Optimal Estimation of Eigenspace of Large Density Matrices of Quantum Systems Based on Pauli Measurements

Tony Cai[1], Donggyu Kim[2]*, and Yazhen Wang[3]

[1]Department of Statistics, Wharton School, University of Pennsylvania, PA 19104

[2]College of Business, Korea Advanced Institute of Science and Technology, Seoul, Korea

[3]Department of Statistics, University of Wisconsin- Madison, WI 53706

September 3, 2017

Abstract

Quantum state tomography, which aims to reconstruct quantum states described by density matrices, is becoming increasingly important in many scientific studies involving quantum systems. This paper considers the reconstruction of high-dimensional low-rank density matrices based on Pauli measurements. In particular it focuses on estimation of eigenspace for a large low-rank density matrix. Both ordinary principal component analysis (PCA) and iterative thresholding sparse PCA (ITSPCA) are studied and optimal rates of convergence are established. Minimax lower bounds for eigenspace estimation under the spectral and Frobenius norms are derived. It is shown that the convergence rates of the ITSPCA algorithm matches the minimax lower bounds and the procedure is thus rate-optimal. With these PCA estimators, we reconstruct the large low-rank density matrix and obtain the optimal convergence rate. A simulation study is carried out to investigate the finite sample performance of the proposed estimators of the density matrices.

Key words and phrases: Iterative thresholding, minimax estimation, principal component analysis, quantum state tomography, quantum measurement, Pauli matrices, sparsity.
Running title: Estimate Eigenspace of Large Density Matrices.

*corresponding author.

1 Introduction

Modern scientific studies often need to learn and engineer quantum systems. Examples include quantum computation, quantum information, and quantum simulation (Nielsen and Chuang (2000) and Wang (2011, 2012)). A quantum system is described by its state, and the state is often characterized by a complex matrix called density matrix. For the study of a quantum system, it is important but often difficult to reconstruct the quantum state (density matrix). The literature refers to this reconstruction of the quantum state as quantum state tomography, which is based on observations obtained from measuring identically prepared quantum systems. Traditionally, quantum state tomography employs classical statistical models and methods to deduce quantum states from quantum measurements. However, the size of the density matrix usually grows exponentially with the size of the quantum system, and thus, quantum state tomography requires the recovery of high-dimensional density matrices. It is well known that classical statistical approaches are neither efficient nor effective in estimating large density matrices.

Cai *et al.* (2016) developed a rate-optimal density matrix estimator based on Pauli measurements, under the assumption that the density matrix has a sparse representation under the Pauli matrices. A thresholding procedure is proposed to recover the density matrix. However, the sparsity assumption does not hold in general under the Pauli representation when the density matrix is of low-rank (see Wang (2013)). That is, the low-rank density matrices are not included in their sparse density matrix class. Spectral decomposition indicates that a low-rank density matrix $\boldsymbol{\rho}$ can be decomposed as follows:

$$\boldsymbol{\rho} = \sum_{\nu=1}^{r} \lambda_\nu \mathbf{q}_\nu \mathbf{q}_\nu^\dagger,$$

where r is a finite rank, $\dagger$ denotes conjugate transpose, λ_ν's are the eigenvalues of $\boldsymbol{\rho}$, and $\mathbf{q}_\nu \in \mathbb{C}^d$ are the eigenvectors corresponding to λ_ν. Recently, Koltchinskii and Xia (2015) investigated the optimal estimation of low-rank density matrices under the general low-rank density matrix class. For example, they established the minimax lower bound for quantum versions of Kullback-Leibler divergence and of Hellinger distance, and Schatten norm and showed that the least squared estimator with von Neumann entropy penalization achieves the minimax lower bound. Their study mainly focused on reconstructing the low-rank density matrix itself. However, we often have an interest in estimating eigenspace, which also plays an important role in reconstructing low-rank density matrices. Furthermore, the recent development in high-dimensional statistics indicates that the optimal rate of estimating eigenspace is depending on the sparsity of eigenvectors (see, for example, Ma (2013), Cai, Ma, and Wu (2013, 2015), Vu and Lei (2013), Johnstone and Lu (2009), and Birnbaum *et al.* (2013)). The low-rank density matrix class considered in Koltchinskii and Xia (2015) is broad, and thus their minimax lower bound may not be sharp under the sparse eigenvector condition.

The present paper considers the problem of eigenspace estimation for a quantum spin system based on Pauli measurements. Since all Pauli matrices have ± 1 eigenvalues, Pauli measurements take discrete values 1 and -1, and their distributions can be characterized by binomial distributions (Wang (2013) and Cai *et al.* (2016)). Statistically, the problem of eigenspace estimation lies in the framework of high-dimensional statistics with binomial distributions, where both the matrix size and sample size diverge to infinity. We first analyze the asymptotic behavior of the ordinary principal component analysis (PCA) estimator and establish the optimal convergence rate when the eigenvectors are dense. We then consider the setting where the eigenvectors are sparse. Under the sparsity condition, we establish the minimax lower bound for the eigenspace estimation problem and show that the iterative thresholding sparse PCA (ITSPCA) (Ma (2013)) can achieve the minimax lower bound, and thus its convergence rate is rate-optimal. The convergence rate and minimax lower bound are obtained by asymptotic analysis for binomial distributions instead of usual normal distributions. Using the ITSPCA estimator, we propose an estimator for the eigenvalues, which leads to the reconstruction of the density matrix. The proposed low-rank density matrix estimators also can obtain the optimal convergence rate.

The rest of paper proceeds as follows. Section 2 reviews the quantum state, density matrix, and Pauli matrices and describes the density matrix representation through Pauli matrices. Section 3 provides the iterative thresholding algorithm and presents a sparse condition. Section 4 establishes the asymptotic theory for the iterative thresholding estimator and the minimax lower bound of eigenspace estimators for spectral and Frobenius norms under the sparsity assumption, where both the matrix size and sample size go to infinity. Section 5 proposes the eigenvalue and low-rank density matrix estimators and derives their convergence rates. Section 6 features a simulation study to illustrate the finite sample performances of the estimators, and in Section 7 we conducted a Monte Carlo simulation to analyze the real density matrices. Section 8 outlines the key ideas and main steps of the proofs, and Appendix details technical proofs.

2 Quantum state tomography with Pauli measurement

2.1 Quantum state and Pauli matrices

For a d-dimensional quantum system, we describe its quantum state by a density matrix $\boldsymbol{\rho}$ on d-dimensional complex space $\mathbb{C}^d$, where density matrix $\boldsymbol{\rho}$ is a d-by-d complex matrix satisfying (1) Hermitian, that is, $\boldsymbol{\rho}$ is equal to its conjugate transpose; (2) positive semi-definite; (3) unit trace, that is, $\mathrm{tr}(\boldsymbol{\rho}) = \sum_{\nu=1}^{r} \lambda_\nu = 1$.

The density matrix $\boldsymbol{\rho}$ can be expressed by the d-dimensional Pauli matrices. Specifically, let

$$\boldsymbol{\sigma}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}, \ \text{ and } \ \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$, and $\boldsymbol{\sigma}_3$ are called Pauli matrices. Tensor products are used to define high-dimensional Pauli matrices. Let $d = 2^b$ for some integer b. We form b-fold tensor products of $\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$, and $\boldsymbol{\sigma}_3$ to obtain d-dimensional Pauli matrices

$$\mathcal{G}_P = \left\{ \mathbf{B}_j = \boldsymbol{\sigma}_{\ell_1} \otimes \boldsymbol{\sigma}_{\ell_2} \otimes \cdots \otimes \boldsymbol{\sigma}_{\ell_b}, \quad (\ell_1, \ell_2, \ldots, \ell_b) \in \{0,1,2,3\}^b \right\},$$

and the cardinality of $\mathcal{G}_P$ is $p = 4^b$. We set $\mathbf{B}_1 = \mathbf{I}_d$, where $\mathbf{I}_d$ is the d-dimensional identity matrix. Denote by $\mathbb{C}^{d\times d}$ the space of all d-by-d complex matrix equipped with the Frobenius norm. Proposition 1 in Cai *et al.* (2016) showed that all Pauli matrices $\mathbf{B}_1, \ldots, \mathbf{B}_p$ form an orthogonal basis for complex Hermitian matrices in $\mathbb{C}^{d\times d}$, and any density matrix $\boldsymbol{\rho}$ can be expanded under the Pauli basis as follows:

$$\boldsymbol{\rho} = d^{-1} \left(\mathbf{I}_d + \sum_{j=2}^{p} \beta_j \mathbf{B}_j \right),$$

where coefficients β_j satisfy $\beta_j = \mathrm{tr}(\boldsymbol{\rho}\mathbf{B}_j)$ and $|\beta_j| \leq 1$.

2.2 Pauli measurements and density matrix estimation

Quantum measurements are often based on observables, where an observable is defined as a Hermitian matrix on $\mathbb{C}^d$. The Pauli matrices are widely used in quantum physics and quantum information science to perform quantum measurements. Suppose that an experiment is conducted to perform measurements on each Pauli observable $\mathbf{B}_j$ independently for n quantum systems which are identically prepared in the same quantum state $\boldsymbol{\rho}$. As $\mathbf{B}_j$ has eigenvalues ± 1, the theory of quantum mechanics indicates that the Pauli measurements take values 1 and -1, and thus are Bernoulli trials. Denote by N_j the average of the n measurement outcomes obtained from measuring $\mathbf{B}_j$, $j = 2, \ldots, p$. Then $n(N_j + 1)/2$ follows a binomial distribution with n trials and cell probability $(1+\beta_j)/2$, where $E(N_j) = \beta_j$ and $Var(N_j) = (1 - \beta_j^2)/n$ (see Cai *et al.* (2016)). The goal of this paper is to estimate eigenspace of $\boldsymbol{\rho}$ based on data $N_2, \ldots, N_p$.

To estimate the eigenspace of the density matrix $\boldsymbol{\rho}$, we need an initial estimator of $\boldsymbol{\rho}$ based on the Pauli measurements $N_2, \ldots, N_p$. From the binomial distribution, we easily derive that each N_j is MLE and UMVUE of β_j. Thus, a natural estimator of $\boldsymbol{\rho}$ is given by

$$\widehat{\boldsymbol{\rho}} = (\widehat{\rho}_{ij})_{i,j=1,\ldots,d} = \frac{1}{d} \left(\mathbf{I}_d + \sum_{j=2}^{p} \widehat{\beta}_j \mathbf{B}_j \right), \tag{2.1}$$

where $\widehat{\beta}_j = N_j$.

3 Eigenspace estimation

3.1 Model set-up

Assume that a density matrix $\boldsymbol{\rho}$ has finite rank r. By the spectral decomposition, we have

$$\boldsymbol{\rho} = \sum_{\nu=1}^{r} \lambda_\nu \mathbf{q}_\nu \mathbf{q}_\nu^\dagger, \tag{3.1}$$

where λ_ν's are eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$, $\sum_{\nu=1}^{r} \lambda_\nu = 1$, and their corresponding eigenvectors are $\mathbf{q}_1, \ldots, \mathbf{q}_r \in \mathbb{C}^d$. In this paper we consider estimation of the eigenspace spanned by the first m eigenvectors of $\boldsymbol{\rho}$, that is, our aim is to estimate the eigenspace generated by $\mathbf{Q} = (\mathbf{q}_1, \ldots, \mathbf{q}_m) \in \mathbb{C}^{d\times m}$, where m is a given integer. To make the eigenspace estimation problem well defined, we need to assume that $m \leq r$ and $\lambda_m - \lambda_{m+1} > C_\lambda$ for some generic positive constant C_λ free of n and d, that is, there is a gap between eigenvalues λ_m and λ_{m+1} so that the corresponding eigenspaces are well separated for investigating asymptotic properties of the eigenspace estimation.

3.2 Ordinary PCA

We define the eigenspace estimator of $\mathbf{Q}$ by the eigenspace spanned by the first m eigenvectors of the density matrix estimator $\widehat{\boldsymbol{\rho}}$ in (2.1). As the m eigenvectors are from ordinary PCA, the defined eigenspace estimator is called the ordinary PCA estimator and denote by $\widehat{\mathbf{Q}}$. Before investigating its asymptotic properties, we first fix some notations. For $\mathbf{x} = (x_1, \ldots, x_d)^T \in \mathbb{C}^d$ and $\mathbf{A} = (A_{ij}) \in \mathbb{C}^{d\times d}$, define the ℓ_α-norms,

$$\|\mathbf{x}\|_\alpha = \left(\sum_{i=1}^{d} |x_i|^\alpha \right)^{1/\alpha}, \quad \|\mathbf{A}\|_\alpha = \sup\{\|\mathbf{A}\mathbf{x}\|_\alpha, \|\mathbf{x}\|_\alpha = 1\}, \quad 1 \leq \alpha \leq \infty.$$

Then $\|\mathbf{A}\|_2$ is called the matrix spectral norm and equal to the square root of the largest eigenvalue of $\mathbf{A}\mathbf{A}^\dagger$, and

$$\|\mathbf{A}\|_1 = \max_{1\leq j\leq d} \sum_{i=1}^{d} |A_{ij}|, \quad \|\mathbf{A}\|_\infty = \max_{1\leq i\leq d} \sum_{j=1}^{d} |A_{ij}|.$$

We have the following inequality,

$$\|\mathbf{A}\|_2^2 \leq \|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty.$$

The Frobenius norm of $\mathbf{A}$ is denoted by $\|\mathbf{A}\|_F = \sqrt{\mathrm{tr}(\mathbf{A}^\dagger\mathbf{A})}$. For a symmetric or complex Hermitian matrix $\mathbf{A}$, $\|\mathbf{A}\|_F$ is the square root of the sum of squared eigenvalues, $\|\mathbf{A}\|_2$ is equal to its largest absolute eigenvalue, and $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_1 = \|\mathbf{A}\|_\infty$. Denote by C a generic constant whose values are free of n and p and may change from appearance to appearance.

To measure the performances of eigenspace estimators, we define a notation of distance between eigenspaces as follows. We define the distance between two eigenspaces spanned by $\mathbf{Q}_1$ and $\mathbf{Q}_2$ by

$$\|\sin(\mathbf{Q}_1, \mathbf{Q}_2)\|_F^2 = \|\mathbf{Q}_1\mathbf{Q}_1^\dagger(\mathbf{Q}_2\mathbf{Q}_2^\dagger)^\perp\|_F^2 \tag{3.2}$$

and

$$\|\sin(\mathbf{Q}_1, \mathbf{Q}_2)\|_2^2 = \|\mathbf{Q}_1\mathbf{Q}_1^\dagger(\mathbf{Q}_2\mathbf{Q}_2^\dagger)^\perp\|_2^2, \tag{3.3}$$

where $\mathbf{Q}_1\mathbf{Q}_1^\dagger$ and $\mathbf{Q}_2\mathbf{Q}_2^\dagger$ are projection matrices on eigenspaces $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively, and for a given projection matrix $\mathbf{P}$, $\mathbf{P}^\perp = \mathbf{I}_d - \mathbf{P}$. The distances are referred to as the canonical angles between $\mathbf{Q}_1$ and $\mathbf{Q}_2$ that generalize the notion of angles between lines.

The following theorem establishes the convergence rate of the ordinary PCA estimator.

Theorem 1 *Suppose that $n^{\alpha_1} \le d \le \exp(n^{\alpha_2})$ for some $\alpha_1 > 1/2$ and $\alpha_2 < 1$. Then we have*

$$\sup_{\mathbf{Q}\in\mathbb{V}_{d,m}} E\left[\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}})\|_2^2\right] \le \sup_{\mathbf{Q}\in\mathbb{V}_{d,m}} E\left[\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}})\|_F^2\right] \le Cn^{-1}, \tag{3.4}$$

where $\mathbb{V}_{d,m} = \{\mathbf{Q} \in \mathbb{C}^{d\times m} : \mathbf{Q}^\dagger\mathbf{Q} = \mathbf{I}_d\}$ is the complex Stiefel manifold of d-by-m orthonormal matrices, and C is a generic constant free of n and d.

Remark 1 From Theorem 1 we can see that ordinary PCA has convergence rate $n^{-1/2}$ regardless of sparsity condition on eigenvectors. As the PCA approach does not utilize any sparse eigenvectors, it can achieve only $n^{-1/2}$ convergence rate even for a density matrix with sparse eigenvectors. We will show later that the convergence rate is suboptimal for the sparse case.

3.3 Sparse eigenvectors and iterative thresholding PCA estimator

As the complexity of a quantum system is exponentially increasing with its components, its dimension d and density matrix grow exponentially in its size and are often very large. As in usual high-dimensional statistics, we may impose sparsity on the eigenvectors of its density matrix and estimate the eigenspace spanned the first m sparse eigenvectors. For $\mathbf{A} \in \mathbb{C}^{d\times m}$, $\mathbf{A}_{IJ}$ denotes the submatrix of $\mathbf{A}$ formed by rows and columns whose indices are in I and J, respectively, where I and J are subsets of $\{1, \cdots, d\}$. When I or J includes all the indices, we replace them with dot. For example, $\mathbf{A}_{\cdot J}$ is the submatrix of $\mathbf{A}$ with all rows and columns indexed by J.

We impose the sparsity condition on the first r eigenvectors of $\boldsymbol{\rho}$ in (3.1) as follows. For each $\nu = 1, \ldots, r$, assume that for some $\delta \in [0, 2)$,

$$\mathbf{q}_\nu \in \Xi_\delta(\pi(d)) \stackrel{\text{def}}{=} \left\{\mathbf{a} = (a_1, \ldots, a_d) : \sum_{\nu=1}^d |a_\nu|^\delta \le \pi(d) \text{ and } \sum_{\nu=1}^d |a_\nu|^2 = 1\right\}, \tag{3.5}$$

where $\pi(d)$ is a deterministic function of d that diverges slowly such as $\log d$. The sparsity condition is usually considered in high-dimensional statistics, for example, sparse covariance matrix estimation (Bickel and Levina (2008), Cai and Liu (2011), and Cai and Zhou (2012)), sparse integrated volatility matrix estimation (Kim *et al.* (2016), Tao *et al.* (2013a, 2013b), and Wang and Zou (2013)), and sparse PCA (Birnbaum *et al.* (2013), Johnstone and Lu (2009), Ma (2013), Vu *et al.* (2013), and Vu and Lei (2013)).

The orthogonal iteration may be used to compute the leading eigenspace of a given Hermitian matrix (Golub and Loan (1996)), which yields the ordinary PCA estimator. As we have seen in Section 3.2, the ordinary PCA estimator has the convergence rate $n^{-1/2}$. However, the ordinary PCA estimation may not be the best way to estimate the sparse eigenvectors in terms of mean squared error (MSE). In order to obtain better eigenspace estimators under sparsity, we adapt the iterative thresholding algorithm known as the iterative thresholding sparse PCA (ITSPCA) proposed by Ma (2013).

Algorithm 1 Iterative thresholding sparse PCA (ITSPCA) (Ma (2013))

Input:

(1) Estimated density matrix $\widehat{\boldsymbol{\rho}}$;

(2) Target subspace dimension m;

(3) Initial orthonormal matrix $\widehat{\mathbf{Q}}^{(0)}$;

(4) Thresholding function $\mathcal{T}(t,\gamma)$, and threshold levels $\gamma_{nj}, j = 1, \dots, m$.

```
repeat
    Multiplication: T^(k) = (t_{νj}^(k)) = ρ̂ Q̂^(k-1);
    Thresholding: T̂^(k) = (t̂_{νj}^(k)), with t̂_{νj}^(k) = T(t_{νj}^(k), γ_{nj});
    QR factorization: Q̂^(k) R̂^(k) = T̂^(k);
until convergence.
```

As described in Algorithm 1, the ITSPCA method has three steps; multiplication, thresholding, and QR factorization. Without the thresholding step, the ITSPCA method returns to the ordinary orthogonal iteration method. The thresholding step removes weak signal elements of $\mathbf{T}^{(k)}$ with a user-specified thresholding function $\mathcal{T}$ which satisfies

$$|\mathcal{T}(t,\gamma) - t| \le \gamma \quad \text{and} \quad \mathcal{T}(t,\gamma)\mathbf{1}_{(|t|\le\gamma)} = 0 \quad \text{for all } t \text{ and all } \gamma > 0, \tag{3.6}$$

where $\mathbf{1}_E$ denotes the indicator function of an event E. Note that both hard thresholding rule $\mathcal{T}_H(t,\gamma) = t\,\mathbf{1}_{(|t|>\gamma)}$ and soft thresholding rule $\mathcal{T}_S(t,\gamma) = e^{\sqrt{-1}\,\theta}\max(0, |t| - \gamma)$ satisfy (3.6), where $t = |t|e^{\sqrt{-1}\,\theta}$, and θ is the phase of complex number t.

To harness the ITSPCA algorithm in Algorithm 1, an appropriate initial orthonormal matrix $\widehat{\mathbf{Q}}^{(0)}$ is required. Johnstone and Lu (2009) introduced a diagonal thresholding sparse

Algorithm 2 Diagonal thresholding sparse PCA (DTSPCA) (Johnstone and Lu (2009))

Input:

(1) Estimated density matrix $\widehat{\boldsymbol{\rho}}$;

(2) Diagonal thresholding parameter α_n.

Output: Orthonormal matrix $\widehat{\mathbf{Q}}_S$.

1: Selection: select the set S of coordinates:

$$S = \{\nu : \widehat{\rho}_{\nu\nu} \geq \alpha_n\};$$

2: Reduced PCA: compute the eigenvectors, $\widehat{\mathbf{q}}_1^S, \ldots, \widehat{\mathbf{q}}_{|S|}^S$, of the submatrix $\widehat{\boldsymbol{\rho}}_{SS}$;

3: Zero-padding: construct $\widehat{\mathbf{Q}}_S = (\widehat{\mathbf{q}}_1, \ldots, \widehat{\mathbf{q}}_{|S|})$ such that

$$\widehat{\mathbf{q}}_{jS} = \widehat{\mathbf{q}}_j^S, \quad \widehat{\mathbf{q}}_{jS^c} = 0, \quad j = 1, \ldots, |S|.$$

PCA (DTSPCA) method to estimate the eigenspace and showed its consistency. We can use the DTSPCA method to obtain $\widehat{\mathbf{Q}}^{(0)}$. The DTSPCA algorithm is described in Algorithm 2. Given the output $\widehat{\mathbf{Q}}_S = (\widehat{\mathbf{q}}_1, \ldots, \widehat{\mathbf{q}}_{|S|})$, we take the first m columns as the initial orthogonal matrix, that is, $\widehat{\mathbf{Q}}^{(0)} = (\widehat{\mathbf{q}}_1, \ldots, \widehat{\mathbf{q}}_m)$.

4 Asymptotic theory for the eigenspace estimator

4.1 Convergence rate

Assume that density matrix $\boldsymbol{\rho}$ belongs to the following class,

$$\mathcal{F}_\delta(\pi(d)) = \left\{ \boldsymbol{\rho} = \sum_{\nu=1}^{r} \lambda_\nu \mathbf{q}_\nu \mathbf{q}_\nu^\dagger : \mathbf{q}_\nu \in \Xi_\delta(\pi(d)) \text{ for all } \nu \in \{1, \ldots, r\} \right\}, \tag{4.1}$$

where $\Xi_\delta(\pi(d))$ is defined in (3.5). We consider the following ITSPCA estimator $\widehat{\mathbf{Q}}^{(R_s)}$ for $\boldsymbol{\rho}$. For a theoretical study, we always stop Algorithm 1 after R_s iterations, where

$$R_s = \frac{1.1\ell_1^S}{\ell_m^S - \ell_{m+1}^S}(\log n + 0.5\log(d \vee n)),$$

$\ell_j^S = \ell_j(\widehat{\boldsymbol{\rho}}_{SS}) \vee 0$, and $\ell_j(\widehat{\boldsymbol{\rho}}_{SS})$ is the j-th largest eigenvalue of $\widehat{\boldsymbol{\rho}}_{SS}$.

The following theorem establishes the convergence rate of the eigenspace estimator $\widehat{\mathbf{Q}}^{(R_s)}$ obtained from Algorithm 1.

Theorem 2 *Assume density matrix $\boldsymbol{\rho}$ given by model* (3.1) *belongs to $\mathcal{F}_\delta(\pi(d))$ defined in* (4.1) *so that for some $\delta \in (0, 2/3)$,*

$$\pi(d)\tau_n^{1/2-3\delta/4} = O(1), \tag{4.2}$$

where $\tau_n = \sqrt{\frac{\log(d\vee n)}{nd}}$. *Take* $\alpha_n = C_\alpha \tau_n$ *in Algorithm 2 and* $\gamma_{nj} = C_\gamma \sqrt{\ell_j^S}\tau_n$ *in Algorithm 1 for some constant* C_α *and* C_γ *free of* n *and* d, *and let*

$$R = \frac{\lambda_1}{\lambda_m - \lambda_{m+1}}(\log n + 0.5\log(d \vee n)).$$

Then there exist constants C_0 *and* C_u *such that for sufficiently large* n, *uniformly over* $\mathcal{F}_\delta(\pi(d))$, *with probability at least* $1 - C_0(d \vee n)^{-2}$, *and* $R_s \in [R, 2R]$, *we have*

$$\| \sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_2^2 \leq \| \sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2 \leq C_u \pi(d)\tau_n^{2-\delta}.$$

Remark 2 When $\delta < 2/3$, condition (4.2) indicates that $\pi(d)$ is at most of order d with some positive power. In high-dimensional statistics, $\pi(d)$ usually grows very slowly in d with an example of $\log d$, and so condition (4.2) is not restrictive.

The result of Theorem 2 can be extended to an upper bound for the mean squared error (MSE). Note that $(d \vee n)^{-2} = o\left(\pi(d)\tau_n^{2-\delta}\right)$ and the loss functions, (3.2) and (3.3), are bounded by r and 1, respectively. The following corollary is a direct consequence of Theorem 2.

Corollary 1 *Under the conditions of Theorem 2, we have*

$$\sup_{\boldsymbol{\rho}\in\mathcal{F}_\delta(\pi(d))} E\left[\| \sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_2^2\right] \leq \sup_{\boldsymbol{\rho}\in\mathcal{F}_\delta(\pi(d))} E\left[\| \sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2\right] \leq C_u \pi(d)\tau_n^{2-\delta}.$$

Remark 3 Since in high-dimensional statistics, there is a constant C free of n and d such that $\log(d \vee n) \leq C \log d$, Theorem 2 and Corollary 1 show that the ITSPCA estimator has the convergence rate $\pi(d)^{1/2}\left[n^{-1}d^{-1}\log d\right]^{1/2-\delta/4}$ under the Frobenius and spectral norms. As d is often much larger than n, the convergence rate is faster than $n^{-1/2}$ convergence rate for the ordinary PCA case.

Although our main objective is to estimate the eigenspace, when individual eigenvector $\mathbf{q}_k$ is identifiable, it is interesting to see whether the ITSPCA method can estimate $\mathbf{q}_k$ well. The theorem below shows that the k-th column of $\widehat{\mathbf{Q}}^{(R_s)}$ provides a good estimator of the well separated $\mathbf{q}_k$.

Corollary 2 *Suppose that for some* $k \leq m$, $\lambda_k - \lambda_{k+1} \geq C_{\lambda 1}$ *and* $\lambda_{k-1} - \lambda_k \geq C_{\lambda 2}$ *for some positive constants* $C_{\lambda 1}$ *and* $C_{\lambda 2}$ *free of* n *and* d. *Under the conditions of Theorem 2, we have that the* k-*th column* $\widehat{\mathbf{q}}_k^{(R_s)}$ *of* $\widehat{\mathbf{Q}}^{(R_s)}$ *satisfies*

$$\sup_{\boldsymbol{\rho}\in\mathcal{F}_\delta(\pi(d))} E\left[\| \sin(\mathbf{q}_k, \widehat{\mathbf{q}}_k^{(R_s)})\|_2^2\right] \leq \sup_{\boldsymbol{\rho}\in\mathcal{F}_\delta(\pi(d))} E\left[\| \sin(\mathbf{q}_k, \widehat{\mathbf{q}}_k^{(R_s)})\|_F^2\right] \leq C_u \pi(d)\tau_n^{2-\delta}.$$

4.2 Optimality of the sparse PCA estimator

This section establishes the minimax lower bound for the problem of estimating the eigenspace spanned by $\mathbf{Q}$ under model (3.1), uniformly over $\mathcal{F}_\delta(\pi(d))$, and shows that the ITSPCA estimator achieves the minimax lower bound, and thus its convergence rate is optimal.

The theorem below provides a minimax lower bound for eigenspace estimation under the Frobenius and spectral norms.

Theorem 3 *For model* (3.1)*, suppose that for some* $\delta \in [0, 2)$*,*

$$\pi(d) = O(d^{(1-\delta/2)-\aleph} n^{-\delta/2} \log^{\delta/2} d), \tag{4.3}$$

where $\aleph$ *is a positive constant free of* n *and* d*. Then there exists a positive constant* C_L *free of* n *and* d *such that for sufficiently large* n*,*

$$\inf_{\check{\mathbf{Q}}} \sup_{\boldsymbol{\rho} \in \mathcal{F}_\delta(\pi(d))} E\left[\|\sin(\mathbf{Q}, \check{\mathbf{Q}})\|_2^2\right] \geq C_L \pi(d) \left[\frac{\log d}{nd}\right]^{1-\delta/2}$$

and

$$\inf_{\check{\mathbf{Q}}} \sup_{\boldsymbol{\rho} \in \mathcal{F}_\delta(\pi(d))} E\left[\|\sin(\mathbf{Q}, \check{\mathbf{Q}})\|_F^2\right] \geq C_L \pi(d) \left[\frac{\log d}{nd}\right]^{1-\delta/2},$$

where $\check{\mathbf{Q}}$ *denotes any estimator of* $\mathbf{Q}$ *based on* $N_2, \ldots, N_p$.

Remark 4 The lower bound in Theorem 3 matches the convergence rate of the ITSPCA estimator in Theorem 2, and so we conclude that the ITSPCA estimator achieves the optimal convergence rate under the Frobenius and spectral norms. That is, under the sparsity condition, the convergence rate, $\pi(d)^{1/2} \left[n^{-1} d^{-1} \log d\right]^{1/2-\delta/4}$, of the ITSPCA estimator is optimal, while the convergence rate, $n^{-1/2}$, of the ordinary PCA estimator is sub-optimal. On the other hand, without the sparsity assumption on eigenspace, that is, $\pi(d) = d$ and $\delta = 0$, we can show that the minimax lower bound for estimating the eigenspace of $\boldsymbol{\rho}$ is $n^{-1/2}$. Thus, the upper bound of the ordinary PCA estimator in Theorem 1 is the optimal rate.

Remark 5 To derive the lower bound in Theorem 3, we consider a special subclass of $\mathbf{Q}$, and take $\boldsymbol{\rho} = m^{-1}\mathbf{Q}\mathbf{Q}^\dagger$, and then as usual we apply Fano's lemma to obtain the minimax lower bound (see also Birnbaum *et al.* (2013) and Vu and Lei (2013)). The key difference between our approach and those in the literatures is that our observations are characterized by binomial distributions instead of usual normal distributions, and as a result more sophisticated proof arguments in Section 8 are needed to obtain the minimax lower bounds.

Remark 6 When $\delta = 0$, condition (4.3) becomes $\pi(d) = O(d^{1-\mathcal{N}})$ with $\mathcal{N} > 0$, and the minimax lower bounds hold for $\pi(d)$ very close to d. Consider $\delta > 0$, and that d typically grows polynomially or exponentially in n. If d grows exponentially in n, that is, $d = e^{n^\kappa}$, then $\pi(d) = O(d^{(1-\delta/2)-\mathcal{N}}(\log d)^{\delta/2(1-1/\kappa)})$, and $\mathcal{N}$ could be chosen very small value such that $\pi(d)$ is of order d with some positive power. In the case of $d = n^\kappa$, as quantum systems often have large d, we may consider the case of $d \geq n$ and take $\kappa \geq 1$, and thus $\pi(d) = O(d^{(1-\delta/2)-\mathcal{N}-\delta/(2\kappa)}(\log d)^{\delta/2})$, which is of order d with some positive power. Therefore, condition (4.3) is not restrictive.

Remark 7 Koltchinskii and Xia (2015) investigated the optimal convergence rate of low-rank density matrix estimators under the general low-rank density matrix class. For example, under the Pauli basis, Theorem 10 in Koltchinskii and Xia (2015) shows that the optimal rate of estimating low-rank density matrices is $n^{-1/2}$ which we can obtain by the ordinary PCA estimator (see Theorem 5 in Section 5). Their low-rank class includes both the dense and sparse eigenvectors, and thus the minimax rate is coming from the dense sub-class. However, as we showed in Theorems 3 and 4, the rate $n^{-1/2}$ is not optimal under the sparse condition (3.5). Also their analysis focused on estimating a low-rank density matrix itself. On the other hand, this paper devotes to investigating the eigenspace estimation problem under the sparse condition (3.5). This analysis implies that the optimal rate of estimating low-rank density matrices under the sparse condition is $\pi(d)^{1/2}\left[n^{-1}d^{-1}\log d\right]^{1/2-\delta/4}$ (see Theorem 4 in Section 5).

5 Reconstruction of low-rank density matrices

This section proposes low-rank density matrix estimators using the ITSPCA and ordinary PCA methods. We first develop estimators for eigenvalues of the low-rank density matrix $\boldsymbol{\rho}$ as follows:

$$\widehat{\lambda}_\nu^{(R_s)} = \frac{\widetilde{\lambda}_\nu^{(R_s)}}{\sum_{j=1}^r \widetilde{\lambda}_j^{(R_s)}} \quad \text{and} \quad \widehat{\lambda}_\nu^* = \frac{\widetilde{\lambda}_\nu}{\sum_{j=1}^r \widetilde{\lambda}_j} \quad \text{for } \nu = 1, \ldots, r,$$

where

$$\widetilde{\lambda}_\nu^{(R_s)} = \max\left[(\widehat{\mathbf{q}}_\nu^{(R_s)})^\dagger \widehat{\boldsymbol{\rho}}\, \widehat{\mathbf{q}}_\nu^{(R_s)}, 0\right], \qquad \widetilde{\lambda}_\nu = \max\left[\widehat{\mathbf{q}}_\nu^\dagger \widehat{\boldsymbol{\rho}}\, \widehat{\mathbf{q}}_\nu, 0\right],$$

and $\widehat{\mathbf{q}}_\nu^{(R_s)}$ and $\widehat{\mathbf{q}}_\nu$ are the ν-th column of $\widehat{\mathbf{Q}}^{(R_S)}$ and $\widehat{\mathbf{Q}}$, respectively. Note that $\widehat{\lambda}_\nu^{(R_s)}$ and $\widetilde{\lambda}_\nu$ are non-negative, and the sum of each set of estimated eigenvalues is 1. Using the eigenvalue and eigenspace estimators, we can reconstruct the low-rank density matrix as follows:

$$\widehat{\boldsymbol{\rho}}^{(R_s)} = \sum_{\nu=1}^r \widehat{\lambda}_\nu^{(R_s)} \widehat{\mathbf{q}}_\nu^{(R_s)} (\widehat{\mathbf{q}}_\nu^{(R_s)})^\dagger \quad \text{and} \quad \widehat{\boldsymbol{\rho}}^* = \sum_{\nu=1}^r \widehat{\lambda}_\nu^* \widehat{\mathbf{q}}_\nu \widehat{\mathbf{q}}_\nu^\dagger.$$

These two density matrix estimators satisfy three properties of the density matrix in Section 2.1.

The following theorems provide the convergence rates of eigenvalue estimators $\widehat{\lambda}_\nu^{(R_s)}$ and $\widehat{\lambda}_\nu^*$, and low-rank density matrix estimators $\widehat{\boldsymbol{\rho}}^{(R_s)}$ and $\widehat{\boldsymbol{\rho}}^*$.

Theorem 4 *Under the assumptions of Theorem 2 for ITSPCA, we have for* $\nu = 1, \ldots, r$,

$$E\left[|\widehat{\lambda}_\nu^{(R_s)} - \lambda_\nu|\right] \leq C\pi(d)^{1/2}\tau_n^{1-\delta/2} \tag{5.1}$$

and

$$E\left[\|\widehat{\boldsymbol{\rho}}^{(R_s)} - \boldsymbol{\rho}\|_F\right] \leq C\pi(d)^{1/2}\tau_n^{1-\delta/2}, \tag{5.2}$$

where C *is a generic constant free of* n *and* d.

Theorem 5 *Under the assumptions of Theorem 1 for ordinary PCA, we have for* $\nu = 1, \ldots, r$,

$$E\left[|\widehat{\lambda}_\nu^* - \lambda_\nu|\right] \leq C(n^{-1} \vee (nd)^{-1/2}) \quad \text{and} \quad E\left[\|\widehat{\boldsymbol{\rho}}^* - \boldsymbol{\rho}\|_F\right] \leq Cn^{-1/2},$$

where C *is a generic constant free of* n *and* d.

Remark 8 When $\delta = 0$, the convergence rate of $E\left[\|\widehat{\boldsymbol{\rho}}^{(R_s)} - \boldsymbol{\rho}\|_F\right]$ is $\pi(d)^{1/2}d^{-1/2}\left(\frac{\log(d\vee n)}{n}\right)^{1/2}$ which is the same as the convergence rate of the optimal density matrix estimator under the sparse representation in Theorem 1 (Cai *et al.* (2016)). Also under the sparse condition (3.5), the minimax lower bound of estimating low-rank density matrices is $\pi(d)^{1/2}\tau_n^{1-\delta/2}$, which can be established using the same sub-class employed in the proof of Theorem 3.

Remark 9 The convergence rate of $\widehat{\boldsymbol{\rho}}$ is $(d/n)^{1/2}$ under the Frobenius norm (see Lemma 3 in Cai *et al.* (2016)). On the other hand, the low-rank density matrix estimator has convergence rate $n^{1/2}$ which is the optimal rate under the general low-rank density matrix class (Koltchinskii and Xia (2015)).

6 A numerical study

We conducted simulations to check the finite sample performances of the proposed estimators.

6.1 Rank one case

We first considered the case where the density matrix has (3.1) with $r = 1$,

$$\boldsymbol{\rho} = \mathbf{Q}\mathbf{Q}^\dagger = d^{-1}\left(\mathbf{I}_d + \sum_{j=2}^{p}\beta_j\mathbf{B}_j\right),$$

where $\mathbf{Q} \in \mathbb{C}^d$ and $\beta_j = \mathrm{tr}(\boldsymbol{\rho}\mathbf{B}_j)$ for $j = 1, \ldots, d^2$. The eigenvector $\mathbf{Q}$ was generated as follows. First, its $\pi(d)$ components were generated by $\pi(d)$ i.i.d. random variables from $U_1 + U_2\sqrt{-1}$, where U_j's are i.i.d. uniform distributions on $[-1, 1]$, and set the rest $d - \pi(d)$ components to be zero. Then normalize the generated vector by dividing its ℓ_2-norm so that the generated $\mathbf{Q}$ satisfies $\|\mathbf{Q}\|_2 = 1$. We varied $\pi(d)$ from $5\log(d)$ to $d-1$ with $d = 64, 128$. The whole procedure repeated 200 times.

For each simulated dataset, we estimated $\mathbf{Q}$ using the ITSPCA with hard threshold (ITS-H), ITSPCA with soft threshold (ITS-S), DTSPCA, and ordinary PCA algorithms. The MSEs of eigenspace estimator $\widehat{\mathbf{Q}}$ and low-rank density matrix estimator $\widehat{\boldsymbol{\rho}}$, $E\|\sin(\widehat{\mathbf{Q}}, \mathbf{Q})\|_F^2$ and $E\|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_F^2$, were calculated by averaging the corresponding squared norms of $\widehat{\mathbf{Q}}$ and $\widehat{\boldsymbol{\rho}}$ over 200 runs. For the ITSPCA and DTSPCA algorithms in Algorithm 1 and 2, respectively, we selected tuning parameters (C_α, C_γ), to be $(0.1, 2), (0.5, 1)$, and (0.1) for ITS-H, ITS-S, and DTSPCA, respectively, by searching in the range of $\{3, 2.5, \ldots, 0.5, 0.1\}^2$ for minimizing MSE. We used hard thresholding $\mathcal{T}_H(t, \gamma) = t\mathbf{1}_{(|t|>\gamma)}$ and soft threshold $\mathcal{T}_S(t, \gamma) = e^{\sqrt{-1}\theta}\max(0, |t| - \gamma)$ for the thresholding step in Algorithm 1 for the ITS-H and ITS-S, respectively, where $t = |t|e^{\sqrt{-1}\theta}$, and stopped iterating once $\|\sin(\widehat{\mathbf{Q}}^{(k)}, \widehat{\mathbf{Q}}^{(k-1)})\|_2 \leq n^{-1}d^{-1}$.

Figure 1 plots the relative efficiencies of the ITS-H, ITS-S, DTSPCA, and PCA estimators with respect to the PCA estimator against the sample size n for different d and $\pi(d)$, and Figure 2 plots their MSEs against $\pi(d)$ for different n and d. The numerical values of the MSEs are reported in Table 1. Figures 1 and 2 show that the MSEs usually decrease in sample size n; for sparse eigenvectors with $\pi(d) = 5\log d$ or $5d^{1/2}$, ITSPCA estimators usually have superior performance over the DTSPCA and PCA estimators; the MSEs of ITSPCA and DTSPCA estimators become worse as $\pi(d)$ increases, while the performance of the PCA estimator is robust against $\pi(d)$ and better than ITSPCA estimators in the non-sparse case where $\pi(d) = d - 1$. MSEs for the density matrix estimators in Table 1 have the similar patterns to the eigenspace estimators, and $\widehat{\boldsymbol{\rho}}$ in (2.1) has much worse performance than the PCA estimator.

6.2 Rank four case

We simulated density matrix using (3.1) with $r = 4$. Specifically, choose arbitrary eigenspace $\mathbf{Q}_0 \in \mathbb{V}_{\pi(d),4}$, where $\mathbb{V}_{h,k}$ is the Stiefel manifold of h-by-k orthonormal matrices. First, we generated a $\pi(d)$-by-$\pi(d)$ positive definite Hermitian matrix from uniform random variables, for example, diagonal elements are 1 and for the off-diagonal elements, the (h, k)-th and (k, h)-th elements are $U_1 + \sqrt{-1}U_2$, where U_i's follow uniform distributions on $(-\sqrt{0.5}, \sqrt{0.5})$. Then form d-by-4 matrix $\mathbf{Q} = (\mathbf{Q}_0^T, 0)^T$. We varied $\pi(d)$ from $5\log(d)$ to $d-1$. Eigenvalues $\boldsymbol{\Lambda}$ are chosen from $(0.25, 0.25, 0.25, 0.25)$, $(0.4, 0.3, 0.2, 0.1)$,

Table 1: The MSEs of ITS-H, ITS-S, DTSPCA, and PCA estimators and corresponding low-rank density matrix estimators for $d = 128$, and $n = 100, 200, 500, 1000, 2000$.

			MSE (eigenspace)$\times 10^2$				MSE (density matrix)$\times 10^2$				
d	$\pi(d)$	n	ITS-H	ITS-S	DTSPCA	PCA	ITS-H	ITS-S	DTSPCA	PCA	$\widehat{\boldsymbol{\rho}}$
64	$5\log(d)$	100	0.3008	0.4383	1.1629	0.9627	0.6016	0.8767	2.3258	1.9254	63.0741
		200	0.1453	0.2155	0.5489	0.4909	0.2906	0.4310	1.0977	0.9819	31.5377
		500	0.0577	0.0878	0.2251	0.1929	0.1153	0.1757	0.4501	0.3858	12.6377
		1000	0.0284	0.0443	0.1163	0.0948	0.0568	0.0885	0.2327	0.1895	6.3216
		2000	0.0142	0.0224	0.0438	0.0485	0.0284	0.0447	0.0877	0.0970	3.1531
	$5d^{1/2}$	100	0.8961	1.0046	4.8367	0.9680	1.7921	2.0093	9.6733	1.9359	63.0148
		200	0.4082	0.5295	2.4086	0.4855	0.8165	1.0590	4.8173	0.9709	31.5230
		500	0.1347	0.2129	1.0878	0.1941	0.2693	0.4259	2.1756	0.3882	12.6007
		1000	0.0593	0.1085	0.4885	0.0961	0.1186	0.2170	0.9771	0.1922	6.2871
		2000	0.0304	0.0576	0.2433	0.0484	0.0607	0.1151	0.4866	0.0969	3.1456
	$d-1$	100	1.3800	1.5670	12.3886	0.9876	2.7600	3.1340	24.7772	1.9751	63.0856
		200	0.6485	0.7903	6.3141	0.4871	1.2970	1.5806	12.6282	0.9742	31.5273
		500	0.1952	0.3325	2.4912	0.1930	0.3905	0.6649	4.9825	0.3860	12.5754
		1000	0.0957	0.1713	1.3645	0.0971	0.1914	0.3427	2.7290	0.1943	6.3095
		2000	0.0485	0.0881	0.5835	0.0493	0.0970	0.1762	1.1671	0.0985	3.1518
128	$5\log(d)$	100	0.2084	0.3270	1.4792	0.9979	0.4169	0.6539	2.9584	1.9958	127.2219
		200	0.0883	0.1600	0.6500	0.4916	0.1765	0.3200	1.3001	0.9832	63.4730
		500	0.0360	0.0691	0.2996	0.1954	0.0721	0.1382	0.5991	0.3909	25.3742
		1000	0.0182	0.0359	0.1593	0.0990	0.0364	0.0718	0.3186	0.1979	12.6988
		2000	0.0090	0.0188	0.0775	0.0498	0.0180	0.0377	0.1550	0.0995	6.3408
	$5d^{1/2}$	100	0.5461	0.7029	4.8951	0.9819	1.0921	1.4058	9.7902	1.9638	126.9431
		200	0.2097	0.3514	2.2088	0.4856	0.4195	0.7029	4.4177	0.9711	63.5284
		500	0.0841	0.1468	0.9327	0.1968	0.1682	0.2937	1.8654	0.3936	25.4182
		1000	0.0430	0.0760	0.5597	0.0988	0.0860	0.1519	1.1195	0.1977	12.7027
		2000	0.0216	0.0388	0.2378	0.0496	0.0433	0.0776	0.4756	0.0993	6.3566
	$d-1$	100	1.6628	1.4878	19.1328	0.9926	3.3256	2.9755	38.2656	1.9851	127.0068
		200	0.7028	0.7814	11.4092	0.4903	1.4056	1.5629	22.8184	0.9807	63.5389
		500	0.2354	0.3478	5.1377	0.1964	0.4708	0.6957	10.2754	0.3929	25.4193
		1000	0.1144	0.1838	2.5418	0.0979	0.2287	0.3676	5.0837	0.1957	12.7010
		2000	0.0553	0.0962	1.2805	0.0488	0.1106	0.1924	2.5610	0.0976	6.3562

$(0.5, 0.3, 0.19, 0.01)$. Then we obtain the density matrix as follows:

$$\boldsymbol{\rho} = \sum_{\nu=1}^{4} \lambda_\nu \mathbf{q}_\nu \mathbf{q}_\nu^\dagger = d^{-1}\left(\mathbf{I}_d + \sum_{j=2}^{d^2} \beta_j \mathbf{B}_j\right),$$

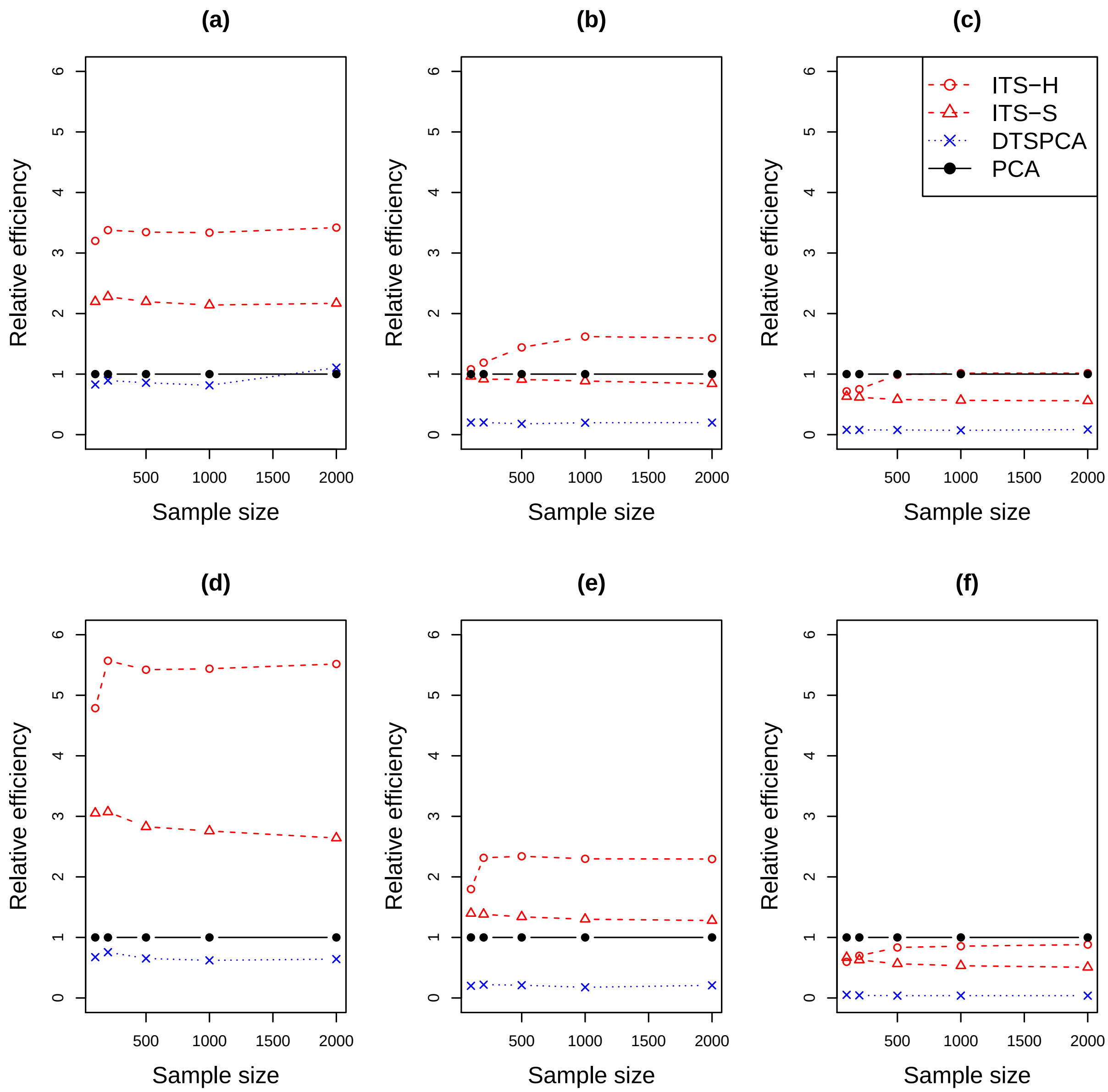

Figure 1: Plots of relative efficiencies against sample size for the ITS-H, ITS-S, DTSPCA, and PCA estimators with respect to the PCA estimator for $\pi(d) = 5\log(d), 5d^{1/2}, d-1$ with $d = 64$ and 128. (a)-(c) are plots of relative efficiencies based on the Frobenius norm for $\pi(d) = 5\log(d), 5d^{1/2}, d-1$, respectively, with $d = 64$. (d)-(f) are plots of relative efficiencies based on the Frobenius norm for $\pi(d) = 5\log(d), 5d^{1/2}, d-1$, respectively, with $d = 128$.

where $\mathbf{Q} = (\mathbf{q}_1, \ldots, \mathbf{q}_4), d = 2^7$. With $\boldsymbol{\rho}$ above, we computed $\beta_j = \mathrm{tr}(\boldsymbol{\rho}\mathbf{B}_j)$ for $j = 1, \ldots, 2^{14}$, where $\mathbf{B}_j$'s are Pauli's matrices. For each simulated dataset, we estimated $\mathbf{Q}$ for $m = 4$ and used the same scheme as the rank one case.

Figure 3 plots the relative efficiencies of the ITS-H, ITS-S, DTSPCA, and PCA esti-

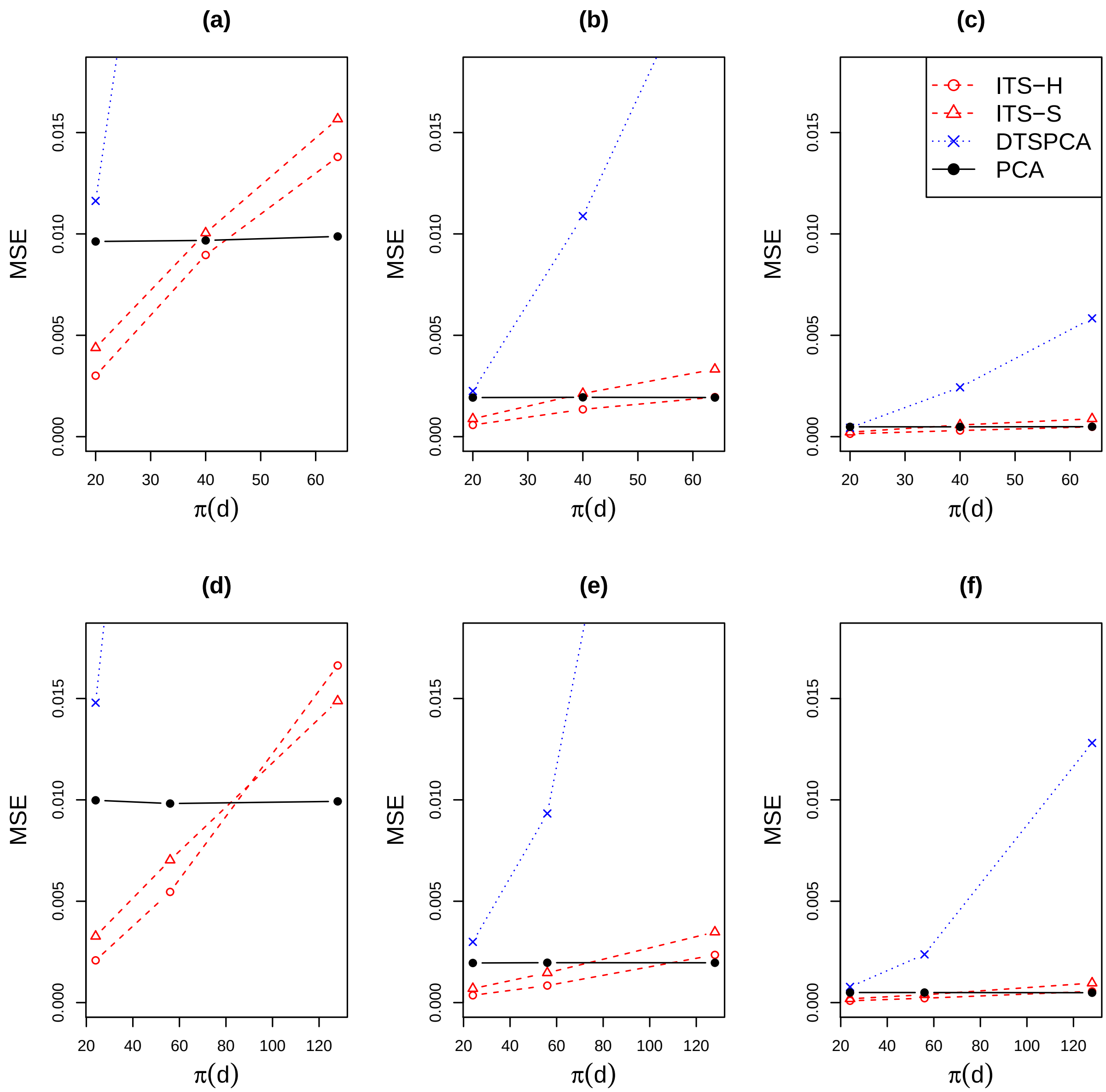

Figure 2: Plots of MSE against $\pi(d)$ for the ITS-H, ITS-S, DTSPCA, and PCA estimators for $n = 100, 500, 2000$ and $d = 64, 128$. (a)-(c) are plots of MSEs based on the Frobenius norm for $n = 100, 500, 2000$, respectively, with $d = 64$. (d)-(f) are plots of MSEs based on the Frobenius norm for $n = 100, 500, 2000$, respectively, with $d = 128$.

mators with respect to the PCA estimator against the sample size n for different $\pi(d)$ and eigenvalues $\mathbf{\Lambda}$, and Figure 2 plots their MSEs against $\pi(d)$ for different sample size and eigenvalues $\mathbf{\Lambda}$. The numerical values of the MSEs are reported in Table 2. As the effects of n and $\pi(d)$ are similar to the rank one case, we focus on the effects of the magnitude of the fourth eigenvalue λ_4. Figures 3 and 4 show that the MSEs decrease in difference of eigenvalues; when the difference of eigenvalues is small like $\lambda_4 = 0.01$, all of the estimators show worse performance, and so the relative efficiencies are very close.

7 An empirical study

We conducted a Monte Carlo simulation to analyze the density matrices which are estimated by Häffner *et al.* (2005). We used two density matrices with $d = 2^7$ and 2^8. Denote by $\boldsymbol{\rho}_7$ and $\boldsymbol{\rho}_8$ density matrices with $d = 2^7$ and $d = 2^8$, respectively. Based on the density matrix $\boldsymbol{\rho}$, we first calculated $\beta_j = \mathrm{tr}(\boldsymbol{\rho}\mathbf{B}_j)$, where $\mathbf{B}_j$'s are Pauli's matrices and then, generated n Pauli measurements for each Pauli's matrix. With generated Pauli measurements, we estimated $\boldsymbol{\rho}$ by ITS-H, ITS-S, DTSPCA, and PCA. The tuning parameters were used $(0.1, 2), (0.5, 1)$, and (0.1) for ITS-H, ITS-S, and DTSPCA, respectively, and we used the result of the rank test in Kim and Wang (2017) to determine the rank r. We varied n from 100 to 2000. The whole procedure repeated 200 times.

Figure 5 plots the absolute values of elements of eigenvectors corresponding to the first six eigenvectors for $\boldsymbol{\rho}_7$ and $\boldsymbol{\rho}_8$, and the first six eigenvalues for $\boldsymbol{\rho}_7$ and $\boldsymbol{\rho}_8$ are $(0.7825, 0.0605, 0.0445, 0.0324, 0.023, 0.0167)$ and $(0.7514, 0.0609, 0.0456, 0.04, 0.0233, 0.0189)$, respectively, which show that the density matrices, $\boldsymbol{\rho}_7$ and $\boldsymbol{\rho}_8$, have a low-rank and sparse eigenvectors. That is, the assumptions in this paper may be satisfied, and the iterative thresholding methods (ITSPCA and DTSPCA) may work well.

Table 3 presents MSEs of ITS-H, ITS-S, DTSPCA, and PCA density estimators. Figure 6 plots relative efficiencies with respect to the PCA estimators against the sample size for $d = 128$ and 256. From Table 3 and Figure 6, we can see that MSEs decrease according to the sample size n, and the iterative thresholding methods usually have smaller MSEs than PCA density matrix estimator or $\widehat{\boldsymbol{\rho}}$.

8 Proofs

Denote by C and C_1 generic constants whose values are free of n and p and may change from appearance to appearance.

Table 2: The MSEs of ITS-H, ITS-S, DTSPCA, and PCA estimators and corresponding low-rank density matrix estimators for $n = 100, 200, 500, 1000, 2000$, $\mathbf{\Lambda} = (0.25, 0.25, 0.25, 0.25)$, $(0.4, 0.3, 0.2, 0.1)$, $(0.5, 0.3, 0.19, 0.01)$, and $\pi(d) = 5\log d, 5d^{1/2}, d-1$ with $d = 128$.

			MSE (eigenspace)				MSE (density matrix)				
$\pi(d)$	$\mathbf{\Lambda}$	n	ITS-H	ITS-S	DTSPCA	PCA	ITS-H	ITS-S	DTSPCA	PCA	$\widehat{\boldsymbol{\rho}}$
$5\log d$	$(0.25, 0.25, 0.25, 0.25)$	100	0.1861	0.1941	0.3532	0.6215	0.0240	0.0248	0.0448	0.0781	1.2792
		200	0.0722	0.0925	0.1724	0.3099	0.0095	0.0119	0.0220	0.0391	0.6386
		500	0.0231	0.0353	0.0648	0.1240	0.0031	0.0046	0.0083	0.0157	0.2555
		1000	0.0109	0.0175	0.0310	0.0622	0.0015	0.0023	0.0040	0.0079	0.1277
		2000	0.0052	0.0085	0.0151	0.0309	0.0007	0.0011	0.0019	0.0039	0.0638
	$(0.4, 0.3, 0.2, 0.1)$	100	0.5935	0.5626	0.7734	1.2608	0.0309	0.0345	0.0519	0.0902	1.2787
		200	0.2444	0.2491	0.3923	0.6903	0.0128	0.0150	0.0249	0.0436	0.6384
		500	0.0846	0.0903	0.1488	0.2758	0.0044	0.0055	0.0092	0.0165	0.2553
		1000	0.0415	0.0440	0.0721	0.1379	0.0019	0.0027	0.0044	0.0080	0.1276
		2000	0.0195	0.0211	0.0344	0.0687	0.0009	0.0013	0.0020	0.0040	0.0638
	$(0.5, 0.3, 0.19, 0.01)$	100	1.1132	1.1137	1.2272	1.4019	0.0257	0.0321	0.0527	0.0851	1.2782
		200	1.0371	1.0486	1.0997	1.1951	0.0143	0.0170	0.0275	0.0455	0.6378
		500	1.0031	1.0125	1.0219	1.0736	0.0070	0.0078	0.0117	0.0194	0.2552
		1000	0.9887	0.9975	0.9873	1.0310	0.0040	0.0044	0.0057	0.0100	0.1275
		2000	0.9741	0.9817	0.9514	1.0041	0.0022	0.0024	0.0029	0.0051	0.0637
$5d^{1/2}$	$(0.25, 0.25, 0.25, 0.25)$	100	0.6163	0.4233	0.5953	0.6201	0.0775	0.0533	0.0747	0.0779	1.2799
		200	0.2571	0.2169	0.3132	0.3085	0.0325	0.0274	0.0394	0.0389	0.6385
		500	0.0813	0.0882	0.1343	0.1239	0.0104	0.0112	0.0170	0.0157	0.2553
		1000	0.0347	0.0442	0.0683	0.0620	0.0044	0.0056	0.0086	0.0079	0.1276
		2000	0.0152	0.0220	0.0336	0.0308	0.0020	0.0028	0.0042	0.0039	0.0638
	$(0.4, 0.3, 0.2, 0.1)$	100	1.1828	0.9932	1.0824	1.2713	0.0779	0.0631	0.0839	0.0899	1.2790
		200	0.5289	0.4585	0.5603	0.6914	0.0329	0.0306	0.0420	0.0435	0.6382
		500	0.1773	0.1701	0.2300	0.2732	0.0113	0.0117	0.0182	0.0164	0.2553
		1000	0.0815	0.0840	0.1172	0.1369	0.0050	0.0058	0.0098	0.0080	0.1276
		2000	0.0385	0.0421	0.0622	0.0688	0.0021	0.0029	0.0054	0.0040	0.0638
	$(0.5, 0.3, 0.19, 0.01)$	100	1.3123	1.2543	1.3840	1.4003	0.1199	0.1089	0.1554	0.1630	1.3185
		200	1.0565	1.1265	1.1983	1.1961	0.0685	0.0723	0.1259	0.1260	0.6779
		500	0.9264	1.0494	1.0758	1.0763	0.0493	0.0508	0.1095	0.1074	0.2952
		1000	0.8809	1.0208	1.0348	1.0346	0.0440	0.0452	0.1082	0.1037	0.1676
		2000	0.8658	1.0051	1.0188	1.0145	0.0412	0.0422	0.1088	0.1035	0.1037
$d-1$	$(0.25, 0.25, 0.25, 0.25)$	100	1.5437	0.8133	1.1214	0.6187	0.1930	0.1020	0.1401	0.0778	1.2761
		200	0.6946	0.4366	0.6917	0.3079	0.0870	0.0549	0.0865	0.0389	0.6394
		500	0.2316	0.1886	0.3512	0.1239	0.0291	0.0237	0.0439	0.0157	0.2553
		1000	0.0996	0.0986	0.1968	0.0618	0.0125	0.0124	0.0246	0.0078	0.1278
		2000	0.0432	0.0503	0.1020	0.0309	0.0055	0.0063	0.0128	0.0039	0.0638
	$(0.4, 0.3, 0.2, 0.1)$	100	2.0636	1.4548	1.6928	1.2713	0.2020	0.1127	0.1640	0.0901	1.2750
		200	1.1570	0.8458	1.0975	0.6977	0.0902	0.0605	0.1019	0.0437	0.6392
		500	0.4181	0.3307	0.5206	0.2738	0.0287	0.0244	0.0507	0.0165	0.2552
		1000	0.1899	0.1702	0.2761	0.1372	0.0120	0.0124	0.0275	0.0080	0.1278
		2000	0.0875	0.0887	0.1406	0.0698	0.0050	0.0063	0.0139	0.0040	0.0637
	$(0.5, 0.3, 0.19, 0.01)$	100	1.9442	1.5199	1.7585	1.3988	0.1611	0.0965	0.1579	0.0847	1.2743
		200	1.4469	1.2706	1.4723	1.1967	0.0792	0.0535	0.1027	0.0456	0.6389
		500	1.1390	1.1081	1.2426	1.0724	0.0271	0.0241	0.0553	0.0195	0.2552
		1000	1.0472	1.0482	1.1359	1.0301	0.0123	0.0128	0.0330	0.0099	0.1277
		2000	1.0068	1.0109	1.0639	1.0034	0.0055	0.0068	0.0185	0.0051	0.0637

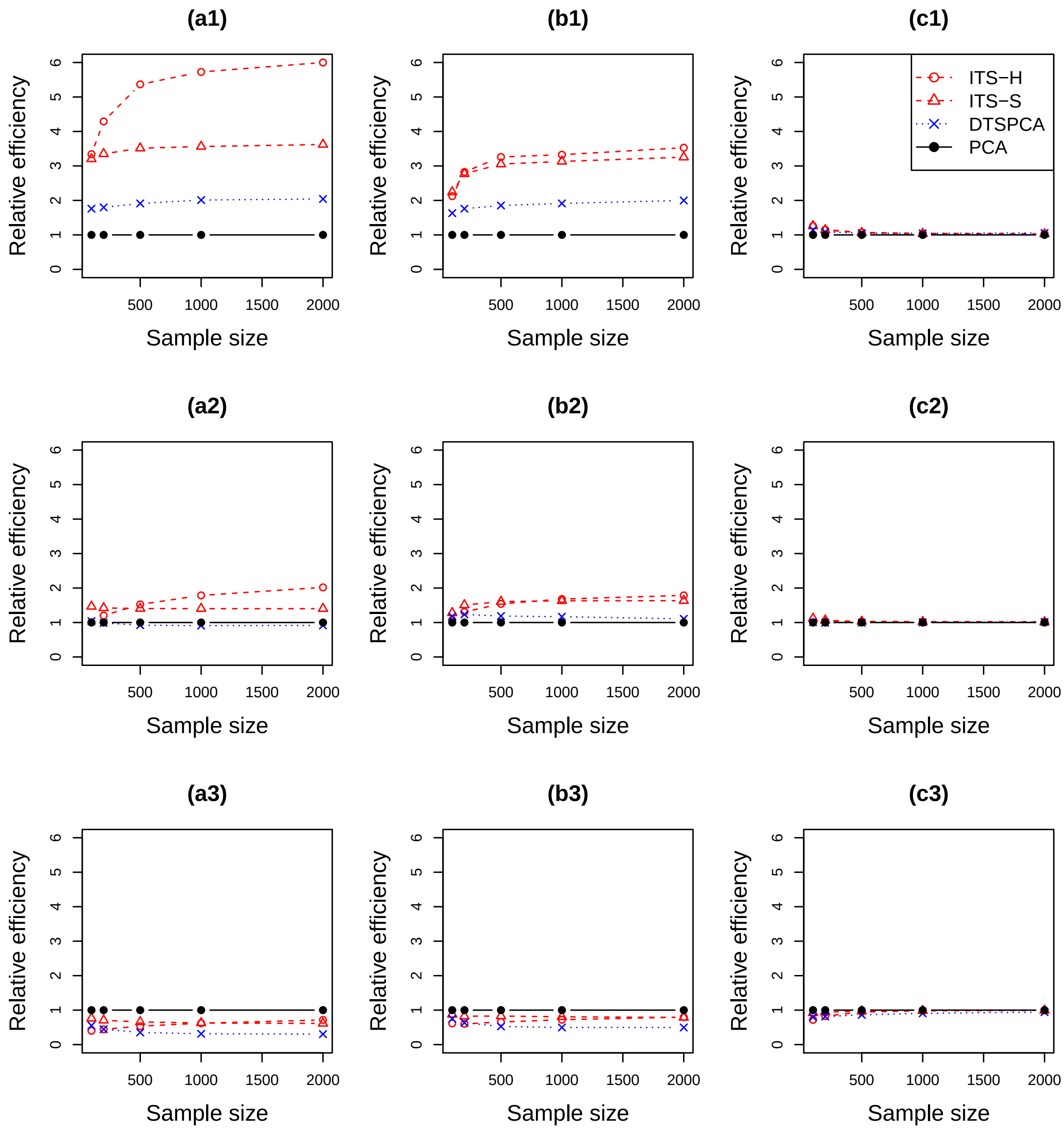

Figure 3: Plots of relative efficiencies against sample size for the ITS-H, ITS-S, DTSPCA, and PCA estimators with respect to the PCA estimator for $\pi(d) = 5\log d, 5d^{1/2}, d-1$ with $d = 128$. (a1)-(a3) are plots of relative efficiencies based on the Frobenius norm for $\pi(d) = 5\log(d), 5d^{1/2}, d-1$, respectively, with $\mathbf{\Lambda} = (0.25, 0.25, 0.25, 0.25)$. (b1)-(b3) are plots of relative efficiencies based on the Frobenius norm for $\pi(d) = 5\log(d), 5d^{1/2}, d-1$, respectively, with $\mathbf{\Lambda} = (0.4, 0.3, 0.2, 0.1)$. (c1)-(c3) are plots of relative efficiencies based on the Frobenius norm for $\pi(d) = 5\log(d), 5d^{1/2}, d-1$, respectively, with $\mathbf{\Lambda} = (0.5, 0.3, 0.19, 0.01)$.

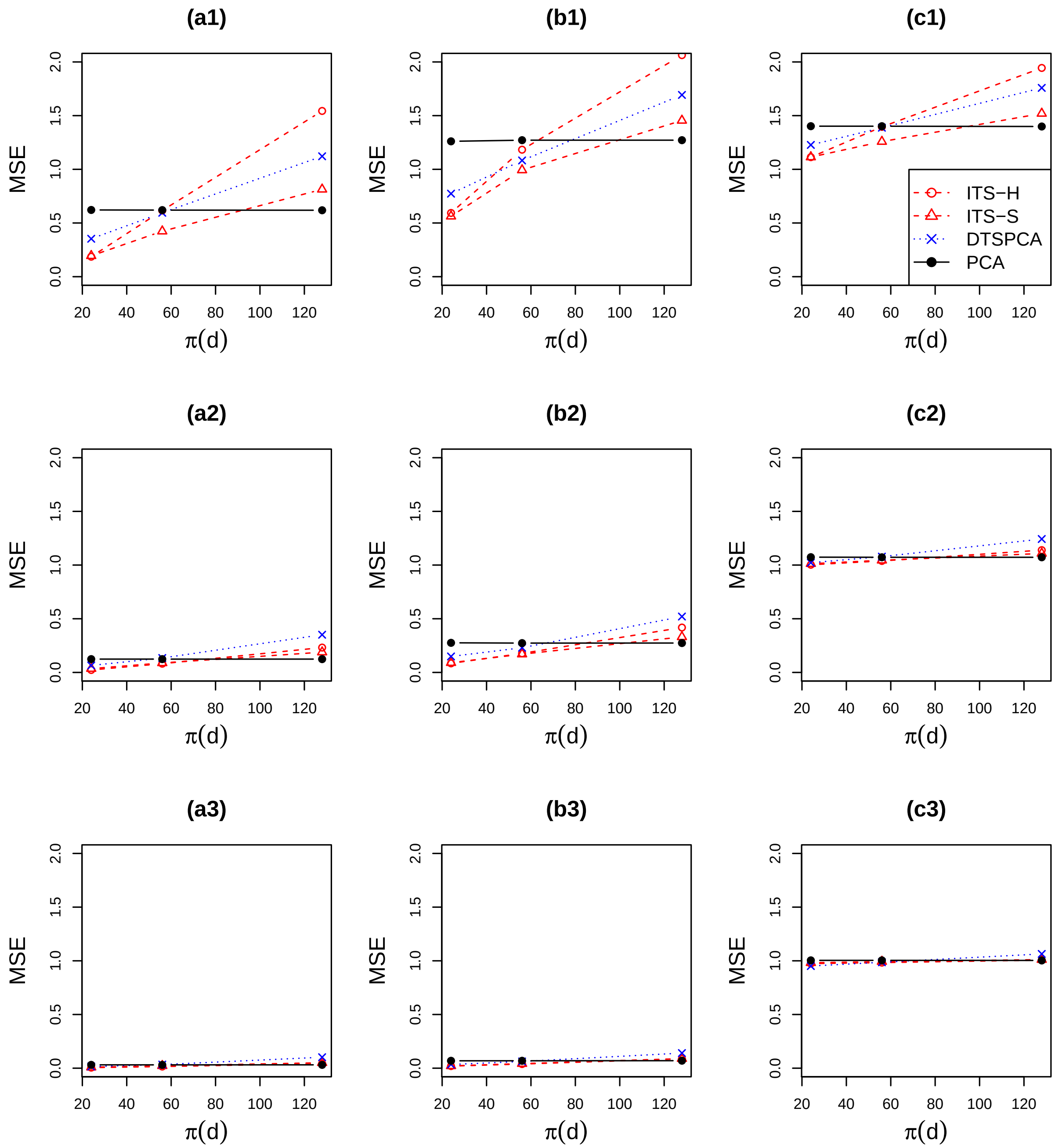

Figure 4: Plots of MSE against $\pi(d)$ for the ITS-H, ITS-S, DTSPCA, and PCA estimators for $n = 100, 500, 2000$ and $d = 128$. (a1)-(a3) are plots of MSEs based on the Frobenius norm for $n = 100, 500, 2000$, respectively, with $\boldsymbol{\Lambda} = (0.25, 0.25, 0.25, 0.25)$. (b1)-(b3) are plots of MSEs based on the Frobenius norm for $n = 100, 500, 2000$, respectively, with $\boldsymbol{\Lambda} = (0.4, 0.3, 0.2, 0.1)$. (c1)-(c3) are plots of MSEs based on the Frobenius norm for $n = 100, 500, 2000$, respectively, with $\boldsymbol{\Lambda} = (0.5, 0.3, 0.19, 0.01)$.

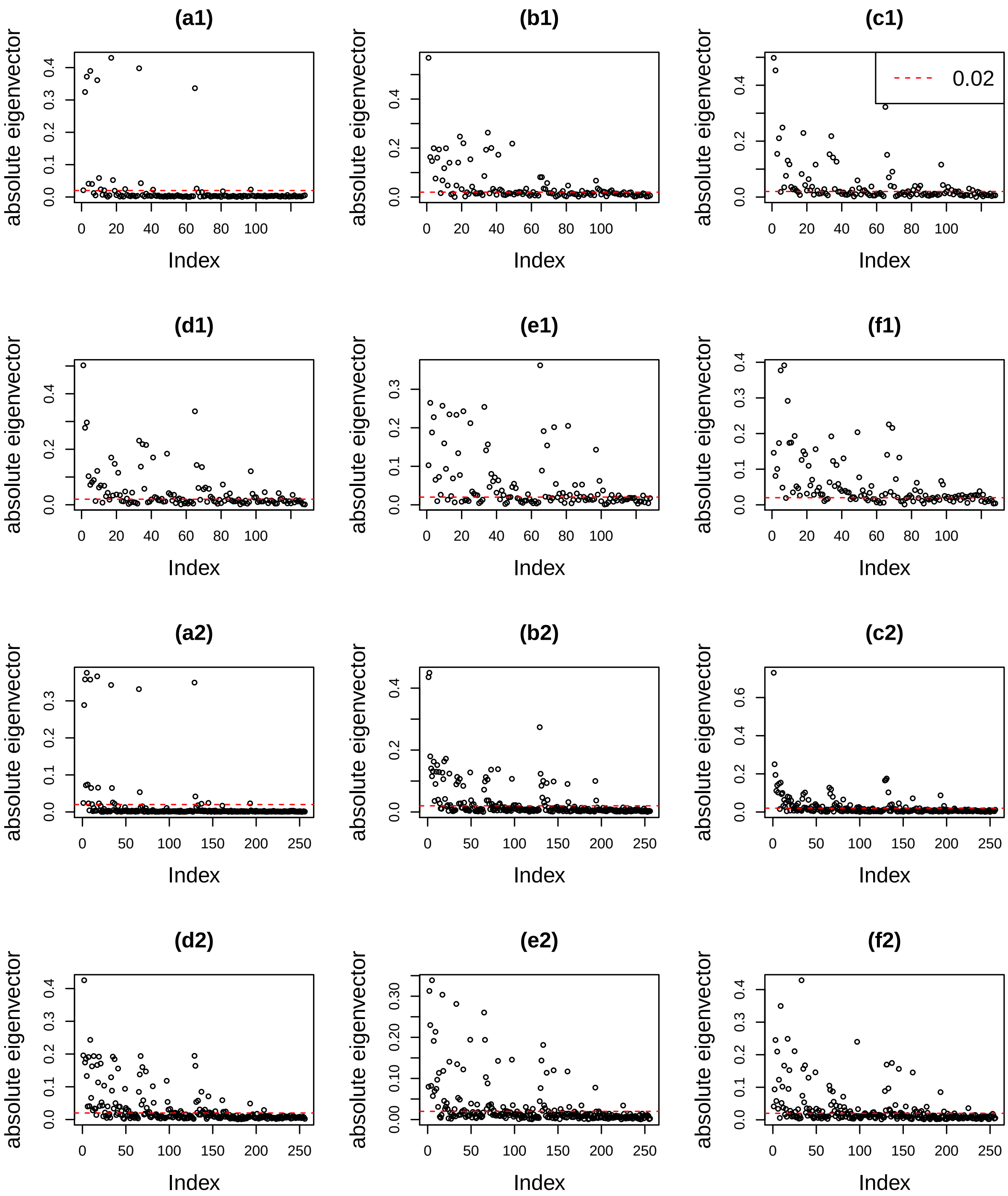

Figure 5: Plots of absolute elements of eigenvectors for the eigenvectors corresponding to the first 6 eigenvalues. (a1)-(f1) are plots for $\boldsymbol{\rho}_7$. (a2)-(f2) are plots for $\boldsymbol{\rho}_8$.

8.1 Proofs of Theorems 1-2

8.1.1 Proof of Theorem 1

Proof of Theorem 1. Consider (3.4). By Davis-Kahn's $\sin\theta$ theorem (Theorem 3.1 in Li (1998b)), we have

$$\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}})\|_F^2 \leq \frac{\|(\boldsymbol{\rho} - \widehat{\boldsymbol{\rho}})\mathbf{Q}\|_F^2}{(\lambda_m - \widehat{\lambda}_{m+1})^2}, \tag{8.1}$$

Table 3: MSEs (Frobenius norm) of ITS-H, ITS-S, DTSPCA, and PCA density matrix estimators for $d = 128, 256$, and $n = 100, 200, 500, 1000, 2000$.

d	n	ITS-H	ITS-S	DTSPCA	PCA	$\widehat{\boldsymbol{\rho}}$
128	100	0.04672	0.04975	0.05157	0.06837	1.27381
	200	0.03360	0.03632	0.03347	0.04897	0.63686
	500	0.02060	0.02060	0.01781	0.02704	0.25442
	1000	0.01233	0.01222	0.01056	0.01557	0.12727
	2000	0.00750	0.00706	0.00630	0.00781	0.06376
256	100	0.04529	0.05616	0.05988	0.09043	2.55323
	200	0.03612	0.03778	0.03966	0.05278	1.27709
	500	0.01995	0.01970	0.01868	0.02876	0.51098
	1000	0.01246	0.01187	0.01041	0.01663	0.25544
	2000	0.00796	0.00758	0.00649	0.00978	0.12770

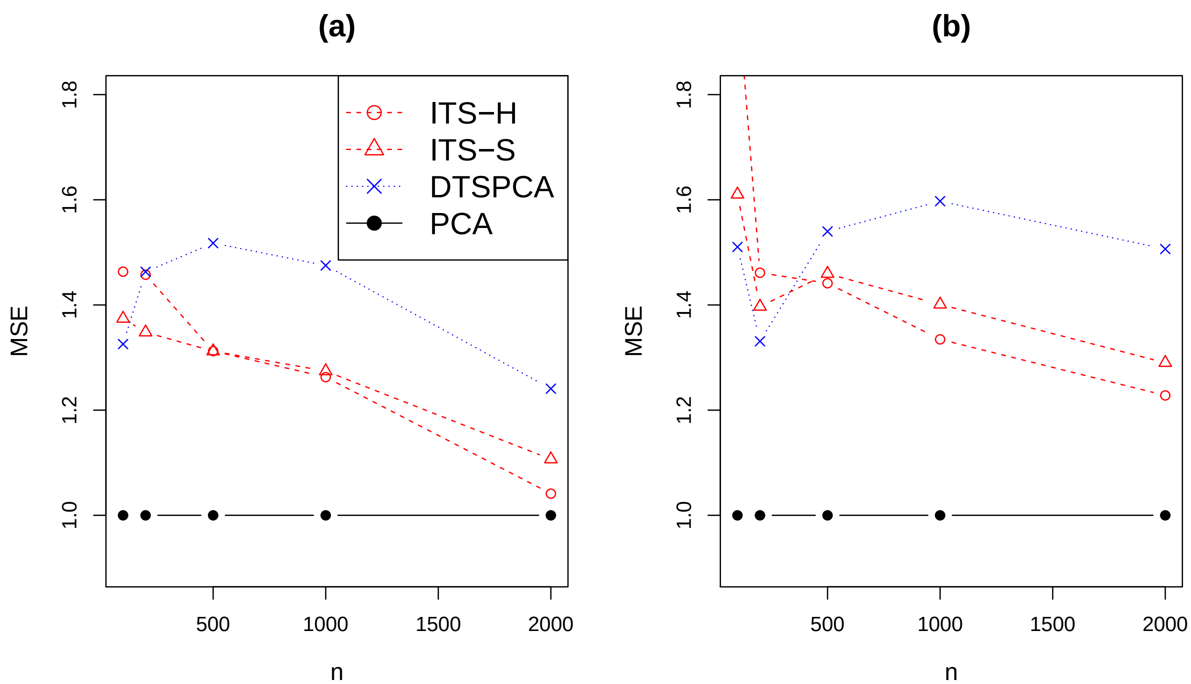

Figure 6: Plots of relative efficiencies against the sample size for the ITS-H, ITS-S, DTSPCA, and PCA estimators with respect to the PCA estimator. (a)-(b) are plots of relative efficiencies based on the Frobenius norm for $d = 128$ and 256, respectively.

where $\widehat{\lambda}_m$ is the m-th eigenvalue of $\widehat{\boldsymbol{\rho}}$. First, consider the denominator on the right hand side. By Weyl's theorem (Theorem 4.3 in Li (1998a)), we have

$$\max_{1\le\nu\le d} |\widehat{\lambda}_\nu - \lambda_\nu| \le \|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_2^2.$$

Simple algebraic manipulations show

$$\max_j \left\| d^{-1}(\widehat{\beta}_j - \beta_j)\mathbf{B}_j \right\|_2 \leq \frac{2}{d}$$

and

$$\left\| d^{-2} \sum_{j=2}^{p} E\left[(\widehat{\beta}_j - \beta_j)^2 \mathbf{B}_j^T \mathbf{B}_j \right] \right\|_2 \leq \frac{1}{n}.$$

Then, by the Matrix Bernstein's inequality (Theorem 6.1 in Tropp (2012)), we have

$$\begin{aligned} P\left(\|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_2 \geq t\right) &= P\left(\left\| d^{-1} \sum_{j=2}^{p} (\widehat{\beta}_j - \beta_j) B_j \right\|_2 \geq t \right) \\ &\leq 2d \exp\left(-\frac{t^2/2}{n^{-1} + 2t/(3d)} \right). \end{aligned}$$

Take $t = \sqrt{6 \log d/n}$. We have

$$P\left(\|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_2 \geq \sqrt{6 \log d/n}\right) \leq 2d^{-2}. \tag{8.2}$$

Consider the numerator in (8.1). Let $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{C}^d$ such that $\|\mathbf{a}\|_2^2 = 1$. Since $(\widehat{\beta}_j - \beta_j)$'s are independent with mean zero, we have

$$\begin{aligned} E\left[\|(\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho})\mathbf{a}\|_2^2\right] &= \frac{1}{d^2} \sum_{j=2}^{p} E\left[(\widehat{\beta}_j - \beta_j)^2 \right] \|\mathbf{B}_j \mathbf{a}\|_2^2 \\ &= \frac{1}{d^2} \sum_{j=2}^{p} \frac{1 - \beta_j^2}{n} \\ &= \frac{1}{n} - \frac{\sum_{\nu=1}^{r} \lambda_\nu^2}{dn}. \end{aligned}$$

Then, since $\|\mathbf{q}_\nu\|_2^2 = 1$, we have

$$\begin{aligned} E\left[\|(\boldsymbol{\rho} - \widehat{\boldsymbol{\rho}})\mathbf{Q}\|_F^2\right] &= \sum_{\nu=1}^{m} E\left[\|(\boldsymbol{\rho} - \widehat{\boldsymbol{\rho}})\mathbf{q}_\nu\|_2^2\right] \\ &= m\left(\frac{1}{n} - \frac{\sum_{\nu=1}^{r} \lambda_\nu^2}{dn} \right). \end{aligned} \tag{8.3}$$

Then, since $\|\sin(\widehat{\mathbf{Q}}, \mathbf{Q})\|_F^2 \leq m$, we have

$$\begin{aligned} E\left[\|\sin(\widehat{\mathbf{Q}}, \mathbf{Q})\|_F^2\right] &= E\left[\|\sin(\widehat{\mathbf{Q}}, \mathbf{Q})\|_F^2 \mathbf{1}_E\right] + E\left[\|\sin(\widehat{\mathbf{Q}}, \mathbf{Q})\|_F^2 \mathbf{1}_{E^c}\right] \\ &\leq \frac{4m}{d^2} + E\left[\|(\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho})\mathbf{Q}\|_F^2\right] \left(\lambda_m - \lambda_{m+1} - \sqrt{\frac{6 \log d}{n}} \right)^{-2} \end{aligned}$$

$$
\begin{aligned}
&\le \frac{4m}{n} + \frac{m}{n}\left(\lambda_m - \lambda_{m+1} - \sqrt{\frac{6\log d}{n}}\right)^{-2} \\
&= O\left(\frac{n^{-1}}{(\lambda_m - \lambda_{m+1})^2}\right),
\end{aligned}
$$

where $E = \{\max_{1\le\nu\le d} |\widehat{\lambda}_\nu - \lambda_\nu| \ge \sqrt{\frac{6\log d}{n}}\}$, and the second and third inequalities are due to (8.2) and (8.3), respectively. ■

8.1.2 Proof of Theorem 2

Proof of Theorem 2. Define the set of high signal coordinates,

$$
H = H(\tau) = \{\nu : |q_{\nu j}| \ge C_\tau \tau_n, \text{ for some } 1 \le j \le r\},
$$

where C_τ is a constant. Then, similar to the proof of Lemma 3.1 (Ma (2013)), we can show

$$
r \le |H| \le C\pi(p)\tau_n^{-\delta}. \tag{8.4}
$$

In addition, let $L = \{1, \dots, d\} \setminus H$. Here and after, we use an extra superscript "o" to indicate oracle quantities. That is, let

$$
\boldsymbol{\rho} = \begin{bmatrix} \boldsymbol{\rho}_{HH} & \boldsymbol{\rho}_{HL} \\ \boldsymbol{\rho}_{LH} & \boldsymbol{\rho}_{LL} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\rho}^o = \begin{bmatrix} \boldsymbol{\rho}_{HH} & 0 \\ 0 & 0 \end{bmatrix}.
$$

$\widehat{\boldsymbol{\rho}}$ and $\widehat{\boldsymbol{\rho}}^o$ are estimators for $\boldsymbol{\rho}$ and $\boldsymbol{\rho}^o$, respectively. Specifically,

$$
\widehat{\boldsymbol{\rho}} = (\widehat{\rho}_{ij})_{i,j=1,\dots,p} \quad \text{and} \quad \widehat{\boldsymbol{\rho}}^o = \begin{bmatrix} \widehat{\boldsymbol{\rho}}_{HH} & 0 \\ 0 & 0 \end{bmatrix}.
$$

Using Algorithm 1, we construct an oracle sequence of d-by-m orthonormal matrices $\{\widehat{\mathbf{Q}}^{(k),o}, k \ge 1\}$ with the initial $\widehat{\mathbf{Q}}^{(0),o}$. To construct $\widehat{\mathbf{Q}}^{(0),o}$, we use an oracle version of Algorithm 2. Specifically, $S^o = S \cap H$. This ensures that $\widehat{\mathbf{Q}}^{(0),o}_{L\cdot} = 0$.

With probability at least $1 - C_0(d \vee n)^{-2}$, we have

$$
\begin{aligned}
&\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2 \\
&\le C\left\{\|\sin(\mathbf{Q}, \mathbf{Q}^o)\|_F^2 + \|\sin(\mathbf{Q}^o, \widehat{\mathbf{Q}}^o)\|_F^2 + \|\sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(R_s),o})\|_F^2 + \|\sin(\widehat{\mathbf{Q}}^{(R_s),o}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2\right\} \\
&\le C\frac{\pi(d)\tau_n^{2-\delta}}{(\lambda_m - \lambda_{m+1})^2},
\end{aligned}
$$

where the fist inequality is due to the triangular inequality and Jensen's inequality, and the last inequality is from Propositions 1-4 below. ■

Proposition 1 *Under assumptions of Theorem 2, we have*

$$
\|\sin(\mathbf{Q}, \mathbf{Q}^o)\|_F^2 \le C\frac{\pi(d)\tau_n^{2-\delta}}{(\lambda_m - \lambda_{m+1})^2}.
$$

Proposition 2 *Under assumptions of Theorem 2, we have with probability at least* $1 - C_0(d \vee n)^{-2}$

$$\|\sin(\mathbf{Q}^o, \widehat{\mathbf{Q}}^o)\|_F^2 \le C \frac{\pi(d)\tau_n^{2-\delta}}{(\lambda_m - \lambda_{m+1})^2}.$$

Proposition 3 *Under assumptions of Theorem 2, we have with probability at least* $1 - C_0(d \vee n)^{-2}$,

$$\|\sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(R_s),o})\|_F^2 \le C \frac{\pi(d)\tau_n^{2-\delta}}{(\lambda_m - \lambda_{m+1})^2}.$$

Proposition 4 *Under assumptions of Theorem 2, we have with probability at least* $1 - C_0(d \vee n)^{-2}$,

$$\widehat{\mathbf{Q}}^{(k),o} = \widehat{\mathbf{Q}}^{(k)} \quad \textit{for } k \ge 0.$$

The proofs of above Propositions 1-4 are given in Appendix.

8.2 Proof of Theorem 3

To obtain the lower bound, we consider the real valued density matrix, $\boldsymbol{\rho}$. That is, β_j's corresponding to complex valued Pauli matrices are zero.

We use the following Fano's lemma (Lemma A.5 in Birnbaum *et al.* (2013)).

Lemma 1 (Fano's Lemma) *Denote by* $\{P_\theta : \theta \in \Theta\}$ *a family of probability distribution on a common measurable space, where* Θ *is an arbitrary parameter set. Then, for any finite subset* $\mathcal{G} = \{\theta_1, \ldots, \theta_M\}$ *of* Θ*, we have*

$$\inf_T \sup_{\theta \in \Theta} P_\theta(T \ne \theta) \ge 1 - \inf_F \frac{M^{-1} \sum_{k=1}^M D(P_k \| F) + \log 2}{\log M},$$

where F *is an arbitrary probability distribution,* $P_k = P_{\theta_k}$, T *denotes an arbitrary estimator of* θ *with values in* Θ*, and* $D(P_k \| F)$ *is the Kullback-Leibler (KL) divergence of* F *from* P_k.

Lemma 2 *For* $k = 1, 2$, *let*

$$\boldsymbol{\rho}_k = \frac{1}{d}\mathbf{B}_1 + \frac{1}{d}\sum_{j=2}^p \beta_j^{(k)} \mathbf{B}_j$$

and P_k *be the product of the binomial probability measures,* $B(n, \frac{1+\beta_2^{(k)}}{2}), \ldots, B(n, \frac{1+\beta_p^{(k)}}{2})$. *Then we have*

$$D(P_1 \| P_2) \le n \sum_{j=2}^p \frac{(\beta_j^{(1)} - \beta_j^{(2)})^2}{1 - (\beta_j^{(2)})^2}.$$

Lemma 3 *For $\epsilon \in [0,1]$, the function $\mathbf{A}_\epsilon : \mathbb{V}_{d-m,m} \mapsto \mathbb{V}_{d,m}$ is defined in block form as*

$$\mathbf{A}_\epsilon(\mathbf{J}) = \begin{pmatrix} (1-\epsilon^2)^{1/2}\mathbf{I}_m \\ \epsilon\mathbf{J} \end{pmatrix},$$

where $\mathbb{V}_{d,h} = \{\mathbf{Q} \in \mathbb{R}^{d\times h} : \mathbf{Q}^\dagger\mathbf{Q} = \mathbf{I}\}$ is the Stiefel manifold of d-by-h orthonormal matrices. For $\mathbf{J}_1, \mathbf{J}_2 \in \mathbb{V}_{d-m,m}$, we have

$$\|\sin(\mathbf{A}_\epsilon(\mathbf{J}_1), \mathbf{A}_\epsilon(\mathbf{J}_2))\|_2^2 \geq \epsilon^2(1-\epsilon^2)\|\mathbf{J}_1 - \mathbf{J}_2\|_2^2,$$

and

$$\epsilon^2(1-\epsilon^2)\|\mathbf{J}_1 - \mathbf{J}_2\|_F^2 \leq \|\sin(\mathbf{A}_\epsilon(\mathbf{J}_1), \mathbf{A}_\epsilon(\mathbf{J}_2))\|_F^2 \leq \epsilon^2\|\mathbf{J}_1 - \mathbf{J}_2\|_F^2.$$

Proof: Similar to the proof of Lemma 3 (Kim and Wang (2016)), we can show this statement. ■

Lemma 4 *Let h be an integer satisfying $e \leq h$, and let $s \in [1,h]$. There exists a subset $\{\mathbf{J}_1, \ldots, \mathbf{J}_M\} \subset \mathbb{V}_{h,1}$ satisfying the following properties:*

(1) $\|\mathbf{J}_j - \mathbf{J}_{j'}\|_2^2 \geq 1/4$ *for all* $j \neq j'$;

(2) $\|\mathbf{J}_j\|_0 \leq s$ *for all* j;

(3) $\log M \geq \max\{cs[1+\log(h/s)], \log h\}$, *where* $c > 1/30$ *is an absolute constant.*

Proof: See the proof of Lemma A.5 (Vu and Lei (2013)). ■

Proof of Theorem 3. Since Pauli matrices form an orthogonal basis for all complex Hermitian matrices, for any given $\mathbf{A} \in \mathbb{V}_{d,m}$, where $\mathbb{V}_{d,m}$ is the Stiefel manifold of d-by-m orthonormal matrices, there are $\beta_j's$ such that

$$\boldsymbol{\rho}(\mathbf{A}) = d^{-1}\left(\mathbf{I}_d + \sum_{j=2}^{p}\beta_j\mathbf{B}_j\right) = m^{-1}\mathbf{A}\mathbf{A}^T.$$

We consider the subclass of $\mathbf{A}$ as follows. Let

$$\mathbf{A}_\epsilon(\mathbf{J}) = \begin{pmatrix} (1-\epsilon^2)^{1/2}\mathbf{I}_m \\ \epsilon\mathbf{J} \end{pmatrix}, \tag{8.5}$$

where $\mathbf{I}_m$ is a m-by-m identity matrix, and $\epsilon \in [0,1]$, and $\mathbf{J} \in \mathbb{V}_{d-m,m}$. Using Lemma 4, we construct the packing set of $\mathbf{J}$ as follows. Define $\mathcal{G}_\tau = \{\mathbf{J}_1, \ldots, \mathbf{J}_{M'}\}$ with $h = \lfloor (d-m)/m \rfloor$ and $s = \varrho h$, where $\varrho \in (1/h, 1)$. Then, from Lemma 4, (i) $\log M' \geq c\max\{d\varrho[1 -$

$\log \varrho], \log d\}$ for some constant c free n and p; (ii) $\|\mathbf{J}_i\|_0 \leq s$ for all $j = 1, \ldots, M'$; (iii) $\|\mathbf{J}_j - \mathbf{J}_{j'}\|_2^2 \geq 1/4$ for all $j \neq j'$. Choose $\mathbf{J}$ in (8.5) as follows:

$$\mathbf{J}(\mathbf{a}_1, \ldots, \mathbf{a}_m) = \begin{pmatrix} \mathbf{a}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{a}_2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{a}_m \end{pmatrix},$$

where $\mathbf{a}_j \in \mathcal{G}_\tau$ for all j. Let $\mathcal{G}(\mathbf{J}) = \{\mathbf{J}(\mathbf{a}_1, \ldots, \mathbf{a}_m), \mathbf{a}_j \in \mathcal{G}_\tau \text{ for } j = 1, \ldots, m\}$. Then, from the construction of $\mathcal{G}_\tau$, $\mathcal{G}(\mathbf{J}) \subset \mathbb{V}_{d-m,m}$, and the cardinality of $\mathcal{G}(\mathbf{J})$ is $M = (M')^m$. Note that $\log M \geq mc \max\{d\varrho[1 - \log \varrho], \log d\}$, and for any $\mathbf{J}_k \in \mathcal{G}(\mathbf{J})$, there exist $\beta_j^{(k)}$'s such that

$$\boldsymbol{\rho}(\mathbf{J}_k) = d^{-1}\left(\mathbf{I}_d + \sum_{j=2}^{p} \beta_j^{(k)} \mathbf{B}_j\right) = m^{-1}\mathbf{A}_\epsilon(\mathbf{J}_k)\mathbf{A}_\epsilon(\mathbf{J}_k)^T.$$

Without loss of generality, we assume that the first d Pauli matrices, $\mathbf{B}_j$'s, correspond to the diagonal Pauli matrices. Define P_0 the product of the binomial probability measures, $B(n, \frac{1+\beta_2^{(0)}}{2}), \ldots, B(n, \frac{1+\beta_p^{(0)}}{2})$ with $\beta_j^{(0)}$'s determined as follows:

$$\beta_{d+1}^{(0)} = \cdots = \beta_p^{(0)} = 0$$

and $\beta_1^{(0)}, \ldots, \beta_d^{(0)}$ are a solution of the following equation,

$$\boldsymbol{\rho}_0 = \frac{1}{d}\sum_{j=1}^{d} \beta_j^{(0)} \mathbf{B}_j = m^{-1}\begin{pmatrix} (1-\epsilon^2)\mathbf{I}_m & 0 \\ 0 & \frac{m\epsilon^2}{d-m}\mathbf{I}_{d-m} \end{pmatrix}.$$

Let $\boldsymbol{\beta}^{(0)} = (\beta_1^{(0)}, \ldots, \beta_d^{(0)})^T$ and $\boldsymbol{\beta}^{(k)} = (\beta_1^{(k)}, \ldots, \beta_d^{(k)})^T$, and $\mathbf{H} = (\mathbf{b}_1, \ldots, \mathbf{b}_d)$, where $\mathbf{b}_j = \text{diag}(\mathbf{B}_j)$ for $j = 1, \ldots, d$. Then, by the construction of the Pauli matrices, $\mathbf{H}$ is d-by-d Hadamard matrix. We have

$$\boldsymbol{\beta}^{(0)} = \mathbf{H}^T \text{diag}(\boldsymbol{\rho}_0) \text{ and } \boldsymbol{\beta}^{(k)} = \mathbf{H}^T \text{diag}(\boldsymbol{\rho}(\mathbf{J}_k)).$$

Then

$$\begin{aligned} \sum_{j=2}^{d} |\beta_j^{(k)} - \beta_j^{(0)}|^2 &= \|\mathbf{H}^T[\text{diag}(\boldsymbol{\rho}(\mathbf{J}_k)) - \text{diag}(\boldsymbol{\rho}_0)]\|_2^2 \\ &= d[\text{diag}(\boldsymbol{\rho}(\mathbf{J}_k)) - \text{diag}(\boldsymbol{\rho}_0)]^T[\text{diag}(\boldsymbol{\rho}(\mathbf{J}_k)) - \text{diag}(\boldsymbol{\rho}_0)] \\ &\leq 2m^{-1}d\epsilon^4, \end{aligned} \tag{8.6}$$

where the second equality is established by the fact that $\mathbf{H}^T\mathbf{H} = d\,\mathbf{I}_d$. Note that $|\beta_j^{(0)}| \leq 1 - \epsilon^2/2$ for all $j = 2, \ldots, d$. For off-diagonal terms, we have for any $k = 1, \ldots, M$,

$$\|\boldsymbol{\rho}(\mathbf{J}_k) - \boldsymbol{\rho}_0\|_F^2 = d\sum_{j=1}^{p} |\beta_j^{(k)} - \beta_j^{(0)}|^2$$

$$
\begin{aligned}
&= m^{-2}\left\|\begin{pmatrix} 0 & (1-\epsilon^2)^{1/2}\epsilon\mathbf{J}_k^T \\ (1-\epsilon^2)^{1/2}\epsilon\mathbf{J}_k & \epsilon^2\mathbf{J}_k\mathbf{J}_k^T - \frac{m\epsilon^2}{d-m}\mathbf{I}_{d-m}\end{pmatrix}\right\|_F^2 \\
&= m^{-1}[2(1-\epsilon^2)\epsilon^2 + \epsilon^4 + m\epsilon^4/(d-m)] \le 2m^{-1}\epsilon^2.
\end{aligned}
$$

So, we have

$$
\sum_{j=d+1}^{p} |\beta_j^{(k)} - \beta_j^{(0)}|^2 \le 2m^{-1}d\epsilon^2. \tag{8.7}
$$

Then, by Lemma 2, we can obtain the upper bound for the KL divergence as follows:

$$
\begin{aligned}
D(P_k\|P_0) &\le n\sum_{j=2}^{p}\frac{(\beta_j^{(k)}-\beta_j^{(0)})^2}{1-(\beta_j^{(0)})^2} = n\left[\sum_{j=2}^{d}\frac{(\beta_j^{(k)}-\beta_j^{(0)})^2}{1-(\beta_j^{(0)})^2} + \sum_{j=d+1}^{p}\frac{(\beta_j^{(k)}-\beta_j^{(0)})^2}{1-(\beta_j^{(0)})^2}\right] \\
&\le n\left[\sum_{j=2}^{d}\frac{(\beta_j^{(k)}-\beta_j^{(0)})^2}{1-(1-\epsilon^2/2)^2} + \sum_{j=d+1}^{p}(\beta_j^{(k)}-\beta_j^{(0)})^2\right] \\
&\le n\left[\frac{4dm^{-1}\epsilon^4}{\epsilon^2} + 2m^{-1}d\epsilon^2\right] = 6m^{-1}nd\epsilon^2,
\end{aligned} \tag{8.8}
$$

where the third inequality is due to (8.6) and (8.7).

By Lemmas 3 and 4, we have for any $k \ne k'$,

$$
\|\sin(\mathbf{A}_\epsilon(\mathbf{J}_k), \mathbf{A}_\epsilon(\mathbf{J}_{k'}))\|_2^2 \ge \epsilon^2(1-\epsilon^2)\|\mathbf{J}_k - \mathbf{J}_{k'}\|_2^2 \ge \frac{1}{4}\epsilon^2(1-\epsilon^2). \tag{8.9}
$$

By Chebyshev's inequality and Lemma 1, we have for all $\epsilon^2 \in [0, 1/2]$,

$$
\begin{aligned}
\max_k E_{P_k}\|\sin(\widehat{\mathbf{A}}, \mathbf{A}_\epsilon(\mathbf{J}_k))\|_2^2 &\ge \frac{\epsilon^2(1-\epsilon^2)}{16}\left[1 - \frac{6m^{-1}dn\epsilon^2 + \log 2}{mc\max\{d\varrho[1-\log\varrho], \log d\}}\right] \\
&\ge \frac{\epsilon^2(1-\epsilon^2)}{16}\left[1 - \frac{6dn\epsilon^2}{cm^2 d\varrho[1-\log\varrho]} - \frac{\log 2}{mc\log d}\right] \\
&\ge \frac{\epsilon^2}{32}\left[\frac{1}{2} - \frac{6dn\epsilon^2}{cm^2 d\varrho[1-\log\varrho]}\right],
\end{aligned}
$$

where the first inequality is due to (8.8) and (8.9). Take

$$
\epsilon^2 = \frac{cm^2}{24}\frac{\varrho d[1-\log\varrho]}{dn} = \frac{cm^2}{24}\frac{\varrho[1-\log\varrho]}{n}.
$$

Then

$$
\max_k E_{P_k}\|\sin(\widehat{\mathbf{A}}, \mathbf{A}_\epsilon(\mathbf{J}_k))\|_2^2 \ge \frac{1}{128}\epsilon^2. \tag{8.10}
$$

To ensure that $\boldsymbol{\rho}(\mathbf{A}_\epsilon(\mathbf{J}_k))$'s are in the sparse subspace, $\mathcal{F}_\delta(\pi(d))$, we need the following condition

$$
1 + \epsilon^\delta s^{(2-\delta)/2} \le \pi(d). \tag{8.11}
$$

Take

$$
\varrho = c_\varrho \pi(d) d^{-1}\left(\frac{\log d}{nd}\right)^{-\delta/2},
$$

where $c_\varrho = \frac{1}{\sqrt{2}}\left(\frac{cm^2}{24}\right)^{-\delta/2}$. Then, $\varrho \in (1/h, 1]$ and

$$\begin{aligned}\epsilon^2 &\leq \frac{c_\varrho cm^2}{24}\pi(d)d^{-1}n^{-1}\left(\frac{\log d}{nd}\right)^{-\delta/2}\left[1+\frac{1}{2}(1-\delta/2)\log d+\delta/2\log\log d\right]\\ &\leq c_\varrho\frac{cm^2}{24}\pi(d)\left(\frac{\log d}{nd}\right)^{1-\delta/2}\leq 1/2.\end{aligned}$$

Simple algebras show

$$\begin{aligned}\epsilon^{2\delta}s^{(2-\delta)} &\leq c_\varrho^2\left(\frac{cm^2}{24}\right)^{\delta}\left(\pi(d)\left(\frac{\log d}{nd}\right)^{1-\delta/2}\right)^{\delta}\left(\pi(d)\left(\frac{\log d}{nd}\right)^{-\delta/2}\right)^{2-\delta}\\ &= \frac{1}{2}\pi(d)^2\left(\frac{\log d}{nd}\right)^{-\delta}\left(\frac{\log d}{nd}\right)^{\delta}\\ &= \frac{1}{2}\pi(d)^2.\end{aligned}$$

Thus, (8.11) holds. Now, from (8.10), we have

$$\begin{aligned}\max_k E_{P_k}\|\sin(\widehat{\mathbf{A}},\mathbf{A}_\epsilon(\mathbf{J}_k))\|_2^2 &\geq C\pi(d)n^{-1}d^{-1}\left(\frac{\log d}{nd}\right)^{-\delta/2}\\ &\quad\times\left[1+\mathcal{N}\log d-\log\left(c_\varrho\pi(d)d^{\mathcal{N}-1}\left(\frac{\log d}{nd}\right)^{-\delta/2}\right)\right]\\ &\geq C\pi(d)n^{-1}d^{-1}\left(\frac{\log d}{nd}\right)^{-\delta/2}\log d\\ &= C\pi(d)\left(\frac{\log p}{nd}\right)^{1-\delta/2},\end{aligned}\tag{8.12}$$

where the second inequality is due to (4.3).

For the Frobenius norm, by Lemmas 3 and 4, we have for any $k \neq k'$,

$$\|\sin(\mathbf{A}_\epsilon(\mathbf{J}_k),\mathbf{A}_\epsilon(\mathbf{J}_{k'}))\|_F^2 \geq \epsilon^2(1-\epsilon^2)\|\mathbf{J}_k-\mathbf{J}_{k'}\|_F^2 \geq \frac{m}{4}\epsilon^2(1-\epsilon^2).$$

Then, similar to the proof of (8.12), we can show

$$\max_k E_{P_k}\|\sin(\widehat{\mathbf{A}},\mathbf{A}_\epsilon(\mathbf{J}_k))\|_F^2 \geq C\pi(d)\left(\frac{\log d}{nd}\right)^{1-\delta/2}.$$

■

Acknowledgments. The research of Tony Cai was supported in part by NSF Grant DMS-1403708 and NIH Grant R01 CA127334. The research of Yazhen Wang was supported in part by NSF Grants DMS-12-65203 and DMS-15-28375.

References

[1] Bickel, P. J., and Levina, E. (2008). Covariance regularization by thresholding, The Annals of Statistics, 36, 2577-2604.

[2] Birgé, L. (2001). A new look at an old result: Fano's lemma. Technical report, Univ. Paris 6.

[3] Birnbaum, A., Johnstone, I. M., Nadler, B., and Paul, D. (2013). Minimax bounds for sparse PCA with noisy high-dimensional data. The Annals of Statistics, 3, 1055-1084.

[4] Cai, T., Kim, D., Wang, Y., Yuan, M., and Zhou, H. (2016). Optimal density matrix estimation for large quantum systems based on Pauli measurements. The Annals of Statistics, 44, 682-712.

[5] Cai, T. T., Ma, Z., and Wu, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. The Annals of Statistics 41, 3074-3110.

[6] Cai, T. T., Ma, Z., and Wu, Y. (2015). Optimal estimation and rank detection for sparse spiked covariance matrices. Probability Theory and Related Fields 161, 781-815.

[7] Cai, T. T., Zhang, C.-H. and Zhou, H. (2010). Optimal rates of convergence for covariance matrix estimation. Annals of Statistics, 38, 2118-2144.

[8] Cai, T. T. and Zhou, H. (2012). Minimax estimation of large covariance matrices under ℓ_1-norm. Statistica Sinica, 22, 1319-1378.

[9] Golub, G.H. and Loan, van C.F. (1996). Matrix Computations. The Johns Hopkins University Press, Baltimore and London, 3rd edition.

[10] Johnstone, I. M. and Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. J. Amer. Statist. Assoc. 104 682-693.

[11] Kim, D.,and Wang, Y. (2017). Hypothesis Tests of Large Density Matrices of Quantum Systems Based on Pauli Measurements. Physica A, 469, 31-51.

[12] Kim, D., Wang, Y., and Zou, J. (2016). Asymptotic theory for large volatility matrix estimation based on high-frequency financial data. Stochastic Processes and Their Applications, 126, 3527-3577.

[13] Kim, D. and Wang, Y. (2016). Sparse PCA Based on High-Dimensional Itô processes with Measurement Errors. Journal of Multivariate Analysis, 152, 172-189.

[14] Koltchinskii, V., and Dong X. (2015) . Optimal estimation of low rank density matrices. Journal of Machine Learning Research, 16, 1757-1792.

[15] Li, R.-C. (1998a). Relative perturbation theory: I. Eigenvalue and singular value variations, SIAM J. Matrix Anal. Appl. 19, 956-982.

[16] Li, R.-C. (1998b). Relative perturbation theory: II. Eigenspace and singular subspace variations, SIAM J. Matrix Anal. Appl. 20, 471-492.

[17] Ma, Z. (2013). Sparse principal component analysis and iterative thresholding. The Annals of Statistics, 41, 772-801.

[18] Nielsen, M. and Chuang, I. (2000). Quantum Computation and Quantum Information. Cambridge: Cambridge University Press.

[19] Tao, M., Wang, Y., and Chen, X. (2013a). Fast convergence rates in estimating large volatility matrices using high-frequency financial data. Econometric Theory, 29, 1-19.

[20] Tao, M., Wang, Y., and Zhou, H. (2013b). Optimal sparse volatility matrix estimation for high-dimensional Itô process with measurement error. The Annals of Statistics, 41, 1816-1864.

[21] Tropp, J. A. (2012). User-friendly tail bounds for sums of random matrices. Foundations of Computational Mathematics, 12(4), 389-434.

[22] Vu, V. Q., Cho, J., Lei, J., and Rohe, K. (2013). Fantope Projection and Selection: A near-optimal convex relaxation of Sparse PCA. Advances in Neural Information Processing Systems (NIPS) 26.

[23] Vu, V. Q. and Lei, J. (2013). Minimax sparse principal subspace estimation in high dimensions. The Annals of Statistics, 41, 2905-2947.

[24] Wang, Y., and Zou, J. (2010). Vast volatility matrix estimation for high-frequency financial data. The Annals of Statistics, 38, 943-978.

[25] Wang, Y. (2011). Quantum Monte Carlo simulation. Annals of Applied Statistics, 5, 669-683.

[26] Wang, Y. (2012). Quantum computation and quantum information. Statistical Science, 27, 373-394.

[27] Wang, Y. (2013). Asymptotic equivalence of quantum state tomography and noisy matrix completion. The Annals of Statistics, 41, 2462-2504.

[28] Yang, Y. and Barron, A. (1999). Information-theoretic determination of minimax rates of convergence. The Annals of Statistics, 27 1564-1599.

A Appendix

A.1 Proofs of Propositions 1-4

A.1.1 Proof of Proposition 1

Proof of Proposition 1. By Davis-Kahn's $\sin\theta$ theorem (Theorem 3.1 in Li (1998b)), we have

$$\|\sin(\mathbf{Q}, \mathbf{Q}^o)\|_F^2 \leq \frac{\|(\boldsymbol{\rho}^o - \boldsymbol{\rho})\mathbf{Q}\|_F^2}{(\lambda_m - \ell_{m+1}^o)^2}, \tag{A.1}$$

where ℓ_j^o is the j-th largest eigenvalues for $\boldsymbol{\rho}^o$. First, consider the denominator of (A.1). By Weyl's theorem (Theorem 4.3 in Li (1998a)), we have

$$|\ell_j^o - \lambda_j| \leq \|\boldsymbol{\rho}^o - \boldsymbol{\rho}\|_2 = \left\| \begin{pmatrix} 0 & -\boldsymbol{\rho}_{HL} \\ -\boldsymbol{\rho}_{LH} & -\boldsymbol{\rho}_{LL} \end{pmatrix} \right\|_2 \leq 2\|\boldsymbol{\rho}_{HL}\|_2 + \|\boldsymbol{\rho}_{LL}\|_2.$$

Simple algebras show

$$\|\boldsymbol{\rho}_{HL}\|_2 \leq \lambda_1 \sum_{j=1}^{r} \|\mathbf{q}_{jL}\|_2 \leq C\pi(d)^{1/2}\tau_n^{1-\delta/2} = o(1), \tag{A.2}$$

where the second inequality can be shown similar to the proof of Lemma A.1 (Ma (2013)). Similarly,

$$\|\boldsymbol{\rho}_{LL}\|_2 \leq C\pi(d)^{1/2}\tau_n^{1-\delta/2} = o(1).$$

Thus, we have

$$|\ell_j^o - \lambda_j| = o(1). \tag{A.3}$$

and

$$(\lambda_m - \ell_{m+1}^o)^2 = (\lambda_m - \lambda_{m+1})^2 + o(1). \tag{A.4}$$

Consider the numerator of (A.1). Simple algebra shows

$$(\boldsymbol{\rho}^o - \boldsymbol{\rho})\mathbf{Q} = \begin{pmatrix} -\boldsymbol{\rho}_{HL}\mathbf{Q}_{L\cdot} \\ -\boldsymbol{\rho}_{LH}\mathbf{Q}_{H\cdot} - \boldsymbol{\rho}_{LL}\mathbf{Q}_{L\cdot} \end{pmatrix}.$$

Then,

$$\begin{aligned} \|(\boldsymbol{\rho}^o - \boldsymbol{\rho})\mathbf{Q}\|_F &\leq \|\boldsymbol{\rho}_{HL}\mathbf{Q}_{L\cdot}\|_F + \|\boldsymbol{\rho}_{LH}\mathbf{Q}_{H\cdot}\|_F + \|\boldsymbol{\rho}_{LL}\mathbf{Q}_{L\cdot}\|_F \\ &= (I) + (II) + (III). \end{aligned}$$

Because of similarity, we provide arguments only for (II). Let

$$\Lambda_0 = \text{diag}(\lambda_1, \ldots, \lambda_m),\ \Lambda_1 = \text{diag}(\lambda_{m+1}, \ldots, \lambda_r), \text{ and } \mathbf{Q}_1 = (\mathbf{q}_{m+1}, \ldots, \mathbf{q}_r).$$

We have

$$\begin{aligned}(II) &= \|\mathbf{Q}_{L\cdot}\Lambda_0\mathbf{Q}_{H\cdot}^{\dagger}\mathbf{Q}_{H\cdot} + \mathbf{Q}_{1,L\cdot}\Lambda_1\mathbf{Q}_{1,H\cdot}^{\dagger}\mathbf{Q}_{H\cdot}\|_F \\ &\leq \|\mathbf{Q}_{L\cdot}\|_F\|\Lambda_0\|_F\|\mathbf{Q}_{H\cdot}^{\dagger}\mathbf{Q}_{H\cdot}\|_F + \|\mathbf{Q}_{1,L\cdot}\|_F\|\Lambda_1\|_F\|\mathbf{Q}_{1,H\cdot}^{\dagger}\mathbf{Q}_{H\cdot}\|_F \\ &\leq C\pi(d)^{1/2}\tau_n^{1-\delta/2}, \end{aligned} \tag{A.5}$$

where the last inequality can be shown similar to the proof of (A.2).

Now, from (A.4) and (A.5), we have

$$\|\sin(\mathbf{Q}, \mathbf{Q}^o)\|_F^2 \leq C\frac{\pi(d)\tau_n^{2-\delta}}{(\lambda_m - \lambda_{m+1})^2}(1 + o(1)).$$

■

A.1.2 Proof of Proposition 2

Proof of Proposition 2. By Davis-Kahn's $\sin\theta$ theorem (Theorem 3.1 in Li (1998b)), we have

$$\|\sin(\mathbf{Q}^o, \widehat{\mathbf{Q}}^o)\|_F^2 \leq \frac{\|(\boldsymbol{\rho}^o - \widehat{\boldsymbol{\rho}}^o)\mathbf{Q}^o\|_F^2}{(\ell_m^o - \widehat{\ell}_{m+1}^o)^2}, \tag{A.6}$$

where ℓ_j^o and $\widehat{\ell}_j^o$ are the j-th largest eigenvalues for $\boldsymbol{\rho}^o$ and $\widehat{\boldsymbol{\rho}}^o$, respectively. First, we show that with probability at least $1 - C_0(d \vee n)^{-2}$,

$$|\widehat{\ell}_j^o - \ell_j^o| = o(1). \tag{A.7}$$

By Weyl's theorem (Theorem 4.3 in Li (1998a)), we have

$$|\widehat{\ell}_j^o - \ell_j^o| \leq \|\widehat{\boldsymbol{\rho}}^o - \boldsymbol{\rho}^o\|_2 = \|\widehat{\boldsymbol{\rho}}_{HH} - \boldsymbol{\rho}_{HH}\|_2.$$

Let $\mathbf{B}_j = (B_j^{l,h})_{l,h=1,\ldots,d}$. We have with $x = Cn^{-1/2}d^{-1/2}|H|\sqrt{\log(d \vee n)}$ for some large C,

$$\begin{aligned} P\left(\|\widehat{\boldsymbol{\rho}}_{HH} - \boldsymbol{\rho}_{HH}\|_2 \geq x\right) &\leq P\left(\max_{h\in H}\sum_{l\in H} d^{-1}\left|\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)B_j^{l,h}\right| \geq x\right) \\ &\leq |H|^2\max_{h,l\in H} P\left(\left|\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)B_j^{l,h}\right| \geq dx/|H|\right) \\ &\leq 2|H|^2\max_{h,l\in H}\exp\left(-\frac{n^2d^2x^2/|H|^2}{2n\sum_{j=2}^{p}(1-\beta_j^2)|B_j^{l,h}| + \frac{4}{3}ndx/|H|}\right) \\ &\leq |H|^2\exp\left(-\frac{n^2d^2x^2/|H|^2}{2nd + \frac{4}{3}ndx/|H|}\right) \\ &\leq C_0(d \vee n)^{-2}, \end{aligned}$$

where third inequality is due to Bernstein's inequality, and the fourth inequality is established by the fact that by the construction of Pauli matrices, we have for any $h, l = 1, \ldots, d$,

$$\sum_{j=1}^{p} |B_j^{h,l}| = d. \tag{A.8}$$

Then we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|\widehat{\boldsymbol{\rho}}_{HH} - \boldsymbol{\rho}_{HH}\|_2^2 \leq C\pi(d)^2 \tau_n^{2-2\delta}. \tag{A.9}$$

Since $\pi(d)^2\tau_n^{2-2\delta} = o(1)$, (A.7) holds. Then, by (A.3), we have

$$|\ell_m^o - \widehat{\ell}_{m+1}^o| = |\lambda_m - \lambda_{m+1}| + o(1). \tag{A.10}$$

Consider the numerator in (A.6). Let $\mathbf{a} = (a_1, \ldots, a_{|H|}) \in \mathbb{C}^{|H|}$ such that $\|\mathbf{a}\|_2^2 = 1$. Let $\mathbf{B}_{jHH}\mathbf{a} = \mathbf{b}_j = \mathbf{b}_j^r + \sqrt{-1}\,\mathbf{b}_j^i$, where $\mathbf{b}_j^r = (b_{j1}^r, \ldots, b_{j|H|}^r)^T, \mathbf{b}_j^i = (b_{j1}^i, \ldots, b_{j|H|}^i)^T \in \mathbb{R}^{|H|}$. We have

$$\begin{aligned} p\|(\widehat{\boldsymbol{\rho}}_{HH} - \boldsymbol{\rho}_{HH})\mathbf{a}\|_2^2 &= \|\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)\mathbf{b}_j\|_2^2 = \sum_{k=1}^{|H|}\left|\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)b_{jk}^r + \sqrt{-1}\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)b_{jk}^i\right|^2 \\ &= \sum_{k=1}^{|H|}\left\{\left[\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)b_{jk}^r\right]^2 + \left[\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)b_{jk}^i\right]^2\right\} \\ &= I + II. \end{aligned}$$

Then we have with $x = Cn^{-1}d^{-1}|H|\log(d \vee n)$ for some large C,

$$\begin{aligned} P\left(d^{-2} \times I \geq x\right) &\leq |H| \max_{k\in H} P\left(\left|\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)b_{jk}^r\right| \geq dx^{1/2}/|H|^{1/2}\right) \\ &\leq 2|H| \max_{k\in H} \exp\left(-\frac{n^2d^2|H|^{-1}x}{2n\sum_{j=2}^{p}(1-\beta_j^2)(b_{jk}^r)^2 + \frac{4}{3}nd|H|^{-1/2}x^{1/2}}\right) \\ &\leq 2|H| \exp\left(-\frac{n^2d^2|H|^{-1}x}{2nd + \frac{4}{3}nd|H|^{-1/2}x^{1/2}}\right) \\ &\leq C_0(d \vee n)^{-2}, \end{aligned}$$

where the second inequality is due to Bernstein's inequality, and the fourth inequality is established by the fact that since each row of $\mathbf{B}_j$ has only one non-zero element (± 1 or $\pm\sqrt{-1}$), we have

$$\begin{aligned} \sum_{j=2}^{p}(b_{jk}^r)^2 &\leq \sum_{j=2}^{p}\left|\sum_{l=1}^{|H|} B_{jHH}^{k,l} a_l\right|^2 \leq \sum_{j=2}^{p}\sum_{l=1}^{|H|}|B_{jHH}^{k,l}||a_l|^2 \\ &\leq \sum_{l=1}^{|H|}|a_l|^2 \sum_{j=2}^{p}|B_{jHH}^{k,l}| \leq d\|\mathbf{a}\|_2^2 = d, \end{aligned}$$

where the last inequality is due to (A.8). Thus, we have with probability at least $1-C_0(d \vee n)^{-2}$,

$$d^{-2} \times I \leq C\pi(d)\tau_n^{2-\delta}.$$

Similarly, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$d^{-2} \times II \leq C\pi(d)\tau_n^{2-\delta}.$$

Thus, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|(\widehat{\boldsymbol{\rho}}_{HH} - \boldsymbol{\rho}_{HH})\mathbf{a}\|_2^2 \leq C\pi(d)\tau_n^{2-\delta} \tag{A.11}$$

for any $\mathbf{a} = (a_1, \ldots, a_{|H|}) \in \mathbb{C}^{|H|}$ such that $\|\mathbf{a}\|_2^2 = 1$. Since $\|\mathbf{q}_\nu^o\|_2^2 = 1$ for $\nu = 1, \ldots, m$, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|(\boldsymbol{\rho}^o - \widehat{\boldsymbol{\rho}}^o)\mathbf{Q}^o\|_F^2 = \sum_{\nu=1}^{m} \|(\boldsymbol{\rho}^o - \widehat{\boldsymbol{\rho}}^o)\mathbf{q}_\nu^o\|_2^2 \leq C\pi(d)\tau_n^{2-\delta}, \tag{A.12}$$

where the last inequality is due to (A.11). From (A.10) and (A.12), we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|\sin(\mathbf{Q}^o, \widehat{\mathbf{Q}}^o)\|_F^2 \leq C\frac{\pi(d)\tau_n^{2-\delta}}{(\lambda_m - \lambda_{m+1})^2}.$$

■

A.1.3 Proof of Proposition 3

Lemma 5 *Under assumptions of Theorem 2, for appropriately chosen C_α and $a_\mp$ ($0 < a_- < 1 < a_+$), when n is sufficiently large, we have with probability at least $1-C_0(d\vee n)^{-2}$,*

$$S_- \subset S \subset S_+ \subset H,$$

where

$$S_\pm = \left\{\nu : \sum_{j=1}^{r} \lambda_j q_{\nu j}^2 > a_\mp \times C_\alpha \tau_n\right\}.$$

Proof: We have

$$\begin{aligned} P\left(S_- \not\subseteq S\right) &\leq P\left(\bigcup_{\nu\in S_-} \{\widehat{\rho}_{\nu\nu} < C_\alpha\tau_n\}\right) \leq \sum_{\nu\in S_-} P\left(\widehat{\rho}_{\nu\nu} < C_\alpha\tau_n\right) \\ &\leq \sum_{\nu\in S_-} P\left(\widehat{\rho}_{\nu\nu} - \rho_{\nu\nu} < -(\alpha_+ - 1)C_\alpha\tau_n\right) \\ &\leq \sum_{\nu\in S_-} P\left(|\widehat{\rho}_{\nu\nu} - \rho_{\nu\nu}| > (\alpha_+ - 1)C_\alpha\tau_n\right) \end{aligned}$$

$$\leq \quad C_0(d \vee n)^{-2},$$

where the last inequality is established by the fact that we have for large C,

$$\begin{aligned} P\left(|\widehat{\rho}_{\nu\nu} - \rho_{\nu\nu}| > C\tau_n\right) &= P\left(\left|\sum_{j=2}^{p}(\widehat{\beta}_j - \beta_j)B_j^{\nu,\nu}\right| > Cp^{1/2}\tau_n\right) \\ &\leq 2\exp\left(-\frac{0.5C^2n^2p\tau_n^2}{n\sum_{j=2}^{p}\frac{1-\beta_j^2}{2}(B_j^{\nu,\nu})^2 + \frac{2}{3}Cnp^{1/2}\tau_n}\right) \\ &\leq 2\exp\left(-\frac{0.5C^2n^2p\tau_n^2}{\frac{1}{2}np^{1/2} + \frac{2}{3}Cnp^{1/2}\tau_n}\right) \\ &\leq C_0(d \vee n)^{-3}, \end{aligned}$$

where the second and third inequalities are due to Bernstein's inequality and (A.8), respectively. Similarly, we can show

$$P\{S \not\subseteq S_+\} \leq C_0(d \vee n)^{-2}.$$

Now, consider $S_+ \subset H$. For any $\nu \in S_+$, by the definition of S_+, there exists $j \in 1, \ldots, r$ such that $\lambda_j q_{\nu j}^2 \geq a_- \times C_\alpha \tau_n / r$. Since $\tau_n \to 0$, for sufficiently large n, $\sqrt{\frac{a_- \times C_\alpha \tau_n}{r\lambda_j}} > C_\tau \tau_n$. Thus, $\nu \in H$. ■

Let

$$\varrho = \widehat{\ell}^o_{m+1} / \widehat{\ell}^o_m,$$

where $\widehat{\ell}^o_j$ is the j-th largest eigenvalue of $\widehat{\boldsymbol{\rho}}^o$.

Lemma 6 *Under assumptions of Theorem 2, uniformly over $\mathcal{F}_\delta(\pi(d))$, with probability at least $1 - C_0(d \vee n)^{-2}$:*

(1) $S^o = S$;

(2) $|\ell_j(\widehat{\boldsymbol{\rho}}_{S^oS^o}) - \widehat{\ell}^o_j| = o(1)$ as $n \to \infty$, for $j = 1, \ldots, r+1$, where $\ell_j(\widehat{\boldsymbol{\rho}}_{S^oS^o})$ is j-th largest eigenvalue for $\widehat{\boldsymbol{\rho}}_{S^oS^o}$;

(3) for sufficiently large n, $\widehat{\mathbf{Q}}^{(0),o}$ has full column rank, and $\|\sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(0),o})\|_F^2 \leq (1 - \varrho)^2/5$;

(4) for sufficiently large n, $R_s \in [R, 2R]$.

Proof: Without loss of generality, we prove the statements on the event such that the conclusions of Lemma 5.

Claim (1). The statement follows from Lemma 5.

Claim (2). Define

$$\widehat{\boldsymbol{\rho}}_0^o = \begin{bmatrix} \widehat{\boldsymbol{\rho}}_{0,HH}^o & 0 \\ 0 & 0 \end{bmatrix} \quad \text{with} \quad \widehat{\boldsymbol{\rho}}_{0,HH}^o = \begin{bmatrix} \widehat{\boldsymbol{\rho}}_{S^oS^o} & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$\boldsymbol{\rho}_0^o = \begin{bmatrix} \boldsymbol{\rho}_{0,HH}^o & 0 \\ 0 & 0 \end{bmatrix} \quad \text{with} \quad \boldsymbol{\rho}_{0,HH}^o = \begin{bmatrix} \boldsymbol{\rho}_{S^oS^o} & 0 \\ 0 & 0 \end{bmatrix},$$

and $D = H \setminus S^o$. By Weyl's theorem (Li (1998a), Theorem 4.3), we have with probability at least $1 - C_0(d \vee n)^{-2}$

$$\begin{aligned} |\ell_j(\widehat{\boldsymbol{\rho}}_{S^oS^o}) - \widehat{\ell}_j^o| &\leq \|\widehat{\boldsymbol{\rho}}_{0,HH}^o - \widehat{\boldsymbol{\rho}}_{HH}^o\|_2 \\ &\leq \|\widehat{\boldsymbol{\rho}}_{0,HH}^o - \boldsymbol{\rho}_{0,HH}^o\|_2 + \|\widehat{\boldsymbol{\rho}}_{HH}^o - \boldsymbol{\rho}_{HH}^o\|_2 + \|\boldsymbol{\rho}_{0,HH}^o - \boldsymbol{\rho}_{HH}^o\|_2 \\ &\leq C\pi(d)\tau_n^{1-\delta} + \|\boldsymbol{\rho}_{0,HH}^o - \boldsymbol{\rho}_{HH}^o\|_2, \end{aligned} \tag{A.13}$$

where the last inequality can be derived similar to the proof of (A.9).

Now, we examine $\|\boldsymbol{\rho}_{0,HH}^o - \boldsymbol{\rho}_{HH}^o\|_2$. Simple algebra manipulations show

$$\|\boldsymbol{\rho}_{0,HH}^o - \boldsymbol{\rho}_{HH}^o\|_2 \leq 2\|\boldsymbol{\rho}_{DS^o}\|_2 + \|\boldsymbol{\rho}_{DD}\|_2.$$

For $\|\boldsymbol{\rho}_{DS^o}\|_2$, since $S \subset S_+ \subset \{\nu \in \{1, \ldots, d\}, |q_{\nu j}| > \left(\frac{a_- C_\alpha}{\lambda_r}\right)^{1/2} \tau_n^{1/2} \text{ for some } 1 \leq j \leq r\}$, similar to the proof of Lemma A.1 (Ma (2013)), we can show for $j = 1, \ldots, r$,

$$\|\mathbf{q}_{S^cj}\|_2^2 \leq C\pi(d)\tau_n^{1-\delta/2}.$$

Thus, we have

$$\|\boldsymbol{\rho}_{DS^o}\|_2 = \|\bar{\mathbf{Q}}_{D\cdot}\Lambda\bar{\mathbf{Q}}_{S^o\cdot}^\dagger\|_2 \leq C\|\bar{\mathbf{Q}}_{D\cdot}\|_F \leq C\pi(d)^{1/2}\tau_n^{1/2-\delta/4}, \tag{A.14}$$

where $\bar{\mathbf{Q}} = (\mathbf{Q}, \mathbf{Q}_1)$. Similarly, we can show

$$\|\boldsymbol{\rho}_{DD}\|_2 = \|\bar{\mathbf{Q}}_{D\cdot}\Lambda\bar{\mathbf{Q}}_{D\cdot}^\dagger\|_2 \leq C\pi(d)^{1/2}\tau_n^{1/2-\delta/4}. \tag{A.15}$$

Combining (A.13)-(A.15), the statement holds.

Claim (3). By Davis-Kahn's $\sin\theta$ theorem (Theorem 3.1 in Li (1998b)), we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\begin{aligned} \|\sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(0),o})\|_F^2 &\leq \frac{\|(\widehat{\boldsymbol{\rho}}^o - \widehat{\boldsymbol{\rho}}_0^o)\widehat{\mathbf{Q}}^o\|_F^2}{(\ell_j(\widehat{\boldsymbol{\rho}}_{S^oS^o}) - \widehat{\ell}_j^o)^2} \leq C\left\|(\widehat{\boldsymbol{\rho}}^o - \widehat{\boldsymbol{\rho}}_0^o)\widehat{\mathbf{Q}}^o\right\|_F^2 \\ &\leq C\bigg(\|(\widehat{\boldsymbol{\rho}}^o - \boldsymbol{\rho}^o)\mathbf{Q}^o\|_F^2 + \|(\widehat{\boldsymbol{\rho}}_0^o - \boldsymbol{\rho}_0^o)\mathbf{Q}^o\|_F^2 + \|(\boldsymbol{\rho}^o - \boldsymbol{\rho}_0^o)\mathbf{Q}^o\|_F^2 \\ &\quad + \|\widehat{\boldsymbol{\rho}}^o - \widehat{\boldsymbol{\rho}}_0^o\|_F^2\|\widehat{\mathbf{Q}}^o - \mathbf{Q}^o\|_F^2\bigg) \end{aligned}$$

$$= \quad (I) + (II) + (III) + (IV),$$

where the second inequality is due to claim (2), (A.4), and (A.10). For (I) and (II), similar to the proof of (A.12), we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$(I) + (II) \leq C\pi(d)\tau_n^{2-\delta}. \tag{A.16}$$

For (III), similar to the proofs of (A.14) and (A.15), we have

$$(III) \leq C\pi(d)\tau_n^{1-\delta/2}. \tag{A.17}$$

Finally, consider (IV). Simple algebras show

$$\|\widehat{\mathbf{Q}}^o - \mathbf{Q}^o\|_F^2 \leq 2r. \tag{A.18}$$

We have

$$\|\widehat{\boldsymbol{\rho}}^o - \widehat{\boldsymbol{\rho}}_0^o\|_F^2 \leq C(\|\widehat{\boldsymbol{\rho}}^o - \boldsymbol{\rho}^o\|_F^2 + \|\widehat{\boldsymbol{\rho}}_0^o - \boldsymbol{\rho}_0^o\|_F^2 + \|\boldsymbol{\rho}^o - \boldsymbol{\rho}_0^o\|_F^2).$$

Then, we have with $x = Cn^{-1}d^{-1}|H|^2\log(d \vee n)$ for large C,

$$\begin{aligned}
P\left(\|\widehat{\boldsymbol{\rho}}^o - \boldsymbol{\rho}^o\|_F^2 \geq x\right) &\leq P\left(\sum_{h\in H}\sum_{l\in H} d^{-2}\left|\sum_{j=2}^p(\widehat{\beta}_j - \beta_j)B_j^{l,h}\right|^2 \geq x\right) \\
&\leq |H|^2 \max_{h,l\in H} P\left(\left|\sum_{j=2}^p(\widehat{\beta}_j - \beta_j)B_j^{l,h}\right| \geq dx^{1/2}/|H|\right) \\
&\leq 2|H|^2 \max_{h,l\in H}\exp\left(-\frac{n^2d^2|H|^{-2}x}{2n\sum_{j=2}^p(1-\beta_j^2)|B_j^{l,h}| + \frac{4}{3}nd|H|^{-1}x^{1/2}}\right) \\
&\leq |H|^2\exp\left(-\frac{n^2d^2|H|^{-2}x}{2nd + \frac{4}{3}nd|H|^{-1}x^{1/2}}\right) \\
&\leq C_0(d \vee n)^{-2},
\end{aligned}$$

where third and fourth inequalities are due to Bernstein's inequality and (A.8), respectively. Thus, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|\widehat{\boldsymbol{\rho}}^o - \boldsymbol{\rho}^o\|_F^2 \leq C\pi(d)^2\tau_n^{2-2\delta}.$$

Similarly, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|\widehat{\boldsymbol{\rho}}_0^o - \boldsymbol{\rho}_0^o\|_F^2 \leq C\pi(d)^2\tau_n^{2-2\delta}.$$

Similar to the proofs of (A.14) and (A.15), we have

$$\|\boldsymbol{\rho}^o - \boldsymbol{\rho}_0^o\|_F^2 \leq C\pi(d)\tau_n^{1-\delta/2}.$$

Therefore, with probability at least $1 - C_0(d \vee n)^{-2}$

$$\|\widehat{\boldsymbol{\rho}}^o - \widehat{\boldsymbol{\rho}}_0^o\|_F^2 \leq C\pi(d)\tau_n^{1-\delta/2}. \tag{A.19}$$

Combining (A.18) and (A.19), we have with probability at least $1 - C_0(d \vee n)^{-2}$

$$(IV) \leq C\pi(d)\tau_n^{1-\delta/2}. \tag{A.20}$$

Now, from (A.16), (A.17), and (A.20), we have with probability at least $1-C_0(d\vee n)^{-2}$,

$$\|\sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(0),o})\|_F^2 \leq C\pi(d)\tau_n^{1-\delta/2} \leq \frac{1}{5}(1-\varrho)^2. \tag{A.21}$$

Claim (4). The statement can be shown by combining the continuous mapping theorem and the results from claims (1) and (2), (A.3), and (A.7). ■

Let $\theta^{(k)} \in [0, \pi/2]$ be the largest canonical angle between the subspaces $\text{ran}(\widehat{\mathbf{Q}}^o)$ and $\text{ran}(\widehat{\mathbf{Q}}^{(k),o})$. Further, let $\phi^{(k)}$ be the largest canonical angle between $\text{ran}(\mathbf{T}^{(k),o})$ and $\text{ran}(\widehat{\mathbf{Q}}^o)$. So, $\|\sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(k),o})\|_F^2 \leq r\sin^2\theta^{(k)}$ and $\|\sin(\mathbf{T}^{(k),o}, \widehat{\mathbf{Q}}^o)\|_F^2 \leq r\sin^2\phi^{(k)}$.

Lemma 7 *Under assumptions of Theorem 2, in the first iteration of oracle Algorithm 1, with probability at least* $1 - C_0(d \vee n)^{-2}$*,*

(1) after the multiplication step, $\mathbf{T}^{(1),o}$ *has full column rank and* $\sin\phi^{(1)} \leq \varrho\tan\theta^{(0)}$*;*

(2) after the thresholding step, $\widehat{\mathbf{Q}}^{(1),o}$ *has full column rank, and* $\|\sin(\mathbf{T}^{(1),o}, \widehat{\mathbf{Q}}^{(1),o})\|_F^2 \leq \omega^2\sec^2\theta^{(0)}$*, where* $\omega = (\widehat{\ell}_m^o)^{-1}\left[|H|\sum_{j=1}^m \gamma_{nj}^2\right]^{1/2}$.

Proof: Without loss of generality, we prove the statements on the event such that the conclusions of Propositions 1 and 2, and Lemma 6. Thus, all the arguments are deterministic.

Claim (1). The statement can be shown similar to the proof of Theorem 8.2.2 (Golub and Loan (1996)) with a Hermitian matrix.

Claim (2). By Wedin's $\sin\theta$ theorem for singular subspace (Theorem 3.3 in Li (1998b)), we have

$$\|\sin(\mathbf{T}^{(1),o}, \widehat{\mathbf{Q}}^{(1),o})\|_F^2 = \|\sin(\mathbf{T}^{(1),o}, \widehat{\mathbf{T}}^{(1),o})\|_F^2 \leq \frac{\|\widehat{\mathbf{T}}^{(1),o} - \mathbf{T}^{(1),o}\|_F^2}{\sigma_m^2(\mathbf{T}^{(1),o})},$$

where $\sigma_m(\mathbf{A})$ is the m-th largest singular value for d-by-m matrix, $\mathbf{A}$. Now, we bound the numerator and denominator on the right hand side.

First, we derive the lower bound for $\sigma_m^2(\mathbf{T}^{(1),o})$. For any unit vector $x \in \mathbb{C}^m$, let $y = \widehat{\mathbf{Q}}^{(0),o}x$, then, y is also a unit vector in $\text{ran}(\mathbf{Q}^{(0),o}) \subset \mathbb{C}^d$. Decompose y as

$$y = \widetilde{y} + \widetilde{y}_c$$

with $\widetilde{y} \in \mathrm{ran}(\widehat{\mathbf{Q}}^o)$ and $\widetilde{y}_c \in \mathrm{ran}(\widehat{\mathbf{Q}}^o_c)$, where $\widehat{\mathbf{Q}}^o$ and $\widehat{\mathbf{Q}}^o_c$ are orthogonal basis. Then we have

$$\|\mathbf{T}^{(1),o}x\|_2^2 = \|\widehat{\boldsymbol{\rho}}^o y\|_2^2 = \|\widehat{\boldsymbol{\rho}}^o \widetilde{y}\|_2^2 + \|\widehat{\boldsymbol{\rho}}^o \widetilde{y}_c\|_2^2 \geq \|\widehat{\boldsymbol{\rho}}^o \widetilde{y}\|_2^2 \geq (\widehat{l}^o_m)^2 \|\widetilde{y}\|_2^2 \geq (\widehat{\ell}^o_m)^2 \cos^2 \theta^{(0)}.$$

Thus, we have

$$\sigma_m^2(\mathbf{T}^{(1),o}) \geq \inf_{\|x\|_2=1} \|\mathbf{T}^{(1),o}x\|_2^2 \geq (\widehat{\ell}^o_m)^2 \cos^2 \theta^{(0)}. \tag{A.22}$$

To bound the numerator, define matrix $\Delta\mathbf{T} \in \mathbb{C}^{d\times m}$, whose (i,j)-th element is given by $(\Delta T)_{i,j} = \gamma_{nj} 1_{(i\in H)}$. Then, by the construction of $\widehat{\mathbf{T}}^{(1),o}$, we have

$$\|\widehat{\mathbf{T}}^{(1),o} - \mathbf{T}^{(1),o}\|_F \leq \|\Delta\mathbf{T}\|_F = \widehat{\ell}^o_m \omega. \tag{A.23}$$

By (A.22) and (A.23), we have

$$\|\sin(\mathbf{T}^{(1),o}, \widehat{\mathbf{Q}}^{(1),o})\|_F^2 \leq \omega^2 \sec^2 \theta^{(0)}.$$

Finally, we show that for sufficiently large n, $\widehat{\mathbf{T}}^{(1),o}$ has full column rank. Now, suppose that $\widehat{\mathbf{T}}^{(1),o}$ does not have full column rank. Then, since $\sigma_m(\widehat{\mathbf{T}}^{(1),o}) = 0$ and (A.22),

$$|\sigma_m(\widehat{\mathbf{T}}^{(1),o}) - \sigma_m(\mathbf{T}^{(1),o})| = |\sigma_m(\mathbf{T}^{(1),o})| \geq \widehat{\ell}^o_m \cos \theta^{(0)} \geq \frac{4}{5} \widehat{\ell}^o_m, \tag{A.24}$$

where the last inequality is due to Lemma 6 (3). On the other hand, by Theorem 4.7 (Li (1998a)), we have

$$|\sigma_m(\widehat{\mathbf{T}}^{(1),o}) - \sigma_m(\mathbf{T}^{(1),o})| \leq \|\widehat{\mathbf{T}}^{(1),o} - \mathbf{T}^{(1),o}\|_2 \leq \widehat{\ell}^o_m \omega = o(1),$$

where the last inequality is established by the fact that

$$\widehat{\ell}^o_m \omega \leq C \left[\pi(d) \tau_n^{2-\delta}\right]^{1/2} = o(1).$$

This result contradicts (A.24). Thus, $\widehat{\mathbf{T}}^{(1),o}$ has full column rank, which implies that $\widehat{\mathbf{Q}}^{(1),o}$ has full column rank. ■

Lemma 8 *Under assumptions of Theorem 2, with probability at least* $1 - C_0(d \vee n)^{-2}$, *uniformly over* $\mathcal{F}_\delta(\pi(d))$, *for all* $k \geq 1$*:*

(1) $\widehat{\mathbf{Q}}^{(k),o}$ *is orthonormal, and* $\theta^{(k)}$ *satisfies*

$$\sin \theta^{(k)} \leq \varrho \tan \theta^{(k-1)} + \omega \sec \theta^{(k-1)},$$

where $\omega = (\widehat{\ell}^o_m)^{-1} \left[|H| \sum_{j=1}^m \gamma_{n,j}^2\right]^{1/2}$*;*

(2) for any $a \in (0, 1/2]$, *if*

$$\sin^2 \theta^{(k-1)} \leq 1.01(1-a)^{-2}(1-\varrho)^{-2}\omega^2,$$

then so is $\sin^2 \theta^{(k)}$. *Otherwise,*

$$\sin^2 \theta^{(k)} / \sin^2 \theta^{(k-1)} \leq \left[1 - a(1-\varrho)\right]^2.$$

Proof: The statement can be shown similar to the proof of Proposition 6.1 (Ma (2013)). ■

Proof of Proposition 3. We prove the statement on the event such that Lemmas 8 and 6, (A.3), and (A.7) hold. If we have for some $a \in (0, 1/2]$,

$$[1 - a(1 - \varrho)]^{2k^*} \leq \frac{1}{nd}, \tag{A.25}$$

then, by Lemma 8, for sufficiently large n,

$$\| \sin(\widehat{\mathbf{Q}}^o, \widehat{\mathbf{Q}}^{(k),o})\|_F^2 \leq r \sin^2 \theta^{(k)} \leq C\pi(d)\tau_n^{2-\delta}, \quad \text{for all } k \geq k^*.$$

Now, it is enough to show that k^* in (A.25) is less than or equal to $R = \frac{\lambda_1}{\lambda_m - \lambda_{m+1}}(\log n + 0.5 \log(d \vee n))$ with probability at least $1 - C_0(d \vee n)^{-2}$. The sufficient condition to hold (A.25) is

$$\frac{\log n + 1/2 \log(d \vee n)}{2\,|\log [1 - a(1 - \varrho)]|} \leq k^*.$$

Since $|\log(1 - x)| \geq x$ for all $x \in (0, 1)$, the upper bound of the left hand side is

$$\frac{\log n + 1/2 \log(d \vee n)}{2a(1 - \varrho)}.$$

Then, if we choose

$$k^* = \frac{1}{1 - \varrho}(\log n + 0.5 \log(d \vee n)) = \frac{\lambda_m}{\lambda_m - \lambda_{m+1}}(\log n + 0.5 \log(d \vee n)) \leq R,$$

(A.25) holds with $a = 1/2$. ■

A.1.4 Proof of Proposition 4

Proof of Proposition 4. We prove the statement on the event which conclusions of Propositions 1-3 and Lemmas 5-8.

Let $\widehat{\mathbf{Q}}^{(0),o} = (\widehat{\mathbf{q}}_1^{(0),o}, \ldots, \widehat{\mathbf{q}}_m^{(0),o})$. Since $S = S^o$, $\widehat{\mathbf{Q}}^{(0)} = \widehat{\mathbf{Q}}^{(0),o}$. Then, for $\nu \in L$ and $h = 1, \ldots, m$, the (ν, h)-th element of $\mathbf{T}^{(1)}$ is $t_{\nu h}^{(1)} = \widehat{\boldsymbol{\rho}}_{\nu\cdot} \widehat{\mathbf{q}}_h^{(0),o} = \widehat{\boldsymbol{\rho}}_{\nu\cdot}^o \widehat{\mathbf{q}}_h^{(0),o}$. Simple algebra manipulations show

$$\begin{aligned} |t_{\nu h}^{(1)}| &\leq |(\widehat{\boldsymbol{\rho}}_{\nu\cdot}^o - \boldsymbol{\rho}_{\nu\cdot}^o)\mathbf{Q}^o(\mathbf{Q}^o)^\dagger \widehat{\mathbf{q}}_h^{(0),o}| + |(\widehat{\boldsymbol{\rho}}_{\nu\cdot}^o - \boldsymbol{\rho}_{\nu\cdot}^o)\mathbf{Q}_c^o(\mathbf{Q}_c^o)^\dagger \widehat{\mathbf{q}}_h^{(0),o}| + |\boldsymbol{\rho}_{\nu H} \widehat{\mathbf{q}}_{hH}^{(0),o}| \\ &\leq \|(\widehat{\boldsymbol{\rho}}_{\nu\cdot}^o - \boldsymbol{\rho}_{\nu\cdot}^o)\mathbf{Q}^o\|_2 \|(\mathbf{Q}^o)^\dagger \widehat{\mathbf{q}}_h^{(0),o}\|_2 + \|\widehat{\boldsymbol{\rho}}_{\nu\cdot}^o - \boldsymbol{\rho}_{\nu\cdot}^o\|_2 \|\mathbf{Q}_c^o(\mathbf{Q}_c^o)^\dagger \widehat{\mathbf{q}}_h^{(0),o}\|_2 + |\boldsymbol{\rho}_{\nu H} \widehat{\mathbf{q}}_{hH}^{(0),o}| \\ &= (I)_{\nu h} + (II)_{\nu h} + (III)_{\nu h}, \end{aligned}$$

where $\mathbf{Q}^o$ and $\mathbf{Q}_c^o$ are orthonormal basis for $\mathbb{R}^d$. For $(I)_{\nu h}$, we have with $x = Cn^{-1/2}d^{-1/2}\sqrt{\log(d \vee n)}$ for sufficiently large C,

$$P\left(\max_{\nu \in L} \|(\widehat{\boldsymbol{\rho}}_{\nu\cdot}^o - \boldsymbol{\rho}_{\nu\cdot}^o)\mathbf{Q}^o\|_2 \geq x\right)$$

$$
\begin{aligned}
&\le m|L| \max_{\nu\in L, h\le m} P\left(\left|p^{-1/2}\sum_{j=2}^{p}(\widehat{\beta}_j-\beta_j)\mathbf{B}_{j\nu H}\mathbf{q}^o_{hH}\right| \ge \sqrt{m}x\right)\\
&\le 2m|L| \max_{\nu\in L, h\le m} \exp\left(-\frac{mn^2d^2x^2}{2n\sum_{j=2}^{p}(1-\beta_j^2)|\mathbf{B}_{j\nu H}\mathbf{q}^o_{hH}|^2+\frac{4}{3}ndm^{1/2}x}\right)\\
&\le 2m|L|\exp\left(-\frac{mn^2d^2x^2}{2nd+\frac{4}{3}ndm^{1/2}x}\right)\\
&\le C_0(d\vee n)^{-2},
\end{aligned}
$$

where the second inequality is due to Bernstein's inequality, and the third inequality is established by the fact that $\sum_{j=2}^{p}|\mathbf{B}_{j\nu H}\mathbf{q}^o_{hH}|^2 \le d$. Then, since $\|\mathbf{Q}^o_c(\mathbf{Q}^o_c)^\dagger\widehat{\mathbf{q}}_h^{(0),o}\|_2 \le 1$, we have with probability at least $1-C_0(d\vee n)^{-2}$,

$$
\max_{\nu\in L, 1\le h\le m}(I)_{\nu h} \le C\tau_n. \tag{A.26}
$$

For $(II)_{\nu h}$, consider $\|\widehat{\boldsymbol{\rho}}_{\nu H}-\boldsymbol{\rho}_{\nu H}\|_2$. We have with $x = Cn^{-1}d^{-1}|H|\log(d\vee n)$ for sufficiently large C,

$$
\begin{aligned}
&P\left(\max_{\nu\in L}\|\widehat{\boldsymbol{\rho}}_{\nu H}-\boldsymbol{\rho}_{\nu H}\|_2^2 \ge x\right)\\
&= P\left(\max_{\nu\in L}\sum_{l\in H} d^{-2}\left|\sum_{j=2}^{p}(\widehat{\beta}_j-\beta_j)B_j^{\nu,l}\right|^2 \ge x\right)\\
&\le |L||H| \max_{\nu\in L, l\in H} P\left(\left|\sum_{j=2}^{p}(\widehat{\beta}_j-\beta_j)B_j^{\nu,l}\right| \ge dx^{1/2}/|H|^{1/2}\right)\\
&\le 2|L||H| \max_{\nu\in L, l\in H}\exp\left(-\frac{n^2d^2|H|^{-1}x}{2n\sum_{j=2}^{p}(1-\beta_j^2)|B_j^{\nu,l}|^2+\frac{4}{3}nd|H|^{-1/2}x^{1/2}}\right)\\
&\le 2|L||H|\exp\left(-\frac{n^2d^2|H|^{-1}x}{2nd+\frac{4}{3}nd|H|^{-1/2}x^{1/2}}\right)\\
&\le C_0(d\vee n)^{-2},
\end{aligned}
$$

where the second inequality is due to Bernstein's inequality. Then, with probability at least $1-C_0(d\vee n)^{-2}$,

$$
\max_{\nu\in L}\|\widehat{\boldsymbol{\rho}}_{\nu H}-\boldsymbol{\rho}_{\nu H}\|_2 \le C\pi(d)^{1/2}\tau_n^{1-\delta/2}. \tag{A.27}
$$

Consider $\|\mathbf{Q}^o_c(\mathbf{Q}^o_c)^\dagger\widehat{\mathbf{q}}_h^{(0),o}\|_2$. Simple algebra manipulations show with probability at least $1-C_0(d\vee n)^{-2}$,

$$
\begin{aligned}
\|\mathbf{Q}^o_c(\mathbf{Q}^o_c)^\dagger\widehat{\mathbf{q}}_h^{(0),o}\|_2^2 &= \mathrm{Tr}\left[\mathbf{Q}^o_c(\mathbf{Q}^o_c)^\dagger\widehat{\mathbf{q}}_h^{(0),o}(\widehat{\mathbf{q}}_h^{(0),o})^\dagger\right] \le \sum_{h=1}^{m}\mathrm{Tr}\left[\mathbf{Q}^o_c(\mathbf{Q}^o_c)^\dagger\widehat{\mathbf{q}}_h^{(0),o}(\widehat{\mathbf{q}}_h^{(0),o})^\dagger\right]\\
&\le \mathrm{Tr}\left[\mathbf{Q}^o_c(\mathbf{Q}^o_c)^\dagger\widehat{\mathbf{Q}}^{(0),o}(\widehat{\mathbf{Q}}^{(0),o})^\dagger\right] \le \|\sin(\mathbf{Q}^o,\mathbf{Q}^{(0),o})\|_F^2\\
&\le 2\left(\|\sin(\mathbf{Q}^o,\widehat{\mathbf{Q}}^o)\|_F^2+\|\sin(\widehat{\mathbf{Q}}^o,\mathbf{Q}^{(0),o})\|_F^2\right)
\end{aligned}
$$

$$\leq \quad C\pi(d)\tau_n^{1-\delta/2}, \tag{A.28}$$

where the last inequality is due to (A.19) and Proposition 2. By (A.27) and (A.28), we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\max_{\nu \in L, 1 \leq h \leq m} (II)_{\nu h} \leq C\tau_n \pi(d)\tau_n^{1/2-3\delta/4} \leq C\tau_n. \tag{A.29}$$

For $(III)_{\nu h}$, simple algebra manipulations show

$$\max_{\nu \in L, 1 \leq h \leq m} (III)_{\nu h} \leq \max_{\nu \in L, 1 \leq h \leq m} \|\boldsymbol{\rho}_{\nu H}\|_2 \leq \max_{\nu \in L, 1 \leq h \leq m} C\sqrt{\sum_{l=1}^{r} \lambda_l^2 |q_{\nu l}|^2} \leq C\tau_n. \tag{A.30}$$

From (A.26), (A.29), and (A.30), by choosing large C_γ, we obtain $\widehat{\mathbf{T}}_{L\cdot}^{(1)} = 0$, and so, $\widehat{\mathbf{T}}^{(1)} = \widehat{\mathbf{T}}^{(1),o}$ with probability at least $1 - C_0(d \vee n)^{-2}$, which implies that $\widehat{\mathbf{Q}}^{(1)} = \widehat{\mathbf{Q}}^{(1),o}$. Note that the above results does not depend on k. Thus, by the induction method, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\widehat{\mathbf{Q}}^{(k)} = \widehat{\mathbf{Q}}^{(k),o}, \quad k \geq 0.$$

■

A.2 Proofs of Corollaries 1 and 2

Proof of Corollary 1. Let E be the event of the consequence of Theorem 2. Then, simple algebras show for large n,

$$\begin{aligned}
E\left[\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2\right] &= E\left[\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2 \mathbf{1}_E\right] + E\left[\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_F^2 \mathbf{1}_{E^c}\right] \\
&\leq \frac{C_u}{(\lambda_m - \lambda_{m+1})^2}\pi(d)\left(\frac{\log(d \vee n)}{nd}\right)^{1-\delta/2} + rP(E^c) \\
&\leq \frac{C_u}{(\lambda_m - \lambda_{m+1})^2}\pi(d)\left(\frac{\log(d \vee n)}{nd}\right)^{1-\delta/2} + rC_0(d \vee n)^{-2} \\
&\leq \frac{2C_u}{(\lambda_m - \lambda_{m+1})^2}\pi(d)\left(\frac{\log(d \vee n)}{nd}\right)^{1-\delta/2},
\end{aligned}$$

where the second inequality is established by the fact that $\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}})\|_F^2 \leq r$ for any $\mathbf{Q}$ and $\widehat{\mathbf{Q}}^{(R_s)}$. Similarly, we have

$$E\left[\|\sin(\mathbf{Q}, \widehat{\mathbf{Q}}^{(R_s)})\|_2^2\right] \leq \frac{2C_u}{(\lambda_m - \lambda_{m+1})^2}\pi(d)\left(\frac{\log(d \vee n)}{nd}\right)^{1-\delta/2}.$$

■

Proof of Corollary 2. Denote by $\widehat{\mathbf{Q}}_k^{(R_S)}$ and $\mathbf{Q}_k$ the first k columns of $\widehat{\mathbf{Q}}^{(R_s)}$ and $\mathbf{Q}$, respectively. Similar to the proof of Theorem 2, we have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\|\sin(\mathbf{Q}_h, \widehat{\mathbf{Q}}_h^{(R_s)})\|_F^2 \le C_u \pi(d) \left(\frac{\log(d \vee n)}{nd} \right)^{1-\delta/2} \quad \text{for } \ h = k, k-1. \tag{A.31}$$

We have with probability at least $1 - C_0(d \vee n)^{-2}$,

$$\begin{aligned} \|\sin(\mathbf{q}_k, \widehat{\mathbf{q}}_k^{(R_s)})\|_F^2 &\le 2\|\sin(\mathbf{Q}_k, \widehat{\mathbf{Q}}_k^{(R_s)})\|_F^2 + 2\|\sin(\mathbf{Q}_{k-1}, \widehat{\mathbf{Q}}_{k-1}^{(R_s)})\|_F^2 \\ &\le C\pi(d) \left(\frac{\log p}{np^{1/2}} \right)^{1-\delta/2}, \end{aligned}$$

where the first inequality is established by the fact that $\mathbf{q}_k \mathbf{q}_k^\dagger = \mathbf{Q}_k \mathbf{Q}_k^\dagger - \mathbf{Q}_{k-1} \mathbf{Q}_{k-1}^\dagger$, and the last inequality is due to (A.31). Then, similar to the proof of Corollary 1, we have

$$\sup_{\boldsymbol{\rho} \in \mathcal{F}_n} E\left[\|\sin(\mathbf{q}_k, \widehat{\mathbf{q}}_k^{(R_s)})\|_F^2 \right] \le C\pi(d) \left[\frac{\log(d \vee n)}{nd} \right]^{1-\delta/2}.$$

Similarly, we have

$$\sup_{\boldsymbol{\rho} \in \mathcal{F}_n} E\left[\|\sin(\mathbf{q}_k, \widehat{\mathbf{q}}_k^{(R_s)})\|_2^2 \right] \le C\pi(d) \left[\frac{\log(d \vee n)}{nd} \right]^{1-\delta/2}.$$

■

A.3 Proof related to the lower bound

Proof of Lemma 2. Simple algebra manipulations show

$$\begin{aligned} & E_{P_1}\left[\log(P_1/P_2)\right] \\ &= \sum_{j=2}^{p} \left\{ E_{P_1}[X_j] \log\left(\frac{1+\beta_j^{(1)}}{1+\beta_j^{(2)}} \right) + E_{P_1}[n - X_j] \log\left(\frac{1-\beta_j^{(1)}}{1-\beta_j^{(2)}} \right) \right\} \\ &= \sum_{j=2}^{p} \left\{ n\frac{1+\beta_j^{(1)}}{2} \log\left(1 + \frac{\beta_j^{(1)} - \beta_j^{(2)}}{1+\beta_j^{(2)}} \right) + n\frac{1-\beta_j^{(1)}}{2} \log\left(1 + \frac{\beta_j^{(2)} - \beta_j^{(1)}}{1-\beta_j^{(2)}} \right) \right\} \\ &\le \sum_{j=2}^{p} \left\{ n\frac{1+\beta_j^{(1)}}{2} \times \frac{\beta_j^{(1)} - \beta_j^{(2)}}{1+\beta_j^{(2)}} + n\frac{1-\beta_j^{(1)}}{2} \times \frac{\beta_j^{(2)} - \beta_j^{(1)}}{1-\beta_j^{(2)}} \right\} \\ &= n \sum_{j=2}^{p} \frac{(\beta_j^{(1)} - \beta_j^{(2)})^2}{1 - (\beta_j^{(2)})^2}, \end{aligned}$$

where the first inequality is established by the fact that $\log(1+x) \le x$ for $x \in (-1, \infty)$. ■

A.4 Proofs of Theorems 4-5

Similar to the proof of Theorem 4, we can prove Theorem 5. So, we omit the proof of Theorem 5.

Proof of Theorem 4. Since (5.2) can be shown imediately from (5.1) and Corollary 1, we provide arguments only for (5.1). Define

$$\mathcal{G}_{\mathcal{O}} = \{\mathcal{O} = (\mathcal{O}_1, \dots, \mathcal{O}_r) \in \mathbb{V}_{r,r} : \mathcal{O}^T \mathbf{Q}^T \boldsymbol{\rho} \mathbf{Q} \mathcal{O} = \boldsymbol{\Lambda}\},$$

where $\mathbb{V}_{r,r}$ is the Stiefel manifold of r-by-r orthonormal matrices, and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_r)$. There is $\mathcal{O} \in \mathcal{G}_{\mathcal{O}}$ such that

$$\|\widehat{\mathbf{Q}}^{(R_s)} - \mathbf{Q}\mathcal{O}\|_F^2 \le 2\|\sin(\widehat{\mathbf{Q}}^{(R_s)}, \mathbf{Q})\|_F^2.$$

Let $\mathbf{Q}^{\mathcal{O}} = \mathbf{Q}\mathcal{O} = (\mathbf{q}_1^{\mathcal{O}}, \dots, \mathbf{q}_r^{\mathcal{O}})$. Simple algebraic manipulations show

$$\begin{aligned} (\widehat{\mathbf{q}}_\nu^{(R_s)})^\dagger \widehat{\boldsymbol{\rho}} \widehat{\mathbf{q}}_\nu^{(R_s)} - \lambda_\nu &= (\mathbf{q}_\nu^{\mathcal{O}})^\dagger (\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \mathbf{q}_\nu^{\mathcal{O}} + \left[(\widehat{\mathbf{q}}_\nu^{(R_s)})^\dagger \widehat{\boldsymbol{\rho}} \widehat{\mathbf{q}}_\nu^{(R_s)} - (\mathbf{q}_\nu^{\mathcal{O}})^\dagger \widehat{\boldsymbol{\rho}} \mathbf{q}_\nu^{\mathcal{O}} \right] \\ &= (I)_\nu + (II)_\nu. \end{aligned}$$

Consider $(I)_\nu$. We have

$$\begin{aligned} E\left[\sum_{\nu=1}^r |(\mathbf{q}_\nu^{\mathcal{O}})^\dagger (\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \mathbf{q}_\nu^{\mathcal{O}}|^2 \right] &\le E\left[\|(\mathbf{Q}^{\mathcal{O}})^\dagger (\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \mathbf{Q}^{\mathcal{O}}\|_F^2 \right] \\ &\le E\left[\|\mathbf{Q}^\dagger (\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \mathbf{Q}\|_F^2 \right] \\ &= \frac{1}{d^2} \sum_{\nu=1}^r \sum_{\nu'=1}^r \sum_{j=2}^p E\left[(\widehat{\beta}_j - \beta_j)^2 \right] |\mathbf{q}_{\nu'}^\dagger \mathbf{B}_j \mathbf{q}_\nu|^2 \\ &\le \frac{1}{d^2 n} \sum_{\nu=1}^r \sum_{\nu'=1}^r \sum_{j=2}^p |\mathbf{q}_{\nu'}^\dagger \mathbf{B}_j \mathbf{q}_\nu|^2 \\ &\le C \frac{r^2}{dn}, \end{aligned} \tag{A.32}$$

where the last inequality is due to (A.33) below. We have

$$\begin{aligned} \sum_{j=2}^p |\mathbf{q}_{\nu'}^\dagger \mathbf{B}_j \mathbf{q}_\nu|^2 &\le \sum_{j=2}^p \left(|\mathbf{q}_\nu^\dagger \mathbf{B}_j \mathbf{q}_\nu|^2 + |\mathbf{q}_{\nu'}^\dagger \mathbf{B}_j \mathbf{q}_{\nu'}|^2 + |(\mathbf{q}_{\nu'} + \mathbf{q}_\nu)^\dagger \mathbf{B}_j (\mathbf{q}_{\nu'} + \mathbf{q}_\nu)|^2 \right) \\ &\le Cd, \end{aligned} \tag{A.33}$$

where the last inequality is due to the fact that by Proposition 1 (Cai *et al.* (2016)), for any $\mathbf{a} \in \mathbb{C}^d$ and $\|\mathbf{a}\|_2 = 1$, we have

$$\begin{aligned} 1 = \|\mathbf{a}\mathbf{a}^\dagger\|_F^2 &= \frac{1}{d^2} \sum_{j=1}^p \mathrm{tr}(\mathbf{B}_j \mathbf{a}\mathbf{a}^\dagger)^2 \mathrm{tr}(\mathbf{B}_j^\dagger \mathbf{B}_j) \\ &= \frac{1}{d} \sum_{j=1}^p \mathrm{tr}(\mathbf{B}_j \mathbf{a}\mathbf{a}^\dagger)^2 = \frac{1}{d} \sum_{j=1}^p (\mathbf{a}^\dagger \mathbf{B}_j \mathbf{a})^2. \end{aligned}$$

For $(II)_\nu$, simple algebraic manipulations show

$$\begin{aligned}|(II)_\nu| &= |(\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}})^\dagger \widehat{\boldsymbol{\rho}}(\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}}) + 2(\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}})^\dagger \widehat{\boldsymbol{\rho}}\mathbf{q}_\nu^{\mathcal{O}}| \\ &\le \|\widehat{\boldsymbol{\rho}}\|_2 \|\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}}\|_2^2 + 2\|\widehat{\boldsymbol{\rho}}\mathbf{q}_\nu^{\mathcal{O}}\|_2 \|\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}}\|_2 \\ &\le (\|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_2 + 1)\|\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}}\|_2^2 + 2(\|(\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho})\mathbf{q}_\nu^{\mathcal{O}}\|_2 + 1)\|\widehat{\mathbf{q}}_\nu^{(R_s)} - \mathbf{q}_\nu^{\mathcal{O}}\|_2.\end{aligned}$$

By (8.2), we can show

$$E\left[\|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_2^2\right] \le C\frac{\log d}{n} = o(1). \tag{A.34}$$

Then we have

$$\begin{aligned}E\left[\sum_{\nu=1}^r |(II)|_\nu\right] &\le E\left[\|\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}\|_2^2\right]^{1/2} E\left[\|\widehat{\mathbf{Q}}^{(R_s)} - \mathbf{Q}^{\mathcal{O}}\|_F^4\right]^{1/2} + E\left[\|\widehat{\mathbf{Q}}^{(R_s)} - \mathbf{Q}^{\mathcal{O}}\|_F^2\right] \\ &\quad +2E\left[\|(\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho})\|_2^2\right]^{1/2} E\left[\|\widehat{\mathbf{Q}}^{(R_s)} - \mathbf{Q}^{\mathcal{O}}\|_F^2\right]^{1/2} + 2E\left[\|\widehat{\mathbf{Q}}^{(R_s)} - \mathbf{Q}^{\mathcal{O}}\|_2\right] \\ &\le C\left[\left(\frac{\log d}{n}\right)^{1/2} \pi(d)\tau_n^{2-\delta} + \pi(d)\tau_n^{2-\delta} + \left(\frac{\log d}{n}\right)^{1/2} \pi(d)^{1/2}\tau_n^{1-\delta/2}\right. \\ &\qquad \left. +\pi(d)^{1/2}\tau_n^{1-\delta/2}\right] \\ &\le C\pi(d)^{1/2}\tau_n^{1-\delta/2}, \end{aligned}\tag{A.35}$$

where the second inequality is due to (A.34) and Corollary 1. Collecting (A.32) and (A.35),

$$E\left[\sum_{\nu=1}^r |(\widehat{\mathbf{q}}_\nu^{(R_s)})^\dagger \widehat{\boldsymbol{\rho}} \widehat{\mathbf{q}}_\nu^{(R_s)} - \lambda_\nu|\right] \le C\pi(d)^{1/2}\tau_n^{1-\delta/2}.$$

From the above results, we obtain

$$E\left[\sum_{\nu=1}^r |\widetilde{\lambda}_\nu^{(R_s)} - \lambda_\nu|\right] \le C\pi(d)^{1/2}\tau_n^{1-\delta/2}.$$

Then, we have

$$\begin{aligned}E\left[\sum_{\nu=1}^r |\widehat{\lambda}_\nu^{(R_s)} - \lambda_\nu|\right] &= E\left[\left|\left(\frac{1}{\sum_{j=1}^r \widetilde{\lambda}_j^{(R_s)}} - 1\right)\widetilde{\lambda}_\nu^{(R_s)} + \widetilde{\lambda}_\nu^{(R_s)} - \lambda_\nu\right|\right] \\ &\le C\sum_{\nu=1}^r E\left[\left|\widetilde{\lambda}_\nu^{(R_s)} - \lambda_\nu\right|\right] \\ &\le C\pi(d)^{1/2}\tau_n^{1-\delta/2}.\end{aligned}$$

■