Journal of Multivariate Analysis 115 (2013) 317–333

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Consistency of sparse PCA in High Dimension, Low Sample Size contexts

Dan Shen *, Haipeng Shen, J.S. Marron

*Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States*

ARTICLE INFO

*Article history:*
Received 23 February 2012
Available online 7 November 2012

*AMS subject classifications:*
primary 62H25
secondary 62F12

*Keywords:*
Sparse PCA
High dimension
Low sample size
Consistency

ABSTRACT

Sparse Principal Component Analysis (PCA) methods are efficient tools to reduce the dimension (or number of variables) of complex data. Sparse principal components (PCs) are easier to interpret than conventional PCs, because most loadings are zero. We study the asymptotic properties of these sparse PC directions for scenarios with fixed sample size and increasing dimension (i.e. High Dimension, Low Sample Size (HDLSS)). We consider the previously studied single spike covariance model and assume in addition that the maximal eigenvector is sparse. We extend the existing HDLSS asymptotic consistency and strong inconsistency results of conventional PCA in an entirely new direction. We find a large set of sparsity assumptions under which sparse PCA is still consistent even when conventional PCA is strongly inconsistent. The consistency of sparse PCA is characterized along with rates of convergence. Furthermore, we clearly identify the mathematical boundaries of the sparse PCA consistency, by showing strong inconsistency for an oracle version of sparse PCA beyond the consistent region, as well as its inconsistency on the boundaries of the consistent region. Simulation studies are performed to validate the asymptotic results in finite samples.



## 1. Introduction

Principal Component Analysis (PCA) is an important visualization and dimension reduction tool for High Dimension, Low Sample Size (HDLSS) data. However, the linear combinations found by PCA typically involve all the variables, with non-zero loadings, which can be challenging to interpret. To overcome this weakness, we will study sparse PCA methods that generate sparse principal components (PCs), i.e. PCs with only a few non-zero loadings. Several sparse PCA methods have been proposed to facilitate the interpretation of HDLSS data, see for example Zou et al. [42], d'Aspremont et al. [12], Shen and Huang [35], Leng and Wang [26], Witten et al. [40], Johnstone and Lu [19], Ma [27], and Lee et al. [23].

Sparse PCA is primarily motivated by modern data sets of very high dimension; hence we prefer the statistical viewpoint of the HDLSS asymptotics. Such asymptotics are based on the limit as the dimension $d \to \infty$ with the sample size $n$ being fixed, as originally studied by Casella and Hwang [11] in the context of James–Stein estimation, and more recently by Hall et al. [17], Ahn et al. [3], Jung and Marron [20], Yata and Aoshima [41], Ahn et al. [2], and Jung et al. [21] in various multivariate analysis contexts. Conventional PCA was first studied using HDLSS asymptotics by [3], and much more thoroughly analyzed by [20,21], with more recent development in [41]. The present paper is the first that studies the HDLSS asymptotic properties of *Sparse* PCA, and takes the statistical lessons learned from this type of asymptotics in a timely and orthogonal direction from that of [20,21].

One main contribution of this paper is a *clear and complete characterization* of HDLSS asymptotic conditions about sparse PCA consistency, inconsistency, as well as strong inconsistency. First of all, we identify in Sections 2 and 3 sparsity conditions where conventional PCA is strongly inconsistent (for scenarios with relatively small population eigenvalues as carefully

* Corresponding author.
*E-mail addresses:* shen.unc@gmail.com, dshen@live.unc.edu (D. Shen), haipeng@email.unc.edu (H. Shen), marron@email.unc.edu (J.S. Marron).

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.10.007

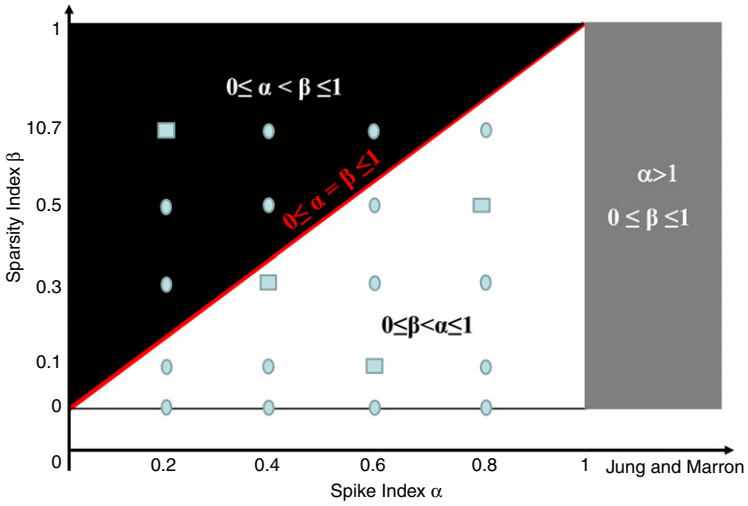


**Fig. 1.** Consistent areas for PCA and sparse PCA, as a function of the spike index $\alpha$ and the sparsity index $\beta$, under the single component spiked model considered in Example 1.1. Conventional PCA is consistent only on the grey rectangle ($\alpha > 1, 0 \leq \beta \leq 1$), while sparse PCA is also consistent on the white triangle ($0 \leq \beta < \alpha \leq 1$). In addition, an oracle sparse PCA procedure is strongly inconsistent on the black triangle ($0 \leq \alpha < \beta \leq 1$), and marginally inconsistent on the red diagonal line ($0 \leq \alpha = \beta \leq 1$). The squares and dots indicate the 20 grid points studied in our simulation studies, where the four squares highlight the four representative scenarios that are discussed in detail in Section 5.

studied in [20,21], and also noted by Lee et al. [25]), yet sparse PCA methods are consistent. Furthermore, the mathematical boundaries of the sparse PCA consistency are clearly established in Section 4, through showing that an oracle version of sparse PCA is marginally inconsistent on the boundaries, and strongly inconsistent beyond the consistent region. The formulation of *strong inconsistency* has not yet been studied in the random matrix literature [19,4]. Following the random matrix work, we also focus on the single component spiked covariance model. Our results depend on a *spike index*, $\alpha$, which measures the dominance of the first eigenvalue, and on a *sparsity index*, $\beta$, which measures the number of non-zero entries of the maximal eigenvector. The two indices $\alpha$ and $\beta$ are formally defined later in Example 1.1 of Section 1.1.

Our results offer major new insights relative to those of [20,21] (when restricting to the maximal eigenvector), who studied HDLSS asymptotic properties of conventional PCA without considering sparsity. Our paper clearly characterizes the benefit of imposing sparsity constraints when the maximal eigenvector is sparse, by revealing a new domain of consistency within the inconsistent region of conventional PCA. In addition, actual rates of convergence are obtained, which are not studied by [20,21]. Our key theoretical findings in connection with the earlier results are intuitively illustrated below for the exemplary model of Example 1.1. (Our results remain valid for more general single component spike models as shown in Sections 2–4.) The consistency and inconsistency results for Example 1.1 are summarized below as functions of $\alpha$ and $\beta$, and illustrated graphically in Fig. 1:

- *Previous results (grey rectangle)*: [20] showed that the first empirical PCA eigenvector is consistent with the maximal eigenvector when the spike index $\alpha$ is greater than 1, only involving the horizontal axis.
- *Consistency (white triangle)*: We will show that sparse PCA is consistent even when $\alpha$ is less than or equal to 1 (hence PCA is either strongly inconsistent [20] or marginally inconsistent [21]), as long as $\alpha$ is greater than the sparsity index $\beta$, involving both horizontal and vertical axes. This is done in Section 2 for a simple thresholding method and in Section 3 for the RSPCA method of [35].
- *Strong inconsistency (black triangle) and Marginal Inconsistency (red diagonal line)*: In Section 4 we show that even an oracle sparse PCA procedure is strongly inconsistent, when $\alpha$ is smaller than $\beta$, and marginally inconsistent when $\alpha = \beta$, the boundary between the consistent and strongly inconsistent regions.

Besides the HDLSS asymptotics, other asymptotic frameworks have been used to study sparse PCA. Under the random matrix framework where both $n$ and $d$ tend to infinity, Johnstone and Lu [19] and Amini and Wainwright [4] considered the single spike covariance model (originally proposed by Johnstone [18]). In particular, [19] showed that conventional PCA is consistent if and only if $d(n)/n \to 0$; furthermore, when $\log(d \vee n)/n \to 0$, they proved that PCA could regain consistency when being performed on a subset of variables with the largest sample variances. Note that our asymptotic framework of $d \to \infty$ with $n$ fixed is not considered by [19], who also modeled the sparsity differently from us. [4] further restricted the maximal eigenvector to have $k$ non-zero entries, and studied support set recovery properties of the thresholding subset PCA procedure of [19] and the DSPCA procedure of [12]. Different from [4], our paper studies asymptotic properties of estimating the actual maximal eigenvector, instead of its support set. Hence, our results are complementary to those of [4], as elaborated below in the context of Example 1.2. Paul and Johnstone [33] developed the augmented sparse PCA procedure along with its optimal rate of convergence property. Ma [27] proposed an iterative thresholding procedure for estimating principal

subspaces, and established its nice theoretical properties. Considering the sample size $n \to \infty$ with dimension $d$ fixed, Leng and Wang [26] proposed an adaptive lasso sparse PCA, and established its consistency for selecting non-zero loadings.

Asymptotics on sparsity have been investigated in other high dimensional settings, such as regression and variable selection by Meinshausen and Bühlmann [29], Candes and Tao [10], van de Geer [39], Bickel et al. [7], Meier et al. [28], Omidiran and Wainwright [31], and Obozinski et al. [30], sparse covariance matrix estimation by Bickel and Levina [5,6], and El Karoui [15], classification by Bühlmann [8], and density estimation by Bunea et al. [9], for example.

### 1.1. Notation and assumptions

All quantities are indexed by the dimension $d$ in this paper. However, when it will not lead to confusion, the subscript $d$ will be omitted for convenience. Let the population covariance matrix be $\Sigma_d$. The eigen-decomposition of $\Sigma_d$ is

$$\Sigma_d = U_d \Lambda_d U_d^T,$$

where $\Lambda_d$ is the diagonal matrix of the population eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ and $U_d$ is the matrix of corresponding population eigenvectors so that $U_d = [u_1, \ldots, u_d]$.

Assume that $X_1, \ldots, X_n$ are random samples from a $d$-dimensional normal distribution $N(0, \Sigma_d)$. Denote the data matrix by $X_{(d)} = [X_1, \ldots, X_n]_{d \times n}$ and the sample covariance matrix by $\hat{\Sigma}_d = n^{-1} X_{(d)} X_{(d)}^T$. Then, the sample covariance matrix $\hat{\Sigma}_d$ can be similarly decomposed as

$$\hat{\Sigma}_d = \hat{U}_d \hat{\Lambda}_d \hat{U}_d^T,$$

where $\hat{\Lambda}_d$ is the diagonal matrix of the sample eigenvalues $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_d$ and $\hat{U}_d$ is the matrix of the corresponding sample eigenvectors so that $\hat{U}_d = [\hat{u}_1, \ldots, \hat{u}_d]$.

Let $\bar{u}_i$ be any sample based estimator of $u_i$, e.g. $\bar{u}_i = \hat{u}_i$ for $i = 1, \ldots, d$. Two important concepts from [20,21] are:

- *Consistency*: The direction $\bar{u}_i$ is *consistent* with its population counterpart $u_i$ if

$$\text{Angle}(\bar{u}_i, u_i) \equiv \arccos(|\langle \bar{u}_i, u_i \rangle|) \xrightarrow{p} 0, \quad \text{as } d \to \infty, \tag{1.1}$$

  where $\langle \cdot, \cdot \rangle$ denotes the inner product between two vectors.
- *Strong inconsistency*: The direction $\bar{u}_i$ is *strongly inconsistent* with its population counterpart $u_i$ if $\text{Angle}(\bar{u}_i, u_i) \xrightarrow{p} \frac{\pi}{2}$, as $d \to \infty$.

In addition, we consider two important concepts in the current paper:

- *Consistency with convergence rate* $d^\iota$: The direction $\bar{u}_i$ is consistent with its population counterpart $u_i$ with the convergence rate $d^\iota$ if $|\langle \bar{u}_i, u_i \rangle| = 1 + o_p(d^{-\iota})$, where the notation $G_d \equiv o_p(d^{-\iota})$ means that $d^\iota G_d \xrightarrow{p} 0$, as $d \to \infty$.
- *Marginal inconsistency*: The direction $\bar{u}_i$ is *marginally inconsistent* with $u_i$ if $\text{Angle}(\bar{u}_i, u_i)$ converges to a (possibly random) quantity in $(0, \frac{\pi}{2})$, as $d \to \infty$.

We now use two illustrative examples to highlight our key theoretical results. The examples are chosen mainly for intuitive illustration. Our theorems cover more general single component spike models (Sections 2–4).

**Example 1.1.** Assume that $X_1, \ldots, X_n$ are random sample vectors from a $d$-dimensional normal distribution $N(0, \Sigma_d)$, where the covariance matrix $\Sigma_d$ has the eigenvalues as

$$\lambda_1 = d^\alpha, \qquad \lambda_2 = \cdots = \lambda_d = 1, \quad \alpha \geq 0. \tag{1.2}$$

This is a special case of the single component spike covariance Gaussian model considered before by, for example, Johnstone [18], Paul [32], Johnstone and Lu [19], Amini and Wainwright [4].

Without loss of generality (WLOG), we further assume that the first eigenvector $u_1$ is proportional to the $d$-dimensional vector

$$\dot{u}_1 = (\overbrace{1, \ldots, 1}^{\lfloor d^\beta \rfloor}, 0, \ldots, 0)^T,$$

where $0 \leq \beta \leq 1$ and $\lfloor d^\beta \rfloor$ denotes the integer part of $d^\beta$. For example, if $\beta = 0$, the first population eigenvector becomes $u_1 = (1, 0, \ldots, 0)^T$. (Note that in general the non-zero entries do not have to be the first $\lfloor d^\beta \rfloor$ elements, nor do they need to have equal values.)

We formally define $\alpha$ as the *spike index* that measures the strength of the spike, and $\beta$ as the *sparsity index* that quantifies the sparsity of the maximal eigenvector $u_1$, where $\lfloor d^\beta \rfloor$ is the number of its non-zero elements. Under Model (1.2), Jung and Marron [20] showed that the first empirical eigenvector (the PC direction) $\hat{u}_1$ is consistent with $u_1$ when $\alpha > 1$; however for $\alpha < 1$, it is strongly inconsistent. Jung, Sen and Marron [21] then showed that $\hat{u}_1$ is in between consistent and strongly inconsistent on the boundary when $\alpha = 1$. Again, one main point of the present paper is an exploration of conditions under which sparse methods can lead to consistency even when the spike index $\alpha \leq 1$, by exploiting *sparsity*.

**Example 1.2.** Our theorems are also applicable to the sparse single-component spike model considered by Amini and Wainwright [4], although we have a different focus from them as illustrated below. The covariance matrix $\Sigma_d$ can be expressed using our notation as

$$\Sigma_d = (\lambda_1 - 1)z^* z^{*T} + \begin{pmatrix} I_{\lfloor d^\beta \rfloor} & 0 \\ 0 & \Gamma_{d-\lfloor d^\beta \rfloor} \end{pmatrix},$$

where the first eigenvalue $\lambda_1 > 1$, the first $\lfloor d^\beta \rfloor$ entries of the maximal eigenvector are non-zero with values of $\pm 1/\sqrt{\lfloor d^\beta \rfloor}$, and $\Gamma_{d-\lfloor d^\beta \rfloor}$ is a symmetric positive semi-definite matrix with the maximal eigenvalue $\lambda_{\max}(\Gamma_{d-\lfloor d^\beta \rfloor}) \leq 1$. For this example, consider cases where all eigenvalues of $\Gamma_{d-\lfloor d^\beta \rfloor}$ are one. Hence, the eigenvalues of $\Sigma_d$ are $\lambda_1 > 1 = \lambda_2 = \cdots = \lambda_d$.[4] focused on the consistent recovery of the support set of the maximal eigenvector $z^*$. We, however, are interested in the consistency of the actual direction vector. We note that the two types of consistency are not equivalent. For the above model, the two results are summarized below.

- Assuming fixed $\lambda_1$ and $n$, $d \to \infty$, [4] showed that the support set can be recovered if $n > c_u(\lfloor d^\beta \rfloor)^2 \log(d - \lfloor d^\beta \rfloor)$, while it cannot be recovered if $n < c_l(\lfloor d^\beta \rfloor)^2 \log(d - \lfloor d^\beta \rfloor)$, where $c_u$ and $c_l$ are two constants.
- For fixed $n$, $\lambda_1 = d^\alpha$ and $d \to \infty$, our Theorems 2.2 and 3.3 show that sparse PCA is consistent when $\alpha > \beta$, although the support set may not be consistently recovered (Theorem 3 of AW); and our Theorem 4.1 indicates that sparse PCA is strongly inconsistent when $\alpha < \beta$, even when one knows the exact support set.

### *1.2. Roadmap of the paper*

The organization of the rest of the paper is as follows. For easy access to the main ideas, Section 2 proves the consistency of a simple thresholding (ST) method that generates sparse PC directions, under the sparsity and small spike conditions where the conventional PCA is strongly inconsistent. Section 3 then generalizes these ideas to establish the consistency of the RSPCA method of [35]. Section 4 identifies the region and its boundaries where the strong inconsistency and marginal inconsistency of an appropriate oracle sparse PCA procedure are proved. Section 5 reports simulation results to illustrate both consistency and strong inconsistency of PCA and sparse PCA. Section 6 concludes the paper with some discussion of future work. Section 7 contains the proofs of the main theorems.

## 2. Consistency of simple thresholding sparse PCA

In Example 1.1, the first eigenvector of the sample covariance matrix $\hat{u}_1$ is strongly inconsistent with $u_1$ when $\alpha < 1$, because it attempts to estimate too many parameters. Sparse data analytic methods assume that many of these parameters are zero, which can allow greatly improved estimation of the first PC direction $u_1$. Here, this issue is explored in the context of sparse PCA. The sample covariance matrix based estimator, $\hat{u}_1$, can be improved by exploiting the fact that $u_1$ has many zero elements.

We first study a natural simple thresholding (ST) method where entries with small absolute values are replaced by zero. (Starting with the ST approach makes it easier to demonstrate the key ideas that are also useful for establishing the consistency of a more sophisticated sparse PCA method in Section 3.) In HDLSS contexts, it is challenging to apply thresholding directly to $\hat{u}_1$, because the number of its entries grows rapidly as $d \to \infty$, which naturally shrinks their magnitudes given that $\hat{u}_1$ has norm one. Thresholding is more conveniently formulated in terms of the *dual covariance matrix* [21].

Denote the dual sample covariance matrix by $S_d = \frac{1}{n} X_{(d)}^T X_{(d)}$ and the first dual eigenvector by $\tilde{v}_1$. The sample eigenvector $\hat{u}_1$ is connected with the dual eigenvector $\tilde{v}_1$ through the following transformation,

$$\tilde{u}_1 = (\tilde{u}_{1,1}, \ldots, \tilde{u}_{d,1})^T = X_{(d)} \tilde{v}_1, \tag{2.1}$$

and the sample estimate is then given by $\hat{u}_1 = \tilde{u}_1 / \|\tilde{u}_1\|$ [21].

Given a sequence of threshold values $\zeta$, define the thresholded entries as

$$\breve{u}_{i,1} = \begin{cases} \tilde{u}_{i,1} & \text{if } |\tilde{u}_{i,1}| > \zeta, \\ 0 & \text{if } |\tilde{u}_{i,1}| \leq \zeta, \end{cases} \quad \text{for } i = 1, \ldots, d. \tag{2.2}$$

Denote $\breve{u}_1 = (\breve{u}_{1,1}, \ldots, \breve{u}_{d,1})^T$ and normalize it to get the simple thresholding (ST) estimator $\hat{u}_1^{\mathrm{ST}} = \breve{u}_1 / \|\breve{u}_1\|$.

For the model in Example 1.1, given an eigenvalue of strength $\alpha \in (0, 1)$, (recall that $\lambda_1 = d^\alpha$ and $\hat{u}_1$ is strongly inconsistent), below we explore conditions on the threshold sequence $\zeta$ under which the ST estimator $\hat{u}_1^{\mathrm{ST}}$ is in fact consistent with $u_1$. First of all, the threshold $\zeta$ cannot be too large; otherwise all the entries will be zeroed out. It will be seen in Theorem 2.1 that a sufficient condition for this is $\zeta \leq d^{\frac{\gamma}{2}}$, where $\gamma \in (0, \alpha)$. Secondly, the threshold $\zeta$ cannot be too small, or pure noise terms will be included. A parallel sufficient condition is shown to be $\zeta \geq \log^\delta(d) \lambda_2^{\frac{1}{2}}$, where $\delta \in (\frac{1}{2}, \infty)$.

Below we formally establish conditions on the eigenvalues of the population covariance matrix $\Sigma_d$ and the thresholding parameter $\zeta$, which give consistency of $\hat{u}_1^{\mathrm{ST}}$ to $u_1$. The proofs are provided in Section 7.

To fix ideas, we first consider the extreme sparsity case $u_1 = (1, 0, \ldots, 0)^T$. Suppose that $\lambda_1 \sim d^\alpha$, in the sense that $0 < c_1 \leq \underline{\lim}_{d\to\infty} \frac{\lambda_1}{d^\alpha} \leq \overline{\lim}_{d\to\infty} \frac{\lambda_1}{d^\alpha} \leq c_2$, for two constants $c_1$ and $c_2$. WLOG, assume $\sum_{i=2}^d \lambda_i \sim d$. We also need the second eigenvalue $\lambda_2$ to be an obvious distance away from the first eigenvalue $\lambda_1$. If not, it is hard to distinguish the first and second empirical eigenvectors as observed by Jung and Marron [20], among others. In that case the appropriate amount of thresholding on the first empirical eigenvector becomes unclear. Therefore, we assume that $\lambda_2 \sim d^\theta$, where $\theta < \alpha$.

**Theorem 2.1.** *Suppose that $X_1, \ldots, X_n$ are random samples from a $d$-dimensional normal distribution $N(0, \Sigma_d)$ and the first population eigenvector $u_1 = (1, 0, \ldots, 0)^T$. If the following conditions are satisfied:*

(a) $\lambda_1 \sim d^\alpha$, $\lambda_2 \sim d^\theta$, *and* $\sum_{i=2}^d \lambda_i \sim d$, *where* $\theta \in [0, \alpha)$ *and* $\alpha \in (0, 1]$,
(b) $\log^\delta(d)\, d^{\frac{\theta}{2}} \leq \zeta \leq d^{\frac{\gamma}{2}}$, *where* $\delta \in (\frac{1}{2}, \infty)$ *and* $\gamma \in (\theta, \alpha)$,

*then the simple thresholding estimator $\hat{u}_1^{\mathrm{ST}}$ is consistent with $u_1$.*

In fact, $u_1 = (1, 0, \ldots, 0)^T$ in Theorem 2.1 is a very extreme case. The following theorem considers the general case $u_1 = (u_{1,1}, \ldots, u_{d,1})^T$, where only $\lfloor d^\beta \rfloor$ elements of $u_1$ are non-zero. WLOG, we assume that the first $\lfloor d^\beta \rfloor$ entries are non-zero just for notational convenience.

Define

$$Z_j \equiv (z_{1,j}, \ldots, z_{d,j})^T = (X_j^T u_1, \ldots, X_j^T u_d)^T, \quad j = 1, \ldots, n. \tag{2.3}$$

We can show that $Z_j$ are i.i.d. $N(0, \mathrm{diag}\{\lambda_1, \ldots, \lambda_d\})$ random vectors. Let

$$W_j \equiv (w_{1,j}, \ldots, w_{d,j})^T = \left(\lambda_1^{-\frac{1}{2}} z_{1,j}, \ldots, \lambda_d^{-\frac{1}{2}} z_{d,j}\right)^T, \quad j = 1, \ldots, n, \tag{2.4}$$

and the $W_j$ are i.i.d. $N(0, I_d)$ random vectors, where $I_d$ is the $d$-dimensional identity matrix.

The following conditions are also needed to ensure the consistency of $\hat{u}_1^{\mathrm{ST}}$:

- The non-zero entries of the population eigenvector $u_1$ need to be a certain distance away from zero. In fact, if the non-zero entries of the first population eigenvector are close to zero, the corresponding entries of the first empirical eigenvector would also be small and look like pure noise entries. Thus, we assume

  $$\max_{1\leq i\leq \lfloor d^\beta \rfloor} |u_{i,1}|^{-1} \sim d^{\frac{\eta}{2}}, \quad \text{where } \eta \in [0, \alpha).$$

- From (2.3), we have

  $$X_j = \sum_{i=1}^d z_{i,j} u_i, \quad j = 1, \ldots, n.$$

  Since $z_{1,j}$ has the largest variance $\lambda_1$, then $z_{1,j}u_1$ contributes the most to the variance of $X_j$, $j = 1, \ldots, n$. Note that $z_{1,j}u_1$ is consistent with $u_1$, and so $z_{1,j}u_1$ is the key to making the simple thresholding method work. So we need to show that the remaining parts

  $$H_j \equiv (h_{1,j}, \ldots, h_{d,j})^T = \sum_{i=2}^d z_{i,j} u_i, \quad j = 1, \ldots, n \tag{2.5}$$

  have a negligible effect on the direction vector $\hat{u}_1^{\mathrm{ST}}$.
- Suppose that the $H_j$ are i.i.d. $N(0, \Delta_d)$, where $\Delta_d = (m_{kl})_{d\times d}$. A sufficient condition to make their effect negligible is the following mixing condition of Leadbetter, Lindgren and Rootzen [22]:

  $$|m_{kl}| \leq m_{kk}^{\frac{1}{2}} m_{ll}^{\frac{1}{2}} \rho_{|k-l|}, \quad 1 \leq k \neq l \leq \lfloor d^\beta \rfloor, \tag{2.6}$$

  where $\rho_t < 1$ for all $t > 1$ and $\rho_t \log(t) \longrightarrow 0$, as $t \to \infty$. This mixing condition can guarantee that $\max_{1\leq j\leq n} |h_{1,j}|$ has a quick convergence rate, as $d \to \infty$. It enables us to neglect the influence of $H_j$ for sufficiently large $d$ and make $z_{i,j}u_1$ the dominant component, which then gives consistency to the first population eigenvector $u_1$. Thus the thresholding estimator $\hat{u}_1^{\mathrm{ST}}$ becomes consistent.
- The mixing condition (2.6) allows the sparsity index to be $0 \leq \beta \leq 1$. For $\beta = 0$, the mixing condition is always satisfied because the set $\{k, l \mid 1 \leq k \neq l \leq \lfloor d^\beta \rfloor = 1\}$ is empty. For $\beta > 0$ and under the sparsity assumption in Example 1.1, $|m_{kl}| = \lfloor d^\beta \rfloor^{-1}$ and $m_{kk} = m_{ll} = 1 - \lfloor d^\beta \rfloor^{-1}$. In this case, we can choose $\rho_t = \frac{1}{t - t\lfloor d^\beta \rfloor^{-1}}$ in order for the mixing condition to be satisfied.

We now state one of the main theorems:

**Theorem 2.2.** *Assume that $X_1, \ldots, X_n$ are random samples from a $d$-dimensional normal distribution $N(0, \Sigma_d)$. Define $Z_j$, $W_j$ and $H_j$ as in (2.3)–(2.5) for $j = 1, \ldots, n$. The first population eigenvector is $u_1 = (u_{1,1}, \ldots, u_{d,1})^T$ with $u_{i,1} \neq 0, i = 1, \ldots, \lfloor d^\beta \rfloor$, and otherwise $u_{i,1} = 0$.*

*If the following conditions are satisfied:*

(a) $\lambda_1 \sim d^\alpha$, $\lambda_2 \sim d^\theta$, *and* $\sum_{i=2}^d \lambda_i \sim d$, *where* $\theta \in [0, \alpha)$ *and* $\alpha \in (0, 1]$,
(b) $\max_{1 \leq i \leq \lfloor d^\beta \rfloor} |u_{i,1}|^{-1} \sim d^{\frac{\eta}{2}}$, *where* $\eta \in [0, \alpha)$,
(c) $H_j$ *satisfies the mixing condition (2.6),* $j = 1, \ldots, n$,
(d) $\log^\delta(d)\, d^{\frac{\theta}{2}} \leq \zeta \leq d^{\frac{\gamma}{2}}$, *where* $\delta \in (\frac{1}{2}, \infty)$ *and* $\gamma \in (\theta, \alpha - \eta)$,

*then the thresholding estimator $\hat{u}_1^{\mathrm{ST}}$ is consistent with $u_1$.*

We offer a couple of remarks regarding Theorem 2.2. First of all, the theorem naturally reduces to Theorem 2.1 if we let the sparsity index $\beta = 0$. Secondly, Condition (b) implies $0 \leq \beta < \alpha$. Considering Example 1.1 or more general cases where all nonzero entries of $u_1$ have the same magnitude so that $u_{i1} \sim u_{11}$, $1 \leq i \leq \lfloor d^\beta \rfloor$, then Condition (b) holds with $\eta = \beta$ and thus it is equivalent to requiring $0 \leq \beta < \alpha$. More generally, let $S$ denote the set of all norm-1 vectors $u$ such that the first $\lfloor d^\beta \rfloor$ entries are nonzero with the rest being zero. Then, Condition (b) would be violated if $\beta \geq \alpha$ since

$$\min_{u_1 \in S} \max_{1 \leq i \leq \lfloor d^\beta \rfloor} |u_{i1}|^{-1} \sim d^{\beta/2}.$$

Finally, this theorem, and the following ones in Sections 2–4, suggest that the concepts depicted in Fig. 1 hold much more generally than just the models in Examples 1.1 and 1.2. In particular, in the above theorem, setting $\theta = 0$ and $\eta = \beta$ would give the results plotted in Fig. 1.

In addition, for different thresholding parameter $\zeta$, the ST estimator $\hat{u}_1^{\mathrm{ST}}$ is consistent with different convergence rate, as stated in the following theorem. The notation $\zeta = o(d^\rho)$ below means that $\zeta d^{-\rho} \to 0$ as $d \to \infty$.

**Theorem 2.3.** *For the thresholding parameter $\zeta = o(d^{\frac{\alpha-\eta-\kappa}{2}})$, where $\kappa \in [0, \alpha - \eta - \theta)$, the corresponding thresholding estimator $\hat{u}_1^{\mathrm{ST}}$ is consistent with $u_1$, with a convergence rate of $d^{\frac{\kappa}{2}}$.*

## 3. Consistency of RSPCA

As noted in Section 1, several sparse PCA methods have been proposed in the literature. Here we perform a detailed HDLSS asymptotic analysis of the sparse PCA procedure developed by Shen and Huang [35]. For simplicity, we refer to it as the *regularized sparse PCA*, or RSPCA for short. All the detailed proofs are again provided in Section 7.

We start with briefly reviewing the methodological details of RSPCA. (For more details, see [35].) Given a $d$-by-$n$ data matrix $X_{(d)}$, consider the following penalized sum-of-squares criterion:

$$\|X_{(d)} - uv^T\|_F^2 + P_\zeta(u), \quad \text{subject to } \|v\| = 1, \tag{3.1}$$

where $u$ is a $d$-vector, $v$ is a unit $n$-vector, $\|\cdot\|_F$ denotes the Frobenius norm, and $P_\zeta(u) = \sum_{i=1}^d p_\zeta\left(|u_{i,1}|\right)$ is a penalty function with $\zeta \geq 0$ being the penalty parameter. The penalty function can be any sparsity-inducing penalty. In particular, Shen and Huang [35] considered the soft thresholding (or $L_1$ or LASSO) penalty of Tibshirani [38], the hard thresholding penalty of Donoho and Johnstone [13], and the smoothly clipped absolute deviation (SCAD) penalty of Fan and Li [16].

Without the penalty term or when $\zeta = 0$, minimization of (3.1) can be obtained via singular value decomposition (SVD) [14], which results in the best rank-one approximation of $X_{(d)}$ as $\tilde{u}_1\tilde{v}_1^T$, where $\tilde{u}_1$ and $\tilde{v}_1^T$ minimize the criterion (3.1). The normalized $\tilde{u}_1$ turns out to be the first empirical PC loading vector. With the penalty term, Shen and Huang [35] define the sparse PC loading vector as $\hat{u}_1 = \tilde{u}_1/\|\tilde{u}_1\|$ where $\tilde{u}_1$ is now the minimizer of (3.1) with the penalty term included. The minimization now needs to be performed iteratively. For a given $\tilde{v}_1$ in the criterion (3.1), we can get the minimizing vector as $\tilde{u}_1 = h_\zeta\left(X_{(d)}\tilde{v}_1\right)$, where $h_\zeta$ is a thresholding function that depends on the particular penalty function used and the penalty (or thresholding) parameter $\zeta$. See [35] for more details. The thresholding is applied to the vector $X_{(d)}\tilde{v}_1$ componentwise.

Shen and Huang [35] proposed the following iterative procedure for minimizing the criterion (3.1):

The RSPCA Algorithm

1. Initialize:
   (a) Use SVD to obtain the best rank-one approximation $\tilde{u}_1\tilde{v}_1^T$ of the data matrix $X_{(d)}$, where $\tilde{v}_1$ is a unit vector.
   (b) Set $\tilde{u}_1^{\mathrm{old}} = \tilde{u}_1$ and $\tilde{v}_1^{\mathrm{old}} = \tilde{v}_1$.
2. Update:
   (a) $\tilde{u}_1^{\mathrm{new}} = h_\zeta\left(X_{(d)}\tilde{v}_1^{\mathrm{old}}\right)$.
   (b) $\tilde{v}_1^{\mathrm{new}} = \frac{X_{(d)}^T\tilde{u}_1^{\mathrm{new}}}{\|X_{(d)}^T\tilde{u}_1^{\mathrm{new}}\|}$.
3. Repeat Step 2 setting $\tilde{u}_1^{\mathrm{old}} = \tilde{u}_1^{\mathrm{new}}$ and $\tilde{v}_1^{\mathrm{old}} = \tilde{v}_1^{\mathrm{new}}$ until convergence.
4. Normalize the final $\tilde{u}_1^{\mathrm{new}}$ to get $\hat{u}_1$, the desired sparse loading vector.

There exists a nice connection between the ST method of Section 2 and RSPCA with hard thresholding. The ST estimator $\hat{u}_1^{ST}$ is exactly the sparse loading vector $\hat{u}_1$ obtained from the first iteration of the RSPCA iterative algorithm, when the hard thresholding penalty is used. Such a connection suggests that we can extend the proofs for the theorems regarding the property of the ST estimator to establish consistency theorems for RSPCA below. Although nicely connected, RSPCA actually performs better numerically over ST as we will illustrate in Section 5.

Below we develop conditions under which the RSPCA estimator $\hat{u}_1$ is consistent with the population eigenvector $u_1$ when a proper thresholding parameter $\zeta$ is used. All three of the soft thresholding, hard thresholding or SCAD penalties are considered. To prove the consistency of RSPCA, we first establish in Theorem 3.1 the consistency of the first-step RSPCA estimator obtained during the initial iteration, which can then be used to show in Theorem 3.3 that the follow-up updated RSPCA estimator remains consistent.

The following Theorem 3.1 states conditions when the first-step RSPCA estimator $\hat{u}_1$ is consistent with $u_1$ under a proper thresholding parameter $\zeta$. The theorem also shows that for different parameters $\zeta$, the RSPCA estimator $\hat{u}_1$ is consistent with different convergence rates. Given the aforementioned connection between ST and RSPCA, the consistency and convergence rate conditions are the same as those needed for ST. See the discussion of Theorems 2.2 and 2.3 for the implications of these conditions.

**Theorem 3.1.** *Under the assumptions and conditions of Theorem 2.2, the first-step sparse loading vector $\hat{u}_1$ is consistent with $u_1$. In addition, for the thresholding parameter $\zeta = o(d^{\frac{\alpha-\eta-\kappa}{2}})$, where $\kappa \in [0, \alpha-\eta-\theta)$, $\hat{u}_1$ is consistent with a convergence rate of $d^{\frac{\kappa}{2}}$.*

To obtain an updated estimate for $v_1$ we set $\hat{u}_1^{\text{old}}$ to be the consistent first-step RSPCA estimate, and obtain $\tilde{v}_1^{\text{new}} = X_{(d)}^T \hat{u}_1^{\text{old}} / \|X_{(d)}^T \hat{u}_1^{\text{old}}\|$. Theorem 3.2 studies the asymptotic properties of $\tilde{v}_1^{\text{new}}$.

**Theorem 3.2.** *Assume that $\hat{u}_1^{\text{old}}$ is consistent with $u_1$ with the convergence rate $d^{\frac{\kappa}{2}}$, where $\kappa \in [1-\alpha, \infty)$. If the $\varepsilon_2$-condition is satisfied, then*

$$\tilde{v}_1^{\text{new}} \xrightarrow{p} \frac{\tilde{W}_1}{\|\tilde{W}_1\|}, \quad \text{as } d \to \infty,$$

*where $\tilde{W}_1 = (w_{1,1}, \ldots, w_{1,n})$ follows a standard $n$-dimensional normal distribution $N(0, I_n)$ and the $w_{i,j}$ are defined in (2.4).*

Given the above established asymptotic properties of $\tilde{v}_1^{\text{new}}$, we can now study the asymptotic properties of the updated RSPCA estimator

$$\hat{u}_1^{\text{new}} = \frac{\tilde{u}_1^{\text{new}}}{\|\tilde{u}_1^{\text{new}}\|}, \quad \text{with } \tilde{u}_1^{\text{new}} = h_\zeta(X_{(d)} \tilde{v}_1^{\text{new}}), \tag{3.2}$$

as defined in the iterative RSPCA algorithm. The following Theorem 3.3 shows that, with a proper choice of $\zeta$, the updated RSPCA estimator $\hat{u}_1^{\text{new}}$ remains to be consistent with the population eigenvector $u_1$, and for different $\zeta$, $\hat{u}_1^{\text{new}}$ is consistent with different convergence rates.

**Theorem 3.3.** *Under the assumptions and conditions of Theorems 2.2 and 3.2, the updated loading vector $\hat{u}_1^{\text{new}}$ (3.2) is consistent with $u_1$. In addition, for the thresholding parameter $\zeta = o(d^{\frac{\alpha-\eta-\kappa}{2}})$, where $\kappa \in [0, \alpha-\eta-\theta)$, $\hat{u}_1^{\text{new}}$ is consistent with $u_1$, with convergence rate $d^{\frac{\kappa}{2}}$.*

Theorems 3.1 and 3.3 imply that, if $\alpha - \eta - \theta > 1 - \alpha$, then by choosing the thresholding parameter to be $\zeta = o(d^{\frac{\alpha-\eta-\kappa}{2}})$, we can make the updated RSPCA loading vector $\hat{u}_1^{\text{new}}$ consistent with $u_1$ at every updating step.

Interestingly, Theorems 2.3 and 3.3 suggest that the ST estimator and the RSPCA estimator share the same rate of convergence. However, we note that the RSPCA estimator has better finite sample performance than ST: RSPCA is more stable due to the multiple iterations involved in the estimation, which reduce the estimation variability. The improvement of RSPCA over ST is illustrated in Fig. 4.

## 4. Strong inconsistency and marginal inconsistency

We have shown that we can attain consistency using sparse PCA, when the spike index $\alpha$ is greater than the sparsity index $\beta$. This motivates the question of consistency using sparse PCA when $\alpha$ is smaller than or equal to $\beta$. To answer this question, we consider an oracle estimator which "knows" the exact positions of the zero entries of the maximal eigenvector $u_1$. We will show that even this oracle estimator is strongly inconsistent when $\alpha$ is smaller than $\beta$, and marginally inconsistent when $\alpha = \beta$. Compared with this oracle sparse PCA, threshold methods can perform no better because they also need to estimate location of the zero entries; hence threshold methods will also be inconsistent.

To make this precise, we study the procedure to generate the oracle estimator for general single component models. Similar to Sections 2 and 3, assume that the first $\lfloor d^\beta \rfloor$ entries of the maximal eigenvector $u_1$ are non-zero and the rest are all zero: $u_1 = (u_{1,1}, \ldots, u_{d,1})^T$, where $u_{i,1} \neq 0$, $i = 1, \ldots, \lfloor d^\beta \rfloor$; otherwise $u_{i,1} = 0$.

Let $X_j^* = (x_{1,j}, \ldots, x_{\lfloor d^\beta \rfloor, j})^T \sim N(0, \Sigma^*_{\lfloor d^\beta \rfloor})$, where $\Sigma^*_{\lfloor d^\beta \rfloor}$ is the covariance matrix of $X_j^*, j = 1, \ldots, n$. Then, the eigen-decomposition of $\Sigma^*_{\lfloor d^\beta \rfloor}$ is

$$\Sigma^*_{\lfloor d^\beta \rfloor} = U^*_{\lfloor d^\beta \rfloor} \Lambda^*_{\lfloor d^\beta \rfloor} (U^*_{\lfloor d^\beta \rfloor})^T,$$

where $\Lambda_d^*$ is the diagonal matrix of eigenvalues $\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda^*_{\lfloor d^\beta \rfloor}$, and $U^*_{\lfloor d^\beta \rfloor}$ is the matrix of the corresponding eigenvectors so that $U^*_{\lfloor d^\beta \rfloor} = [u_1^*, \ldots, u^*_{\lfloor d^\beta \rfloor}]$.

Since the last $d - \lfloor d^\beta \rfloor$ entries of the maximal eigenvector $u_1$ equal zero, it follows that the first eigenvector $u_1^*$ of $\Sigma^*_{\lfloor d^\beta \rfloor}$ is formed by the non-zero entries of $u_1$, i.e. $u_1^* = (u_{1,1}, \ldots, u_{\lfloor d^\beta \rfloor, 1})^T$. So we have

$$u_1 = ((u_1^*)^T, \overbrace{0, \ldots, 0}^{d - \lfloor d^\beta \rfloor})^T. \tag{4.1}$$

Consider the following data matrix $X^*_{\lfloor d^\beta \rfloor} = [X_1^*, \ldots, X_n^*]$, and denote the sample covariance matrix by $\hat{\Sigma}^*_{\lfloor d^\beta \rfloor} = n^{-1} X^*_{\lfloor d^\beta \rfloor} X^T_{\lfloor d^\beta \rfloor}$. Then, the sample covariance matrix $\hat{\Sigma}^*_{\lfloor d^\beta \rfloor}$ can be similarly decomposed as

$$\hat{\Sigma}^*_{\lfloor d^\beta \rfloor} = \hat{U}^*_{\lfloor d^\beta \rfloor} \hat{\Lambda}^*_{\lfloor d^\beta \rfloor} (\hat{U}^*_{\lfloor d^\beta \rfloor})^T,$$

where $\hat{\Lambda}^*_{\lfloor d^\beta \rfloor}$ is the diagonal matrix of the sample eigenvalues $\hat{\lambda}_1^* \geq \hat{\lambda}_2^* \geq \cdots \geq \hat{\lambda}^*_{\lfloor d^\beta \rfloor}$, and $\hat{U}^*_{\lfloor d^\beta \rfloor}$ is the matrix of the corresponding sample eigenvectors so that $\hat{U}^*_{\lfloor d^\beta \rfloor} = [\hat{u}_1^*, \ldots, \hat{u}_d^*]$.

Then, we define the oracle (OR) estimator as

$$\hat{u}_1^{\mathrm{OR}} = ((\hat{u}_1^*)^T, \overbrace{0, \ldots, 0}^{d - \lfloor d^\beta \rfloor})^T. \tag{4.2}$$

The following theorem formally states its inconsistency properties.

**Theorem 4.1.** *Assume that $X_1, \ldots, X_n$ are random samples from a $d$-dimensional normal distribution $N(0, \Sigma_d)$.*

(a) *If $\lambda_1 \sim d^\alpha$, $\lambda_2 \sim d^\theta$, $\lambda_d \sim 1$ and $\sum_{i=2}^d \lambda_i \sim d$, where $\alpha < \beta$ and $\theta \in [0, \frac{\beta}{2})$, then the oracle estimator $\hat{u}_1^{\mathrm{OR}}$ in (4.2) is strongly inconsistent with $u_1$.*

(b) *If $\lambda_1 / d^\alpha \to c \in (0, \infty)$ and $\lambda_i \to c_\lambda, i = 2, \ldots, d$, where $\alpha = \beta$, then the oracle estimator $\hat{u}_1^{\mathrm{OR}}$ in (4.2) is marginally inconsistent with $u_1$ in that*

$$|\langle \hat{u}_1^{\mathrm{OR}}, u_1 \rangle| \Rightarrow \left(1 + \frac{c_\lambda}{c \chi_n^2}\right)^{-\frac{1}{2}}, \tag{4.3}$$

*where $\Rightarrow$ denotes convergence in distribution, and $\chi_n^2$ denotes the chi-squared distribution with $n$ degrees of freedom.*

## 5. Simulations for sparse PCA

We perform simulation studies to illustrate the performance of the ST method and the RSPCA with the hard thresholding penalty. We fix the sample size at $n = 25$ and consider a range of the dimension $d = 500, 1000, 2500, 5000, 10\,000$. The diverging $d$ allows us to study the convergence behavior of the methods. To generate the data matrix $X_{(d)}$, we first need to construct the population covariance matrix for $X_{(d)}$ that approximates the conditions of Theorems 2.2 and 3.1 when the spike index $\alpha$ is greater than the sparsity index $\beta$.

For the population covariance matrix, we consider the motivating model in Example 1.1 for the first population eigenvector and the eigenvalues, where the first eigenvalue $\lambda_1 = d^\alpha$ and the rest equal one, i.e. $\lambda_i = 1, i \geq 2$. For the additional population eigenvectors $u_i$, $2 \leq i \leq \lfloor d^\beta \rfloor$, let the last $d - \lfloor d^\beta \rfloor$ entries of these eigenvectors be zero. In particular, let the eigenvectors $u_i$, $2 \leq i \leq \lfloor d^\beta \rfloor$, be proportional to

$$\dot{u}_i = (\overbrace{1, \ldots, 1}^{i-1}, -i + 1, 0, \ldots, 0)^T.$$

After normalizing $\dot{u}_i$, we get the $i$-th eigenvector $u_i = \dot{u}_i / \|\dot{u}_i\|$. For $i > \lfloor d^\beta \rfloor$, let the $i$-th eigenvector have just one non-zero entry in the $i$-th position such that $u_i = (\overbrace{0, \ldots, 0}^{i-1}, 1, 0, \ldots, 0)^T$.

Then the data matrix is generated as

$$X_{(d)} = d^{\frac{\alpha}{2}} u_1 z_1^T + \sum_{i=2}^d u_i z_i^T,$$

where the $z_i$ follows the $n$-dimensional standard normal distribution.

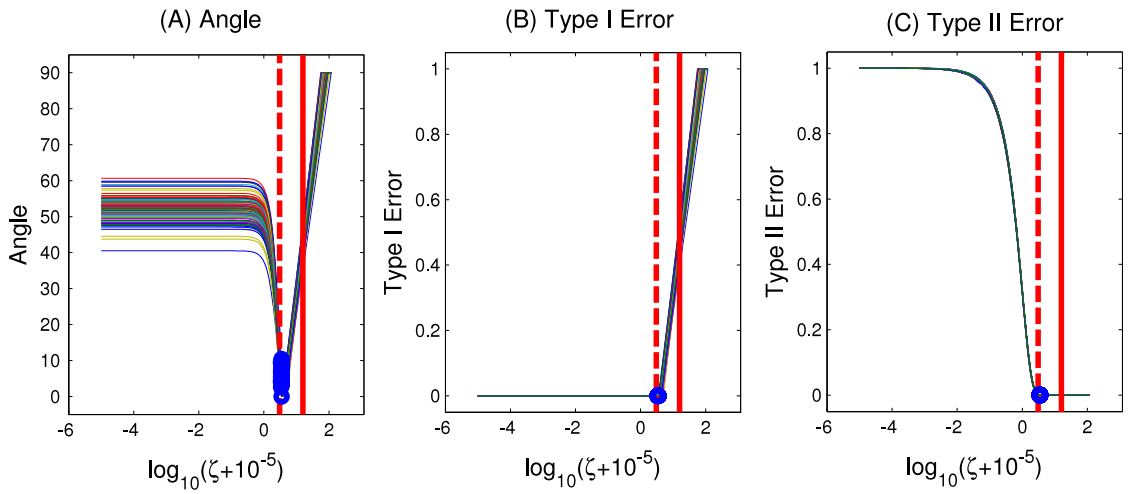


**Fig. 2.** Performance summary of RSPCA for spike index $\alpha = 0.6$ and sparsity index $\beta = 0.1$ where consistency is expected. Panel (A) shows angle to the maximal eigenvector as a function of thresholding parameter $\zeta$. Panel (B) and (C) are Type I Error and Type II Error as functions of $\zeta$. The vertical dashed and solid lines indicate the range of the thresholding parameter that leads to the consistency of RSPCA. The RSPCA performs nicely within the indicated range, which empirically confirms our asymptotic calculation. The circles indicate the BIC choice of $\zeta$.

We select twenty spike and sparsity pairs $(\alpha, \beta)$ with $\alpha = \{0.2, 0.4, 0.6, 0.8\}$ and $\beta = \{0, 0.1, 0.3, 0.5, 0.7\}$, which are shown in Fig. 1. We perform simulations for all twenty pairs under each specific dimension $d$, a total of 100 simulation setups. For each triplet $(\alpha, \beta, d)$, we generate 100 realizations of $X_{(d)}$. Results for four representative $(\alpha, \beta)$ pairs (indicated by the squares in Fig. 1) are reported below for the case of $d = 10\,000$, unless indicated otherwise. Additional simulation results can be found in an online supplement [36].

First of all, the plots in Fig. 2 summarize the results for the pair $(\alpha, \beta) = (0.6, 0.1)$, corresponding to one of the square dots in the white (consistent) triangular area of Fig. 1. For each replication of $X_{(d)}$ and a range of the thresholding parameter $\zeta$, we obtained the ST estimator $\hat{u}_1^{\mathrm{ST}}$ and the RSPCA estimator $\hat{u}_1$. Then we calculate the angle between the estimates $\hat{u}_1^{\mathrm{ST}}$ (or $\hat{u}_1$) and the true eigenvector $u_1$ through (1.1). Plotting this angle as a function of the thresholding parameter $\zeta$ gives the curve in Panel (A) of Fig. 2. Since ST and RSPCA perform very similarly in this case, only the RSPCA plots are shown in Fig. 2. The 100 simulation realizations generate the one-hundred curves in the panel. We rescale the thresholding parameter as $\log_{10}(\zeta + 10^{-5})$, to help reveal clearly the tendency of the angle curves as the thresholding parameter increases.

In these angle curves, the angles with $\zeta = 0$ (essentially the left edge of each curve) correspond to those obtained by the conventional PCA. Note that these angles are all over 40° which confirms the results of Jung and Marron [20] that when the spike index $\alpha < 1$, the conventional PCA is not consistent for $u_1$. As $\zeta$ increases, the angle remains stable for a while, then decreases to almost 0°, before eventually starting to increase to 90°. The plot suggests that RSPCA does improve over PCA for a range of $\zeta$. When $\zeta$ is large enough, all the entries of the RSPCA estimator will be zeroed out, so the estimator eventually becomes a $d$-dimensional zero vector, and the angle to $u_1$ goes to 90°.

Theorem 3.3 suggests a *consistent* range of the thresholding parameter $\zeta$, i.e. when $\zeta$ is inside this range, RSPCA would give a consistent estimator for $u_1$. The boundaries of this consistent range are indicated by the dashed and solid vertical lines in the angle plot. The range is very reasonable in the current case as the angles inside it are all small, and this offers an empirical validation for our asymptotic results.

Note that the consistent range of $\zeta$ depends on several unknown parameters: $\alpha$, $\beta$ and $\theta$. In practice, we propose to use the Bayesian Information Criterion (BIC) [34] to perform data-driven selection of the thresholding parameter $\zeta$. Zou, Hastie and Tibshirani [43] suggest the use of BIC to select the number of the non-zero coefficients for a lasso regression. Lee et al. [24] apply this idea to the sparse PCA context. Below we want to investigate how BIC performs numerically.

According to [24], for a fixed $\tilde{v}_1$, minimization of (3.1) with respect to $\tilde{u}_1$ is equivalent to minimizing the following penalized regression criterion:

$$\|X_{(d)} - \tilde{u}_1\tilde{v}_1^T\|_F^2 + P_\zeta(\tilde{u}_1) = \|Y - (I_d \bigotimes \tilde{v}_1)\tilde{u}_1\|^2 + P_\zeta(\tilde{u}_1), \tag{5.1}$$

where $Y = (X^1, \ldots, X^d)^T$, with $X^i$ being the $i$-th row of $X_{(d)}$, and $\bigotimes$ is the Kronecker product. Following [24], we define the BIC for (5.1) as

$$\mathrm{BIC}(\zeta) = \frac{\|Y - \hat{Y}\|^2}{nd\hat{\sigma}^2} + \frac{\log(nd)}{nd}\widehat{df}(\zeta), \tag{5.2}$$

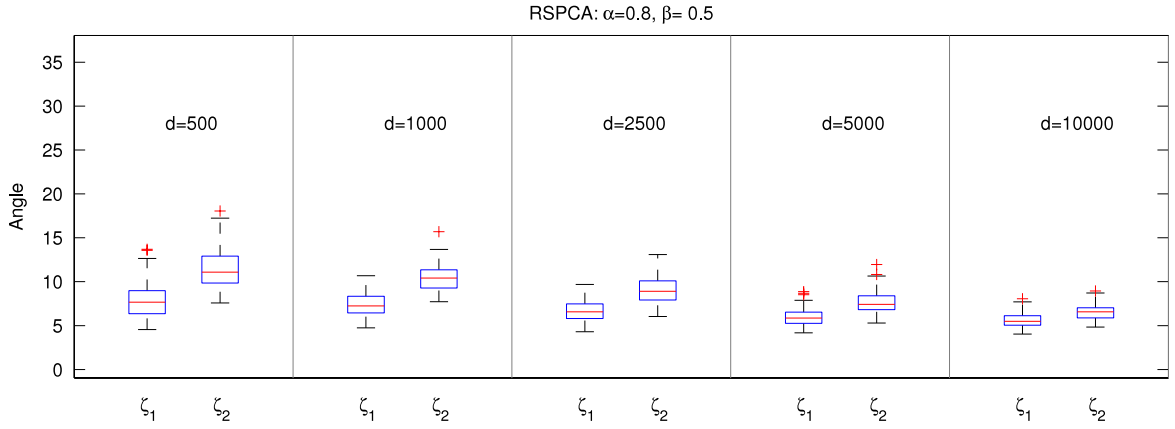


**Fig. 3.** Demonstration of convergence for growing $d$ in the case of $\alpha = 0.8$ and $\beta = 0.5$. Angle boxplots are shown for two parameter values $\zeta_1$ and $\zeta_2$ chosen within the RSPCA consistency region, under a sequence of increasing dimensions $d = 500$, 1000, 2500, 5000 and 10 000. One can see that $\zeta_1$ corresponds to a faster convergence rate than $\zeta_2$, which confirms the theoretical results in Theorem 3.3.

where $\hat{\sigma}^2$ is the ordinary-least squares estimate of the error variance, and $\widehat{df}(\zeta)$ is the degree of sparsity for the thresholding parameter $\zeta$, i.e. the number of non-zero entries in $\tilde{u}_1$. For every step of the iterative procedure of RSPCA, we can use BIC (5.2) to select the thresholding parameter and then obtain the corresponding sparse PC direction, until the algorithm converges.

For every angle curve in the angle plots of Fig. 2, we use a blue circle to indicate the thresholding parameter $\zeta$ that is selected by BIC during the last iterative step of RSPCA, and the corresponding angle. In the current $\alpha = 0.6$, $\beta = 0.1$ context, BIC works well, and all the BIC-selected $\zeta$ values are very close, so the 100 circles are essentially over plotted on each other. BIC also works well for the other spike and sparsity pairs $(\alpha, \beta)$ we considered where $\alpha > \beta$, as shown in [36].

Another measure of the success of a sparse estimator is in terms of which entries are zeroed. Type I Error is the proportion of non-zero entries in $u_1$ that are mistakenly estimated as zero. Type II Error is the proportion of zero entries in $u_1$ that are mistakenly estimated as non-zero. Similar to the angle curves in Panel (A), there are one hundred Type I Error and Type II Error curves in Panels (B) and (C) of Fig. 2, respectively. The dashed and solid vertical lines are the same as those in Panel (A). Note that for all the thresholding parameters in the range indicated by the lines, the errors are very small, which is again consistent with the asymptotic results of Theorem 3.3. The circles again are selected by BIC and they correspond to the same thresholding parameter, as in the angle plots. Thus, BIC works well here. BIC also generates similarly very small errors for the other spike and sparsity pairs $(\alpha, \beta)$ in Fig. 1 that satisfy $\alpha > \beta$.

In addition, our Theorem 3.3 shows that $\hat{u}_1$ converges at a different rate for different thresholding parameter $\zeta$. We empirically demonstrate this below in Fig. 3 for the case of $\alpha = 0.8$ and $\beta = 0.5$. We choose two threshold values $\zeta_1$ and $\zeta_2$ just inside the RSPCA consistency region. Theorem 3.3 suggests that $\hat{u}_1$ corresponding to $\zeta_1$ would converge faster as the dimension $d$ increases, which is exactly the case in Fig. 3. Similar phenomena can be observed for the other $(\alpha, \beta)$ scenarios as well as for the ST estimator, as shown in our online supplement [36].

Next we compare the performance among PCA, ST and RSPCA. In almost all cases, ST and RSPCA give better results than PCA and in some extreme cases, the three methods have similar poor performance. Although in most cases both ST and RSPCA have similar performance, however, there are some cases (for example when $\alpha = 0.4$ and $\beta = 0.3$), where RSPCA performs better than ST. For every simulation replication, we use BIC to select the thresholding parameter and obtain the ST and RSPCA estimators. We then calculate the angle, Type I Error and Type II Error for the three estimators, as well as the difference between ST and RSPCA (ST minus RSPCA). The 100 values of each measure are summarized using box plots in Fig. 4.

Panel (A) of Fig. 4 shows the box plots of the angles between the maximal eigenvector and the estimates obtained by PCA, ST and RSPCA, as well as the differences between the angles corresponding to ST and RSPCA. Note that the PCA angles are large, compared with ST and RSPCA, indicating the worse performance of PCA. The angle of ST seems larger than RSPCA. For a deeper view of this comparison, the pairwise differences are studied in the fourth box plot of the panel. The angle differences are almost always positive, with some differences bigger than 50°, which suggests that RSPCA has a better performance than ST. Similar conclusions can be made from the box plots of the errors in Panels (B) and (C). The box plot for PCA is not shown in Panel (C) because the corresponding Type II Error almost always equals one, which is far outside the shown range of interest.

The observed improvement of RSPCA over ST deserves some discussion, in connection with Theorems 2.3 and 3.3 that suggest the same rate of convergence for the two. This simulation compares finite sample performance between the two methods, and we believe the observed (and anticipated) improvement is due to the multiple iterations involved in RSPCA that reduce the variability and make the estimation more stable. Hence such improvement appears asymptotically in the constant coefficient, rather the rate.

In addition, as shown in the supplement [36], when $\alpha$ is small (e.g. $\alpha = 0.2$), even when $\alpha > \beta$ and $d = 10,000$, RSPCA and ST still perform badly. The reason is that the spike signal in this case is so weak that the theoretically consistent range of

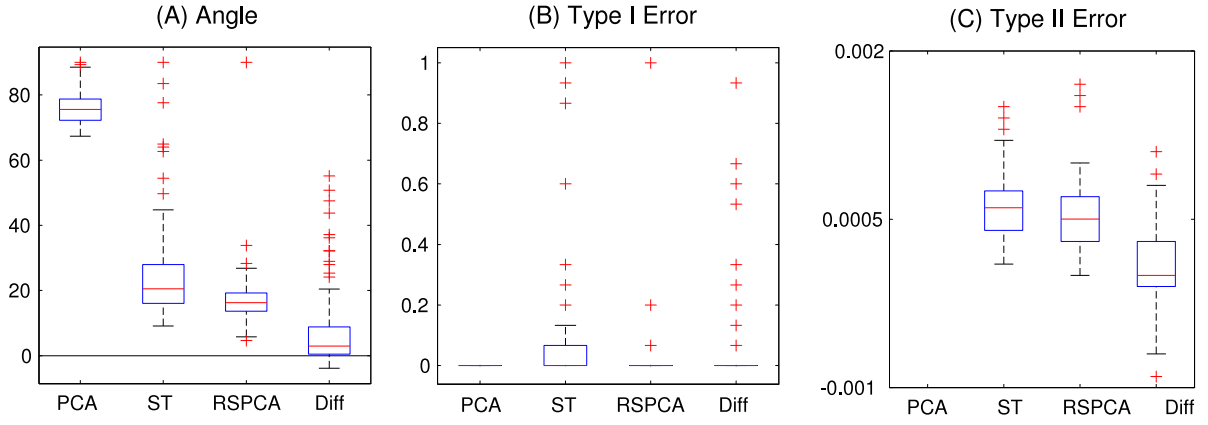


**Fig. 4.** Comparison of PCA, ST and RSPCA for spike index $\alpha = 0.4$ and sparsity index $\beta = 0.3$. Panels (A), (B) and (C) respectively contain four angle, Type I Error and Type II Error box plots: (i) conventional PCA; (ii) and (iii) ST and RSPCA with BIC; (iv) the difference between ST and RSPCA. In Panel (A), angles for conventional PCA are generally larger than ST and RSPCA which indicates the worse performance of contentional PCA. In addition, the angles and Type I Errors for ST are larger than RSPCA and their difference box plots furthermore confirm this point, which indicates the better performance of RSPCA in this case. ST and RSPCA has similar Type II Errors.

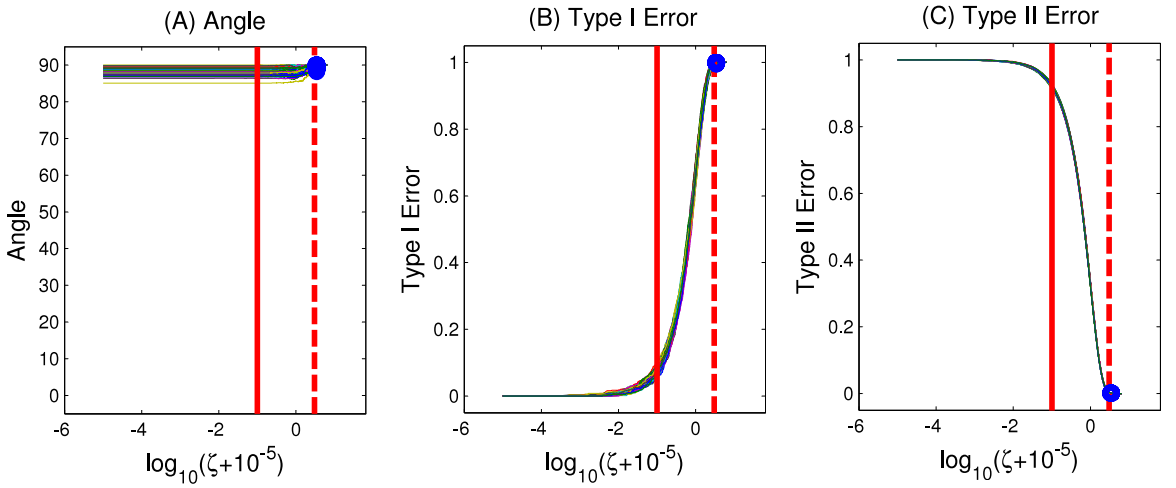


**Fig. 5.** Performance summary of RSPCA for spike index $\alpha = 0.2$ and sparsity index $\beta = 0.7$ where strong inconsistency is expected. Panel (A) shows the angle curves, between the RSPCA estimator and the first population eigenvector. Same format as Fig. 2. Since the spike index $\alpha = 0.2$ is smaller than the sparsity index $\beta = 0.7$, it follows that the right bound (solid line) is smaller than the left bound (dashed line). Thus the theorems do not give a meaningful range of the thresholding parameter. As expected, performance is very poor for any thresholding parameter $\zeta$.

the thresholding parameter $\zeta$ turns out to be empty, for example when $\alpha = 0.2$, $\beta = 0$ or 0.1, and $d = 10,000$, because the left bound in Condition (b) of Theorem 2.1 and Condition (d) of Theorem 2.2 is greater than the right bound. This suggests that, for small $\alpha$, we need an extremely large $d$ in order for RSPCA and ST to work.

Finally, Theorems 2.2 and 3.3 consider the condition that the spike index $\alpha$ is greater than the sparsity index $\beta$. When $\alpha$ is smaller than $\beta$, neither ST nor RSPCA is expected to give consistent estimation for the first population eigenvector $u_1$, as discussed in Section 4. For the spike and sparsity pairs $(\alpha, \beta)$ such that $\alpha < \beta$, the simulation results also confirm this point. Here, we display the simulation plots for the spike and sparsity pair $(\alpha, \beta) = (0.2, 0.7)$ in Fig. 5 as a representative of such simulations. Since ST and RSPCA have very similar performance here, we just show the simulation results for RSPCA. Similar to Fig. 2, the circles in Fig. 5 correspond to the thresholding parameter selected by BIC. From the angle plots, we can see that the angles, selected by BIC, are close to 90°, which suggests the failure of BIC in this case. In fact, all the angle curves are above 80°. Thus, neither ST nor RSPCA generates a reasonable sparse estimator. This is a common phenomenon when the spike index $\alpha$ is smaller than the sparsity index $\beta$. It is consistent with the theoretical investigation in Section 4.

Furthermore, the corresponding Type I Error, generated by ST or RSPCA with BIC, is close to one. This further confirms that BIC doesn't work when the spike index $\alpha$ is smaller than the sparsity index $\beta$. ST and RSPCA with $\zeta = 0$ is just the conventional PCA, and typically will not generate a sparse estimator. This entails that the Type I Error and Type II Error, corresponding to $\zeta = 0$, respectively equals zero and one. As the thresholding parameter increases, more and more entries are thresholded out; hence Type I Error increases to one and Type II Error decreases to zero.

## 6. Discussion

In this paper, we consider single-component Gaussian spike models under the HDLSS asymptotic framework where the sample size $n$ is fixed, the dimension $d$ and the maximal eigenvalue $\lambda_1$ both go to infinity. As a comparison, Johnstone and Lu [19] consider cases that $n$ and $d$ go to infinity together with $n/d$ converging to a constant, while fixing $\lambda_1$. The sizes of $n$ and $\lambda_1$ contribute positively to (i.e. encourage) the consistency of PCA and sparse PCA, while the dimension $d$ contributes negatively to (i.e. discourages) the consistency. It is interesting in future work to explore the transition from the random matrix asymptotic domain of [19] to our HDLSS domain of asymptotics.

It is also interesting to consider a multiple component spike model as in Paul and Johnstone [33], Jung and Marron [20] and Ma [27]. One theoretical challenge is to figure out what kind of sparsity assumptions (in terms of both strength and location) can be simultaneously imposed on the multiple components to make the model setup meaningful in practice. The interplay of the multiple sparsity indices and spike indices would make the theoretical analysis very casewise. An interesting future research challenge is to find a way to organize the many possible cases into a single coherent framework. Simultaneous selection of the multiple thresholding parameters involved is both important and will be challenging as well. Shen and Huang [35] suggested to apply RSPCA to extract the leading sparse PCs in a sequential manner, to allow different amounts of sparsity for each PC, which has been shown to have practical advantages.

We also hope to extend our theorems to more general distributions. However, this will be challenging as sparse PCA methods may not work in some extreme non-Gaussian cases, as illustrated in the following example.

**Example 6.1.** Let $\alpha \in (0, 1)$ and $X = (x_1, \ldots, x_d)^T$, where $\{x_i, i = 1, \ldots, d\}$ are independent discrete random variables distributed as follows:

$$x_1 = \begin{cases} d^{\frac{\alpha}{2}}, & \text{with probability } \frac{1}{2}, \\ -d^{\frac{\alpha}{2}}, & \text{with probability } \frac{1}{2}; \end{cases}$$

$$x_i = \begin{cases} d^{\frac{\alpha+1}{4}}, & \text{with probability } d^{-\frac{\alpha+1}{2}}, \\ -d^{\frac{\alpha+1}{4}}, & \text{with probability } d^{-\frac{\alpha+1}{2}}, \quad \text{for } i = 2 \ldots, d \\ 0, & \text{with probability } 1 - 2d^{-\frac{\alpha+1}{2}}. \end{cases}$$

Then $X$ has mean 0 and variance–covariance matrix as

$$\Sigma_d = d^{\alpha} u_1 u_1^T + \sum_{k=2}^{d} u_k u_k^T, \quad \text{with } u_i = (0, \ldots, 0, \overbrace{1}^{i}, 0, \ldots, 0)^T.$$

Suppose that we only have sample size $n = 1$, i.e. $X_1 = (x_{i,1}, \ldots, x_{d,1})^T$, then the first empirical eigenvector

$$\hat{u}_1 = \left(\hat{u}_{1,1}, \ldots, \hat{u}_{d,1}\right)^T = \frac{1}{\sqrt{\sum_{i=1}^{d} x_{i,1}^2}} (x_{i,1}, \ldots, x_{i,d})^T.$$

Under this condition, we can prove that $\mathrm{P}\left(\mathrm{argmax}_i |\hat{u}_{i,d}| = 1\right)$ goes to zero as $d \to \infty$. This suggests that the absolute value of the first entry of $\hat{u}_1$ cannot be greater than the others with probability 1, so we cannot always threshold out the right entries which would result in the failure of sparse PCA. We perform simulation studies to numerically demonstrate this point, which can be found in our online supplement [36].

## 7. Proofs

### 7.1. Proofs of Theorems 2.1–2.3

To prove these theorems, we need the dependent extreme value results from Leadbetter, Lindgren and Rootzen [22], in particular their Lemma 6.1.1 and Theorem 6.1.3.

An immediate consequence of those results is the following proposition.

**Proposition 7.1.** *Suppose that the standard normal sequence $\{\xi_i, i = 1, \ldots, \lfloor d^{\beta} \rfloor\}$ satisfies the mixing condition (2.6). Let the positive constants $\{c_i\}$ be such that $\sum_{i=1}^{\lfloor d^{\beta} \rfloor} (1 - \Phi(c_i))$ is bounded and such that $C_{\lfloor d^{\beta} \rfloor} = \min_{1 \le i \le \lfloor d^{\beta} \rfloor} c_i \ge c(\log(\lfloor d^{\beta} \rfloor))^{\frac{1}{2}}$ for some $c > 0$.*

*Then the following holds:*

$$P\left[\bigcap_{i=1}^{\lfloor d^\beta \rfloor}\{\xi_i \le c_i\}\right] - \prod_{i=1}^{\lfloor d^\beta \rfloor}\Phi(c_i) \longrightarrow 0, \quad \text{as } d \to \infty,$$

*where $\Phi$ is the standard normal distribution function. Furthermore, if for some $\jmath \ge 0$, we have*

$$\sum_{i=1}^{\lfloor d^\beta \rfloor}(1-\Phi(c_i)) \longrightarrow \jmath, \quad \text{as } d \to \infty,$$

*then*

$$P\left[\bigcap_{i=1}^{\lfloor d^\beta \rfloor}\{\xi_i \le c_i\}\right] \longrightarrow e^{-\jmath}, \quad \text{as } d \to \infty.$$

Proposition 7.1 is used to control the right side of (2.5) through the following lemma, the proof of which can be found in our online supplement [37].

**Lemma 7.1.** *Suppose that $\xi_i \sim N(0, \delta_{i,i})$ satisfies the mixing condition (2.6), where $\delta_{ij}$ is the covariance of the normal sequence $\{\xi_i\}$, $i, j = 1, \ldots, \lfloor d^\beta \rfloor$. If $C_{\lfloor d^\beta \rfloor} \ge (\log(\lfloor d^\beta \rfloor))^\delta \max_{1\le i\le \lfloor d^\beta \rfloor} \delta_{ii}^{\frac{1}{2}}$, where $\delta \in (\frac{1}{2}, \infty)$, then*

$$C_{\lfloor d^\beta \rfloor}^{-1} \max_{1\le i\le \lfloor d^\beta \rfloor} |\xi_i| \xrightarrow{p} 0, \quad \text{as } d \to \infty.$$

We now begin the proofs of the theorems. Denote $\tilde{X}_i = (x_{i,1}, \ldots, x_{i,n})^T$, $\tilde{Z}_i = (z_{i,1}, \ldots, z_{i,n})^T$, $\tilde{W}_i = (w_{i,1}, \ldots, w_{i,n})^T$ and $\tilde{H}_i = (h_{i,1}, \ldots, h_{i,n})$, $i = 1, \ldots, d$.

Note that

$$|\langle \hat{u}_1^{\mathrm{ST}}, u_1\rangle| = \frac{\left|\sum_{i=1}^{\lfloor d^\beta \rfloor} \breve{u}_{i,1}u_{i,1}\right|}{\sqrt{\sum_{i=1}^{d}(\breve{u}_{i,1})^2}} = \frac{\lambda_1^{-\frac{1}{2}}\left|\sum_{i=1}^{\lfloor d^\beta \rfloor} \breve{u}_{i,1}u_{i,1}\right|}{\lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^{d}(\breve{u}_{i,1})^2}}. \tag{7.1}$$

Below we need to bound the denominator and the numerator of (7.1).

We start with the numerator. Since $\tilde{X}_i = u_{i,1}\tilde{Z}_1 + \tilde{H}_i$, $i = 1, \ldots, d$, it follows that $\tilde{v}_1^T\tilde{X}_i = u_{i,1}\tilde{v}_1^T\tilde{Z}_1 + \tilde{v}_1^T\tilde{H}_i$, which yields

$$\begin{aligned}\breve{u}_{i,1} &= u_{i,1}\tilde{v}_1^T\tilde{Z}_1 1_{\{|\tilde{v}_1^T\tilde{X}_i|>\zeta\}} + \tilde{v}_1^T\tilde{H}_i 1_{\{|\tilde{v}_1^T\tilde{X}_i|>\zeta\}}\\ &= u_{i,1}\tilde{v}_1^T\tilde{Z}_1 + u_{i,1}\tilde{v}_1^T\tilde{Z}_1 1_{\{|\tilde{v}_1^T\tilde{X}_i|\le\zeta\}} + \tilde{v}_1^T\tilde{H}_i 1_{\{|\tilde{v}_1^T\tilde{X}_i|>\zeta\}},\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^{\lfloor d^\beta \rfloor} \breve{u}_{i,1}u_{i,1} &= \sum_{i=1}^{\lfloor d^\beta \rfloor} u_{i,1}^2\tilde{v}_1^T\tilde{Z}_1 1_{\{|\tilde{v}_1^T\tilde{X}_i|>\zeta\}} + \sum_{i=1}^{\lfloor d^\beta \rfloor} u_{i,1}\tilde{v}_1^T\tilde{H}_i 1_{\{|\tilde{v}_1^T\tilde{X}_i|>\zeta\}}\\ &= \tilde{v}_1^T\tilde{Z}_1 + \sum_{i=1}^{\lfloor d^\beta \rfloor} u_{i,1}^2\tilde{v}_1^T\tilde{Z}_1 1_{\{|\tilde{v}_1^T\tilde{X}_i|\le\zeta\}} + \sum_{i=1}^{\lfloor d^\beta \rfloor} u_{i,1}\tilde{v}_1^T\tilde{H}_i 1_{\{|\tilde{v}_1^T\tilde{X}_i|>\zeta\}}.\end{aligned}$$

It follows that

$$\begin{aligned}\lambda_1^{-\frac{1}{2}}\left|\sum_{i=1}^{\lfloor d^\beta \rfloor} \breve{u}_{i,1}u_{i,1}\right| &\le \lambda_1^{-\frac{1}{2}}\sum_{i=1}^{\lfloor d^\beta \rfloor} u_{i,1}^2|\tilde{v}_1^T\tilde{Z}_1| + \lambda_1^{-\frac{1}{2}}\sum_{i=1}^{\lfloor d^\beta \rfloor}|u_{i,1}\tilde{v}_1^T\tilde{H}_i|\\ &= |\tilde{v}_1^T\tilde{W}_1| + \sum_{i=1}^{\lfloor d^\beta \rfloor}\sum_{j=1}^{n}\lambda_1^{-\frac{1}{2}}|u_{i,1}h_{i,j}|,\end{aligned} \tag{7.2}$$

and

$$\lambda_1^{-\frac{1}{2}}\left|\sum_{i=1}^{\lfloor d^\beta\rfloor}\breve{u}_{i,1}u_{i,1}\right| \geq \lambda_1^{-\frac{1}{2}}|\tilde{v}_1^T\tilde{Z}_1| - \lambda_1^{-\frac{1}{2}}\sum_{i=1}^{\lfloor d^\beta\rfloor}u_{i,1}^2|\tilde{v}_1^T\tilde{Z}_1|1_{\{|\tilde{v}_1^T\tilde{X}_i|\leq\zeta\}} - \lambda_1^{-\frac{1}{2}}\sum_{i=1}^{\lfloor d^\beta\rfloor}|u_{i,1}\tilde{v}_1^T\tilde{H}_i|$$

$$\geq |\tilde{v}_1^T\tilde{W}_1| - |\tilde{v}_1^T\tilde{W}_1|\sum_{i=1}^{\lfloor d^\beta\rfloor}u_{i,1}^2 1_{\{|\tilde{v}_1^T\tilde{X}_i|\leq\zeta\}} - \sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}\lambda_1^{-\frac{1}{2}}|u_{i,1}h_{i,j}|. \tag{7.3}$$

Next we will show that

$$\sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}\lambda_1^{-\frac{1}{2}}|u_{i,1}h_{i,j}| = o_p\left(d^{-\frac{\kappa}{2}}\right), \quad \text{where } \kappa\in[0,\alpha-\eta-\theta). \tag{7.4}$$

Since $H_j=(h_{1,j},\ldots,h_{d,j})^T=\sum_{k=2}^d z_{k,j}u_k$, $j=1,\ldots,n$, it follows that $h_{i,j}=\sum_{k=2}^d u_{i,k}z_{k,j}=\sum_{k=2}^d u_{i,k}\lambda_k^{\frac{1}{2}}w_{k,j}\sim \mathrm{N}(0,\sigma_{i,j}^2)$, where $\sigma_{i,j}^2\leq\lambda_2$, $i=1,\ldots,\lfloor d^\beta\rfloor$, $j=1,\ldots,n$. Thus, for fix $\tau$

$$P\left[\sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}d^{\frac{\kappa}{2}}\lambda_1^{-\frac{1}{2}}|u_{i,1}h_{i,j}|\geq\tau\right] \leq P\left[\bigcup_{i=1}^{\lfloor d^\beta\rfloor}\left\{\sum_{j=1}^{n}|u_{i,1}h_{i,j}|\geq d^{-\frac{\kappa}{2}}\lambda_1^{\frac{1}{2}}\tau u_{i,1}^2\right\}\right]$$

$$\leq \sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}P\left[|h_{i,j}|\geq n^{-1}d^{-\frac{\kappa}{2}}\lambda_1^{\frac{1}{2}}\tau|u_{i,1}|\right] \leq \sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}P\left[|h_{i,j}\sigma_{i,j}^{-1}|\geq cd^{\frac{\alpha-\eta-\theta-\kappa}{2}}\right]$$

$$= 2n\lfloor d^\beta\rfloor\int_{cd^{(\alpha-\eta-\theta-\kappa)/2}}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}dx\longrightarrow 0, \quad \text{as } d\to\infty,$$

where $c$ is constant. Since

$$P\left[\sum_{i=1}^{\lfloor d^\beta\rfloor}d^{\frac{\kappa}{2}}\lambda_1^{-1}\left(\sum_{j=1}^{n}|h_{i,j}|\right)^2\geq\tau\right] \leq P\left[\bigcup_{i=1}^{\lfloor d^\beta\rfloor}\left\{\sum_{j=1}^{n}|h_{i,j}|\geq d^{-\frac{\kappa+\beta}{2}}\lambda_1^{\frac{1}{2}}\tau^{\frac{1}{2}}\right\}\right]$$

$$\leq \sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}P\left[|h_{i,j}|\geq n^{-1}d^{-\frac{\kappa+\beta}{2}}\lambda_1^{\frac{1}{2}}\tau^{\frac{1}{2}}\right] \leq \sum_{i=1}^{\lfloor d^\beta\rfloor}\sum_{j=1}^{n}P\left[|h_{i,j}\sigma_{i,j}^{-1}|\geq cd^{\frac{\alpha-\eta-\theta-\kappa}{2}}\right]$$

$$= 2n\lfloor d^\beta\rfloor\int_{cd^{(\alpha-\eta-\theta-\kappa)/2}}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}dx\longrightarrow 0, \quad \text{as } d\to\infty,$$

it follows that

$$\sum_{i=1}^{\lfloor d^\beta\rfloor}\lambda_1^{-1}\left(\sum_{j=1}^{n}|h_{i,j}|\right)^2 = o_p\left(d^{\frac{\kappa}{2}}\right). \tag{7.5}$$

Similarly, since

$$P\left[\sum_{i=\lfloor d^\beta\rfloor+1}^{d}\sum_{j=1}^{n}d^{\frac{\kappa}{2}}\lambda_1^{-\frac{1}{2}}|h_{i,j}|1_{\left\{\sum_{j=1}^{n}|h_{i,j}|>\zeta\right\}}\right] \leq P\left[\bigcup_{i=\lfloor d^\beta\rfloor+1}^{d}\left\{\sum_{j=1}^{n}|h_{i,j}|>\zeta\right\}\right]$$

$$= 2nd\int_{c\log^\delta(d)}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}dx\longrightarrow 0, \quad \text{as } d\to\infty,$$

it follows that

$$\sum_{i=\lfloor d^\beta\rfloor+1}^{d}\sum_{j=1}^{n}\lambda_1^{-\frac{1}{2}}|h_{i,j}|1_{\left\{\sum_{j=1}^{n}|h_{i,j}|>\zeta\right\}} = o_p\left(d^{-\frac{\kappa}{2}}\right). \tag{7.6}$$

Finally, we want to show that

$$\sum_{i=1}^{\lfloor d^\beta\rfloor}u_{i,1}^2 1_{\{|\tilde{v}_1^T\tilde{X}_i|\leq\zeta\}} = o_p\left(d^{-\frac{\kappa'}{2}}\right), \tag{7.7}$$

where $\kappa'$ satisfies that $d^{\frac{\kappa'+\eta-\alpha}{2}}\zeta = o(1)$. From the condition (a) of Theorem 2.2, we have $\overline{\lim}_{d\to\infty}\frac{\lambda_1}{\sum_{i=2}^d\lambda_i} = c < \infty$. Since we can always find a subsequence of $\{\frac{\lambda_1}{\sum_{i=2}^d\lambda_i}\}$ and make it convergent to a nonnegative constant, for simplicity, we just assume that $\lim_{d\to\infty}\frac{\lambda_1}{\sum_{i=2}^d\lambda_i} = c$. If $c = 0$, then the spike index $\alpha < 1$, and Jung and Marron [20] showed that

$$c_d^{-1}S_d \xrightarrow{p} I_n, \quad \text{as } d \to \infty,$$

where $c_d = n^{-1}\sum_{i=1}^d\lambda_i$. Since the eigenvector $\tilde{v}_1^T$ of $c_d^{-1}S_d$ can be chosen so that they are continuous according to Acker [1], it follows that $\tilde{v}_1^T \Rightarrow v_1$, as $d \to \infty$, where $\Rightarrow$ denotes the convergence in distribution and $v_1$ is the first eigenvector of $n$-dimensional identity matrix. If $c > 0$, then the spike index $\alpha = 1$ and Jung, Sen and Marron [21] showed that $\tilde{v}_1^T \Rightarrow \frac{\tilde{W}_1}{\|\tilde{W}_1\|}$, as $d \to \infty$. Therefore, we have

$$|\tilde{v}_1^T\tilde{W}_1| \Rightarrow |v_1^T\tilde{W}_1| \quad \text{or} \quad \|\tilde{W}_1\|, \quad \text{as } d \to \infty. \tag{7.8}$$

Since $d^{\frac{\kappa'+\eta-\alpha}{2}}\zeta = o(1)$, $d^{\frac{\kappa'+\eta-\alpha}{2}}\sum_{j=1}^n\max_{1\le i\le\lfloor d^\beta\rfloor}|h_{i,j}| = o_p(1)$, and

$$\begin{aligned}
\sum_{i=1}^{\lfloor d^\beta\rfloor} d^{\frac{\kappa'}{2}}u_{i,1}^2 1_{\left\{|\tilde{v}_1^T\tilde{X}_i|\le\zeta\right\}} &\le \sum_{i=1}^{\lfloor d^\beta\rfloor} d^{\frac{\kappa'}{2}}u_{i,1}^2 1_{\left\{|u_{i,1}\tilde{v}_1^T\tilde{Z}_1|\le|\tilde{v}_1^T\tilde{H}_i|+\zeta\right\}} \\
&\le \sum_{i=1}^{\lfloor d^\beta\rfloor} d^{\frac{\kappa'}{2}}u_{i,1}^2 1_{\left\{|\tilde{v}_1^T\tilde{W}_1|\le\lambda_1^{-\frac{1}{2}}\max_{1\le i\le\lfloor d^\beta\rfloor}|u_{i,1}|^{-1}\left(\sum_{j=1}^n\max_{1\le i\le\lfloor d^\beta\rfloor}|h_{i,j}|+\zeta\right)\right\}} \\
&\le \frac{cd^{\frac{\kappa'+\eta-\alpha}{2}}\sum_{j=1}^n\max_{1\le i\le\lfloor d^\beta\rfloor}|h_{i,j}| + cd^{\frac{\kappa'+\eta-\alpha}{2}}\zeta}{|\tilde{v}_1^T\tilde{W}_1|},
\end{aligned}$$

where $c$ is a constant, it follows that (7.7) is established.

Then, from (7.2)–(7.4) and (7.7), we obtain the following result about the numerator

$$\lambda_1^{-\frac{1}{2}}\left|\sum_{i=1}^{\lfloor d^\beta\rfloor}\breve{u}_{i,1}u_{i,1}\right| = |\tilde{v}_1^T\tilde{W}_1| + o_p\left(d^{-\frac{\min\{\kappa,\kappa'\}}{2}}\right). \tag{7.9}$$

Similarly for the denominator, we have

$$\begin{aligned}
\lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^d\breve{u}_{i,1}^2} &\le \lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}\breve{u}_{i,1}^2} + \lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=\lfloor d^\beta\rfloor+1}^d\breve{u}_{i,1}^2} \\
&\le \lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}\left(u_{i,1}\tilde{v}_1^T\tilde{Z}_1\right)^2} + \lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}(\tilde{v}_1^T\tilde{H}_i)^2} + \lambda_1^{-\frac{1}{2}}\sum_{i=\lfloor d^\beta\rfloor+1}^d|\breve{u}_{i,1}| \\
&= |\tilde{v}_1^T\tilde{W}_1| + \sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}\lambda_1^{-1}\left(\sum_{j=1}^n|h_{i,j}|\right)^2} + \sum_{i=\lfloor d^\beta\rfloor+1}^d\sum_{j=1}^n\lambda_1^{-\frac{1}{2}}|h_{i,j}|1_{\left\{\sum_{j=1}^n|h_{i,j}|>\zeta\right\}},
\end{aligned} \tag{7.10}$$

and

$$\begin{aligned}
\lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^d\breve{u}_{i,1}^2} &\ge \lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}\breve{u}_{i,1}^2} \\
&\ge |\tilde{v}_1^T\tilde{W}_1| - |\tilde{v}_1^T\tilde{W}_1|\sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}u_{i,1}^2 1_{\{|\tilde{v}_1^T\tilde{X}_i|\le\zeta\}}} - \sqrt{\sum_{i=1}^{\lfloor d^\beta\rfloor}\lambda_1^{-1}\left(\sum_{j=1}^n|h_{i,j}|\right)^2}.
\end{aligned} \tag{7.11}$$

Combining (7.10), (7.11), (7.5)–(7.7), we have

$$\lambda_1^{-\frac{1}{2}}\sqrt{\sum_{i=1}^d\breve{u}_{i,1}^2} = |\tilde{v}_1^T\tilde{W}_1| + o_p\left(d^{-\frac{\min\{\kappa,\kappa'\}}{2}}\right). \tag{7.12}$$

Furthermore, (7.1), (7.8), (7.9) and (7.12) suggest that

$$|\langle \hat{u}_1^{\mathrm{ST}}, u_1\rangle| = \frac{|\tilde{v}_1^T\tilde{W}_1| + o_p\left(d^{-\frac{\min\{\kappa,\kappa'\}}{2}}\right)}{|\tilde{v}_1^T\tilde{W}_1| + o_p\left(d^{-\frac{\min\{\kappa,\kappa'\}}{2}}\right)} = 1 + o_p\left(d^{-\frac{\min\{\kappa,\kappa'\}}{2}}\right),$$

which means that $\hat{u}_1^{\mathrm{ST}}$ is consistent with $u_1$ with convergence rate $d^{-\frac{\min\{\kappa,\kappa'\}}{2}}$. This concludes the proof for Theorem 2.2. Note that Theorem 2.1 is a special case of Theorem 2.2.

In addition, note that $d^{\frac{\kappa'+\eta-\alpha}{2}}\zeta = o(1)$. If $\zeta = o(d^{\frac{\alpha-\eta-\kappa}{2}})$, then we can take $\kappa' = \kappa$. Then $\hat{u}_1^{\mathrm{ST}}$ is consistent with $u_1$ with convergence rate $d^{\frac{\kappa}{2}}$. This finishes the proof of Theorem 2.3.

### 7.2. Proofs of Theorems 3.1–3.3

The proofs of Theorems 3.1 and 3.3 are modifications of the proofs of Theorems 2.2 and 2.3. These are provided as supplementary materials, available online at [37]. The proof of Theorem 3.2 is also given in the supplement.

### 7.3. Proof of Theorem 4.1

Since $X_j^* = (I_{\lfloor d^\beta\rfloor}, (0)_{\lfloor d^\beta\rfloor\times(d-\lfloor d^\beta\rfloor)})X_j$, where $I_{\lfloor d^\beta\rfloor}$ denotes the $\lfloor d^\beta\rfloor$-dimensional identity matrix and $(0)_{\lfloor d^\beta\rfloor\times(d-\lfloor d^\beta\rfloor)}$ is the $\lfloor d^\beta\rfloor$-by-$(d-\lfloor d^\beta\rfloor)$ zero matrix, $j = 1,\ldots,n$, it follows that

$$\Sigma^*_{\lfloor d^\beta\rfloor} = (I_{\lfloor d^\beta\rfloor}, (0)_{\lfloor d^\beta\rfloor\times(d-\lfloor d^\beta\rfloor)})\Sigma_d(I_{\lfloor d^\beta\rfloor}, (0)_{\lfloor d^\beta\rfloor\times(d-\lfloor d^\beta\rfloor)})^T,$$

which yields

$$\lambda_1^* = \lambda_1, \lambda_2 \geq \lambda_i^* \geq \lambda_d, \quad i = 2,\ldots,\lfloor d^\beta\rfloor. \tag{7.13}$$

Therefore,

$$\frac{\sum_{i=2}^{\lfloor d^\beta\rfloor}\lambda_i^{*2}}{\left(\sum_{i=2}^{\lfloor d^\beta\rfloor}\lambda_i^*\right)^2} \leq \frac{\lfloor d^\beta\rfloor\lambda_2^2}{(\lfloor d^\beta\rfloor)^2\lambda_d^2} = \frac{O(d^\beta)O(d^{2\theta})}{O(d^{2\beta})} = o(1), \tag{7.14}$$

and

$$\frac{\lambda_1^*}{\sum_{i=2}^{\lfloor d^\beta\rfloor}\lambda_i^*} \leq \frac{\lambda_1}{\lfloor d^\beta\rfloor\lambda_d} = \frac{O(d^\alpha)}{O(d^\beta)} = o(1). \tag{7.15}$$

If we rescale $\lambda_i^*, i = 1,\ldots,\lfloor d^\beta\rfloor$, (7.14) satisfies the $\varepsilon_2$ assumption of Jung and Marron [20] and (7.15) satisfies the assumption $\lambda_1 = O(d^\alpha)$ and $\sum_{i=2}^d \lambda_i = O(d)$, where $\alpha < 1$. For this case, Jung and Marron [20] have shown that $\hat{u}_1^*$ is strongly inconsistent with $u_1^*$. This means that the oracle estimator $\hat{u}_1^{\mathrm{OR}}$ is strongly inconsistent with $u_1$.

For the boundary case of $\alpha = \beta$, WLOG, we assume that $\lambda_1 = cd^\alpha$ and $\lambda_i = c_\lambda$, $i = 2,\ldots,d$. Then (7.13) becomes

$$\lambda_1^* = \lambda_1 = cd^\alpha, \qquad \lambda_i^* = c_\lambda, \quad j = 2,\ldots,\lfloor d^\beta\rfloor.$$

Note that the dimension of $X_j^*$ is $\lfloor d^\beta\rfloor = \lfloor d^\alpha\rfloor$. If we treat the effective dimension $\lfloor d^\alpha\rfloor$ as $d$ in Jung, Sen and Marron [21], it then follows that

$$|\langle \hat{u}_1^*, u_1^*\rangle| \Rightarrow \left(1 + \frac{c_\lambda}{c\chi_n^2}\right)^{-\frac{1}{2}}.$$

The above, together with (4.1) and (4.2), leads to the inconsistency of $\hat{u}_1^{\mathrm{OR}}$ (4.3).

## Acknowledgments

*Supplement* A: detailed technical proofs of Theorems 3.1–3.3 are provided, available online at [37].

*Supplement* B: additional simulation results for (1) the twenty spike and sparsity index pairs (Fig. 1), under a range of dimension $d$; (2) non-Gaussian distributions, available online at [36].

## Appendix. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jmva.2012.10.007.

## References


[1] A.F. Acker, Absolute continuity of eigenvectors of time-varying operators, Proceedings of the American Mathematical Society 42 (1) (1974) 198–201.
[2] J. Ahn, M.H. Lee, Y.J. Yoon, Clustering high dimension, low sample size data using the maximal data piling distance, Statistica Sinica 22 (2) (2012) 443–464.
[3] J. Ahn, J.S. Marron, K.M. Muller, Y.Y. Chi, The high-dimension, low-sample-size geometric representation holds under mild conditions, Biometrika 94 (3) (2007) 760–766.
[4] A.A. Amini, M.J. Wainwright, High-dimensional analysis of semidefinite relaxations for sparse principal components, The Annals of Statistics 37 (5B) (2009) 2877–2921.
[5] P.J. Bickel, E. Levina, Covariance regularization by thresholding, The Annals of Statistics 36 (6) (2008) 2577–2604.
[6] P.J. Bickel, E. Levina, Regularized estimation of large covariance matrices, The Annals of Statistics 36 (1) (2008) 199–227.
[7] P.J. Bickel, Y. Ritov, A.B. Tsybakov, Simultaneous analysis of Lasso and Dantzig selector, The Annals of Statistics 37 (4) (2009) 1705–1732.
[8] P. Bühlmann, Boosting for high-dimensional linear models, The Annals of Statistics 34 (2) (2006) 559–583.
[9] F. Bunea, A.B. Tsybakov, M.H. Wegkamp, A. Barbu, Spades and mixture models, The Annals of Statistics 38 (4) (2010) 2525–2558.
[10] E. Candes, T. Tao, The Dantzig selector: statistical estimation when $p$ is much larger than $n$, The Annals of Statistics 35 (6) (2007) 2313–2351.
[11] G. Casella, J.T. Hwang, Limit expressions for the risk of James–Stein estimators, Canadian Journal of Statistics 10 (4) (1982) 305–309.
[12] A. d'Aspremont, L. El Ghaoui, M.I. Jordan, G.R.G. Lanckriet, A direct formulation for sparse PCA using semidefinite programming, SIAM Review 49 (3) (2007) 434–448.
[13] D.L. Donoho, J.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (3) (1994) 425–455.
[14] C. Eckart, G. Young, The approximation of one matrix by another of lower rank, Psychometrika 1 (3) (1936) 211–218.
[15] N. El Karoui, Operator norm consistent estimation of large-dimensional sparse covariance matrices, The Annals of Statistics 36 (6) (2008) 2717–2756.
[16] J. Fan, R. Li, Variable selection via nonconcave penalized likelihood and its oracle properties, Journal of the American Statistical Association 96 (456) (2001) 1348–1360.
[17] P. Hall, J.S. Marron, A. Neeman, Geometric representation of high dimension, low sample size data, Journal of the Royal Statistical Society. Series B 67 (3) (2005) 427–444.
[18] I.M. Johnstone, On the distribution of the largest eigenvalue in principal components analysis, The Annals of Statistics 29 (2) (2001) 295–327.
[19] I.M. Johnstone, A.Y. Lu, On consistency and sparsity for principal components analysis in high dimensions, Journal of the American Statistical Association 104 (486) (2009) 682–693.
[20] S. Jung, J.S. Marron, PCA consistency in high dimension. Low sample size context, The Annals of Statistics 37 (6B) (2009) 4104–4130.
[21] S. Jung, A. Sen, J.S. Marron, Boundary behavior in high dimension, low sample size asymptotics of PCA, Technical Report, UNC-CH, 2012.
[22] M.R. Leadbetter, G. Lindgren, H. Rootzén, Extremes and Related Properties of Random Sequences and Processes, Springer-Verlag, New York, 1983.
[23] Y.K. Lee, E.R. Lee, B.U. Park, Principal component analysis in very high-dimensional spaces, Statistica Sinica 22 (2012) 933–956.
[24] M. Lee, H. Shen, J.Z. Huang, J.S. Marron, Biclustering via sparse singular value decomposition, Biometrics 66 (4) (2010) 1087–1095.
[25] S. Lee, F. Zou, F.A. Wright, Convergence and prediction of principal component scores in high-dimensional settings, Annals of Statistics 38 (6) (2010) 3605–3629.
[26] C. Leng, H. Wang, On general adaptive sparse principal component analysis, Journal of Computational and Graphical Statistics 18 (1) (2009) 201–215.
[27] Z. Ma, Sparse principal component analysis and iterative thresholding, Technical Report, UPenn, 2012.
[28] L. Meier, S. van de Geer, P. Bühlmann, High-dimensional additive modeling, The Annals of Statistics 37 (6B) (2009) 3779–3821.
[29] N. Meinshausen, P. Bühlmann, High-dimensional graphs and variable selection with the lasso, The Annals of Statistics 34 (3) (2006) 1436–1462.
[30] G. Obozinski, M.J. Wainwright, M.I. Jordan, Support union recovery in high-dimensional multivariate regression, The Annals of Statistics 39 (1) (2011) 1–47.
[31] D. Omidiran, M.J. Wainwright, High-dimensional variable selection with sparse random projections: measurement sparsity and statistical efficiency, The Journal of Machine Learning Research 99 (2010) 2361–2386.
[32] D. Paul, Asymptotics of sample eigenstructure for a large dimensional spiked covariance model, Statistica Sinica 17 (4) (2007) 1617–1642.
[33] D. Paul, I. Johnstone, Augmented sparse principal component analysis for high-dimensional data, Technical Report, UC Davis, 2007.
[34] G. Schwarz, Estimating the dimension of a model, The Annals of Statistics 6 (2) (1978) 461–464.
[35] H. Shen, J.Z. Huang, Sparse principal component analysis via regularized low rank matrix approximation, Journal of Multivariate Analysis 99 (6) (2008) 1015–1034.
[36] D. Shen, H. Shen, J.S. Marron, Consistency of sparse PCA in high dimension, low sample size contexts: additional simulation results, 2012. Available online at: http://www.unc.edu/~dshen/RSPCA/SupplementB.html.
[37] D. Shen, H. Shen, J.S. Marron, Consistency of sparse PCA in high dimension, low sample size contexts: supplement materials, 2012. Available online at: http://www.unc.edu/~dshen/RSPCA/SupplementA.pdf.
[38] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society. Series B 58 (1996) 267–288.
[39] S.A. van de Geer, High-dimensional generalized linear models and the lasso, The Annals of Statistics 36 (2) (2008) 614–645.
[40] D.M. Witten, R. Tibshirani, T. Hastie, A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis, Biostatistics 10 (3) (2009) 515–534.
[41] K. Yata, M. Aoshima, Effective PCA for high-dimension, low-sample-size data with noise reduction via geometric representations, Journal of Multivariate Analysis 105 (1) (2012) 193–215.
[42] H. Zou, T. Hastie, R. Tibshirani, Sparse principal component analysis, Journal of Computational and Graphical Statistics 15 (2) (2006) 265–286.
[43] H. Zou, T. Hastie, R. Tibshirani, On the "degrees of freedom" of the lasso, The Annals of Statistics 35 (5) (2007) 2173–2192.