evaluation Metric for Top-k set recovery: Under both the parametric and nonparametric models, we evaluate the performance of top-k recovery by one minus the size of overlap between the estimator and the truth, $1 - \frac{|\widehat{\mathcal{S}}_k \cap \mathcal{S}_k|}{k}$, where $\mathcal{S}_k$ is the true top-k set and $\widehat{\mathcal{S}}_k$ is an estimator.

4.1.2 Experiments

Experiment 1. We study the number of items n's effect on the accuracy of our estimator (Figure 1). The sampling probability p is fixed at 1, and we consider four privacy levels $\varepsilon \in \{0.5, 1, 2.5, \infty\}$. All loss functions decrease as n increases, demonstrating the consistency of our suggested approaches. It is noteworthy that, for top-k set recovery, the nonparametric Copeland algorithm outperforms the penalized-MLE in both the private and non-private regimes.

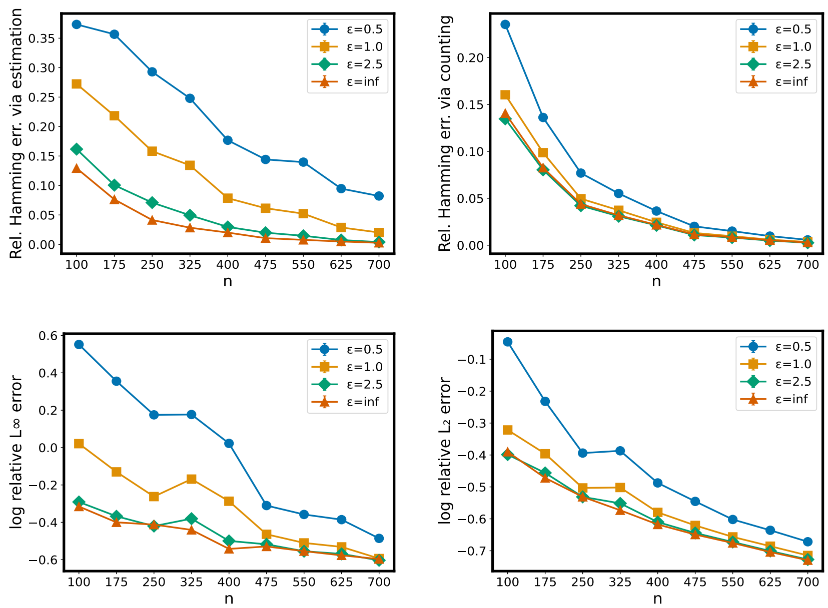

Figure 1: Edge DP estimation errors versus the number of items n at various privacy levels.

Experiment 2. We investigate the effect of edge probability p on the accuracy of the proposed methods (Figure 2). The sample size is fixed at $n = 300$, and ε varies across four different levels $\{0.5, 1, 2.5, \infty\}$. As p increases, we observe more pairwise comparisons, effectively increasing the sample size and leading to better performance.

Experiment 3. Here we investigate the effect of privacy parameter ε on the accuracy of our methods (Figure 3). The sample size is fixed at $n = 300$, and the sampling probability p varies across four levels $\{0.25, 0.5, 0.75, 1\}$. Increasing ε reduces the errors.

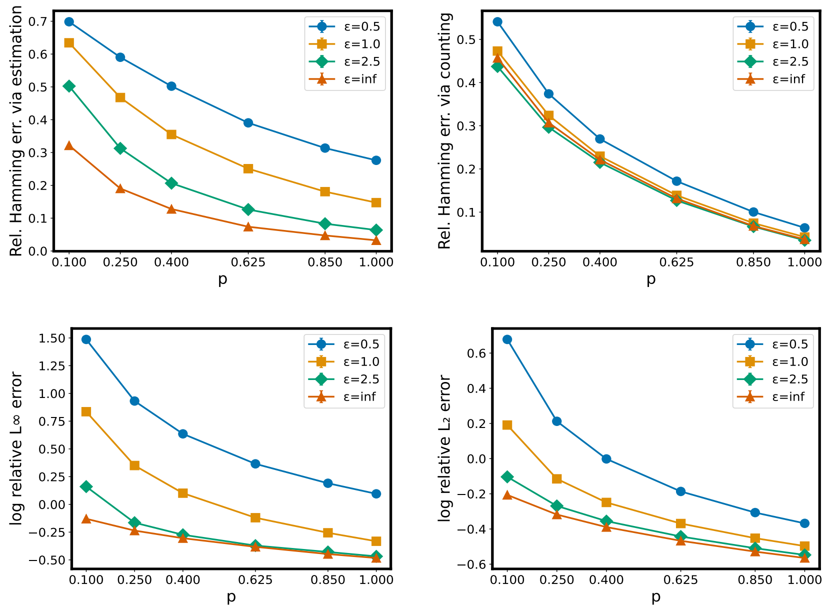

Figure 2: Edge DP estimation errors versus the edge probability p at various privacy levels.

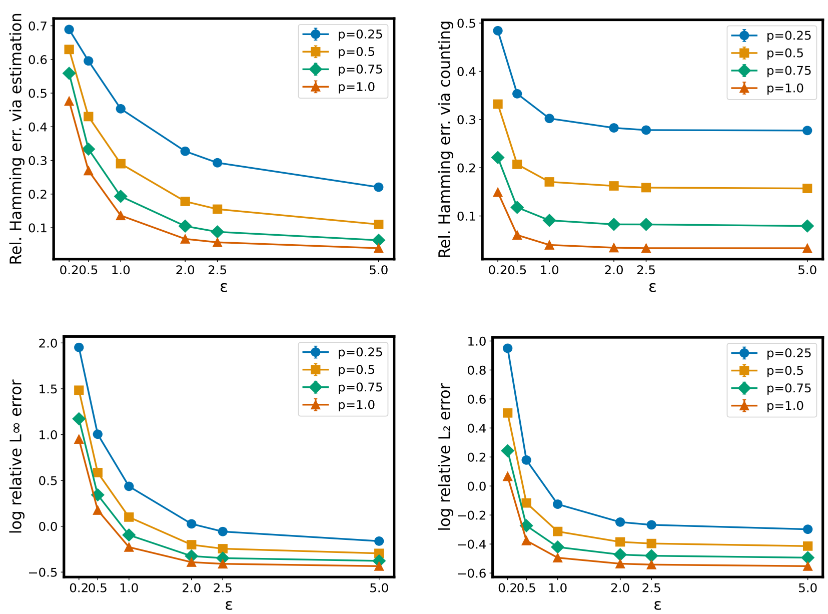

Figure 3: Edge DP errors versus the privacy parameter ε at various edge probability values.

Experiment 4. We compare the non-parametric and parametric algorithms with the one-shot algorithm proposed by [40] at fixed $p = 1$ and various n and ε values (Figure 4). In terms of relative Hamming errors, our algorithms significantly outperform the one-shot algorithm.

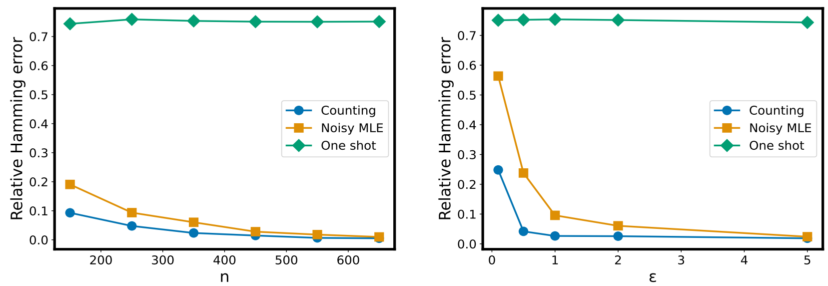

Figure 4: Relative Hamming errors of our algorithms and the one-shot algorithm in [40] at $p = 1$ and various levels of n and ε.

4.2 Simulated Data Experiments under Individual DP

For individual DP simulations, we keep the generation of θ and evaluation metrics the same as in Section 4.1, and study the accuracy impact of varying the number of items n, the number of individuals m, as well as the privacy level ε. The number of comparisons contributed by each individual, consistent with the assumption of our theoretical analysis, is chosen to be an absolute constant $L = 5$.

Similar to the edge DP case, the theoretical dependence of estimation errors in n, m and ε is corroborated by the numerical results. In contrast to Experiment 1 in Section 4.1, the estimation errors under individual DP increases with the number of items n. The role of m in individual DP is similar to that of edge probability p in edge DP; the behavior of estimation errors versus ε is also similar to the edge DP case.

Experiment 5. We study the number of items n's effect on the estimation errors (Figure 5). The number of individuals is fixed at $m = 1000$, and we consider four privacy levels $\varepsilon \in \{0.5, 1, 2.5, \infty\}$.

Experiment 6. We study the effect of varying the number of individuals m (Figure 6). The sample size is fixed at $n = 16$, and ε varies across four different levels $\{0.5, 1, 2.5, \infty\}$.

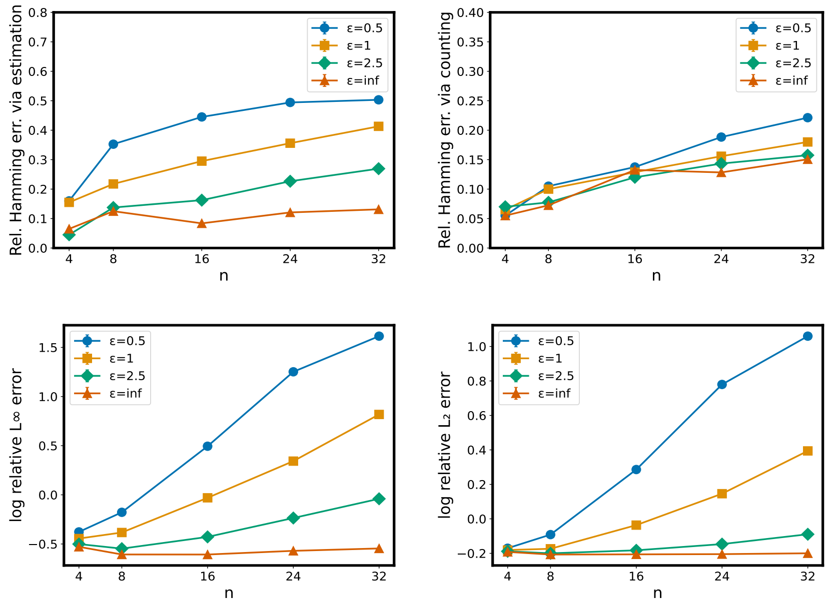

Figure 5: Individual DP errors versus the number of items n at various privacy levels.

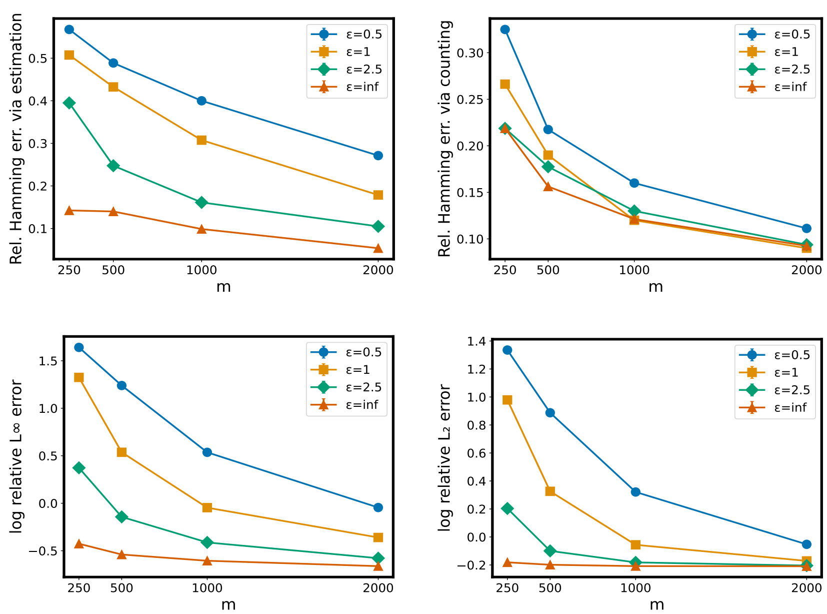

Figure 6: Individual DP errors versus the number of individuals m at various privacy levels.

Experiment 7. We investigate the effect of privacy parameter ε on the accuracy (Figure 7). The sample size is fixed at $n = 16$, and the sampling probability m varies across four levels $\{250, 500, 1000, 2000\}$.

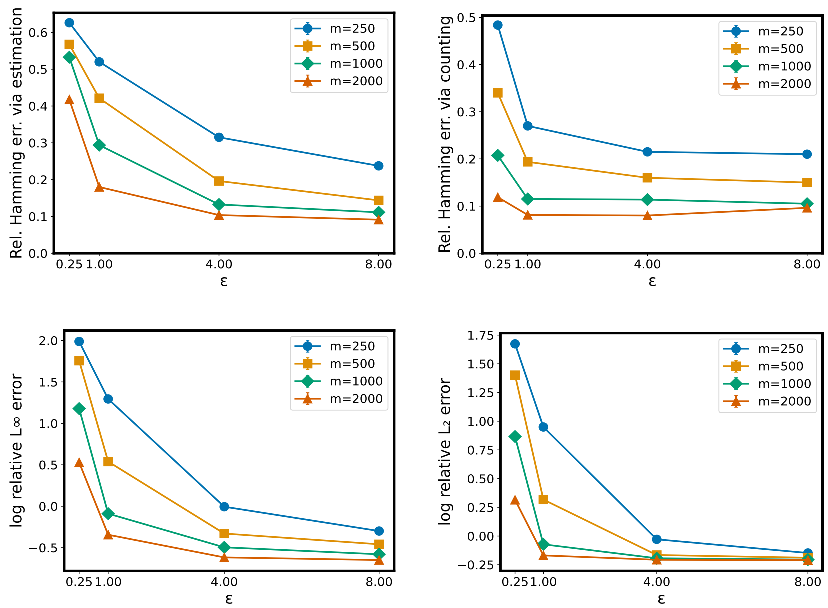

Figure 7: Individual DP errors versus the privacy level ε at various levels of m.

4.3 Real Data Analysis under Individual DP

Now we move beyond simulations and study the effectiveness of our algorithms on real data sets. As real data sets do not have "true" rankings, we measure the accuracy of our algorithms by the average difference in ranks produced by the DP algorithm versus those by its non-DP counterpart, which quantifies the loss of accuracy attributable to differential privacy constraints in the practical task of ranking items.

4.3.1 Data Sets

University Preferences The university preference data set [22] is collected in a survey conducted among students in the "Community of European Management Schools" (CEMS) program by the Vienna University of Economics. The data set consists of observations from 303 students ($m = 303$) and records their preference between pairs of European universities for their semester abroad. For each student, a total of 15 ($L = 15$) pairwise comparisons between 6 universities

($n = 6$) were asked for, and then an overall ranking of all universities was derived using the comparison outcomes.

Student Attitudes on Immigration This dataset is collected in a survey conducted by [53] to understand public opinions on immigration. The survey collected responses from 98 students ($m = 98$), each agreed to answer at least one paired comparison drawn from a pool of four extreme statements about immigrants ($n = 4, L = \binom{4}{2} = 6$).

4.3.2 Results

In both examples, the individual DP algorithms can produce ranks close to the non-DP estimated ranks. As the non-DP, noiseless algorithms are known to be optimal without differential privacy [45, 19], it is further implied that the DP ranks are also of high quality. Although our theoretical results do not encompass the mean absolute difference in ranks, this metric's behavior versus the privacy parameter ε is as expected: lower ε, namely stronger privacy level, results in noisier ranks.

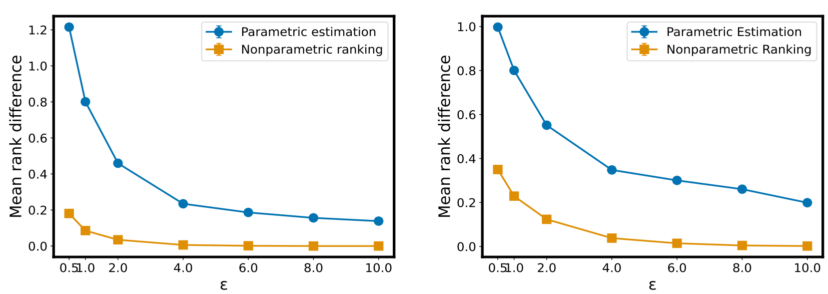

Figure 8: Mean rank difference versus ε. Left: CEMS data. Right: immigration data.

Similar to the simulated data case, the real data results exhibit clear advantages of the non-parametric ranks over the ranks based on parametric estimation at the same privacy level. It therefore appears that, if differentially private ranking is of the main interest, one should generally prefer the non-parametric method. However, the parametric method offers a notable benefit: it provides differentially private estimates of pairwise winning probabilities at no additional cost, since these are deterministic functions of the parameters under the BTL model.

5 Discussion

In this paper, we proposed differentially private algorithms for ranking based on pairwise comparisons, analyzed their rates of convergence, and established their optimality among all differentially private ranking procedures. Our theoretical results were supported by numerical experiments on both real and simulated datasets. We conclude by outlining several directions for future research.

Further theoretical analysis of individual DP. A comparison between edge DP and individual DP reveals intriguing differences in how privacy constraints impact statistical accuracy. For example, under edge DP, the rate of convergence improves as the number of items n increases, whereas under individual DP, it worsens. This divergence highlights fundamentally different statistical costs associated with the two privacy models. However, our theoretical treatment of individual DP remains less comprehensive than that of edge DP, largely due to the simplifying assumption that the number of comparisons L contributed by each individual is fixed and known. Removing or relaxing this assumption may lead to a deeper understanding of individual DP in more realistic settings.

The cost of approximate and pure differential privacy. Interestingly, the optimal (ε, δ)-DP ranking algorithms proposed in this paper also satisfy the stronger $(\varepsilon, 0)$-DP, possibly implying that, for our ranking problem, the cost of "pure" differential privacy is not higher than that of "approximate" differential privacy, when measured by the minimax rate of convergence and ignoring the exact constants in the minimax risk. This phenomenon stands in contrast with differentially private (Gaussian) mean estimation in high dimensions [6, 48, 13], where the optimal rate of convergence with (ε, δ)-DP explicitly depends on δ. It is an interesting theoretical question to understand the conditions under which approximate (ε, δ)-DP is strictly less costly to statistical accuracy than pure $(\varepsilon, 0)$-DP, in terms of either the minimax rate or better constants at the same minimax rate.

Lower bound technique for entry-wise parametric estimation. Under the parametric model, a single dose of noise added to the objective function simplifies the privacy analysis, and avoids potentially higher privacy cost incurred by iterative noise addition required by methods such as noisy gradient descent. The entry-wise error analysis of the perturbed MLE is potentially applicable in other statistical problems where entry-wise or ℓ_∞ errors are of primary interest. On the lower bound side, the entry-wise version of score attack in Section 2.1.2 results in a $O(\log n)$ gap from the optimal lower bound in Section 2.2.2. One would wonder if this method can be strengthened to eliminate such gaps.

Data Availability Statement

The real datasets used in this study are publicly available in the `prefmod` package for R. Specifically, the university preference dataset from [22] is available as `cemspc`, and the student attitudes on immigration dataset from [53] is available as `immig`.

References

[1] Supplement to "optimal differentially private ranking from pairwise comparisons", 2023. Supplemental Material.

[2] J. Acharya, Z. Sun, and H. Zhang. Differentially private testing of identity and closeness of discrete distributions. In *Adv. Neural Inf. Process. Syst.*, pages 6878–6891, 2018.

[3] J. Acharya, Z. Sun, and H. Zhang. Differentially private Assouad, Fano, and Le Cam. In *Algorithmic Learning Theory*, pages 48–78. PMLR, 2021.

[4] J. Bailie, R. Gong, and X.-L. Meng. A refreshment stirred, not shaken (iii): Can swapping be differentially private? In *Data Privacy Protection and the Conduct of Applied Research: Methods, Approaches and Their Consequences*. University of Chicago Press, May 2025.

[5] S. Balakrishnan and S. Chopra. Two of a kind or the ratings game? adaptive pairwise preferences and latent factor models. *Frontiers of Computer Science*, 6(2):197–208, 2012.

[6] R. F. Barber and J. C. Duchi. Privacy and statistical risk: Formalisms and minimax bounds. *arXiv preprint arXiv:1412.4451*, 2014.

[7] R. Bassily, V. Feldman, K. Talwar, and A. G. Thakurta. Private stochastic convex optimization with optimal rates. In *Adv. Neural Inf. Process. Syst.*, pages 11282–11291, 2019.

[8] R. Bassily, A. Smith, and A. Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *FOCS 2014*, pages 464–473. IEEE, 2014.

[9] R. A. Bradley and M. E. Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

[10] M. Bun, J. Ullman, and S. Vadhan. Fingerprinting codes and the price of approximate differential privacy. In *STOC 2014*, pages 1–10. ACM, 2014.

[11] T. T. Cai, A. Chakraborty, and L. Vuursteen. Optimal federated learning for nonparametric regression with heterogeneous distributed differential privacy constraints. *arXiv preprint arXiv:2406.06755*, 2024.

[12] T. T. Cai, Y. Wang, and L. Zhang. The cost of privacy in generalized linear models: Algorithms and minimax lower bounds. *arXiv preprint arXiv:2011.03900*, 2020.

[13] T. T. Cai, Y. Wang, and L. Zhang. The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, 49(5):2825–2850, 2021.

[14] T. T. Cai, Y. Wang, and L. Zhang. Score attack: A lower bound technique for optimal differentially private learning. *arXiv preprint arXiv:2303.07152*, 2023.

[15] C. L. Canonne, G. Kamath, A. McMillan, A. Smith, and J. Ullman. The structure of optimal private tests for simple hypotheses. *arXiv preprint arXiv:1811.11148*, 2018.

[16] K. Chaudhuri and C. Monteleoni. Privacy-preserving logistic regression. In *Adv. Neural Inf. Process. Syst.*, pages 289–296, 2009.

[17] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(Mar):1069–1109, 2011.

[18] X. Chen, S. Gopi, J. Mao, and J. Schneider. Competitive analysis of the top-k ranking problem. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1245–1264. SIAM, 2017.

[19] Y. Chen, J. Fan, C. Ma, and K. Wang. Spectral method and regularized mle are both optimal for top-k ranking. *Ann. Statist.*, 47(4):2204, 2019.

[20] Y. Chen and C. Suh. Spectral mle: Top-k rank aggregation from pairwise comparisons. In *ICML 2015*, pages 371–380. PMLR, 2015.

[21] D. Desfontaines. A list of real-world uses of differential privacy. https://desfontain.es/blog/real-world-differential-privacy.html, 2021. Ted is writing things (personal blog).

[22] R. Dittrich, R. Hatzinger, and W. Katzenbeisser. Modelling the effect of subject-specific covariates in paired comparison studies with an application to university rankings. *J. R. Statist. Soc. C*, 47(4):511–525, 1998.

[23] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 486–503. Springer, 2006.

[24] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *TCC 2006*, pages 265–284. Springer, 2006.

[25] C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[26] C. Dwork, A. Smith, T. Steinke, J. Ullman, and S. Vadhan. Robust traceability from trace amounts. In *FOCS 2015*, pages 650–669. IEEE, 2015.

[27] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang. Analyze gauss: optimal bounds for privacy-preserving principal component analysis. In *STOC 2014*, pages 11–20. ACM, 2014.

[28] L. R. Ford Jr. Solution of a ranking problem from binary comparisons. *The American Mathematical Monthly*, 64(8P2):28–33, 1957.

[29] M. Hay, L. Elagina, and G. Miklau. Differentially private rank aggregation. In *Proceedings of the 2017 SIAM International Conference on Data Mining*, pages 669–677. SIAM, 2017.

[30] S. Heldsinger and S. Humphry. Using the method of pairwise comparison to obtain reliable teacher assessments. *The Australian Educational Researcher*, 37(2):1–19, 2010.

[31] G. Kamath, J. Li, V. Singhal, and J. Ullman. Privately learning high-dimensional distributions. *arXiv preprint arXiv:1805.00216*, 2018.

[32] V. Karwa and S. Vadhan. Finite sample differentially private confidence intervals. *arXiv preprint arXiv:1711.03908*, 2017.

[33] S. P. Kasiviswanathan, K. Nissim, S. Raskhodnikova, and A. Smith. Analyzing graphs with node differential privacy. In *Theory of Cryptography Conference*, pages 457–476. Springer, 2013.

[34] D. Kifer, A. Smith, and A. Thakurta. Private convex empirical risk minimization and high-dimensional regression. In *COLT 2012*, pages 25.1–25.40, 2012.

[35] J. Lei. Differentially private M-estimators. In *NeurIPS 2011*, pages 361–369, 2011.

[36] R. D. Luce. *Individual Choice Behavior.* John Wiley, 1959.

[37] S. Negahban, S. Oh, and D. Shah. Rank centrality: Ranking from pairwise comparisons. *Operations Research*, 65(1):266–287, 2017.

[38] K. Nissim, S. Raskhodnikova, and A. Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 75–84, 2007.

[39] A. Pananjady, C. Mao, V. Muthukumar, M. J. Wainwright, and T. A. Courtade. Worst-case versus average-case design for estimation from partial pairwise comparisons. *Ann. Statist.*, 48(2):1072–1097, 2020.

[40] G. Qiao, W. Su, and L. Zhang. Oneshot differentially private top-k selection. In *International Conference on Machine Learning*, pages 8672–8681. PMLR, 2021.

[41] A. Rajkumar and S. Agarwal. A statistical convergence perspective of algorithms for rank aggregation from pairwise data. In *ICML 2014*, pages 118–126. PMLR, 2014.

[42] N. Shah, S. Balakrishnan, J. Bradley, A. Parekh, K. Ramchandran, and M. Wainwright.

Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. In *Artificial Intelligence and Statistics*, pages 856–865. PMLR, 2015.

[43] N. Shah, S. Balakrishnan, A. Guntuboyina, and M. Wainwright. Stochastically transitive models for pairwise comparisons: Statistical and computational issues. In *ICML 2016*, pages 11–20. PMLR, 2016.

[44] N. B. Shah, S. Balakrishnan, and M. J. Wainwright. Feeling the bern: Adaptive estimators for bernoulli probabilities of pairwise comparisons. *IEEE Trans. Inf. Theory*, 65(8):4854–4874, 2019.

[45] N. B. Shah and M. J. Wainwright. Simple, robust and optimal ranking from pairwise comparisons. *J. Mach. Learn. Res.*, 18(1):7246–7283, 2017.

[46] S. Shang, T. Wang, P. Cuff, and S. Kulkarni. The application of differential privacy for rank aggregation: Privacy and accuracy. In *17th International Conference on Information Fusion (FUSION)*, pages 1–7. IEEE, 2014.

[47] B. Song, Q. Lan, Y. Li, and G. Li. Distributed differentially private ranking aggregation. *arXiv preprint arXiv:2202.03388*, 2022.

[48] T. Steinke and J. Ullman. Between pure and approximate differential privacy. *Journal of Privacy and Confidentiality*, 7(2), 2017.

[49] T. Steinke and J. Ullman. Tight lower bounds for differentially private selection. In *FOCS 2017*, pages 552–563. IEEE, 2017.

[50] L. L. Thurstone. A law of comparative judgment. *Psychol. Rev.*, 34(4):273, 1927.

[51] L. Wasserman and S. Zhou. A statistical framework for differential privacy. *J. Am. Stat. Assoc.*, 105(489):375–389, 2010.

[52] F. Wauthier, M. Jordan, and N. Jojic. Efficient ranking from pairwise comparisons. In *ICML 2013*, pages 109–117. PMLR, 2013.

[53] D. Weber and R. Hatzinger. A novel approach for modelling paired comparisons data with

non-ignorable missing values on student's attitudes towards foreigners. *Data Analysis Bulletin*, 12:11–22, 2011.

[54] Z. Yan, G. Li, and J. Liu. Private rank aggregation under local differential privacy. *International Journal of Intelligent Systems*, 35(10):1492–1519, 2020.