

# Asymptotic properties of principal component analysis and shrinkage-bias adjustment under the generalized spiked population model

Rounak Dey, Seunggeun Lee\*

Department of Biostatistics, University of Michigan School of Public Health, 1415 Washington Heights, Ann Arbor, MI 48109-2029, USA

## ARTICLE INFO

### Article history:

Received 19 January 2018

Received in revised form 14 February 2019

Accepted 14 February 2019

Available online 19 February 2019

### AMS 2010 subject classifications:

primary 62H25

62H12

secondary 15A18

15A52

### Keywords:

Consistent estimation

High-dimensional data

PC scores

Random matrix

## ABSTRACT

With the development of high-throughput technologies, principal component analysis (PCA) in the high-dimensional regime is of great interest. Most of the existing theoretical and methodological results for high-dimensional PCA are based on the spiked population model in which all the population eigenvalues are equal except for a few large ones. Due to the presence of local correlation among features, however, this assumption may not be satisfied in many real-world datasets. To address this issue, we investigate the asymptotic behavior of PCA under the generalized spiked population model. Based on our theoretical results, we propose a series of methods for the consistent estimation of population eigenvalues, angles between the sample and population eigenvectors, correlation coefficients between the sample and population principal component (PC) scores, and the shrinkage bias adjustment for the predicted PC scores. Using numerical experiments and real data examples from the genetics literature, we show that our methods can greatly reduce bias and improve prediction accuracy.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

Principal component analysis (PCA) is a very popular tool for analyzing high-dimensional biomedical data, where the number  $p$  of features is often substantially larger than the number  $n$  of observations. PCA is widely used to adjust for population stratification in genome-wide association studies [21] and to identify overall expression patterns in transcriptome analysis [23]. However, the asymptotic properties of PCA in high-dimensional data are profoundly different from the properties in low-dimensional ( $p$  finite,  $n \rightarrow \infty$ ) settings. In high-dimensional settings, the sample eigenvalues and eigenvectors are not consistent estimators of the population eigenvalues and eigenvectors [12,20], and the predicted principal component (PC) scores based on the sample eigenvectors can be systematically biased toward zero [14].

There has been extensive effort to investigate the asymptotic behavior of PCA in high-dimensional settings. To provide a statistical framework for PCA in these settings, Johnston introduced a spiked population model, which assumes that all the eigenvalues are equal except for finitely many large ones called the spikes. A spiked population covariance matrix is basically a finite-rank perturbation of a scalar multiple of the identity matrix. A typical example of a spiked population with two spikes is shown in Fig. 1a. This two-spike eigenvalue structure arises if the population consists of three sub-populations, and the features are largely independent with equal variances. Under this model, convergence of sample eigenvalues, eigenvectors and PC scores have been extensively studied [4,11,14,20].

\* Corresponding author.

E-mail address: [leeshawn@umich.edu](mailto:leeshawn@umich.edu) (S. Lee).

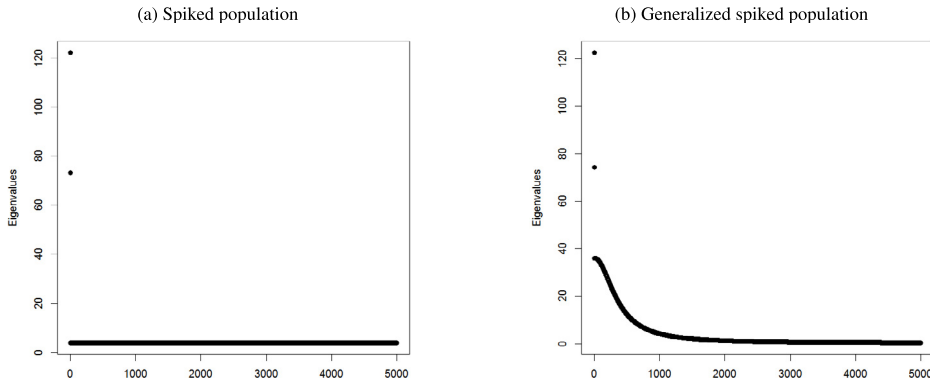


Fig. 1. Eigenvalue structures for SP and GSP models.

In many biomedical data, however, the assumption of the equality of non-spiked eigenvalues can be violated due to the presence of local correlation among features. In genome-wide association studies, for example, the genetic variants are locally correlated due to linkage disequilibrium. In gene-expression data, since genes in the same pathway are often expressed together, their expression measurements are often correlated. These local correlations can cause substantial differences in non-spiked eigenvalues. To illustrate this phenomenon, we obtained eigenvalues with an autoregressive within-group correlation structure rather than the independent structure of the previous example. Fig. 1b shows that the equality assumption is clearly violated. Thus, if methods developed under the equality assumption are applied to these types of data, we will obtain biased results.

The generalized spiked population model [3] has been proposed to address this problem. The condition that the non-spikes have to be equal is removed in this generalization. In this model, the set of population eigenvalues consists of finitely many large eigenvalues called the generalized spikes, which are well separated from infinitely many small eigenvalues. Although the generalized spiked population model has a great potential to provide more accurate inference in high-dimensional biomedical data, only limited literature is available on the asymptotic properties of PCA under this model and their application to real data. Bai and Yao [3] and Ding [8] provided results regarding convergence of eigenvalues and eigenvectors. However, their work remained largely theoretical. Moreover, to the best of our knowledge, no method has been developed for estimating the correlations between the sample and population PC scores, and adjusting biases in the predicted PC scores under the generalized spiked population model.

In this paper, we systematically investigate the asymptotic behavior of PCA under the generalized spiked population model, and develop methods to estimate the population eigenvalues and adjust for the bias in the predicted PC scores. We first propose two different approaches to consistently estimate the population eigenvalues, the angles between the sample and population eigenvectors, and the correlation coefficients between the sample and population PC scores. We compare these two methods and show the asymptotic equivalence of the estimators across them. Finally, we propose a method to reduce the bias in the predicted PC scores based on the estimated population eigenvalues.

The paper is organized as follows. We begin in Section 2 by providing the definition of the generalized spiked population model and present existing theoretical results. We develop our methods to consistently estimate the population spikes in Section 3. In Section 4, we construct consistent estimators of the angles between the sample and population eigenvectors, and the correlation coefficients between the sample and population PC scores. We also propose the bias-reduction technique for the predicted PC scores. Section 5 presents the algorithm [9] to estimate the population limiting spectral distribution and the non-spiked eigenvalues. In Section 6, we present results from simulation studies and an example from the Hapmap project to demonstrate the improved performance of our method over the existing one. Finally, we conclude the paper with a discussion.

## 2. Generalized spiked population model

In order to formally define generalized spiked population model, we require the concept of spectral distribution. In the random matrix literature, it is natural to associate a probability measure to the set of eigenvalues as the dimension  $p$  goes to infinity. More explicitly, if a Hermitian matrix  $\Sigma_p$  has eigenvalues  $\lambda_1, \dots, \lambda_p$ , we can define the empirical spectral distribution (ESD) of  $\Sigma_p$  to be  $H_p$  based on the probability measure

$$dH_p(x) = \frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i}(x),$$

where  $\delta_{\lambda_i}(x)$  is unity when  $x = \lambda_i$ , and otherwise zero. Now, for a sequence  $(\Sigma_p)$  of covariance matrices, if the corresponding sequence  $(H_p)$  of ESDs converges weakly to a non-random probability distribution  $H$  as  $p \rightarrow \infty$ , then we define  $H$  as the limiting spectral distribution (LSD) of the sequence  $(\Sigma_p)$ .

The generalized spiked population model [3] is defined as follows. Suppose that  $H_p$  is the ESD corresponding to the population covariance matrix  $\Sigma_p$  and that it converges weakly to a non-random probability distribution  $H$ . Let  $\Gamma_H$  be the support of  $H$  and  $d(x, A) = \inf_{y \in A} |x - y|$  be the distance metric from a point  $x$  to a set  $A$ . Then the set of eigenvalues of  $\Sigma_p$  comprises of two subsets of eigenvalues  $\alpha_1 \geq \dots \geq \alpha_m$  and  $\beta_{p,1} \geq \dots \geq \beta_{p,p-m}$ , as follows:

- (i) there exists  $\delta > 0$  such that  $d(\alpha_i, \Gamma_H) > \delta$  for all  $i \in \{1, \dots, m\}$ ;  $\alpha_1, \dots, \alpha_m$  are called the generalized spikes.
- (ii)  $\max_{1 \leq i \leq p-m} d(\beta_{p,i}, \Gamma_H) = \epsilon_p \rightarrow 0$ ;  $\beta_{p,1}, \dots, \beta_{p,p-m}$  are called the non-spikes.

It is obvious from the definition that the generalized spikes are measure zero points of the population LSD. For Johnstone's spiked population model [11], the population LSD is  $H = \delta_{\{1\}}$ , indicating  $\Gamma_H = \{1\}$ . From the definition above, all eigenvalues larger than 1 are spikes. Hence, Johnstone's spiked population model is a special case of the generalized spiked population model.

Suppose that the population covariance matrix  $\Sigma_p$  has eigenvalues  $\lambda_1 \geq \dots \geq \lambda_p$ , and the sample covariance matrix  $S_p = X^\top X/n$  has eigenvalues  $d_1 \geq \dots \geq d_p$ , where  $X$  is an  $n \times p$  data matrix. Let  $\Sigma_p = E \Lambda_p E^\top$  and  $S_p = U D_p U^\top$  be the spectral decompositions of  $\Sigma_p$  and  $S_p$ , respectively. Further, we will assume the following throughout the paper.

- (A)  $p \rightarrow \infty, n \rightarrow \infty, p/n \rightarrow \gamma < \infty$ .
- (B)  $X = Y \Lambda_p^{1/2} E^\top$ , where  $Y$  is an  $n \times p$  random matrix with iid elements such that  $E(Y_{ij}) = 0$ ,  $E(|Y_{ij}|^2) = 1$ ,  $E(|Y_{ij}|^4) < \infty$ .
- (C) The population eigenvalues follow the generalized spiked population (GSP) model with  $m$  generalized spikes. Let  $H_p$  be the population ESD,  $H$  be the population LSD, and  $\Gamma_H$  be the support of  $H$ . Moreover, the sequence  $(\|\Sigma_p\|)$  of spectral norms is bounded.
- (D) The generalized spikes are larger than  $\sup \Gamma_H$ .

Even though we will develop our estimation methods based on the asymptotic regime where  $p/n \rightarrow \gamma < \infty$ , we will discuss the applicability of our methods when  $p$  is greatly larger than  $n$  in Section 4.6.

In this paper, we will derive and discuss asymptotic properties of different functions of eigenvalues and angles between the sample and population eigenvectors, both of which are rotation-invariant, i.e., if we rotate  $X$  by some  $p \times p$  orthogonal matrix  $P$ , the eigenvalues (both sample and population) and angles between the sample and population eigenvectors will remain the same. The new population covariance matrix  $\tilde{\Sigma}_p = P^\top \Sigma_p P$  will have the same eigenvalues as  $\Sigma_p$ , and the new sample covariance matrix  $\tilde{S}_p = P^\top S_p P$  will also have the same eigenvalues as  $S_p$ . The angles between the eigenvectors of  $\tilde{S}_p$  and  $\tilde{\Sigma}_p$  will be the same as the angles between the eigenvectors of  $S_p$  and  $\Sigma_p$ , both given by the elements of the matrix  $E^\top U$ . Therefore, without loss of generality, by using  $P = E$ , we can assume the population covariance matrix to be diagonal.

From the Marčenko–Pastur theorem [16], the sample ESD  $F_p$  converges weakly to a non-random probability distribution  $F$  with support  $\Gamma_F$ . For  $\alpha \notin \Gamma_H$ ,  $\alpha \neq 0$  and  $x > 0$ , we define the following two functions:

$$\psi(\alpha) = \alpha + \gamma \alpha \int \frac{\lambda dH(\lambda)}{\alpha - \lambda}, \quad f_F(x) = x / \left\{ 1 + \gamma \int \frac{\tau dF(\tau)}{x - \tau} \right\}. \quad (1)$$

The following result by Bai and Yao [3] provides the almost sure limits of the sample eigenvalues corresponding to the population generalized spikes.

**Result 1.** Suppose assumptions (A)–(D) hold. Let  $\lambda_k$  be a generalized spike of multiplicity 1 and the corresponding sample eigenvalue is  $d_k$ . Moreover, let  $\psi'$  denote the first derivative of the function  $\psi$ . Then,

- (i) If  $\psi'(\lambda_k) > 0$ , then the sample eigenvalue  $d_k$  converges almost surely to  $\psi(\lambda_k)$ , i.e.,  $|d_k - \psi(\lambda_k)| \xrightarrow{a.s.} 0$ .
- (ii) If  $\psi'(\lambda_k) \leq 0$ , then let  $(u_k, v_k) \subset (\sup \Gamma_H, \infty)$  be the maximal interval on which  $\psi' > 0$ . The sample eigenvalue  $d_k$  converges almost surely to  $\psi(w)$ , where  $w$  is a boundary of  $[u_k, v_k]$  that is nearest to  $\lambda_k$ .

Given that  $\psi'(\alpha)$  is a strictly increasing function for  $\alpha > \sup \Gamma_H$ , if a generalized spike  $\lambda_k$  is large enough such that  $\psi'(\alpha) > 0$ , according to Result 1 the corresponding sample eigenvalue will converge almost surely to  $\psi(\lambda_k)$ . However if the generalized spike lies close enough, i.e.,  $\psi'(\lambda_k) \leq 0$ , to the set of non-spikes, then the convergence of the corresponding sample eigenvalue is given by the second part of the result. We will denote a generalized spike  $\lambda_k$  as a “distant spike” if  $\psi'(\lambda_k) > 0$ ; otherwise we will call it a “close spike”.

### 3. Consistent estimation of the generalized spikes

The following theorem provides two different consistent estimators of the distant spikes.

**Theorem 1.** Let  $\lambda_k$  be a distant spike of multiplicity 1 and the corresponding sample eigenvalue is  $d_k$ . If the assumptions (A)–(D) hold, then

$$|\psi^{-1}(d_k) - \lambda_k| \xrightarrow{P} 0,$$

where  $\psi^{-1}$  is the left inverse of  $\psi$ . Also,  $|f_F(d_k) - \lambda_k| \xrightarrow{P} 0$ .

This theorem shows that for any distant spike  $\lambda_k$  we have two consistent estimators  $\psi^{-1}(d_k)$  and  $f_F(d_k)$ . Notice that the function  $f_F$  depends only on the sample LSD which can be approximated by the sample ESD. Thus,  $f_F(d_k)$  can be approximated directly using the sample eigenvalues. More explicitly,  $f_F(d_k)$  can be closely approximated as

$$f_F(d_k) \approx d_k / \left\{ 1 + \frac{\gamma}{p-m} \sum_{i=m+1}^p \frac{d_i}{d_k - d_i} \right\}.$$

In contrast, the  $\psi$  function depends on the population LSD which is unknown. We can estimate the  $\psi$  function using the algorithm described in Section 5 and then find the inverse function  $\psi^{-1}$  using a Newton–Raphson type algorithm.

#### 4. Consistent estimators of the asymptotic shrinkage in predicting the PC scores

In this section, we investigate the convergence of sample eigenvectors, PC scores, and shrinkage factors in predicting the PC scores. Let  $e_i$  and  $E_i$  be the  $i$ th sample and population eigenvectors, respectively. In addition to assumptions (A)–(D), we further assume that the distant spikes are of multiplicity 1. This assumption is to restrict the dimension of the corresponding eigenspaces to 1, as otherwise the angle between sample and population eigenvectors, or shrinkage in predicted PC scores cannot be well defined.

##### 4.1. Angle between sample and population eigenvectors

We first present the following theorem on the convergence of the quadratic forms of the sample eigenvectors.

**Theorem 2.** Let  $\lambda_k$  be a distant spike of multiplicity 1, and the assumptions (A)–(D) hold. Consider the quadratic form  $\hat{\eta}_k = s_1^\top e_k e_k^\top s_2$ , where  $s_1$  and  $s_2$  are non-random vectors with uniformly bounded norm for all  $p$ . Then

$$|\hat{\eta}_k - \eta_k| \xrightarrow{a.s.} 0,$$

where  $\eta_k = \lambda_k \psi'(\lambda_k) s_1^\top E_k E_k^\top s_2 / \psi(\lambda_k)$ .

Mestre [19] showed similar asymptotic properties of the quadratic forms under the assumption that the number of spikes increases with the dimension. Theorem 2 shows the convergence of the angle between sample and population eigenvectors. Suppose  $s_1 = s_2 = E_k$ . Then

$$\hat{\eta}_k = E_k^\top e_k e_k^\top E_k = \langle e_k, E_k \rangle^2, \quad \eta_k = \lambda_k \psi'(\lambda_k) / \psi(\lambda_k).$$

Combining them, we can show

$$|\langle e_k, E_k \rangle^2 - \lambda_k \psi'(\lambda_k) / \psi(\lambda_k)| \xrightarrow{a.s.} 0.$$

Therefore,  $\{\lambda_k \psi'(\lambda_k) / \psi(\lambda_k)\}^{1/2}$  is a consistent estimator of the cosine of the angle, i.e., the absolute value of the inner product, between the  $k$ th sample and population eigenvectors. In order to obtain this estimator, we first need to estimate the  $\psi$  function using the algorithm described in Section 5.

The following result by [8] provides another consistent estimator for the angle between the  $k$ th sample and population eigenvectors. The proof of the asymptotic equivalence of these two estimators is given in the Appendix A.

**Result 2.** Let  $\lambda_k$  be a distant spike of multiplicity 1, and  $d_k$  be the corresponding sample eigenvalue. Assume that (A)–(D) hold. Define,

$$g_F(x) = \left\{ 1 + \gamma f_F(x) \int \frac{\tau}{(x - \tau)^2} dF(\tau) \right\}^{-1}.$$

Then  $|\langle e_k, E_k \rangle^2 - g_F(d_k)| \xrightarrow{p} 0$ .

Hence  $g_F(d_k)^{1/2}$  also works as a consistent estimator of  $|\langle e_k, E_k \rangle|$ . Since the function  $g_F$  depends only on sample LSD, it can be approximated directly using sample eigenvalues. More explicitly, if there are  $m$  spikes in the population, the function  $g_F$  can be closely approximated as

$$g_F(d_k) \approx \left\{ 1 + \frac{\gamma f_F(d_k)}{p-m} \sum_{i=m+1}^p \frac{d_i}{(d_k - d_i)^2} \right\}^{-1}.$$

The above equation can be used to estimate the angle between the sample and population eigenvectors.

#### 4.2. Correlation between sample and population PC scores

The sample and population PC scores are the projections of the data on the sample and population eigenvectors, respectively. The correlation between them can be perceived as a measure of accuracy of the PCA. The squared correlation can also be interpreted as the proportion of variance in the population PC scores that can be explained by corresponding sample PC scores. The following theorem provides the consistent estimators of the correlation between the sample and population PC scores corresponding to a distant spike.

**Theorem 3.** Suppose  $\lambda_k$  is a distant spike of multiplicity 1,  $d_k$  is the corresponding sample eigenvalue, and the assumptions (A)–(D) hold. Let the normalized  $k$ th population PC score be  $P_k = XE_k/(n\lambda_k)^{1/2}$  and the normalized  $k$ th sample PC score be  $p_k = Xe_k/(nd_k)^{1/2}$ . Then

$$|\langle P_k, p_k \rangle^2 - \psi'(\lambda_k)| \xrightarrow{p} 0$$

and

$$|\langle P_k, p_k \rangle^2 - d_k g_F(d_k)/f_F(d_k)| \xrightarrow{p} 0,$$

where the function  $g_F$  is as defined in [Result 2](#).

Since  $P_k$  and  $p_k$  are normalized random vectors, the absolute value of the inner product  $\langle P_k, p_k \rangle$  is identical to the absolute value of their correlation coefficient. Since correlation is scale invariant, this is also the correlation between  $k$ th sample and population PC scores. Therefore we can consider both  $\psi'(\lambda_k)^{1/2}$  and  $\{d_k g_F(d_k)/f_F(d_k)\}^{1/2}$  to be consistent estimators of the correlation between the  $k$ th sample and population PC scores.

#### 4.3. Asymptotic shrinkage factor

Suppose  $\lambda_k$  is a distant spike. Let the  $k$ th sample PC score for the  $j$ th observation  $x_j$  be  $p_{kj} = x_j^\top e_k$ , and the  $k$ th predicted PC score for a new observation  $x_{new}$  be  $q_k = x_{new}^\top e_k$ . Then the quantity  $\rho_k = \lim_{p \rightarrow \infty} \{E(q_k^2)/E(p_{kj}^2)\}^{1/2}$  describes the asymptotic shrinkage in the  $k$ th predicted PC score for a new observation. As both  $p_{kj}$  and  $q_k$  are centered, i.e.,  $E(p_{kj}) = E(q_k) = 0$ ,  $\rho_k$  represents the limiting ratio of the standard deviations of the predicted PC scores and the sample PC scores. Therefore, if we can estimate  $\rho_k$ , then the shrinkage bias in the  $k$ th predicted PC scores can be easily adjusted by rescaling the predicted scores by the factor  $\rho_k^{-1}$ . The following theorem provides the consistent estimator of the asymptotic shrinkage factor  $\rho_k$ .

**Theorem 4.** Suppose  $\lambda_k$  is a distant spike of multiplicity 1,  $d_k$  is the corresponding sample eigenvalue, and the assumptions (A)–(D) hold. Let  $p_{kj}$  and  $q_k$  be as defined above. Then

$$\left| \sqrt{E(q_k^2)/E(p_{kj}^2)} - \lambda_k/d_k \right| \xrightarrow{p} 0.$$

This is a surprising result in which the asymptotic shrinkage factor is expressed as a simple ratio of the population and sample eigenvalues. Recall that we already constructed the consistent estimators for population eigenvalues in the previous sections. Using these results, the asymptotic shrinkage factor  $\rho_k$  can be consistently estimated by  $\hat{\lambda}_k/d_k$ , where  $\hat{\lambda}_k$  is any consistent estimator of  $\lambda_k$ .

#### 4.4. Comparison of the two different estimators

For each of the quantities discussed above, we proposed two asymptotically equivalent estimators. In terms of practical applications, they have their own advantages and disadvantages. One of them can be approximated directly based only on the sample eigenvalues, while the other one requires to estimate the LSD of the population eigenvalues to obtain the  $\psi$  function. For ease of discourse, we will call the former “ $d$ -estimator” and the later “ $\lambda$ -estimator”.

If the number of spikes is known, estimating the  $d$ -estimator is computationally more efficient than estimating the  $\lambda$ -estimator as it does not involve estimating the population LSD. However, by estimating the population LSD the  $\lambda$ -estimation procedure can verify whether an estimated eigenvalue is actually a distant spike by checking if  $\psi' > 0$ . Thus it can be used to estimate the number of distant spikes when it is unknown; see [Section 5](#). In contrast, the  $d$ -estimation procedure provides no information on the population LSD and thus cannot distinguish among distant spikes, close spikes and non-spikes.

To summarize, when the number of spikes is known or we only want to estimate few of the largest eigenvalues which are known to be distant spikes, then the  $d$ -estimation procedure has the advantage of a faster computation, while the  $\lambda$ -estimation procedure is more useful when the number of spikes is unknown or the distribution of the non-spikes is of interest.

#### 4.5. Comparison of the generalized spiked population (GSP) model and the spiked population (SP) model

As mentioned before, the SP model [11] is a special case of the GSP model. It is easy to verify that when the population eigenvalues follow the SP model, our consistent estimators for the spiked eigenvalues, the angles between the eigenvectors, the correlation coefficients between the PC scores and the shrinkage factors conform to the consistent estimators derived by [4,14,20]. For an SP model where all the non-spikes are equal to 1, the LSD  $H$  is a degenerate distribution at 1, and

$$\psi(\alpha) = \alpha\{1 + \gamma/(\alpha - 1)\}, \quad \psi'(\alpha) = 1 - \gamma/(\alpha - 1)^2.$$

Now,  $\psi'(\alpha) > 0$  if and only if  $\alpha > 1 + \gamma^{1/2}$ . If  $\alpha > 1 + \gamma^{1/2}$  and  $d$  is the corresponding sample eigenvalue, then the consistent estimator of  $\alpha$  is given by  $\psi^{-1}(d)$ , and

$$\frac{\alpha\psi'(\alpha)}{\psi(\alpha)} = \frac{1 - \gamma/(\alpha - 1)^2}{1 + \gamma/(\alpha - 1)}, \quad \frac{\alpha}{\psi(\alpha)} = \frac{\alpha - 1}{\alpha + \gamma - 1},$$

which show that all our results match with the results from [14].

It is of interest to investigate how closely methods developed under the SP model can approximate the consistent estimators for the distant spikes when the population eigenvalues actually follow a GSP model. Suppose the population eigenvalues  $\lambda_1 \geq \dots \geq \lambda_p$  follow the GSP model with  $m$  distant spikes. The sample eigenvalues are  $d_1 \geq \dots \geq d_p$ . Let  $\lambda_k$  be a distant spike with multiplicity 1, and the corresponding sample eigenvalue is  $d_k$ . Then according to Result 1,  $d_k \rightarrow \psi(\lambda_k)$  almost surely. From the definition of  $\psi$ ,

$$\psi(\lambda_k) = \lambda_k \left\{ 1 + \gamma \int \frac{\lambda}{\lambda_k - \lambda} dH(\lambda) \right\} = \lambda_k + \gamma \int \frac{\lambda}{1 - \lambda/\lambda_k} dH(\lambda).$$

If  $H$  is almost degenerate, i.e., the non-spikes are nearly identical, then

$$\psi(\lambda_k) \approx \lambda_k + \frac{\gamma \bar{\lambda}}{1 - \bar{\lambda}/\lambda_k}, \quad (2)$$

where  $\bar{\lambda} = \int \lambda dH(\lambda)$  is the mean of the population LSD which can be closely approximated by the mean of the non-spikes. By contrast, if the spike  $\lambda_k$  is very large compared to all the non-spikes such that  $\lambda/\lambda_k \approx 0$  for any  $\lambda \in \Gamma_H$ , then

$$\psi(\lambda_k) \approx \lambda_k + \gamma \bar{\lambda}.$$

Now, suppose that instead of using the GSP assumption, we use the SP assumption to estimate the distant spikes. We assume that under the SP model, the population covariance matrix is scaled by a factor  $\zeta$  and the population eigenvalues are  $\beta_1 \geq \dots \geq \beta_m > \zeta = \dots = \zeta$ . If  $\beta_k$  is the population eigenvalue corresponding to  $d_k$ , then  $d_k \rightarrow \psi(\beta_k)$  almost surely, where

$$\psi(\beta_k) = \beta_k \left( 1 + \gamma \frac{\zeta}{\beta_k - \zeta} \right) = \beta_k + \frac{\gamma \zeta}{1 - \zeta/\beta_k}.$$

Here  $\zeta$  is estimated as the mean of the non-spikes as they are all assumed to be equal to  $\zeta$ . Notice that this expression is approximately equal to the expression in (2) with  $\beta_k = \lambda_k$  and  $\zeta = \bar{\lambda}$ . Therefore, the asymptotic limit of  $d_k$  under both the GSP and the SP model are approximately equal when the non-spikes are nearly identical. In contrast, when the spike  $\beta_k$  is very large compared to all the non-spikes such that  $\zeta/\beta_k \approx 0$ , then  $\psi(\beta_k) \approx \beta_k + \gamma \zeta$ . In this case also, the asymptotic limit of  $d_k$  under both the GSP and the SP model are approximately equal with  $\beta_k = \lambda_k$  and  $\zeta = \bar{\lambda}$ . Therefore, if a generalized spike is very far away from the support of the population LSD, then the estimate of the spike based on an SP model will closely approximate the estimate based on a GSP model. However, the SP model will provide potentially biased estimates if the non-spikes are not similar and the ratio between the largest non-spike and the spike of interest is substantially larger than zero.

#### 4.6. Comparison with ultra high-dimensional regime-based results when $p/n$ is large

Our methods are developed under the high-dimensional regime  $p/n \rightarrow \gamma < \infty$ , and do not theoretically warrant them to be applied in ultra-high dimensional (UHD) regime where  $p/n \rightarrow \infty$ . However, often times in real-world applications, we only have data with large  $p$  and large  $n$ , but the relative rate of their asymptotic divergence is unknown. Therefore, we do not know whether the true asymptotic regime is high-dimensional ( $p/n \rightarrow \gamma < \infty$ ) or ultra high-dimensional ( $p/n \rightarrow \infty$ ).

Suppose that the true asymptotic regime is high-dimensional with  $\gamma$  finite but large compared to  $n$ , and the eigenvalues follow the GSP model. In such situations, we can either correctly assume the high-dimensional regime and apply the results discussed in this paper, or we can falsely assume the ultra high-dimensional asymptotic regime and employ the theoretical results derived under this regime [15]. In this section, we will investigate whether it is prudent to assume the UHD regime in such situations. In other words, we will try to answer how large  $\gamma$  can be considered to be diverging to infinity for practical applications.



We first show that for large enough  $\gamma$ , the theoretical results based on the UHD regime become nearly identical to the results under the correctly assumed GSP (under high-dimensional regime) model. The UHD-based results presented in [15] require weaker conditions for the non-spiked eigenvalues than those for the spiked population model. Instead of assuming that the non-spiked eigenvalues are the same, Lee et al. [15] assume certain conditions on the moments of the non-spiked eigenvalues. Since the population LSD has a finite support and all of its central moments are finite, the condition on their moments, i.e., Condition 2 in [15] is satisfied with an additional assumption that  $n^3/p^2 = o(1)$ . Without loss of generality, we assume that the mean of the non-spikes is unity. Then, under the UHD regime,

$$d/\lambda \xrightarrow{P} \gamma/\lambda + 1 \text{ when } \lambda \geq O(\gamma), \quad d/\gamma \xrightarrow{P} 1 \text{ when } \lambda = o(\gamma), \quad (3)$$

where  $\lambda$  and  $d$  are a spiked population eigenvalue and its corresponding sample eigenvalue, respectively. Here  $\lambda \geq O(\gamma)$  means  $\lambda/\gamma$  is bounded away from zero, and  $\lambda = o(\gamma)$  means  $\lambda/\gamma \rightarrow 0$ . Lee et al. [15] also showed the convergence of sample eigenvectors and PC scores.

Alternatively, under the GSP model,  $d \rightarrow \psi(\lambda)$  when  $\lambda$  is a distance spike. From Result 1, a distant spike  $\lambda$  must satisfy

$$1 - \gamma \int \frac{x^2}{(\lambda - x)^2} dH(x) > 0,$$

where  $H$  is the population LSD. Since  $f_\lambda(x) = x^2(\lambda - x)^{-2}$  is a continuous function for  $\lambda > \sup \Gamma_H$  and  $x \in \Gamma_H$ , where  $\Gamma_H$  is the support of  $H$ , there exists  $x^* \in (\inf \Gamma_H, \sup \Gamma_H)$  such that  $\int x^2(\lambda - x)^{-2} dH(x) = x^{*2}(\lambda - x^*)^{-2}$ . Then

$$1 - \gamma x^{*2}/(\lambda - x^*)^2 > 0,$$

which implies  $\lambda > x^* + x^* \sqrt{\gamma}$ . Thus, for any  $\lambda > x^* + x^* \sqrt{\gamma}$ ,  $d/\lambda - \psi(\lambda)/\lambda$  converges to zero.

Now, under the true asymptotic regime (high-dimensional)  $\lambda$  and  $\gamma$  are both finite and non-zero, and thus  $\lambda = O(\gamma)$ . However, under the falsely assumed UHD regime, one can further assume  $\lambda \geq O(\gamma)$  or  $\lambda = o(\gamma)$  depending on whether  $\lambda$  is large or small compared to  $\gamma$ . If one assumes  $\lambda \geq O(\gamma)$ , then the difference between the convergence of  $d/\lambda$  from the two models is

$$\psi(\lambda)/\lambda - \gamma/\lambda - 1 = \frac{\gamma}{\lambda} \left\{ \int \frac{x}{1 - x/\lambda} dH(x) - 1 \right\}. \quad (4)$$

Since  $\gamma/\lambda = O(1)$  and  $\int x(1 - x/\lambda)^{-1} dH(x) - 1 = O(1/\lambda)$  as the mean of the non-spikes is unity, (4) becomes almost identical to zero when  $\lambda$  is sufficiently large.

Now, suppose one assumes  $\lambda = o(\gamma)$ . Let  $\lambda \simeq a + b\gamma^k$  for some finite  $a, b$  and  $k \in [1/2, 1)$ . Then, the difference between our result and the UHD result is

$$|\psi(\lambda)/\gamma - 1| = \left| \lambda/\gamma + \lambda \int \frac{x}{\lambda - x} dH(x) - 1 \right| = O(\gamma^{k-1}). \quad (5)$$

Thus, in this case also (5) becomes almost identical to zero when  $\gamma$  is sufficiently large. We can also show the similar results for eigenvectors and PC scores.

Although both GSP and UHD eventually provide nearly identical results when  $\gamma$  is sufficiently large, the GSP model can provide substantially better estimates. The difference can be large when  $\lambda$  is small compared to  $\gamma$ , i.e.,  $k < 1$ , since the difference in (5) is of the order  $O(\gamma^{k-1})$ . The difference will be at least as large as  $O(1/\sqrt{\gamma})$  in such cases. In simulation studies, we show this numerically.

One important thing to note here is that the results under the ultra high-dimensional regime do not provide any consistent estimators for the spiked eigenvalues, albeit there still can be methods to estimate shrinkage. When the spike  $\lambda = o(\gamma)$ , the limiting behavior of the corresponding sample eigenvalue does not depend on  $\lambda$ , and when  $\lambda \geq O(\gamma)$ ,  $d/\lambda \xrightarrow{P} \gamma/\lambda + 1$  does not imply  $d - \gamma$  to be a consistent estimator of  $\lambda$  because  $\lambda$  itself is divergent.

In the simulation studies and real data applications, we further evaluated the performance of  $d - \gamma$  as an estimator. Lee et al. [15] also provided an estimator of the asymptotic shrinkage which we have evaluated in the simulation studies. A recent paper by Cai et al. [7] provided limiting laws for spiked eigenvalues and corresponding eigenvectors when the spikes are divergent, under more general assumptions than the ultra high-dimensional assumptions of [15]. In terms of eigenvalue estimation, this model will still have the problem of the spikes not being estimable. However, it can be an interesting future research direction to find estimators of the shrinkage under this model.

## 5. Estimation of the population LSD

The  $\lambda$ -estimators rely on  $\psi$ , that is a function of the unknown population LSD  $H$ . To use the  $\lambda$ -estimators, it is thus required to estimate  $H$ . Using the Stieltjes transformation and the Marčenko–Pastur theorem, El Karoui [9] developed a general algorithm to estimate the population LSD from the sample ESD,  $F_p$ . We propose to use El Karoui's method to estimate the population LSD  $H$  and then use it to estimate  $\psi$ .

### 5.1. El Karoui's algorithm

Suppose  $v_{F_p}$  is the Stieltjes transformation of the set of eigenvalues in the sample covariance matrix in which

$$v_{F_p}(z) = \frac{1}{n} \sum_{i=1}^n \frac{1}{d_i - z}$$

for any  $z \in \mathbb{C}^+$ ,  $\mathbb{C}^+ = \{x \in \mathbb{C} : \text{Im}(x) > 0\}$ . According to the Marčenko–Pastur theorem [16], when assumptions (A)–(D) hold,  $v_{F_p}$  converges point-wise almost surely to a non-random limit  $v_F$ , which uniquely satisfies the following equation:

$$v_F(z) = - \left\{ z - \gamma \int \frac{\lambda}{1 + \lambda v_F(z)} dH(\lambda) \right\}^{-1}.$$

El Karoui's method first calculates  $v_{F_p}$  for a grid of values  $z_1, \dots, z_J$ , and then finds

$$\hat{H} = \arg_H \min L \left[ \left\{ \frac{1}{v_{F_p}(z_j)} + z_j - \frac{p}{n} \int \frac{\lambda}{1 + \lambda v_{F_p}(z_j)} dH(\lambda) \right\}_{j=1}^J \right],$$

where  $L$  is any pre-defined convex loss function. In order to approximate the integral inside the loss function, the algorithm discretizes  $H$  as

$$dH(\lambda) \simeq \sum_{k=1}^K w_k \delta_{t_k}(\lambda),$$

where  $\delta_{t_k}(\lambda) = 1$  if  $\lambda = t_k$  and 0 otherwise,  $w_1 + \dots + w_K = 1$  with  $w_k > 0$  for all  $k \in \{1, \dots, K\}$ , and  $t_1, \dots, t_K$  is a grid of points on the support of  $H$ . This is basically approximating  $H$  by a discrete distribution with support  $t_1, \dots, t_K$ . Then the integral is approximated by

$$\int \frac{\lambda}{1 + \lambda v_F(z)} dH(\lambda) \simeq \sum_{k=1}^K w_k \frac{t_k}{1 + t_k v_{F_p}(z_j)},$$

and the minimization problem transforms into

$$\hat{H} = \arg \min_H L \left[ \left\{ \frac{1}{v_{F_p}(z_j)} + z_j - \frac{p}{n} \sum_{k=1}^K w_k \frac{t_k}{1 + t_k v_{F_p}(z_j)} \right\}_{j=1}^J \right]. \quad (6)$$

El Karoui [9] has shown the weak convergence of  $\hat{H}$  to  $H$ , i.e.,  $\hat{H} \rightarrow H$ .

Some examples of the convex loss function  $L$  are

- (i)  $L_\infty(e_1, \dots, e_J) = \max_j \max\{|\text{Re}(e_j)|, |\text{Im}(e_j)|\}$ ;
- (ii)  $L_1(e_1, \dots, e_J) = |e_1| + \dots + |e_J|$ ;
- (iii)  $L_2(e_1, \dots, e_J) = |e_1|^2 + \dots + |e_J|^2$ .

For the convex loss functions described above, the estimation of  $H$  in (6) reduces to a convex optimization problem [6]. El Karoui also provided a translation of this problem into a linear programming (LP) problem when the  $L_\infty$  loss function is used. Further details can be found in [9].

### 5.2. Implementing El Karoui's algorithm when the number of spikes is known

Since the generalized spikes fall outside the support of the population LSD, El Karoui's algorithm cannot be directly applied to estimate the spikes. Furthermore, Bai and Silverstein [2] showed that the probability of a sample eigenvalue falling outside the support of the sample LSD will go to zero as  $p$  increases, which implies that the sample eigenvalues corresponding to the population generalized spikes will be measure zero points in the sample LSD. Since the spikes behave like measure zero points (or outliers) when we are concerned about estimating the population LSD, we can exclude the sample eigenvalues corresponding to the population generalized spikes while calculating  $v_{F_p}$  and that will lead to a more robust estimation of  $H$ . Therefore, we will apply El Karoui's algorithm in the following way:

- (I) Suppose the population covariance matrix possesses  $m$  generalized spikes. We exclude the top  $m$  sample eigenvalues while calculating  $v_{F_p}$ , viz.

$$v_{F_p}(z) = \frac{1}{n-m} \sum_{i=1}^n \frac{1}{d_i - z}.$$



- (II) Apply El Karoui's algorithm to obtain  $\hat{H}$ . Additionally, if it is reasonable to assume that the true population LSD is a continuous or piecewise continuous distribution function, suitable kernel smoothing algorithm can be used on  $\hat{H}$  to obtain a more continuous approximation of  $H$ .
- (III) The quantiles of  $\hat{H}$  can be considered as the estimators of the non-spikes.
- (IV) Suppose  $\hat{\lambda}_{m+1}, \dots, \hat{\lambda}_p$  are the estimated non-spikes. Then the  $\psi$  function is estimated by

$$\hat{\psi}(\alpha) = \alpha + \frac{\gamma\alpha}{p-m} \sum_{i=m+1}^p \frac{\hat{\lambda}_i}{\alpha - \hat{\lambda}_i}.$$

Due to the weak convergence  $\hat{H} \rightarrow H$ ,  $\hat{\psi}$  will also converge to  $\psi$  point-wise. Thus, all the estimates provided in Sections 3 and 4 will still be consistent if we replace  $\psi$  with  $\hat{\psi}$ .

A computationally challenging part in El Karoui's algorithm is solving the LP problem, for which there are many fast algorithms available that can solve it in polynomial (in  $K$ , the grid size of  $w_1, \dots, w_K$ ) computation time. An important property of El Karoui's algorithm is that the complexity of the LP step does not depend on  $p$ , which is especially useful for estimating the LSDs of large covariance matrices. In our implementation, we used the `lpSolve` R package [5] to solve the LP problem.

### 5.3. Estimating the number of spikes

Our application of El Karoui's algorithm to the GSP model depends on the number of spikes  $m$ , which is usually unknown. If we have some knowledge of the underlying structure of the data, we can use it to estimate  $m$  roughly. Suppose we know that the data are coming from a mixture of  $K$  subpopulations, and within each subpopulation the observations are iid  $\mathcal{N}(\mu_k, \Sigma)$ , where  $\mu_k$  represents the mean for the  $k$ th subpopulation, and  $\Sigma$  is the common within-group population covariance matrix. Then, as the spikes represent the between-group differences, the number of spikes should be the same as the rank of the between-group covariance matrix which is  $K - 1$ . However in real data, it is often hard to accurately assess the number of such homogeneous subpopulations. In those cases, we can use the following algorithm to estimate  $m$ .

- (I) Start with a reasonable finite upper bound  $m_{\max}$  of the number of spikes. The upper bound can be selected based on prior information on the subpopulations, or by examining the sample eigenvalues. Set  $m = m_{\max}$ .
- (II) Use El Karoui's algorithm to estimate the population LSD and the non-spikes. Suppose the estimated non-spikes are  $\hat{\lambda}_{m+1} \geq \dots \geq \hat{\lambda}_p$ , and the  $\psi$  function is estimated by

$$\hat{\psi}(\alpha) = \alpha + \frac{\gamma\alpha}{p-m} \sum_{i=m+1}^p \frac{\hat{\lambda}_i}{\alpha - \hat{\lambda}_i}.$$

- (III) Find  $S_\psi > \lambda_{m+1}$  using the Newton–Raphson algorithm such that

$$\hat{\psi}'(S_\psi) = 1 - \frac{\gamma}{p-m} \sum_{i=m+1}^p \left( \frac{\hat{\lambda}_i}{S_\psi - \hat{\lambda}_i} \right)^2 = 0.$$

- (IV) Since any distant spike must be larger than  $S_\psi$ , and  $\hat{\psi}, \hat{\psi}'$  are both continuous and strictly increasing functions on  $(S_\psi, \infty)$ , the equation  $\hat{\psi}(\lambda) - d_k = 0$  has a root in  $(S_\psi, \infty)$  if and only if  $\hat{\psi}(S_\psi) - d_k < 0$ . Therefore, find the smallest index  $i^* \in \{1, \dots, m\}$  such that  $d_{i^*} \leq \hat{\psi}(S_\psi)$ . If all  $d_1, \dots, d_m$  are larger than  $\hat{\psi}(S_\psi)$ , then stop and select  $m$  as the number of distant spikes. Otherwise, set  $m = i^* - 1$  and repeat steps (II)–(IV).

Note that the close spikes occur so close to the support of the population LSD that they cannot be distinguished separately from the non-spikes when the number of spikes is unknown.

The selection of  $m_{\max}$  is subjective. It can be selected based on the prior knowledge of the number of subpopulations or by investigating the sample eigenvalues. In real data applications, we are usually interested in only a few large eigenvalues. In such situations,  $m_{\max}$  can also be selected to be slightly larger than the number of eigenvalues we are interested in. As seen from our simulation studies, this spike selection algorithm can overestimate the number of spikes if the upper bound  $m_{\max}$  is too large or underestimate the number of spikes if there are close spikes present; see the Online Supplement. However, as long as  $m_{\max}$  is small compared to  $n$  and  $p$ , the estimation of the true population distant spikes will still remain consistent.

## 6. Simulation studies and real data example

### 6.1. Simulation studies: Compare GSP and SP-based methods

In this section we will present simulation studies of four different scenarios to compare the performances of the proposed GSP-based methods and the existing SP-based method proposed by Lee et al. [14]. For each study, we simulated

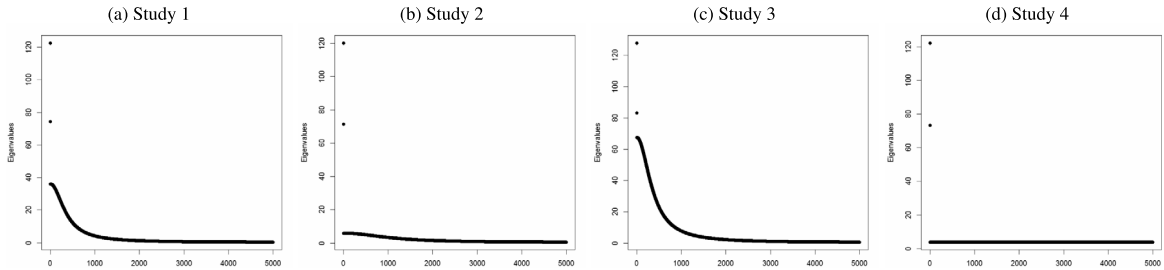


Fig. 2. Eigenvalue structures in simulation studies comparing GSP-based and SP-based methods.

Table 1

Simulation results for GSP-based and SP-based methods for estimating the population eigenvalues, cosine of the angles between sample and population eigenvectors, correlations between sample and population PC scores, and the asymptotic shrinkage factors. Each cell has empirical bias (%) with coefficients of variations (%) in parentheses.

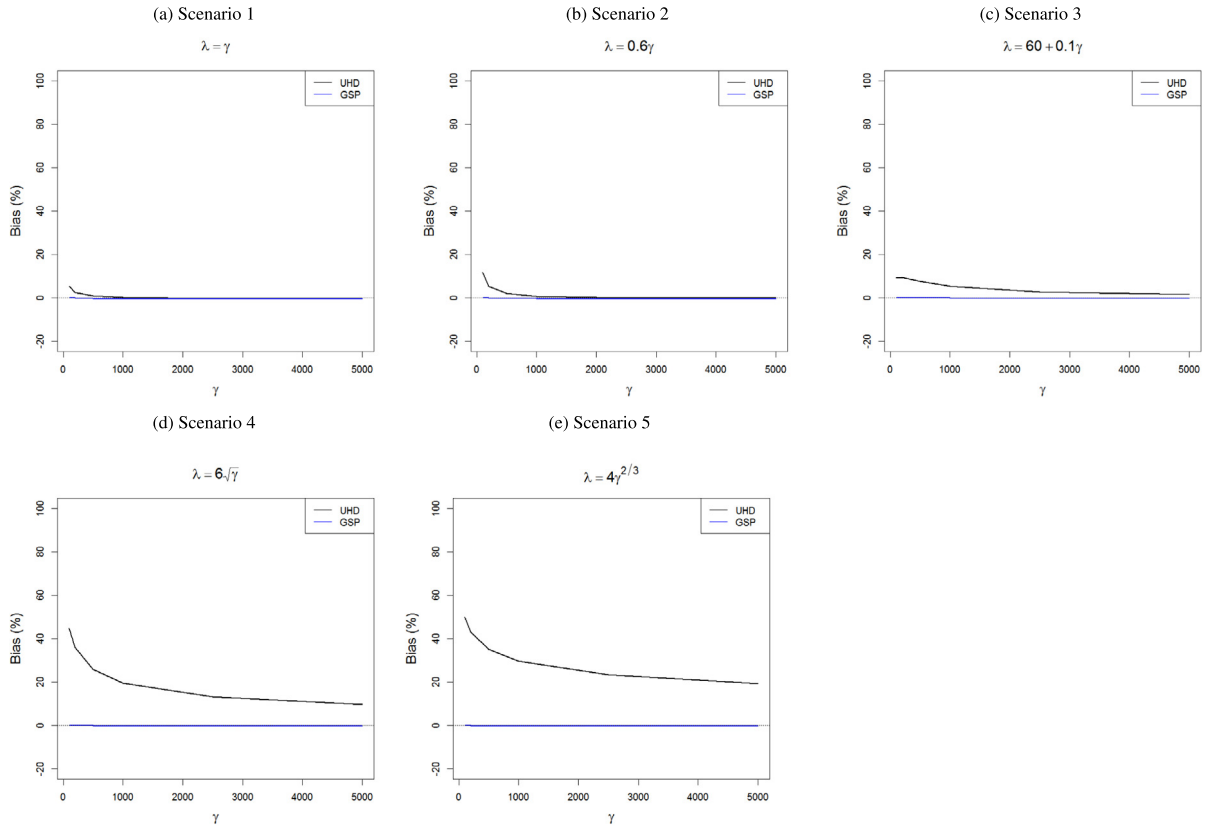
Settings		Method	Eigenvalue		Angle		Correlation		Shrinkage	
No.			1	2	1	2	1	2	1	2
1	$n = 500$ $p = 5000$ $\sigma^2 = 4$ $\rho = 0.8$	SP	5.27 (2.37)	18.27 (3.11)	6.52 (0.32)	34.07 (0.60)	3.83 (0.03)	23.33 (0.08)	5.32 (0.60)	17.88 (1.06)
		$\lambda$ -GSP	0.43 (2.67)	0.95 (5.27)	0.53 (0.77)	3.28 (6.26)	0.33 (0.31)	2.79 (4.69)	0.47 (0.92)	0.58 (3.31)
		$d$ -GSP	0.47 (2.67)	0.69 (5.45)	0.47 (0.77)	2.48 (6.70)	0.24 (0.31)	2.11 (5.07)	0.51 (0.92)	0.31 (3.51)
2	$n = 500$ $p = 5000$ $\sigma^2 = 1$ $\rho = 0.7$	SP	0.10 (0.90)	0.46 (1.27)	0.16 (0.04)	0.44 (0.08)	0.08 (0.001)	0.24 (0.003)	0.18 (0.08)	0.39 (0.16)
		$\lambda$ -GSP	−0.04 (0.90)	0.04 (1.28)	0.01 (0.04)	0.004 (0.10)	0.01 (0.03)	0.01 (0.01)	0.03 (0.08)	−0.03 (0.18)
		$d$ -GSP	−0.004 (0.90)	0.10 (1.28)	0.03 (0.04)	0.03 (0.10)	0.004 (0.03)	0.01 (0.01)	0.07 (0.08)	0.03 (0.18)
3	$n = 500$ $p = 5000$ $\sigma^2 = 7.5$ $\rho = 0.8$	SP	25.68 (2.54)	—	64.06 (0.52)	—	46.50 (0.07)	—	26.41 (0.90)	—
		$\lambda$ -GSP	2.92 (5.7)	—	12.62 (11.90)	—	10.95 (10.13)	—	3.47 (4.20)	—
		$d$ -GSP	2.45 (5.74)	—	12.25 (10.52)	—	10.87 (8.58)	—	3.00 (4.24)	—
4	$n = 500$ $p = 5000$ $\sigma^2 = 4$ $\rho = 0$	SP	0.05 (1.58)	−0.26 (2.35)	0.06 (0.23)	−0.06 (0.53)	0.03 (0.02)	0.05 (0.08)	0.07 (0.43)	−0.22 (0.90)
		$\lambda$ -GSP	0.03 (1.58)	−0.35 (2.35)	0.02 (0.24)	−0.18 (0.54)	0.01 (0.02)	−0.02 (0.09)	0.04 (0.43)	−0.31 (0.91)
		$d$ -GSP	0.16 (1.58)	−0.12 (2.35)	0.10 (0.23)	−0.03 (0.53)	0.01 (0.02)	0.02 (0.09)	0.18 (0.42)	−0.08 (0.90)

a training dataset with  $n = 500$  individuals and  $p = 5000$  features. The data were generated from three subpopulations with sample sizes 100, 150 and 250. For each subpopulation we first selected a mean vector  $\mu_i$  by drawing its elements randomly with replacement from  $\{-0.3, 0, 0.3\}$ . Then samples in the  $i$ th subpopulation were drawn from  $\mathcal{N}_p(\mu_i, V)$  where  $V$  is the AR(1) covariance matrix with variance  $\sigma^2$  and autocorrelation  $\rho$ . The  $(\sigma^2, \rho)$  pairs used for the four studies were (4, 0.8), (1, 0.7), (7.5, 0.8) and (4, 0). The population eigenvalue plots for all the studies are shown in Fig. 2.

We also generated test datasets for each study with the same settings as the training datasets. Then we applied our GSP-based methods and the existing SP-based method to estimate the population spikes, the angles between the sample and population eigenvectors, the correlations between the sample and population PC scores, and the asymptotic shrinkage factors. For all of the studies, we used the upper bound  $m_{\max} = 5$  to estimate the number of distant spikes using the algorithm described in Section 5.3. We simulated each study 200 times to calculate the empirical biases and standard errors of the estimates. The results are presented in Table 1.

It is clear from Table 1 that for Studies 1, 2 and 3, our methods reduced the bias in all the estimates while having similar standard errors as the existing method. The positive empirical biases in all the SP estimates suggest that the SP method tends to overestimate all the quantities. In Study 4, since the underlying population satisfied the SP assumption, all methods provided very similar and almost unbiased estimates ( $< 1\%$ ). The results also verify that the  $\lambda$ -estimates and  $d$ -estimates are asymptotically equivalent. The performances of the  $\lambda$ -estimates and the  $d$ -estimates are nearly identical in all the simulation studies.

In Study 1, the ratio of the largest non-spike with the two spikes is 0.29 and 0.48, which is substantially larger than zero. Thus according to the discussion in Section 4, the SP model does not closely approximate the GSP model. The results



**Fig. 3.** Relative errors (%) in the convergence results of the largest sample eigenvalue derived under GSP and UHD regimes. The population eigenvalues and the rate of increment of the largest population eigenvalue are assumed to be known.

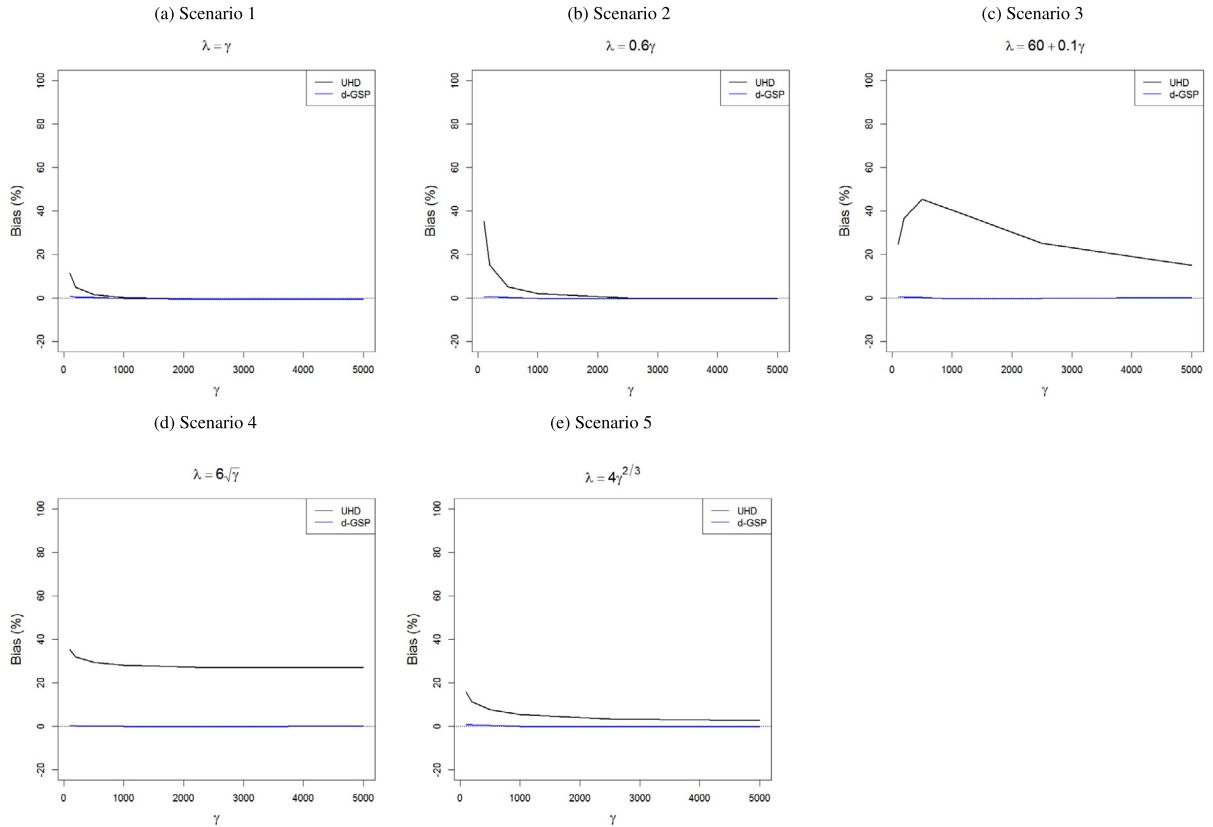
support this assertion as the SP model-based estimates are highly biased whereas the estimates based on our methods have very little empirical bias. In Study 2, the largest non-spike is very small compared to the smallest spike (ratio 0.08). Thus the estimates based on the SP model closely approximate the estimates based on the GSP model, and we find very little empirical bias ( $< 1\%$ ) in all of the SP model-based estimates. In Study 3, even though there were two spikes present, only the largest population eigenvalue was a distant spike. So we presented only the estimates corresponding to the largest population eigenvalue. Since the ratio of the largest non-spike and the largest spike is substantially larger than zero (0.53) in this study, we observe very high empirical bias in the SP model-based estimates. However, our methods provided negligible empirical biases even in the presence of a close spike. We also presented the estimated number of distant spikes in each of the simulation studies in the Online Supplement. Note that in some cases, our algorithm over-estimates the number of distant spikes. However, as the over-estimation is only finite, the estimates of the distant spikes still remain consistent.

## 6.2. Simulation studies: Compare GSP and UHD-based methods

In Section 4.6, we compared the asymptotic results under the UHD regime and the results based on the high-dimensional GSP model when  $p$  is greatly larger than  $n$ , but  $p/n$  is large but finite. We theoretically established that the results from the two regimes become almost identical when  $p/n = \gamma$  is sufficiently large. However, given large but finite  $\gamma$  in the data, the difference can be substantial when the spike is smaller compared to  $\gamma$ .

In this section, we will assess that result by numerically comparing the GSP and UHD-based estimates for different values of  $\gamma$ . We considered five different scenarios where the largest population eigenvalue  $\lambda \in \{\gamma, 0.6\gamma, 60 + 0.1\gamma, 6\sqrt{\gamma}, 4\gamma^{2/3}\}$ . For the first three scenarios, under the UHD regime,  $\lambda/\gamma$  can be assumed to be bounded away from zero, and for the last two,  $\lambda/\gamma \rightarrow 0$  as  $\gamma \rightarrow \infty$ . To compare the performances as  $\gamma$  increases, we selected six different values for  $\gamma$ , viz.  $\gamma \in \{100, 200, 500, 1000, 2500, 5000\}$ . For each combination of  $\gamma$  and  $\lambda$ , we simulated 200 datasets, each with  $n = 200$  samples from a population with only one spike  $\lambda$ , and the non-spikes generated from the AR(1) covariance structure with  $(\sigma^2, \rho) = (1, 0.9)$ .

First, we compare the convergence results of the largest sample eigenvalue  $d$  from Result 1 and (3). For this purpose, we assume the population eigenvalues and the rate of increment of  $\lambda$  are known, and we compare the relative errors



**Fig. 4.** Empirical biases (%) in estimating the largest population eigenvalue for GSP-based and UHD-based methods. The population eigenvalues and the rate of increment of the largest population eigenvalue are assumed to be unknown.

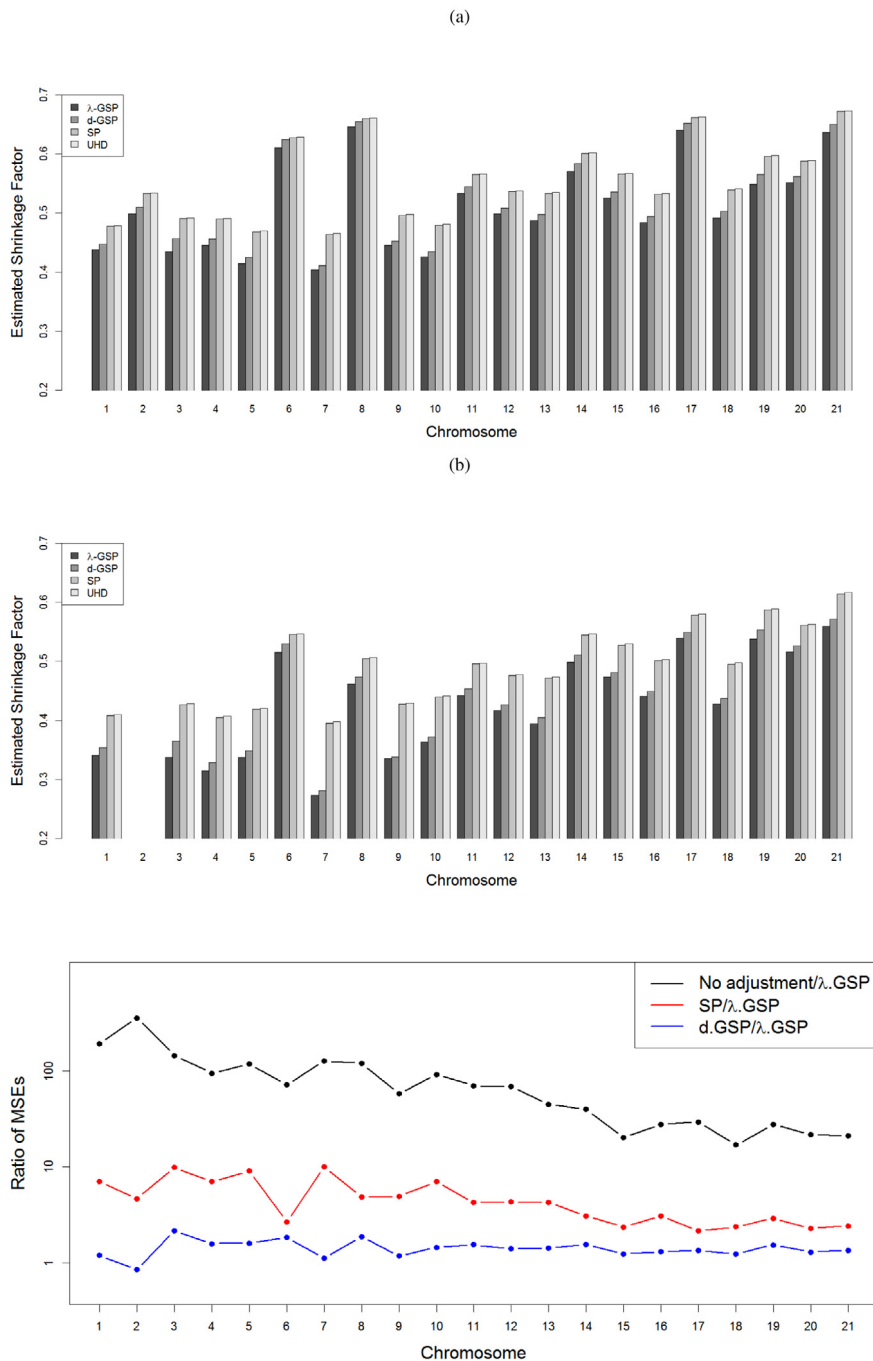
$\epsilon_{GSP} = \{d - \psi(\lambda)\}/d$  and  $\epsilon_{UHD} = (d - \lambda - \gamma)/d$  or  $(d - \gamma)/d$  depending on whether  $\lambda/\gamma$  is assumed to be bounded away from zero or not. Fig. 3 shows that for all combinations of  $(\gamma, \lambda)$ , the GSP-based convergence result (Result 1) has very negligible relative errors. In contrast, the UHD-based convergence result (3) has substantially large relative errors even for  $\gamma$  as large as 5000 in Scenarios 3, 4 and 5. For Scenarios 1 and 2, since  $\lambda$  increases at a faster rate with  $\gamma$  than in other scenarios, the relative errors based on the two results converge much faster. However, for relatively smaller values of  $\gamma$  (100, 200, 500), the differences are substantial even though  $\gamma$  is large compared to  $n = 200$ . This suggests that we need  $\gamma$  to be large in an absolute sense, and not only in a relative sense compared to  $n$  in order to assume  $\gamma \rightarrow \infty$  and apply UHD-based results.

Next, we compare the estimates of the spike  $\lambda$  using GSP-based and UHD-based methods assuming the population eigenvalues and the rate of increment of the population spike  $\lambda$  to be unknown. Among the GSP-based methods, we only used the  $d$ -GSP method for this purpose due to the computational burden associated with applying the  $\lambda$ -GSP method on such a large number of simulated datasets. One thing to note here is that the UHD results do not provide any consistent estimators for  $\lambda$ , as it is assumed to be divergent when  $\lambda \geq O(\gamma)$ , and the asymptotic properties of the sample eigenvalues do not depend on  $\lambda$  when  $\lambda/\gamma \rightarrow 0$ . Thus, in order to compare these methods, we estimate  $\lambda$  by  $\hat{\lambda} = d - \gamma$  when considering the UHD regime.

From Fig. 4 we can see that our proposed  $d$ -GSP method provides almost negligible biases for all combinations of  $(\gamma, \lambda)$ , whereas the UHD-based estimates have substantial biases even for  $\gamma$  as large as 5000 in Scenarios 3, 4 and 5. For Scenarios 1 and 2, both methods provide almost unbiased estimates when  $\gamma \geq 1000$  and  $\gamma \geq 2000$ , respectively. Further, we compare the estimated shrinkage factors based on these two methods in the Online Supplement. They also show very similar patterns as the estimated spikes.

### 6.3. Application on Hapmap III data

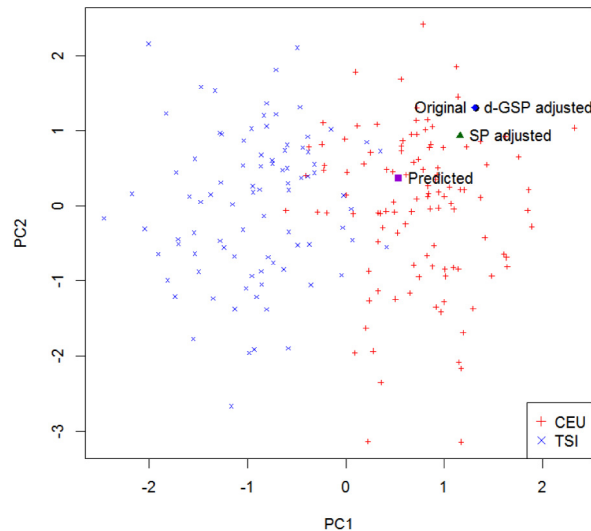
For this demonstration, we used genetic data from the Hapmap Phase III project (<http://hapmap.ncbi.nlm.nih.gov/>). Our sample consisted of unrelated individuals sampled from two different populations: a) Utah residents with Northern and Western European ancestry (CEU) and b) Tuscans in Italy (TSI). We only included genomic markers that are on chromosomes 1–22, have less than 5% missing values, and those with minor allele frequency more than 0.05. We also



**Fig. 5.** Estimated shrinkage factors for (a) PC1 and (b) PC2 across chromosomes 1–21 based on three different methods. (c) Comparison of the mean squared errors (MSE) of the unadjusted and adjusted PC scores based on the  $d$ -GSP and SP methods with the adjusted PC scores based on the  $\lambda$ -GSP method. The ratios of the MSEs are presented for chromosomes 1–21 using the threshold  $\epsilon = 1$ . The y-axis is presented in a logarithmic scale.

excluded two samples (both from CEU) with outlier PC scores (more than six standard deviations away from the mean PC score corresponding to at least one distant spike). We then mean-centered and variance-standardized the data for each marker. The final sample consisted of 198 individuals (110 from CEU and 88 from TSI). Total number of markers selected across chromosomes 1–22 was 1,389,511.

To evaluate the performance of the proposed methods with different  $p$ , we performed PCA on each chromosome separately. The number of markers varied from 19,331 (chromosome 21) to 116,582 (chromosome 2). The distribution



**Fig. 6.** PC1 vs PC2 plot of the Hapmap III CEU and TSI samples based on chromosome 7. The predicted PC scores for the illustrative individual, and its bias-adjusted PC scores are also presented. Since the  $d$ -GSP and the  $\lambda$ -GSP adjusted scores are nearly the same, the  $\lambda$ -GSP adjusted scores are not presented.

of the number of markers across different chromosomes is presented in the Online Supplement. We first estimated the number of distant spikes using the algorithm described in Section 5.3. We found no distant spike in chromosome 22 and only one distant spike in chromosome 2. Then we applied our GSP-based methods, the existing SP-based method [14] and the UHD-based method [15] to estimate the asymptotic shrinkage factors corresponding to the distant spikes. Fig. 5a, 5b compares the estimated asymptotic shrinkage factors for the first two PCs across different chromosomes. The plots show that for all the chromosomes,  $\lambda$ -GSP and  $d$ -GSP methods provided almost equal estimates while the SP and UHD estimates are larger than both the GSP estimates. This suggests that the SP method would over-estimate the shrinkage factors when the population eigenvalues deviate from the assumption that the non-spiked eigenvalues are the same. Moreover, the UHD method over-estimated the shrinkage factors even for  $p/n$  nearly as large as 600 (chromosome 2).

To investigate whether the proposed shrinkage-bias adjustment can improve the prediction accuracy, we performed a leave-one-out cross-validation. In each iteration we removed one individual (test sample) and performed PCA on the remaining individuals (training samples) to predict the PC score of the test sample. For each predicted PC score, we adjusted the shrinkage-bias using the GSP-based, SP-based and UHD-based shrinkage factor estimates.

One important issue with this cross-validation is that the exclusion of one individual can substantially change the PC-coordinates, in which the PC score plots from the training sample-based and complete sample-based PCA can be substantially different. To circumvent this problem, in each iteration we first rescaled the PC scores based on their corresponding sample eigenvalues to make the PCs comparable. In addition, we obtained the mean squared difference of the training sample PC1-2 scores with and without the exclusion of the test sample (for chromosome 2, only PC1 is used), and excluded the test sample from the prediction error estimation if the mean squared difference was above a threshold  $\epsilon$ . We used four different values 0.5, 1, 5 and 10 for the threshold parameter  $\epsilon$ , and for each value of  $\epsilon$  we calculated the mean squared errors (MSE) of the unadjusted and adjusted predicted PC scores of the test samples. The sample sizes of the test samples that were finally included in the prediction error estimation for different values of  $\epsilon$  are shown in the Online Supplement.

Fig. 5 shows the estimated MSEs for  $\epsilon = 1$ . It is clear that both the  $\lambda$ -GSP and  $d$ -GSP methods have much smaller MSEs than the SP method. The UHD-based method had almost identical MSEs as the SP-based method for all the chromosomes, hence we omitted the UHD-based results in this plot. As expected, the unadjusted predicted PC scores have substantially larger MSE than all the proposed adjustments. The plots are very similar for the other values of  $\epsilon$ , and they can be found in the Online Supplement.

Fig. 6 illustrates the shrinkage-bias adjustment for the PC1 and PC2 scores of an individual based on the markers on chromosome 7. The plot clearly shows that the bias-adjusted PC score based on the SP model is still biased toward zero, whereas the bias-adjusted PC score based on the GSP model is very close to the original sample PC score. We only showed the  $d$ -GSP adjusted score in the plot as the  $d$ -GSP and  $\lambda$ -GSP adjusted scores were almost equal.

## 7. Conclusions and discussion

In this paper, we investigated the asymptotic properties of PCA under the Generalized Spiked Population model and derived estimators of the population eigenvalues, the angles between the sample and population eigenvectors, and the



correlation coefficients between the sample and population PC scores. We also proposed methods to adjust the shrinkage bias in the predicted PC scores. Further, theoretically and using simulation studies, we compared our results with the results developed under the ultra-high dimensional regime [15], and showed that our methods provide more accurate results under the high-dimensional regime even when  $p$  is greatly larger than  $n$ . Since the proposed methods do not require the equality of the non-spiked eigenvalues, they can be widely used in high-dimensional biomedical data analysis. We implemented all our algorithms in the R package `hdPCA`.

We note that Mestre [18,19] proposed an asymptotic setting similar to the generalized spiked population model but with a different assumption on the number of spikes in which the number of spikes increases with the dimension. Under this assumption, he provided asymptotic properties of sample eigenvalues and eigenvectors. However, in many biomedical data, the number of spikes is usually finite as the spikes represent the difference between finitely many underlying subpopulations. Therefore we believe that the generalized spiked population model is more appropriate in such cases.

In some special cases, even though the features exhibit strong local correlation, one can use the spiked population model based methods after some suitable data manipulation. In genome-wide association studies, SNP pruning [1] can be used to remove locally correlated SNPs to satisfy the spiked population model. For example, Lee et al. [14] reported good performance of the spiked population model-based methods with the SNP-pruned Hapmap III dataset. This approach, however, can lead to a considerable loss of information; the SNP-pruning in Hapmap III data removed nearly 90% of the SNPs. Since the proposed approach does not require this additional step, it can use most of the information present in the data.

## Acknowledgments

We thank the Editor-in-Chief, Christian Genest, and referees for their helpful comments and suggestions, which led to substantial improvements of the manuscript. The work was supported by National Institutes of Health, USA Grants R00HL113164 and R01HG008773.

## Appendix A

**Proof of Theorem 1.** The first part of the proof follows directly from Result 1 along with the fact that on the domain of the distant spikes, the  $\psi$  function is strictly increasing, and hence is left invertible. Since  $\psi''(\alpha) > 0$  for any  $\alpha > \sup \Gamma_H$ ,  $\psi'(\alpha)$  is a strictly increasing function for  $\alpha > \sup \Gamma_H$ . Let  $S_\psi > \sup \Gamma_H$  be a solution for  $\psi'(\alpha) = 0$ . Then for any  $\alpha > \sup \Gamma_H$ ,  $\psi'(\alpha) > 0$  if and only if  $\alpha > S_\psi$ . Therefore the interval  $(S_\psi, \infty)$  is the domain of the distance spikes, and  $\psi$  is a strictly increasing function on this interval. The second part follows from Lemma B.  $\square$

**Proof of Theorem 2.** The proof closely follows the proof of Theorem 2 in [18]. However, contrary to [18], we do not assume that the population LSD contains the generalized spikes. Thus, some of the derivation steps and results are substantially different from [18]. We start the derivation by first noting that the quadratic forms  $\hat{\eta}_k$  can be expressed as contour integrals of a special class of Stieltjes transforms of the sample covariance matrix. Let us define, for all  $z \in \mathbb{C}^+$ ,

$$\hat{m}_p(z) = s_1^\top (S_p - zI_p)^{-1} s_2 = \sum_{j=1}^p s_1^\top e_j e_j^\top s_2 / (d_j - z),$$

where  $s_1$  and  $s_2$  are non-random vectors with uniformly bounded norms. Girko [10] and Mestre [17] showed that under the assumption that the population LSD contains the generalized spikes, one has, for all  $z \in \mathbb{C}^+$ ,

$$|\hat{m}_p(z) - m_p(z)| \xrightarrow{a.s.} 0, \quad (\text{A.1})$$

where

$$m_p(z) = s_1^\top \{w(z)\Sigma_p - zI_p\}^{-1} s_2 = \sum_{j=1}^p \frac{s_1^\top E_j E_j^\top s_2}{w(z)\lambda_j - z}.$$

The function  $w(z)$  is defined as  $w(z) = 1 - \gamma - \gamma z b_F(z)$ , where  $b_F(z) = \int (\tau - z)^{-1} dF(\tau)$  is the Stieltjes transform of the sample LSD. It is easy to check, by the same arguments provided in [17], that the result still holds when the generalized spikes are considered lying outside the support of the population LSD.

The functions  $\hat{m}_p$ ,  $m_p$  and  $b_F$  can be extended to  $\mathbb{C}^- = \{z \in \mathbb{C} : \text{Im}(z) < 0\}$  by defining  $\hat{m}_p(z) = \hat{m}_p^*(z^*)$ ,  $m_p(z) = m_p^*(z^*)$  and  $b_F(z) = b_F^*(z^*)$  for  $z \in \mathbb{C}^-$ , where  $z^*$  is the complex conjugate of  $z$ .

With this definition,  $|\hat{m}_p(z) - m_p(z)| \xrightarrow{a.s.} 0$  even when  $z \in \mathbb{C}^-$ . Now  $\hat{\eta}_k$  can be expressed as an integral of  $\hat{m}_p$ , viz.

$$\hat{\eta}_k = \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_Y^-(k)} \hat{m}_p(z) dz,$$

where  $i = \sqrt{-1}$ ,  $y > 0$  and  $\partial \hat{\mathbb{R}}_y^-(k)$  is the negatively (clockwise) oriented boundary of the rectangle

$$\hat{\mathbb{R}}_y(k) = \{z \in \mathbb{C} : \hat{a}_1 \leq \operatorname{Re}(z) \leq \hat{a}_2, |\operatorname{Im}(z)| \leq y\}.$$

The values of  $\hat{a}_1$  and  $\hat{a}_2$  can be arbitrarily chosen provided that  $\hat{\mathbb{R}}_y(k)$  contains only the sample eigenvalue  $d_k$  and no other sample eigenvalue. Then the following lemma gives the almost sure limit of  $\hat{\eta}_k$ .

**Lemma A.**

$$\left| \frac{1}{2\pi i} \oint_{\partial \hat{\mathbb{R}}_y^-(k)} \hat{m}_p(z) dz - \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} m_p(z) dz \right| \xrightarrow{\text{a.s.}} 0,$$

where  $y > 0$  and  $\partial \mathbb{R}_y^-(k)$  is the negatively (clockwise) oriented boundary of the rectangle

$$\mathbb{R}_y(k) = \{z \in \mathbb{C} : a_1 \leq \operatorname{Re}(z) \leq a_2, |\operatorname{Im}(z)| \leq y\}.$$

The constants  $a_1$  and  $a_2$  can be arbitrarily chosen so that  $\psi(\lambda_k) \in [a_1, a_2]$  and  $[a_1, a_2] \subset \psi(S_\psi, \infty)$ , where  $S_\psi > \sup \Gamma_H$ ,  $\psi'(S_\psi) = 0$ . Here,  $\psi(S_\psi, \infty)$  denotes the image of the interval  $(S_\psi, \infty)$  under  $\psi$ .

**Lemma A** implies

$$\left| \hat{\eta}_k - \sum_{j=1}^p \left\{ \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \frac{dz}{w(z)\lambda_j - z} \right\} s_1^\top E_j E_j^\top s_2 \right| \xrightarrow{\text{a.s.}} 0. \quad (\text{A.2})$$

Now we need to evaluate the integral in (A.2) in order to get the almost sure limit of the random variable  $\hat{\eta}_k$ . First, we extend the  $\psi$  function to  $\mathbb{R}_y(k)$  as follows. For all  $z \in \mathbb{R}_y(k)$ ,

$$\psi(z) = z \left\{ 1 + \gamma \int \frac{\lambda}{z - \lambda} dH(\lambda) \right\}.$$

According to [16], for all  $z \in \mathbb{C}^+$ ,  $b_F(z) = b$  is the unique solution to the equation

$$b = \int \frac{1}{\lambda(1 - \gamma - \gamma z b) - z} dH(\lambda) \quad (\text{A.3})$$

in the set  $\{b \in \mathbb{C} : \gamma b - (1 - \gamma)/z \in \mathbb{C}^+\}$ . It is easy to see that  $b_F$  also satisfies (A.3) when  $z \in \mathbb{C}^-$ . Now we formally define the  $f_F$  function introduced in (1). For all  $z \in \mathbb{C} \setminus \mathbb{R}$ , set

$$f_F(z) = \frac{z}{w(z)} = \frac{z}{1 - \gamma - \gamma z b_F(z)}. \quad (\text{A.4})$$

Then  $b_F$  can be expressed in terms of  $f_F$  as

$$b_F(z) = \frac{(1 - \gamma)f_F(z) - z}{\gamma z f_F(z)}.$$

By replacing  $b$  with  $\{(1 - \gamma)f - z\}/\gamma z f$  in (A.3), we get

$$f \left\{ 1 + \gamma \int \frac{\lambda}{f - \lambda} dH(\lambda) \right\} = z. \quad (\text{A.5})$$

It is easy to see that  $b_F$  is a solution to (A.3) if and only if  $f_F$  is a solution to (A.5). Therefore, for all  $z \in \mathbb{C}^+$  (similarly for  $z \in \mathbb{C}^-$ ),  $f_F(z) = f$  is the unique solution to (A.5) on  $\mathbb{C}^+$  (respectively,  $\mathbb{C}^-$ ). This implies  $\psi\{f_F(z)\} = z$  for all  $z \in \mathbb{R}_y(k) \setminus [a_1, a_2]$ .

Now we focus on the case when  $z \in \mathbb{R} \setminus \{0\}$ . According to [22], we can extend  $b_F$  to  $\mathbb{R} \setminus \{0\}$  by defining  $b_F(z) = \lim_{y \rightarrow 0^+} b_F(z + iy)$  for any  $z \in \mathbb{R} \setminus \{0\}$ . The definition of  $f_F$  can also be extended in a similar fashion. In Lemma B we have shown that  $f_F$  is the inverse function of  $\psi$  on  $(S_\psi, \infty)$ , and there exists  $M_f > \sup \Gamma_F$  for which  $\psi(S_\psi, \infty) = (M_f, \infty)$ . Thus,  $[a_1, a_2] \subset \psi(S_\psi, \infty)$  implies  $\psi\{f_F(z)\} = z$  for all  $z \in \mathbb{R}_y(k)$ . Furthermore, the function  $\psi$  is continuous and differentiable on  $\mathbb{R}_y(k)$ , and the derivative is given by

$$\psi'(z) = 1 - \gamma \int \left( \frac{\lambda}{z - \lambda} \right)^2 dH(\lambda).$$

Then the integral in (A.2) can be expressed in terms of  $\psi$  and  $f_F$  as follows:

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \frac{dz}{w(z)\lambda_j - z} &= \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \frac{dz}{\frac{z}{f_F(z)}\lambda_j - z} \\ &= \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \frac{1}{\lambda_j - f_F(z)} \times \frac{f_F(z)}{\psi\{f_F(z)\}} dz. \end{aligned} \quad (\text{A.6})$$

The integrand in the final expression is holomorphic on  $\mathbb{R}_y^-(k)$  when  $j \neq k$  and possesses a simple pole  $\psi(\lambda_k)$  when  $j = k$ . Therefore, when  $j \neq k$  the integral in (A.6) is zero. When  $j = k$ , applying the residue theorem on the final integral, we get

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \frac{dz}{w(z)\lambda_k - z} &= \lim_{z \rightarrow \psi(\lambda_k)} \frac{\psi(\lambda_k) - z}{\lambda_k - f_F(z)} \times \frac{f_F(z)}{\psi\{f_F(z)\}} \\ &= \lim_{z \rightarrow \psi(\lambda_k)} \frac{\psi(\lambda_k) - \psi\{f_F(z)\}}{\lambda_k - f_F(z)} \times \frac{f_F(z)}{\psi\{f_F(z)\}} = \frac{\lambda_k \psi'(\lambda_k)}{\psi(\lambda_k)}. \end{aligned}$$

This implies

$$\frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \frac{dz}{w(z)\lambda_j - z} = \begin{cases} \lambda_k \psi'(\lambda_k) / \psi(\lambda_k) & \text{if } j = k, \\ 0 & \text{if } j \neq k, \end{cases}$$

and the proof of Theorem 2 is complete.  $\square$

**Proof of Lemma A.** First, we show that  $\hat{a}_1, \hat{a}_2, a_1, a_2$  can be chosen satisfying  $\hat{a}_1 \rightarrow a_1$  and  $\hat{a}_2 \rightarrow a_2$ . This is possible due to the fact that  $d_k \xrightarrow{a.s.} \psi(\lambda_k)$  and  $\psi(\lambda_k) \subset \psi(S_\psi, \infty) = (M_f, \infty)$ , where  $M_f > \sup \Gamma_F$ . Therefore, we can choose a neighborhood  $[a_1, a_2]$  around  $\psi(\lambda_k)$  so that  $[a_1, a_2] \subset (M_f, \infty)$ . Moreover, as  $M_f$  is bounded away from the support of the sample LSD  $F$  and  $d_k \xrightarrow{a.s.} \psi(\lambda_k)$ , we can select a neighborhood  $[\hat{a}_1, \hat{a}_2]$  around  $d_k$  which does not contain any other eigenvalue for which  $\hat{a}_1 \rightarrow a_1, \hat{a}_2 \rightarrow a_2$ . Then

$$\begin{aligned} &\left| \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} \hat{m}_p(z) dz - \frac{1}{2\pi i} \oint_{\partial \mathbb{R}_y^-(k)} m_p(z) dz \right| \\ &\leq \frac{1}{2\pi} \left\{ \sup_{z \in \partial \mathbb{R}_y^-(k) \cup \partial \mathbb{R}_y^+(k)} |\hat{m}_p(z)| \right\} (|\hat{a}_1 - a_1| + |\hat{a}_2 - a_2|) \\ &\quad + \frac{1}{2\pi} \oint_{\partial \mathbb{R}_y^-(k)} |\hat{m}_p(z) - m_p(z)| \times |dz|. \end{aligned} \quad (\text{A.7})$$

From the Cauchy–Schwarz inequality, we can obtain the following upper bound for  $\hat{m}_p$ :

$$|\hat{m}_p(z)| \leq \|s_1\| \times \|s_2\| / d(z, \Gamma_{F_p}).$$

Here,  $d(z, \Gamma_{F_p}) = \inf_{y \in \Gamma_{F_p}} |z - y|$ . Since  $F_p \rightarrow F$  point-wise and  $[a_1, a_2]$  is bounded away from  $\Gamma_F$ ,  $d(z, \Gamma_{F_p})$  is bounded away from zero with probability 1 for large enough  $p$  and  $n$ . Therefore  $|\hat{m}_p(z)|$  is finite for  $z \in \mathbb{R}_y(k)$  with probability 1 for large enough  $p$  and  $n$ . Moreover, since  $[\hat{a}_1, \hat{a}_2] \rightarrow [a_1, a_2]$ , the interval  $[\hat{a}_1, \hat{a}_2]$  will eventually be bounded away from  $\Gamma_F$ . Thus, eventually the upper bound for  $|\hat{m}_p(z)|$  will also be finite for  $z \in \mathbb{R}_y(k)$ . Therefore, the first term on the right-hand side of (A.7) will go to zero as  $\hat{a}_1 \rightarrow a_1, \hat{a}_2 \rightarrow a_2$ .

Now, as  $\hat{m}_p(z)$  and  $m_p(z)$  are holomorphic functions on the compact set  $\partial \mathbb{R}_y^-(k)$ , we have

$$\sup_{z \in \partial \mathbb{R}_y^-(k)} |\hat{m}_p(z) - m_p(z)| < \infty.$$

Also from (A.1),  $|\hat{m}_p(z) - m_p(z)| \xrightarrow{a.s.} 0$  point-wise for all  $z \in \mathbb{C} \setminus \mathbb{R}$ . Therefore, by Lebesgue's Dominated Convergence Theorem, the second term on the right-hand side of (A.7) also converges to zero almost surely. This completes the proof of Lemma A.  $\square$

We can show the asymptotic equivalence of the limits derived in Theorem 2 and Result 2 as a direct application of the following lemma.

**Lemma B.** Suppose assumptions (A)–(D) hold. If  $\lambda_k$  is a distant spike with multiplicity 1, and  $d_k$  is the corresponding sample eigenvalue, then

$$f_F(d_k) \xrightarrow{p} \lambda_k, \quad \frac{d_k g_F(d_k)}{f_F(d_k)} \xrightarrow{p} \psi'(\lambda_k).$$

**Proof.** We have already established in the proof of Theorem 2 that for all  $z \in \mathbb{C}^+$  (similarly for  $z \in \mathbb{C}^-$ ),  $f_F(z) = f$  is the unique solution to (A.5) on  $\mathbb{C}^+$  (respectively,  $\mathbb{C}^-$ ). When  $z$  is restricted to  $\mathbb{C} \setminus \mathbb{R}$ , using (A.4) and the fact that  $b_F(z) = \int (\tau - z)^{-1} dF(\tau)$ , we can write

$$f_F(z) = z \left\{ 1 + \gamma \int \frac{\tau}{z - \tau} dF(\tau) \right\}.$$

Now suppose  $z = x \in \mathbb{R} \setminus \{0\}$ . Then both Eqs. (A.3) and (A.5) will have multiple roots, both real and complex valued depending on  $x$  and  $H$ . If we look at (A.5) closely, we can see for real-valued  $x$ , it can be represented as  $\psi\{f(x)\} = x$ , where

the  $\psi$  function is as defined in (1). As we have seen in the proof of Theorem 1,  $\psi$  is strictly increasing in the interval  $(S_\psi, \infty)$ , where  $S_\psi > \sup \Gamma_H$  and  $\psi'(S_\psi) = 0$ . Therefore, any real-valued solution  $f$  of  $\psi\{f(x)\} = x$  in  $(S_\psi, \infty)$  has to be the inverse of  $\psi$ , which is unique due to the strict monotonicity of  $\psi$  on  $(S_\psi, \infty)$ . Now suppose  $\Gamma_F$  is the support of the sample LSD  $F$ . We will show that there exists  $M_f > \sup \Gamma_F$  such that for any  $x > M_f$ , the function  $f_F$  is real-valued and it is a solution to (A.5) in the interval  $(S_\psi, \infty)$ . Thus it is also the unique such solution and the inverse of the  $\psi$  function in  $(S_\psi, \infty)$ .

Let  $x \in \mathbb{R}$ ,  $x > \sup \Gamma_F$  and  $z = x + iy \in \mathbb{C}^+$ . Now, as  $z \in \mathbb{C}^+$ ,  $f_F(z)$  is the unique solution to (A.5) in  $\mathbb{C}^+$ . Therefore, if we express  $f_F(z)$  as  $u(z) + iv(z)$ , then  $v(z) > 0$ . Also, the imaginary part of (A.5) can be written as

$$v(z) \left[ 1 - \gamma \int \frac{\lambda^2}{\{u(z) - \lambda\}^2 + v(z)^2} dH(\lambda) \right] = y.$$

Both  $v(z)$  and  $y$  being positive implies that

$$1 - \gamma \int \frac{\lambda^2}{\{u(z) - \lambda\}^2 + v(z)^2} dH(\lambda) > 0. \quad (\text{A.8})$$

Due to the continuity of  $f_F$  on the set  $\{z \in \mathbb{C}^+ : z = x + iy, x > \sup \Gamma_F\}$ ,

$$f_F(x) = \lim_{y \rightarrow 0^+} \frac{x + iy}{1 + \gamma \int \tau(x + iy - \tau)^{-1} dF(\tau)} = \frac{x}{1 + \gamma \int \tau(x - \tau)^{-1} dF(\tau)},$$

which is real-valued. Thus  $u(z) \rightarrow f_F(x)$  and  $v(z) \rightarrow 0$  as  $y \downarrow 0$ . Therefore as  $y \downarrow 0$ , the inequality (A.8) becomes

$$1 - \gamma \int \frac{\lambda^2}{\{f_F(x) - \lambda\}^2} dH(\lambda) > 0,$$

which implies  $\psi'\{f_F(x)\} > 0$ .

We can see that  $f_F(x)$  attains zero at  $\sup \Gamma_F$  and it is strictly and unboundedly increasing for  $x > \sup \Gamma_F$ . This ensures the existence of a threshold  $M_F > \sup \Gamma_F$  such that the function  $f_F$  maps the interval  $(M_F, \infty)$  to  $(S_\psi, \infty)$ . Therefore,  $f_F$  and  $\psi$  are both strictly increasing, continuous and bijective mappings between the intervals  $(M_F, \infty)$  and  $(S_\psi, \infty)$ . Since  $f_F(z)$  is the unique solution to (A.5) in  $\mathbb{C}^+$  when  $z \in \mathbb{C}^+$ ,  $f_F$  is also a solution to (A.5) in  $(S_\psi, \infty)$  when  $x > M_F$  due to the continuity of the left-hand side of (A.5) on the set  $\{f \in \mathbb{C}^+ : f = u + iv, u > S_\psi\}$ , which further implies that  $f_F$  is the inverse function of  $\psi$  on  $(S_\psi, \infty)$ .

The first part of Lemma B is proved as a corollary to Result 1 as  $\psi^{-1} = f_F$  on the domain of distant spikes, i.e.,  $(S_\psi, \infty)$ . For the second part, we first need to derive the expression of  $f'_F$ , and then derive the expression of  $\psi'$  in terms of  $f_F$  and  $F$ , viz.

$$f'_F(x) = f(x)\{1 + \gamma v_F(x)\}/x; \quad v_F(x) = \int \frac{\tau}{(x - \tau)^2} dF(\tau).$$

For a distant spike  $\lambda_k$ , using the expression of  $f'_F$ , we get

$$\frac{\lambda_k \psi'(\lambda_k)}{\psi(\lambda_k)} = \frac{\lambda_k}{\psi(\lambda_k) f'_F\{\psi(\lambda_k)\}} = \frac{1}{1 + \gamma f_F\{\psi(\lambda_k)\} \int \tau \{\psi(\lambda_k) - \tau\}^{-2} dF(\tau)} = g_F\{\psi(\lambda_k)\}.$$

As  $\psi(\lambda_k) > M_f$ ,  $g_F$  is continuous at  $\psi(\lambda)$ . Since  $d_k \xrightarrow{p} \psi(\lambda_k)$ ,

$$g_F(d_k) \xrightarrow{p} g_F\{\psi(\lambda_k)\} = \lambda_k \psi'(\lambda_k) / \psi(\lambda_k), \quad d_k g_F(d_k) / f_F(d_k) \xrightarrow{p} \psi'(\lambda_k).$$

This concludes the proof of Lemma B.  $\square$

**Proof of Theorem 3.** We have

$$\langle P_k, p_k \rangle^2 = \frac{1}{n^2 \lambda_k d_k} \langle X E_k, X e_k \rangle^2 = \frac{1}{\lambda_k d_k} (E_k^\top X^\top X e_k / n)^2 = \frac{1}{\lambda_k d_k} \left\{ E_k^\top \left( \sum_{i=1}^p d_i e_i e_i^\top \right) e_k \right\}^2 = \frac{d_k}{\lambda_k} \langle e_k, E_k \rangle^2.$$

Using the limits derived in Theorem 2 and Result 1, we deduce that

$$\left| \frac{d_k}{\lambda_k} \langle e_k, E_k \rangle^2 - \psi'(\lambda_k) \right| \xrightarrow{p} 0.$$

Using Lemma B, we conclude that

$$\left| \frac{d_k}{\lambda_k} \langle e_k, E_k \rangle^2 - \frac{d_k g_F(d_k)}{f_F(d_k)} \right| \xrightarrow{p} 0.$$

$\square$

**Proof of Theorem 4.** We show that the denominator  $E(p_{kj}^2)$  converges to  $\psi(\lambda_k)$  and the numerator  $E(q_k^2)$  converges to  $\lambda_k^2/\psi(\lambda_k)$ . The proof will be complete using the fact that  $d_k \xrightarrow{p} \psi(\lambda_k)$ . For the denominator, we have

$$\begin{aligned} E(p_{kj}^2) &= \frac{1}{n} E \left( \sum_{i=1}^n p_{ki}^2 \right) = \frac{1}{n} E \left\{ \sum_{i=1}^n (x_i^\top e_k)^2 \right\} \\ &= E(e_k^\top X^\top X e_k / n) = E \left\{ e_k^\top \left( \sum_{i=1}^p d_i e_i e_i^\top \right) e_k \right\} = E(d_k) \rightarrow \psi(\lambda_k). \end{aligned}$$

For the numerator, we have

$$E(q_k^2) = E\{(x_{new}^\top e_k)^2\} = E\{E(x_{new}^\top e_k)^2 | e_k\} = E\{\text{var}(x_{new}^\top e_k) | e_k\} = E(e_k^\top \Sigma_p e_k).$$

Now, using the notations in the proof of Theorem 2 and Lemma B, we have  $b_F(z) = \int (\tau - z)^{-1} dF(\tau)$  as the Stieltjes transform of the sample LSD and the function  $f_F$  defined as  $f_F(z) = z\{1 - \gamma - \gamma z b_F(z)\}^{-1}$ . Therefore,

$$b_F(z) = \frac{(1 - \gamma)f_F(z) - z}{\gamma z f_F(z)}.$$

The functions  $b_F$  and  $f_F$  can be extended to the real axis by defining the extensions as shown in the proof of Lemma B. Thus, for the sample eigenvalue  $d_k$  corresponding to the distant spike  $\lambda_k$  we have

$$b_F(d_k) = \frac{(1 - \gamma)f_F(d_k) - d_k}{\gamma d_k f_F(d_k)}.$$

According to Theorem 4 in [13], the limit of  $e_k^\top \Sigma_p e_k$  is given by  $d_k\{1 - \gamma - \gamma d_k b_F(d_k)\}^{-2}$ . Replacing the expression of  $b_F(d_k)$  in this limit, we get

$$|e_k^\top \Sigma_p e_k - f_F^2(d_k)/d_k| \xrightarrow{p} 0.$$

Using Result 1 and Lemma B, we have  $f_F^2(d_k)/d_k \xrightarrow{p} \lambda_k^2/\psi(\lambda_k)$ . Therefore, the limit of the numerator is given by

$$E(q_k^2) = E(e_k^\top \Sigma_p e_k) \rightarrow \lambda_k^2/\psi(\lambda_k).$$

This completes the proof of Theorem 4.  $\square$

## Appendix B. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jmva.2019.02.007>.

## References

- [1] C.A. Anderson, F.H. Pettersson, G.M. Clarke, L.R. Cardon, A.P. Morris, K.T. Zondervan, Data quality control in genetic case-control association studies, *Nat. Protoc.* 5 (2010) 1564–1573.
- [2] Z.D. Bai, J.W. Silverstein, No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices, *Ann. Probab.* 26 (1998) 316–345.
- [3] Z.D. Bai, J. Yao, On sample eigenvalues in a generalized spiked population model, *J. Multivariate Anal.* 106 (2012) 167–177.
- [4] J. Baik, J.W. Silverstein, Eigenvalues of large sample covariance matrices of spiked population models, *J. Multivariate Anal.* 97 (2006) 1382–1408.
- [5] M. Berkelaar, et al., IpSolve: Interface to 'Lp\_solve' v. 5.5 to Solve Linear/Integer Programs, 2015. R package version 5.6.13.
- [6] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
- [7] T.T. Cai, X. Han, G. Pan, Limiting laws for divergent spiked eigenvalues and largest non-spiked eigenvalue of sample covariance matrices, 2017. ArXiv e-prints.
- [8] X. Ding, Convergence of sample eigenvectors of spiked population model, *Commun. Statist. Theory Methods* 44 (2015) 3825–3840.
- [9] N. El Karoui, Spectrum estimation for large dimensional covariance matrices using random matrix theory, *Ann. Statist.* 36 (2008) 2757–2790.
- [10] V.L. Girko, Strong law for the eigenvalues and eigenvectors of empirical covariance matrices, *Random Oper. Stoch. Equ.* 4 (1996) 176–204.
- [11] I.M. Johnstone, On the distribution of the largest eigenvalue in principal components analysis, *Ann. Statist.* 29 (2001) 295–327.
- [12] I.M. Johnstone, A.Y. Lu, On consistency and sparsity for principal components analysis in high dimensions, *J. Amer. Statist. Assoc.* 104 (2009) 682–693.
- [13] O. Ledoit, S. P     , Eigenvectors of some large sample covariance matrix ensembles, *Probab. Theory Related Fields* 151 (2010) 233–264.
- [14] S. Lee, F. Zou, F.A. Wright, Convergence and prediction of principal component scores in high-dimensional settings, *Ann. Statist.* 38 (2010) 3605–3629.
- [15] S. Lee, F. Zou, F.A. Wright, Convergence of sample eigenvalues, eigenvectors, and principal component scores for ultra-high dimensional data, *Biometrika* 101 (2014) 484–490.
- [16] V.A. Mar     , L.A. Pastur, Distribution of eigenvalues for some sets of random matrices, *Math. USSR-Sb.* 1 (1967) 457.
- [17] X. Mestre, On the Asymptotic Behavior of Quadratic Forms of the Resolvent of Certain Covariance-Type Matrices, Technical Report, CTTC/RC/2006-001, Centre Tecnologic de Telecomunicacions de Catalunya, Barcelona, Spain, 2006.
- [18] X. Mestre, Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates, *IEEE Trans. Inform. Theory* 54 (2008) 5113–5129.
- [19] X. Mestre, On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices, *IEEE Trans. Signal Process.* 56 (2008) 5353–5368.

- [20] D. Paul, Asymptotics of sample eigenstruture for a large dimensional spiked covariance model, *Statist. Sinica* 17 (2007) 1617–1642.
- [21] A.L. Price, N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies, *Nature Genet.* 38 (2006) 904–909.
- [22] J. Silverstein, S. Choi, Analysis of the limiting spectral distribution of large dimensional random matrices, *J. Multivariate Anal.* 54 (1995) 295–309.
- [23] J.D. Storey, W. Xiao, J.T. Leek, R.G. Tompkins, R.W. Davis, Significance analysis of time course microarray experiments, *Proc. Nat. Acad. Sci. U.S.A.* 102 (2005) 12837–12842.