

Article

Differentially Private Sparse Covariance Matrix Estimation under Lower-Bounded Moment Assumption

Huimin Li and Jinru Wang *

Department of Mathematics, Beijing University of Technology, Beijing 100124, China; lihuim@emails.bjut.edu.cn

* Correspondence: wangjinru@bjut.edu.cn

Abstract: This paper investigates the problem of sparse covariance matrix estimation while the sampling set contains sensitive information, and both the differentially private algorithm and locally differentially private algorithm are adopted to preserve privacy. It is worth noting that the requirement of the distribution assumption in our work is only the existing bounded $4 + \epsilon$ ($\epsilon > 0$) moment. Meanwhile, we reduce the error bounds by modifying the threshold of the existing differentially private algorithms. Finally, the numerical simulations and results from a real data application are presented to support our theoretical claims.

Keywords: differential privacy; lower-bounded moment; sparse covariance matrix; probability estimation

MSC: 62F30; 62H12; 62J10



Citation: Li, H.; Wang, J. Differentially Private Sparse Covariance Matrix Estimation under Lower-Bounded Moment Assumption. *Mathematics* **2023**, *11*, 3670. <https://doi.org/10.3390/math11173670>

Received: 14 June 2023

Revised: 2 August 2023

Accepted: 23 August 2023

Published: 25 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The covariance matrix plays an important role in modern multivariate analysis. Many methodologies, including regression analysis, principal component analysis and discriminant analysis, rely on the estimation of the covariance matrices. Such estimation problems arise in various fields of science; for example, risk management [1], gene expressions [2], social networks [3], among others. Up to now, many methods have been developed to estimate the covariance matrix. One of the most popular approaches is to impose the sparse structure on the covariance matrix (see [4,5]).

In the big data era, sampling databases usually include personal financial or health information, such as those used in social science, biomedicine and genomics, so it is inevitable that one must deal with sensitive data. In recent years, the differentially private algorithm (DPA) and locally differentially private algorithm (LDPA) have become widely used methods that can prevent privacy leakage and defend against differential attacks, and these methods have been developed in the real world by Apple [6], Microsoft [7] and Google [8]. The DPA and LDPA aim at hiding the true information while keeping the basic property of the whole dataset. A popular idea to achieve this goal is to add some special noise into the original model [9–11].

This paper focuses on efficient estimation for a form of general sparse covariance matrix based on both DPA and LDPA in order to protect sensitive information. Although the study of datasets with the DPA and LDPA brings difficulties, it also makes it possible to use some special methods to break this bottleneck. In this paper, we adopt the Gaussian mechanism to keep the differentially private property of algorithms, and use the thresholding method to deal with the covariance matrix.

1.1. Related Work and Our Contributions

There are several papers developing the theory of covariance matrix estimation based on the DPA or LDPA to provide private protection. Jiang et al. [12] added Wishart-distributed noise to construct a DPA, and applied it to keep privacy while establishing the

covariance matrix estimation; however, the Wishart distribution seems a strong condition. Amin et al. [13] used the Laplace mechanism to construct a DPA, and they utilized it to protect sensitive samples while estimating the covariance matrix; nonetheless, their method is just for the low-dimensional case. Kamath et al. [14] considered high-dimensional covariance matrix estimation with a DPA to protect privacy, but they did not assume the sparse structure of the covariance matrix, so it led to a large error bound. Moreover, the distribution assumption is required to be Gaussian. Recently, Wang and Xu [10] reduced the distribution assumption from Gaussian to sub-Gaussian and imposed the sparse structure on the covariance matrix, which led to a low convergence rate, while using a DPA and LDPA to preserve privacy.

Our contributions are as follows: (i) For the random vector $\mathbf{X} = (X_1, \dots, X_p)^T \in \mathbb{R}^p$, the requirement for the distribution of X_i is only the existing bounded $4 + \varepsilon$ ($\varepsilon > 0$) moment. However, previous work requires that \mathbf{X} is Gaussian [14] or sub-Gaussian [10], which is equivalent to any existing bounded moment of X_i ($i = 1, \dots, p$) (see [15]). Moreover, if X_i ($i = 1, \dots, p$) is subjected to a heavy-tailed distribution, i.e., the distribution of X_i satisfies $\int_{\mathbb{R}} e^{tx} dF_i(x) = \infty$ for any $t > 0$, then it is obvious that this distribution can be covered by our distribution, and cannot be handled by the existing literature. (ii) We adjust the threshold in the original algorithms of Wang and Xu [10], which leads to the error bound becoming smaller under both the DPA and LDPA (see Remarks in 3 and 7). (iii) We measure the error bounds not only by the spectral norm and matrix operator norm, but also by the Frobenius norm, which has not previously been considered as far as we know. (iv) The sparse structure of the covariance matrix in the present paper contains that of Wang and Xu under an extra condition as a special case (see Remark 1).

1.2. Notations

For a random variable Z , $\mathbb{E}Z$ and $\mathbb{D}Z$ denote the expectation and variance of Z , respectively. For a vector $\mathbf{X} = (X_1, \dots, X_p)^T \in \mathbb{R}^p$, we define its l_w norm by $\|\mathbf{X}\|_{l_w} = (\sum_{i=1}^p |X_i|^w)^{1/w}$ with $w \in [1, \infty)$ and $\|\mathbf{X}\|_{l_\infty} = \max_i |X_i|$. For a matrix $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{p \times p}$, the spectral norm is defined as $\|\mathbf{A}\|_2 = \sup_{\|\mathbf{x}\|_2 \leq 1} \|\mathbf{A}\mathbf{x}\|_2$; the matrix l_1 norm and the Frobenius norm are defined by $\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^p |a_{ij}|$ and $\|\mathbf{A}\|_F = \sqrt{\sum_{i,j=1}^p |a_{ij}|^2}$, respectively. Moreover, the matrix l_w operator norm is given by $\|\mathbf{A}\|_w = \sup_{\|\mathbf{x}\|_{l_w} \leq 1} \|\mathbf{A}\mathbf{x}\|_{l_w}$. For two sequences of real numbers $\{a_n\}_{n \geq 1}$ and $\{b_n\}_{n \geq 1}$, “ $a_n = O(b_n)$ ” stands for $a_n \leq Cb_n$ for some constant $C > 0$ independent of n .

1.3. Organization of this Paper

The remainder of this paper is outlined as follows. Section 2 introduces some important definitions and lemmas. In Section 3, we firstly establish the sparse covariance matrix estimation based on the DPA, while any component of the random vector just has a bounded $4 + \varepsilon$ ($\varepsilon > 0$) moment. Then, the estimation problem is extended to utilize the LDPA. Finally, the results from several numerical experiments and a real data example are presented to support our theoretical results. All proofs are relegated to Appendix A.

2. Preliminaries

2.1. Sparse Covariance Estimation

In this paper, we always assume the real random vector $\mathbf{X} = (X_1, \dots, X_p)^T \in \mathbb{R}^p$ satisfying the following conditions, denoted as $\mathbf{X} \sim \mathcal{P}(K, \tau_0, s_{n,p}, M)$.

Condition (i) For a polynomial-type moment, assume $\mathbb{E}\mathbf{X} = 0$, $\|\mathbf{X}\|_2 = 1$, and for some positive constants γ , τ_0 , K and $\varepsilon = c(\gamma)$, it holds $p \leq n^\gamma$,

$$\mathbb{E} \left| \frac{X_i}{\sqrt{\mathbb{D}X_i}} \right|^{4+\varepsilon} \leq K \text{ for } 1 \leq i \leq p$$

$$\text{and } \min_{1 \leq i, j \leq p} \mathbb{D} \left(\frac{X_i X_j}{\sqrt{\mathbb{D}X_i \mathbb{D}X_j}} \right) \geq \tau_0.$$

Condition (ii) The covariance matrix Σ of the random vector \mathbf{X} belongs to the l_q -sparse space as follows:

$$\mathcal{U}_q(s_{n,p}, M) := \left\{ \Sigma = (\sigma_{ij})_{p \times p} > 0 : \max_j \sum_{i=1}^p |\sigma_{ij}|^q \leq s_{n,p}, \max_j \sigma_{jj} \leq M \right\}.$$

Remark 1. Condition (i) mainly means that the distribution of the any component of \mathbf{X} satisfies the existing bounded $4 + \varepsilon$ ($\varepsilon > 0$) moment. Note that if \mathbf{X} is subjected to a multivariate normal distribution $N_p(\mu, \Sigma)$ or a multivariate t -distribution $T_p(\alpha; \mu, \Sigma)$ with freedom $\alpha \geq 4 + \varepsilon$, then Condition (i) holds. Moreover, the relationship $p \leq n^\gamma$ is necessary for proofing Lemma 1. Meanwhile, the parameter ε controls the level of moment; K, M can be determined if the distribution of \mathbf{X} and the covariance matrix Σ are given.

Condition (ii) is a common assumption for the high-dimensional covariance matrix Σ of \mathbf{X} (see [4] and [5]). Moreover, we find that the parameter space defined in Wang and Xu [10] is

$$\mathcal{G}_0(s_{n,p}) := \left\{ \Sigma = (\sigma_{ij})_{p \times p} > 0 : \max_j \sum_{i \neq j} |\sigma_{ij}|^0 \leq s_{n,p} \right\}. \quad (1)$$

Then, it is obvious that there exists a constant $C > 0$, such that

$$\mathcal{G}_0(s_{n,p}, M) := \left\{ \Sigma \in \mathcal{G}_0(s_{n,p}), 0 < \sigma_{jj} \leq M (j = 1, \dots, p) \right\} \subseteq \mathcal{U}_0(Cs_{n,p}, M), \quad (2)$$

i.e., the sparse space (1) with the extra condition of $0 < \sigma_{jj} \leq M (j = 1, \dots, p)$ can be seen as a special case of $\mathcal{U}_0(s_{n,p}, M)$.

Let $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{i.i.d.}{\sim} \mathcal{P}(K, \tau_0, s_{n,p}, M)$, where $\mathbf{X}_i = (X_{1i}, \dots, X_{pi})$, then, the sample covariance matrix is given as

$$\hat{\Sigma} := (\hat{\sigma}_{ij})_{p \times p} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T. \quad (3)$$

The following conclusion about the difference between $\hat{\sigma}_{ij}$ and σ_{ij} was acquired by [4].

Lemma 1. For any $\eta \geq 2$ and some $\varepsilon > 0$,

$$\mathbb{P} \left\{ |\hat{\sigma}_{ij} - \sigma_{ij}| \geq \eta \sqrt{(\hat{\theta}_{ij} \log p) / n}, \exists 1 \leq i, j \leq p \right\} = O \left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\varepsilon} \right),$$

where $\hat{\theta}_{ij} = \frac{1}{n} \sum_{k=1}^n (X_{ik} X_{jk} - \hat{\sigma}_{ij})^2$.

2.2. Differential Privacy: Gaussian Mechanism and Post-Processing Property

This paper adopts the differentially private mechanism to protect sensitive data, which requires no significant change in the outcome if a single data point of a dataset varies.

Definition 1 ([9]). Two datasets, \mathbf{X} and $\mathbf{X}' \in \mathcal{X}^p$, are called neighbors if they only differ in only one entry, denoted as $\mathbf{X} \sim \mathbf{X}'$.

A similar explanation can be found in Cai et al. [11].

Definition 2 ([13]). A randomized algorithm \mathcal{M} is (ϵ, δ) -differentially private (DP) if

$$\mathbb{P}(\mathcal{M}(\mathbf{X}) \in \mathcal{O}) \leq e^\epsilon \mathbb{P}(\mathcal{M}(\mathbf{X}') \in \mathcal{O}) + \delta$$

holds for every pair of neighbors, i.e., \mathbf{X} and $\mathbf{X}' \in \mathcal{X}^p$, where \mathcal{O} is any measurable event in the output space of \mathcal{M} .

This paper adopts the Gaussian mechanism to guarantee the algorithm being (ϵ, δ) -DP.

Definition 3 ([16]). For any algorithm f mapping a dataset \mathcal{X}^p to \mathbb{R}^p , the Gaussian mechanism is defined as

$$\mathcal{M}_G(\mathbf{X}) := f(\mathbf{X}) + \mathbf{Y},$$

where $\mathbf{Y} = (Y_1, \dots, Y_p)^T$ with $Y_i \stackrel{i.i.d.}{\sim} N(0, 2(\Delta_2(f)/\epsilon)^2 \log(1.25/\delta))$ and $\Delta_2(f) = \sup_{\mathbf{X} \sim \mathbf{X}'} \|f(\mathbf{X}) - f(\mathbf{X}')\|_{l_2}$.

In fact, post-processing a differentially private algorithm preserves privacy.

Lemma 2. Post-processing property [11]: If \mathcal{M}_1 is a (ϵ, δ) -differentially private algorithm and \mathcal{M}_2 is any deterministic algorithm, then the composition $\mathcal{M}_2(\mathcal{M}_1(\cdot))$ is (ϵ, δ) -differentially private.

3. Two Methods for Privacy Protection

3.1. Differentially Private Algorithm

Firstly, we introduce the (ϵ, δ) -differentially private algorithm, which post-processes the perturbed sample covariance matrix by the modified threshold. After thresholding, eigenvalue selection is conducted to guarantee a positive semi-definite property. For details, see the following description of Algorithm 1. Moreover, we provide Theorem 1 to bound the errors between $\Sigma \in \mathcal{U}_q(s_{n,p}, M)$ and the output $\tilde{\Sigma}_+^\tau$ of Algorithm 1.

Algorithm 1 Modified DP-Thresholding

Input: $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{i.i.d.}{\sim} \mathcal{P}(K, \tau_0, s_{n,p}, M)$, $\epsilon, \delta \in (0, 1)$, $\eta, K, M > 0$ and $\eta \geq 2$

1: Compute

$$\tilde{\Sigma} = (\tilde{\sigma}_{ij})_{1 \leq i, j \leq p} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T + \mathbf{N}, \quad (4)$$

where $\mathbf{N} = (n_{ij})_{1 \leq i, j \leq p}$ is a symmetric matrix for $i \leq j$, $n_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_1^2)$ with $\sigma_1^2 = 4 \log(1.25/\delta) / (n\epsilon)^2$.

2: Define the thresholding estimator $\tilde{\Sigma}^\tau = (\tilde{\sigma}_{ij}^\tau)_{1 \leq i, j \leq p}$ with

$$\tilde{\sigma}_{ij}^\tau = \tilde{\sigma}_{ij} \cdot I \left\{ |\tilde{\sigma}_{ij}| > \frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\epsilon}} M \sqrt{\frac{\log p}{n}} + 4\sigma_1 (\log p)^{1/2} \right\}.$$

3: Compute the eigen-decomposition $\tilde{\Sigma}^\tau = \sum_{i=1}^p \tilde{\lambda}_i v_i v_i^T$ and $\tilde{\lambda}_i^+ = \max\{\tilde{\lambda}_i, 0\}$, then let $\tilde{\Sigma}_+^\tau = \sum_{i=1}^p \tilde{\lambda}_i^+ v_i v_i^T$.

4: **return** $\tilde{\Sigma}_+^\tau$.

In Algorithm 1, the parameters ϵ and δ control the level of privacy protection, and the privacy constraint becomes more stringent as ϵ and δ tend to 0. The parameter η is usually chosen by cross-validation in the simulation experiment in order to decrease the error.

Remark 2. Algorithm 1 is (ϵ, δ) -differentially private. Since Theorem 1 in [17] states that Step 1 in Algorithm 1 preserves (ϵ, δ) -differential privacy, then this, along with Lemma 2, implies that Algorithm 1 is (ϵ, δ) -differentially private. Note that the process of verifying differential privacy is long and elementary [18], so we omit it here.

Theorem 1. For any $\eta \geq 2$ and some $\varepsilon > 0$, there exist positive constants $C_1 := C_1(q, \eta, K, M, \varepsilon)$, $C_2 := C_2(q, \eta, K, M, \varepsilon)$, such that the output $\tilde{\Sigma}_+^\tau$ of Algorithm 1 satisfies

$$\begin{aligned} (i). \quad & \inf_{P \in \mathcal{P}(K, \tau_0, s_{n,p}, M)} \mathbb{P}_{\mathbf{X}|P} \left\{ \|\tilde{\Sigma}_+^\tau - \Sigma\|_2 \leq C_1 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)^{1-q} \right\} \\ & \geq 1 - O\left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\varepsilon} + p^{-\frac{5}{2}}\right); \\ (ii). \quad & \inf_{P \in \mathcal{P}(K, \tau_0, s_{n,p}, M)} \mathbb{P}_{\mathbf{X}|P} \left\{ \frac{1}{\sqrt{p}} \|\tilde{\Sigma}_+^\tau - \Sigma\|_F \leq C_2 \sqrt{s_{n,p}} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)^{1-q/2} \right\} \\ & \geq 1 - O\left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\varepsilon} + p^{-\frac{5}{2}}\right). \end{aligned}$$

Remark 3. In Theorem 1, the condition $\eta \geq 2$ is necessary. Otherwise, $O((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\varepsilon})$ may be larger than 1 for some $\varepsilon > 0$, which leads to results (i)–(ii) being trivial. According to the results in Theorem 1, we find that when the levels of sparsity (i.e., $s_{n,p}$) and the dimension p of the random vector increase and the sample size n decreases, the errors between $\Sigma \in \mathcal{U}_q(s_{n,p}, M)$ and the output $\tilde{\Sigma}_+^\tau$ of Algorithm 1 become larger. Moreover, if ε and δ become smaller, which means stronger privacy protection and noise disturbance, the errors become larger. This theoretical analysis is consistent with the results of the simulation experiments in Section 4.1.

Note that the error bound under the spectral norm over the parameter space $\mathcal{G}_0(s_{n,p})$ (see (1)) in Wang and Xu [10] is $O\left(s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right) + \frac{\log^2 \delta^{-1}}{n^2 \varepsilon^4}\right)$, while guaranteeing DP. Theorem 1(i), together with (2), states that our error bound over $\mathcal{G}_0(s_{n,p}, M)$ is $O\left(s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)\right)$, which performs better than that of Wang and Xu if the space $\mathcal{G}_0(s_{n,p})$ has an extra condition. However, the distribution assumption in our work is weaker than that of Wang and Xu, and we also consider the case with $0 < q < 1$.

Remark 4. In a non-private case, result (i) of Theorem 1 reduces to

$$\inf_{P \in \mathcal{P}(K, \tau_0, s_{n,p}, M)} \mathbb{P}_{\mathbf{X}|P} \left\{ \|\tilde{\Sigma}_+^\tau - \Sigma\|_2 \leq C_1 s_{n,p} \left(\frac{\log p}{n} \right)^{\frac{1-q}{2}} \right\} \geq 1 - O\left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\varepsilon} + p^{-\frac{5}{2}}\right),$$

which is consistent with Theorem 1(ii) in [4], i.e., we extend the work of [4] from a non-private mechanism to a (ε, δ) -differentially private mechanism.

Remark 5. From the proving process of Theorem 1(i), we find that

$$\|\tilde{\Sigma}^\tau - \Sigma\|_1 \leq C_1 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)^{1-q}$$

holds under event E defined by (A1).

On the other hand, the Riesz–Thorin interpolation theorem [19] states that the l_ω operator norm of the symmetric matrix \mathbf{A} satisfies $\|\mathbf{A}\|_\omega \leq \|\mathbf{A}\|_1$. Hence, we obtain

$$\|\tilde{\Sigma}^\tau - \Sigma\|_\omega \leq C_1 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)^{1-q}$$

under event E . This, along with Lemma A1, implies that

$$\begin{aligned} & \inf_{P \in \mathcal{P}(K, \tau_0, s_{n,p}, M)} \mathbb{P}_{\mathbf{X}|P} \left\{ \|\tilde{\Sigma}^\tau - \Sigma\|_\omega \leq C_1 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)^{1-q} \right\} \\ & \geq 1 - O\left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\varepsilon} + p^{-\frac{5}{2}}\right), \end{aligned}$$

i.e., we can provide the upper error bound under the l_ω operator norm of the matrix.

3.2. Locally Differentially Private Algorithm

Differential privacy in the local model. In the locally differentially private (LDP) model, there exists a data universe \mathcal{D} , n users who hold private data $\mathbf{x} \in \mathcal{D}$ and a server. In each round, the server sends a message, each user selects a differentially private algorithm and runs it on their data in order to encrypt it, and then sends the output back to the server.

Definition 4 ([20]). A randomized algorithm \mathcal{M} is (ϵ, δ) -locally differentially private if

$$\mathbb{P}(\mathcal{M}(\mathbf{x}) \in \mathcal{S}) \leq e^\epsilon \mathbb{P}(\mathcal{M}(\mathbf{x}') \in \mathcal{S}) + \delta$$

holds for all pairs of $\mathbf{x}, \mathbf{x}' \in \mathcal{D}$, where \mathcal{S} is any measurable event in the output space of \mathcal{M} .

In fact, the idea of Algorithm 2 is that each \mathbf{X}_i perturbs its sample covariance matrix in order to aggregate the noise disturbance in the algorithm.

Algorithm 2 Modified LDP-Thresholding

Input: $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{i.i.d.}{\sim} \mathcal{P}(K, \tau_0, s_{n,p}, M)$, $\epsilon, \delta \in (0, 1)$, $\epsilon, K, M > 0$ and $\eta \geq 2$

1: **for** each $k \in \{1, \dots, n\}$ **do**
 2:

$$\mathbf{X}_k \mathbf{X}_k^T + \mathbf{R}^k,$$

where $\mathbf{R}^k = (r_{ij}^k)_{1 \leq i, j \leq p}$ is a symmetric matrix and for $i \leq j$, $r_{ij}^k \stackrel{i.i.d.}{\sim} N(0, \sigma_2^2)$ with $\sigma_2^2 = 2 \log(1.25/\delta)/\epsilon^2$.

3: **end for**

4: Compute

$$\check{\Sigma} = (\check{\sigma}_{ij})_{p \times p} = \frac{1}{n} \sum_{k=1}^n (\mathbf{X}_k \mathbf{X}_k^T + \mathbf{R}^k).$$

5: Define the thresholding estimator $\check{\Sigma}^\tau = (\check{\sigma}_{ij}^\tau)_{1 \leq i, j \leq p}$ with

$$\check{\sigma}_{ij}^\tau = \check{\sigma}_{ij} \cdot I \left\{ |\check{\sigma}_{ij}| > \frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\epsilon}} M \sqrt{\frac{\log p}{n}} + 4\sigma_2 (\log p)^{1/2} \right\}.$$

6: Compute the eigen-decomposition $\check{\Sigma}^\tau = \sum_{i=1}^p \check{\lambda}_i v_i v_i^T$ and $\check{\lambda}_i^+ = \max\{\check{\lambda}_i, 0\}$, then let $\check{\Sigma}_+^\tau = \sum_{i=1}^p \check{\lambda}_i^+ v_i v_i^T$.

7: **return** $\check{\Sigma}_+^\tau$.

Remark 6. In fact, Step 2 in Algorithm 2 can be viewed as a special case of Algorithm 1 in [21] that keeps (ϵ, δ) -locally differentially private. Then, using the post-processing property leads to Algorithm 2 being (ϵ, δ) -locally differentially private. Here, we omit Step 2 to verify the preservation of local differential privacy.

Theorem 2. For any $\eta \geq 2$ and some $\epsilon > 0$, there exist positive constants $C_3 := C_3(q, \eta, K, M, \epsilon)$, $C_4 := C_4(q, \eta, K, M, \epsilon)$, such that the output $\check{\Sigma}_+^\tau$ of Algorithm 2 satisfies

$$\begin{aligned} (i). \quad & \inf_{P \in \mathcal{P}(K, \tau_0, s_{n,p}, M)} \mathbb{P}_{\mathbf{X}|P} \left\{ \|\check{\Sigma}_+^\tau - \Sigma\|_2 \leq C_3 s_{n,p} \left(\frac{\log \delta^{-1} \log p}{n \epsilon^2} \right)^{\frac{1-q}{2}} \right\} \\ & \geq 1 - O \left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\epsilon} + p^{-\frac{5}{2}} \right); \end{aligned}$$

$$(ii). \inf_{P \in \mathcal{P}(K, \tau_0, s_{n,p}, M)} \mathbb{P}_{\mathbf{X}|P} \left\{ \frac{1}{\sqrt{p}} \|\tilde{\Sigma}_+^{\tau} - \Sigma\|_F \leq C_4 \sqrt{s_{n,p}} \left(\frac{\log \delta^{-1} \log p}{n\epsilon^2} \right)^{\frac{1}{2} - \frac{q}{4}} \right\} \\ \geq 1 - O\left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\epsilon} + p^{-\frac{5}{2}}\right).$$

Remark 7. The reason for setting $\eta \geq 2$ and $\epsilon > 0$ is the same as in Theorem 1. Moreover, the parameters $s_{n,p}$ and p are positively correlated with error bounds, and n , ϵ and δ are negatively correlated with error bounds, respectively, and these results are consistent with the simulation studies in Section 4.1.

It is easy to see that the error bound under the spectral norm over $\mathcal{G}_0(s_{n,p})$ in Wang and Xu [10] is $O\left(s_{n,p} \sqrt{\frac{\log \delta^{-1} \log p}{n\epsilon^2}}\right)$ while maintaining LDP. Moreover, our Theorem 2(i) and (2) imply that the error bound under the spectral norm over $\mathcal{G}_0(s_{n,p}, M)$ differs with that of Wang and Xu only in theory. However, according to the numerical studies in Section 4, the output of our Algorithm 2 outperforms that of Wang and Xu. In addition, our assumption of the distribution is weaker than Wang and Xu's condition, and the parameter space with $0 < q < 1$ is considered in our work.

Remark 8. According to the proving process of Theorem 2, we observe that the error bounds under spectral and Frobenius norms can be $O\left(s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n\epsilon^2}}\right)^{1-q}\right)$ and $O\left(\sqrt{s_{n,p}} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n\epsilon^2}}\right)^{1-q/2}\right)$, respectively. Comparing these results with the conclusions of Theorem 1, it is easy to see that the upper bound of Theorem 2 is larger than that of Theorem 1. This is reasonable since Algorithm 2 enforces a stronger privacy protection mechanism, so it has to pay a larger price.

4. Numerical Experiments

4.1. Simulation Studies

In this subsection, we investigate the numerical performances of the outputs $\tilde{\Sigma}_+^{\tau}$ and $\tilde{\Sigma}_+^{\tau}$ of Algorithms 1 and 2, respectively. Moreover, we also compare these two estimators with the outputs $\tilde{\Sigma}_1^+$ and $\tilde{\Sigma}_2^+$ of the DP-thresholding and LDP-thresholding algorithms proposed by Wang and Xu [10], respectively.

Data generation

The following models for the covariance matrix $\Sigma = (\sigma_{ij})_{p \times p}$ are considered:

Model 1. $\sigma_{ij} = 0.6^{|i-j|}$;

Model 2. $\sigma_{ii} = 1$, $\sigma_{i,i+1} = \sigma_{i+1,i} = 0.6$, $\sigma_{i,i+2} = \sigma_{i+2,i} = 0.3$ and $\sigma_{ij} = 0$ for $|i-j| \geq 3$.

Then, we generate samples $\mathbf{X}_1, \dots, \mathbf{X}_n$ in two different ways:

(i) $\mathbf{X}_k \stackrel{i.i.d.}{\sim} N(0, \Sigma_{p \times p})$;

(ii) \mathbf{X}_k are independent from multivariate t -distribution $T_p(\alpha; 0, \Sigma_{p \times p})$ with freedom $\alpha = 5$.

Experimental settings

(i) Set $p = \{50, 100, 200\}$, $n = \{200, 300\}$, $\epsilon = 1/2$, $\epsilon = 0.5$ and $\delta = 1/400$;

(ii) Set $p = 100$, $n = 200$, $\epsilon = 1/2$, $\delta = 1/400$ and $\epsilon = \{0.1, 0.4, 0.6, 0.7\}$;

(iii) Set $p = 100$, $n = 200$, $\epsilon = 1/2$, $\epsilon = 0.5$ and $\delta = \{1/50, 1/100, 1/200, 1/500\}$.

In each setting, we choose the tuning parameter η in the threshold by 10-fold cross-validation, as proposed by Cai and Liu [4]. Moreover, we measure the errors by the spectral and Frobenius norms, respectively. We run each experiment 50 times and take the average errors as the final result (with standard errors in parentheses).

Experimental results

From Tables 1 and 2, we find that the output $\tilde{\Sigma}_+^{\tau}$ of Algorithm 1 (differentially private algorithm) performs better than the output $\tilde{\Sigma}_1^+$ of the DP-thresholding algorithm proposed by Wang and Xu [10]. The above property also holds for our Algorithm 2 (locally differentially private algorithm) compared with the LDP-thresholding algorithm of Wang and Xu. Moreover, we observe that the dimension p increase leads to the errors becoming larger. Meanwhile, if the sample size n becomes larger, then the errors become smaller. It can

also be seen that the errors under Model 1 are larger than those under Model 2, since the covariance matrix in Model 1 is more dense.

Tables 3 and 4 demonstrate that if the privacy parameter ϵ increases (resulting in weaker privacy protection), the errors will become smaller. Moreover, Tables 5 and 6 show that when the privacy parameter δ decreases, the errors will become larger. These phenomena are also consistent with our theoretical results. In addition, the following numerical results are in line with the theoretical findings indicated by Remark 8, i.e., the error estimation of Algorithm 1 is smaller than that of Algorithm 2.

Table 1. Comparison of different outputs of the algorithms with normal distributions.

| (p, n) | Spectral Norm | | | | Frobenius Norm | | | |
|-----------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|----------------------|
| | $\tilde{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_1^+$ | $\check{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_2^+$ | $\tilde{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_1^+$ | $\check{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_2^+$ |
| Model 1 | | | | | | | | |
| (50,200) | 1.92 (0.12) | 2.50 (0.14) | 4.31 (0.23) | 4.95 (0.48) | 4.41 (0.17) | 5.90 (0.31) | 8.15 (0.36) | 9.77 (0.40) |
| (50,300) | 1.52 (0.13) | 1.89 (0.15) | 3.70 (0.29) | 4.18 (0.25) | 3.74 (0.16) | 4.65 (0.13) | 6.58 (0.30) | 7.65 (0.44) |
| (100,200) | 2.13 (0.13) | 2.79 (0.08) | 5.44 (0.29) | 7.67 (0.43) | 6.83 (0.17) | 9.34 (0.23) | 10.71 (0.25) | 14.11 (0.39) |
| (100,300) | 1.76 (0.08) | 2.18 (0.11) | 4.73 (0.17) | 6.73 (0.43) | 5.86 (0.17) | 7.13 (0.26) | 8.81 (0.37) | 11.07 (0.28) |
| (200,300) | 1.89 (0.07) | 2.56 (0.14) | 6.08 (0.25) | 8.81 (0.41) | 8.73 (0.10) | 10.10 (0.23) | 11.68 (0.27) | 15.28 (0.19) |
| Model 2 | | | | | | | | |
| (50,200) | 1.01 (0.16) | 1.56 (0.15) | 3.46 (0.29) | 4.11 (0.29) | 3.32 (0.21) | 4.90 (0.28) | 6.42 (0.29) | 8.95 (0.46) |
| (50,300) | 0.74 (0.16) | 0.92 (0.15) | 3.13 (0.27) | 3.51 (0.24) | 2.87 (0.06) | 3.23 (0.13) | 5.20 (0.26) | 7.16 (0.37) |
| (100,200) | 1.28 (0.16) | 1.78 (0.06) | 4.19 (0.28) | 5.04 (0.23) | 4.99 (0.15) | 8.07 (0.31) | 8.03 (0.33) | 13.35 (0.30) |
| (100,300) | 0.82 (0.13) | 1.43 (0.09) | 3.60 (0.26) | 4.30 (0.18) | 4.29 (0.07) | 5.30 (0.25) | 5.75 (0.41) | 9.53 (0.39) |
| (200,300) | 0.93 (0.12) | 1.63 (0.11) | 4.00 (0.17) | 5.35 (0.17) | 6.28 (0.09) | 8.93 (0.34) | 8.94 (0.31) | 13.84 (0.24) |

Table 2. Comparison of different outputs of the algorithms with t -distributions.

| (p, n) | Spectral Norm | | | | Frobenius Norm | | | |
|-----------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|----------------------|
| | $\tilde{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_1^+$ | $\check{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_2^+$ | $\tilde{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_1^+$ | $\check{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_2^+$ |
| Model 1 | | | | | | | | |
| (50,200) | 4.48 (0.10) | 5.31 (0.10) | 8.64 (0.26) | 10.45 (0.13) | 9.44 (0.21) | 11.41 (0.20) | 13.75 (0.26) | 15.44 (0.42) |
| (50,300) | 3.69 (0.20) | 4.63 (0.08) | 7.79 (0.48) | 9.36 (0.37) | 7.95 (0.28) | 9.80 (0.14) | 12.74 (0.46) | 13.63 (0.33) |
| (100,200) | 4.81 (0.08) | 5.56 (0.04) | 9.98 (0.35) | 13.75 (0.12) | 14.10 (0.25) | 16.78 (0.09) | 20.32 (0.28) | 22.97 (0.28) |
| (100,300) | 4.35 (0.09) | 5.08 (0.10) | 8.73 (0.35) | 11.69 (0.44) | 12.53 (0.27) | 14.84 (0.21) | 18.72 (0.48) | 19.42 (0.32) |
| (200,300) | 4.59 (0.04) | 5.42 (0.03) | 10.62 (0.30) | 14.61 (0.31) | 18.91 (0.21) | 20.59 (0.24) | 23.09 (0.20) | 24.10 (0.14) |
| Model 2 | | | | | | | | |
| (50,200) | 2.81 (0.27) | 4.23 (0.44) | 7.34 (0.48) | 9.42 (0.40) | 6.06 (0.33) | 7.29 (0.28) | 10.85 (0.44) | 11.78 (0.27) |
| (50,300) | 2.27 (0.21) | 3.35 (0.42) | 6.19 (0.49) | 7.47 (0.43) | 4.86 (0.26) | 5.88 (0.30) | 9.53 (0.35) | 10.77 (0.42) |
| (100,200) | 3.91 (0.36) | 4.61 (0.39) | 8.46 (0.27) | 10.95 (0.30) | 9.68 (0.25) | 13.17 (0.40) | 14.58 (0.21) | 16.69 (0.32) |
| (100,300) | 2.94 (0.17) | 3.73 (0.26) | 6.69 (0.39) | 8.52 (0.46) | 7.63 (0.29) | 10.83 (0.31) | 11.95 (0.37) | 13.65 (0.20) |
| (200,300) | 3.56 (0.27) | 4.46 (0.34) | 9.36 (0.24) | 11.48 (0.22) | 12.25 (0.36) | 15.81 (0.25) | 15.46 (0.32) | 18.42 (0.22) |

Table 3. Comparison of different privacy levels ϵ with normal distributions under Model 2.

| ϵ | Spectral Norm | | | | Frobenius Norm | | | |
|------------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|----------------------|---------------------------|----------------------|
| | $\tilde{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_1^+$ | $\check{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_2^+$ | $\tilde{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_1^+$ | $\check{\Sigma}_+^{\tau}$ | $\tilde{\Sigma}_2^+$ |
| Model 1 | | | | | | | | |
| 0.1 | 2.60 (0.06) | 3.04 (0.02) | 7.49 (0.23) | 9.84 (0.22) | 7.69 (0.21) | 11.14 (0.13) | 12.74 (0.26) | 17.54 (0.35) |
| 0.4 | 1.51 (0.15) | 2.12 (0.06) | 5.25 (0.33) | 6.88 (0.38) | 5.95 (0.29) | 9.11 (0.17) | 9.82 (0.46) | 14.72 (0.32) |
| 0.6 | 1.08 (0.14) | 1.70 (0.07) | 3.31 (0.25) | 4.07 (0.17) | 4.66 (0.14) | 7.64 (0.31) | 7.67 (0.09) | 12.67 (0.15) |
| 0.7 | 1.03 (0.16) | 1.65 (0.10) | 3.03 (0.14) | 3.34 (0.16) | 4.45 (0.11) | 7.02 (0.33) | 7.16 (0.04) | 11.66 (0.06) |

Table 4. Comparison of different privacy levels ϵ with t -distributions under Model 2.

| ϵ | Spectral Norm | | | | Frobenius Norm | | | |
|------------|-----------------------------|------------------------|-----------------------------|------------------------|-----------------------------|------------------------|-----------------------------|------------------------|
| | $\tilde{\Sigma}_{+}^{\tau}$ | $\tilde{\Sigma}_1^{+}$ | $\check{\Sigma}_{+}^{\tau}$ | $\check{\Sigma}_2^{+}$ | $\tilde{\Sigma}_{+}^{\tau}$ | $\tilde{\Sigma}_1^{+}$ | $\check{\Sigma}_{+}^{\tau}$ | $\check{\Sigma}_2^{+}$ |
| | Model 1 | | | | | | | |
| 0.1 | 6.98 (0.28) | 8.29 (0.28) | 15.03 (0.16) | 17.59 (0.17) | 12.99 (0.21) | 16.58 (0.47) | 19.70 (0.35) | 22.18 (0.24) |
| 0.4 | 4.71 (0.39) | 5.72 (0.36) | 10.37 (0.42) | 12.88 (0.31) | 10.78 (0.43) | 14.34 (0.45) | 16.14 (0.18) | 18.16 (0.14) |
| 0.6 | 3.02 (0.30) | 3.82 (0.37) | 6.90 (0.39) | 9.15 (0.21) | 8.89 (0.40) | 12.39 (0.42) | 13.25 (0.47) | 15.38 (0.22) |
| 0.7 | 1.84 (0.41) | 2.55 (0.31) | 4.54 (0.19) | 6.28 (0.24) | 7.38 (0.41) | 10.65 (0.39) | 11.52 (0.30) | 13.22 (0.44) |

Table 5. Comparison of different privacy levels δ with normal distributions under Model 1.

| ϵ | Spectral Norm | | | | Frobenius Norm | | | |
|------------|-----------------------------|------------------------|-----------------------------|------------------------|-----------------------------|------------------------|-----------------------------|------------------------|
| | $\tilde{\Sigma}_{+}^{\tau}$ | $\tilde{\Sigma}_1^{+}$ | $\check{\Sigma}_{+}^{\tau}$ | $\check{\Sigma}_2^{+}$ | $\tilde{\Sigma}_{+}^{\tau}$ | $\tilde{\Sigma}_1^{+}$ | $\check{\Sigma}_{+}^{\tau}$ | $\check{\Sigma}_2^{+}$ |
| | Model 1 | | | | | | | |
| 1/50 | 1.85 (0.14) | 2.47 (0.11) | 4.08 (0.21) | 4.92 (0.22) | 6.37 (0.16) | 8.67 (0.24) | 6.33 (0.31) | 8.62 (0.34) |
| 1/100 | 1.99 (0.14) | 2.58 (0.09) | 4.49 (0.19) | 5.51 (0.27) | 6.47 (0.47) | 8.89 (0.36) | 7.66 (0.26) | 11.14 (0.15) |
| 1/200 | 2.08 (0.11) | 2.62 (0.09) | 4.92 (0.27) | 6.43 (0.45) | 6.68 (0.21) | 9.12 (0.29) | 9.17 (0.18) | 13.09 (0.17) |
| 1/500 | 2.32 (0.14) | 2.96 (0.08) | 5.63 (0.31) | 7.97 (0.25) | 7.04 (0.17) | 9.66 (0.22) | 11.15 (0.17) | 15.68 (0.21) |

Table 6. Comparison of different privacy levels δ with t -distributions under Model 1.

| ϵ | Spectral Norm | | | | Frobenius Norm | | | |
|------------|-----------------------------|------------------------|-----------------------------|------------------------|-----------------------------|------------------------|-----------------------------|------------------------|
| | $\tilde{\Sigma}_{+}^{\tau}$ | $\tilde{\Sigma}_1^{+}$ | $\check{\Sigma}_{+}^{\tau}$ | $\check{\Sigma}_2^{+}$ | $\tilde{\Sigma}_{+}^{\tau}$ | $\tilde{\Sigma}_1^{+}$ | $\check{\Sigma}_{+}^{\tau}$ | $\check{\Sigma}_2^{+}$ |
| | Model 1 | | | | | | | |
| 1/50 | 4.40 (0.07) | 5.29 (0.06) | 6.95 (0.35) | 10.31 (0.28) | 13.42 (0.19) | 16.22 (0.21) | 17.94 (0.40) | 19.17 (0.36) |
| 1/100 | 4.64 (0.06) | 5.38 (0.05) | 8.01 (0.29) | 11.24 (0.14) | 13.69 (0.20) | 16.39 (0.13) | 18.70 (0.37) | 20.56 (0.39) |
| 1/200 | 5.73 (0.08) | 5.45 (0.03) | 8.91 (0.35) | 13.00 (0.21) | 13.94 (0.18) | 16.55 (0.13) | 19.10 (0.18) | 21.96 (0.28) |
| 1/500 | 5.05 (0.04) | 5.78 (0.02) | 10.51 (0.39) | 14.47 (0.26) | 14.33 (0.17) | 16.93 (0.07) | 20.66 (0.38) | 23.37 (0.17) |

4.2. Real Data Application

To demonstrate the performances of the outputs $\tilde{\Sigma}_{+}^{\tau}$ and $\check{\Sigma}_{+}^{\tau}$ of Algorithms 1 and 2, and compare them with the outputs $\tilde{\Sigma}_1^{+}$ and $\check{\Sigma}_2^{+}$ of the DP-thresholding and LDP-thresholding algorithms proposed by Wang and Xu [10], respectively, we use quadratic discriminant analysis (QDA) presented by Liang et al. [22]. We apply these estimators on the human gut microbiome dataset collected by Wu et al. [23], which contains 98 healthy individuals at the University of Pennsylvania. In fact, the bacterial community in the dataset was categorized into 87 genera, which appeared in at least one sample. We select $p = 40$ bacterial genera that appeared in at least four samples. Then, according to the body mass index (BMI), we divide the dataset with 40 bacterial genera into a lean group ($\text{BMI} < 25$, $n = 63$) and an obese group ($\text{BMI} \geq 25$, $n = 35$).

We randomly select 13 lean subjects and 7 obese subjects via the stratified sampling approach in order to constitute the testing set (roughly 1/5 of the subjects in each group); the remaining subjects form the training set. Moreover, a two-sample t test is performed between the two groups for each bacterial genus in the training set. In the our analysis, $p = 25$ and 40 are considered. The data are assumed to be normally distributed as $N(\mu_k, \Sigma)$, and the two groups are assumed to have the same covariance matrix Σ but different means μ_k , $k = 1$ for the obese group and $k = 2$ for the lean group. Note that the parameters δ , ϵ of the additive noise in Algorithms 1 and 2 are set as 1/196 and 0.5, respectively. In addition, we set $\theta = 4\sqrt{2}\eta K^{\frac{2}{4+\epsilon}} M/3$ in Algorithms 1 and 2 as the tuning parameter, which is selected by cross-validation conducted on the real data.

In order to compare the advantage, the outputs $\tilde{\Sigma}_+^\tau, \tilde{\Sigma}_+^\tau, \tilde{\Sigma}_1^+$ and $\tilde{\Sigma}_2^+$ are used to replace $\hat{\Sigma}$ in the following QDA score functions:

$$\delta_k(y) = -\frac{1}{2} \log(\det(\hat{\Sigma})) - \frac{1}{2} (y - \hat{\mu}_k)^T \hat{\Sigma}^{-1} (y - \hat{\mu}_k) + \log(\hat{\pi}_k),$$

where $\hat{\pi}_k = n_k/n$ is the proportion of group k subjects in the training set and $\hat{\mu}_k = \frac{1}{n_k} \sum_{kth\text{-group}} y_i$ is the within-group average vector in the training set. The classification rule is

$$\hat{k}(y) = \arg \max_k \delta_k(y), k = 1, 2.$$

We use the testing set to evaluate the estimation performance through the average number of misclassifications. Our result is based on 20 replications of the above procedure. Table 7 presents the average number of misclassifications for every output of the corresponding algorithms. It can be seen that the output $\tilde{\Sigma}_+^\tau$ of our Algorithm 1 performs better than the output $\tilde{\Sigma}_1^+$ of the DP-thresholding algorithm proposed by Wang and Xu [10]. Similarly, the output $\tilde{\Sigma}_+^\tau$ of Algorithm 2 is more effective than the output $\tilde{\Sigma}_2^+$ of the LDP-thresholding algorithm [10].

Table 7. Comparison of different outputs of algorithms.

| p | Average Number of Misclassifications | | | |
|-----|--------------------------------------|----------------------|-------------------------|----------------------|
| | $\tilde{\Sigma}_+^\tau$ | $\tilde{\Sigma}_1^+$ | $\tilde{\Sigma}_+^\tau$ | $\tilde{\Sigma}_2^+$ |
| 25 | 2.25 | 5.70 | 4.30 | 6.45 |
| 40 | 3.20 | 4.95 | 5.15 | 6.90 |

5. Discussion

In this paper, we study the problem of estimating the sparse covariance matrix under lower-bounded moment assumption while guaranteeing DP and LDP, and measure the error bounds by spectral and Frobenius norms, respectively. Furthermore, we conduct numerical experiments to support our theoretical analysis.

However, we do not derive the lower bound of the error estimation. This is possible to achieve, since there are many well-known technique methods that can be utilized, such as Fano's lemma [24], Assuouad's lemma [25] and Le Cam's method [26]. Moreover, a noteworthy open problem is to consider the estimation in terms of expectation with the lower-bounded moment assumption while preserving DP and LDP. We leave these aspects as future work.

Author Contributions: H.L. conceived the study and wrote the original draft. J.W. conducted the theoretical analysis and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is supported by the National Natural Science Foundation of China (No. 12171016).

Data Availability Statement: The data presented in this study are available within this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In order to present Theorems 1 and 2, we introduce the following lemma:

Lemma A1. Define event E as

$$E = \left\{ |\tilde{\sigma}_{ij} - \sigma_{ij}| \leq \frac{3}{4} \left(\xi \sqrt{\frac{\log p}{n}} + 4\sigma_1 (\log p)^{1/2} \right), \forall 1 \leq i, j \leq p \right\}, \quad (\text{A1})$$

and $\xi := \frac{4\sqrt{2}}{3}\eta K^{\frac{2}{4+\varepsilon}}M$, where K, γ and M are given in Conditions (i)–(ii). Then, for any $\eta \geq 2$ and some $\varepsilon > 0$,

$$\mathbb{P}(E) \geq 1 - O\left((\log p)^{-\frac{1}{2}}p^{-\eta+2} + n^{-\varepsilon} + p^{-\frac{5}{2}}\right).$$

Proof. Note that (3) and (4) in Algorithm 1 show that

$$\tilde{\sigma}_{ij} = \hat{\sigma}_{ij} + n_{ij},$$

thus, we know

$$\begin{aligned} \mathbb{P}(E^c) &= \mathbb{P}\left\{|\hat{\sigma}_{ij} + n_{ij} - \sigma_{ij}| \geq \sqrt{2}\eta K^{\frac{2}{4+\varepsilon}}M\sqrt{\frac{\log p}{n}} + 3\sigma_1\sqrt{\log p}, \exists 1 \leq i, j \leq p\right\} \\ &\leq \mathbb{P}\left\{|\hat{\sigma}_{ij} - \sigma_{ij}| \geq \sqrt{2}\eta K^{\frac{2}{4+\varepsilon}}M\sqrt{\frac{\log p}{n}}, \exists 1 \leq i, j \leq p\right\} \\ &\quad + \mathbb{P}\{|n_{ij}| \geq 3\sigma_1\sqrt{\log p}, \exists 1 \leq i, j \leq p\} := R_1 + R_2. \end{aligned} \quad (\text{A2})$$

Since $n_{ij} \sim N(0, \sigma_1^2)$, it holds that

$$\mathbb{P}\{|n_{ij}| \geq 3\sigma_1\sqrt{\log p}, \exists 1 \leq i, j \leq p\} \leq 2p^{-\frac{5}{2}} \quad (\text{A3})$$

due to $\mathbb{P}\{|X| > t\} \leq 2e^{-\frac{t^2}{2\sigma^2}}$ for $X \sim N(0, \sigma^2)$ in [27].

According to Condition (i), we have $\mathbb{E}|X_i|^{4+\varepsilon} \leq K\sigma_{ii}^{2+\varepsilon/2}$. Therefore,

$$\mathbb{D}(X_i X_j) \leq \mathbb{E}|X_i X_j|^2 \leq \sqrt{\mathbb{E}|X_i|^4 \mathbb{E}|X_j|^4} \leq K^{\frac{4}{4+\varepsilon}}\sigma_{ii}\sigma_{jj}.$$

By applying the $\Sigma \in \mathcal{U}_q(s_{n,p}, M)$ in Condition (ii), it follows that

$$\mathbb{D}(X_i X_j) \leq K^{\frac{4}{4+\varepsilon}}M^2. \quad (\text{A4})$$

On the other hand, Condition (i) implies that

$$\mathbb{D}(X_i X_j) \geq \tau_0\sigma_{ii}\sigma_{jj}. \quad (\text{A5})$$

Define event E_1 as

$$E_1 = \{\hat{\theta}_{ij} \leq 2\mathbb{D}(X_i X_j), \forall 1 \leq i, j \leq p\}$$

and

$$\begin{aligned} \mathbb{P}(E_1^c) &\leq \mathbb{P}\{|\hat{\theta}_{ij} - \mathbb{D}(X_i X_j)| \geq \mathbb{D}(X_i X_j), \exists 1 \leq i, j \leq p\} \\ &\leq \mathbb{P}\{|\hat{\theta}_{ij} - \mathbb{D}(X_i X_j)| \geq \tau_0\sigma_{ii}\sigma_{jj}, \exists 1 \leq i, j \leq p\} \end{aligned}$$

holds thanks to (A5). Using Lemma 2(ii) in [4], we know that for some $\varepsilon > 0$,

$$\mathbb{P}(E_1^c) \leq O(n^{-\varepsilon}). \quad (\text{A6})$$

Hence, we obtain that for any $\eta \geq 2$ and some $\varepsilon > 0$,

$$\begin{aligned} R_1 &\leq \mathbb{P}\left\{|\hat{\sigma}_{ij} + n_{ij} - \sigma_{ij}| \geq \sqrt{2}\eta\sqrt{(\mathbb{D}(X_i X_j)\log p)/n}, \exists 1 \leq i, j \leq p\right\} \\ &\leq \mathbb{P}\left(\left\{|\hat{\sigma}_{ij} + n_{ij} - \sigma_{ij}| \geq \sqrt{2}\eta\sqrt{(\mathbb{D}(X_i X_j)\log p)/n}, \exists 1 \leq i, j \leq p\right\} \cap E_1\right) + \mathbb{P}(E_1^c) \\ &\leq O\left((\log p)^{-\frac{1}{2}}p^{-\eta+2} + n^{-\varepsilon}\right), \end{aligned}$$

where the first inequality holds because of (A4), and the last inequality follows from Lemma 1 and (A6). \square

Proof of Theorem 1. (i) Considering that Σ is a positive definite matrix, we have

$$\begin{aligned}\|\tilde{\Sigma}^{\tau} - \Sigma\|_2 &\leq \|\tilde{\Sigma}^{\tau} - \tilde{\Sigma}^{\tau}\|_2 + \|\tilde{\Sigma}^{\tau} - \Sigma\|_2 \leq \max_{\{\tilde{\lambda}_i \leq 0\}} |\tilde{\lambda}_i| + \|\tilde{\Sigma}^{\tau} - \Sigma\|_2 \\ &\leq \max_{\{\tilde{\lambda}_i \leq 0\}} |\tilde{\lambda}_i - \lambda_i(\Sigma)| + \|\tilde{\Sigma}^{\tau} - \Sigma\|_2 \leq 2\|\tilde{\Sigma}^{\tau} - \Sigma\|_2.\end{aligned}$$

Hence, it suffices to show that there exists the constant $C_1 := C_1(q, \eta, K, M, \varepsilon)$, such that

$$\|\tilde{\Sigma}^{\tau} - \Sigma\|_2 \leq C_1 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \varepsilon^2}} \right)^{1-q} \quad (\text{A7})$$

under event E .

Since $\tilde{\Sigma}^{\tau} - \Sigma$ is symmetric, this, along with the Gersgorin theorem, yields

$$\|\tilde{\Sigma}^{\tau} - \Sigma\|_2 \leq \|\tilde{\Sigma}^{\tau} - \Sigma\|_1 \leq \|\tilde{\Sigma}^{\tau} - \Sigma^{\tau}\|_1 + \|\Sigma^{\tau} - \Sigma\|_1 := R_1 + R_2, \quad (\text{A8})$$

where $\Sigma^{\tau} = (\sigma_{ij}^{\tau})_{p \times p}$ with $\sigma_{ij}^{\tau} = \sigma_{ij} \cdot I\{|\sigma_{ij}| > \lambda\}$ and

$$\lambda = \frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\varepsilon}} M \sqrt{\frac{\log p}{n}} + 4\sigma_1 (\log p)^{1/2}. \quad (\text{A9})$$

We first estimate R_2 . Clearly, it follows that

$$\begin{aligned}R_2 &= \max_j \sum_{i=1}^p [|\sigma_{ij}^{\tau} - \sigma_{ij}| I(|\sigma_{ij}| \leq \lambda) + |\sigma_{ij}^{\tau} - \sigma_{ij}| I(|\sigma_{ij}| > \lambda)] \\ &= \max_j \sum_{i=1}^p |\sigma_{ij}| I(|\sigma_{ij}| \leq \lambda) \leq \max_j \sum_{i=1}^p |\sigma_{ij}|^q \lambda^{1-q}\end{aligned} \quad (\text{A10})$$

from the definition of σ_{ij}^{τ} .

The remaining work sets out to estimate R_1 . It is easy to see that

$$\begin{aligned}R_1 &\leq \max_j \sum_{i=1}^p |\tilde{\sigma}_{ij} - \sigma_{ij}| I(|\tilde{\sigma}_{ij}| > \lambda, |\sigma_{ij}| > \lambda) + \max_j \sum_{i=1}^p |\sigma_{ij}| I(|\tilde{\sigma}_{ij}| \leq \lambda, |\sigma_{ij}| > \lambda) \\ &\quad + \max_j \sum_{i=1}^p |\tilde{\sigma}_{ij}| I(|\tilde{\sigma}_{ij}| > \lambda, |\sigma_{ij}| \leq \lambda) = R_{11} + R_{12} + R_{13}.\end{aligned} \quad (\text{A11})$$

When event E occurs, we obtain

$$R_{11} \leq \frac{3}{4} \max_j \sum_{i=1}^p \lambda I(|\tilde{\sigma}_{ij}| > \lambda, |\sigma_{ij}| > \lambda) \leq \frac{3}{4} \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q \quad (\text{A12})$$

and

$$\begin{aligned}R_{12} &\leq \max_j \sum_{i=1}^p |\tilde{\sigma}_{ij} - \sigma_{ij}| I(|\sigma_{ij}| > \lambda) + \max_j \sum_{i=1}^p |\tilde{\sigma}_{ij}| I(|\tilde{\sigma}_{ij}| \leq \lambda, |\sigma_{ij}| > \lambda) \\ &\leq \frac{7}{4} \max_j \sum_{i=1}^p \lambda I(|\sigma_{ij}| > \lambda) \leq \frac{7}{4} \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q.\end{aligned} \quad (\text{A13})$$

For R_{13} , it holds that

$$\begin{aligned}
R_{13} &\leq \max_j \sum_{i=1}^p |\sigma_{ij}| I(|\sigma_{ij}| \leq \lambda) + \max_j \sum_{i=1}^p |\tilde{\sigma}_{ij} - \sigma_{ij}| I(|\tilde{\sigma}_{ij}| > \lambda, |\sigma_{ij}| \leq \lambda) \\
&\leq \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q + R_{131}.
\end{aligned} \tag{A14}$$

Through the condition $\delta \in (0, 1)$ in Algorithm 1, we know that constant $c_1 > 0$ exists, such that $\log(1.25/\delta) \leq c_1 \log \delta^{-1}$ and $\sigma_1 \leq 2\sqrt{c_1 \log \delta^{-1}/n\epsilon}$. This, along with the definition of λ in (A9), leads to

$$\lambda \leq \frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\epsilon}} M \sqrt{\frac{\log p}{n}} + 8\sqrt{c_1} \sqrt{\frac{\log \delta^{-1}}{\epsilon^2}} \sqrt{\frac{\log p}{n^2}}. \tag{A15}$$

On the other hand, we obtain

$$\begin{aligned}
R_{131} &\leq \frac{3}{4} \max_j \sum_{i=1}^p \lambda I\left(|\tilde{\sigma}_{ij}| > \lambda, \frac{1}{4}\lambda \leq |\sigma_{ij}| < \lambda\right) + \frac{3}{4} \max_j \sum_{i=1}^p \lambda I\left(|\tilde{\sigma}_{ij}| > \lambda, |\sigma_{ij}| < \frac{1}{4}\lambda\right) \\
&\leq \frac{3}{4} \max_j \sum_{i=1}^p \lambda I\left(\frac{1}{4}\lambda \leq |\sigma_{ij}| < \lambda\right) + \frac{3}{4} \max_j \sum_{i=1}^p \lambda I\left(|\tilde{\sigma}_{ij} - \sigma_{ij}| > \frac{3}{4}\lambda\right) \\
&= T_1 + T_2
\end{aligned} \tag{A16}$$

under event E .

Obviously, T_1 in (A16) is controlled by $4^q \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q$. For T_2 , (A15) and the assumption in Condition (i) imply λ is a bounded quantity. In addition, $I(|\tilde{\sigma}_{ij} - \sigma_{ij}| > \frac{3}{4}\lambda) = 0$ if event E occurs. Thus, T_2 in (A16) is 0 under event E . Therefore,

$$R_{131} \leq 4^q \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q.$$

This, along with (A14), reveals that

$$R_{13} \leq (4^q + 1) \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q \tag{A17}$$

under event E . Then, we have

$$R_1 \leq \left(4^q + \frac{7}{2}\right) \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q, \tag{A18}$$

thanks to (A11)–(A13) and (A17).

Furthermore, substituting (A18) and (A10) into (A8) indicates that

$$\|\tilde{\Sigma}^\tau - \Sigma\|_2 \leq \left(4^q + \frac{9}{2}\right) \max_j \sum_{i=1}^p \lambda^{1-q} |\sigma_{ij}|^q.$$

Hence, we derive

$$\begin{aligned}
\|\tilde{\Sigma}^\tau - \Sigma\|_2 &\leq \left(4^q + \frac{9}{2}\right) \left(\frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\epsilon}} M \sqrt{\frac{\log p}{n}} + 8\sqrt{c_1} \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \epsilon^2}}\right)^{1-q} \max_j \sum_{i=1}^p |\sigma_{ij}|^q \\
&\leq C_{1s_n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \epsilon^2}}\right)^{1-q}
\end{aligned}$$

due to (A15) and $\Sigma \in \mathcal{U}_q(s_{n,p}, M)$, which reaches the conclusion of (A7).

(ii) Similar to the discussions in (i), we derive that

$$\begin{aligned} \|\tilde{\Sigma}_+^\tau - \Sigma\|_F^2 &\leq 2(\|\tilde{\Sigma}_+^\tau - \tilde{\Sigma}^\tau\|_F^2 + \|\tilde{\Sigma}^\tau - \Sigma\|_F^2) \leq 2\left(\sum_{i:\tilde{\lambda}_i \leq 0} \tilde{\lambda}_i^2 + \|\tilde{\Sigma}^\tau - \Sigma\|_F^2\right) \\ &\leq 2\left\{\sum_{i=1}^p [\tilde{\lambda}_i - \lambda_i(\Sigma)]^2 + \|\tilde{\Sigma}^\tau - \Sigma\|_F^2\right\} = 4\|\tilde{\Sigma}^\tau - \Sigma\|_F^2. \end{aligned}$$

Therefore, the remaining work is conducted to show that the constant $C_5 := C_5(q, \eta, K, M, \epsilon)$ exists, such that

$$\frac{1}{p}\|\tilde{\Sigma}^\tau - \Sigma\|_F^2 \leq C_5 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \epsilon^2}} \right)^{2-q} \quad (\text{A19})$$

under event E .

Obviously, it holds that

$$\begin{aligned} \frac{1}{p}\|\tilde{\Sigma}^\tau - \Sigma\|_F^2 &= \frac{1}{p} \sum_{j=1}^p \sum_{i=1}^p |\tilde{\sigma}_{ij}^\tau - \sigma_{ij}|^2 \leq \max_j \sum_{i=1}^p |\tilde{\sigma}_{ij}^\tau - \sigma_{ij}|^2 \\ &\leq 2 \left(\max_j \sum_{i=1}^p |\tilde{\sigma}_{ij}^\tau - \sigma_{ij}^\tau|^2 + \max_j \sum_{i=1}^p |\sigma_{ij}^\tau - \sigma_{ij}|^2 \right). \end{aligned}$$

Using similar discussion techniques as those applied for estimating R_1, R_2 in (A11) and (A10), we know that

$$\frac{1}{p}\|\tilde{\Sigma}^\tau - \Sigma\|_F^2 \leq C_q \max_j \sum_{i=1}^p \lambda^{2-q} |\sigma_{ij}|^q$$

under event E , where C_q denotes a constant related to q . Thus, we have

$$\begin{aligned} \frac{1}{p}\|\tilde{\Sigma}^\tau - \Sigma\|_F^2 &\leq C_q \left(\frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\epsilon}} M \sqrt{\frac{\log p}{n}} + 8\sqrt{c_1} \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \epsilon^2}} \right)^{2-q} \max_j \sum_{i=1}^p |\sigma_{ij}|^q \\ &\leq C_5 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n^2 \epsilon^2}} \right)^{2-q} \end{aligned}$$

due to (A15) and $\Sigma \in \mathcal{U}_q(s_{n,p}, M)$. This is the expected conclusion of (A19). \square

Proof of Theorem 2. To show the conclusions (i)–(ii) of Theorem 2, we define the event as

$$E_2 = \left\{ |\tilde{\sigma}_{ij} - \sigma_{ij}| \leq \frac{3}{4} \left(\xi \sqrt{\frac{\log p}{n}} + 4\sigma_2 (\log p)^{1/2} \right), \forall 1 \leq i, j \leq p \right\}$$

and $\xi := \frac{4\sqrt{2}}{3} \eta K^{\frac{2}{4+\epsilon}} M$.

Based on Lemma 1 and $\mathbb{P}\{|X| > t\} \leq 2e^{-\frac{t^2}{2\sigma^2}}$ for $X \sim N(0, \sigma^2)$ in [27], and using the similar proving process of Lemma A1, we know that for any $\eta \geq 2$ and some $\epsilon > 0$,

$$\mathbb{P}(E_2) \geq 1 - O\left((\log p)^{-\frac{1}{2}} p^{-\eta+2} + n^{-\epsilon} + p^{-\frac{5}{2}}\right). \quad (\text{A20})$$

By considering the similar discussions used to prove Theorem 1, we deduce that constants $C_6 := C_6(q, \eta, K, M, \epsilon)$, $C_7 := C_7(q, \eta, K, M, \epsilon)$ exist, such that

$$\|\check{\Sigma}_+^\tau - \Sigma\|_2 \leq C_6 s_{n,p} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n\epsilon^2}} \right)^{1-q}$$

and

$$\frac{1}{\sqrt{p}} \|\check{\Sigma}_+^\tau - \Sigma\|_F \leq C_7 \sqrt{s_{n,p}} \left(\sqrt{\frac{\log p}{n}} + \sqrt{\frac{\log \delta^{-1} \log p}{n\epsilon^2}} \right)^{1-q/2}$$

under event E_2 . Furthermore, there exists $c_2 > 0$, such that $(\log p)/n \leq c_2(\log \delta^{-1} \log p)/n\epsilon^2$. Hence, the above two inequalities reduce to

$$\|\check{\Sigma}_+^\tau - \Sigma\|_2 \leq C_3 s_{n,p} \left(\frac{\log \delta^{-1} \log p}{n\epsilon^2} \right)^{\frac{1-q}{2}}$$

and

$$\frac{1}{\sqrt{p}} \|\check{\Sigma}_+^\tau - \Sigma\|_F \leq C_4 \sqrt{s_{n,p}} \left(\frac{\log \delta^{-1} \log p}{n\epsilon^2} \right)^{\frac{1-q}{2} - \frac{q}{4}}$$

under event E_2 . This, along with (A20), enables us to reach the desired conclusions of Theorem 2. \square

References

1. Sifaou, H.; Kammoun, A.; Alouini, M.-S. High-dimensional linear discriminant analysis classifier for spiked covariance model. *J. Mach. Learn. Res.* **2020**, *21*, 4508–4531.
2. Chaudhuri, S.; Drton, M.; Richardson, T. Estimation of a covariance matrix with zeros. *Biometrika* **2007**, *94*, 199–216. [\[CrossRef\]](#)
3. Xu, K.; Hero, A. Dynamic stochastic blockmodels for time-evolving social networks. *IEEE J. Sel. Topics Signal Process.* **2014**, *8*, 552–562. [\[CrossRef\]](#)
4. Cai, T.; Liu, W. Adaptive thresholding for sparse covariance matrix estimation. *J. Amer. Statist. Assoc.* **2011**, *106*, 672–684. [\[CrossRef\]](#)
5. Belomestny, D.; Mathias, T.; Tsybakov, A. Sparse covariance matrix estimation in high-dimensional deconvolution. *Bernoulli* **2019**, *25*, 1901–1938. [\[CrossRef\]](#)
6. Tang, J.; Korolova, A.; Bai, X.; Wang, X.; Wang, X. Privacy loss in Apple’s implementation of differential privacy on MacOS 10.12. *arXiv* **2017**, arXiv:1709.02753.
7. Ding, B.; Kulkarni, J.; Yekhanin, S. Collecting Telemetry Data Privately. In Proceedings of the Thirty-First Annual Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 3571–3580.
8. Erlingsson, U.; Pihur, V.; Korolova, A. Randomized aggregatable privacy-preserving ordinal response. In Proceedings of the Twenty-First ACM Conference on Computer and Communications Security, Scottsdale, AZ, USA, 3–7 November 2014; pp. 1054–1067.
9. Dwork, C.; McSherry, F.; Nissim, K.; Smith, A. Calibrating noise to sensitivity in private data analysis. In Proceedings of the Theory of Cryptography Conference, New York, NY, USA, 4–7 March 2006; pp. 265–284.
10. Wang, D.; Xu, J. Differentially private high dimensional sparse covariance matrix estimation. *Theoret. Comput. Sci.* **2021**, *865*, 119–130. [\[CrossRef\]](#)
11. Cai, T.; Wang, Y.; Zhang, L. The cost of privacy: Optimal rates of convergence for parameter estimation with differentially privacy. *Ann. Statist.* **2021**, *49*, 2825–2850. [\[CrossRef\]](#)
12. Jiang, W.; Xie, C.; Zhang, Z. Wishart mechanism for differentially private principal components analysis. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 1730–1736.
13. Amin, K.; Dick, T.; Kulesza, A.; Munoz, A.; Vassilvitskii, S. Differentially private covariance estimation. In Proceedings of the Thirty-Third Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 14213–14222.
14. Kamath, G.; Li, J.; Singhal, V.; Ullman, J. Privately learning high-dimensional distributions. In Proceedings of the Thirty-Second Annual Conference on Learning Theory, Phoenix, AZ, USA, 4–8 March 2019; pp. 1853–1902.
15. Vershynin, R. *Introduction to the Non-Asymptotic Analysis of Random Matrices*; Cambridge University Press: Cambridge, UK, 2012; pp. 210–268.

16. Dwork, C.; Kenthapadi, K.; McSherry, F.; Mironov, I.; Naor, M. Our data, ourselves: Privacy via distributed noise generation. In Proceedings of the Twenty-Fourth Annual International Conference on Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, 28 May–1 June 2006; pp. 486–503.
17. Su, W.; Guo, X.; Zhang, H. Differentially private precision matrix estimation. *Acta Math. Sin.* **2020**, *36*, 1107–1124. [[CrossRef](#)]
18. Dwork, C.; Roth, A. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* **2014**, *9*, 211–407. [[CrossRef](#)]
19. Thorin, G. Generalizing those of M. riesz and hadamard with some applications. *Comm. Sem. Math. Univ. Lund* **1948**, *9*, 1–58.
20. Kasiviswanathan, S.; Lee, H.; Nissim, K.; Raskhodnikova, S.; Smith, A. What can we learn privately? *SIAM J. Comput.* **2011**, *40*, 793–826. [[CrossRef](#)]
21. Dwork, C.; Talwar, K.; Thakurta, A.; Zhang, L. Analyze gauss: Optimal bounds for privacy-preserving principal component analysis. In Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing, New York, NY, USA, 31 May–3 June 2014; pp. 11–20.
22. Liang, W.; Wu, Y.; Chen, H. Sparse covariance matrix estimation for ultrahigh dimensional data. *Stat* **2022**, *11*, e479. [[CrossRef](#)]
23. Wu, G.; Chen, J.; Hoffmann, C.; Bittinger, K.; Chen, Y.Y.; Keilbaugh, S.A.; Bewtra, M.; Knights, D.; Walters, W.A.; Knight, R.; et al. Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science* **2011**, *335*, 105–108. [[CrossRef](#)] [[PubMed](#)]
24. Tsybakov, A. *Introduction to Nonparametric Estimation*; Springer: New York, NY, USA, 2009.
25. Assouad, P. Deux remarques sur l'estimation. *C. R. Acad. Sci. Paris Sér.* **1983**, *296*, 1021–1024.
26. Le Cam, L. *Asymptotic Methods in Statistical Decision Theory*; Springer: New York, NY, USA, 1986.
27. Vershynin, R. *High-Dimensional Probability: An Introduction with Applications in Data Science*; Cambridge University Press: Cambridge, UK, 2018; pp. 672–684.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.