

---

# Optimal and Private Learning from Human Response Data

---

Duc Nguyen

Anderson Y. Zhang

Department of Computer & Information Science

Department of Statistics & Data Science

University of Pennsylvania

University of Pennsylvania

## Abstract

Item response theory (IRT) is the study of how people make probabilistic decisions, with diverse applications in education testing, recommendation systems, among others. The Rasch model of *binary response data*, one of the most fundamental models in IRT, remains an active area of research with important practical significance. Recently, Nguyen and Zhang (2022) proposed a new spectral estimation algorithm that is efficient and accurate. In this work, we extend their results in two important ways.

Firstly, we obtain a refined *entrywise error bound* for the spectral algorithm, complementing the ‘average error’  $\ell_2$  bound in their work. Notably, under mild sampling conditions, the spectral algorithm achieves the minimax optimal error bound (modulo a log factor). Building on the refined analysis, we also show that the spectral algorithm enjoys optimal sample complexity for top- $K$  recovery (e.g., identifying the best  $K$  items from approval/disapproval response data), explaining the empirical findings in the previous work.

Our second contribution addresses an important but understudied topic in IRT: privacy. Despite the human-centric applications of IRT, there has not been any proposed privacy-preserving mechanism in the literature. We develop a private extension of the spectral algorithm, leveraging its unique Markov chain formulation and the discrete Gaussian mechanism (Canonne et al., 2020). Experiments show that our approach is significantly more

accurate than the baselines in the low-to-moderate privacy regime.

## 1 INTRODUCTION

Item response theory (IRT) was originally developed within the psychometric community in the early 60s (Rasch, 1960; Van der Linden and Hambleton, 1997) as a tool to model patients’ responses to psychological tests. Since its conception, item response theory has been applied to diverse settings including education testing (Lord, 2012), crowdsourcing (Whitehill et al., 2009), recommendation systems (Chen et al., 2005), finance (Schellhorn and Sharma, 2013), marketing (Brzezińska et al., 2016) and more recently the evaluation of machine learning algorithms (Moraes et al., 2020; Chen and Ahn, 2020; Martínez-Plumed et al., 2019).

One of the most fundamental models in IRT is the Rasch model (Rasch, 1960). It assumes that the *binary response*  $X_{li} \in \{0, 1\}$  of person  $l$  with latent characteristic  $\theta_l^* \in \mathbb{R}$  to item  $i$  with latent parameter  $\beta_i^* \in \mathbb{R}$  is given as follows.

$$\mathbb{P}(X_{li} = 1) = \frac{1}{1 + \exp(-(\theta_l^* - \beta_i^*))}. \quad (1)$$

As an example in education testing,  $\theta_l^*$  corresponds to the ability of student  $l$  and  $\beta_i^*$  the difficulty of problem  $i$ . The random response variable  $X_{li}$  is whether the student correctly solves the problem.

Recently, Nguyen and Zhang (2022) proposed a new spectral algorithm for *one-sided* parameter estimation under the Rasch model (i.e., estimate  $\beta^*$ ). Experimentally, the spectral algorithm is significantly faster than its competitors while being comparatively accurate. As the spectral algorithm is entirely new, we identify two important open directions from the previous work.

**Entrywise Error Guarantee.** The guarantee in

Nguyen and Zhang (2022) is an  $\ell_2$  error bound for parameter estimation (cf., Theorem 3.3), which can be roughly considered the average estimation error. In certain domains, we might desire a small entrywise error guarantee. For example, consider a conference reviewing system where reviewers vote to reject or accept papers. We would like to accurately identify the top papers from reject/accept data. One approach is to estimate the quality parameters of the papers and accept  $K$  papers with the highest parameter values<sup>1</sup>. The  $\ell_2$  error does not indicate how the estimation error distributes among the papers. The average error may still be low if all the error is concentrated in a single estimate. Really good papers could get rejected or sub-par papers get accepted. On the other hand, a small entrywise parameter estimation error ensures that the quality for each paper is accurately measured.

**Privacy.** The study of privacy-preserving mechanisms in IRT is understudied. This is quite concerning because many item response theory models and algorithms are deployed on human response data containing sensitive information (e.g., a student passing a test, a voter supporting a legislation). One may consider the binary nature of the response data and propose randomized response (Warner, 1965) as a natural solution. However, as we will show in our experiments, even when the number of responses per person is moderate  $m \approx 10$ , the amount of noise that randomized response needs to inject to ensure sufficient privacy is too large for any downstream algorithm to obtain accurate parameter estimates. On the other hand, analyzing the sensitivity (cf. Definition 3.1 in Dwork et al. (2014)) of the algorithms used in the IRT literature including optimization based algorithms (Andersen, 1973; Fischer, 1981; Haberman, 1977; Hambleton et al., 1991) is substantially more difficult. This might explain the seeming lack of progress towards developing a private algorithm in IRT.

**Our Contributions.** In this work, we make progress in both of the above open directions.

- In section 3, we obtain refined entrywise analysis of the spectral algorithm, coupled with a matching lower bound, showing that under mild sampling conditions, the spectral algorithm achieves optimal entrywise error guarantee (up to a log factor). As a result of this optimality, we also show that the spectral algorithm can accurately identify all of the top  $K$  items using only a constant factor more than the information theoretically minimal sample size.

<sup>1</sup>To adapt the Rasch model to this application, define  $X_{li} = 0$  if reviewer  $l$  accepts paper  $i$ .  $\beta_i^*$  corresponds to the quality of paper  $i$  and  $\theta_l^*$  the ‘harshness’ of reviewer  $l$ .

- In Section 4, we propose a privacy-preserving mechanism that is designed with the discrete nature of human response data in mind. We take advantage of the Markov chain formulation underlying the spectral algorithm and the discrete Gaussian mechanism to develop a conceptually simple yet performant private algorithm. The algorithm provides strong and controllable privacy guarantee and outperforms other approaches, especially in the low-to-moderate privacy regime.

## 2 PROBLEM DESCRIPTION & THE SPECTRAL ALGORITHM

We follow the notations of Nguyen and Zhang (2022). Consider a universe with  $n$  people and  $m$  items. Each person  $l$  has parameter  $\theta_l^* \in \mathbb{R}$  and each item  $i$  has parameter  $\beta_i^* \in \mathbb{R}$ . To overcome a fundamental identifiability issue associated with parameter translation, we assume a normalization constraint on the item parameters  $\beta^{*\top} \mathbf{1}_m = 0$ . We also assume that  $\beta_i^* \in [\beta_{\min}^*, \beta_{\max}^*] \forall i \in [m]$  for some constants  $\beta_{\min}^*, \beta_{\max}^*$ . Similarly, we assume that  $\theta_l^* \in [\theta_{\min}^*, \theta_{\max}^*]$  for some constants  $\theta_{\min}^*, \theta_{\max}^*$ . Let  $A \in \{0, 1\}^{n \times m}$  denote the assignment matrix where  $A_{li} = 1$  if person  $l$  responds to item  $i$  and 0 indicates missing data. The observed data is  $X \in \{0, 1, *\}^{n \times m}$  where  $*$  denotes missing data. For entries where  $A_{li} = 1$ ,  $X_{li}$  is independently distributed per Equation (1). Let us consider the uniform sampling model where  $A_{li} = 1$  with probability  $p$  for some  $p > 0$ .

**The Spectral Algorithm.** We summarize the original formulation of the spectral algorithm here. For more details on its implementation and analysis, we refer the reader to Nguyen and Zhang (2022). For each item pair  $i \neq j$ , define the *pairwise differential measurement* as

$$Y_{ij} = \sum_{l=1}^n A_{li} A_{lj} X_{li} (1 - X_{lj}) \quad \forall i \neq j \in [m]. \quad (2)$$

Note that this is a discrete-valued quantity. Given the pairwise differential measurements, the algorithm constructs a Markov chain  $\hat{M} \in [0, 1]^{m \times m}$  whose transition probabilities are defined as follows:

$$\hat{M}_{ij} = \begin{cases} \frac{1}{d} Y_{ij} & \text{if } i \neq j \\ 1 - \sum_{k \neq i} \frac{1}{d} Y_{ik} & \text{if } i = j \end{cases}, \quad (3)$$

where  $d$  is a sufficiently large normalization factor chosen so that the resulting pairwise transition probability matrix  $\hat{M}$  does not contain any negative entries. Typically,  $d = O(\max_{i \in [m]} \sum_{k \neq i} B_{ik})$  where  $B_{ik} = \sum_{l=1}^n A_{li} A_{lk}$ . The algorithm then computes

the stationary distribution  $\pi$  of  $\hat{M}$  and recovers  $\beta$  using a post-processing step. Algorithm 1 summarizes the spectral algorithm.

---

**Algorithm 1** The Spectral Algorithm
 

---

**Input:** Binary response data  $X \in \{0, 1, *\}^{n \times m}$ .

**Output:** Item parameter estimate  $(\beta_i)_{i=1}^m$ .

- 1: Construct a Markov chain  $\hat{M}$  per Equation (3).
- 2: Compute the stationary distribution of  $\hat{M}$ :  
 Initialize  $\pi^{(0)} = [\frac{1}{m}, \dots, \frac{1}{m}]$ .  
 For  $t = 1, 2, \dots$  until convergence, compute

$$\pi^{(t)\top} = \frac{\pi^{(t-1)\top} \hat{M}}{\|\pi^{(t-1)\top} \hat{M}\|_1}.$$

- 3: Compute  $z = \pi/d$  and  $\bar{\beta} = \log(z)$  and return the normalized item parameters, i.e.,  $\beta = \bar{\beta} - \bar{\beta}^\top \mathbf{1}/m$ .
- 

### 3 FINE-GRAIN ANALYSIS OF THE SPECTRAL ALGORITHM

#### 3.1 Entrywise Error Bound

In this section we provide refined entrywise error bound for the spectral algorithm. Recall the example with conference reviewing and that our practical motivation is to show that the estimation error incurred by the spectral algorithm ‘spreads out’ among the parameters. This is to prevent the case where the estimation error is concentrated on a few items.

**Theorem 3.1.** *Consider the uniform sampling model described in Section 2. There exist constants  $C_1, C_2$  such that if  $np^2 \geq C_1 \log m$  and  $mp \geq C_2 \log n$  then the output of the spectral algorithm satisfies*

$$\|\beta - \beta^*\|_\infty \leq \frac{C\sqrt{\log m}}{\sqrt{np}}$$

with probability at least  $1 - \mathcal{O}(m^{-10}) - \mathcal{O}(n^{-10})$  where  $C$  is an absolute constant.

The above theorem is a refinement of Theorem 3.3 in Nguyen and Zhang (2022). In that paper, the authors show that the output of the spectral algorithm satisfies  $\|\beta - \beta^*\|_2 = \mathcal{O}\left(\frac{\sqrt{m}}{\sqrt{np}}\right)$ . Comparing this  $\ell_2$  bound with Theorem 3.1 reveals that the magnitude of the entrywise error is no more than  $\mathcal{O}\left(\sqrt{\frac{\log m}{m}}\right)$  the size of the  $\ell_2$  error. This shows that the cumulative error indeed distributes evenly among the parameters and no two error terms differ in magnitude by more than a log factor.

Readers who are familiar with the learning to rank literature might see the parallel between this result and

that in the Bradley-Terry-Luce (BTL) model where a similar algorithm – spectral ranking – enjoys a favorable entrywise error guarantee. However, the results are only superficially familiar and are based on similar fundamental theorems of Markov chain perturbation. Our results for the entrywise error bound and top- $K$  recovery in the next section are not simple extensions of the known results in the BTL literature. Firstly, the sampling model considered here fundamentally differs from the Erdos-Reyni comparison graph sampling model considered in the BTL model analysis. Additionally, the construction of the spectral algorithm is also different from the spectral ranking algorithm. Specifically, the pairwise transition probabilities and the construction of the Markov chain in Algorithm 1 and the spectral ranking algorithm are different.

Besides the spectral algorithm, we are only aware of entrywise error guarantee for joint maximum likelihood estimate (JMLE) (Andersen, 1973; Haberman, 1977), formally stated in Theorem 2 of Chen et al. (2021). However, it is well known that JMLE produces an inconsistent estimate of  $\beta^*$  when  $m$  is constant (Andersen, 1973; Ghosh, 1995). In the supplementary materials, we provide a complementary entrywise error bound for the spectral algorithm where we remove the assumption  $mp = \Omega(\log n)$  from Theorem 3.1, showing that the spectral method always gives a consistent estimation of individual item parameters. On a practical note, the spectral algorithm has been shown to be similar in terms of accuracy but is significantly faster than JMLE (Nguyen and Zhang, 2022).

#### 3.2 Top- $K$ Recovery Guarantee

In this section, we analyze the performance of the spectral algorithm in terms of top- $K$  recovery, building on the entrywise error bound obtained earlier. Curiously, Nguyen and Zhang (2022) show through experimental results that the spectral algorithm often outperforms the baseline IRT algorithms in terms of top- $K$  accuracy. In this section, we provide a theoretical justification for its strong performance by furnishing an upper bound on the sample complexity for top- $K$  recovery of the spectral algorithm. Under mild sampling conditions, this upper bound turns out to have a matching lower bound which we will show in Section 3.3.

To ground our discussion, we first define a theoretical quantity that captures the model-dependent difficulty of the top- $K$  recovery problem. Define  $\Delta_K = \beta_{[K]}^* - \beta_{[K+1]}^*$  to be the gap between the  $K$ -th and the  $K+1$ -th best items. The smaller the gap, the harder it is to separate the top  $K$  items from the remaining items.

The reader could intuitively see that if we have a finite-sample entrywise error bound on  $\beta$ , we can obtain

a finite sample error guarantee for top- $K$  recovery. Suppose that we have a sufficiently large sample size such that  $\|\beta - \beta^*\|_\infty \leq \frac{\Delta_K}{2}$ . Then we separate all of the top  $K$  items from the bottom  $n - K$  items and achieve perfect top  $K$  recovery. Building on this intuition, we have the following top- $K$  recovery guarantee for the spectral algorithm.

**Theorem 3.2.** *Consider the setting of Theorem 3.1. Consider a top- $K$  estimator that first runs the spectral algorithm on the response data and returns the  $K$  items with the highest parameter values. There exists a constant  $C_K$  such that if  $np \geq \frac{C_K \log m}{\Delta_K^2}$ , then this top- $K$  estimator correctly identifies all of the top  $K$  items with probability at least  $1 - \mathcal{O}(m^{-10}) - \mathcal{O}(n^{-10})$ .*

To summarize, the above theorem states that so long as  $m, n, p$  satisfy

$$np = \Omega\left(\frac{\log m}{\Delta_K^2}\right), np^2 = \Omega(\log m) \text{ and } mp = \Omega(\log n),$$

then the spectral method accurately identifies all of the top  $K$  items. To the best of our knowledge, this guarantee (and the lower bound shown in the next section) is the first formal result for the ranking performance of any estimation algorithm in the literature.

### 3.3 Lower Bounds

In this section, we complement the upper bounds obtained in previous sections with corresponding information theoretic lower bounds.

Firstly, **for parameter estimation**, we obtain a lower bound result that generalizes Theorem 3.5 of Nguyen and Zhang (2022). Note that Theorem 3.5 in the previous work is a Cramer-Rao bound that only applies to *unbiased estimators*. However, our *new* information theoretic lower bound is strictly more general and is applicable to *all statistical estimators*.

**Theorem 3.3.** *Consider the sampling model described in Section 2 and further assume that  $np = \Omega(1)$ . There exists a class of Rasch models  $\mathcal{B}$  such that for any statistical estimator, the minimax risk is lower bounded as*

$$\inf_{\hat{\beta}} \sup_{\beta^* \in \mathcal{B}} \mathbb{E} \|\hat{\beta} - \beta^*\|_2^2 = \Omega\left(\frac{m}{np}\right).$$

As a consequence of the normed inequality  $\|\beta - \beta^*\|_\infty \geq \frac{1}{\sqrt{m}} \|\beta - \beta^*\|_2$ , we can lower bound the minimax entry-wise error for any statistical estimator as

$$\inf_{\hat{\beta}} \sup_{\beta^* \in \mathcal{B}} \mathbb{E} \|\hat{\beta} - \beta^*\|_\infty^2 = \Omega\left(\frac{1}{np}\right).$$

From Theorem 3.1, the spectral algorithm satisfies  $\|\beta - \beta^*\|_\infty^2 = \mathcal{O}\left(\frac{\log m}{np}\right)$ . This directly establishes the

optimality (modulo a log factor) of the spectral algorithm.

Moving on to **top- $K$  recovery guarantee**, the following theorem establishes the minimum sample complexity required to accurately identify all of the top  $K$  items from user-item binary response data.

**Theorem 3.4.** *Consider the sampling model described in Section 2. There is a class of Rasch model such that if  $np \leq \frac{c_K \log m}{\Delta_K^2}$ , where  $c_K$  is a constant, then any estimator will fail to identify all of the top  $K$  items with probability at least  $\frac{1}{2}$ .*

We compare the above lower bound to the upper bound obtained in the previous section. The condition  $mp = \Omega(\log n)$  is a mild one, requiring that the number of items grows slowly and is appropriate in applications such as recommendation systems or crowdsourcing. We see that  $np = \Omega\left(\frac{\log m}{\Delta_K^2}\right)$  is the necessary sample size to accurately identify all of the top- $K$  items and the spectral algorithm needs the same condition. The spectral algorithm also requires  $np^2 = \mathcal{O}(\log m)$ . This is a consequence of the design of the algorithm that operates on pairwise differential measurements.

In general,  $\Delta_K$  may be arbitrarily small and change with  $m$ . For example, if each  $\beta_i^*$  is uniformly sampled from  $[\beta_{\min}^*, \beta_{\max}^*]$  then  $\mathbb{E}[\Delta_K] = \mathcal{O}\left(\frac{1}{m}\right)$ . This makes comparisons between the two conditions on  $n$  difficult. In some real life situations, however, one can assume that  $p = \mathcal{O}(1)$ . This is appropriate in settings such as education testing where each student is shown a constant fraction of all questions from a question bank, or in recommendation systems where the users are shown a constant fraction of a product catalogue. In this regime of  $p$ , the lower bound simplifies to  $n = \Omega\left(\frac{\log m}{\Delta_K^2}\right)$ . Assuming that  $\Delta_K = \mathcal{O}(1)$ , the spectral algorithm needs  $n = \mathcal{O}\left(\max\left\{\frac{\log m}{\Delta_K^2}, \log m\right\}\right) = \mathcal{O}\left(\frac{\log m}{\Delta_K^2}\right)$  to accurately identify all of the top  $K$  items – the optimal sample complexity for top- $K$  recovery.

## 4 THE PRIVATE SPECTRAL ALGORITHM

### 4.1 A Brief Introduction to Differential Privacy

In this section, we propose, for the first time in the IRT literature, a privacy-preserving estimation algorithm. We focus on differential privacy (DP), which has become the de facto framework for analyzing the privacy guarantees of machine learning algorithms. To simplify the description, we consider the *full response setting* – a person responds to all items. The starting point is the definition of two neighboring datasets.

**Definition 1.** Let  $X, X'$  be two binary responses data as described in Section 2. We write  $X \simeq X'$  to denote that  $X$  and  $X'$  are neighboring datasets if  $X$  and  $X'$  differ by the entries of exactly one row.

Note that we are trying to provide *user privacy* in our application. Intuitively, a randomized function or a statistical query of a dataset (e.g., counting the number of people who respond positively to an item) provides privacy to the users if its output is not sensitive to changes in a single user's data. This is formalized by the definition (approximate) differential privacy.

**Definition 2** (Approximate Differential Privacy (Dwork et al., 2014)). For any input domain  $\mathcal{X}$  and output domain  $\mathcal{Y}$ , a randomized function  $\tilde{f} : \mathcal{X} \rightarrow \mathcal{Y}$  satisfies  $(\epsilon, \delta)$ -differential privacy if for any two  $X \simeq X'$  and any  $C \subseteq \mathcal{Y}$ ,  $\mathbb{P}(\tilde{f}(X') \in C) \leq e^\epsilon \mathbb{P}(\tilde{f}(X) \in C) + \delta$ .

The quantity  $\epsilon$  is the *privacy budget* or desired privacy level. The smaller its value, the stronger the privacy guarantee but the private output tends to be less accurate. When an algorithm makes multiple private functions evaluation or queries on the same dataset, we call it *query composition*. In general, privacy guarantee *decays linearly* with the number of queries (Dwork et al., 2014; Dwork and Rothblum, 2016). That is, for a composition of  $k$  queries to satisfy overall  $\epsilon^*$  privacy level, the privacy budget for each query is approximately  $\epsilon^*/k$ . When  $k$  is large, the accuracy of DP algorithms tend to suffer significantly. Motivated by this practical consideration, researchers have developed refined notions of privacy that enjoys better accuracy under query composition than DP. One of these is concentrated differential privacy (CDP).

**Definition 3** (Concentrated Differential Privacy (Dwork and Rothblum, 2016; Bun and Steinke, 2016)). For any input domain  $\mathcal{X}$  and output domain  $\mathcal{Y}$ , a randomized function  $\tilde{f} : \mathcal{X} \rightarrow \mathcal{Y}$  satisfies  $\frac{\epsilon^2}{2}$ -concentrated differential privacy if for all  $X \simeq X'$  and  $\alpha \in (1, \infty)$ ,  $D_\alpha(\tilde{f}(X) \parallel \tilde{f}(X')) \leq \frac{\epsilon^2 \alpha}{2}$ , where  $D_\alpha(P \parallel Q) = \frac{1}{\alpha-1} \log \sum_{y \in \mathcal{Y}} P(y)^\alpha Q(y)^{1-\alpha}$  is the Rényi divergence between the distributions  $P$  and  $Q$ .

## 4.2 The Discrete Gaussian Mechanism

A common approach to preserve privacy is to add random noise to the output of non-private functions. Generally, adding a larger the amount of noise gives better the privacy guarantee at the expense of downstream estimation accuracy. Within this class of mechanisms, the discrete Gaussian mechanism (Canonne et al., 2020) is a recently proposed noise adding mechanism utilizing a discretized generalization of the Gaussian distribution. We deliberately choose a discrete noise mechanism because it has been found that even innocuous

finite-precision calculations can be exploited by malicious users to recover the original discrete data exactly (Mironov, 2012). Given the discrete nature of the binary response data, we believe that this approach is pertinent.

Secondly, we prefer the discrete Gaussian mechanism to the discrete Laplace mechanism which adds discrete Laplace noise (Ghosh et al., 2009). The discrete Gaussian mechanism has over the Laplace mechanism is that the Gaussian distribution has lighter tail than the Laplace distribution. As our experiments will show, when we have to repeatedly make a large number of queries, the total variance of the Gaussian noise is lower than that of the Laplace noise needed to achieve a comparable privacy level and the resulting estimate is more accurate.

For the design of our private spectral algorithm, recall the definition of a pairwise differential measurement in Equation (2). For a fixed privacy budget  $\epsilon$ , define the *privatized pairwise differential measurement* as

$$\tilde{Y}_{ij} = \sum_{l=1}^n A_{li} A_{lj} X_{li} (1 - X_{lj}) + Z_{ij} \quad \forall i \neq j \in [m], \quad (4)$$

where each  $Z_{ij} \sim \mathcal{N}_{\mathbb{Z}}(0, \frac{1}{\epsilon^2})$  is an independent random variable drawn from the 0-mean discrete Gaussian distribution with variance  $\frac{1}{\epsilon^2}$  (see Algorithm 1 of Canonne et al. (2020)).

There are  $m(m-1)$  pairwise differential measurements. Hence, a privatized spectral algorithm that adds  $\mathcal{N}_{\mathbb{Z}}(0, \frac{1}{\epsilon^2})$  to each pairwise differential measurement  $Y_{ij}$  achieves  $\frac{1}{2}m(m-1)\epsilon^2$ -CDP overall (Canonne et al., 2020). Conversely, if we desire an overall privacy level  $\epsilon^*$ , we can compute the noise variance needed for each pairwise differential measurement.

**Lemma 4.1.** Consider a modified spectral algorithm that first adds  $\mathcal{N}_{\mathbb{Z}}\left(0, \frac{\sqrt{m(m-1)}}{\epsilon^{*2}}\right)$ -distributed noise to each pairwise differential measurement per Equation (4). The algorithm then proceeds to use the private differential measurements to construct the Markov chain as in the original algorithm. Then the modified spectral algorithm satisfies  $\frac{1}{2}\epsilon^{*2}$ -CDP.

## 4.3 Changing the Privacy-Accuracy Tradeoff via Sparsification

From Lemma 4.1, one can see that the variance of the additive Gaussian noise scales as the *square root of the number of pairwise differential measurements*. When  $n \gg m$ , the additive noise will be dominated by the signal from the user data and the private spectral algorithm will produce an accurate estimate (e.g., see Figure 1). On the other hand, suppose we have a large

$m$  (say  $m \geq 100$ ) but a relatively small  $n$  ( $n \leq 500$ ). Adding  $\mathcal{N}_{\mathbb{Z}}\left(0, \frac{\sqrt{m(m-1)}}{\epsilon^2}\right)$ -distributed noise to each pairwise differential measurement may introduce a lot of noise, dominating the signal and leading to a less accurate final estimate.

To overcome this limitation, we propose a heuristic modification to the private spectral algorithm. Instead of constructing a “dense” Markov chain where every pairwise transition probabilities are non-zero, we construct a *sparse approximation* of this Markov chain. We first generate an Erdős-Rényi graph with  $m$  vertices and edge probability  $p_0 = (1 + o(1))\frac{\log m}{m}$ . Note that the condition on the edge probability is to ensure that the resulting comparison graph is connected with high probability (Erdos et al., 1960). If we further ensure that all pairwise transition probabilities are positive<sup>2</sup>, the resulting Markov chain will admit a unique stationary distribution and subsequently a unique parameter estimate. The advantage here is that we only compute  $(1 + o(1))m \log m$  pairwise differential measurements. Hence for each subsampled pairwise differential measurement, we only need to add  $\mathcal{N}_{\mathbb{Z}}\left(0, \frac{\sqrt{\mathcal{O}(m \log m)}}{\epsilon^{*2}}\right)$ -distributed noise and achieve the same CDP guarantee, a  $\sqrt{m}$  factor reduction in variance compared to the dense Markov chain construction.

However, sparsification might also come with its own limitation. By removing some of the pairwise transition probabilities information from the Markov chain, the final estimate obtained from the sparse Markov chain may deviate from that using full information. As our experiments will show, sparsification tradeoffs between additive noise and approximation error. It is appropriate to use sparsification to achieve a low-to-moderate privacy guarantee with good estimation accuracy when the number of items is large and the number of user is small. Finally, Algorithm 2 summarizes our private spectral algorithm.

---

**Algorithm 2** The Private Spectral Estimator
 

---

**Input:** Response data  $X \in \{0, 1\}^{n \times m}$ ,  $p_0$  (default 1) and desired privacy level  $\epsilon^*$ .

**Output:** Private item estimate  $\tilde{\beta}$

- 1: Construct a private Markov chain  $\tilde{M}$  with underlying comparison graph drawn from an Erdos-Reyni distribution with edge probability  $p_0$  per Equation (4) and desired privacy level  $\epsilon^*$ .
  - 2: Follow steps 2-3 in Algorithm 1 using  $\tilde{M}$ .
- 

<sup>2</sup>This can be achieved by adding regularization as described in Nguyen and Zhang (2022)

## 5 EXPERIMENTS

In this section, we evaluate the privacy-accuracy trade-off of the private spectral algorithm on synthetic and real-life datasets. We use randomized response (on the user data) and the discrete Laplace mechanism as baseline methods. Recently it has been shown that one can amplify the privacy guarantee of randomized response by shuffling the randomized user data (cf., Theorem III.1 of Feldman et al. (2022)). We compare the algorithms in terms of their approximate differential privacy guarantee<sup>3</sup>. Specifically, we fix a common  $\delta = 10^{-4}$  and evaluate the accuracy at varying privacy levels  $\epsilon^*$ . For readers who are not familiar with privacy-preserving mechanisms, we describe these approaches in the supplementary materials and their implementation.

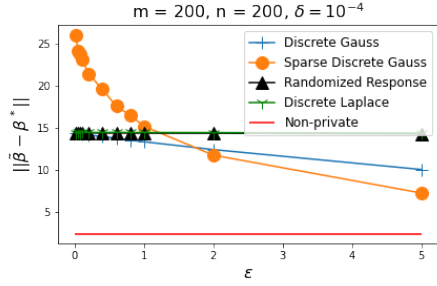
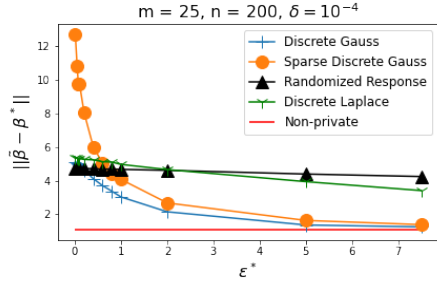
### 5.1 Synthetic Datasets.

Figure 1 shows the performance of the algorithms on synthetic datasets where we know the ground truth  $\beta^*$ . The amount of noise introduced by the discrete Gaussian mechanism gets proportionally smaller than the signal as  $n$  grows. Expectedly, given a sufficiently large  $n$  (relative to  $m$ ), we see almost identical performance between the non-private estimator and the private estimator using the discrete Gaussian mechanism. At the very high privacy regime  $\epsilon^* \approx 0.01$ , all methods produce very inaccurate results. Sparsification is significantly worse than the other approaches in this regime because of the combined effect of a large amount of additive noise and the approximation error introduced by sparsification. The differences among the methods become more apparent in the low-to-moderate privacy regime. As  $\epsilon^*$  increases, the accuracy of the discrete Gaussian mechanism improves significantly. On the other hand, randomized response and the Laplace mechanism would require either a very large  $n$  or a very large privacy budget  $\epsilon^*$  to produce a reasonably accurate estimate, neither of which is desirable.

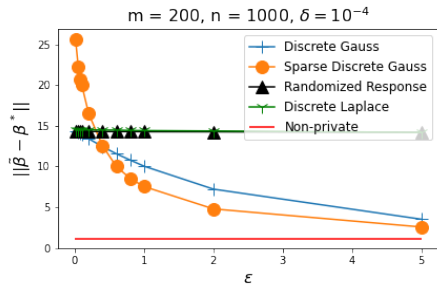
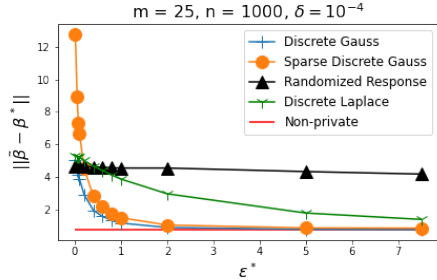
### 5.2 Real-life Datasets.

For real life datasets, we do not have a ground truth  $\beta^*$ . However, previous work has shown that the spectral algorithm is competitive with the most commonly used IRT algorithms. Also considering that we focus on privacy-preservation in this work, we evaluate  $\|\tilde{\beta} - \beta\|$ , where  $\beta$  is the output of the non-private spectral algorithm and  $\tilde{\beta}$  that of the corresponding private algorithm.

<sup>3</sup>One can convert concentrated differential privacy guarantee of the spectral algorithm to approximate differential privacy guarantee using Lemma 3.5 of Bun and Steinke (2016).



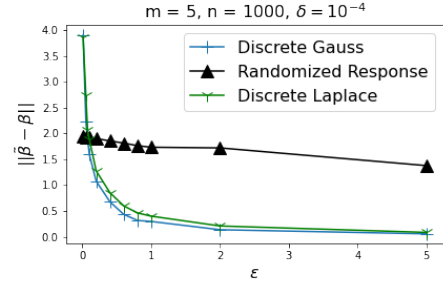
(a) Given the same number of users, increasing  $m$  expectedly leads to a worse privacy-accuracy tradeoff for all methods. The discrete Gaussian mechanism generally outperforms the baselines at the low-to-moderate privacy regime.



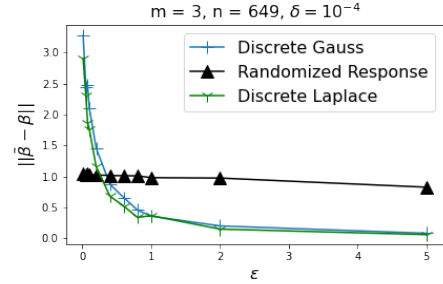
(b) The privacy-accuracy tradeoff for the discrete Gaussian mechanism gets better as  $n$  increases (from 200 in the previous subfigure to 1000) while randomized response and the Laplace mechanism see little gain from a larger  $n$ .

**Figure 1: Synthetic datasets.** The discrete Gaussian mechanism outperforms the baselines at low-to-moderate privacy regime and enjoys substantial improvement from a larger sample size  $n$ .

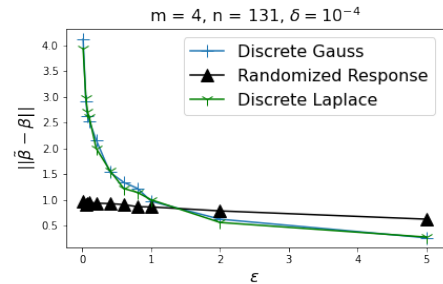
**Education Datasets.** We include commonly used education testing datasets including LSAT (McDonald, 2013), UCI (Hussain et al., 2018) and the Three Grades dataset (Cortez and Silva, 2008). These datasets have a very small number of items. At very high privacy level  $\epsilon \approx 0.01$ , all methods produce very inaccurate estimate and randomized response is the most accurate method. However, at a moderate privacy level  $\epsilon^* \approx 1$ , the Laplace mechanism and the Gaussian mechanism outperform the randomized response approach. On the other hand, to produce a reasonably accurate estimate, randomized response would require a very large  $\epsilon^*$ , leading to a vacuous privacy guarantee.



(a) LSAT dataset.



(b) Three grades dataset.



(c) UCI student dataset.

**Figure 2: Education datasets.** The Gaussian mechanism and the Laplace mechanism outperforms randomized response at low-to-moderate privacy levels. The results here suggest that for many education testing application where the number of items is small, the spectral algorithm with the discrete Gaussian (or Laplace) mechanism should be the preferred privacy-preserving algorithm.

**ML-100K Recommendation Systems Datasets.** For experiments on recommendation systems datasets

(Harper and Konstan, 2015), we follow a similar approach to previous work to generate binary response data from ratings data (Lan et al., 2018; Davenport et al., 2014). We first perform a low rank matrix estimation of the user-item ratings matrix and use the estimation as the user-item preference scores. For each user, we convert scores that are lower than the average to 1 and scores that are higher than the average to 0. We then select a random subset of  $m$  items and extract the corresponding submatrix as the response matrix. From Figure 3, one can see that the discrete Gaussian mechanism provides superior privacy-accuracy tradeoff compared to the other approaches. Notably, sparsification provides a more accurate parameter estimation by reducing the amount of noise needed to ensure the same level of privacy and this improvement is significant when  $m$  is large.

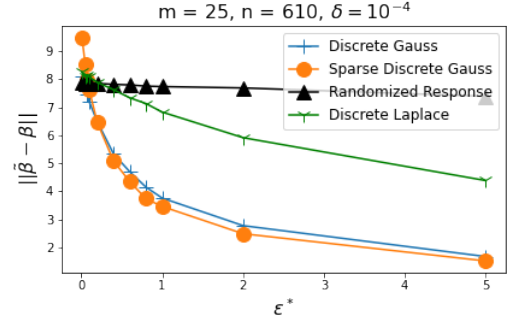
From the experiments on synthetic and real-life datasets, one can see that when we desire high accuracy and a moderate level of privacy, the discrete Gaussian mechanism should be the preferred approach. Additionally, when the number of items  $m$  is large while the number of users  $n$  is small, one should use the sparse Markov chain construction.

## 6 ETHICAL DISCUSSIONS

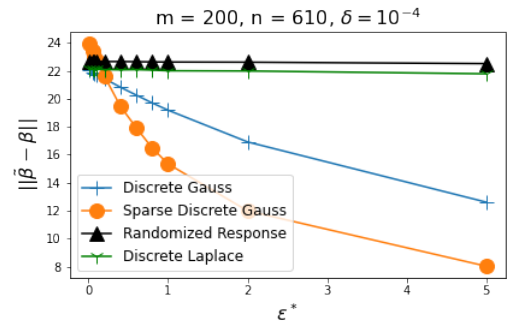
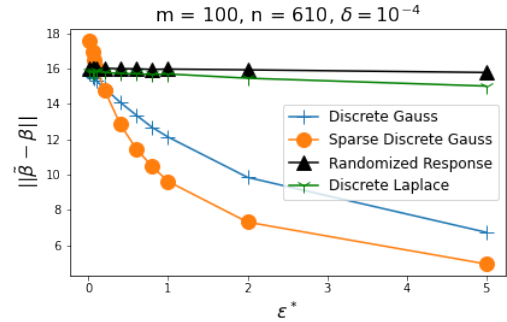
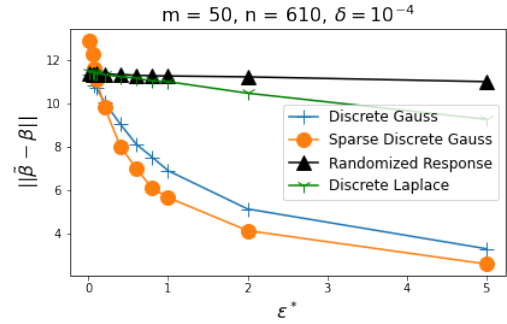
Our work explores privacy matters for an algorithm of which real-life applications often involve human data with sensitive information. For example, the Rasch model is often studied in the context of education testing and psychological testing where the subjects are students and patients. Therefore, deploying any algorithms in such capacity needs to be accompanied by thoughtful and thorough ethical considerations. In this work, we provide the algorithmic tool to quantify the error associated with the parameter estimates and to preserve the privacy of the human subjects.

## 7 CONCLUSION

In this work, we made two contributions to the understanding of the spectral algorithm for parameter estimation under the Rasch model. Firstly, we show that the spectral algorithm achieves the optimal entry-wise error bound, complementing its optimal  $\ell_2$  error guarantee obtained in the previous work. Secondly, we propose, for the first time in the item response theory literature, a differentially private algorithm. We complement our theoretical and algorithmic contributions with experimental results showing that the discrete Gaussian mechanism provides accurate parameter estimation, outperforming the baselines. A possible future extension of this work is a deeper analysis of the ran-



(a) When  $m$  is small, sparsification doesn't significantly improve the performance of the private spectral algorithm.



(b) With the same  $n$  and a large number of items  $m$ , the 'sparse Markov chain' approach is significantly better than the 'dense Markov chain' approach.

**Figure 3: ML-100K datasets.** The discrete Gaussian mechanism generally outperforms the baselines, especially in the low-to-moderate privacy regime  $\epsilon^* \geq 1$ .



dom subsampling of the comparison graph in the sparse discrete Gaussian mechanism. We believe that the randomized sparsification step could provide some form of privacy amplification. Furthermore, designing an accurate estimation algorithm for the high privacy regime ( $\epsilon^* \approx 0.01$ ) remains an unresolved challenge.

## 8 ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their thoughtful suggestions and comments. We thank Aaron Roth and Emily Diana of the University of Pennsylvania for fruitful discussions on the topic of differential privacy. The authors are supported by NSF Grant DMS-2112099. A.Z. acknowledges financial support from the Alfred H. Williams Faculty Scholar award. Any opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

**Addressing the Reviewers’ Feedback.** In response to the anonymous reviewers’ feedback during the submission process, we have performed additional theoretical and experimental analyses and included the additional findings in the supplementary materials. Specifically, we have studied how the performance of the sparse spectral algorithm depends on the edge sampling probability  $p_0$ . We have also provided a high-level analysis of the entrywise error of the *private* spectral estimate, complementing Theorem 3.1.

## References

- Andersen, E. B. (1973). Conditional inference for multiple-choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 26(1):31–44.
- Brzezińska, J. et al. (2016). Latent variable modelling and item response theory analyses in marketing research. *Folia Oeconomica Stetinensia*, 16(2):163–174.
- Bun, M. and Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658. Springer.
- Canonne, C. L., Kamath, G., and Steinke, T. (2020). The discrete gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33:15676–15688.
- Chen, C.-M., Lee, H.-M., and Chen, Y.-H. (2005). Personalized e-learning system using item response theory. *Computers & Education*, 44(3):237–255.
- Chen, Y., Li, C., and Xu, G. (2021). A note on statistical inference for noisy incomplete 1-bit matrix. *arXiv preprint arXiv:2105.01769*.
- Chen, Z. and Ahn, H. (2020). Item response theory based ensemble in machine learning. *International Journal of Automation and Computing*, 17(5):621–636.
- Cortez, P. and Silva, A. M. G. (2008). Using data mining to predict secondary school student performance.
- Davenport, M. A., Plan, Y., Van Den Berg, E., and Wootters, M. (2014). 1-bit matrix completion. *Information and Inference: A Journal of the IMA*, 3(3):189–223.
- Dwork, C., Roth, A., et al. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.
- Dwork, C. and Rothblum, G. N. (2016). Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*.
- Erdos, P., Renyi, A., et al. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5(1):17–60.
- Feldman, V., McMillan, A., and Talwar, K. (2022). Hiding among the clones: A simple and nearly optimal analysis of privacy amplification by shuffling. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 954–964. IEEE.
- Fischer, G. H. (1981). On the existence and uniqueness of maximum-likelihood estimates in the rasch model. *Psychometrika*, 46(1):59–77.
- Ghosh, A., Roughgarden, T., and Sundararajan, M. (2009). Universally utility-maximizing privacy mechanisms. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 351–360.
- Ghosh, M. (1995). Inconsistent maximum likelihood estimators for the rasch model. *Statistics & Probability Letters*, 23(2):165–170.
- Haberman, S. J. (1977). Maximum likelihood estimates in exponential response models. *The annals of statistics*, 5(5):815–841.
- Hambleton, R. K., Swaminathan, H., and Rogers, H. J. (1991). *Fundamentals of item response theory*, volume 2. Sage.
- Harper, F. M. and Konstan, J. A. (2015). The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)*, 5(4):1–19.
- Hussain, S., Dahan, N. A., Ba-Alwib, F. M., and Ribata, N. (2018). Educational data mining and analysis of students’ academic performance using weka. *Indonesian Journal of Electrical Engineering and Computer Science*, 9(2):447–459.
- Lan, A., Chiang, M., and Studer, C. (2018). An estimation and analysis framework for the rasch model. In *International Conference on Machine Learning*, pages 2883–2891. PMLR.

- Lord, F. M. (2012). *Applications of item response theory to practical testing problems*. Routledge.
- Martínez-Plumed, F., Prudêncio, R. B., Martínez-Usó, A., and Hernández-Orallo, J. (2019). Item response theory in ai: Analysing machine learning classifiers at the instance level. *Artificial intelligence*, 271:18–42.
- McDonald, R. P. (2013). *Test theory: A unified treatment*. psychology press.
- Mironov, I. (2012). On significance of the least significant bits for differential privacy. In *Proceedings of the 2012 ACM conference on Computer and communications security*, pages 650–661.
- Moraes, J. V., Reinaldo, J. T., Prudencio, R. B., and Silva Filho, T. M. (2020). Item response theory for evaluating regression algorithms. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.
- Nguyen, D. and Zhang, A. (2022). A spectral approach to item response theory. *Advances in neural information processing systems*, 35.
- Rasch, G. (1960). Studies in mathematical psychology: I. probabilistic models for some intelligence and attainment tests.
- Schellhorn, C. and Sharma, R. (2013). Using the rasch model to rank firms by managerial ability. *Managerial Finance*.
- Van der Linden, W. J. and Hambleton, R. (1997). Handbook of item response theory. *Taylor & Francis Group*. Citado na pág, 1(7):8.
- Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69.
- Whitehill, J., Wu, T.-f., Bergsma, J., Movellan, J., and Ruvolo, P. (2009). Whose vote should count more: Optimal integration of labels from labelers of unknown expertise. *Advances in neural information processing systems*, 22.

# Supplementary Materials: Optimal and Private Learning from Human Response Data

Duc Nguyen & Anderson Zhang

## Contents

<b>1</b>	<b>Proofs of Entrywise Error Guarantee</b>	<b>1</b>
1.1	Preliminaries . . . . .	1
1.2	Basic Entrywise Analysis . . . . .	3
1.3	Improved Rates for the Regime $mp \gtrsim \log n$ . . . . .	9
<b>2</b>	<b>Proofs of Lower Bounds</b>	<b>14</b>
2.1	Estimation Error Lower Bound . . . . .	14
2.2	Top- $K$ Recovery Lower Bound . . . . .	16
<b>3</b>	<b>Differentially Private Spectral Algorithm</b>	<b>18</b>
<b>4</b>	<b>Additional Experimental Details</b>	<b>20</b>

## 1 Proofs of Entrywise Error Guarantee

### 1.1 Preliminaries

Define  $\pi_i^* = \frac{e^{\beta_i^*}}{\sum_{k=1}^m e^{\beta_k^*}}$  for  $i \in [m]$ ,  $\pi_{\max}^* = \max_{i \in [m]} \{\pi_i^*\}$ ,  $\pi_{\min}^* = \min_{i \in [m]} \{\pi_i^*\}$ . Define  $\kappa := \beta_{\max}^* - \beta_{\min}^*$ . We have  $\pi_{\max}^* / \pi_{\min}^* = e^\kappa$ . Let  $\gamma = \min_{l \in [n], i, j \in [m]} \mathbb{E}[X_{li}(1 - X_{lj})]$ . Define  $B = A^\top A$  (i.e.,  $B_{ij} = \sum_{l=1}^n A_{li}A_{lj}$ ) and the following two events.

$$\mathcal{A} = \left\{ \frac{np^2}{2} \leq B_{ij} \leq \frac{3np^2}{2} \forall i \neq j \in [m] \right\},$$

$$\mathcal{A}^+ = \mathcal{A} \cap \left\{ \frac{mp}{2} \leq A_l^\top \mathbf{1} \leq \frac{3mp}{2} \forall l \in [n] \right\}.$$

Both events happen with high probability under mild sampling conditions, as summarized by two lemmas below. As a convenient notation, we use  $a \gtrsim b$  to mean  $a = \Omega(b)$ .

**Lemma 1.1.** *Consider the random sampling scheme described in Section 2. Suppose that  $np^2 \geq C_0 \log m$ , for a sufficiently large constant  $C_0$  (e.g.,  $C_0 \geq 20$ ). We have*

$$\mathbb{P}(\mathcal{A}) \geq 1 - \exp\left(-(1 - o(1))\frac{np^2}{10}\right)$$

*Proof.* Invoking Chernoff bound, we have for a pair  $i \neq j$ ,

$$\mathbb{P}\left(\left|\sum_{l=1}^n A_{li}A_{lj} - \mathbb{E}[A_{li}A_{lj}]\right| > \frac{1}{2}np^2\right) \leq 2\exp\left(-\frac{\frac{1}{4}np^2}{\frac{1}{2} + 2}\right) = \exp\left(-\frac{np^2}{10} + \ln 2\right).$$

Applying union bound over all pairs, we have

$$\begin{aligned}
& \mathbb{P} \left( \exists i \neq j : \left| \sum_{l=1}^n A_{li} A_{lj} - \mathbb{E}[A_{li} A_{lj}] \right| > \frac{1}{2} np^2 \right) \\
& \leq \binom{m}{2} \cdot \exp \left( -\frac{np^2}{10} + \ln 2 \right) \\
& \leq \exp \left( -\frac{np^2}{10} + \ln 2 - 2 \ln m \right) \\
& \approx \exp \left( -(1 - o(1)) \frac{np^2}{10} \right).
\end{aligned}$$

■

**Lemma 1.2.** *Consider the random sampling scheme described in Section 2. Suppose that  $np^2 \geq C_0 \log m$  and  $mp \geq C'_0 \log n$  for sufficiently large constants  $C_0, C'_0$  (e.g.,  $C_0 \geq 20, C'_0 \geq 10$ ), then*

$$\mathbb{P}(\mathcal{A}^+) \geq 1 - \exp \left( -(1 - o(1)) \frac{np^2}{10} \right) - \exp \left( -(1 - o(1)) \frac{mp}{10} \right).$$

*Proof.* The first term is obtained using the same argument as in the proof of Lemma 1.1. For the second term, we again invoke Chernoff bound.

$$\mathbb{P} \left( \left| \sum_{l=1}^n A_{li} A_{lj} - \mathbb{E}[A_{li} A_{lj}] \right| > \frac{1}{2} mp \right) \leq 2 \exp \left( -\frac{\frac{1}{2} mp}{\frac{1}{2} + 2} \right) = \exp \left( -\frac{mp}{10} + \ln 2 \right).$$

Applying union bound over all users  $n$  gives

$$\begin{aligned}
& \mathbb{P} \left( \exists l \in [n] : \left| \sum_{i=1}^m A_{li} A_{lj} - \mathbb{E}[A_{li} A_{lj}] \right| > \frac{1}{2} mp \right) \\
& \leq n \cdot \exp \left( -\frac{mp}{10} + \ln 2 \right) \\
& = \exp \left( -\frac{mp}{10} + \ln 2 - \log n \right) \\
& \approx \exp \left( -(1 - o(1)) \cdot \frac{mp}{10} \right).
\end{aligned}$$

This completes the proof. ■

Under either event  $\mathcal{A}$  (or  $\mathcal{A}^+$ ) happening, one can set  $d = \frac{3mnp^2}{2}$ . One can verify the following useful inequalities:

$$\pi_{\min}^* \geq \frac{\sum_{i=1}^m \pi_i^*}{me^\kappa} = \frac{1}{me^\kappa}.$$

$$\pi_{\max}^* \leq \frac{e^\kappa \sum_{i=1}^m \pi_i^*}{m} = \frac{e^\kappa}{m}.$$

## 1.2 Basic Entrywise Analysis

Pick an index  $i \in [m]$ . We have the following deterministic decomposition.

$$\begin{aligned}
\pi_i - \pi_i^* &= (\pi^\top M)_i - (\pi^{*\top} M^*)_i \\
&= \sum_j \pi_j M_{ji} - \sum_j \pi_j^* M_{ji}^* \\
&= \sum_j \pi_j M_{ji} - \sum_j \pi_j^* (M_{ji}^* - M_{ji} + M_{ji}) \\
&= \sum_j (\pi_j - \pi_j^*) M_{ji} - \sum_j \pi_j^* (M_{ji}^* - M_{ji}) \\
&= \sum_j (\pi_j - \pi_j^*) M_{ji} + \sum_j \pi_j^* (M_{ji} - M_{ji}^*) \\
&= \underbrace{\sum_{j \neq i} (\pi_j - \pi_j^*) M_{ji}}_{I_1} + \underbrace{\sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*)}_{I_3} + \underbrace{\pi_i^* (M_{ii} - M_{ii}^*)}_{I_4}.
\end{aligned}$$

Rearranging the terms gives

$$\pi_i - \pi_i^* = \underbrace{\frac{1}{1 - M_{ii}}}_{I_2} \cdot \left( \sum_{j \neq i} (\pi_j - \pi_j^*) M_{ji} + \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ii} - M_{ii}^*) \right).$$

The goal in this section and the next is to provide a bound on the absolute deviation of the RHS by bounding each of the terms separately. The main difference between the analysis in this section versus that in the next section is the added assumption that  $m$  grows with  $\log n$ ,  $mp \gtrsim \log n$ . In this section, we do not assume anything about  $m$ . The goal here is to show that when  $m$  is a constant or grows very slowly with  $n$ , the spectral algorithm produces consistent entrywise estimate of the parameter  $\beta^*$ . This is in sharp contrast to JMLE which produces an inconsistent estimate of  $\beta^*$  under this regime (Andersen, 1973; Ghosh, 1995). In the next section, we obtain better (in fact optimal) entrywise error bound when  $m$  grows at least logarithmically with  $n$ .

For a single pair  $(i, j)$ , note that

$$M_{ij} - M_{ij}^* = \frac{1}{d} \sum_{l=1}^n A_{li} A_{lj} \left( X_{li}(1 - X_{lj}) - \mathbb{E}[X_{li}(1 - X_{lj})] \right)$$

is a sum of  $n$  independent random variables. Using Chernoff bound, we have, for  $\alpha \in (0, 1)$ ,

$$\begin{aligned}
&\mathbb{P} \left( |M_{ij} - M_{ij}^*| > \alpha M_{ij}^* \mid \mathcal{A} \right) \\
&\leq 2 \exp \left( -\frac{\alpha^2 M_{ij}^*}{3} \right) \leq 2 \exp \left( -\frac{\alpha^2 (B_{ij} \gamma)}{3} \right) \\
&\leq 2 \exp \left( -\frac{\alpha^2 \gamma np^2}{12} \right).
\end{aligned}$$

Define the following event

$$\mathcal{E}_C = \left\{ |M_{ij} - M_{ij}^*| \leq \frac{\sqrt{12} \sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}} \cdot M_{ij}^* \quad \forall i \neq j \right\}.$$

We have

$$\mathbb{P}(\neg \mathcal{E}_C \mid \mathcal{A}) \leq \frac{2m^2}{(np^2)^C}.$$

We first upper bound  $I_2, I_3, I_4$ .

**Lemma 1.3.** *Suppose events  $\mathcal{A}$  and  $\mathcal{E}_C$  hold for some constant  $C$ . Then*

$$I_2 \geq \left(1 - \frac{\sqrt{12}\sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}}\right) \cdot \frac{\gamma}{3},$$

$$|I_3|, |I_4| \leq \frac{\sqrt{12}\sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}} \cdot \|\pi^*\|_\infty.$$

*Proof.* To avoid cluttering the notation, define  $\alpha_C = \frac{\sqrt{12}\sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}}$ . When both events  $\mathcal{A}$  and  $\mathcal{E}$  hold, recall that  $B_{ij} = \sum_l A_{li} A_{lj} \leq \frac{3np^2}{2}$  and we set  $d = \frac{3mnp^2}{2}$ . We have

$$\begin{aligned} 1 - M_{ii} &= \sum_{j \neq i} M_{ij} \geq \sum_{j \neq i} (1 - \alpha_C) M_{ij}^* \\ &= (1 - \alpha_C) \cdot \sum_{j \neq i} \frac{1}{d} \sum_{l=1}^n A_{li} A_{lj} \mathbb{E}[X_{li}(1 - X_{lj})] \\ &\geq (1 - \alpha_C) \cdot \frac{1}{2d} (m-1) np^2 \gamma \approx (1 - \alpha_C) \cdot \frac{\gamma}{3}. \end{aligned}$$

We will now bound the term  $I_3$ .

$$\begin{aligned} &\sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) \\ &\leq \alpha_C \cdot \sum_{j \neq i} \pi_j^* M_{ji}^* = \alpha_C \cdot \pi_i^* \leq \alpha_C \cdot \|\pi^*\|_\infty. \end{aligned}$$

Next, we bound  $I_4$ , using a similar argument as that used to bound  $I_3$ , we have

$$\begin{aligned} \pi_i^* (M_{ii} - M_{ii}^*) &= \pi_i^* \cdot \sum_{j \neq i} (M_{ij}^* - M_{ij}) \\ &\leq \pi_i^* \cdot \sum_{j \neq i} \alpha_C M_{ij}^* = \alpha_C \cdot \sum_{j \neq i} \pi_i^* M_{ij}^* \\ &= \alpha_C \cdot \sum_{j \neq i} \pi_j^* M_{ji}^* = \alpha_C \pi_i^* \leq \alpha_C \cdot \|\pi^*\|_\infty, \end{aligned}$$

where between the second and third line we use the fact that the Markov chain  $M^*$  is time-reversible (Proposition 2.1 of Nguyen and Zhang (2022)). That is,  $\pi_i^* M_{ij}^* = \pi_j^* M_{ji}^*$ . ■

What remains is to bound the term  $I_1$ .

**Lemma 1.4.** *Suppose that events  $\mathcal{A}$  and  $\mathcal{E}_C$  hold for some constant  $C$ . Then for a constant  $C' > 1$ ,*

$$\mathbb{P} \left( |I_1| \geq \frac{1 + \alpha_C}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C} \cdot \frac{\|\pi^*\|_\infty \sqrt{4C' \max\{1, \log np^2/m\}}}{\sqrt{3np^2}} \right) \leq \min \left\{ \frac{1}{\exp(4(C' - 1)m)}, \frac{1}{(np^2)^{4(C' - 1)}} \right\},$$

where  $\alpha_C = \frac{\sqrt{12}\sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}}$ .

*Proof.* Using Cauchy-Schwarz inequality, we have

$$I_1 \leq \|\pi - \pi^*\|_2 \cdot \sqrt{\sum_{j \neq i} M_{ji}^2}.$$

There are two terms here. For the second term, we have

$$\begin{aligned} \sqrt{\sum_{j \neq i} M_{ji}^2} &\leq \sqrt{\sum_{j \neq i} (M_{ji}^* (1 + \alpha_C))^2} \\ &\leq (1 + \alpha_C) \cdot \sqrt{\sum_{j \neq i} M_{ji}^{*2}} \leq \frac{1 + \alpha_C}{\sqrt{m}}. \end{aligned}$$

The last inequality comes from the fact that conditioned on  $\mathcal{A}$  and  $\mathcal{E}$ ,  $M_{ji}^* \leq \frac{B_{ij}}{d} \leq \frac{3/2np^2}{3/2mnp^2} = \frac{1}{m}$ .

For the first term, we invoke the  $\ell_2$  analysis of the spectral method (cf. Lemma A.3 of Nguyen and Zhang (2022)), we have the following useful deterministic inequality

$$\|\pi - \pi^*\|_2 \leq \frac{\|\pi^{*\top} (M^* - M)\|_2}{\mu(M^*) - \|M - M^*\|_2},$$

where  $\mu(M^*)$  is the spectral gap of  $M^*$ . From Lemma A.6 of Nguyen and Zhang (2022), we have

$$\mu(M^*) \geq \frac{\gamma}{3e^{2\kappa}}.$$

Define  $\alpha_C = \frac{\sqrt{12}\sqrt{C \log np^2}}{\gamma \sqrt{np^2}}$ . We have already established that  $M_{ij}^* \leq \frac{1}{m}$  for any  $i \neq j$ . We have

$$\begin{aligned} \|M - M^*\|_2 &\leq \|\text{diag}(M)I - \text{diag}(M^*)I\|_2 + \|[M - M^*]_{i \neq j}\|_2 \\ &\leq \max_i |M_{ii} - M_{ii}^*| + \max_{u,v: \|u\|=\|v\|=1} \sum_{i \neq j} u_i (M_{ij} - M_{ij}^*) v_j \\ &\leq \max_i \left| \sum_{j \neq i} M_{ij} - M_{ij}^* \right| + \max_{i \neq j} |M_{ij} - M_{ij}^*| \cdot \sum_{i \neq j} |u_i| |v_j| \\ &\leq 2m \cdot \max_{i \neq j} |M_{ij} - M_{ij}^*| \leq 2m \alpha_C \max_{i \neq j} M_{ij}^* \leq 2\alpha_C. \end{aligned}$$

Then

$$\mu(M^*) - \|M - M^*\|_2 \geq \frac{\gamma}{3e^{2\kappa}} - 2\alpha_C.$$

For the numerator term  $\|\pi^{*\top} (M^* - M)\|_2$ , we follow a similar argument as in the proof of Lemma 8.4 in Chen et al. (2020) in turning the normed term into a linear term. This has also appeared in the analysis of the  $\ell_2$  error of the spectral algorithm (cf., Lemma A.9 of Nguyen and Zhang (2022)). For completeness we reproduce this argument here. We first have

$$\begin{aligned} \|\pi^{*\top} (M - M^*)\|_2 &= \sqrt{\sum_{i=1}^n \left( \sum_{j=1}^n \pi_j^* (M_{ji} - M_{ji}^*) \right)^2} \\ &= \sqrt{\sum_{i=1}^n \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ii} - M_{ii}^*) \right)^2} \\ &= \sqrt{\sum_{i=1}^n \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (\sum_{j \neq i} M_{ij}^* - M_{ij}) \right)^2} \\ &= \sqrt{\sum_{i=1}^n \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right)^2}. \end{aligned}$$

Let  $\mathcal{B}$  denote the unit norm ball in  $\mathbb{R}^m$  and  $\mathcal{V}$  denote a  $1/2$ -net of  $\mathcal{B}$ . That is, for every  $u \in \mathcal{B}$ , there exists  $v \in \mathcal{V}$  such that  $\|u - v\|_2 \leq \frac{1}{2}$ . For any  $u \in \mathcal{B}$  and any corresponding  $v$ , we have

$$\begin{aligned} & \sum_{i=1}^n u_i \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right) \\ &= \sum_{i=1}^n v_i \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right) + \sum_{i=1}^n (u_i - v_i) \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right) \\ &\leq \sum_{i=1}^n v_i \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right) + \frac{1}{2} \cdot \sqrt{\sum_{i=1}^n \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right)^2}. \end{aligned}$$

Maximizing both sides of the above inequality with respect to  $u$  and rearranging the terms gives

$$\sqrt{\sum_{i=1}^n \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right)^2} \leq 2 \max_{v \in \mathcal{V}} \sum_{i=1}^n v_i \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right).$$

In summary, we can upper bound the normed term the following linear term.

$$\begin{aligned} \|\pi^{*\top} (M - M^*)\|_2 &= \sqrt{\sum_{i=1}^n \left( \sum_{j=1}^n \pi_j^* (M_{ji} - M_{ji}^*) \right)^2} \\ &\leq 2 \max_{v \in \mathcal{V}} \sum_{i=1}^n v_i \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right). \end{aligned} \tag{1}$$

We now expand on the linear term.

$$\begin{aligned} & \sum_{i=1}^m v_i \left( \sum_{j \neq i} \pi_j^* (M_{ji} - M_{ji}^*) + \pi_i^* (M_{ij}^* - M_{ij}) \right) \\ &= \frac{1}{d} \sum_{i=1}^m v_i \left( \sum_{j \neq i} \pi_j^* \left[ \sum_{l=1}^n A_{li} A_{lj} (X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]) \right] \right. \\ &\quad \left. - \sum_{j \neq i} \pi_i^* \left[ \sum_{l=1}^n A_{li} A_{lj} (X_{li}(1 - X_{lj}) - \mathbb{E}[X_{li}(1 - X_{lj})]) \right] \right) \\ &= \frac{1}{d} \sum_{l=1}^n \sum_{i=1}^m \left( \sum_{j \neq i} v_i \pi_j^* A_{li} A_{lj} [X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]] \right) \\ &\quad - \frac{1}{d} \sum_{l=1}^n \sum_{i=1}^m \left( \sum_{j \neq i} v_i \pi_i^* A_{li} A_{lj} [X_{li}(1 - X_{lj}) - \mathbb{E}[X_{li}(1 - X_{lj})]] \right) \\ &= \frac{1}{d} \sum_{l=1}^n \sum_{i=1}^m \sum_{j \neq i} \left( (v_i - v_j) \pi_j^* A_{li} A_{lj} [X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]] \right). \end{aligned}$$

We will use the method of bounded difference to obtain a high probability bound on the above sum. Note that this sum is essentially a function  $f$  of  $n \times m$  independent Bernoulli random variables  $\{X_{li}\}$ . Let  $X$  and  $X'$  be identical copies except for  $X_{li} \neq X'_{li}$ . That is, when one changes the outcome of a single random variable  $X_{li}$ . We have the following bound on the deviation of  $f$  following such changes.

$$|f(X) - f(X')| = \frac{1}{d} \left| \sum_{j \neq i} A_{li} A_{lj} (v_i - v_j) [X_{lj}(\pi_i^* - \pi_j^*) - \pi_i^*] \right|. \tag{2}$$



Using Cauchy-Schwarz, we can upper bound the absolute difference term as

$$\begin{aligned} & \frac{1}{d} \sum_{j \neq i} A_{li} A_{lj} (v_i - v_j) \underbrace{[X_{lj}(\pi_i^* - \pi_j^*) - \pi_i^*]}_{\leq \max\{\pi_i^*, \pi_j^*\} \leq \|\pi^*\|_\infty} \\ & \leq \frac{1}{d} \|\pi^*\|_\infty \cdot \sqrt{m} \cdot \sqrt{\sum_{j \neq i} A_{li} A_{lj} (v_i - v_j)^2}. \end{aligned}$$

At this point we can invoke concentration inequalities based on bounded difference (e.g., Hoeffding's inequality). We have

$$\begin{aligned} & \mathbb{P}\left(\|\pi^{*\top}(P - M^*)\|_2 > t \mid \mathcal{A}\right) \\ & \leq \mathbb{P}\left(\frac{1}{d} \sum_{l=1}^n \sum_{i=1}^m \sum_{j \neq i} (v_i - v_j) \pi_j^* A_{li} A_{lj} [X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]] > t \quad \forall v \in \mathcal{V} \mid \mathcal{A}\right) \\ & \leq 2 \cdot 5^m \cdot \exp\left(-\frac{4d^2 t^2}{3\|\pi^*\|_\infty^2 m^2 n p^2}\right) \leq \exp\left(-\frac{3np^2 t^2}{\|\pi^*\|_\infty^2} + 4m\right), \end{aligned}$$

where the  $5^m$  term comes from applying union bound over all  $v \in \mathcal{V}$  and we know that  $|\mathcal{V}| \approx 5^m$  (cf. Corollary 4.2.13 Vershynin (2018)). Set

$$t = \frac{\|\pi^*\|_\infty \sqrt{4C' \max\{m, \log np^2\}}}{\sqrt{3np^2}}.$$

We have

$$\begin{aligned} & \mathbb{P}\left(\|\pi^{*\top}(P - M^*)\|_2 > \frac{\|\pi^*\|_\infty \sqrt{4C' \max\{m, \log np^2\}}}{\sqrt{3np^2}} \mid \mathcal{A}\right) \leq \exp(-4(C' - 1) \max\{m, \log np^2\}) \\ & = \min\left\{\frac{1}{\exp(4(C' - 1)m)}, \frac{1}{(np^2)^{4(C' - 1)}}\right\} \end{aligned}$$

Combining the bounds for  $\mu(M^*) - \|M - M^*\|_2 \geq \frac{\gamma}{3e^{2\kappa}} - 2\alpha_C$  and  $\sqrt{\sum_{j \neq i} M_{ji}^2} \leq \frac{1+\alpha_C}{\sqrt{m}}$  obtained earlier finishes the proof. ■

**Theorem 1.5.** *There exist a constant  $C$  such that if  $np^2 \geq C \log m$ . For any  $C' > 1$ , with probability at least*

$$1 - m \left( \min \left\{ \frac{1}{\exp(4(C' - 1)m)}, \frac{1}{(np^2)^{4(C' - 1)}} \right\} \right) + \frac{2m^2}{(np^2)^C} + \exp\left(-\frac{(1 - o(1))np^2}{10}\right),$$

*the following holds true.*

$$\frac{\|\pi - \pi^*\|_\infty}{\|\pi\|_\infty} \leq \frac{3}{\gamma(1 - \alpha_C)} \cdot \left( \frac{1 + \alpha_C}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C} \cdot \frac{\sqrt{4C' \max\{1, \log np^2/m\}}}{\sqrt{3np^2}} + 2\alpha_C \right),$$

where

$$\alpha_C = \frac{\sqrt{12} \sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}}.$$

*Proof.* Suppose that events  $\mathcal{A}, \mathcal{E}_C$  hold, then

$$\begin{aligned} |\pi - \pi_i^*| &\leq \frac{1}{1 - M_{ii}} \cdot (|I_1| + |I_3| + |I_4|) \\ &\leq \frac{3}{\gamma(1 - \alpha_C)} \cdot \left( \frac{1 + \alpha_C}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C} \cdot \frac{\|\pi^*\|_\infty \sqrt{4C' \max\{1, \log np^2/m\}}}{\sqrt{3np^2}} + \frac{2\sqrt{12}\sqrt{C \log np^2}}{\gamma \sqrt{np^2}} \cdot \|\pi^*\|_\infty \right) \end{aligned}$$

with probability at least  $1 - \exp(-4(C' - 1))$ . In other words

$$\mathbb{P} \left( \left| \pi_i - \pi_i^* \right| \geq \underbrace{\frac{3}{\gamma(1 - \alpha_C)} \cdot \left( \frac{1 + \alpha_C}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C} \cdot \frac{\sqrt{4C' \max\{1, \log np^2/m\}}}{\sqrt{3}} + \frac{2\sqrt{12}\sqrt{C \log np^2}}{\gamma} \right)}_{\alpha_{C,C'}^*} \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np^2}} \middle| \mathcal{A}, \mathcal{E}_C \right) \leq \exp(-4(C' - 1)m).$$

We have

$$\begin{aligned} &\mathbb{P} \left( \exists i \in [m] : |\pi_i - \pi_i^*| \geq \alpha_{C,C'}^* \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np^2}} \right) \\ &\leq \mathbb{P} \left( \exists i \in [m] : |\pi_i - \pi_i^*| \geq \alpha_{C,C'}^* \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np^2}} \middle| \mathcal{A}, \mathcal{E}_C \right) + \mathbb{P}(\neg \mathcal{A}) + \mathbb{P}(\neg \mathcal{E}_C) \\ &\leq \mathbb{P} \left( \exists i \in [m] : |\pi_i - \pi_i^*| \geq \alpha_{C,C'}^* \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np^2}} \middle| \mathcal{A}, \mathcal{E}_C \right) + 2\mathbb{P}(\neg \mathcal{A}) + \mathbb{P}(\neg \mathcal{E}_C | \mathcal{A}) \\ &\leq m \cdot \min \left\{ \frac{1}{\exp(4(C' - 1)m)}, \frac{1}{(np^2)^{4(C' - 1)}} \right\} + \exp \left( -\frac{(1 - o(1))np^2}{10} \right) + \frac{2m^2}{(np^2)^C}. \end{aligned}$$

The last inequality in the derivation comes applying union bound over all  $i \in [m]$ . ■

The result we have obtained provides a bound on  $\frac{\|\pi - \pi^*\|_\infty}{\pi_{\max}^*}$ . Recall that we are ultimately interested in  $\|\beta - \beta^*\|_\infty$ .

**Lemma 1.6.** *Let  $\pi$  be the stationary distribution of the empirical Markov chain constructed in Algorithm 1 and  $\beta$  the final output. Then*

$$\|\beta - \beta^*\|_\infty \leq 2me^\kappa \cdot \|\pi - \pi^*\|_\infty.$$

*Proof.* Let  $L$  be a factor such that  $|\log x - \log x'| \leq L|x - x'|$  for all  $x, x' \in [\pi_{\min}^*, \pi_{\max}^*]$ . Note that  $L$  may change with  $m$  as we will show later. The output of the spectral algorithm is essentially

$$\beta_i = \log \pi_i - \frac{1}{m} \sum_{k=1}^m \log \pi_k,$$

where  $\pi$  is the stationary distribution of the Markov chain constructed in Algorithm 1. And  $\beta^*$  is similarly related to  $\pi^*$ :

$$\beta_i^* = \log \pi_i^* - \frac{1}{m} \sum_{k=1}^m \log \pi_k^*.$$

One can see that

$$\begin{aligned}
|\beta_i - \beta_i^*| &\leq |\log \pi_i - \log \pi_i^*| + \frac{1}{m} \sum_{k=1}^m |\log \pi_k - \log \pi_k^*| \\
&\leq 2 \max_k |\log \pi_k - \log \pi_k^*| \\
&\leq 2L \max_k |\pi_k - \pi_k^*| \\
&\leq 2L \cdot \|\pi - \pi^*\|_\infty.
\end{aligned}$$

Observe that  $\pi_{\min}^* \leq \frac{1}{me^\kappa}$ . One can thus easily see that the log function within the dynamic range has gradient absolutely bounded by  $me^\kappa$ . Therefore  $L \leq me^\kappa$ . Substituting this upper bound on  $L$  into the inequality obtained above finishes the proof. ■

**Theorem 1.7.** *There exists a constant  $C$  such that if  $np^2 \geq C \log m$ , then with probability at least*

$$1 - m \left( \min \left\{ \frac{1}{\exp(4(C' - 1)m)}, \frac{1}{(np^2)^{4(C' - 1)}} \right\} \right) + \frac{2m^2}{(np^2)^C} + \exp \left( -\frac{(1 - o(1))np^2}{10} \right),$$

*the following holds true for any constant  $C' > 1$ .*

$$\|\beta - \beta^*\|_\infty \leq \frac{6e^{2\kappa}}{\gamma(1 - \alpha_C)} \cdot \left( \frac{1 + \alpha_C}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C} \cdot \frac{\sqrt{4C' \max\{1, \log np^2/m\}}}{\sqrt{3np^2}} + 2\alpha_C \right),$$

where

$$\alpha_C = \frac{\sqrt{12}\sqrt{C \log np^2}}{\sqrt{\gamma} \sqrt{np^2}}.$$

*Proof.* The proof combines Theorem 1.5, Lemma 1.6 and the fact that  $\|\pi^*\|_\infty \leq \frac{e^\kappa}{m}$ . ■

### 1.3 Improved Rates for the Regime $mp \gtrsim \log n$

In this section, we present the proof of Theorem 3.1 in the main paper. For a single pair  $(i, j)$ , note that

$$M_{ij} - M_{ij}^* = \frac{1}{d} \sum_{l=1}^n A_{li} A_{lj} \left( X_{li}(1 - X_{lj}) - \mathbb{E}[X_{li}(1 - X_{lj})] \right)$$

is a sum of  $n$  independent random variables. Again using Chernoff bound, we have, for  $\alpha \in (0, 1)$ ,

$$\begin{aligned}
&\mathbb{P} \left( |M_{ij} - M_{ij}^*| > \alpha M_{ij}^* \mid \mathcal{A}^+ \right) \\
&\leq 2 \exp \left( -\frac{\alpha^2 M_{ij}^*}{3} \right) \leq 2 \exp \left( -\frac{\alpha^2 (B_{ij} \gamma)}{3} \right) \\
&\leq 2 \exp \left( -\frac{\alpha^2 \gamma np^2}{12} \right)
\end{aligned}$$

Define the following event

$$\mathcal{E}_C^+ = \left\{ |M_{ij} - M_{ij}^*| \leq \frac{\sqrt{12}\sqrt{C \log m}}{\sqrt{\gamma} \sqrt{np^2}} \cdot M_{ij}^* \quad \forall i \neq j \right\}.$$

We have

$$\mathbb{P}(\neg \mathcal{E}_C^+ \mid \mathcal{A}) \leq \frac{2}{m^{C-2}}.$$

**Lemma 1.8.** Suppose events  $\mathcal{A}^+$  and  $\mathcal{E}_C^+$  hold for some constant  $C$ . There exist a constant  $C^+$  such that

$$I_2 \geq \left(1 - \frac{\sqrt{12}\sqrt{C \log m}}{\sqrt{\gamma} \sqrt{np^2}}\right) \cdot \frac{\gamma}{3},$$

$$|I_3|, |I_4| \leq \frac{\sqrt{12}\sqrt{C^+ \log m}}{\sqrt{2} \sqrt{np}} \cdot \|\pi^*\|_\infty$$

with probability at least  $1 - \frac{4}{m^{C^+}}$ .

*Proof.* Note that event  $\mathcal{A}^+$  is a special case of event  $\mathcal{A}$ . Therefore, the first statement is the same as that in Lemma 1.3. More precisely, conditioned on event  $\mathcal{A}^+$ , the first statement holds. It remains to obtain a sharper bound for  $I_3$  and  $I_4$ .

We first tackle the term  $I_3$  using the method of bounded difference.

$$I_3 = \sum_{j \neq i} \pi_j^* (P_{ji} - P_{ji}^*) = \frac{1}{d} \cdot \sum_{j \neq i} \pi_j^* \cdot \sum_{l=1}^n A_{li} A_{lj} [X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]].$$

For a fixed  $l$  if we change  $X_{li}$ , the sum above changes by at most  $\frac{1}{d} \cdot \sum_{j \neq i} \pi_j^* A_{li} A_{lj}$ . On the other hand, for a fixed  $l$  and  $j \neq i$ , if we change  $X_{lj}$ , the sum above changes by at most  $\frac{1}{d} \cdot \pi_j^* A_{li} A_{lj} \leq \frac{1}{d} \cdot \|\pi^*\|_\infty A_{li} A_{lj}$ . Note also that under the growing  $m$  regime and by Cauchy-Schwarz inequality, we have

$$\begin{aligned} \sum_{j \neq i} \frac{1}{d} \pi_j^* A_{li} A_{lj} &\leq \|\pi^*\|_\infty \cdot \sum_{j \neq i} A_{lj} \cdot A_{li} A_{lj} \\ &\leq \frac{1}{d} \|\pi^*\|_\infty \cdot \sqrt{\sum_{j \neq i} A_{lj}} \cdot \sqrt{\sum_{j \neq i} A_{li} A_{lj}} \\ &\leq \frac{1}{d} \|\pi^*\|_\infty \cdot \sqrt{3/2} \cdot \sqrt{mp} \cdot \sqrt{\sum_{j \neq i} A_{li} A_{lj}}. \end{aligned}$$

Invoking Hoeffding's inequality, we have

$$\begin{aligned} &\mathbb{P}\left(\sum_{j \neq i} \pi_j^* \cdot \sum_{l=1}^n A_{li} A_{lj} [X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]] > t \mid \mathcal{A}^+\right) \\ &\leq 2 \exp\left(-\frac{2t^2}{\frac{1}{d^2} \|\pi^*\|_\infty^2 \left(\sum_{l=1}^n (3/2)mp \sum_{j \neq i} A_{li} A_{lj} + \sum_{j \neq i} \sum_{l=1}^n A_{li} A_{lj}\right)}\right) \\ &= 2 \exp\left(-\frac{2t^2}{\frac{1}{d^2} \|\pi^*\|_\infty^2 (3/2 mp + 1) \left(\sum_{l=1}^n \sum_{j \neq i} A_{li} A_{lj}\right)}\right) \\ &\leq 2 \exp\left(-\frac{2d^2 t^2}{\|\pi^*\|_\infty^2 (3/2 mp + 1) (m-1) \frac{3}{2} np^2}\right) \\ &\leq 2 \exp\left(-\frac{2d^2 t^2}{\|\pi^*\|_\infty^2 9/4 mp \cdot mnp^2}\right) \\ &\quad \left[\text{Setting } d = \frac{3}{2} mnp^2\right] \\ &= 2 \exp\left(-\frac{2mnp^2 t^2}{\|\pi^*\|_\infty^2 mp}\right) = 2 \exp\left(-\frac{2npt^2}{\|\pi^*\|_\infty^2}\right). \end{aligned}$$

We have, conditioned on event  $\mathcal{A}^+$  and event  $\mathcal{E}_C^+$ ,

$$|I_3| \leq \frac{C_3 \sqrt{\log m} \|\pi^*\|_\infty}{\sqrt{2} \sqrt{np}}$$

with probability at least  $1 - \frac{2}{m^{C_3}}$ .

Similarly, we can bound the term  $I_4$  using the method of bounded difference. Specifically, we have

$$\begin{aligned} I_4 &= \pi_i^*(M_{ii} - M_{ii}^*) \\ &= \pi_i^* \sum_{j \neq i} \sum_{l=1}^n \frac{1}{d} (A_{li} A_{lj} X_{li} (1 - X_{lj}) - \mathbb{E}[X_{li} (1 - X_{lj})]) \\ &\leq \|\pi^*\|_\infty \sum_{j \neq i} \sum_{l=1}^n \frac{1}{d} (A_{li} A_{lj} X_{li} (1 - X_{lj}) - \mathbb{E}[X_{li} (1 - X_{lj})]) . \end{aligned}$$

One can see that this can be bounded in the same way as  $I_3$  shown above. Therefore,

$$|I_4| \leq \frac{C_4 \sqrt{\log m} \|\pi^*\|_\infty}{\sqrt{2} \sqrt{np}}$$

with probability at least  $1 - \frac{2}{m^{C_4}}$ . To obtain the statement in the theorem we pick a common  $C^+ = C_3 = C_4$  and apply union bound. Namely, conditioned on event  $A^+$ ,

$$|I_3|, |I_4| \leq \frac{\sqrt{C^+ \log m} \|\pi^*\|_\infty}{\sqrt{2} \sqrt{np}}$$

with probability at least  $1 - \frac{4}{m^{C^+}}$ . ■

We can obtain sharper bound for the term  $I_1$  under the assumption that  $mp \gtrsim \log n$  as well.

**Lemma 1.9.** *Suppose that events  $\mathcal{A}^+$  and  $\mathcal{E}_C^+$  hold for some constant  $C$ . Then for  $C_1^+ > 1$ ,*

$$\mathbb{P} \left( |I_1| \geq \frac{1 + \alpha_C^+}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+} \cdot \frac{\|\pi^*\|_\infty \sqrt{4C_1^+ \log m}}{\sqrt{3np}} \right) \leq \exp(-4(C_1^+ - 1)m) ,$$

where  $\alpha_C^+ = \frac{\sqrt{12} \sqrt{C \log m}}{\sqrt{\gamma} \sqrt{np^2}}$ .

*Proof.* Using Cauchy-Schwarz inequality, we have

$$I_1 \leq \|\pi - \pi^*\|_2 \cdot \sqrt{\sum_{j \neq i} M_{ji}^2} .$$

Following the same argument in the proof of Lemma 1.4, we have

$$\sqrt{\sum_{j \neq i} M_{ji}^2} \leq \frac{1 + \alpha_C^+}{\sqrt{m}}$$

and

$$\mu(M^*) - \|M - M^*\|_2 \geq \frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+ .$$

Now, we follow a similar argument as in the proof of Lemma 1.8 but with an added assumption that  $mp \gtrsim \log n$  (event  $\mathcal{A}^+$ ) to obtain a sharper bound for  $\|\pi - \pi^*\|_2$ . Firstly, follow the same argument as in the proof of Lemma 1.4. The key difference is that under conditioned  $\mathcal{A}^+$  we could obtain better bound also using the

bounded difference method. Namely, one can invoke Cauchy-Schwarz on the absolute difference term in Equation (2) as follows:

$$\begin{aligned}
& \frac{1}{d} \sum_{j \neq i} A_{li} A_{lj} (v_i - v_j) \underbrace{[X_{lj}(\pi_i^* - \pi_j^*) - \pi_i^*]}_{\leq \max\{\pi_i^*, \pi_j^*\} \leq \frac{e^\kappa}{m}} \\
& \leq \frac{1}{d} \frac{e^\kappa}{m} \cdot \sum_{j \neq i} A_{li} A_{lj} (v_i - v_j) = \frac{1}{d} \frac{e^\kappa}{m} \cdot \sum_{j \neq i} A_{lj} A_{li} A_{lj} (v_i - v_j) \\
& \leq \frac{1}{d} \frac{e^\kappa}{m} \cdot \sqrt{\underbrace{\sum_j A_{lj}}_{\leq \frac{3}{2}mp}} \cdot \sqrt{\sum_{j \neq i} A_{li} A_{lj} (v_i - v_j)^2} = \frac{1}{d} \sqrt{3/2} \cdot \frac{e^\kappa \sqrt{p}}{\sqrt{m}} \cdot \sqrt{\sum_{j \neq i} A_{li} A_{lj} (v_i - v_j)^2}.
\end{aligned}$$

Define  $\mathcal{V}$  to be the  $\frac{1}{2}$ -net of the  $\ell_2$  ball. We have

$$\begin{aligned}
& \mathbb{P}\left(\|\pi^{*\top}(P - M^*)\|_2 > t \mid \mathcal{A}^+\right) \\
& \leq \mathbb{P}\left(\frac{1}{d} \sum_{l=1}^n \sum_{i=1}^m \sum_{j \neq i} (v_i - v_j) \pi_j^* A_{li} A_{lj} [X_{lj}(1 - X_{li}) - \mathbb{E}[X_{lj}(1 - X_{li})]] > t \quad \forall v \in \mathcal{V} \mid \mathcal{A}^+\right) \\
& \leq 2 \cdot 5^m \cdot \exp\left(-\frac{2d^2 t^2}{\|\pi^*\|_\infty^2 \cdot \sum_{l=1}^n \sum_{i=1}^m mp \sum_{j \neq i} A_{li} A_{lj} (v_i - v_j)^2}\right) \\
& \leq 2 \cdot 5^m \cdot \exp\left(-\frac{2d^2 t^2}{mp \|\pi^*\|_\infty^2 \sum_{i \neq j} B_{ij} (v_i - v_j)^2}\right) \\
& \leq \exp\left(-\frac{4/3 d^2 t^2}{\|\pi^*\|_\infty^2 mp np^2} + 4m\right) \\
& = \exp\left(-\frac{3mnp t^2}{\|\pi^*\|_\infty^2} + 4m\right).
\end{aligned}$$

Set

$$t = \frac{\|\pi^*\|_\infty \sqrt{4/3 C_1^+ \log m}}{\sqrt{np}}.$$

Simple algebraic manipulations show that

$$\mathbb{P}\left(\|\pi^{*\top}(P - M^*)\|_2 > \frac{\|\pi^*\|_\infty \sqrt{4/3 C_1^+ \log m}}{\sqrt{np}} \mid \mathcal{A}^+\right) \leq \exp(-4(C_1^+ - 1)m).$$

Combining the bounds for  $\mu(M^*) - \|M - M^*\|_2 \geq \frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+$  and  $\sqrt{\sum_{j \neq i} M_{ji}^2}$  obtained earlier finishes the proof. ■

We are now ready to put all of the bounds on  $I_1, I_3, I_4$  together.

**Theorem 1.10.** *There exist constants  $C, C'$  such that if  $np^2 \geq C \log m$  and that  $mp \geq C' \log n$ , then with probability at least*

$$1 - m \left( \exp(-4(C_1^+ - 1)m) + \frac{4}{m^{C^+}} \right) + \exp\left(-\frac{(1 - o(1))np^2}{10}\right) + \exp\left(-\frac{(1 - o(1))mp}{10}\right) + \frac{2}{m^{C-2}},$$

the following holds true for any  $C_1^+, C^+ > 1$ .

$$\frac{\|\pi - \pi^*\|_\infty}{\|\pi\|_\infty} \leq \frac{3}{\gamma(1 - \alpha_C^+)} \cdot \left( \frac{1 + \alpha_C^+}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+} \cdot \sqrt{4C_1^+ \log m} + \frac{2\sqrt{12}\sqrt{C^+ \log m}}{\sqrt{\gamma}} \right) \cdot \frac{1}{\sqrt{np}},$$

where

$$\alpha_C^+ = \frac{\sqrt{12}\sqrt{C\log m}}{\sqrt{\gamma}\sqrt{np^2}}.$$

*Proof.* We combine the implications of Lemma 1.5 and Lemma 1.9. Suppose that events  $\mathcal{A}^+, \mathcal{E}_C^+$  hold for some constant  $C$ , then

$$\begin{aligned} |\pi - \pi_i^*| &\leq \frac{1}{1 - M_{ii}} \cdot (|I_1| + |I_3| + |I_4|) \\ &\leq \frac{3}{\gamma(1 - \alpha_C^+)} \cdot \left( \frac{1 + \alpha_C^+}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+} \cdot \frac{\|\pi^*\|_\infty \sqrt{4C_1^+ \log m}}{\sqrt{3np}} + \frac{2\sqrt{12}\sqrt{C^+ \log m}}{\sqrt{\gamma}\sqrt{np}} \cdot \|\pi^*\|_\infty \right) \end{aligned}$$

with probability at least  $1 - \exp(-4(C_1^+ - 1)) - \frac{4}{m^{C^+}}$ . We invert the probability to have

$$\begin{aligned} \mathbb{P} \left( |\pi_i - \pi_i^*| \geq \underbrace{\frac{3}{\gamma(1 - \alpha_C^+)} \cdot \left( \frac{1 + \alpha_C^+}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+} \cdot \frac{\sqrt{4C_1^+ \log m}}{\sqrt{3}} + \frac{2\sqrt{12}\sqrt{C^+ \log m}}{\sqrt{\gamma}} \right)}_{\alpha_{C^+, C_1^+}^*} \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np}} \middle| \mathcal{A}^+, \mathcal{E}_C^+ \right) \\ \leq \exp(-4(C_1^+ - 1)m) + \frac{4}{m^{C^+}}. \end{aligned}$$

We have

$$\begin{aligned} \mathbb{P} \left( \exists i \in [m] : |\pi_i - \pi_i^*| \geq \alpha_{C^+, C_1^+}^* \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np}} \right) \\ \leq \mathbb{P} \left( \exists i \in [m] : |\pi_i - \pi_i^*| \geq \alpha_{C^+, C_1^+}^* \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np}} \middle| \mathcal{A}^+, \mathcal{E}_C^+ \right) + \mathbb{P}(\neg \mathcal{A}^+) + \mathbb{P}(\neg \mathcal{E}_C^+) \\ \leq \mathbb{P} \left( \exists i \in [m] : |\pi_i - \pi_i^*| \geq \alpha_{C^+, C_1^+}^* \cdot \frac{\|\pi^*\|_\infty}{\sqrt{np}} \middle| \mathcal{A}^+, \mathcal{E}_C^+ \right) + 2\mathbb{P}(\neg \mathcal{A}^+) + \mathbb{P}(\neg \mathcal{E}_C^+ | \mathcal{A}^+) \\ \leq m \left( \exp(-4(C_1^+ - 1)m) + \frac{4}{m^{C^+}} \right) + \exp\left(-\frac{(1 - o(1))np^2}{10}\right) + \exp\left(-\frac{(1 - o(1))mp}{10}\right) + \frac{2}{m^{C-2}}. \end{aligned}$$

The last step in the derivation applies union bound over all  $i \in [m]$ . ■

Combining Lemma 1.6 with Theorem 1.10 gives the following extended version of Theorem 3.1.

**Theorem 3.1.** *There exist constants  $C, C' > 1$  such that if  $np^2 \geq C \log m$  and that  $mp \geq C' \log n$ , then with probability at least*

$$1 - m \left( \exp(-4(C_1^+ - 1)m) + \frac{4}{m^{C^+ \log m}} \right) + \exp\left(-\frac{(1 - o(1))np^2}{10}\right) + \exp\left(-\frac{(1 - o(1))mp}{10}\right),$$

the following holds true for any  $C_1^+, C^+ > 1$ .

$$\|\beta - \beta^*\|_\infty \leq \frac{6e^{2\kappa}}{\gamma(1 - \alpha_C^+)} \cdot \left( \frac{1 + \alpha_C^+}{\frac{\gamma}{3e^{2\kappa}} - 2\alpha_C^+} \cdot \sqrt{4C_1^+ \log m} + \frac{2\sqrt{12}\sqrt{C^+ \log m}}{\sqrt{\gamma}} \right) \cdot \frac{1}{\sqrt{np}},$$

where

$$\alpha_C^+ = \frac{\sqrt{12}\sqrt{C\log m}}{\sqrt{\gamma}\sqrt{np^2}}.$$

*Proof.* The proof is similar to that of Theorem 1.7. ■

To simplify the upper bound in theorem statement into the form we include in the main paper, observe that for sufficiently large  $C_1^+$  and  $C^+$ , we can simplify

$$\frac{6e^{2\kappa}}{\gamma(1 - \alpha_C^+)} = \mathcal{O}(1).$$

For the two terms inside the paranthesis, the first term reduces to a constant factor. While the second term scale as  $\mathcal{O}(\sqrt{\log m})$ . Also for sufficiently large  $C_1^+$  and  $C^+$ , the probability bound can be further simplified as

$$1 - \mathcal{O}(\exp(-4m)) - 4/m^{10} - \exp(-np^2/10) - \frac{1}{n^{10}}.$$

Note that under the assumption  $np^2 \gtrsim \log m$ , the third term can be absorbed into the  $\mathcal{O}(m^{-10})$  term.

**Theorem 3.2.** *Consider the setting of Theorem 3.1. Consider a top- $K$  estimator that first runs the spectral algorithm on the response data and returns the  $K$  items with the highest parameter values. There exists a constant  $C_K$  such that if  $np \geq \frac{C_K \log m}{\Delta_K^2}$ , then this top- $K$  estimator correctly identifies all of the top  $K$  items with probability at least  $1 - \mathcal{O}(m^{-10}) - \exp(-\mathcal{O}(np^2)) - \mathcal{O}(n^{-10})$  where  $C$  is an absolute constant.*

*Proof.* One can see that given sufficiently large  $n$  such that

$$\|\beta - \beta^*\|_\infty < \Delta_K,$$

then the top  $K$  items with the highest parameter values per the spectral algorithm is also the true top  $K$  items. ■



## 2 Proofs of Lower Bounds

### 2.1 Estimation Error Lower Bound

In this section we prove the lower bound for the estimation error presented in the main paper. We emphasize that this lower bound is different from the Cramer Rao lower bound obtained in Nguyen and Zhang (2022). Specifically, Cramer-Rao lower bounds apply only to the class of unbiased estimators. However, the information theoretic lower bound presented here applies to all statistical estimators, a strictly larger class of estimators. We first restate the theorem statement and defer the proof to later.

**Theorem 3.3.** *Consider the sampling model described in Section 2 and further assume that  $np = \Omega(1)$ . There exists a class of Rasch models  $\mathcal{B}$  such that for any statistical estimator, the minimax risk is lower bounded as*

$$\inf_{\hat{\beta}} \sup_{\theta^* \in \mathcal{T}, \beta^* \in \mathcal{B}} \mathbb{E} \|\hat{\beta} - \beta^*\|_2^2 = \Omega\left(\frac{m}{np}\right).$$

The proof of Theorem 3.3, which we defer to later, relies on the pairwise Fano's inequality (Thomas and Joy, 2006).

**Lemma 2.1.** (Pairwise Fano minimax lower bound) *Suppose that we can construct a set  $\mathcal{W} = \{\beta^{*1}, \dots, \beta^{*M}\}$  with cardinality  $M$  such that*

$$\max_{a \neq b \in [M]} \|\beta^{*a} - \beta^{*b}\|_2^2 \leq \delta^2 \quad \text{and} \quad \max_{a \neq b \in [M]} \mathbb{D}_{KL}(\mathbb{P}^a(X) \|\mathbb{P}^b(X)) \leq \xi,$$

where  $\mathbb{P}^a(X)$  denotes the distribution over  $X$  under a model parametrized by  $\beta^{*a}$ . Then the minimax risk is lower bounded as

$$\inf_{\hat{\beta}} \sup_{\beta^* \in \mathcal{B}} \mathbb{E} [\|\hat{\beta} - \beta^*\|_2^2] \geq \frac{\delta^2}{2} \left(1 - \frac{\xi + \log 2}{\log M}\right),$$

where  $\hat{\beta}$  is the output of any statistical estimator.

Intuitively, the theorem says that if we can construct a set of models where every pair of model is sufficiently different (measured in  $\ell_2^2$  distance) yet they parametrize ‘similar’ distributions (measured in KL divergence), then any statistical estimator will fail to identify the correct model and thus suffer from a minimum estimation error. To construct the set of models, we follow the construction similar to that in the proof of Theorem 1 of Shah et al. (2015). We first restate a coding theoretic due to Shah et al. (2015).

**Lemma 2.2.** (Lemma 7 of Shah et al. (2015)) *For any  $\alpha \in (0, \frac{1}{4})$ , there exists a set of  $M(\alpha) = \Theta(\exp(m))$  binary vectors  $\{z^1, \dots, z^{M(\alpha)}\} \subset \{0, 1\}^m$  such that*

$$\alpha m \leq \|z^a - z^b\|_2^2 \leq m \quad \forall a \neq b \in [1, \dots, M(\alpha)].$$

Consider a set  $\{z^1, \dots, z^{M(\alpha)}\} \subset \{0, 1\}^m$  of  $m$ -dimensional binary vectors given by as Lemma 2.2. Let

$$\beta^{*a} = \frac{\delta}{\sqrt{m}} \cdot z^a \quad \forall a \in [M(\alpha)],$$

where  $\delta$  is to be determined later. It is easy to see that for  $a \neq b \in [M(\alpha)]$

$$\|\beta^{*a} - \beta^{*b}\|_2^2 = \frac{\delta^2}{m} \cdot \|z^a - z^b\|_2^2 \leq \delta^2.$$

**Lemma 2.3** (Reverse Pinsker's inequality). *Consider two probability measures  $P$  and  $Q$  defined on the same measure space  $(\mathcal{A}, \mathcal{F})$  the KL divergence between  $P$  and  $Q$  can be bounded as*

$$\mathbb{D}_{KL}(P \| Q) \leq \left(\frac{\log e}{Q_{\min}}\right) \cdot \|P - Q\|_{TV}^2,$$

where  $\|P - Q\|_{TV}$  is the total variation distance between the two distributions and  $Q_{\min} := \min_{x \in \mathcal{A}} Q(x)$ .

For our lower bound, it suffices to consider the case where  $\theta_1 = \dots = \theta_n = \frac{1}{2}$ . Denote  $\mathcal{T} = \{\theta : \theta_1 = \dots = \theta_n = \frac{1}{2}\}$ .

**Lemma 2.4.** *Consider  $\theta \in \mathcal{T}$  and any two alternative parameter sets  $\beta^a$  and  $\beta^b$ . There exists a constant  $c^* = c^*(\beta_{\min}^*, \beta_{\max}^*)$  such that*

$$\mathbb{D}_{\text{KL}}(\mathbb{P}^a(X) \parallel \mathbb{P}^b(X)) \leq \frac{e^2 \log e}{2c^*} \cdot np\delta^2.$$

*Proof.* As a short hand notation, define  $\mathbb{P}_i^a := \frac{e^\theta}{e^\theta + e^{\beta_i^a}}$  and analogously for  $\mathbb{P}_i^b$ .

$$\begin{aligned} \mathbb{D}_{\text{KL}}(\mathbb{P}^a(X) \parallel \mathbb{P}^b(X)) &= \sum_{l=1}^n \sum_{i=1}^m \mathbb{D}_{\text{KL}}(\mathbb{P}^a(X_{li}) \parallel \mathbb{P}^b(X_{li})) \\ &= np \sum_{i=1}^m \mathbb{D}_{\text{KL}}(\mathbb{P}^a \parallel \mathbb{P}^b) \\ &\quad [\text{Using the inverse Pinsker's inequality}] \\ &\leq np \sum_{i=1}^m \frac{\log e}{c^*} \cdot 2(\mathbb{P}_i^a - \mathbb{P}_i^b)^2 \\ &\leq \frac{np 2 \log e}{c^*} \cdot \sum_{i=1}^m \frac{e^2 \delta^2}{4m} \\ &\leq \frac{e^2 \log e}{2c^*} \cdot np\delta^2. \end{aligned}$$

The second line follows from

$$\begin{aligned} \mathbb{D}_{\text{KL}}(\mathbb{P}^a(X_{li}) \parallel \mathbb{P}^b(X_{li})) &= \mathbb{P}^a(X_{li} = *) \cdot \log \left( \frac{\mathbb{P}^a(X_{li} = *)}{\mathbb{P}^b(X_{li} = *)} \right) + \mathbb{P}^a(X_{li} = 1) \cdot \log \left( \frac{\mathbb{P}^a(X_{li} = 1)}{\mathbb{P}^b(X_{li} = 1)} \right) \\ &\quad + \mathbb{P}^a(X_{li} = 0) \cdot \log \left( \frac{\mathbb{P}^a(X_{li} = 0)}{\mathbb{P}^b(X_{li} = 0)} \right) \\ &= (1-p) \cdot 0 + p\mathbb{P}_i^a \cdot \log \left( \frac{p\mathbb{P}_i^a}{p\mathbb{P}_i^b} \right) + p(1-\mathbb{P}_i^a) \cdot \log \left( \frac{p(1-\mathbb{P}_i^a)}{p(1-\mathbb{P}_i^b)} \right) \\ &= p \cdot \left( \mathbb{P}_i^a \cdot \log \left( \frac{\mathbb{P}_i^a}{\mathbb{P}_i^b} \right) + (1-\mathbb{P}_i^a) \cdot \log \left( \frac{(1-\mathbb{P}_i^a)}{(1-\mathbb{P}_i^b)} \right) \right) \\ &= p \mathbb{D}_{\text{KL}}(\mathbb{P}^a \parallel \mathbb{P}^b). \end{aligned}$$

The second inequality comes from applying inverse Pinsker's inequality (for binary random variable) and the observation that

$$\begin{aligned} \min_{X_{li} \in \{0,1\}} \mathbb{P}^b(X_{li}) &= \min \left\{ \frac{\sqrt{e}}{\sqrt{e} + e^{\beta_i^b}}, \frac{e^{\beta_i^b}}{\sqrt{e} + e^{\beta_i^b}} \right\} \\ &\geq \min \left\{ \frac{\sqrt{e}}{\sqrt{e} + e^{\beta_{\max}^*}}, \frac{e^{\beta_{\min}^*}}{\sqrt{e} + e^{\beta_{\min}^*}} \right\} =: c^*. \end{aligned}$$

The third inequality comes from

$$\begin{aligned}
(\mathbb{P}_i^a - \mathbb{P}_i^b)^2 &= \left( \frac{e^{1/2}}{e^{1/2} + e^{\beta_i^a}} - \frac{e^{1/2}}{e^{1/2} + e^{\beta_i^b}} \right)^2 \\
&= \left( \frac{\sqrt{e} \cdot (e^{\beta_i^a} - e^{\beta_i^b})}{(\sqrt{e} + e^{\beta_i^a})(\sqrt{e} + e^{\beta_i^b})} \right)^2 \\
&\leq \left( \frac{\sqrt{e}(e^{\delta/\sqrt{m}} - 1)}{(\sqrt{e} + e^{\delta/\sqrt{m}})(\sqrt{e} + 1)} \right)^2 \\
&\leq \left( \frac{e\delta}{2\sqrt{m}} \right)^2 = \frac{e^2\delta^2}{4m}.
\end{aligned}$$

The second last inequality comes from the observation that  $\beta_i^a, \beta_i^b \in \{0, \frac{\delta}{\sqrt{m}}\}$ . To see why the last inequality holds, one can verify the following useful inequality.

$$\frac{e^x - 1}{(\sqrt{e} + e^x)(\sqrt{e} + 1)} \leq \frac{x}{2e} \quad \forall x > 0.$$

■

Finally, the proof for Theorem 3.3 follows from Lemma 2.4 and Lemma 2.2.

*Proof. (Of Theorem 3.3)* We can now apply Lemma 2.1. The minimax risk is lower bounded as

$$\inf_{\hat{\beta}} \sup_{\beta^* \in \mathcal{B}} \mathbb{E}[\|\hat{\beta} - \beta^*\|_2^2] \geq \frac{\delta^2}{2} \left( 1 - \frac{Cnp\delta^2 + \log 2}{m} \right),$$

where  $C$  is  $\frac{e^2 \log e}{2c^*}$ . Setting  $\delta^2 = \frac{m}{2Cnp}$  gives us

$$\inf_{\hat{\beta}} \sup_{\beta^* \in \mathcal{B}} \mathbb{E}[\|\hat{\beta} - \beta^*\|_2^2] \geq \frac{m}{4Cnp} \cdot \left( 1 - \frac{m/2 + \log 2}{m} \right) \gtrsim \frac{m}{np}.$$

This completes the proof. ■

## 2.2 Top- $K$ Recovery Lower Bound

We first restate a different version of Fano's inequality Thomas and Joy (2006) which will be useful to the construction of our lower bound. Note that this Fano's inequality is designed for hypothesis testing questions which is different from the pairwise Fano's inequality used to prove the lower bound for parameter estimation. An algorithm does not have to produce parameter estimates in order to produce top- $K$  estimate. Therefore, we need to develop a separate theorem proving the lower bound for top- $K$  recovery.

**Lemma 2.5** (Fano's inequality). *Consider a set of  $L$  distributions  $\{\mathbb{P}^1, \dots, \mathbb{P}^L\}$ . Suppose that we observe a random variable (or a set of random variables)  $Y$  that was generated by first picking an index  $A \in \{1, \dots, L\}$  uniformly at random and then  $Y \sim \mathbb{P}^A$ . Fano's inequality states that any hypothesis test  $\phi$  for this problem has an error probability lower bounded as*

$$\mathbb{P}[\phi(Y) \neq A] \geq 1 - \frac{\max_{a,b \in [L], a \neq b} \mathbb{D}_{KL}(\mathbb{P}^a(Y) \parallel \mathbb{P}^b(Y)) + \log 2}{\log L}.$$

The intuition behind the above version of Fano's inequality is the same as that for the version used in proving the estimation error lower bound. Suppose that we can construct a set of models where the distributions parametrized by any pair of models are similar (as measured in terms of KL divergence). Then no statistical estimator can accurately identify a single model that has been uniformly chosen from this set.

**Construction of the models.** Let us consider the following constructions for  $m - K + 1$  models. For simplicity, let us consider the unnormalized parameter space for now. This is valid because we are only concerned about the KL divergence between any pair of models, which is not affected by the normalization step and only by the relative difference among the items. For model  $a \in [K, K + 1, \dots, m]$ ,

$$\beta_i^{*a} = \begin{cases} \delta & \text{if } i \leq K - 1 \\ \delta & \text{if } i - K = a \\ 0 & \text{otherwise} \end{cases}$$

In other words, the  $n - K + 1$  models differ exactly by the identity of the  $K$ -th best item. It is also clear that  $\Delta_K = \delta$  in this context. For top- $K$  recovery lower bound, it also suffices to consider the case where  $\theta_1 = \dots = \theta_n = \frac{1}{2}$ .

**Lemma 2.6.** *Consider the model construction described above. There exists a constant  $c$  such that for any two models  $a \neq b$ ,*

$$\mathbb{D}_{KL}(\mathbb{P}^a \| \mathbb{P}^b) \leq c np \Delta_K^2.$$

*Proof.* We follow the same argument as used in the proof of Lemma 2.4. We have

$$\begin{aligned} \mathbb{D}_{KL}(\mathbb{P}^a(X) \| \mathbb{P}^b(X)) &= \sum_{l=1}^n \sum_{i=1}^m \mathbb{D}_{KL}(\mathbb{P}^a(X_{li}) \| \mathbb{P}^b(X_{li})) \\ &= np \sum_{i=1}^m \mathbb{D}_{KL}(\mathbb{P}^a \| \mathbb{P}^b) \\ &\quad [\text{Using the inverse Pinsker's inequality}] \\ &\leq np \sum_{i=1}^m \frac{\log e}{c^*} \cdot 2(\mathbb{P}_i^a - \mathbb{P}_i^b)^2 \\ &\leq \frac{np 2 \log e}{c^*} \cdot \sum_{i=1}^m \frac{e^2 \delta^2}{4m} \\ &\leq \frac{e^2 \log e}{2c^*} \cdot np \delta^2, \end{aligned}$$

where  $c^* = \min \left\{ \frac{\sqrt{e}}{\sqrt{e} + e^{\beta_{\max}^*}}, \frac{e^{\beta_{\min}^*}}{\sqrt{e} + e^{\beta_{\min}^*}} \right\}$  and the constant  $c$  in the lemma statement is  $\frac{e^2 \log e}{2c^*}$ . ■

**Theorem 3.4.** *Consider the sampling model described in Section 2. There is a class of Rasch model such that if  $np \leq \frac{c_K \log m}{\Delta_K^2}$ , where  $c_K$  is a constant, then any estimator will fail to identify all of the top  $K$  items with probability at least  $\frac{1}{2}$ .*

*Proof.* Applying Lemma 2.5 and Lemma 2.6, we have

$$\begin{aligned} \mathbb{P}[\phi(Y) \neq A] &\geq 1 - \frac{\max_{a,b \in [L], a \neq b} \mathbb{D}_{KL}(\mathbb{P}^a(Y) \| \mathbb{P}^b(Y)) + \log 2}{\log L} \\ &\geq 1 - \frac{\frac{4np\delta^2}{c} + \log 2}{\log(m - K + 1)} \\ &\approx 1 - \frac{4np\delta^2}{c \log m}, \end{aligned}$$

where  $c = \frac{e^2 \log e}{2c^*}$ . The above lower bound tells us that if

$$np \leq \frac{c \log m}{8 \Delta_K^2}$$

then  $\mathbb{P}[\phi(Y) \neq A] \geq \frac{1}{2}$ . This finishes the proof. ■

### 3 Differentially Private Spectral Algorithm

The following lemma establishes the desirable property of concentrated differential privacy that both post-processing and composition preserve concentrated differential-privacy. Intuitively, one can see that the construction of the private pairwise differential measurement in Equation (4) is analogous to a data-preprocessing step. By ensuring that this step is private, the final output of the algorithm remains private as well. An analogous result for DP can be found in Dwork et al. (2014).

**Lemma 3.1** (Lemma 2.3 of Bun and Steinke (2016)). *Consider domains  $\mathcal{X}$ ,  $\mathcal{Y}$ ,  $\mathcal{Z}$  and two functions  $f : \mathcal{X} \rightarrow \mathcal{Y}$ ,  $f' : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$ . Suppose that  $f$  satisfies  $\xi$ -concentrated differential privacy and  $f'$  satisfies  $\xi'$ -concentrated differential privacy. Define  $f'' : \mathcal{X} \rightarrow \mathcal{Z}$  as  $f''(x) = f'(x, f(x))$ . Then  $f''$  satisfies  $(\xi + \xi')$ -concentrated differential privacy.*

The following lemma establishes the privacy guarantee of the discrete Gaussian mechanism.

**Lemma 3.2** (Theorem 4 of Canonne et al. (2020)). *Let  $\Delta, \epsilon > 0$ . For an input domain  $\mathcal{X}$ , let  $f : \mathcal{X} \rightarrow \mathbb{Z}$  satisfy  $|f(x) - f(x')| \leq \Delta$  for all  $x \simeq x'$  differing on a single entry. Define a randomized algorithm  $M : \mathcal{X} \rightarrow \mathbb{Z}$  as  $M(x) = q(x) + Y$  where  $Y \sim \mathcal{N}_{\mathbb{Z}}(0, \frac{\Delta^2}{\epsilon^2})$ . Then  $M$  satisfies  $\frac{1}{2}\epsilon^2$ -concentrated differential privacy.*

The quantity  $\Delta$  in the theorem statement is often referred to as the sensitivity of a query  $f$ . Simply put, it is the maximum change in output of a query  $f$  when one changes the data of a single user. With the above two Lemmas, we are ready to prove the privacy guarantee of the private spectral algorithm. This quantity will reappear in the analysis of the discrete Laplace mechanism.

**Lemma 4.1.** *Consider a modified spectral algorithm that first adds  $\mathcal{N}_{\mathbb{Z}}\left(0, \frac{\sqrt{m(m-1)}}{\epsilon^{*2}}\right)$ -distributed noise to each pairwise differential measurement per Equation (4). This algorithm then proceeds to use the private differential measurements in the same way as in the original algorithm. This modified spectral algorithm satisfies  $\frac{1}{2}\epsilon^{*2}$ -concentrated differential privacy.*

*Proof.* Recall the construction of the pairwise differential measurement for a pair of items  $i, j$  as

$$Y_{ij}(X) = \sum_{l=1}^n A_{li} A_{lj} X_{li} (1 - X_{lj}). \quad (3)$$

It is clear that for any neighboring datasets  $X \simeq X'$ ,  $|Y_{ij}(X) - Y_{ij}(X')| \leq 1$  (sensitivity). Invoking Lemma 3.2, we conclude that the privatized function

$$\tilde{Y}_{ij}(X) = Y_{ij}(X) + Z_{ij},$$

where  $Z_{ij} \sim \mathcal{N}_{\mathbb{Z}}(0, \frac{1}{\epsilon^2})$  satisfies  $\frac{1}{2}\epsilon^2$ -concentrated differential privacy. Invoking lemma 3.1 with  $\Delta = 1$ , we see that the privatized spectral algorithm that constructs  $m(m-1)$  pairwise differential measurements where each non-private measurement is added with  $\mathcal{N}_{\mathbb{Z}}(0, \frac{1}{\epsilon^2})$ -distributed noise satisfies  $\frac{m(m-1)}{2}\epsilon^2$ -concentrated differential privacy.

Using simple algebra, one can check that for  $\epsilon = \frac{\epsilon^*}{\sqrt{m(m-1)}}$ ,  $m(m-1)$  compositions of  $\frac{1}{2}\epsilon^2$ -concentrated differentially private queries (using  $\mathcal{N}_{\mathbb{Z}}\left(0, \frac{\sqrt{m(m-1)}}{\epsilon^2}\right)$ -distributed noise for each query) satisfies  $\frac{1}{2}\epsilon^{*2}$ -concentrated differential privacy overall. Now, since concentrated differential privacy is closed under post-processing, the entire private spectral algorithm also satisfies the same privacy guarantee as the private pairwise measurements computation step, which is  $\frac{1}{2}\epsilon^{*2}$ -concentrated differentially private. ■

**Entrywise Error Analysis of the Private Spectral Algorithm.** During the submission process of this paper, an anonymous reviewer has kindly pointed out that our theoretical analysis would be more complete with a study of the entrywise error guarantee of the spectral algorithm. It is indeed possible to obtain some

rough analysis as follows. Firstly, note that the spectral algorithm constructs a Markov chain. Per our analysis, the statistical noise for the pairwise transition probabilities can be bounded as

$$|M_{ij} - M_{ij}^*| = O\left(\frac{\sqrt{\log m}}{m\sqrt{np^2}}\right) \quad \forall i \neq j \in [m],$$

where  $M_{ij}$  is the non-private estimate and  $M_{ij}^*$  is the idealized estimate. On the other hand, the discrete Gaussian distribution ‘behaves’ similarly to a continuous Gaussian distribution Canonne et al. (2020) (i.e., light-tailed). That is, with high probability, the amount of noise introduced by the Gaussian mechanism to the pairwise transition probabilities (cf. Equation (4)) is bounded as

$$\left|\frac{Z_{ij}}{d}\right| = O\left(\frac{\sqrt{\log m}}{np^2\epsilon^*}\right).$$

Hence, the deviation of the *private* pairwise transition probabilities from the idealized pairwise transition probabilities can be bounded as

$$|\tilde{M}_{ij} - M_{ij}^*| = O\left(\frac{\sqrt{\log m}}{m\sqrt{np^2}} + \frac{\sqrt{\log m}}{np^2\epsilon^*}\right).$$

One can see that given sufficiently large sample size,  $np^2 = \Omega\left(\frac{m^2}{\epsilon^{*2}}\right)$ , the additive noise introduced by the Gaussian mechanism is on the same order as or dominated by the statistical noise and we obtain the same error guarantees as the non-private spectral algorithm (up to a constant factor). Formalizing this argument, one obtains the following entrywise error guarantee of the private spectral algorithm.

**Theorem 3.3.** *Consider the uniform sampling model described in Section 2. There exist constants  $C_1, C_2$  such that if  $np^2 \geq C_1 \left\{\log m, \frac{m^2}{(\epsilon^*)^2}\right\}$  and  $mp \geq C_2 \log n$  then the output of the spectral algorithm satisfies*

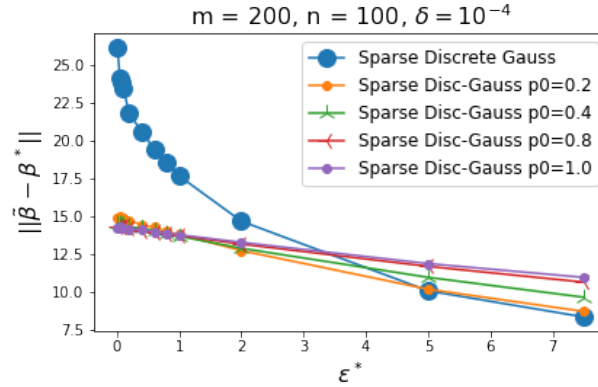
$$\|\beta - \beta^*\|_\infty \leq \frac{C' \sqrt{\log m}}{\sqrt{np}}$$

*with probability at least  $1 - \mathcal{O}(m^{-10}) - \mathcal{O}(n^{-10})$  where  $C'$  is an absolute constant.*

*Proof.* The proof is relatively simple and is thus omitted. ■

## 4 Additional Experimental Details

**Additional Experiments.** During the submission process, an anonymous reviewer has kindly suggested that we can perform additional experiments examining other aspects of the algorithm such as time complexity and how the performance varies with  $p_0$  (the sampling probability when applying the sparse Gaussian mechanism). Firstly, in terms of time complexity, both randomized response and the Gaussian mechanism are essentially data processing algorithms and both very fast compared to underlying non-private algorithm. The time complexity for the Gaussian mechanism is  $O(m^2)$  while that for randomized response is  $O(mn)$  where  $m$  is the number of items and  $n$  the number of users. In many real life data sets,  $n \gg m$ , the Gaussian mechanism is a more efficient procedure. The effect of  $p_0$  on the performance of the Gaussian mechanism depends on the privacy regime. The figure below shows how the performance of the sparse Discrete Gaussian mechanism varies with  $p_0$ . Note that the line in blue is the proposed sparse method in our paper with  $p_0 = \frac{\log n}{n}$ . While



**Figure 1:** In the high-privacy regime (small  $\epsilon^*$ ), a larger  $p_0$  generally leads to better performance. In the low-privacy regime (larg  $\epsilon^*$ ), using a sparse sampling graph results in more accurate estimates. In our work, we propose  $p_0 = \frac{\log n}{n}$  as a basic starting point for designing a sparse Gaussian mechanism that performs decently well.

the experiments in our work focus primarily on privacy mechanisms and parameter estimation under the Rasch model, we can also follow the experimental setup in previous works such as Nguyen and Zhang (2022) to explore private ranking applications. We also like to point out that a commonly studied algorithm in the ranking literature is spectral ranking Negahban et al. (2017) also constructs a Markov chain in its procedure. The application of Gaussian mechanism in our algorithm can be extended to that setting to construct a differentially private algorithm specifically for ranking applications.

**Privacy Guarantee of Randomized Response with Shuffling.** Firstly, we recall the procedure behind randomized response Warner (1965) for binary responses. For each user  $l$  and a response  $X_{li}$  for  $i \in [m]$ . The following procedure ensures  $(\epsilon, 0)$ -differential privacy (cf. Wang et al. (2016)).

1. The participant flips a coin with bias  $\frac{e^\epsilon}{1+e^\epsilon}$  toward heads (outcome unknown to the statistician).
2. If the outcome is heads, the participant answers truthfully. If the outcome is tails, the participant reports the answer opposite to the truth.

We include here our implementation for randomized response.

---

```
import numpy as np

def randomized_response(data, epsilon):
    # With probability e^epsilon/(1+e^epsilon), flip the label
    # For each user, we need to compute the number of responses, then divide epsilon by that number
    num_responses = np.sum(data != INVALID_RESPONSE, 0)
    m, n = data.shape
```

```

privatized_data = np.copy(data)

for l in range(n):
    effective_epsilon = epsilon / num_responses[l]
    # The epsilon we should be using for each user to flip their responses
    rand_p = np.random.rand(m) # Flip the binary response with probability 1/(1+e^epsilon)
    privatized_data[:, l] = np.where(np.logical_and(rand_p < 1/(1+np.exp(effective_epsilon)), \
        data[:, l] != INVALID_RESPONSE), 1 - data[:, l], data[:, l])

return privatized_data

```

---

Recall that the privacy guarantee of DP decays linearly with the number of queries,  $m$  in our application. Hence, to ensure a  $\epsilon^*$  privacy level overall for each user, each response  $X_{li}$  by the same user  $l$  needs to be randomized using the above procedure with  $\epsilon = \epsilon^*/m$ . Randomized response satisfies a special notion of DP known as local DP (Kasiviswanathan et al., 2011). It is also well known that local DP implies global DP. Furthermore, it has been shown recently that one can further amplify the privacy guarantee of local DP mechanisms such as randomized response simply by permuting the randomized user data. In our application, since we are concerned primarily with estimating the item parameters, the algorithm is completely agnostic of this shuffling procedure.

**Lemma 4.1** (Theorem III.1 of Feldman et al. (2022)). *For any domain  $\mathcal{D}$ , let  $\mathbb{R}^{(i)} : \mathcal{S}^{(1)} \times \dots \times \mathcal{S}^{(i-1)} \times \mathcal{D} \rightarrow \mathcal{S}^{(i)}$  for  $i \in [n]$  be a sequence of algorithms such that  $\mathbb{R}^{(i)}(z_{1:i-1}, \cdot)$  is an  $\epsilon_0$ -DP local randomizer for all values of auxiliary inputs  $z_{1:i-1} \in \mathcal{S}^{(1)} \times \dots \times \mathcal{S}^{(i-1)}$ . Let  $\mathcal{A}_s : \mathcal{D}^n \rightarrow \mathcal{S}^{(1)} \times \dots \times \mathcal{S}^{(n)}$  be the algorithm that given a dataset  $x_{1:n} \in \mathcal{D}^n$ , samples a uniform random permutation  $\sigma$  over  $[n]$ , then sequentially computes  $z_i = \mathbb{R}^{(i)}(z_{1:i-1}, x_{\sigma(i)})$  for  $i \in [n]$  and outputs  $z_{1:n}$ . Then for any  $\delta \in [0, 1]$  such that  $\epsilon_0 \leq \log \left( \frac{n}{16 \log(2/\delta)} \right)$ ,  $\mathcal{A}_s$  is  $(\epsilon, \delta)$ -DP, where*

$$\epsilon \leq \log \left( 1 + \frac{e^{\epsilon_0} - 1}{e^{\epsilon_0} + 1} \left( \frac{8\sqrt{e^{\epsilon_0} \log(4/\delta)}}{\sqrt{n}} \right) \right).$$

Since the bound on  $\epsilon$  in the theorem statement is a monotonous function in  $\epsilon_0$ , we can find  $\epsilon_0$  using a simple bisection search algorithm. We include the code for computing  $\epsilon_0$  below.

---

```

def find_effective_epsilon0_rr(overall_epsilon, overall_delta, n):
    # epsilon0 is the maximum epsilon0 such that
    # e^eps = 1 + (e^eps0 - 1)/(e^eps0 + 1) (8 * sqrt(e^eps0 * log(4/delta)))/sqrt(n) + 8e^eps/n
    # Solve for eps0

    def eps_estimate(eps0):
        eeps0 = np.exp(eps0)
        temp1 = (8 * np.sqrt(eeps0 * np.log(4/overall_delta)))/(np.sqrt(n))
        temp2 = 8 * eeps0/n
        return np.log(1 + (eeps0-1)/(eeps0+1) * (temp1 + temp2))

    # Just do bi-section search
    lower = 0.000001
    upper = np.log(n/(16*np.log(2/overall_delta)))

    if upper < 0:
        return overall_epsilon

    while eps_estimate(lower) > overall_epsilon:
        lower /= 2

```



---

```

while upper - lower > 1e-8:
    mid = (upper + lower)/2
    if eps_estimate(mid) < overall_epsilon:
        lower = mid
    else:
        upper = mid

return mid

```

---

**The discrete Laplace Mechanism.** The discrete Laplace mechanism is another noise-adding mechanism to ensure privacy that we consider in this work. Firstly, we restate the definition of the discrete Laplace mechanism (Ghosh et al., 2009).

**Definition 1.** Let  $t > 0$ . The discrete Laplace distribution with scale parameter  $t$  is denoted  $Lap_{\mathbb{Z}}(t)$ . It is a probability distribution supported on the integers and defined by

$$\mathbb{P}_{X \leftarrow Lap_{\mathbb{Z}}(t)}[X = x] = \frac{e^{1/t} - 1}{e^{1/t} + 1} \cdot e^{-|x|/t}$$

for all  $x \in \mathbb{Z}$ .

The Laplace mechanism adds discrete noise to a non-private query, with variance proportional to the sensitivity of the query.

**Lemma 4.2** (Theorem 1 of Ghosh et al. (2009)). Let  $\Delta, \epsilon > 0$ . For an input domain  $\mathcal{X}$ , let  $f : \mathcal{X} \rightarrow \mathbb{Z}$  satisfy  $|f(x) - f(x')| \leq \Delta$  for all  $x \simeq x'$  differing on a single entry. Define a randomized algorithm  $M : \mathcal{X} \rightarrow \mathbb{Z}$  as  $M(x) = q(x) + Y$  where  $Y \sim Lap_{\mathbb{Z}}(0, \frac{\Delta}{\epsilon})$ . Then  $M$  satisfies  $\epsilon$ -differential privacy.

It is possible to draw the discrete Laplace distribution using the following code.

---

```

import numpy as np

def laplace_noise(eps):
    def zero_prob(alpha):
        return 0 if np.random.rand() < (1.0 - alpha)/(1.0 + alpha) else 1

    def sign_prob():
        return -1 if np.random.rand() < 0.5 else 1

    alpha = np.exp(-eps) # Between 0 and 1
    # This implementation is based on the usual geometric distribution
    # With probability (1-alpha)/(1+alpha), return 0
    # Otherwise, draw from the geometric distribution but with random sign flipping

    return zero_prob(alpha) * np.random.geometric(1-alpha) * sign_prob()

```

---

**Drawing from the discrete Gaussian distribution.** The following code can be used to draw from the discrete Gaussian distribution. This is based on Algorithm 1 in Canonne et al. (2020).

---

```

import numpy as np

def exponential_bernoulli_sample(gamma):
    assert(gamma >= 0)
    if 0 <= gamma <= 1:

```

```

K = 1
while(1):
    A = 1 if np.random.rand() < gamma/K else 0
    if A == 0:
        break
    else:
        K += 1
    return 1 if K % 2 == 1 else 0
else:
    for k in range(1, int(np.floor(gamma) + 1)):
        B = exponential_bernoulli_sample(1)
        if B == 0:
            return 0
    return exponential_bernoulli_sample(gamma - np.floor(gamma))

def discrete_gaussian_sample(sigma):
    t = int(np.floor(sigma) + 1)

    while(1):
        U = np.random.choice(t)
        D = exponential_bernoulli_sample(U/t)
        if D == 0:
            # Restart
            return discrete_gaussian_sample(sigma)

    V = 0
    while(1):
        A = exponential_bernoulli_sample(1)
        if A == 0:
            break
        V += 1

    B = 1 if np.random.rand() < 1./2 else 0
    if B == 1 and U == 0 and V == 0:
        return discrete_gaussian_sample(sigma)
    Z = (1 - 2*B) * (U + t * V)
    C = exponential_bernoulli_sample((np.abs(Z)-sigma**2/t)**2/(2*sigma**2))

    if C == 0:
        return discrete_gaussian_sample(sigma)
    return Z

```

---

**The private spectral algorithm.** We now have all the necessary components to construct the private spectral algorithm. We include the implementation of the private spectral algorithm including the function to subsample for a random graph in the construction of the sparse Markov chain. For the code to run and produce our experiment results, refer to the attached python script in the supplementary materials.

---

```

def subsampl_graph(m, p):
    subsample_graph = np.zeros((m, m))

    for i in range(m-1):
        for j in range(i+1, m):
            if np.random.rand() < p:

```

```

        subsample_graph[i, j] = 1
        subsample_graph[j, i] = 1

    return subsample_graph

def construct_markov_chain_private(performances, lambd=1., epsilon=0.1, subsample_graph=None):
    m = len(performances)

    if subsample_graph is None:
        subsample_graph = np.ones((m, m))

    D = np.ma.masked_where(performances == INVALID_RESPONSE, performances)
    D_compl = 1. - D
    M = np.ma.dot(D, D_compl.T)

    A = np.ma.masked_where(performances == INVALID_RESPONSE, np.ones_like(performances))
    B = np.ma.dot(A, A.T)

    np.fill_diagonal(M, 0)
    np.nan_to_num(M, False)
    M = np.round(M)

    M_non_priv = np.copy(M)
    M_add = np.copy(M)

    # Add discrete Gaussian noise
    for i in range(m):
        for j in range(m):
            if j != i and M[i, j] != 0 and np.abs(subsample_graph[i, j] - 1) < 1e-6:
                noise = discrete_gaussian_sample((1./epsilon))
                M[i, j] = max(1, M[i, j] + noise)
                M_add[i, j] = noise

            if j != i and np.abs(subsample_graph[i, j] - 0) < 1e-6:
                M_non_priv[i, j] = 0
                M[i, j] = 0
                M_add[i, j] = 0

    # Add regularization to the 'missing' entries
    M = np.where(np.logical_or((M != 0), (M.T != 0)), M+lambd, M)

    # d = np.ma.sum(M, 1) + 1
    d = np.ones((m,)) * np.ma.max(np.ma.sum(M, 1) + 1)

    for i in range(m):
        di = d[i]
        M[i, :] /= di
        M[i, i] = 1. - np.sum(M[i, :])

    return M_non_priv, M_add, M, d

```

```

def spectral_estimate_private(A, lamdb=1., epsilon=0.1, subsample_graph=None,
                             max_iters=1000, eps=1e-5):
    _, _, M, d = construct_markov_chain_private(A, lamdb=lamdb, epsilon=epsilon,
                                                subsample_graph=subsample_graph)

    assert(not np.any(np.isnan(M)))
    from scipy.sparse import csc_matrix

    M = csc_matrix(M)
    m = len(A)

    pi = np.ones((m,)).T
    for _ in range(max_iters):
        pi_next = (pi @ M)
        pi_next /= np.sum(pi_next)
        if np.linalg.norm(pi_next - pi) < eps:
            pi = pi_next
            break
    pi = pi_next

    pi = pi.T
    # pi = np.maximum(pi, 1e-12)
    pi /= np.sum(pi)
    pi = (pi/d)/np.sum(pi/d)
    assert(not np.any(np.isnan(pi)))
    beta = np.log(pi)
    beta = beta - np.mean(beta)
    assert(not np.any(np.isnan(beta)))
    return beta

```

---

## Converting Concentrated Differential Privacy to Approximate Differential Privacy.

**Lemma 4.3** (Lemma 3.5 of Bun and Steinke (2016)). *Let  $M$  be a  $\rho$ -concentrated differentially private algorithm. Then  $M$  also satisfies  $(\epsilon, \delta)$ -approximate differential privacy for all  $\delta > 0$  and*

$$\epsilon = \rho + \sqrt{4\rho \log(1/\delta)}.$$

Generally, to achieve a desired  $(\epsilon, \delta)$ -differential privacy guarantee, it suffices to satisfy  $\rho$ -CDP with

$$\rho \approx \frac{\epsilon^2}{4 \log(1/\delta)}.$$

Recall that the discrete Gaussian mechanism that adds  $\mathcal{N}_{\mathbb{Z}}\left(0, \frac{1}{\epsilon_0^2}\right)$  noise to a non-private query satisfies  $\frac{1}{2}\epsilon_0^2$ -CDP. By setting  $\frac{1}{2}\epsilon_0^2$  to the approximation in the theorem statement above, we can solve for  $\epsilon_0$  using the following one-line code.

```

import numpy as np

def find_effective_epsilon0_zgauss(overall_epsilon, overall_delta):
    return overall_epsilon/np.sqrt(2 * np.log(1./overall_delta))

```

---

## References

Andersen, E. B. (1973). Conditional inference for multiple-choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 26(1):31–44.

- Bun, M. and Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658. Springer.
- Canonne, C. L., Kamath, G., and Steinke, T. (2020). The discrete gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33:15676–15688.
- Chen, P., Gao, C., and Zhang, A. Y. (2020). Partial recovery for top- $k$  ranking: Optimality of mle and sub-optimality of spectral method. *arXiv preprint arXiv:2006.16485*.
- Dwork, C., Roth, A., et al. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.
- Feldman, V., McMillan, A., and Talwar, K. (2022). Hiding among the clones: A simple and nearly optimal analysis of privacy amplification by shuffling. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 954–964. IEEE.
- Ghosh, A., Roughgarden, T., and Sundararajan, M. (2009). Universally utility-maximizing privacy mechanisms. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 351–360.
- Ghosh, M. (1995). Inconsistent maximum likelihood estimators for the rasch model. *Statistics & Probability Letters*, 23(2):165–170.
- Kasiviswanathan, S. P., Lee, H. K., Nissim, K., Raskhodnikova, S., and Smith, A. (2011). What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826.
- Negahban, S., Oh, S., and Shah, D. (2017). Rank centrality: Ranking from pairwise comparisons. *Operations Research*, 65(1):266–287.
- Nguyen, D. and Zhang, A. (2022). A spectral approach to item response theory. *Advances in neural information processing systems*, 35.
- Shah, N., Balakrishnan, S., Bradley, J., Parekh, A., Ramchandran, K., and Wainwright, M. (2015). Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. In *Artificial intelligence and statistics*, pages 856–865. PMLR.
- Thomas, M. and Joy, A. T. (2006). *Elements of information theory*. Wiley-Interscience.
- Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press.
- Wang, Y., Wu, X., and Hu, D. (2016). Using randomized response for differential privacy preserving data collection. In *EDBT/ICDT Workshops*, volume 1558, pages 0090–6778.
- Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69.