Estimating Sparse Covariance Matrix Under Differential Privacy via Thresholding

Di Wang*, Jinhui Xu*, Yang He†
*Deparment of Computer Science and Engineering
State University of New York at Buffalo, Buffalo, USA, 14260
Email: dwang45@buffalo.edu, jinhui@buffalo.edu
†Department of Economics
State University of New York at Buffalo, Buffalo, USA, 14260
Email: yhe6@buffalo.edu

Abstract—**In this paper, we study the problem of estimating the covariance matrix under differential privacy, where the underlying covariance matrix is assumed to be sparse and of high dimensions. We propose a new method, called DP-Thresholding, to achieve a non-trivial ℓ_2-norm based error bound, which is significantly better than the existing ones from adding noise directly to the empirical covariance matrix. Experiments on the synthetic datasets show consistent results with our theoretical claims.**

I. Introduction

Machine Learning and Statistical Estimation have made profound impact in recent years to many applied domains such as social sciences, genomics, and medicine. During their applications, a frequently encountered challenge is how to deal with the high dimensionality of the datasets, especially for those in genomics, educational and psychological research. A commonly adopted strategy for dealing with such an issue is to assume that the underlying structures of parameters are sparse.

Another often encountered challenge is how to handle sensitive data, such as those in social science, biomedicine and genomics. A promising approach is to use some differentially private mechanisms for the statistical inference and learning tasks. Differential Privacy (DP) [1] is a widely-accepted criterion that provides provable protection against identification and is resilient to arbitrary auxiliary information that might be available to attackers. Since its introduction over a decade ago, a rich line of works are now available [2]–[5], which have made differential privacy a compelling privacy enhancing technology for many organizations, such as Google [6], Apple [7].

Estimating or studying the high dimensional datasets while keeping them (locally) differentially private could be quite challenging for many problems, such as sparse linear regression [8]. However, there are also evidences showing that the loss of some problems under the privacy constraints can be quite small compared with their non-private counterparts. Examples of such nature include high dimensional sparse PCA [9], sparse inverse covariance estimation [10], and high-dimensional distributions estimation [11]. Thus, it is desirable to determine which high dimensional problem can be learned or estimated efficiently in a private manner.

In this paper, we try to give an answer to this question for a simple but fundamental problem in machine learning and statistics, called estimating the underlying sparse covariance matrix of bounded sub-Gaussian distribution. For this problem, we propose a simple but nontrivial (ϵ, δ)-DP method, DP-Thresholding, and show that the squared ℓ_2-norm error is bounded by $O(\frac{s \log p}{n\epsilon^2})$, where s is the sparsity of each row in the underlying covariance matrix. Experiments on synthetic datasets confirm the theoretical claims. To our best knowledge, this is the first paper studying the problem of estimating high dimensional sparse covariance matrix under (local) differential privacy.

II. Related Work

Recently, there are several papers studying private distribution estimation, such as [11]–[15]. For distribution estimation under the central differential privacy model, [13] considers the 1-dimensional private mean estimation of a Gaussian distribution with (un)known variance. The work that is probably most related to ours is [11], which studies the problem of privately learning a multivariate Gaussian and product distributions. The following are the main differences with ours. Firstly, our goal is to estimate the covariance of a sub-Gaussian distribution. Even though the class of distributions considered in our paper is larger than the one in [11], it has an additional assumption which requires the ℓ_2 norm of a sample of the distribution to be bounded by 1. This means that it does not include the general Gaussian distribution. Secondly, although [11] also considers the high dimensional case, it does not assume the sparsity of the underlying covariance matrix. Thus, its error bound depends on the dimensionality p polynomially, which is large in the high dimensional case ($p \gg n$), while the dependence in our paper is only logarithmically (*i.e.,* $\log p$). Thirdly, the error in [11] is measured by the total variation distance, while it is by ℓ_2-norm in our paper. Thus, the two results are not comparable. [15] recently also studies the covariance matrix estimation via iterative eigenvector sampling. However, their method is just for the low dimensional case and with Frobenious norm as the error measure.

In this paper, we mainly use Gaussian mechanism to the covariance matrix, which has been studied in [9], [10], [16].

However, as it will be shown later, simply outputting the perturbed covariance can cause big error and thus is insufficient for our problem. Compared to these problems, ours is clearly more complicated.

III. Preliminaries

A. Differential Privacy

Differential privacy [1] is by now a defacto standard for statistical data privacy which constitutes a strong standard for privacy guarantees for algorithms on aggregate databases. One likely reason that it gains so much popularity is its guarantee of no significant change on the outcome distribution when there is one entry change to the dataset. We say that two datasets D, D' are neighbors if they differ by only one entry, denoted as $D \sim D'$.

Definition 1 (Differentially Private [1]). A randomized algorithm $\mathcal{A}$ is (ϵ, δ)-differentially private (DP) if for all neighboring datasets D, D' and for all events S in the output space of $\mathcal{A}$, the following holds

$$\mathbb{P}(\mathcal{A}(D) \in S) \leq e^{\epsilon}\mathbb{P}(\mathcal{A}(D') \in S) + \delta.$$

When $\delta = 0$, $\mathcal{A}$ is ϵ-differentially private.

We will use Gaussian Mechanism [1] to guarantee (ϵ, δ)-DP.

Definition 2 (Gaussian Mechanism). Given any function $q : \mathcal{X}^n \to \mathbb{R}^p$, the Gaussian Mechanism is defined as:

$$\mathcal{M}_G(D, q, \epsilon) = q(D) + Y,$$

where Y is drawn from Gaussian Distribution $\mathcal{N}(0, \sigma^2 I_p)$ with $\sigma \geq \frac{\sqrt{2\ln(1.25/\delta)}\Delta_2(q)}{\epsilon}$. Here $\Delta_2(q)$ is the ℓ_2-sensitivity of the function q, i.e.

$$\Delta_2(q) = \sup_{D \sim D'} ||q(D) - q(D')||_2.$$

Gaussian Mechanism preservers (ϵ, δ)-differential privacy.

B. Private Sparse Covariance Estimation

Let $x_1, x_2, \cdots, x_n$ be n random samples from a p-variate distribution with covariance matrix $\Sigma = (\sigma_{ij})_{1\leq i,j\leq p}$, where the dimensionality p is assumed to be high, *i.e.,* $p \gg n \geq \text{Poly}(\log p)$.

We define the parameter space of s-sparse covariance matrices as the following:

$$\mathcal{G}_0(s) = \{\Sigma = (\sigma_{ij})_{1\leq i,j\leq p} : \sigma_{-j,j} \text{ is } s\text{-sparse } \forall j \in [p]\}, \quad (1)$$

where $\sigma_{-j,j}$ means the j-th column of Σ with the entry σ_{jj} removed. That is, a matrix in $\mathcal{G}_0(s)$ has at most s non-zero off-diagonal elements in each column.

We assume that each x_i is sampled from a 0-mean and sub-Gaussian distribution with parameter σ^2, that is,

$$\mathbb{E}[x_i] = 0, \mathbb{P}\{|v^T x_i| > t\} \leq e^{-\frac{t^2}{2\sigma^2}}, \forall t > 0 \text{ and } \|v\|_2 = 1. \quad (2)$$

This means that all the one-dimensional marginals of x_i have sub-Gaussian tails. We also assume that with probability 1, $\|x_i\|_2 \leq 1$. We note that such assumptions are quite common in the differential privacy literature, such as [9].

Let $\mathcal{P}_d(\sigma^2, s)$ denote the set of distributions of x_i satisfying all the above conditions (*ı.e.,* (2) and $\|x_i\|_2 \leq 1$) and with the covariance matrix $\Sigma \in \mathcal{G}_0(s)$. The goal of private covariance estimation is to obtain an estimator Σ^{priv} of the underlying covariance matrix Σ based on $\{x_1, \cdots, x_n\} \sim P \in \mathcal{P}_d(\sigma^2, s)$ while keeping it differnetially private. In this paper, we will focus on the (ϵ, δ)-differential privacy. We use the ℓ_2 norm to measure the difference between Σ^{priv} and Σ, *i.e.,* $\|\Sigma^{\text{priv}} - \Sigma\|_2$.

Lemma 1. Let $\{x_1, \cdots, x_n\}$ be n random variables sampled from the Gaussian distribution $\mathcal{N}(0, \sigma^2)$. Then

$$\mathbb{E} \max_{1\leq i\leq n} |x_i| \leq \sigma\sqrt{2\log 2n}, \quad (3)$$

$$\mathbb{P}\{\max_{1\leq i\leq n} |x_i| \geq t\} \leq 2ne^{-\frac{t^2}{2\sigma^2}}. \quad (4)$$

Particularly, if $n = 1$, we have $\mathbb{P}\{|x_i| \geq t\} \leq 2e^{-\frac{t^2}{2\sigma^2}}$.

Lemma 2 ([17]). If $\{x_1, x_2, \cdots, x_n\}$ are sampled form a sub-Gaussian distribution in (2) and $\Sigma^* = (\sigma^*)_{1\leq i,j\leq p} = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$ is the empirical covariance matrix, then there exist constants C_1 and $\gamma > 0$ such that $\forall i, j \in [p]$

$$\mathbb{P}(|\sigma^*_{ij} - \sigma_{ij}| > t) \leq C_1 e^{-nt^2\frac{8}{\gamma^2}} \quad (5)$$

for all $|t| \leq \delta$, where C_1 and γ are constants and depend only on σ^2. Specifically,

$$\mathbb{P}\{|\sigma^*_{ij} - \sigma_{ij}| > \gamma\sqrt{\frac{\log p}{n}}\} \leq C_1 p^{-8}. \quad (6)$$

IV. Method

A. A First Approach

A direct way to obtain a private estimator is to perturb the empirical covariance matrix by symmetric Gaussian matrices, which has been used in previous work on private PCA, such as [9], [16]. However, as we can see bellow, this method will introduce big error.

By [16], for any give $0 < \epsilon, \delta \leq 1$ and $\{x_1, x_2, \cdots, x_n\} \sim P \in \mathcal{P}_p(\sigma^2, s)$, the following perturbing procedure is (ϵ, δ)-differentially private:

$$\tilde{\Sigma} = \Sigma^* + N = (\tilde{\sigma}_{ij})_{1\leq i,j\leq p} = \frac{1}{n}\sum_{i=1}^n x_i x_i^T + N, \quad (7)$$

where N is a symmetric matrix with its upper triangle (including the diagonal) being i.i.d samples from $\mathcal{N}(0, \sigma_1^2)$; here $\sigma_1^2 = \frac{2\ln(1.25/\delta)}{n^2\epsilon^2}$, and each lower triangle entry is copied from its upper triangle counterpart. By [18], we know that $\|N\|_2 \leq O(\sqrt{p}\sigma_1) = O(\frac{\sqrt{p}\sqrt{\log\frac{1}{\delta}}}{n\epsilon})$. We can easily get that

$$\|\tilde{\Sigma} - \Sigma\|_2 \leq \|\Sigma^* - \Sigma\|_2 + \|N\|_2 \leq O(\frac{\sqrt{p\log\frac{1}{\delta}}}{n\epsilon}), \quad (8)$$

where the second inequality is due to [19]. However, we can see that the upper bound of the error in (8) is quite large in the high dimensional case.

Another issue of the private estimator in (7) is that it is not clear whether it is positive-semidefinite, a property that is normally expected from an estimator.

B. Post-processing via Thresholding

We note that one of the reasons that the private estimator $\tilde{\Sigma}$ in (7) fails is due to the fact that some entries are quite large which make $\|\tilde{\Sigma}_{ij} - \Sigma_{ij}\|_2$ large for some i, j. To see it more precisely, by (4) and (5) we can get the following, with probability at least $1 - Cp^{-6}$, for all $1 \leq i, j \leq p$,

$$|\tilde{\sigma}_{ij} - \sigma_{ij}| \leq \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln\frac{1.25}{\delta}}\sqrt{\log p}}{n\epsilon} = O(\gamma\sqrt{\frac{\log p}{n\epsilon^2}}). \tag{9}$$

Thus, to reduce the error, it is natural to think of the following way. For those σ_{ij} with larger values, we keep the corresponding $\tilde{\sigma}_{ij}$ in order to make their difference less than some threshold. For those σ_{ij} with smaller values compared with (9), since the corresponding $\tilde{\sigma}_{ij}$ may still be large, if we threshold $\tilde{\sigma}_{ij}$ to 0, we can lower the error on $\tilde{\sigma}_{ij} - \sigma_{ij}$.

Following the above thinking and the thresholding methods in [17] and [20], we propose the following DP-Thresholding method, which post-processes the perturbed covariance matrix in (7) with the threshold $\gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}$. After thresholding, we further threshold the eigenvalues of $\hat{\Sigma}$ in order to make it positive semi-definite. See Algorithm 1 for detail.

Algorithm 1 DP-Thresholding

```
Input: $\epsilon, \delta$ are privacy parameters, $\{x_1, x_2, \cdots, x_n\} \sim P \in \mathcal{P}(\sigma^2, s)$.
1: Compute
       $$\tilde{\Sigma} = (\tilde{\sigma}_{ij})_{1\leq i,j\leq p} = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T + N,$$
   where $N$ is a symmetric matrix with its upper triangle (including the diagonal) being i.i.d samples from $\mathcal{N}(0, \sigma_1^2)$; here $\sigma_1^2 = \frac{2\ln(1.25/\delta)}{n^2\epsilon^2}$, and each lower triangle entry is copied from its upper triangle counterpart.
2: Define the thresholding estimator $\hat{\Sigma} = (\hat{\sigma}_{ij})_{1\leq i,j\leq n}$ as
       $$\hat{\sigma}_{ij} = \tilde{\sigma}_{ij} \cdot I[|\tilde{\sigma}_{ij}| > \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}]. \tag{10}$$
3: Let the eigen-decomposition of $\hat{\Sigma}$ as $\hat{\Sigma} = \sum_{i=1}^{p} \lambda_i v_i v_i^T$. Let $\lambda^+ = \max\{\lambda_i, 0\}$ be the positive part of $\lambda_i$, then define $\Sigma^+ = \sum_{i=1}^{p} \lambda^+ v_i v_i^T$.
4: return $\Sigma^+$.
```

Theorem 1. For any $0 < \epsilon, \delta \leq 1$, Algorithm 1 is (ϵ, δ)-differentially private.

Proof. By [9] and [16], we know that Step 1 keeps the matrix (ϵ, δ)-differentially private. Thus, Algorithm 1 is (ϵ, δ)-differentially private due to the post-processing property of differential privacy [1]. □

For the matrix $\hat{\Sigma}$ in (10) after the first step of thresholding, we have the following key lemma.

Lemma 3. For every fixed $1 \leq i, j \leq p$, there exists a constant $C_1 > 0$ such that with probability at least $1 - C_1 p^{-\frac{9}{2}}$, the following holds:

$$|\hat{\sigma}_{ij} - \sigma_{ij}| \leq 4\min\{|\sigma_{ij}|, \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}\}. \tag{11}$$

Proof of Lemma 3. Let $\Sigma^* = (\sigma^*_{ij})_{1\leq i,j\leq p}$ and $N = (n_{ij})_{1\leq i,j\leq p}$. Define the event $A_{ij} = \{|\tilde{\sigma}_{ij}| > \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}\}$. We have:

$$|\hat{\sigma}_{ij} - \sigma_{ij}| = |\sigma_{ij}| \cdot I(A^c_{ij}) + |\tilde{\sigma}_{ij} - \sigma_{ij}| \cdot I(A_{ij}). \tag{12}$$

By the triangle inequality, it is easy to see that

$$\begin{aligned} A_{ij} &= \{|\tilde{\sigma}_{ij} - \sigma_{ij} + \sigma_{ij}| > \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}\} \\ &\subset \{|\tilde{\sigma}_{ij} - \sigma_{ij}| > \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon} - |\sigma_{ij}|\} \end{aligned}$$

and

$$\begin{aligned} A^c_{ij} &= \{|\tilde{\sigma}_{ij} - \sigma_{ij} + \sigma_{ij}| \leq \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}\} \\ &\subset \{|\tilde{\sigma}_{ij} - \sigma_{ij}| > |\sigma_{ij}| - (\gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon})\}. \end{aligned}$$

Depending on the value of σ_{ij}, we have the following three cases.

a) **Case 1:** $|\sigma_{ij}| \leq \frac{\gamma}{4}\sqrt{\frac{\log p}{n}} + \frac{\sqrt{2\log 1.25/\delta}\sqrt{\log p}}{n\epsilon}$. For this case, we have

$$\begin{aligned} \mathbb{P}(A_{ij}) \leq \mathbb{P}(|\tilde{\sigma}_{ij} - \sigma_{ij}| > \frac{3\gamma}{4}\sqrt{\frac{\log p}{n}} + \frac{3\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}) \\ \leq C_1 p^{-\frac{9}{2}} + 2p^{-\frac{9}{2}}. \end{aligned} \tag{13}$$

This is due to the followings:

$$\mathbb{P}\big(|\tilde{\sigma}_{ij}-\sigma_{ij}| > \frac{3\gamma}{4}\sqrt{\frac{\log p}{n}} + \frac{3\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}\big) \tag{14}$$

$$\leq \mathbb{P}\big(|\sigma^*_{ij}-\sigma_{ij}| > \frac{3\gamma}{4}\sqrt{\frac{\log p}{n}} + \frac{3\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}) - |n_{ij}|\big) \tag{15}$$

$$= \mathbb{P}\big(B_{ij} \bigcap \{\frac{3\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}) - |n_{ij}| > 0\}\big) \tag{16}$$

$$+ \mathbb{P}\big(B_{ij} \bigcap \{\frac{3\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}) - |n_{ij}| \leq 0\}\big) \tag{17}$$

$$\leq \mathbb{P}(|\sigma^*_{ij}-\sigma_{ij}| > \frac{3\gamma}{4}\sqrt{\frac{\log p}{n}}) + \mathbb{P}(\frac{2\sqrt{3\ln 1.25/\delta}\log p}{n\epsilon}) \leq |n_{ij}|) \tag{18}$$

$$\leq C_1 P^{-\frac{9}{2}} + 2p^{-\frac{9}{2}}, \tag{19}$$

where event B_{ij} denotes $B_{ij} = \{|\sigma^*_{ij}-\sigma_{ij}| > \frac{3\gamma}{4}\sqrt{\frac{\log p}{n}} + \frac{2\sqrt{2\ln 1.25/\delta}\log p}{n\epsilon}) - |n_{ij}|\}$, and the last inequality is due to (4) and (5).

Thus by (12), with probability at least $1 - C_1 p^{-\frac{9}{2}} - 2p^{-\frac{9}{2}}$, we have

$$|\hat{\sigma}_{ij}-\sigma_{ij}| = |\sigma_{ij}|,$$

which satisfies (11).

b) **Case 2:** $|\sigma_{ij}| \geq 2\gamma\sqrt{\frac{\log p}{n}} + \frac{8\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon})$. For this case, we have

$$\mathbb{P}(A^c_{ij}) \leq \mathbb{P}(|\tilde{\sigma}_{ij}-\sigma_{ij}| \geq \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}) \\ \leq C_1 p^{-8} + 2p^{-8},$$

where the proof is the same as (13-17). Thus, with probability at least $1 - C_1 p^{-\frac{9}{2}} - 2p^{-8}$, we have

$$|\hat{\sigma}_{ij}-\sigma_{ij}| = |\tilde{\sigma}_{ij}-\sigma_{ij}|. \tag{20}$$

Also, by (9), (11) also holds.

c) **Case 3:** Otherwise,

$$\frac{\gamma}{4}\sqrt{\frac{\log p}{n}} + \frac{\sqrt{2\log 1.25/\delta}\sqrt{\log p}}{n\epsilon} \leq |\sigma_{ij}| \\ \leq 2\gamma\sqrt{\frac{\log p}{n}} + \frac{8\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}).$$

For this case, we have

$$|\hat{\sigma}_{ij}-\sigma_{ij}| = |\sigma_{ij}| \text{ or } |\tilde{\sigma}_{ij}-\sigma_{ij}|. \tag{21}$$

When $|\sigma_{ij}| \leq \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}$, we can see from (9) that with probability at least $1 - 2p^{-6} - C_1 p^{-8}$,

$$|\tilde{\sigma}_{ij}-\sigma_{ij}| \leq \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon} \leq 4|\sigma_{ij}|.$$

Thus, (11) also holds.

Otherwise when $|\sigma_{ij}| \leq \gamma\sqrt{\frac{\log p}{n}} + \frac{4\sqrt{2\ln 1.25/\delta}\sqrt{\log p}}{n\epsilon}$, (11) also holds. Thus, Lemma 3 is true. □

By Lemma 3, we have the following upper bound on the ℓ_2-norm error of Σ^+.

Theorem 2. The output Σ^+ of Algorithm 1 satisfies:

$$\mathbb{E}\|\hat{\Sigma}-\Sigma\|_2^2 = O(\frac{s\log p\log\frac{1}{\delta}}{n\epsilon^2}), \tag{22}$$

where the expectation is taken over the coins of the Algorithm and the randomness of $\{x_1, x_2, \cdots, x_n\}$.

Proof. Due the space limit, we leave the proof in the full version of the paper. □

Comparing the bound in the above corollary with the optimal minimax rate $\Theta(\frac{s\log p}{n})$ in [17] for the non-private case, we can see that the impact of the differential privacy is to make the number of efficient sample from n to $n\epsilon^2$. It is an open problem to determine whether the bound in Theorem 2 is tight.

V. Experiments

In this section, we evaluate the performance of Algorithm 1 practically on synthetic datasets.

a) Data Generation: We first generate a symmetric sparse matrix $\tilde{U}$ with the sparsity ratio sr, that is, there are $sr \times p \times p$ non-zero entries of the matrix. Then, we let $U = \tilde{U} + \lambda I_p$ for some constant λ to make U positive semi-definite and then scale it to $U = \frac{U}{c}$ by some constant c which makes the norm of samples less than 1 (with high probability)[1]. Finally, we sample $\{x_1, \cdots, x_n\}$ from the multivariate Gaussian distribution $\mathcal{N}(0, U)$. In this paper, we will use set $\lambda = 50$ and $c = 200$.

b) Experimental Settings: To measure the performance, we compare the ℓ_2 norm of relative error, respectively. That is, $\frac{\|\Sigma^+ - U\|_2}{\|U\|_2}$ with the sample size n in three different settings: 1) we set $p = 100$, $\epsilon = 1$, $\delta = \frac{1}{n}$ and change the sparse ratio $sr = \{0.1, 0.2, 0.3, 0.5\}$. 2) We set $\epsilon = 1$, $\delta = \frac{1}{n}$, $sr = 0.2$, and let the dimensionality p vary in $\{50, 100, 200, 500\}$. 3) We fix $p = 200$, $\delta = \frac{1}{n}$, $sr = 0.2$ and change the privacy level as $\epsilon = \{0.1, 0.5, 1, 2\}$. We run each experiment 20 times and take the average error as the final one.

c) Experimental Results: Figure 1 is the result of DP-Thresholding (Algorithm 1) with ℓ_2 relative error, respectively. From the figure we can see that: 1) if the sparsity ratio is large *i.e.,* the underlying covairance matrix is more dense, the relative error will be larger, this is due to the fact showed in Theorem 2 that the error depends on the sparsity s. 2) The dimensionality only slightly affects the relative error. That is, even if we double the value of p, the error increases only slightly. This is consistent with our theoretical analysis in Theorem 2 which says that the error of our private estimators is only logarithmically depending on p (*i.e.,* $\log p$). 3) With

[1] Although the distribution is not bounded by 1, actually, as we see from previous section, we can obtain the same result as long as the ℓ_2 norm of the samples is bounded by 1.

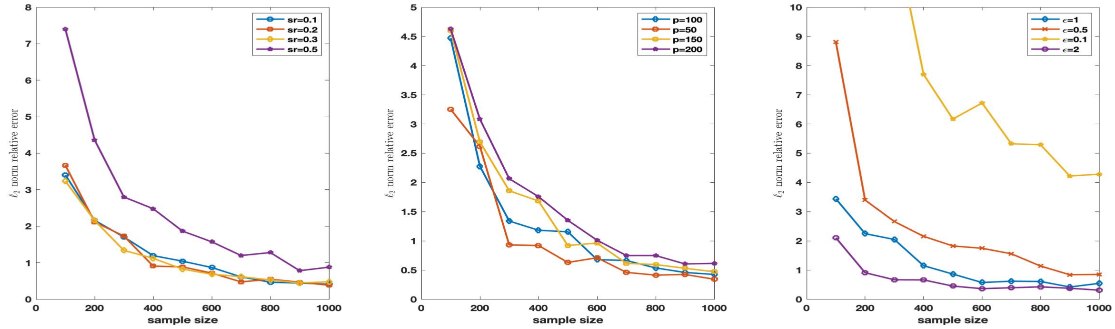

Fig. 1. Experiment results of Algorithm 1 for ℓ_2-norm relative error. The left one is for different sparsity levels, the middle one is for different dimensionality p, and the right one is for different privacy level ϵ.

the privacy parameter ϵ increases (which means more private), the error will become larger. This has also been showed in previous theorems.

In summary, all the experimental results support our theoretical analysis.

VI. Conclusion and Discussion

In the paper, we study the problem of estimating the sparse covariance matrix of a bounded sub-Gaussian distribution under differential privacy model and propose a method called DP-Threshold, which achieves a non-trivial error bound. Experiments on synthetic datasets yield consistent results with the theoretical analysis.

There are still some open problems for this problem. Firstly, although the thresholding method can achieve non-trivial error bound for our private estimator, in practice it is hart to find the best threshold. Thus, an open problem is how to get the best threshold. Secondly, as mentioned in the related work section, there are many recent results on private Gaussian estimation, which may make the ℓ_2 norm of the samples greater than 1. Thus, it is an interesting problem to extend our method to a general Gaussian distribution.

Acknowledgment

This research was supported in part by the National Science Foundation (NSF) through grants CCF-1422324 and CCF-1716400.

References

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography Conference*. Springer, 2006, pp. 265–284.

[2] D. Wang, A. Smith, and J. Xu, "Differentially private empirical risk minimization in non-interactive local model via polynomial of inner product approximation," in *Algorithmic Learning Theory, ALT 2019, 22-24 March 2019, Chicago, IL, USA*, 2019.

[3] D. Wang, M. Gaboardi, and J. Xu, "Empirical risk minimization in non-interactive local differential privacy revisited," in *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada.*, 2018, pp. 973–982.

[4] D. Wang, M. Ye, and J. Xu, "Differentially private empirical risk minimization revisited: Faster and more general," in *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, 2017, pp. 2719–2728.

[5] D. Wang and J. Xu, "Differentially private empirical risk minimization with smooth non-convex loss functions: A non-stationary view," *Thirty-Third AAAI Conference on Artificial Intelligence, (AAAI-19), Honolulu, Hawaii, USA, January 27-February 1, 2019*, 2019.

[6] Ú. Erlingsson, V. Pihur, and A. Korolova, "Rappor: Randomized aggregatable privacy-preserving ordinal response," in *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*. ACM, 2014, pp. 1054–1067.

[7] J. Tang, A. Korolova, X. Bai, X. Wang, and X. Wang, "Privacy loss in apple's implementation of differential privacy on macos 10.12," *CoRR*, vol. abs/1709.02753, 2017.

[8] D. Wang, A. Smith, and J. Xu, "High dimensional sparse linear regression under local differential privacy: Power and limitations," *2018 NIPS workshop in Privacy-Preserving Machine Learning*, 2018.

[9] J. Ge, Z. Wang, M. Wang, and H. Liu, "Minimax-optimal privacy-preserving sparse pca in distributed systems," in *International Conference on Artificial Intelligence and Statistics*, 2018, pp. 1589–1598.

[10] D. Wang, M. Huai, and J. Xu, "Differentially private sparse inverse covariance estimation," in *2018 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2018, Anaheim, CA, USA, November 26-29, 2018*.

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

[12] M. Joseph, J. Kulkarni, J. Mao, and Z. S. Wu, "Locally private gaussian estimation," *arXiv preprint arXiv:1811.08382*, 2018.

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

[14] M. Gaboardi, R. Rogers, and O. Sheffet, "Locally private mean estimation: Z-test and tight confidence intervals," *arXiv preprint arXiv:1810.08054*, 2018.

[15] K. Amin, T. Dick, A. Kulesza, A. M. Medina, and S. Vassilvitskii, "Private covariance estimation via iterative eigenvector sampling," *2018 NIPS workshop in Privacy-Preserving Machine Learning*, 2018.

[16] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang, "Analyze gauss: optimal bounds for privacy-preserving principal component analysis," in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*. ACM, 2014, pp. 11–20.

[17] T. T. Cai, H. H. Zhou *et al.*, "Optimal rates of convergence for sparse covariance matrix estimation," *The Annals of Statistics*, vol. 40, no. 5, pp. 2389–2420, 2012.

[18] T. Tao, *Topics in random matrix theory*. American Mathematical Soc., 2012, vol. 132.

[19] J. A. Tropp *et al.*, "An introduction to matrix concentration inequalities," *Foundations and Trends® in Machine Learning*, vol. 8, no. 1-2, pp. 1–230, 2015.

[20] P. J. Bickel, E. Levina *et al.*, "Covariance regularization by thresholding," *The Annals of Statistics*, vol. 36, no. 6, pp. 2577–2604, 2008.