Author Posting. (c) 'Copyright Holder', 2009.

This is the author's version of the work. It is posted here by permission of 'Copyright Holder'

for personal use, not for redistribution.

The definitive version was published in Communications in Statistics - Theory and Methods,

Volume 38 Issue 16, January 2009.

doi:10.1080/03610910902936083 (<http://dx.doi.org/10.1080/03610910902936083>)

PCA Consistency for Non-Gaussian Data in High Dimension, Low Sample Size Context

Kazuyoshi Yata

Graduate School of Pure and Applied Sciences
University of Tsukuba
Ibaraki 305-8571, Japan

Makoto Aoshima

Institute of Mathematics
University of Tsukuba
Ibaraki 305-8571, Japan

Abstract In this paper, we investigate both sample eigenvalues and Principal Component (PC) directions along with PC scores when the dimension d and the sample size n both grow to infinity in such a way that n is much lower than d . We consider general settings that include the case when the eigenvalues are all in the range of sphericity. We do not assume either the normality or a ρ -mixing condition. We attempt finding a difference among the eigenvalues by choosing n with a suitable order of d . We give the consistency properties for both the sample eigenvalues and the PC directions along with the PC scores. We also show that the sample eigenvalue has a Gaussian limiting distribution when the population counterpart is of multiplicity one.

Key Words: Consistency; Dual covariance matrix; Eigenvalue distribution; HDLSS; Large p small n ; Principal component analysis; Random matrix theory; Sample size.

1. Introduction

High Dimension, Low Sample Size (HDLSS) data are emerging in various areas of modern science such as genetic microarrays, medical imaging, text recognition, finance, chemometrics, and so on. The asymptotic studies of this type of data are becoming increasingly relevant. Principal Component Analysis (PCA) is an important tool of dimension reduction especially when the dimension is very high. PCA visualizes important structures in the data by approximating the data with the first few principal components. Thus it is important to

study the asymptotic behaviors of both eigenvalues and Principal Component (PC) directions along with PC scores. In this paper, we assume $d > n$ and that the dimension d and the sample size n both grow to infinity in such a way that n is much lower than d .

In recent years, substantial work has been done on the asymptotic behavior of eigenvalues of the sample covariance matrix in the limit as $d \rightarrow \infty$, see Johnstone (2001), Baik et al. (2005) and Paul (2007) for Gaussian assumptions, and Baik and Silverstein (2006) for non-Gaussian but i.i.d. assumptions when d and n increase at the same rate, i.e. $n/d \rightarrow c > 0$. On the other hand, Johnstone and Lu (2008) have shown that the estimate of the leading principal component vector is consistent if and only if $d(n)/n \rightarrow 0$. Many of these focus on the spiked covariance model introduced by Johnstone. The spiked covariance model assumes that the first few eigenvalues of the population covariance matrix are greater than 1 and the rest are set to be 1 for all d . The HDLSS asymptotics, where only $d \rightarrow \infty$ while n is fixed, have been studied by Hall et al. (2005) and Ahn et al. (2007). They explored conditions to give the geometric representation of HDLSS data. Jung and Marron (2008) have investigated conditions for consistency and strong inconsistency of eigenvectors of the sample covariance matrix in the HDLSS data situations. For a fixed number κ , they assumed the first κ eigenvalues are much larger than the others. The rest of eigenvalues are assumed to satisfy a condition related to the sphericity. They showed that when $\kappa = 1$, the first sample eigenvector is consistent and the others are strongly inconsistent.

However, the HDLSS asymptotics usually regulate either the population distribution by the normality or the dependency of the random variables in the sphered data matrix by the ρ -mixing condition (see, for example, Sec.5.1 on p.440 in Hall et al. 2005). Those assumptions are somewhat too strict and have some obvious shortcomings. It is common to have severe collinearity among variables and the ρ -mixing condition critically depends on the ordering of variables. For many interesting data types, such as microarray data, there is clear dependence but no natural ordering of the variables. In order to relax those regulations, Yata and Aoshima (2008) have developed the HDLSS asymptotics in more general settings without assuming either the normality or the ρ -mixing condition.

In this paper, we study PCA for HDLSS data in more general settings that include the case when the eigenvalues are all in the range of sphericity. We give the consistency properties of both the eigenvalues and the PC directions along with the PC scores. We also show that the sample eigenvalue has a Gaussian limiting distribution when the population counterpart is of multiplicity one. We do not impose the normality and the ρ -mixing condition on the main results given in this paper.

2. HDLSS setting

Suppose we have a $d \times n$ data matrix $\mathbf{X}_{(d)} = [\mathbf{x}_{1(d)}, \dots, \mathbf{x}_{n(d)}]$ with $d > n$, where $\mathbf{x}_{j(d)} = (x_{1j(d)}, \dots, x_{dj(d)})^T$, $j = 1, \dots, n$, are independent and identically distributed as a d -dimensional multivariate distribution with mean zero and nonnegative definite covariance matrix Σ_d . The eigen-decomposition of Σ_d is $\Sigma_d = \mathbf{H}_d \Lambda_d \mathbf{H}_d^T$, where Λ_d is a diagonal matrix of eigenvalues $\lambda_{1(d)} \geq \dots \geq \lambda_{d(d)} \geq 0$ and $\mathbf{H}_d = [\mathbf{h}_{1(d)}, \dots, \mathbf{h}_{d(d)}]$ is a matrix of corresponding eigenvectors. Then, $\mathbf{Z}_{(d)} = \Lambda_d^{-1/2} \mathbf{H}_d^T \mathbf{X}_{(d)}$ is a $d \times n$ sphered data matrix from a distribution with the identity covariance matrix. Here, we write $\mathbf{Z}_{(d)} = [\mathbf{z}_{1(d)}, \dots, \mathbf{z}_{d(d)}]^T$ and $\mathbf{z}_{i(d)} = (z_{i1(d)}, \dots, z_{in(d)})^T$, $i = 1, \dots, d$. Hereafter, the subscript d will be omitted for the sake of simplicity when it does not cause any confusion. We assume that the fourth moments of each variable in $\mathbf{Z}$ are uniformly bounded and $\|\mathbf{z}_i\| \neq 0$ for $i = 1, \dots, d$, where $\|\cdot\|$ denotes the Euclidean norm. We emphasize that the multivariate distribution assumed here does not have to be Gaussian and the random variables in $\mathbf{Z}$ do not have to be regulated by the ρ -mixing condition.

We consider a general setting as follows:

$$\lambda_i = a_i d^{\alpha_i} \quad (i = 1, \dots, m) \quad \text{and} \quad \lambda_j = c_j \quad (j = m + 1, \dots, d). \quad (1)$$

Here, $a_i (> 0)$, $c_j (\geq 0)$ and $\alpha_i (\alpha_1 \geq \dots \geq \alpha_m > 0)$ are unknown constants preserving the ordering that $\lambda_1 \geq \dots \geq \lambda_d$, and m is an unknown positive integer. The sample covariance matrix is $\mathbf{S} = n^{-1} \mathbf{X} \mathbf{X}^T$. We consider the $n \times n$ dual sample covariance matrix defined by $\mathbf{S}_D = n^{-1} \mathbf{X}^T \mathbf{X}$. Note that $\mathbf{S}_D$ and $\mathbf{S}$ share non-zero eigenvalues and $E\{(n / \sum_{i=1}^d \lambda_i) \mathbf{S}_D\} =$

$\mathbf{I}_n$. Ahn et al. (2007), and Jung and Marron (2008) claimed that when the eigenvalues of Σ are sufficiently diffused in the sense that

$$\sum_{i=1}^d \tilde{\lambda}_i^2 \rightarrow 0 \text{ as } d \rightarrow \infty, \quad \text{where } \tilde{\lambda}_i = \lambda_i / (\sum_{i=1}^d \lambda_i), \quad (2)$$

the sample eigenvalues behave as if they are from a scaled identity covariance matrix. When $\mathbf{X}$ is Gaussian or the components of $\mathbf{Z}$ are ρ -mixing, they showed that it follows that $(n / \sum_{i=1}^d \lambda_i) \mathbf{S}_D \rightarrow \mathbf{I}_n$ in probability as $d \rightarrow \infty$ under (2). That is, the set of sample eigenvectors tends to be an arbitrary choice. When $\mathbf{X}$ is non-Gaussian and non- ρ -mixing, Yata and Aoshima (2008) showed that the eigenvalues are of the same order of d under (2) for a fixed n . Therefore, in both the situations, it is difficult to find a difference among the eigenvalues under (2). We emphasize that the formulation (1), provided that $\alpha_1 < 1$ and $c_d > 0$, includes the case satisfying (2). That is, it is quite difficult to estimate eigenvalues in such a situation. Actually, Jung and Marron (2008) found it strongly inconsistent for estimating PC directions of HDLSS data satisfying (1) and (2). In this paper, we will attempt finding a difference among the eigenvalues by choosing the sample size n with a suitable order of d . The estimation, given in this paper, enjoys consistency properties in the situations where Jung and Marron (2008) found it strongly inconsistent for both the eigenvalues and the PC directions.

3. Consistency and asymptotic normality of eigenvalues

The study of asymptotic behavior of the sample eigenvalues is an important part in the study of PC directions, and also could be of independent interest. Jung and Marron (2008) gave an excellent overview for $\mathbf{S}_D$ with the help of the ρ -mixing condition. Their findings are summarized: The large sample eigenvalues with power $\alpha_i > 1$ increase at the same speed as their population counterpart and the relatively small eigenvalues for the other i tend to be of order of d as d tends to infinity. Refer to Jung and Marron (2008) for the details.

Let $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_n \geq 0$ be the eigenvalues of $\mathbf{S}_D$. Let us write the eigen-decomposition

of $\mathbf{S}_D$ as $\mathbf{S}_D = \sum_{i=1}^n \hat{\lambda}_i \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^T$. For the eigenvalues of $\mathbf{S}_D$, we obtain the following theorems without assuming either the ρ -mixing condition or the normality.

Theorem 1. *For $i = 1, \dots, m$, we have that*

$$\frac{\hat{\lambda}_i}{\lambda_i} = 1 + o_p(1) \quad (3)$$

under the conditions:

- (i) $d \rightarrow \infty$ and $n \rightarrow \infty$ for i such that $\alpha_i > 1$,
- (ii) $d \rightarrow \infty$ and $d^{2-2\alpha_i}/n \rightarrow 0$ for i such that $\alpha_i \in (0, 1]$.

Theorem 2. Let $V(z_{ij}^2) = M_i$ ($< \infty$) for $i = 1, \dots, m$. Assume that the first m population eigenvalues are distinct. Then, under the conditions (i)–(ii) in Theorem 1, we have for $i = 1, \dots, m$, that

$$\sqrt{\frac{n}{M_i}} \left(\frac{\hat{\lambda}_i}{\lambda_i} - 1 \right) \Rightarrow N(0, 1), \quad (4)$$

where “ $\Rightarrow$ ” denotes the convergence in distribution and $N(0, 1)$ denotes a random variable distributed as the Standard normal distribution.

Let us consider the situation in which Baik and Siliverstein (2006) investigated the asymptotic behaviors of eigenvalues. They assumed that z_{ij} , $i, j = 1, 2, \dots$, are independent and identically distributed and considered the condition that $d \rightarrow \infty$ and $d/n \rightarrow c > 0$. In view of Theorem 1 in the HDLSS context with $d > n$, one would be interested in the case that $\alpha_i > 1/2$ from the condition (ii). If the independence is assumed in Theorem 1, we can not only relax the convergence condition with respect to n but also extend the range of allowable α thresholds in the HDLSS context to any $\alpha_i > 0$ as in the following

Corollary 1. *Assume further in Theorem 1 that z_{ij} , $i = 1, \dots, d$ ($j = 1, \dots, n$) are independent. Then, for $i = 1, \dots, m$, we have (3) under the conditions:*

- (i) $d \rightarrow \infty$ and $n \rightarrow \infty$ for i such that $\alpha_i > 1$,

(ii) $d \rightarrow \infty$ and $d^{1-\alpha_i}/n \rightarrow 0$ for i such that $\alpha_i \in (0, 1]$.

Remark 1. When $\mathbf{X}$ is Gaussian, we can claim (4) with $M_i = 2$ ($i = 1, \dots, m$) in Theorem 2. Then, the corresponding results have been given by Jung and Marron (2008, Corollary 3) in the case that $d \rightarrow \infty$ and n is fixed when $\alpha_i > 1$, and by Paul (2007, Theorem 3) in the case that $n, d \rightarrow \infty$ and $d/n = o(n^{-1/2})$ when $\alpha_i \rightarrow 0$. When $\alpha_i \rightarrow 0$, Corollary 1 (ii) corresponds to the results given by Baik and Siliverstein (2006, Theorems 1.1–1.3).

4. Consistency of PC directions

In this section, we investigate the sample PC direction vectors. Jung and Marron (2008) considered a covariance structure with several spikes, in which s (≥ 1) eigenvalues are much larger than the others. In order to have consistency of the first s eigenvectors, they required that each of s eigenvalues has a distinct order of magnitude and the sum of the rest is order of d . Let $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \dots, \hat{\mathbf{h}}_d]$ such that $\hat{\mathbf{H}}^T \mathbf{S} \hat{\mathbf{H}} = \hat{\mathbf{\Lambda}}$ and $\hat{\mathbf{\Lambda}} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_d)$. Then, Jung and Marron (2008) concluded under the ρ -mixing condition that the first s sample eigenvectors converge to their population counterparts (*consistency*), while the rest of sample eigenvectors tend to be perpendicular to their population counterparts (*strong inconsistency*). Their result is summarized as follows: For a fixed n , let $d = n + 1, n + 2, \dots$. Assume that

- (a) The components of $\mathbf{Z}$ are ρ -mixing,
- (b) $\alpha_1 > \dots > \alpha_s > 1$ for some s ($\leq m$) in (1).

Then, the first s sample eigenvectors are consistent in the sense that

$$\text{Angle}(\hat{\mathbf{h}}_i, \mathbf{h}_i) \xrightarrow{p} 0 \quad (5)$$

for $i = 1, \dots, s$, as $d \rightarrow \infty$. In addition, suppose the condition that

$$(c) \quad \frac{\sum_{i=s+1}^d \lambda_i^2}{(\sum_{i=s+1}^d \lambda_i)^2} \rightarrow 0 \quad \text{and} \quad \sum_{i=s+1}^d \lambda_i = O(d) \quad \text{as } d \rightarrow \infty.$$

Then, the rest of sample eigenvectors are strongly inconsistent in the sense that

$$\text{Angle}(\hat{\mathbf{h}}_i, \mathbf{h}_i) \xrightarrow{p} \frac{\pi}{2} \text{ for } i = s+1, \dots, n, \text{ as } d \rightarrow \infty.$$

If both d and n tend to infinity, replace the assumption (b) with

(b') The first κ eigenvalues are distinct such as $\lambda_1 > \dots > \lambda_\kappa$ and $\alpha_\kappa > 1$ for some $\kappa (\leq m)$ in (1).

Then, the assertion (5) is claimed for $i = 1, \dots, \kappa$, as $d \rightarrow \infty$ and $n \rightarrow \infty$.

Remark 2. The condition (c) is satisfied for the case that $\alpha_{s+1} < 1$ and $\lambda_d > 0$ in (1).

The following theorem separates PC directions better without assuming either the ρ -mixing condition or the normality when both d and n tend to infinity. By choosing n as $n = n(d)$ with a suitable order of d , PC directions are separated for $\alpha_i \in (0, 1]$ in (1) as well.

Theorem 3. *Assume that the first m population eigenvalues are distinct such as $\lambda_1 > \dots > \lambda_m$. Then, the first m sample eigenvectors are consistent in the sense of (5) under the conditions (i)–(ii) in Theorem 1.*

Corollary 2. *Assume further that z_{ij} , $i = 1, \dots, d$ ($j = 1, \dots, n$) are independent. Then, the first m sample eigenvectors are consistent in the sense of (5) under the conditions (i)–(ii) in Corollary 1.*

Remark 3. In the case that $\mathbf{X}$ is Gaussian and $\alpha_i \rightarrow 0$, Corollary 2 (ii) corresponds to the results given by Johnstone and Lu (2008, Theorems 1–3) and Paul (2007, Theorem 4).

Remark 4. One has a sample eigenvector as $\hat{\mathbf{h}}_i = (n\hat{\lambda}_i)^{-1/2} \mathbf{X} \hat{\mathbf{u}}_i$. Note that $\hat{\mathbf{h}}_i$ can be calculated by using a unit-norm eigenvector, $\hat{\mathbf{u}}_i$, of $\mathbf{S}_D$ whose size is much smaller than $\mathbf{S}$ especially for a HDLSS data matrix.

Remark 5. Jung and Marron (2008) discussed *subspace consistency* as $d \rightarrow \infty$ for the sample eigenvectors corresponding to the population eigenvalues given by (b') that is described

above. The *subspace consistency* is defined as follows:

$$\text{Angle}(\hat{\mathbf{h}}_{s_{j-1}+i}, \text{span}\{\mathbf{h}_{s_{j-1}+1}, \dots, \mathbf{h}_{s_j}\}) \xrightarrow{p} 0 \quad \text{as } d \rightarrow \infty$$

$$(i = 1, \dots, s_j - s_{j-1}; j = 1, \dots, p)$$

for the first κ eigenvalues that are distinct but $\alpha_1 = \dots = \alpha_{s_1} > \alpha_{s_1+1} = \dots = \alpha_{s_2} > \dots > \alpha_{s_{p-1}+1} = \dots = \alpha_{s_p} > 1$, where $p \leq m$, $s_0 = 0$ and $s_p = \kappa$. When the population eigenvalues are not distinct such as $\lambda_1 \geq \dots \geq \lambda_m$, we can develop the notion of subspace consistency with the help of the conditions (i)–(ii) in Theorem 1 as follows:

$$\text{Angle}(\hat{\mathbf{h}}_{t_{j-1}+i}, \text{span}\{\mathbf{h}_{t_{j-1}+1}, \dots, \mathbf{h}_{t_j}\}) \xrightarrow{p} 0 \quad (i = 1, \dots, t_j - t_{j-1}; j = 1, \dots, r)$$

under the conditions (i)–(ii) in Theorem 1

for the first m eigenvalues such that $\lambda_1 = \dots = \lambda_{t_1} > \lambda_{t_1+1} = \dots = \lambda_{t_2} > \dots > \lambda_{t_{r-1}+1} = \dots = \lambda_{t_r}$, where $r \leq m$, $t_0 = 0$ and $t_r = m$.

5. Principal component scores

The estimation of principal component scores (Pcs) is an important issue in PCA. The i -th Pcs of $\mathbf{x}_j$ is given by $\mathbf{h}_i^T \mathbf{x}_j = \sqrt{\lambda_i} z_{ij}$ ($= s_{ij}$, say). However, since $\mathbf{h}_i$ is unknown, one estimates $\mathbf{h}_i^T \mathbf{x}_j$ by using some reasonable estimate of $\mathbf{h}_i$. In view of Remark 4, one has an estimate of $\mathbf{h}_i$ as $\hat{\mathbf{h}}_i = (n\hat{\lambda}_i)^{-1/2} \mathbf{X} \hat{\mathbf{u}}_i$. Thus, the i -th Pcs of $\mathbf{x}_j$ is estimated by $\hat{\mathbf{h}}_i^T \mathbf{x}_j = \sqrt{n\hat{\lambda}_i} \hat{u}_{ij}$ ($= \hat{s}_{ij}$, say), where $\hat{\mathbf{u}}_i^T = (\hat{u}_{i1}, \dots, \hat{u}_{in})$. Let us define a sample variance of the i -th Pcs by $V(\hat{s}_i) = n^{-1} \sum_{j=1}^n (\hat{s}_{ij} - s_{ij})^2$. Then, we obtain the following

Theorem 4. *Assume that the first m population eigenvalues are distinct such that $\lambda_1 > \dots > \lambda_m$. Then, for $i = 1, \dots, m$, we have that*

$$\frac{V(\hat{s}_i)}{\lambda_i} = o_p(1) \tag{6}$$

under the conditions (i)–(ii) in Theorem 1.

Corollary 3. *Assume further that z_{ij} , $i = 1, \dots, d$ ($j = 1, \dots, n$) are independent. Then, we*

have (6) under the conditions (i)–(ii) in Corollary 1.

Remark 6. When the population eigenvalues are not distinct such as $\lambda_1 \geq \dots \geq \lambda_m$ but $\lambda_i \neq \lambda_{i'}$ ($i \neq i'$) for some i , we can still claim (6) under the conditions (i)–(ii) in Theorem 1.

Corollary 4. Suppose the assumption in Theorem 4. Then, the first m eigenvectors of $\mathbf{S}_D$ are consistent in the sense that

$$\text{Angle}(\hat{\mathbf{u}}_i, n^{-1/2} \mathbf{z}_i) \xrightarrow{p} 0$$

for $i = 1, \dots, m$, under the conditions (i)–(ii) in Theorem 1.

Remark 7. When the population eigenvalues are not distinct such as $\lambda_1 \geq \dots \geq \lambda_m$, we can claim *subspace consistency* under the conditions (i)–(ii) in Theorem 1 as follows:

$$\text{Angle}(\hat{\mathbf{u}}_{t_{j-1}+i}, \text{span}\{n^{-1/2} \mathbf{z}_{t_{j-1}+1}, \dots, n^{-1/2} \mathbf{z}_{t_j}\}) \xrightarrow{p} 0 \quad (i = 1, \dots, t_j - t_{j-1}; j = 1, \dots, r)$$

for the first m eigenvalues such that $\lambda_1 = \dots = \lambda_{t_1} > \lambda_{t_1+1} = \dots = \lambda_{t_2} > \dots > \lambda_{t_{r-1}+1} = \dots = \lambda_{t_r}$, where $r \leq m$, $t_0 = 0$ and $t_r = m$.

Appendix

Throughout this section, let us write $\mathbf{R}_n = \{\mathbf{e}_n \in \mathbf{R}^n : \|\mathbf{e}_n\| = 1\}$ and let $\mathbf{e}_{jn}$, $j = 1, 2$, be arbitrary elements of $\mathbf{R}_n$. Let us write that $\mathbf{U}_1 = n^{-1} \sum_{i=1}^m \lambda_i \mathbf{z}_i \mathbf{z}_i^T$ and $\mathbf{U}_2 = n^{-1} \sum_{s=m+1}^d \lambda_s \mathbf{z}_s \mathbf{z}_s^T$. Let us write $u_i = n^{-1} \sum_{s=m+1}^d \lambda_s z_{si}^2$ as a diagonal element of $\mathbf{U}_2$. Let $\mathbf{U}_{2(t)} = (u_{ij(t)})$, $t = 1, 2, \dots$, be $n \times n$ matrices such that

$$u_{ij(t)} = \begin{cases} n^{-1} \sum_{s=m+1}^d \lambda_s^{2^{t-1}} z_{si} z_{sj} & (i \neq j), \\ 0 & (i = j). \end{cases}$$

Let $\mathbf{V}_{2(1)} = (v_{ij(1)})$ be $n \times n$ matrices such that

$$v_{ij(1)} = \begin{cases} n^{-1} \sum_{s=1}^d z_{si} z_{sj} & (i \neq j), \\ 0 & (i = j). \end{cases}$$

Let $\mathbf{V}_{2(2)} = (v_{ij(2)})$ be $n \times n$ matrices such that

$$v_{ij(2)} = \begin{cases} n^{-2} \sum_{s_1 \neq s_2}^d z_{s_1 i} z_{s_2 j} \sum_{k=1(\setminus i, j)}^n z_{s_1 k} z_{s_2 k} & (i \neq j), \\ 0 & (i = j), \end{cases}$$

where $(\setminus i, j, k, \dots)$ excludes numbers $i, j, k, \dots$. Let $\omega_{i, j, k(i', j', \dots)} = \sum_{k=1(\setminus i', j', \dots)}^n z_{ik} z_{jk}$. Then, let $\mathbf{V}_{2(r)} = (v_{ij(r)})$, $r = 3, 4, \dots (< d)$, be $n \times n$ matrices such that

$$v_{ij(r)} = \begin{cases} n^{-r} \sum_{s_1 \neq s_2 \neq \dots \neq s_r}^d z_{s_1 i} z_{s_r j} \omega_{s_1, s_2, k_1(i, j)} \left(\prod_{l=2}^{r-1} \omega_{s_l, s_{l+1}, k_l(i, j, k_{l-1})} \right) & (i \neq j), \\ 0 & (i = j). \end{cases}$$

Suppose that $\alpha_1 = \dots = \alpha_{s_1} > \alpha_{s_1+1} = \dots = \alpha_{s_2} > \dots > \alpha_{s_{l-1}+1} = \dots = \alpha_{s_l} (= \alpha_m)$, where $l \leq m$. For every $i (= 1, \dots, l)$, let $\mathbf{U}_{1i} = n^{-1} \sum_{j=1}^{s_i} \lambda_j \mathbf{z}_j \mathbf{z}_j^T$. Let $\tilde{\lambda}_{i1} \geq \dots \geq \tilde{\lambda}_{is_i}$ be eigenvalues of $\mathbf{U}_{1i}$ and let $\tilde{\mathbf{u}}_{ij} \in \mathbf{R}_n$ be an eigenvector corresponding to $\tilde{\lambda}_{ij}$ ($j = 1, \dots, s_i$). Then, we have the eigen-decomposition as $\mathbf{U}_{1i} = \sum_{j=1}^{s_i} \tilde{\lambda}_{ij} \tilde{\mathbf{u}}_{ij} \tilde{\mathbf{u}}_{ij}^T$. Let $\tilde{\mathbf{z}}_i = (||n^{-1/2} \mathbf{z}_i||)^{-1} n^{-1/2} \mathbf{z}_i$, $i = 1, \dots, m$.

Lemma 1. *It holds for $i = 1, \dots, m$, that*

$$||d^{-2^{t-1}\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t)}||^2 = d^{-2^t\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t+1)} \mathbf{e}_{1n} + o_p(1), \quad t = 1, 2, \dots$$

under the conditions:

- (i) $d \rightarrow \infty$ and n is fixed for i such that $\alpha_i > 1/2$,
- (ii) $d \rightarrow \infty$ and $d^{2-2\alpha_i}/n \rightarrow 0$ for i such that $\alpha_i \in (0, 1/2]$.

Proof. Let $\mathbf{e}_{1n} = (e_{11}, \dots, e_{1n})^T$, where $\sum_{i=1}^n e_{1i}^2 = 1$. For every $t (= 1, 2, \dots)$, we write that

$$\begin{aligned} ||\mathbf{e}_{1n}^T \mathbf{U}_{2(t)}||^2 &= \sum_{k=1}^n \sum_{i'=1(\setminus k)}^n e_{1i'}^2 u_{i'k(t)}^2 + \sum_{k=1}^n \sum_{i' \neq j'(\setminus k)}^n e_{1i'} e_{1j'} u_{i'k(t)} u_{j'k(t)} \\ &= \sum_{i'=1}^n e_{1i'}^2 \sum_{k=1(\setminus i')}^n u_{i'k(t)}^2 + \sum_{i' \neq j'}^n e_{1i'} e_{1j'} \sum_{k=1(\setminus i', j')}^n u_{i'k(t)} u_{j'k(t)}. \end{aligned} \quad (7)$$

We first consider the second term in (7). Let us write that

$$y_{ij(1)} = n^{-1} \sum_{s=m+1}^d \lambda_s^{2^t} z_{si} z_{sj} \left(n^{-1} \sum_{k=1(\setminus i,j)}^n z_{sk}^2 - 1 \right),$$

$$y_{ij(2)} = n^{-2} \sum_{s_1 \neq s_2 (\geq m+1)} \lambda_{s_1}^{2^{t-1}} \lambda_{s_2}^{2^{t-1}} z_{s_1 i} z_{s_2 j} \sum_{k=1(\setminus i,j)}^n z_{s_1 k} z_{s_2 k}.$$

Then, it holds that

$$\sum_{i' \neq j'} e_{1i'} e_{1j'} \sum_{k=1(\setminus i', j')}^n u_{i'k(t)} u_{j'k(t)} = \sum_{i' \neq j'} e_{1i'} e_{1j'} (y_{i'j'(1)} + y_{i'j'(2)} + u_{i'j'(t+1)}). \quad (8)$$

Here, by using Markov's inequality, for any $\tau (> 0)$ and the uniform bound M for the fourth moments condition of z_{ij} , one has for $i' \neq j'$ under (i)–(ii) that

$$P\left(\sum_{i' \neq j'} |d^{-2^t \alpha_i} y_{i'j'(1)}|^2 > \tau\right) \leq \tau^{-1} d^{-2^{t+1} \alpha_i} n^{-1} M \sum_{s=m+1}^d \lambda_s^{2^{t+1}} \leq \tau^{-1} M d^{1-2^{t+1} \alpha_i} n^{-1} \lambda_{m+1}^{2^{t+1}} = o(1). \quad (9)$$

Thus we have that

$$d^{-2^t \alpha_i} \|0, y_{12(1)}, y_{13(1)}, \dots, y_{1n(1)}, y_{21(1)}, 0, y_{23(1)}, \dots, y_{n1(1)}, \dots, y_{nn-1(1)}, 0\| = o_p(1).$$

From the fact that $\sum_{i' \neq j'} e_{1i'}^2 e_{1j'}^2 \leq 1$, since it holds that

$$\|0, e_{11}e_{12}, e_{11}e_{13}, \dots, e_{11}e_{1n}, e_{12}e_{11}, 0, e_{12}e_{13}, \dots, e_{1n}e_{11}, \dots, e_{1n}e_{1n-1}, 0\| \leq 1,$$

we obtain that

$$d^{-2^t \alpha_i} \sum_{i' \neq j'} e_{1i'} e_{1j'} y_{i'j'(1)} = o_p(1). \quad (10)$$

On the other hand, one has for $i' \neq j'$ under (i)–(ii) that

$$P\left(\sum_{i' \neq j'} |d^{-2^t \alpha_i} y_{i'j'(2)}|^2 > \tau\right) \leq \tau^{-1} M d^{2-2^{t+1} \alpha_i} n^{-1} \lambda_{m+1}^{2^{t+1}} = o(1).$$

Then, similarly to (9)–(10), we have that

$$d^{-2^t \alpha_i} \sum_{i' \neq j'} e_{1i'} e_{1j'} y_{i'j'(2)} = o_p(1). \quad (11)$$

By combining (10) and (11) with (8), we obtain under (i)–(ii) that

$$d^{-2^t \alpha_i} \sum_{i' \neq j'} e_{1i'} e_{1j'} \sum_{k=1(\setminus i', j')}^n u_{i'k(t)} u_{j'k(t)} = d^{-2^t \alpha_i} \sum_{i' \neq j'} e_{1i'} e_{1j'} u_{i'j'(t+1)} + o_p(1). \quad (12)$$

Next, we consider the first term in (7). When $\alpha_i > 1/2$, by using Markov's inequality for any $\tau (> 0)$, one has for $i' \neq j'$ under (i) that

$$\sum_{i'=1}^n P\left(d^{-2^t \alpha_i} \sum_{k=1(\setminus i')}^n u_{i'k(t)}^2 > \tau\right) \leq \tau^{-1} n d^{-2^t \alpha_i} \sum_{k=1(\setminus i')}^n E(u_{i'k(t)}^2) = O(d^{1-2\alpha_i}) = o(1).$$

When $\alpha_i < 1/2$, by using Chebyshev's inequality for any $\tau (> 0)$, one has that

$$\begin{aligned} \sum_{i'=1}^n P\left(d^{-2^t \alpha_i} \sum_{k=1(\setminus i')}^n u_{i'k(t)}^2 > \tau\right) &\leq \tau^{-2} n d^{-2^{t+1} \alpha_i} E\left(\left(\sum_{k=1(\setminus i')}^n u_{i'k(t)}^2\right)^2\right) \\ &= d^{4-2^{t+1} \alpha_i} / n^2 + d^{2-2^{t+1} \alpha_i} / n = o(1) \end{aligned}$$

for $i' \neq j'$ under (ii). Thus we obtain under (i)–(ii) that

$$d^{-2^t \alpha_i} \sum_{i'=1}^n e_{1i'}^2 \sum_{k=1(\setminus i')}^n u_{i'k(t)}^2 = o_p(1). \quad (13)$$

By combining (12) and (13) with (7), we conclude the result. $\square$

Lemma 2. *It holds for $i = 1, \dots, m$, that $\|d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(1)}\|^2 = o_p(1)$ under the conditions (i)–(ii) in Lemma 1.*

Proof. Since $\alpha_i > 0$ in (1), there is at most one positive integer $t_o (\geq 2)$ satisfying $1 - 2^{t_o} \alpha_i < 0$. Then, we have for $i' \neq j'$ under (i)–(ii) in Lemma 1 that

$$P\left(d^{-2^{t_o} \alpha_i} \sum_{i' \neq j'} |u_{i'j'(t_o)}|^2 > \tau\right) \leq \tau^{-1} d^{1-2^{t_o} \alpha_i} \lambda_{m+1}^{2^{t_o}} = o_p(1).$$

Then, similarly to (9)–(10), it holds that $d^{-2^{t_o-1} \alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t_o)} \mathbf{e}_{1n} = o_p(1)$. Thus we have that

$$\|d^{-2^{t_o-2} \alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t_o-1)}\|^2 = d^{-2^{t_o-1} \alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t_o)} \mathbf{e}_{1n} + o_p(1) = o_p(1),$$

so that $d^{-2^{t_o-2} \alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t_o-1)} \mathbf{e}_{1n} = o_p(1)$. Hence, we obtain that

$$\|d^{-2^{t_o-3} \alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t_o-2)}\|^2 = d^{-2^{t_o-2} \alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(t_o-1)} \mathbf{e}_{1n} + o_p(1) = o_p(1).$$

Similarly, we claim until $\mathbf{U}_{2(2)}$ to obtain that

$$||d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(1)}||^2 = d^{-2\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(2)} \mathbf{e}_{1n} + o_p(1) = o_p(1),$$

which concludes the result. $\square$

Lemma 3. *It holds for $i = 1, \dots, m$, that $d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_2 \mathbf{e}_{2n} = o_p(1)$ under the conditions:*

- (i) $d \rightarrow \infty$ and n is fixed for i such that $\alpha_i > 1$,
- (ii) $d \rightarrow \infty$ and $d^{2-2\alpha_i}/n \rightarrow 0$ for i such that $\alpha_i \in (0, 1]$.

Proof. By using Chebyshev's inequality, for any $\tau > 0$ and the uniform bound M for the fourth moments condition, one has under (i)–(ii) that

$$\begin{aligned} \sum_{k=1}^n P\left(d^{-\alpha_i} u_k > \tau\right) &= \sum_{k=1}^n P\left((nd^{\alpha_i})^{-1} \sum_{s=m+1}^d \lambda_s z_{sk}^2 > \tau\right) \\ &\leq (\tau n^{1/2} d^{\alpha_i})^{-2} M \left(\sum_{s=m+1}^d \lambda_s \right)^2 = O(d^{2-2\alpha_i}/n) = o(1). \end{aligned}$$

Thus it holds that $d^{-\alpha_i} u_k = o_p(1)$ for all k ($= 1, \dots, n$). From Lemma 2, we have under (i)–(ii) in Lemma 1 that $d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(1)} \mathbf{e}_{2n} = o_p(1)$, $i = 1, \dots, m$. Hence, we obtain that

$$d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_2 \mathbf{e}_{2n} = d^{-\alpha_i} \left(\mathbf{e}_{1n}^T \mathbf{U}_{2(1)} \mathbf{e}_{2n} + \mathbf{e}_{1n}^T \text{diag}(u_1, \dots, u_n) \mathbf{e}_{2n} \right) = o_p(1) \quad (i = 1, \dots, m)$$

under (i)–(ii). $\square$

Lemma 4. *It holds for $i = 1, \dots, m$, that*

$$d^{-\alpha_i} n^{-1} \mathbf{z}_{i'}^T \mathbf{U}_2 \mathbf{z}_{j'} = o_p(n^{-1/2}) \quad (i' = 1, \dots, m; j' = 1, \dots, m)$$

under the conditions (i)–(ii) in Theorem 1.

Proof. One can write that

$$d^{-\alpha_i} n^{-1} \mathbf{z}_{i'}^T \mathbf{U}_{2(1)} \mathbf{z}_{j'} = d^{-\alpha_i} \sum_{k_1 \neq k_2} n^{-1} z_{i'k_1} z_{j'k_2} u_{k_1 k_2(1)}.$$

We first consider the case of $i' = j'$. Note that $E(n^{-2}z_{i'k_1}^2 z_{i'k_2} z_{i'k_3} u_{k_1k_2(1)} u_{k_1k_3(1)}) = 0$ ($k_1 \neq k_2 \neq k_3$) and $E(n^{-2}z_{i'k_1}^2 z_{i'k_2}^2 u_{k_1k_2(1)}^2) \leq M^2(\sum_{s=m+1}^d \lambda_s)^2$ ($k_1 \neq k_2$) with the uniform bound M for the fourth moments condition. Then, for any $\tau > 0$ and the uniform bound M for the fourth moments condition, one has under (i)–(ii) that

$$\begin{aligned} P\left(|d^{-\alpha_i} \sum_{k_1 \neq k_2} n^{-1} z_{i'k_1} z_{i'k_2} u_{k_1k_2(1)}| > n^{-1/2}\tau\right) &\leq \tau^{-2} d^{-2\alpha_i} n^{-1} \sum_{k_1 \neq k_2} E(z_{i'k_1}^2 z_{i'k_2}^2 u_{k_1k_2(1)}^2) \\ &= d^{2-2\alpha_i}/n = o(1), \\ P\left(d^{-\alpha_i} n^{-1} \sum_{k=1}^n z_{i'k}^2 u_k > n^{-1/2}\tau\right) &\leq \tau^{-1} n^{-1/2} d^{-\alpha_i} M \sum_{s=m+1}^d \lambda_s = O(d^{1-\alpha_i}/n^{1/2}) = o(1). \end{aligned}$$

Hence, we obtain that

$$\begin{aligned} d^{-\alpha_i} n^{-1} \mathbf{z}_{i'}^T \mathbf{U}_2 \mathbf{z}_{i'} &= d^{-\alpha_i} (n^{-1} \mathbf{z}_{i'}^T \mathbf{U}_{2(1)} \mathbf{z}_{i'} + n^{-1} \mathbf{z}_{i'}^T \text{diag}(u_1, \dots, u_n) \mathbf{z}_{i'}) \\ &= o_p(n^{-1/2}) \quad (i = 1, \dots, m). \end{aligned}$$

As for the case of $i' \neq j'$, note that

$$P\left(|d^{-\alpha_i} \sum_{k_1 \neq k_2} n^{-1} z_{i'k_1} z_{j'k_2} u_{k_1k_2(1)}| > n^{-1/2}\tau\right) = o(1)$$

and

$$\begin{aligned} P\left(d^{-\alpha_i} n^{-1} \left| \sum_{k=1}^n z_{i'k} z_{j'k} u_k \right| > n^{-1/2}\tau\right) &\leq P\left(d^{-\alpha_i} n^{-1/2} \sum_{k=1}^n |z_{i'k} z_{j'k}| u_k > \tau\right) \\ &\leq \tau^{-1} n^{1/2} d^{-\alpha_i} E(|z_{i'k} z_{j'k}| u_k) \\ &\leq \tau^{-1} n^{1/2} d^{-\alpha_i} (E(z_{i'k}^2 z_{j'k}^2) E(u_k^2))^{1/2} \\ &= O(d^{1-\alpha_i}/n^{1/2}) = o(1). \end{aligned}$$

Therefore, we can conclude the result. $\square$

Lemma 5. Assume that z_{ij} , $i = 1, \dots, d$ ($j = 1, \dots, n$) are independent. It holds for $\alpha_i \in (0, 1/2]$ that

$$\begin{aligned} \|d^{-2^{t-1}\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(r)}\|^2 &= d^{-2^t\alpha_i} \mathbf{e}_{1n}^T (\mathbf{V}_{2(2r-1)} + \mathbf{V}_{2(2r)}) \mathbf{e}_{1n} + o_p(1) \\ (t = 1, 2, \dots; r = 1, \dots, 2^{t-1}) \end{aligned}$$

under the condition: (i) $d \rightarrow \infty$ and $d^{1-2\alpha_i}/n \rightarrow 0$.

Proof. By using Chebyshev's inequality, for any $\tau (> 0)$ and the uniform bound M for the fourth moments condition, one has under (i) that

$$\begin{aligned} & \sum_{i'=1}^n P\left(d^{-2^t\alpha_i} \left| \sum_{k=1(\setminus i')}^n v_{i'k(1)}^2 - n^{-1}d \right| > \tau\right) \\ & \leq \tau^{-2} M^2 (d^{2-2^{t+1}\alpha_i}/n^2 + d^{1-2^{t+1}\alpha_i}/n) = o(1), \quad t = 1, 2, \dots \end{aligned}$$

Thus it holds for every $t (= 1, 2, \dots)$ under (i) that

$$d^{-2^t\alpha_i} \sum_{i'=1}^n e_{1i'}^2 \sum_{k=1(\setminus i')}^n v_{i'k(1)}^2 = d^{-2^t\alpha_i} \sum_{i'=1}^n e_{1i'}^2 n^{-1}d + o_p(1) = o_p(1), \quad (14)$$

where e_{1i} 's are components of $\mathbf{e}_{1n} = (e_{11}, \dots, e_{1n})^T$ with $\sum_{i=1}^n e_{1i}^2 = 1$. Then, similarly to (7)–(10), we obtain that

$$\begin{aligned} \|d^{-2^{t-1}\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(1)}\|^2 &= d^{-2^t\alpha_i} \sum_{i' \neq j'} e_{1i'} e_{1j'} (v_{i'j'(1)} + v_{i'j'(2)}) + o_p(1) \\ &= d^{-2^t\alpha_i} \mathbf{e}_{1n}^T (\mathbf{V}_{2(1)} + \mathbf{V}_{2(2)}) \mathbf{e}_{1n} + o_p(1), \quad t = 1, 2, \dots \end{aligned} \quad (15)$$

Now, note that

$$\|\mathbf{e}_{1n}^T \mathbf{V}_{2(2)}\|^2 = \sum_{i'=1}^n e_{1i'}^2 \sum_{k=1(\setminus i')}^n v_{i'k(2)}^2 + \sum_{i' \neq j'} e_{1i'} e_{1j'} \sum_{k=1(\setminus i', j')}^n v_{i'k(2)} v_{j'k(2)}.$$

By using Chebyshev's inequality for any $\tau (> 0)$, we have for every $t (= 2, 3, \dots)$ under (i) that

$$\sum_{i'=1}^n P\left(d^{-2^t\alpha_i} \left| \sum_{k=1(\setminus i')}^n v_{i'k(2)}^2 - n^{-2}d(d-1) \right| > \tau\right) = o(1),$$

so that $d^{-2^t\alpha_i} \sum_{i'=1}^n e_{1i'}^2 \sum_{k=1(\setminus i')}^n v_{i'k(2)}^2 = o_p(1)$ in a way similar to (14). Let us write that

$\sum_{k=1(\setminus i', j')}^n v_{i'k(2)} v_{j'k(2)} = y_{i'j'(3)} + y_{i'j'(4)} + v_{i'j'(3)} + v_{i'j'(4)}$, where

$$\begin{aligned} y_{ij(3)} &= n^{-4} \sum_{s_2 \neq s_1, s_3}^d z_{s_1 i} z_{s_3 j} \omega_{s_2, s_2, k_1(i, j)} \omega_{s_1, s_2, k_2(k_1, i)} \omega_{s_2, s_3, k_3(k_1, j)} - v_{ij(3)}, \\ y_{ij(4)} &= n^{-4} \sum_{s_1 \neq s_2}^d \sum_{s_3(\setminus s_2) \neq s_4}^d z_{s_1 i} z_{s_4 j} \omega_{s_2, s_3, k_1(i, j)} \omega_{s_1, s_2, k_2(k_1, i)} \omega_{s_3, s_4, k_3(k_1, j)} - v_{ij(4)}. \end{aligned}$$

Note that $\omega_{s_1, s_2, k_1(i_1, i_2)} = \omega_{s_1, s_2, k_1(i_1, i_2, k_2)} + z_{s_1 k_2} z_{s_2 k_2}$. Note that $\omega_{s_1, s_2, k_1(i_1, i_2)}$ and $\omega_{s_3, s_4, k_2(i_3, i_4)}$ are independent for any $s_1 \neq s_2 \neq s_3 \neq s_4$, and $\omega_{s_1, s_2, k_1(i_1, i_2)}$ and $\omega_{s_3, s_4, k_2(i_3, i_4, k_1)}$ are independent for any s_1, s_2, s_3 and s_4 . Then, similarly to (9), it holds for every $t (= 2, 3, \dots)$ under (i) that

$$P\left(\sum_{i' \neq j'} |d^{-2^t \alpha_i} y_{i'j'(3)}|^2 > \tau\right) = o(1) \quad \text{and} \quad P\left(\sum_{i' \neq j'} |d^{-2^t \alpha_i} y_{i'j'(4)}|^2 > \tau\right) = o(1).$$

Thus we obtain that

$$\|d^{-2^{t-1} \alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(2)}\|^2 = d^{-2^t \alpha_i} \mathbf{e}_{1n}^T (\mathbf{V}_{2(3)} + \mathbf{V}_{2(4)}) \mathbf{e}_{1n} + o_p(1), \quad t = 2, 3, \dots$$

in a way similar to (15). Similarly, we can claim until $r = 2^{t-1}$ that

$$\|d^{-2^{t-1} \alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(2^{t-1})}\|^2 = d^{-2^t \alpha_i} \mathbf{e}_{1n}^T (\mathbf{V}_{2(2^{t-1}-1)} + \mathbf{V}_{2(2^t)}) \mathbf{e}_{1n} + o_p(1), \quad t = 1, 2, \dots$$

It concludes the result. $\square$

Lemma 6. Assume that z_{ij} , $i = 1, \dots, d$ ($j = 1, \dots, n$) are independent. It holds for $\alpha_i \in (0, 1/2]$ that $\|d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(1)}\|^2 = o_p(1)$ under the condition:

(i) $d \rightarrow \infty$ and there exists a positive constant ε_i satisfying $d^{1-2\alpha_i}/n < d^{-\varepsilon_i}$.

Proof. First, let us show that one can claim that $\|d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(1)}\|^2 = o_p(1)$ under (i). We first consider the case that $\varepsilon_i \leq 2\alpha_i$. Let $\beta_t = 1 - 2\alpha_i - (2^{t-1} - 1)\varepsilon_i$. Then, we have for every $t (= 1, 2, \dots)$ that

$$n^2 E\{(d^{-2^{t-1} \alpha_i} v_{ij(r)})^2\} = O(d^{r-2^t \alpha_i}/n^{r-1}) = O(d^{\beta_t}), \quad r = 1, \dots, 2^{t-1}.$$

Here, in view of (i), there is at most one positive integer $t_\varepsilon (\geq 2)$ satisfying $\beta_{t_\varepsilon} < 0$. Then, it holds for $i' \neq j'$ under (i) that

$$P\left(\sum_{i' \neq j'} |d^{-2^{t_\varepsilon-1} \alpha_i} v_{i'j'(r)}|^2 > \tau\right) \leq \tau^{-1} n^2 E\{(d^{-2^{t_\varepsilon-1} \alpha_i} v_{i'j'(r)})^2\} = O(d^{\beta_{t_\varepsilon}}) = o(1).$$

Then, similarly to (9)–(10), we have under (i) that

$$d^{-2^{t_\varepsilon-1} \alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(r)} \mathbf{e}_{1n} = d^{-2^{t_\varepsilon-1} \alpha_i} \sum_{i' \neq j'} e_{1i'} e_{1j'} v_{i'j'(r)} = o_p(1), \quad r = 1, \dots, 2^{t_\varepsilon-1},$$

so that

$$\sum_{r=1}^{2^{t_\varepsilon-2}} \|d^{-2^{t_\varepsilon-2}\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(r)}\|^2 = \sum_{r=1}^{2^{t_\varepsilon-1}} d^{-2^{t_\varepsilon-1}\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(r)} \mathbf{e}_{1n} = o_p(1).$$

Similarly to the proof of Lemma 2, we claim until $t = 1$ to obtain that

$$\|d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(1)}\|^2 = \sum_{r=1}^2 d^{-2\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(r)} \mathbf{e}_{1n} + o_p(1) = o_p(1).$$

Next, we consider the case that $\varepsilon > 2\alpha_i$. There is at most one positive integer t_o (≥ 2) satisfying $1 - 2^{t_o}\alpha_i < 0$. Thus it holds that $n^2 E\{(d^{-2^{t_o-1}\alpha_i} v_{ij(r)})^2\} = O(d^{r-2^{t_o}\alpha_i}/n^{r-1}) = O(d^{1-2^{t_o}}) = o(1)$. Hence, similarly to the former case, we obtain that $\|d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{V}_{2(1)}\|^2 = o_p(1)$ under (i). Finally, by replacing $\mathbf{V}_{2(1)}$ with $\mathbf{U}_{2(1)}$, we can claim that $\|d^{-\alpha_i} \mathbf{e}_{1n}^T \mathbf{U}_{2(1)}\|^2 = o_p(1)$ under (i) in a similar fashion. $\square$

Lemma 7. Let $\mathbf{U}_1 = \sum_{j=1}^m \tilde{\lambda}_j \tilde{\mathbf{u}}_j \tilde{\mathbf{u}}_j^T$ be the eigen-decomposition of $\mathbf{U}_1$, where $\tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_m$ are eigenvalues and $\tilde{\mathbf{u}}_j \in \mathbf{R}_n$ is an eigenvector corresponding to $\tilde{\lambda}_j$ ($j = 1, \dots, m$). Assume that the first m population eigenvalues are distinct as $\lambda_1 > \dots > \lambda_m$. Then, it holds as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\frac{\tilde{\lambda}_i}{\lambda_i} = \|n^{-1/2} \mathbf{z}_i\|^2 + o_p(n^{-1/2}) = 1 + o_p(1), \quad \tilde{\mathbf{u}}_i^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}) \quad (i = 1, \dots, m).$$

Proof. By using Chebyshev's inequality, for any τ (> 0) and the uniform bound M for the fourth moments condition, one has as $n \rightarrow \infty$ that

$$P(|n^{-1} \mathbf{z}_i^T \mathbf{z}_j| > n^{-1/4} \tau) = P\left(\left|n^{-1} \sum_{k=1}^n z_{ik} z_{jk}\right| > n^{-1/4} \tau\right) \leq \tau^{-2} M n^{-1/2} \rightarrow 0 \quad (i \neq j).$$

Thus we claim as $n \rightarrow \infty$ that $n^{-1} \mathbf{z}_i^T \mathbf{z}_j = o_p(n^{-1/4})$ ($i \neq j$). Note that $\|n^{-1/2} \mathbf{z}_i\|^2 = 1 + o_p(1)$ as $n \rightarrow \infty$. First, let us consider λ_i ($i = 1, \dots, s_1$) that holds power α_{s_1} . By noting as $n \rightarrow \infty$ that $\lambda_1 \|n^{-1/2} \mathbf{z}_1\|^2 > \dots > \lambda_m \|n^{-1/2} \mathbf{z}_m\|^2$ w.p.1, we have that

$$\begin{aligned} \frac{\tilde{\lambda}_1}{\lambda_1} &= \tilde{\mathbf{u}}_1^T \frac{\mathbf{U}_1}{\lambda_1} \tilde{\mathbf{u}}_1 = \tilde{\mathbf{u}}_1^T \left(\sum_{j=1}^m \frac{\lambda_j}{\lambda_1 n} \mathbf{z}_j \mathbf{z}_j^T \right) \tilde{\mathbf{u}}_1 = \tilde{\mathbf{u}}_1^T \left(\sum_{j=1}^m \frac{\lambda_j}{\lambda_1} \|n^{-1/2} \mathbf{z}_j\|^2 \tilde{\mathbf{z}}_j \tilde{\mathbf{z}}_j^T \right) \tilde{\mathbf{u}}_1 \\ &= \|n^{-1/2} \mathbf{z}_1\|^2 + o_p(n^{-1/2}) = 1 + o_p(1). \end{aligned} \tag{16}$$

Then, it holds that $\tilde{\mathbf{u}}_1^T \tilde{\mathbf{z}}_1 = 1 + o_p(n^{-1/2})$, so that $\tilde{\mathbf{u}}_2^T \tilde{\mathbf{z}}_1 = o_p(n^{-1/4})$. Since we have as $n \rightarrow \infty$ that

$$\frac{\tilde{\lambda}_2}{\lambda_2} = \tilde{\mathbf{u}}_2^T \frac{\mathbf{U}_1}{\lambda_2} \tilde{\mathbf{u}}_2 = \tilde{\mathbf{u}}_2^T \left(\sum_{j=2}^m \frac{\lambda_j}{\lambda_2} \|n^{-1/2} \mathbf{z}_j\|^2 \tilde{\mathbf{z}}_j \tilde{\mathbf{z}}_j^T \right) \tilde{\mathbf{u}}_2 = \|n^{-1/2} \mathbf{z}_2\|^2 + o_p(n^{-1/2}) = 1 + o_p(1),$$

it holds that $\tilde{\mathbf{u}}_2^T \tilde{\mathbf{z}}_2 = 1 + o_p(n^{-1/2})$ by noting that $\tilde{\mathbf{z}}_1^T \tilde{\mathbf{z}}_2 = o_p(n^{-1/4})$. Similarly, we claim until s_1 to obtain as $n \rightarrow \infty$ that

$$\frac{\tilde{\lambda}_i}{\lambda_i} = \|n^{-1/2} \mathbf{z}_i\|^2 + o_p(n^{-1/2}) = 1 + o_p(1), \quad \tilde{\mathbf{u}}_i^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}) \quad (i = 1, \dots, s_1). \quad (17)$$

Then, we obtain for $\mathbf{U}_{11} = \sum_{j=1}^{s_1} \tilde{\lambda}_{1j} \tilde{\mathbf{u}}_{1j} \tilde{\mathbf{u}}_{1j}^T$ that

$$\frac{\tilde{\lambda}_{1i}}{\lambda_i} = 1 + o_p(1), \quad \tilde{\mathbf{u}}_{1i}^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}) \quad (i = 1, \dots, s_1). \quad (18)$$

Next, let us consider λ_i ($i = s_1 + 1, \dots, s_2$) that holds power α_{s_2} . From (18), note that $\tilde{\mathbf{u}}_{1j}^T \tilde{\mathbf{z}}_{j'} = o_p(n^{-1/4})$ ($j = 1, \dots, s_1$; $j' = s_1 + 1, \dots, m$). Thus we have as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_1}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_i = \tilde{\mathbf{u}}_{1j}^T \frac{\sum_{j'=1}^m \lambda_{j'} \mathbf{z}_{j'} \mathbf{z}_{j'}^T}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_i = \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_{11}}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}) \quad (j = 1, \dots, s_1). \quad (19)$$

Hence, from (18) and (19), we obtain as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\begin{aligned} \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_1}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_i &= \frac{\tilde{\lambda}_i}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{1j}^T \tilde{\mathbf{u}}_i, \\ \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_1}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_i &= \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_{11}}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}) = \frac{\tilde{\lambda}_{1j}}{d^{\alpha_{s_1}}} \tilde{\mathbf{u}}_{1j}^T \tilde{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}) \\ &(j = 1, \dots, s_1; \quad i = s_1 + 1, \dots, s_2). \end{aligned} \quad (20)$$

From the fact that $\text{rank}(d^{-\alpha_{s_1}} \mathbf{U}_1) = \text{rank}(d^{-\alpha_{s_1}} \mathbf{U}_{11}) = s_1$ w.p.1. as $d \rightarrow \infty$, it holds that $d^{-\alpha_{s_1}} \tilde{\lambda}_i = o_p(1)$ as $d \rightarrow \infty$ for $i = s_1 + 1, \dots, m$. Then, one has from (20) that

$$\begin{aligned} \left(\frac{\tilde{\lambda}_{1j}}{d^{\alpha_{s_1}}} + o_p(1) \right) \tilde{\mathbf{u}}_{1j}^T \tilde{\mathbf{u}}_i &= o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}), \quad \text{i.e.} \quad \tilde{\mathbf{u}}_{1j}^T \tilde{\mathbf{u}}_i = o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}) \\ &(j = 1, \dots, s_1; \quad i = s_1 + 1, \dots, s_2). \end{aligned}$$

So, we have as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\tilde{\mathbf{u}}_i^T \frac{\mathbf{U}_{11}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i = \sum_{j=1}^{s_1} \frac{\tilde{\lambda}_{1j}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i^T \tilde{\mathbf{u}}_{1j} \tilde{\mathbf{u}}_{1j}^T \tilde{\mathbf{u}}_i = o_p(n^{-1/2} d^{\alpha_{s_2} - \alpha_{s_1}}) \quad (i = s_1 + 1, \dots, s_2). \quad (21)$$

Then, we obtain that

$$\begin{aligned} \tilde{\mathbf{u}}_i^T \frac{\mathbf{U}_1}{\lambda_i} \tilde{\mathbf{u}}_i &= \tilde{\mathbf{u}}_i^T \frac{\mathbf{U}_1 - \mathbf{U}_{11}}{\lambda_i} \tilde{\mathbf{u}}_i + o_p(n^{-1/2}) \\ &= \tilde{\mathbf{u}}_i^T \left(\sum_{j=s_1+1}^m \frac{\lambda_j}{\lambda_i} \|n^{-1/2} \mathbf{z}_j\|^2 \tilde{\mathbf{z}}_j \tilde{\mathbf{z}}_j^T \right) \tilde{\mathbf{u}}_i + o_p(n^{-1/2}) \quad (i = s_1 + 1, \dots, s_2). \end{aligned} \quad (22)$$

Similarly to (16), it holds (17) for $i = s_1 + 1, \dots, s_2$, as $d \rightarrow \infty$ and $n \rightarrow \infty$. Then, we obtain for $\mathbf{U}_{12} = \sum_{j=1}^{s_2} \tilde{\lambda}_{2j} \tilde{\mathbf{u}}_{2j} \tilde{\mathbf{u}}_{2j}^T$ that

$$\frac{\tilde{\lambda}_{2i}}{\lambda_i} = 1 + o_p(1), \quad \tilde{\mathbf{u}}_{2i}^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}) \quad (i = 1, \dots, s_2). \quad (23)$$

As for λ_i ($i = s_2 + 1, \dots, s_3$) that holds power α_{s_3} , note that $\tilde{\mathbf{u}}_{2j}^T \tilde{\mathbf{z}}_{j'} = o_p(n^{-1/4})$ ($j = 1, \dots, s_2$; $j' = s_2 + 1, \dots, m$) in view of (23). Thus we have that

$$\tilde{\mathbf{u}}_{2j}^T \frac{\mathbf{U}_1}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i = \tilde{\mathbf{u}}_{2j}^T \frac{\mathbf{U}_{12}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_3} - \alpha_{s_2}}) \quad (j = 1, \dots, s_2).$$

Similarly to (20), we have as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\begin{aligned} \tilde{\mathbf{u}}_{2j}^T \frac{\mathbf{U}_1}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i &= \frac{\tilde{\lambda}_i}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_{2j}^T \tilde{\mathbf{u}}_i, \\ \tilde{\mathbf{u}}_{2j}^T \frac{\mathbf{U}_1}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i &= \tilde{\mathbf{u}}_{2j}^T \frac{\mathbf{U}_{12}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_3} - \alpha_{s_2}}) = \frac{\tilde{\lambda}_{2j}}{d^{\alpha_{s_2}}} \tilde{\mathbf{u}}_{2j}^T \tilde{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_3} - \alpha_{s_2}}) \\ &\quad (j = 1, \dots, s_2; i = s_2 + 1, \dots, s_3). \end{aligned}$$

Since it holds for $i = s_2 + 1, \dots, s_3$, that

$$\tilde{\mathbf{u}}_{2j}^T \tilde{\mathbf{u}}_i = \begin{cases} o_p(n^{-1/4} d^{\alpha_{s_3} - \alpha_{s_1}}) & (j = 1, \dots, s_1), \\ o_p(n^{-1/4} d^{\alpha_{s_3} - \alpha_{s_2}}) & (j = s_1 + 1, \dots, s_2), \end{cases}$$

we obtain (17) for $i = s_2 + 1, \dots, s_3$, in a way similar to (21)–(22).

As for λ_i ($i = s_{l-1} + 1, \dots, s_l$) that holds power α_{s_l} ($l \geq 4$) as well, we can obtain (17).

Therefore, for every i ($= 1, \dots, m$), we claim as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\frac{\tilde{\lambda}_i}{\lambda_i} = \|n^{-1/2} \mathbf{z}_i\|^2 + o_p(n^{-1/2}) = 1 + o_p(1), \quad \tilde{\mathbf{u}}_i^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}). \quad (24)$$

It concludes the results. $\square$

Remark 8. When the population eigenvalues are not distinct such as $\lambda_1 \geq \dots \geq \lambda_m$, we consider the case as follows: Suppose that $\lambda_1 = \dots = \lambda_{t_1} > \lambda_{t_1+1} = \dots = \lambda_{t_2} > \dots > \lambda_{t_{r-1}+1} = \dots = \lambda_{t_r}$ ($= \lambda_m$), where $r \leq m$. We can claim as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\frac{\tilde{\lambda}_{t_{j-1}+i}}{\lambda_{t_{j-1}+i}} = 1 + o_p(1), \quad \tilde{\mathbf{u}}_{t_{j-1}+i} \in \left\{ \sum_{i'=1}^{t_j-t_{j-1}} b_{i'} \tilde{\mathbf{z}}_{t_{j-1}+i'} : \sum_{i'=1}^{t_j-t_{j-1}} b_{i'}^2 = 1 \right\}$$

$$(i = 1, \dots, t_j - t_{j-1}; j = 1, \dots, r),$$

where $t_0 = 0$. Thus it holds as $d \rightarrow \infty$ and $n \rightarrow \infty$ that

$$\text{Angle}(\tilde{\mathbf{u}}_{t_{j-1}+i}, \text{span}\{\tilde{\mathbf{z}}_{t_{j-1}+1}, \dots, \tilde{\mathbf{z}}_{t_j}\}) \xrightarrow{p} 0$$

$$(i = 1, \dots, t_j - t_{j-1}; j = 1, \dots, r).$$

Proof of Theorems 1 and 2. Note that $\mathbf{U}_2$ is a symmetric matrix. From Lemmas 3 and 4, we have under (i)–(ii) in Theorem 1 that

$$d^{-\alpha_i} n^{-1} \mathbf{z}_{i'}^T \mathbf{U}_2 \mathbf{z}_{i'} \geq d^{-2\alpha_i} n^{-1} \mathbf{z}_{i'}^T \mathbf{U}_2 \mathbf{U}_2 \mathbf{z}_{i'} = \|d^{-\alpha_i} n^{-1/2} \mathbf{z}_{i'}^T \mathbf{U}_2\|^2 = o_p(n^{-1/2}) \quad (i' = 1, \dots, m).$$

Thus it holds under (i)–(ii) that

$$d^{-\alpha_i} n^{-1/2} \mathbf{z}_{i'}^T \mathbf{U}_2 \mathbf{e}_{1n} = o_p(n^{-1/4}) \quad (i' = 1, \dots, m). \quad (25)$$

Let us first consider the case when $\lambda_1 > \dots > \lambda_m$. For λ_i ($i = 1, \dots, s_1$) that holds power α_{s_1} , we have from Lemma 3 under (i)–(ii) that $\lambda_i^{-1} \mathbf{e}_{1n}^T \mathbf{U}_2 \mathbf{e}_{1n} = o_p(1)$. Then, it holds that $\lambda_1 \|n^{-1/2} \mathbf{z}_1\|^2 > \dots > \lambda_m \|n^{-1/2} \mathbf{z}_m\|^2$ and $\lambda_{s_1} \|n^{-1/2} \mathbf{z}_{s_1}\|^2 > \mathbf{e}_{1n}^T \mathbf{U}_2 \mathbf{e}_{1n}$ w.p.1. Let $\eta_i = \lambda_i^{-1} \hat{\mathbf{u}}_i^T \mathbf{U}_2 \hat{\mathbf{u}}_i$, $i = 1, \dots, m$. Then, in a way similar to (16), with the help of Lemma 4 and (25), it holds under (i)–(ii) that

$$\begin{aligned} \frac{\hat{\lambda}_1}{\lambda_1} &= \hat{\mathbf{u}}_1^T \frac{\mathbf{S}_D}{\lambda_1} \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_1^T \frac{\mathbf{U}_1 + \mathbf{U}_2}{\lambda_1} \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_1^T \left(\sum_{j=1}^m \frac{\lambda_j}{\lambda_1} \|n^{-1/2} \mathbf{z}_j\|^2 \tilde{\mathbf{z}}_j \tilde{\mathbf{z}}_j^T + \frac{\mathbf{U}_2}{\lambda_1} \right) \hat{\mathbf{u}}_1 \\ &= \hat{\mathbf{u}}_1^T \left(\sum_{j=1}^m \frac{\lambda_j}{\lambda_1} \|n^{-1/2} \mathbf{z}_j\|^2 \tilde{\mathbf{z}}_j \tilde{\mathbf{z}}_j^T \right) \hat{\mathbf{u}}_1 + \eta_1 = \|n^{-1/2} \mathbf{z}_1\|^2 + \eta_1 + o_p(n^{-1/2}), \end{aligned} \quad (26)$$

so that $\hat{\mathbf{u}}_1^T \tilde{\mathbf{z}}_1 = 1 + o_p(n^{-1/2})$ and $\eta_1 = \lambda_1^{-1} \tilde{\mathbf{z}}_1^T \mathbf{U}_2 \tilde{\mathbf{z}}_1 + o_p(n^{-1/2}) = o_p(n^{-1/2})$. Under (i)–(ii), in a way similar to Lemma 7 and (26), we claim that

$$\frac{\hat{\lambda}_i}{\lambda_i} = \|n^{-1/2} \mathbf{z}_i\|^2 + o_p(n^{-1/2}) = 1 + o_p(1), \quad \hat{\mathbf{u}}_i^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}) \quad (i = 1, \dots, s_1). \quad (27)$$

Here, as for Theorem 2, recall that $V(z_{ij}^2) = M_i$, $i = 1, \dots, m$. For each i , by using the central limiting theorem, one has as $n \rightarrow \infty$ that $(nM_i)^{-1/2}(\sum_{j=1}^n z_{ij}^2 - n) \Rightarrow N(0, 1)$. Hence, under (i)–(ii), we have from (27) that

$$\sqrt{\frac{n}{M_i}} \left(\frac{\hat{\lambda}_i}{\lambda_i} - 1 \right) \Rightarrow N(0, 1) \quad (i = 1, \dots, s_1). \quad (28)$$

For λ_i ($i = s_1 + 1, \dots, s_2$) that holds power α_{s_2} , let us denote $\eta_{ij} = \lambda_i^{-1} \tilde{\mathbf{u}}_{1j}^T \mathbf{U}_2 \hat{\mathbf{u}}_i$, $j = 1, \dots, s_1$. Then, from (18) and (25), it holds under (i)–(ii) that

$$\eta_{ij} = \lambda_i^{-1} \tilde{\mathbf{z}}_j^T \mathbf{U}_2 \hat{\mathbf{u}}_i + o_p(n^{-1/4}) = o_p(n^{-1/4}) \quad (j = 1, \dots, s_1; i = s_1 + 1, \dots, s_2).$$

Thus we have under (i)–(ii) that

$$\begin{aligned} \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{S}_D}{d^{\alpha_{s_1}}} \hat{\mathbf{u}}_i &= \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_{11}}{d^{\alpha_{s_1}}} \hat{\mathbf{u}}_i + \eta_{ij} O(d^{\alpha_{s_2} - \alpha_{s_1}}) + o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}) \\ &= \tilde{\mathbf{u}}_{1j}^T \frac{\mathbf{U}_{11}}{d^{\alpha_{s_1}}} \hat{\mathbf{u}}_i + o_p(n^{-1/4} d^{\alpha_{s_2} - \alpha_{s_1}}) \quad (j = 1, \dots, s_1; i = s_1 + 1, \dots, s_2). \end{aligned}$$

Then, similarly to (20)–(21), we claim that

$$\hat{\mathbf{u}}_i^T \frac{\mathbf{U}_{11}}{d^{\alpha_{s_2}}} \hat{\mathbf{u}}_i = o_p(n^{-1/2} d^{\alpha_{s_2} - \alpha_{s_1}}) = o_p(n^{-1/2}) \quad (i = s_1 + 1, \dots, s_2).$$

Hence, it holds for $i = s_1 + 1, \dots, s_2$, that

$$\begin{aligned} \hat{\mathbf{u}}_i^T \frac{\mathbf{S}_D}{\lambda_i} \hat{\mathbf{u}}_i &= \hat{\mathbf{u}}_i^T \frac{\mathbf{U}_1 - \mathbf{U}_{11}}{\lambda_i} \hat{\mathbf{u}}_i + \eta_i + o_p(n^{-1/2}) \\ &= \hat{\mathbf{u}}_i^T \left(\sum_{j=s_1+1}^m \frac{\lambda_j}{\lambda_i} \|n^{-1/2} \mathbf{z}_j\|^2 \tilde{\mathbf{z}}_j \tilde{\mathbf{z}}_j^T \right) \hat{\mathbf{u}}_i + \eta_i + o_p(n^{-1/2}). \end{aligned}$$

Under (i)–(ii), in a way similar to (26), it holds (27) and (28) for $i = s_1 + 1, \dots, s_2$, as well. Similarly, for λ_i ($i = s_{l-1} + 1, \dots, s_l$) that holds power α_{s_l} ($l \geq 3$), we can obtain (27) and

(28). Hence, for each i ($= 1, \dots, m$), we can claim under (i)–(ii) that

$$\begin{aligned} \frac{\hat{\lambda}_i}{\lambda_i} &= \|n^{-1/2} \mathbf{z}_i\|^2 + o_p(n^{-1/2}) = 1 + o_p(1), \quad \hat{\mathbf{u}}_i^T \tilde{\mathbf{z}}_i = 1 + o_p(n^{-1/2}) \\ \text{and } \sqrt{\frac{n}{M_i}} \left(\frac{\hat{\lambda}_i}{\lambda_i} - 1 \right) &\Rightarrow N(0, 1). \end{aligned} \quad (29)$$

It concludes the result in Theorem 2.

Next, we consider the case when $\lambda_1 \geq \dots \geq \lambda_m$. One may refer to Remark 8. Since we can claim that $\hat{\lambda}_i/\lambda_i = 1 + o_p(1)$ for $i = 1, \dots, m$, under (i)–(ii), in a way similar to (29), it concludes the result in Theorem 1. $\square$

Proof of Corollary 1. By using Chebyshev's inequality, for any $\tau > 0$ and the uniform bound M for the fourth moments condition, one has under (i)–(ii) in Corollary 1 that

$$\sum_{k=1}^n P\left((nd^{\alpha_i})^{-1} \left| \sum_{s=m+1}^d \lambda_s z_{sk}^2 - \sum_{s=m+1}^d \lambda_s \right| > \tau\right) = O(d^{1-2\alpha_i}/n) = o(1).$$

Thus it holds that $d^{-\alpha_i} u_k = d^{-\alpha_i} n^{-1} \sum_{s=m+1}^d \lambda_s + o_p(1) = d^{1-\alpha_i}/n + o_p(1) = o_p(1)$ for all k ($= 1, \dots, n$). Let us write under (ii) that $d^{1-\alpha_i}/n = d^{\alpha_i} d^{1-2\alpha_i}/n \rightarrow 0$. One has under (ii) that $d^{1-2\alpha_i}/n \leq d^{-\alpha_i}$ to claim Lemma 4 under (ii). From Lemmas 2 and 6, we claim Lemma 3 under (i)–(ii). Here, it should be noted that Lemma 4 cannot be claimed under (i)–(ii) in Corollary 1. Hence, we obtain under (i)–(ii) that

$$\frac{\hat{\lambda}_i}{\lambda_i} = \|n^{-1/2} \mathbf{z}_i\|^2 + o_p(1) = 1 + o_p(1) \quad (i = 1, \dots, m)$$

in a way similar to the proof of Theorem 1. $\square$

Proof of Theorem 3. From Remark 4, we note that $\hat{\mathbf{h}}_i = (n\hat{\lambda}_i)^{-1/2} \mathbf{X} \hat{\mathbf{u}}_i$. Then, we have that

$$\mathbf{h}_i^T \hat{\mathbf{h}}_i = (\hat{\lambda}_i n)^{-1/2} \lambda_i^{1/2} \mathbf{z}_i^T \hat{\mathbf{u}}_i = \left(\frac{\lambda_i}{\hat{\lambda}_i} \right)^{1/2} \frac{\mathbf{z}_i^T}{\sqrt{n}} \hat{\mathbf{u}}_i.$$

From the proof of Theorem 1, we claim (29) for $i = 1, \dots, m$, under (i)–(ii) in Theorem 1.

Thus we obtain that

$$\mathbf{h}_i^T \hat{\mathbf{h}}_i = \left(\frac{\lambda_i}{\hat{\lambda}_i} \right)^{1/2} \frac{\mathbf{z}_i^T}{\sqrt{n}} \hat{\mathbf{u}}_i = 1 + o_p(1) \quad (i = 1, \dots, m),$$

so that $\text{Angle}(\mathbf{h}_i, \hat{\mathbf{h}}_i) = o_p(1)$ for $i = 1, \dots, m$. □

Proof of Theorem 4. For each i ($= 1, \dots, n$), let us write that

$$\begin{aligned} V(\hat{s}_i) &= \lambda_i n^{-1} \sum_{k=1}^n \left(z_{ik} - \sqrt{n \frac{\hat{\lambda}_i}{\lambda_i}} \hat{u}_{ik} \right)^2 \\ &= \lambda_i \left(n^{-1} \sum_{k=1}^n z_{ik}^2 + \frac{\hat{\lambda}_i}{\lambda_i} \sum_{k=1}^n \hat{u}_{ik}^2 - 2 \sqrt{\frac{\hat{\lambda}_i}{\lambda_i}} \frac{\mathbf{z}_i^T}{\sqrt{n}} \hat{\mathbf{u}}_i \right). \end{aligned}$$

We have (29) under (i)–(ii) in Theorem 1. Noting that $n^{-1} \sum_{k=1}^n z_{ik}^2 = 1 + o_p(1)$ as $n \rightarrow \infty$ for each i ($= 1, \dots, m$), it concludes the result. □

Proof of Corollaries 2 and 3. When z_{ij} , $i = 1, \dots, d$ ($i = 1, \dots, n$) are independent and $\lambda_1 > \dots > \lambda_m$, one can claim for i ($= 1, \dots, m$) under (i)–(ii) in Corollary 1 that

$$\frac{\hat{\lambda}_i}{\lambda_i} = 1 + o_p(1), \quad \hat{\mathbf{u}}_i^T \tilde{\mathbf{z}}_i = 1 + o_p(1).$$

The results are straightforwardly obtained in a way similar to the proofs of Theorems 3 and 4. □

Proof of Corollary 4. In view of (29), we obtain the result straightforwardly. □

Acknowledgements The authors would like to thank anonymous referees for their valuable comments. Research of the second author was partially supported by Grant-in-Aid for Scientific Research (B), Japan Society for the Promotion of Science (JSPS), under Contract Number 18300092.

References

- Ahn, J., Marron, J. S., Muller, K. M. and Chi, Y.-Y. (2007). The high-dimension, low-sample-size geometric representation holds under mild conditions. *Biometrika* **94**, 760–766.

- Baik, J., Ben Arous, G., and P, S. (2005). Phase transition of the largest eigenvalue for non-null complex covariance matrices. *Ann. Probab.* **33**, 1643–1697.
- Baik, J. and Silverstein, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Mult. Anal.* **97**, 1382–1408.
- Hall, P., Marron, J. S. and Neeman, A. (2005). Geometric representation of high dimension, low sample size data. *J. R. Statist. Soc.*, **B 67**, 427–444.
- Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, **29**, 295–327.
- Johnstone, I. and Lu, A. (2008). Sparse principal component analysis. *J. Amer. Statist. Assoc.*, to appear.
- Jung, S. and Marron, J. S. (2008). PCA consistency in high dimension, low sample size context. *Ann. Statist.*, revised.
- Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica* **17**, 1617–1642.
- Yata, K. and Aoshima, M. (2008). Intrinsic dimensionality estimation of high dimension, low sample size data with d -asymptotics. *Commun. Statist.-Theory Meth.*, revised.