DIFFERENTIALLY PRIVATE SPARSE INVERSE COVARIANCE ESTIMATION

Di Wang *Mengdi Huai* *Jinhui Xu*

State University of New York at Buffalo,
Department of Computer Science and Engineering,
Buffalo, New York.

ABSTRACT

In this paper, we present the first results on the sparse inverse covariance estimation problem under the differential privacy model. We first gave an ϵ-differentially private algorithm using output perturbation strategy, which is based on the sensitivity of the optimization problem and the Wishart mechanism. To further improve this result, we then introduce a general covariance perturbation method to achieve both ϵ-differential privacy and (ϵ, δ)-differential privacy. For ϵ-differential privacy, we analyze the performance of Laplacian and Wishart mechanisms, and for (ϵ, δ)-differential privacy, we examine the performance of Gaussian and Wishart mechanisms. Experiments on both synthetic and benchmark datasets confirm our theoretical analysis.

Index Terms— differential privacy, sparse inverse covariance estimation

1. INTRODUCTION

Nowadays, machine learning and signal processing algorithms are often required to deal with sensitive data or information. This means that the algorithm needs to not only learn effectively from the data or information but also provide a certain level of guarantee on privacy preserving. Differential privacy [7] is a rigorous notion for defining data privacy and has received a great deal of attentions in recent years.

Estimating the inverse covariance matrix (also called precision matrix) in high dimensional space is a fundamental problem in statistics and finds applications in many fields such as machine leaning, signal processing, computational biology, etc [21] [10]. It provides a good way for discovering the interactions among variables in high dimensional datasets, especially those from genetics, medicine, and healthcare. The inverse covariance matrix is also a natural way for parameterizing the Gaussian graphical model. One problem that often occurs in applying such a model is how to deal with sensitive data. For example, datasets related to gene expression may contain private information of individuals. Thus, it becomes a challenge for estimating the inverse covariance while preserving privacy. In this paper, we study the problem under the differential privacy model, and provide some results on this problem.

Definition 1. Let $\{x_1, \cdots, x_n\}$ be n instances sampled from a Gaussian distribution $\mathcal{N}(0, \Sigma)$, where each instance $x_i \in \mathbb{R}^d$ for $i \in [n]$ and $\Sigma \in \mathbb{R}^{d\times d}$ is the covariance matrix. The inverse covariance problem is to recover Σ^{-1} in a high dimensional setting, where $n \ll d$. Note that if $n \geq d$, we can solve the problem by optimizing $\Theta^* = S^{-1} = \arg\min_{\Theta\in\mathcal{S}^d_{++}} -\log\det\Theta + \langle S, \Theta\rangle$, where $S = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$ is the empirical covariance. But in a high dimensional setting, the above optimization problem is ill-posed, since S is rank-deficient. To make it well-defined, we borrow an idea in LASSO and use an ℓ_1 norm regularization in the objective function, which assumes that Θ^* is sparse. Thus, the objective function becomes the following:

$$\Theta^*_\rho = \arg\min_{\Theta\in\mathcal{S}^d_{++}} \{-\log\det\Theta + \langle S, \Theta\rangle + \rho\|\Theta\|_1\}, \quad (1)$$

where $\rho > 0$ is the penalty parameter, $\langle S, \Theta\rangle = \mathrm{tr}(S\Theta^T)$, and $\|\Theta\|_1 = \sum_{i,j}|\Theta_{i,j}|$.

Under the differential privacy model, our problem is to obtain Θ_{priv} under differential privacy so that $\|\Theta_{\text{priv}} - \Theta^*_\rho\|_F$ is as small as possible.

We first present an output-perturbation algorithm (See Algorithm 1) based on the sensitivity of (1). Unlike the commonly used Laplacian or Gaussian mechanisms in differential privacy [8], we adopt the Wishart distribution to preserve the positive definite property for the resulting matrix. To reduce the error bound of the above algorithm, we then introduce a general method by perturbing the covariance matrix, and analyze the error upper bound for different perturbing matrices. Finally, we evaluate the performance of our algorithms using both synthetic and real world datasets.

2. RELATED WORKS

The most closely related work to ours is differentially private PCA, since it also relies on random matrices to preserve privacy. For example, [12, 13] used the Wishart mechanism to achieve ϵ-differentially private PCA, and [4, 9] adopted the Gaussian mechanism to analyze the optimal bound of PCA under the (ϵ, δ)-differential privacy model. Note that although

our paper uses the same mechanisms (as the aforementioned results), the way for analyzing the error bound is quite different. While the above results mainly relied on techniques in linear algebra, ours is based on some optimization techniques (due to the ℓ_1 regularization and the positive definite requirement for the resulting matrix). Thus, existing approaches/techniques cannot be used to analyze our problem.

3. PRELIMINARIES

Definition 2 ([7]). A randomized algorithm $\mathcal{A}$ is (ϵ, δ)-differentially private if for all neighboring datasets $D, D' \in \mathcal{X}^n$ and for all events S in the output space of $\mathcal{A}$, the following holds $\Pr(\mathcal{A}(D) \in S) \leq e^{\epsilon}\Pr(\mathcal{A}(D') \in S) + \delta$. When $\delta = 0$, $\mathcal{A}$ is ϵ-differentially private.

Definition 3. A $d \times d$ random symmetric positive definite matrix W is said to have a Wishart distribution $W \sim \mathcal{W}_d(m, C)$ if its probability density function is

$$p(W) = \frac{(\det W)^{\frac{m-d-1}{2}}}{2^{\frac{md}{2}}(\det C)^{\frac{m}{2}}\Gamma_d(\frac{m}{2})}\exp(-\frac{1}{2}\operatorname{tr}(C^{-1}W)), \quad (2)$$

where $m > d-1$ and C is a $d \times d$ positive definite matrix.

Next we show how to select the parameters m and C to ensure differential privacy and their tail bounds.

Lemma 1 ((ϵ, δ)-differential privacy [18]). Fix $\epsilon \in (0, 1)$ and $\delta \in (0, \frac{1}{e})$. For a fixed constant $B > 0$, let A be an $n \times d$ matrix, where each row of A has bounded ℓ_2-norm B. Let N be a matrix sampled from $\mathcal{W}(m, B^2 I_d)$ for $m \geq d + \frac{14}{\epsilon^2}\ln(\frac{4}{\delta})$. Then, outputting $X = A^T A + N$ is (ϵ, δ)-differentially private.

Lemma 2 (ϵ-differential privacy [13]). Fix $\epsilon > 0$ and let A be an $n \times d$ matrix, where each row of A has bounded ℓ_2-norm of B. Let $N \sim \mathcal{W}_d(d+1, C)$, where $C = \frac{3}{2n\epsilon}B^2 I_d$. Then, outputting $X = A^T A + N$ is ϵ-differentially private.

Lemma 3 ([18]). Fix $\delta' \in (0, \frac{1}{e})$, and a random matrix $X \sim \mathcal{W}_d(m, V)$, where $m > (\sqrt{d} + \sqrt{2\log\frac{2}{\delta'}})^2$. Then, with probability at least $1 - \delta'$, the following holds for every $j = 1, \cdots, d$, $\sigma_j(X) \in (\sqrt{m} \pm (\sqrt{d} + \sqrt{2\log\frac{2}{\delta'}}))^2\sigma_j(V)$.

Lemma 4 ([22]). If $X \sim \mathcal{W}_d(m, V)$, then with probability at least $1 - 2d\exp(-\theta)$ for any $\theta \geq 0$, we have for each $l = 1, \cdots, d$, $|\sigma_l(\frac{1}{m}X) - \sigma_l(V)| \leq (\sqrt{\frac{2\theta k_l^2(r+1)}{m}} + \frac{2\theta k_l r}{m})\sigma_l(V)$, where $r = \frac{\operatorname{tr}(V)}{\sigma_1(V)}$ and $k_l = \frac{\sigma_1(V)}{\sigma_l(V)}$.

If taking $V = B^2 I_d, m = d+1$, and $\theta = \log\frac{2d}{\delta'}$, Lemma 4 tells us that with probability at least $1-\delta'$, we have $\sigma_l(X) \leq O(d\log\frac{d}{\delta'}B^2)$ for each $l = 1 \cdots, d$.

For $\rho > 0$, the problem is strongly convex and thus has a unique optimal solution Θ_ρ^*, which satisfies the following.

Lemma 5 ([6, 15]). The solution of (1), Θ_ρ^*, satisfies that $\alpha I_d \preceq \Theta_\rho^* \preceq \beta I_d$, for $\alpha = \frac{1}{\|S\|_2 + \rho d}, \beta = \frac{d - \alpha\operatorname{tr}(S)}{\rho}$.

4. OUTPUT PERTURBATION METHOD

In this section, we present an ϵ-differentially private algorithm based on the output perturbation strategy (see Algorithm 1 for details), and analyze the sensitivity and stability of the problem (1). Although the method has some undesirable features, the error bound analysis and the guarantee of differential privacy are useful for our later methods.

Algorithm 1 Output Perturbation

Input: $D = \{x_i\}_{i=1}^n$, $S = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$, where the ℓ_2-norm of each row x_i is bounded by 1, $\rho > 0$.

1: Compute $\Theta_\rho^* = \arg\min_{\Theta \in \mathcal{S}_{++}^d}\{-\log\det\Theta + \langle S, \Theta\rangle + \rho\|\Theta\|_1\}$,
2: **return** $\tilde{\Theta}_\rho^* = \Theta_\rho^* + N$, where $N \sim \mathcal{W}_d(d+1, C), C = \frac{d^{\frac{5}{2}}}{n\epsilon\rho^2}I_d$.

Theorem 1 (Privacy guarantee). For any $\epsilon > 0$, Algorithm 1 is ϵ-differentially private.

Proof. For convenience, we denote Step 1 of Algorithm 1 as $\mathcal{A}$. That is, $\Theta_\rho^* = \mathcal{A}(D)$. Also, we let D' be a neighboring dataset, and $S' = S - \frac{1}{n}vv^T + \frac{1}{n}v'v'^T$, $\Theta_\rho'^* = \mathcal{A}(D')$. Then by the optimality of $\Theta_\rho'^*, \Theta_\rho^*$ of the optimization problem, we have, for any $\eta > 0$, $\|\Theta_\rho^* - \Theta_\rho'^*\|_F = \|\operatorname{prox}_{\eta g}(\Theta_\rho^* - \eta(S - \Theta_\rho^{*-1})) - \operatorname{prox}_{\eta g}(\Theta_\rho'^* - \eta(S' - \Theta_\rho^{'*-1}))\|_F$, here prox is the proximal operator with respect to $\|\cdot\|_1$ [5]. Then, by the non-expansive property of the proximal operator, we have $\|\Theta_\rho^* - \Theta_\rho'^*\|_F \leq \|(\Theta_\rho^* - \eta(S - \Theta_\rho^{*-1})) - (\Theta_\rho'^* - \eta(S' - \Theta_\rho^{'*-1}))\|_F$. If let $f(\Theta_\rho^*) = \Theta_\rho^* + \eta\Theta_\rho^{*-1}$, we have $\|\Theta_\rho^* - \Theta_\rho'^*\|_F \leq \|f(\Theta_\rho^*) - f(\Theta_\rho'^*)\|_F + \eta\|S - S'\|_F$.

For the last term, we have $\|S - S'\|_F = \|\frac{1}{n}(vv^T - v'v'^T)\|_F \leq \frac{2}{n}$. In order to bound the first term, we need the following lemma, which has been proved in [16].

Lemma 6. [16] For $\Theta_1, \Theta_2 \in \mathcal{S}_{++}^d$, $\eta > 0$, we have $\|f(\Theta_1) - f(\Theta_2)\|_F \leq \max\{|1 - \frac{\eta}{a^2}|, |1 - \frac{\eta}{b^2}|\}\|\Theta_1 - \Theta_2\|_F$, where $a = \max\{\sigma_{\max}(\Theta_1), \sigma_{\max}(\Theta_2)\}$, $b = \min\{\sigma_{\min}(\Theta_1), \sigma_{\min}(\Theta_2)\}$.

Take $\Theta_\rho^*, \Theta_\rho'^*$ into Lemma 6 and set $0 < \eta < b^2$, we now have $\|\Theta_\rho^* - \Theta_\rho'^*\|_F \leq \frac{2\beta^2}{n}$, where $\beta = \max\{\|\Theta_\rho^*\|_2, \|\Theta_\rho'^*\|_2\}$. Now we will show the ϵ-differential privacy. Since for every W,

$$\frac{\Pr[\Theta_\rho^* + N = W]}{\Pr[\Theta_\rho'^* + N = W]} = \frac{\Pr[N = W - \Theta_\rho^*]}{\Pr[N = W - \Theta_\rho'^*]} =$$

$$\frac{\exp(-\frac{1}{2}\operatorname{tr}(C^{-1}(W - \Theta_\rho^*))}{\exp(-\frac{1}{2}\operatorname{tr}(C^{-1}(W - \Theta_\rho'^*))} = \exp(-\frac{1}{2}\operatorname{tr}(C^{-1}(\Theta_\rho^* - \Theta_\rho'^*))$$

$$\leq \exp(\frac{1}{2}\|C^{-1}\|_F\|\Theta_\rho^* - \Theta_\rho'^*\|_F) \leq \exp(\frac{1}{2}\sqrt{d}\frac{n\epsilon\rho^2}{d^{\frac{5}{2}}}\frac{2\beta^2}{n})$$

$$\leq \exp(\epsilon).$$

Where the last inequality comes from Lemma 5. □

By Lemma 4, we have the following error upper bound.

Theorem 2. For Algorithm 1, with probability at least $1-\delta$ for any $0 < \delta < 1$, we have $\|\tilde{\Theta}^*_\rho - \Theta^*_\rho\|_F \leq O(\frac{\log \frac{d}{\delta} d^4}{n\epsilon\rho^2})$, where Θ^*_ρ is the optimal solution of the original problem (1).

Remark 1. Note that in Algorithm 1, a Wishart matrix needs to be added to the output to ensure that the resulting matrix is positive definite (as required by problem (1)). Since other random matrices, such as symmetric Laplacian matrices, may not be positive definite [11], adding them to the output may not yield the desired solution.

Although Algorithm 1 provides an ϵ-differentially private algorithm for the inverse covariance estimation problem. It also leaves quite a few unresolved issues. Firstly, from Theorem 2, we know that the error bound heavily depends on the dimensionality (*i.e.,* $d^4 \log d$), which could be too large for high dimensional datasets. Thus, a natural question is whether the error bound can be further reduced. Secondly, for many problems, the error bound of an (ϵ, δ)-differentially private algorithm is often lower than that of an ϵ-differentially private algorithm (*e.g.,* Differentially Private Empirical Risk Minimization [2, 20]). Thus, an interesting question is whether the problem considered in this paper also follows the same pattern. Below we will address the two issues by proposing a covariance perturbation method.

5. COVARIANCE PERTURBATION METHODS

As shown in Theorem 1, the sensitivity of problem (1) is high (since β is often large). This means that we need to add a large amount of noise in Algorithm 1 to ensure the ϵ-differential privacy. To deal with this problem, along with the aforementioned issues, we propose in this section a general method which perturbs the empirical covariance S (see Algorithm 2), instead of the output. This allows us to significantly reduce the amount of noise that needs to be added. Also, it can be implemented by using different kinds of random matrices N. To compare the performance for different mechanisms, we analyze the error bound for each of them.

Algorithm 2 Covariance Perturbation

Input: $D = \{x_i\}_{i=1}^n$, where the ℓ_2-norm of each row x_i is bounded by 1, $\rho > 0$. $\epsilon, \delta \geq 0$ are the privacy parameters.

1: Let $S = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$; sample a symmetric matrix $N \in \mathbb{R}^{d\times d} \sim \mathcal{P}$, which makes $S+N$ ϵ- or (ϵ,δ)-differentially private. Let $\tilde{S} = S + N$.
2: Return $\hat{\Theta}^*_\rho = \arg\min_{\Theta\in\mathcal{S}^d_{++}}\{-\log\det\Theta + \langle\tilde{S},\Theta\rangle + \rho\|\Theta\|_1\}$.

Below, we consider those random matrices that ensure ϵ-differential privacy. The first one is due to Lemmas 2 and 4.

Theorem 3. In Algorithm 2, for any $\epsilon > 0$, if choose $\mathcal{P} = \mathcal{W}_d(m, C)$ with $C = \frac{3}{2\epsilon n} I_d$ and $m = d+1$, it is ϵ-differentially private for any $\epsilon > 0$. Moreover, with probability at least $1-\delta'$, the following holds

$$\|\hat{\Theta}^*_\rho - \Theta^*_\rho\|_F \leq O(\frac{\log\frac{d}{\delta'} d^{\frac{3}{2}} \max\{\|\Theta^*_\rho\|_2^2, \|\hat{\Theta}^*_\rho\|_2^2\}}{n\epsilon}).$$

Next, we consider the case that N is sampled from a Laplacian distribution. Since the covariance matrix is symmetric, the added noise also needs to be symmetric, the following lemma is due to [19].

Theorem 4. In Algorithm 2, for any $\epsilon > 0$, if N is a symmetric Laplacian matrix N whose entries are i.i.d drawn from Lap$(0, \frac{2d}{n\epsilon})$, then it is ϵ-differentially private. Moreover, with high probability, the following holds

$$\|\hat{\Theta}^*_\rho - \Theta^*_\rho\|_F \leq O(\frac{d^2 \max\{\|\Theta^*_\rho\|_2^2, \|\hat{\Theta}^*_\rho\|_2^2\}}{n\epsilon}).$$

Remark 2. Comparing Theorems 3 and 4, we can see that the error in Theorem 3 is less than that in Theorem 4 (if we omit the term $\max\{\|\Theta^*_\rho\|_2^2, \|\hat{\Theta}^*_\rho\|_2^2\}$). Another advantage is that adding Wishart matrix not only preserves the symmetry property, but also guarantees the positive semi-definite property of the covariance matrix. Thus, for ϵ-differential privacy, it is better to use Wishart mechanism theoretically.

Next, we consider (ϵ, δ)-differential privacy and also start with adding Wishart matrices. The following theorem is due to Lemmas 1 and 3.

Theorem 5. For any $\epsilon \in (0,1)$ and $\delta \in (0, \frac{1}{e})$, if choose $\mathcal{P} = \mathcal{W}_d(m, C)$ with $C = \frac{1}{n} I_d$ and $m = d + \frac{14}{\epsilon^2}\ln(\frac{4}{\delta})$ in Algorithm 2, it is (ϵ,δ)-differentially private. Moreover, if $m > (\sqrt{d} + \sqrt{2\log\frac{2}{\delta'}})^2$ for $0 < \delta' < 1$, then with probability at least $1-\delta'$, we have

$$\|\hat{\Theta}^*_\rho - \Theta^*_\rho\|_F \leq O(\frac{\max\{\|\Theta^*_\rho\|_2^2, \|\hat{\Theta}^*_\rho\|_2^2\}\ln(1/\delta)\ln(1/\delta')d^{\frac{3}{2}}}{n\epsilon^2}).$$

Now, we consider adding symmetric Gaussian matrices.

Theorem 6. In Algorithm 2, for any $\epsilon > 0$ and $0 < \delta < 1$, if N is a symmetric Gaussian matrix N whose entries are i.i.d drawn from $\mathcal{N}(0, \beta^2)$, where $\beta = \frac{\sqrt{2\ln(\frac{1.25}{\delta})}}{n\epsilon}$, then it is (ϵ,δ)-differentially private. Moreover, with high probability, we have

$$\|\hat{\Theta}^*_\rho - \Theta^*_\rho\|_F \leq O(\frac{d\sqrt{\ln(\frac{1}{\delta})}\max\{\|\Theta^*_\rho\|_2^2, \|\hat{\Theta}^*_\rho\|_2^2\}}{\epsilon n}).$$

Remark 3. From the above two theorems, we can see that although the Wishart mechanism preserves the positive definite property of $\tilde{S}$, which is not the case for the Gaussian mechanism [11], it has an additional factor of $\sqrt{d}$ in its error bound compared with the Gaussian mechanism (if we omit the term $\max\{\|\Theta_\rho^*\|_2^2, \|\hat{\Theta}_\rho^*\|_2^2\}$). Thus, if we need a more accurate solution, Gaussian mechanism is a better choice from the theoretical view.

Now, we address the three issues raised in previous part. Firstly, for the large error bound in Theorem 2, we know from Theorem 3 that the covariance perturbation based ϵ-differentially private algorithm always has a lower error bound than that of an output perturbation based algorithm (since $\max\{\|\Theta_\rho^*\|_2^2, \|\hat{\Theta}_\rho^*\|_2^2\} \leq \frac{d^2}{\rho^2}$ by Lemma 5). Secondly, if we view ϵ as a constant and omit the term of $\max\{\|\Theta_\rho^*\|_2^2, \|\hat{\Theta}_\rho^*\|_2^2\}$, the error bound of the (ϵ, δ)-differentially private algorithm with covariance perturbation strategy is lower than it under ϵ-differential privacy, and Gaussian mechanism achieves the lowest error bound.

6. EXPERIMENTS

In this section, we present some numerical results on both real-world and synthetic datasets to evaluate the performance of our proposed differentially private algorithms. More experiments are left to the full paper.

We first introduce the algorithms that we are going to compare. For ϵ-differentially private algorithm, we will compare with output perturbation, Laplace and Wishart covariance perturbation methods. For (ϵ, δ)-differentially private algorithm, we will compare with SULQ Framework [3], Wishart and Gaussian covariance perturbation methods.

We let $\hat{\Theta}_\rho^*$ denote the output of the differentially private algorithm and Θ_ρ^* denote the optimal solution of the original problem. To evaluate the performance of the proposed methods, we choose **Relative Error**, which is defined as $\frac{\|\hat{\Theta}_\rho^* - \Theta_\rho^*\|_F}{\|\Theta_\rho^*\|_F}$. If the relative error is greater than 200, we use NA to indicate.

For synthetic datasets, we first fix the dimensionality d and create a sparse matrix U with nonzero entries equal to -1 or 1 with equal probability. Then, we compute $S = (U * U^T)^{-1}$ as the true covariance matrix. The inverse covariance matrix $S^{-1} = UU^T$ is, thus, sparse. Given the inverse covariance matrix $S^{-1} = UU^T$, we then draw $n = r \times d$ samples from the Gaussian distribution $\mathcal{N}(0, S)$ to simulate the high-dimensional settings, where r denotes the ratio of n (*i.e.,* the sample size) over d (*i.e.,* the dimensionality of the samples). We test our proposed methods for $d = 400$ and $r = 0.5, 1.0, 1.5$.

For real-world datasets, we use the colon cancer dataset [1] and the Parkinson's disease dataset [14] to evaluate our proposed methods. The colon cancer dataset contains information of 69 individuals with 2000 attributes. We choose 300 variables for the experiment. The size of Parkinson's disease dataset is $(192, 22)$. The datasets are normalized before processing.

For each experiment, we choose $\epsilon = 0.5, 1, 1.5$, respectively. For (ϵ, δ)-DP, we let $\delta = 0.0001$. Note that these privacy parameters are often chosen in other works. To solve the optimization problem (1), we set $\rho = 0.001$ and use the method in [17]. All experiments run in MATLAB.

Table 1: Performance comparisons of the ϵ-differentially private algorithms on both synthetic and real-world datasets.

		Synthetic Datasets			Real-world Datasets	
ϵ	Methods	$r = 0.5$	$r = 1.0$	$r = 1.5$	Colon	Parkinson's
0.5	**Wishart**	**0.993**	**0.9918**	**0.9914**	**0.995**	**0.9140**
	Output	NA	NA	NA	NA	NA
	Laplace	101.4	52.85	35.42	190.57	9.950
1.0	**Wishart**	**0.986**	**0.9863**	**0.9856**	**0.993**	**0.8899**
	Output	NA	NA	NA	NA	NA
	Laplace	49.44	25.41	16.83	95.01	4.690
1.5	**Wishart**	**0.9817**	**0.9815**	**0.9806**	**0.9907**	**0.8796**
	Output	NA	NA	NA	NA	NA
	Laplace	32.30	16.41	10.76	63.67	3.913

Table 2: Performance comparisons of the (ϵ, δ)-differentially private algorithms on both synthetic and real-world datasets.

		Synthetic Datasets			Real-world Datasets	
ϵ	Methods	$r = 0.5$	$r = 1.0$	$r = 1.5$	Colon	Parkinson's
0.5	Wishart	0.9999	0.9997	0.9993	1.636	1.00
	SQLU	NA	NA	NA	NA	0.7419
	Gaussian	**0.1285**	**0.1607**	**0.1759**	**0.3039**	**0.1527**
1.0	Wishart	0.9982	0.9947	0.9906	1.1155	0.990
	SQLU	NA	NA	NA	NA	0.7318
	Gaussian	**0.1254**	**0.1605**	**0.1737**	**0.1081**	**0.1514**
1.5	Wishart	0.9954	0.9895	0.9837	1.0474	0.9992
	SQLU	NA	NA	NA	NA	0.7065
	Gaussian	**0.1242**	**0.1585**	**0.1701**	**0.0833**	**0.1474**

The experimental results of the ϵ-differentially private algorithms on both synthetic and real-world datasets are shown in Table 1. From the table, we can see that in all the cases, Wishart and Laplacian mechanisms achieve better performance than the output perturbation method. Furthermore, Wishart mechanism is the best among the three types of methods. From Table 2, we can see that the Gaussian method has the lowest relative error among all (ϵ, δ)-differentially private algorithms. Also, Gaussian mechanism has the lowest relative error among all the compared methods. In general, the relative error becomes smaller for larger ϵ. But in some cases, the relative error are almost same for different ϵ values. This could be due to the fact the difference of these ϵ values is small.

Thus, we can see all the experimental results support our theoretical analysis.

7. REFERENCES

[1] U. Alon, N. Barkai, D. A. Notterman, K. Gish, S. Ybarra, D. Mack, and A. J. Levine, "Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays," *Proceedings of the National Academy of Sciences*, vol. 96, no. 12, pp. 6745–6750, 1999.

[2] R. Bassily, A. Smith, and A. Thakurta, "Private empirical risk minimization: Efficient algorithms and tight error bounds," in *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*. IEEE, 2014, pp. 464–473.

[3] A. Blum, C. Dwork, F. McSherry, and K. Nissim, "Practical privacy: the sulq framework," in *Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 2005, pp. 128–138.

[4] K. Chaudhuri, A. Sarwate, and K. Sinha, "Near-optimal differentially private principal components," in *Advances in Neural Information Processing Systems*, 2012, pp. 989–997.

[5] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[6] A. d'Aspremont, O. Banerjee, and L. El Ghaoui, "First-order methods for sparse covariance selection," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 1, pp. 56–66, 2008.

[7] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *TCC*, vol. 3876. Springer, 2006, pp. 265–284.

[8] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[9] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang, "Analyze gauss: optimal bounds for privacy-preserving principal component analysis," in *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*. ACM, 2014, pp. 11–20.

[10] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.

[11] Z. Füredi and J. Komlós, "The eigenvalues of random symmetric matrices," *Combinatorica*, vol. 1, no. 3, pp. 233–241, 1981.

[12] H. Imtiaz and A. D. Sarwate, "Symmetric matrix perturbation for differentially-private principal component analysis," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2339–2343.

[13] W. Jiang, C. Xie, and Z. Zhang, "Wishart mechanism for differentially private principal components analysis." in *AAAI*, 2016, pp. 1730–1736.

[14] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, L. O. Ramig *et al.*, "Suitability of dysphonia measurements for telemonitoring of parkinson's disease," *IEEE transactions on biomedical engineering*, vol. 56, no. 4, pp. 1015–1022, 2009.

[15] Z. Lu, "Smooth optimization approach for sparse covariance selection," *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1807–1827, 2009.

[16] B. Rolfs, B. Rajaratnam, D. Guillot, I. Wong, and A. Maleki, "Iterative thresholding algorithm for sparse inverse covariance estimation," in *Advances in Neural Information Processing Systems*, 2012, pp. 1574–1582.

[17] K. Scheinberg, S. Ma, and D. Goldfarb, "Sparse inverse covariance selection via alternating linearization methods," in *Advances in neural information processing systems*, 2010, pp. 2101–2109.

[18] O. Sheffet, "Private approximations of the 2nd-moment matrix using existing techniques in linear regression," *arXiv preprint arXiv:1507.00056*, 2015.

[19] T. Tao, *Topics in random matrix theory*. American Mathematical Society Providence, RI, 2012, vol. 132.

[20] D. Wang, M. Ye, and J. Xu, "Differentially private empirical risk minimization revisited: Faster and more general," in *NIPS-2017*, 2017.

[21] M. Yuan and Y. Lin, "Model selection and estimation in the gaussian graphical model," *Biometrika*, vol. 94, no. 1, pp. 19–35, 2007.

[22] S. Zhu, "A short note on the tail bound of wishart distribution," *arXiv preprint arXiv:1212.5860*, 2012.