Testing in high-dimensional spiked models

Iain M. Johnstone and Alexei Onatski.

September 21, 2015

Abstract

We consider five different classes of multivariate statistical problems identified by James (1964). Each of these problems is related to the eigenvalues of $E^{-1}H$ where H and E are proportional to high-dimensional Wishart matrices. Under the null hypothesis, both Wisharts are central with identity covariance. Under the alternative, the non-centrality or the covariance parameter of H has a single eigenvalue, a spike, that stands alone. When the spike is larger than a case-specific phase transition threshold, one of the eigenvalues of $E^{-1}H$ separates from the bulk. This makes the alternative easily detectable, so that reasonable statistical tests have asymptotic power one. In contrast, when the spike is sub-critical, that is lies below the threshold, none of the eigenvalues separates from the bulk, which makes the testing problem more interesting from the statistical perspective. In such cases, we show that the log likelihood ratio processes parameterized by the value of the sub-critical spike converge to Gaussian processes with logarithmic correlation. We use this result to derive the asymptotic power envelopes for tests for the presence of a spike in the data representing each of the five cases in James' classification.

1 Introduction

High-dimensional multivariate models and methods, such as regression, principal components, and canonical correlation analysis, have become subject of much recent research. In contrast to the classical framework where the dimensionality is fixed, the current focus is on situations where the dimensionality diverges to infinity together with the sample size. In this context, *spiked models* that deviate from a reference model along a small fixed number of unknown directions have proven to be a fruitful research tool. A basic statistical question that arises in the analysis of spiked models is how to test for the presence of spikes in the data.

James (1964) arranges multivariate statistical problems in five different groups with broadly similar features. His classification corresponds to the five types of the hypergeometric functions ${}_pF_q$ that often occur in multivariate distributions. In this paper, we describe spiked models that represent each of James' classes, and derive the asymptotic behavior of the corresponding likelihood ratios, that is the ratios of the joint densities of the relevant data under the alternative hypothesis, which assumes the presence of the spikes, to that under the null of no spikes. In each of the cases, the relevant data consist of the maximal invariant statistic represented by eigenvalues of a large random matrix. We consider the asymptotic regime where the dimensionality of the data and the number of observations go to infinity proportionally.

We find that the measures corresponding to the joint distributions of the eigenvalues under the alternative hypothesis and under the null are mutually contiguous when the values of the spikes are below a phase transition threshold. The value of the threshold depends on the problem's type. Furthermore, we find that the log likelihood ratio processes parametrized by the values of the spikes are asymptotically Gaussian, with logarithmic mean and autocovariance functions. These findings allow us to compute the asymptotic power envelopes for the tests for the presence of spikes in five multivariate models representing each of James' classes.

Our analysis is based on the classical results that assume Gaussianity. All the likelihood ratios that we study correspond to the joint densities of the solutions to the basic equation of classical multivariate statistics,

$$\det (H - \lambda E) = 0, \tag{1}$$

where H and E are proportional to Wishart matrices.

The five different cases that we study are: 1) E is a known deterministic matrix, and H is a central Wishart matrix with covariance equal to a low-rank perturbation of E; 2) both E and H are central Wisharts with unknown covariance matrices that differ by a matrix of low rank; 3) E is a known deterministic matrix, and H is a non-central Wishart matrix with covariance equal to E and with a low-rank non-centrality; 4) E is a central Wishart matrix, while H is a non-central one with the same unknown covariance matrix and with a low-rank non-centrality; 5) E is a central Wishart, while H is a non-central Wishart conditionally on a random low-rank non-centrality parameter. These five cases can be linked via sufficiency and invariance arguments to a principal components problem, a signal detection problem, hypotheses testing in multivariate regression with known and with unknown error covariance, and a canonical correlation problem, respectively. We briefly discuss the links in the next section of the paper.

The main steps of our asymptotic analysis are the same for all the five cases. The likelihood ratios have explicit forms that involve hypergeometric functions of two high-dimensional matrix arguments. However, the low-rank nature of the alternatives that we consider ensures that one of the arguments have low rank. For tractability, we focus on the special case of rank-one alternatives. In such case, using the recent result of Dharmawansa and Johnstone (2014), we represent the hypergeometric function of two high-dimensional matrix arguments in the form of a contour integral that involves a scalar hypergeometric function of the same type. Then we deform the contour of integration so that the integral becomes amenable to Laplace approximation analysis (see Olver (1997), chapter 4).

Using the Laplace approximation technique, we show that the log likelihood ratios are asymptotically equivalent to random quadratic functions of the spike parameters. The randomness in the quadratic function enters via a linear spectral statistic of a large random matrix of either sample covariance or F-ratio type. Using CLT for the linear spectral statistics, established by Bai and Silverstein (2004) for the sample-covariance-type random matrices and by Zheng (2012) for the F-ratio-type random matrices, we derive the asymptotic Gaussianity and obtain the mean and the autocovariance functions of the log likelihood ratio processes.

The derived asymptotics of the log likelihood processes shows that the corresponding statistical experiments do not converge to Gaussian shift experiments. In other words, the experiments that consist of observing the solutions to equation (1) parameterized by the values of the spikes under the alternative hypothesis are not of the Locally Asymptotically Normal (LAN) type. This implies that there are no ready-to-use optimality results associated with LAN experiments that can be applied in our setting. However at the fundamental level, the derived asymptotics of the log likelihood ratio processes is all that is needed for the asymptotic analysis of the risk of the corresponding statistical decisions.

In this paper, we use the derived asymptotics together with the Neyman-Pearson lemma and Le Cam's third lemma (see van der Vaart (1998)), to find simple analytic expressions for the asymptotic power envelopes for the statistical tests of the null hypothesis of no spikes in the data. The form of the envelope is different depending on whether both H and E in the corresponding equation (1) are Wisharts or only H is Wishart whereas E is deterministic.

For most of the cases, as the value of the spike under the alternative increases, the envelope, at first, rises very slowly. Then, as the spike approaches the phase transition, the rise quickly accelerates and the envelope 'hits' unity at the threshold. However, in cases of two Wisharts and when the dimensionality is not much smaller than the degrees of freedom of E, the envelope rises much faster. In such cases, the information in all the eigenvalues of $E^{-1}H$ might be useful for detecting population spikes which lie far below the phase transition threshold.

A type of the analysis performed in this paper has been previously implemented in the study of the principal components case by Onatski et al (2013). Our work extends theirs to the remaining four cases in James' classification of multivariate statistical problems. One of the hardest challenges in such an extension is the rigorous implementation of the Laplace approximation step. With this goal in mind, we have developed asymptotic approximations to the hypergeometric functions ${}_1F_1$ and ${}_2F_1$ which are uniform in certain domains of the complex plane.

A trivial observation that the solutions to equation (1) can be interpreted as the eigenvalues of random matrix $E^{-1}H$ relates our work to the vast literature on the spectrum of large random matrices. We refer the reader to Bai and Silverstein (2006) for a recent book-long treatment of the subject. Three extensively studied classical ensembles of random matrices are the Gaussian, Laguerre and Jacobi ensembles (see Mehta (2004)). However, only the Laguerre and Jacobi ensembles are relevant for the five scenarios for (1) that correspond to James' five-fold classification of multivariate statistical problems. This prompts us to search for a "missing" class in James' system that could be linked to the Gaussian ensemble.

Such a class is easy to obtain by taking the limit of $\sqrt{n_1}\,(H - I_p)$ as $n_1 \to \infty$, where n_1 and p are H's degrees of freedom and dimensionality, respectively. The corresponding statistical problem can be called "symmetric matrix denoising". Under the null hypothesis, the observations are given by a $p \times p$ matrix $Z/\sqrt{p}$ with Z from the Gaussian Orthogonal Ensemble. Under the alternative, the observations are given by $Z/\sqrt{p} + \Phi$, where Φ is a deterministic symmetric matrix of low rank. We call this situation "case zero", and add it to James' classification. We derive the asymptotics of the corresponding log likelihood ratio and obtain the related asymptotic power envelope.

Many existing results in the random matrix literature do not require that the data are Gaussian. This suggests that some results about tests for the presence of the spikes in the data may remain valid without the Gaussianity. One may for example consider H and E in (1) that, although have the form of sample covariance matrices, do not come from the underlying Gaussian distribution, and study the properties of the corresponding tests. We leave this line of research to the future.

Since the explicit form of the joint distribution of the solutions to (1) is only known in the Gaussian case, it seems unlikely that one would be able to completely summarize the asymptotic behavior of the corresponding non-Gaussian statistical experiments. We hope that the results of this paper, that provide such a summary under the Gaussianity, can serve as a useful benchmark for the future studies that would relax our assumptions.

The rest of the paper is organized as follows. In the next section, we relate the five different cases of equation (1) to the classical multivariate statistical problems representing different cells of James' (1964) five-fold classification system. In Section 3, we obtain explicit expressions for the likelihood ratios. Section 4 represents the likelihood ratios in the form of contour integrals. Section 5 performs the Laplace approximation analysis. Section 6 derives the asymptotic power envelopes. Section 7 concludes. Technical proofs are given in the Appendix.

2 Links to classical statistical problems

Case 1 corresponds to the problem of using n_1 i.i.d. $N_p(0, \Omega)$ (p-dimensional Gaussian) observations to test the null hypothesis that the population covariance Ω equals a given matrix Σ. The alternative of interest is

$$\Omega = \Sigma + \psi\theta\psi'$$

with unknown $\theta > 0$ and ψ, where ψ is normalized so that $\left\|\Sigma^{-1/2}\psi\right\| = 1$.

Without loss of generality, we may assume that $\Sigma = I_p$. Then under the null, the data are isotropic noise, whereas under the alternative, the first principal component explains a larger portion of the variation than the other principal components. We therefore label Case 1 as the 'principal components analysis' (PCA) case.

The null and the alternative hypotheses can be formulated in terms of the spectral 'spike' parameter θ as

$$H_0 : \theta_0 = 0 \text{ and } H_1 : \theta_0 = \theta > 0, \tag{2}$$

where θ_0 is the true value of the 'spike'. This testing problem remains invariant under the multiplication of the $p \times n_1$ data matrix from the left and from the right by orthogonal matrices, and under the corresponding transformation in the parameter space. A maximal invariant statistic consists of the solutions $\lambda_1 \geq ... \geq \lambda_p$ of equation (1) with H equal to the sample covariance matrix and $E = \Sigma$. We restrict attention to the invariant tests. Therefore, the relevant data are summarized by $\lambda_1, ..., \lambda_p$.

Case 2 is represented by the problem of testing the equality of covariance matrices, Ω and Σ, corresponding to two independent p-dimensional zero-mean Gaussian samples of sizes n_1 and n_2. Throughout the paper, we shall assume that

$$p \leq \min \{n_1, n_2\}.$$

The assumption $p \leq n_2$ is made to ensure the almost sure invertibility of matrix E in (1), whereas the assumption $p \leq n_1$ is made to reduce the number of various situations which need to be considered. Such a reduction makes our exposition more concise.

Returning to Case 2, the alternative hypothesis is the same as in the PCA case. Similar invariance considerations lead to tests based on the eigenvalues of the F-ratio of the sample covariance matrices. Matrix H from (1) equals the sample covariance corresponding to the observations that might contain a 'signal' responsible for the covariance spike, whereas matrix E equals the other sample covariance matrix. We label Case 2 as the 'signal detection' (SigD) case. In this case, we find it more convenient to work with the p solutions to the equation

$$\det \left(H - \lambda \left(E + \frac{n_1}{n_2} H \right) \right) = 0, \tag{3}$$

which we also denote $\lambda_1 \geq ... \geq \lambda_p$ to make the notations as uniform across the different cases as possible. Note that as the number of observations in the second sample, n_2, diverges to infinity while n_1 and p are held constant, equation (3) reduces to equation (1), E converges to Σ, and SigD reduces to PCA.

Cases 3 and 4 occur in multivariate regression

$$Y = X\beta + \varepsilon$$

when the goal is to test linear restrictions on the matrix of coefficients β. Case 3 corresponds to the situation where the covariance matrix Σ of the i.i.d. Gaussian rows of the error matrix ε is known. We label this case as 'regression with known variance' (REG_0). Case 4 corresponds to the unknown Σ, and we label it as 'regression with unknown variance' (REG).

As explained in Muirhead (1982), pp. 433-434, the problem of testing linear restrictions on β can be cast in the canonical form, where the matrix of transformed response variables is split into three parts, Y_1^*, Y_2^*, and Y_3^*. Matrix Y_1^* is $n_1 \times p$, where p is the number of response variables and n_1 is the number of restrictions. Under the null hypothesis, $\mathbb{E}Y_1^* = 0$, whereas under the alternative,

$$\mathbb{E}Y_1^* = \sqrt{n_1 \theta} \varphi \psi', \tag{4}$$

where $\theta > 0$, $\left\| \Sigma^{-1/2} \psi \right\| = 1$, and $\|\varphi\| = 1$. Matrices Y_2^* and Y_3^* are $(q - n_1) \times p$ and $(T - q) \times p$, respectively, where q is the number of regressors and T is the number of observations. These matrices have, respectively, unrestricted and zero means under both the null and the alternative.

For REG_0, sufficiency and invariance arguments lead to tests based on the solutions $\lambda_1, ..., \lambda_p$ of (1) with

$$H = Y_1^{*\prime} Y_1^* / n_1 \text{ and } E = \Sigma.$$

These solutions represent a multivariate analog of the difference between the sum of squared residuals in the restricted and unrestricted regressions. For REG, similar arguments lead to tests based on the p solutions $\lambda_1, ..., \lambda_p$ of (3) with

$$H = Y_1^{*\prime} Y_1^* / n_1 \text{ and } E = Y_3^{*\prime} Y_3^* / n_2,$$

where $n_2 = T - q$. These solutions represent a multivariate analog of the ratio of the difference between the sum of squared residuals in the restricted and unrestricted regressions to the sum of squared residuals in the restricted regression. Note that, as $n_2 \to \infty$ while n_1 and p are held constant, REG reduces to REG_0.

Case 5 occurs in situations where the researcher would like to test for the independence between Gaussian vectors $x_t \in \mathbb{R}^p$ and $y_t \in \mathbb{R}^{n_1}$, given zero mean observations with $t = 1, ..., n_1 + n_2$. Partition the population and sample covariance matrices of the observations $(x_t', y_t')'$ as

$$\begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \text{ and } \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix},$$

respectively. Under the null hypothesis, $\Sigma_{xy} = 0$. The alternative of interest is

$$\Sigma_{xy} = \sqrt{\frac{n_1\theta}{n_1\theta + n_1 + n_2}}\psi\varphi',$$

where the vectors of nuisance parameters $\psi \in \mathbb{R}^p$ and $\varphi \in \mathbb{R}^{n_1}$ are normalized so that

$$\left\|\Sigma_{xx}^{-1/2}\psi\right\| = \left\|\Sigma_{yy}^{-1/2}\varphi\right\| = 1.$$

The peculiar form of the expression under the square root is chosen so as to simplify various expressions in the analysis that follow.

The test can be based on the sample canonical correlations $\lambda_1, ..., \lambda_p$, which are solutions to (1) with

$$H = S_{xy}S_{yy}^{-1}S_{yx} \text{ and } E = S_{xx}.$$

We label Case 5 as the 'canonical correlation analysis' (CCA) case. Remarkably, the sample canonical correlations also solve (3) with different H and E, such that E is a central Wishart matrix and H is a non-central Wishart matrix conditionally on a random non-centrality parameter (for details, see Theorem 11.3.2 of Muirhead (1982)).

Finally, as discussed in the Introduction, we also consider Case 0, which we label as the 'symmetric matrix denoising' (SMD) case. Given a $p \times p$ matrix $X = \Phi + Z/\sqrt{p}$, where Z is a noise matrix from the Gaussian Orthogonal Ensemble (GOE), a researcher would like to make inference about a symmetric rank-one "signal" matrix $\Phi = \psi\theta\psi'$. Recall, that a symmetric matrix Z belongs to GOE if its diagonal and sub-diagonal entries are independently distributed as

$$Z_{ii} \sim N(0, 2) \text{ and } Z_{ij} \sim N(0, 1) \text{ if } i > j.$$

The null and the alternative hypotheses are given by (2). The nuisance $\psi \in \mathbb{R}^p$ is normalized so that $\|\psi\| = 1$. The problem remains invariant under the multiplication of X from the left by an orthogonal matrix, and from the right by its transpose. A maximal invariant statistic consists of the solutions $\lambda_1, ..., \lambda_p$ to (1) with $H = X$ and $E = I_p$. We consider tests based on $\lambda_1, ..., \lambda_p$.

The SMD case can be viewed as a degenerate version of all of the above cases. For example, consider REG_0 with

$$\mathbb{E}Y_1^* = \sqrt{(p/n_1)^{1/2}\, n_1 \theta \varphi \psi'},$$

so that the original value of the spike θ (see equation (4)) is scaled by $(p/n_1)^{1/2}$. Suppose now that n_1 diverges to infinity while p is held constant. Then, by a Central Limit Theorem (CLT),

$$\Sigma^{-1/2} H \Sigma^{-1/2} - I_p = Z/\sqrt{n_1} + \sqrt{p/n_1}\eta\theta\eta' + o_{\mathrm{P}}\left(n_1^{-1/2}\right), \tag{5}$$

where Z belongs to GOE and $\eta = \Sigma^{-1/2}\psi$. On the other hand, equation (1) is equivalent to

$$\det\left(\Sigma^{-1/2} H \Sigma^{-1/2} - \lambda I_p\right) = 0. \tag{6}$$

Multiplying it by $\sqrt{n_1/p}$ and using (5), we see that equation (6) degenerates to

$$\det\left(Z/\sqrt{p} + \eta\theta\eta' - \mu I_p\right) = 0 \text{ with } \mu = \sqrt{n_1/p}\,(\lambda - 1)\,.$$

Hence, REG_0 degenerates to SMD.

For the reader's convenience, we summarize links between the different cases and the definitions of the corresponding matrices H and E in Figure 1. We denote the p-dimensional Wishart distribution with n degrees of freedom, covariance parameter Σ, and non-centrality parameter Ψ as $W_p(n, \Sigma, \Psi)$. Recall that, if $A = B'B$, where the $n \times p$ matrix B is $N(M, I_n \otimes \Sigma)$, then $A \sim W_p(n, \Sigma, \Psi)$ with the non-centrality $\Psi = \Sigma^{-1}M'M$. Notation $W_p(n, \Sigma)$ is used for the central Wishart distribution. Without loss of generality, we assume that $\Sigma = I_p$.

All the cases eventually degenerate to SMD via sequential asymptotic links. Cases SMD, PCA, and REG_0, forming the upper half of the diagram, correspond to random H and deterministic E. The cases in the lower half of the diagram correspond to both H and E being random. Cases PCA and SigD are "parallel" to cases REG_0 and REG in the sense that the alternative hypothesis is characterized by a rank one perturbation of the covariance and of the non-centrality parameter of H for the former and for the latter two cases, respectively. Case CCA "stands alone" because of the different structure of H and E. As discussed above, CCA can be reinterpreted in terms of H and E such that E is Wishart, but H is a non-central Wishart only after conditioning on a random non-centrality parameter.

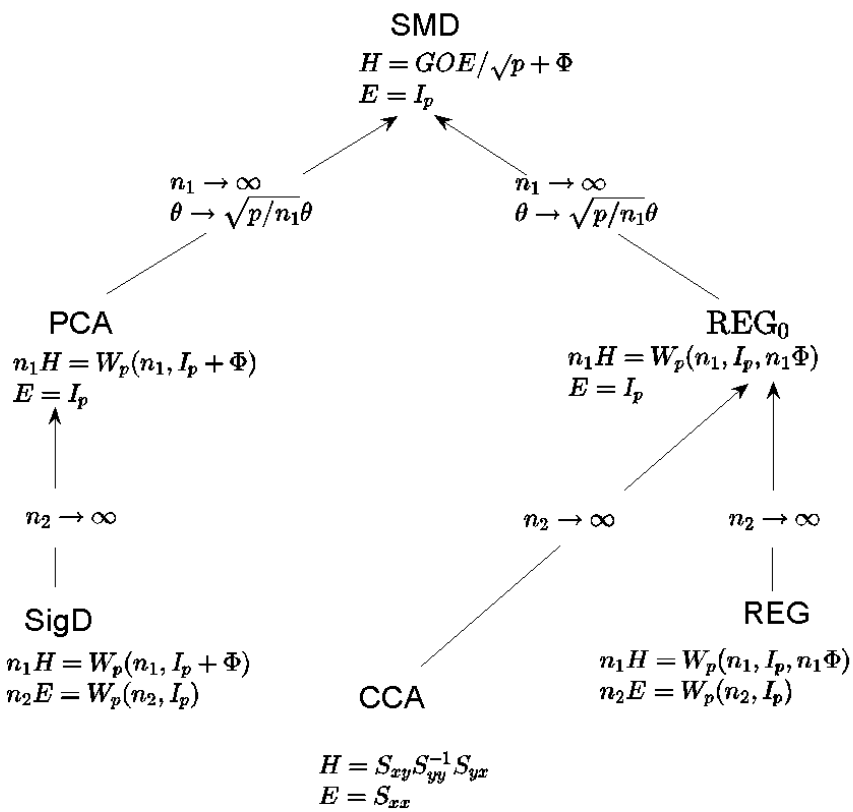

Figure 1: Matrices H and E, and links between the different cases. Matrix Φ has the form $\psi\theta\psi'$ with $\theta \geq 0$ and $\|\psi\| = 1$.

Case	${}_{\bar{p}}F_{\bar{q}}$	$\alpha(\theta)$	a	b	Ψ_{11}
SMD	${}_0F_0$	$\exp(-p\theta^2/4)$	_	_	$\theta p/2$
PCA	${}_0F_0$	$(1+\theta)^{-n_1/2}$	_	_	$\theta n_1/(2(1+\theta))$
SigD	${}_1F_0$	$(1+\theta)^{-n_1/2}$	$n/2$	_	$\theta n_1/(n_2(1+\theta))$
REG$_0$	${}_0F_1$	$\exp(-n_1\theta/2)$	_	$n_1/2$	$\theta n_1^2/4$
REG	${}_1F_1$	$\exp(-n_1\theta/2)$	$n/2$	$n_1/2$	$\theta n_1^2/(2n_2)$
CCA	${}_2F_1$	$(1+n_1\theta/n)^{-n/2}$	$(n/2, n/2)$	$n_1/2$	$\theta n_1^2/(n_2^2+n_2 n_1(1+\theta))$

Table 1: Parameters of the explicit expression (7) for the likelihood ratios. Here $n \equiv n_1 + n_2$.

3 The likelihood ratios

Our goal is to study the asymptotic behavior of the likelihood ratios, which are defined as the ratios of the joint density of $\lambda_1, ..., \lambda_p$ under the alternative to that under the null hypothesis, where both densities are evaluated at the observed values of the λ's. Let

$$\Lambda = \text{diag}\{\lambda_1, ..., \lambda_p\},$$

and let us denote the likelihood ratio corresponding to particular case 'Case' = 'SMD', 'PCA', etc. as $L^{(Case)}(\theta; \Lambda)$. Then

$$L^{(Case)}(\theta; \Lambda) = \alpha(\theta)\, {}_{\bar{p}}F_{\bar{q}}(a, b; \Psi, \Lambda), \tag{7}$$

where Ψ is a p-dimensional matrix diag$\{\Psi_{11}, 0, ..., 0\}$, and the values of Ψ_{11}, $\alpha(\theta)$, $\bar{p}$, $\bar{q}$, a, and b are as given in Table 1.

We prove that $L^{(SMD)}(\theta; \Lambda)$ is as in (7) in the Appendix. For PCA, the explicit form of the likelihood ratio is derived in Onatski et al (2013). For SigD, REG$_0$, and REG, the expressions (7) with the parameters given in Table 1 follow, respectively, from equations (65), (68), and (73) of James (1964). For CCA, the expression is a corollary of Theorem 11.3.2 of Muirhead (1982).

Recall that hypergeometric functions of two matrix arguments Ψ and Λ are defined as

$${}_{\bar{p}}F_{\bar{q}}(a, b; \Psi, \Lambda) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\kappa \vdash k} \frac{(a_1)_\kappa \ldots (a_{\bar{p}})_\kappa}{(b_1)_\kappa \ldots (b_{\bar{q}})_\kappa} \frac{C_\kappa(\Psi) C_\kappa(\Lambda)}{C_\kappa(I_p)},$$

where $a = (a_1, ..., a_{\bar{p}})$ and $b = (b_1, ..., b_{\bar{q}})$ are parameters, κ are partitions of the integer k, $(a_j)_\kappa$ and $(b_i)_\kappa$ are the generalized Pochhammer symbols, and C_κ are the zonal polynomials (see Muirhead (1982), Definition 7.3.2). As mentioned in the Introduction, James' (1964) classification of the multivariate statistical problems is based on the type of ${}_{\bar{p}}F_{\bar{q}}$ that occur in related probability distributions. The function ${}_0F_0$ of *exponential type* corresponds to the first class represented by PCA; the function ${}_1F_0$ of *binomial type* corresponds to the second class represented by SigD; the function ${}_0F_1$ of *Bessel type* is associated with the third class represented by REG_0; the *confluent hypergeometric function* ${}_1F_1$ is associated with the fourth class represented by REG; and the *Gaussian hypergeometric function* ${}_2F_1$ corresponds to the fifth class represented by CCA. Note that some links between the cases illustrated in Figure 1 can also be established via asymptotic relations between the hypergeometric functions in the different rows of Table 1. For example, the links $\text{REG}\mapsto\text{REG}_0$ and $\text{SigD}\mapsto\text{PCA}$ as $n_2 \to \infty$ while p and n_1 are held constant follow from the confluence relations (see, for example, chapter 3.5 of Luke (1969))

$$
\begin{aligned}
{}_0F_1(b; \Psi, \Lambda) &= \lim_{a\to\infty} {}_1F_1\left(a, b; a^{-1}\Psi, \Lambda\right) \text{ and} \\
{}_0F_0(\Psi, \Lambda) &= \lim_{a\to\infty} {}_1F_0\left(a; a^{-1}\Psi, \Lambda\right).
\end{aligned}
$$

In the next section, we shall study the asymptotic behavior of the likelihood ratios (7) as n_1, n_2, and p go to infinity so that

$$
c_1 \equiv \frac{p}{n_1} \to \gamma_1 \in (0, 1) \text{ and } c_2 \equiv \frac{p}{n_2} \to \gamma_2 \in (0, 1]. \tag{8}
$$

We denote this asymptotic regime as $\mathbf{n}, p \to_\gamma \infty$, where $\mathbf{n} = \{n_1, n_2\}$ and $\gamma = \{\gamma_1, \gamma_2\}$. To make our exposition as uniform as possible, we use this notation for all the cases, even though the simpler ones, such as SMD, do not refer to $\mathbf{n}$. In the Conclusion, we briefly discuss possible extensions of our analysis to the situations with $\gamma_1 \geq 1$.

We are interested in the asymptotics of the likelihood ratios under the null hypothesis, that is when the true value of the spike, θ_0, equals zero. Before turning to the next section, let us provide a relevant background on the asymptotics of Λ. Under the null, $\lambda_1, ..., \lambda_p$ are the eigenvalues of $GOE/\sqrt{p}$ in SMD case; of $W_p(n_1, I_p)/n_1$ in PCA and REG_0 cases; and of a scaled (by a factor of n_2/n_1) p-dimensional multivariate beta matrix with parameters $n_1/2$ and $n_2/2$ in SigD,

Case	$F_\gamma^{\lim}$	density for $\beta_- \le \lambda \le \beta_+$	$\beta_\pm$	Threshold $\bar{\theta}$
SMD	SC	$\frac{1}{2\pi}\sqrt{(\beta_+ - \lambda)(\lambda - \beta_-)}$	± 2	1
PCA REG$_0$	MP	$\frac{1}{2\pi\gamma_1\lambda}\sqrt{(\beta_+ - \lambda)(\lambda - \beta_-)}$	$\left(1 \pm \sqrt{\gamma_1}\right)^2$	$\sqrt{\gamma_1}$
SigD REG CCA	W	$\frac{\gamma_1+\gamma_2}{2\pi\gamma_1\lambda(\gamma_1-\gamma_2\lambda)}\sqrt{(\beta_+ - \lambda)(\lambda - \beta_-)}$	$\gamma_1\left(\frac{\rho\pm1}{\rho\pm\gamma_2}\right)^2$	$\frac{\gamma_2+\rho}{1-\gamma_2}$

Table 2: The semi-circle, Marchenko-Pastur, and (scaled) Wachter distributions. Here $\rho = \sqrt{\gamma_1 + \gamma_2 - \gamma_1\gamma_2}$. In the case where $\gamma_1 > 1$, which is not considered in this paper, the Marchenko-Pastur and Wachter distributions will also have mass $(\gamma_1 - 1)/\gamma_1$ at zero. Column 'Threshold $\bar{\theta}$' reports the values of the phase transition thresholds.

REG, and CCA cases. For a definition of the multivariate beta, see Muirhead (1982), p. 110.

Let

$$\hat{F}^{(Case)}(\lambda) = \frac{1}{p}\sum_{j=1}^{p} 1\{\lambda_j \le \lambda\}$$

be the empirical distribution of $\lambda_1, ..., \lambda_p$. As is well known (see Bai (1999)), as $\mathbf{n}, p \to_\gamma \infty$, $\hat{F}^{(Case)}$ almost surely (a.s.) weakly converges

$$\hat{F}^{(Case)} \Rightarrow F_\gamma^{\lim},$$

where $F_\gamma^{\lim}$ is the semi-circle distribution F^{SC} in SMD case; the Marchenko-Pastur distribution $F_{\gamma_1}^{MP}$ in PCA and REG$_0$ cases; and the (scaled) Wachter distribution F_γ^W in SigD, REG, and CCA cases. Table 2 reports the explicit forms of these limiting distributions. Note that the cumulative distribution functions $F_\gamma^{\lim}(\lambda)$ are linked in the sense that $F_\gamma^W(\lambda) \to F_{\gamma_1}^{MP}(\lambda)$ when $\gamma_2 \to 0$ and $F_{\gamma_1}^{MP}\left(\sqrt{\gamma_1}\lambda + 1\right) \to F^{SC}(\lambda)$ when $\gamma_1 \to 0$.

For what follows it will be important that the centered *linear spectral statistics*

$$\sum_{j=1}^{p}\varphi(\lambda_j) - p\int \varphi(\lambda)\, \mathrm{d}F_{\mathbf{c}}^{\lim}(\lambda), \tag{9}$$

where φ is a 'well-behaved' function, converge in distribution to Gaussian random variables. The corresponding CLTs are established in Bai and Yao (2005), Bai and Silverstein (2004), and Zheng (2012) for the cases of the semi-circle, Marchenko-

Pastur, and Wachter limiting distributions, respectively. Note that the centering constant is defined in terms of $F_{\mathbf{c}}^{\lim}$, where $\mathbf{c} = \{c_1, c_2\}$. That is, the "correct centering" can be computed using the densities from Table 2, where γ_1 and γ_2 are replaced by $c_1 \equiv p/n_1$ and $c_2 \equiv p/n_2$, respectively.

Finally, let us note the behavior of the largest eigenvalue λ_1 under the alternative hypothesis. As is well known, λ_1 a.s. converges to the upper boundary of support of $F_{\gamma}^{\lim}$ as long as θ remains below the phase transition threshold $\bar{\theta}$. The value of the threshold is reported in the last column of Table 2. When $\theta > \bar{\theta}$, λ_1 separates from 'the bulk' of the other eigenvalues and a.s. converges to a point strictly above the upper boundary of the support of $F_{\gamma}^{\lim}$. For details, we refer the reader to Maïda (2007), Baik and Silverstein (2006), Nadakuditi and Silverstein (2010), Onatski (2007), Dharmawansa et al (2014a), and Bao et al (2014) for cases SMD, PCA, SigD, REG_0, REG, and CCA, respectively.

The fact that λ_1 converges to different limits under the null and under the alternative hypothesis sheds light on the behavior of the likelihood ratio when θ is above the phase transition threshold. In such cases, which can be called the cases of super-critical θ, the likelihood ratio degenerates. The sequences of measures corresponding to the distributions of Λ under the null and under super-critical alternatives are asymptotically mutually singular as $\mathbf{n}, p \to_{\gamma} \infty$ (see Montanari et al (2014) and Onatski et al (2013) for a detailed analysis of SMD and PCA cases). In contrast, as we shall show below, the sequences of measures corresponding to the distributions of Λ under the null and under sub-critical alternatives (θ is below the threshold) are mutually contiguous, and the likelihood ratio converges to a Gaussian process.

4 Contour integral representation

Asymptotic behavior of the likelihood ratios (7) depends on that of ${}_{\bar{p}}F_{\bar{q}}\,(a, b; \Psi, \Lambda)$. There is a large and well established literature on the asymptotics of ${}_{\bar{p}}F_{\bar{q}}\,(a, b; \Psi, \Lambda)$ when the parameters and the norm of the matrix arguments grow while the dimensionality of the latter remains fixed (see Muirhead (1978) for a review). In contrast, relatively little is known about the asymptotic regime that allows the dimensionality of the matrix arguments Ψ, Λ diverge to infinity. In this paper, we investigate such an asymptotic regime. We exploit the fact that, since we study single-spiked models, the matrix argument Ψ has rank one. This allows us to represent ${}_{\bar{p}}F_{\bar{q}}\left(a,b;\Psi,\Lambda\right)$ in the form of a contour integral of a hypergeometric function with a single scalar argument. Such a representation implies contour integral representations for the corresponding likelihood ratios, which we summarize in the following lemma. The results of the lemma are used below to derive the asymptotics of the likelihood ratios via the Laplace approximation.

In what follows, we omit the superscripts '(Case)' and 'lim' for quantities such as $L^{(Case)}\left(\theta;\Lambda\right)$, $\hat{F}^{(Case)}\left(\lambda\right)$, and $F_{\mathbf{c}}^{\lim}\left(\lambda\right)$ to simplify our notation. However, we shall use these superscripts to identify particular instances, when necessary.

Lemma 1 *Assume that* $p\leq\min\left\{n_1,n_2\right\}$. *Let* $\mathcal{K}$ *be a contour in the complex plane* $\mathbb{C}$ *that starts at* $-\infty$, *encircles* 0 *and* $\lambda_1,...,\lambda_p$ *counterclockwise, and returns to* $-\infty$. *Then*

$$L\left(\theta;\Lambda\right)=\frac{\Gamma\left(s+1\right)\alpha\left(\theta\right)q_s}{\Psi_{11}^s2\pi\mathrm{i}}\int_{\mathcal{K}}{}_{\bar{p}}F_{\bar{q}}\left(a-s,b-s;\Psi_{11}z\right)\prod_{j=1}^{p}\left(z-\lambda_j\right)^{-1/2}\mathrm{d}z,\quad(10)$$

where $s=p/2-1$, *the values of* $\alpha\left(\theta\right)$, Ψ_{11}, a, b, $\bar{p}$, *and* $\bar{q}$ *for the different cases are given in Table 1;* $a-s$ *and* $b-s$ *denote vectors with elements* a_j-s *and* b_j-s, *respectively; the hypergeometric function under the integral is the standard hypergeometric function of a scalar argument; and*

$$q_s=\prod_{j=1}^{\bar{p}}\frac{\Gamma\left(a_j-s\right)}{\Gamma\left(a_j\right)}\prod_{i=1}^{\bar{q}}\frac{\Gamma\left(b_i\right)}{\Gamma\left(b_i-s\right)}.$$

In cases SigD and CCA, we require, in addition, that the contour $\mathcal{K}$ *does not intersect* $\left[\Psi_{11}^{-1},\infty\right)$, *which ensures the analyticity of the integrand in an open subset of* $\mathbb{C}$ *that includes* $\mathcal{K}$.

The statement of the lemma immediately follows from Proposition 1 of Dharmawansa and Johnstone (2014) and from equation (7). Our next step is to apply the Laplace approximation to integrals (10). To this end, we shall transform the right hand side of (10) so that it has a "Laplace form"

$$L\left(\theta;\Lambda\right)=\sqrt{\pi p}\frac{1}{2\pi\mathrm{i}}\int_{\mathcal{K}}\exp\left\{-pf(z)\right\}g(z)\mathrm{d}z.\quad(11)$$

Leaving $\sqrt{\pi p}/\left(2\pi\mathrm{i}\right)$ separate from $g(z)$ allows us to choose $f(z)$ and $g(z)$ that are bounded in probability, and makes some of the expressions below more compact.

Case	$2f_I$	$g_I/(1+o(1))$
SMD	$1+\theta^2/2+\ln\theta$	θ
PCA	$1+\frac{1-c_1}{c_1}\ln(1+\theta)+\ln\frac{\theta}{c_1}$	$\theta c_1^{-1}(1+\theta)^{-1}$
SigD	$2f_I^{(PCA)}-1+\ln\frac{c_1+c_2}{c_1}-\frac{r^2}{c_1c_2}\ln\frac{r^2}{c_1+c_2}$	$\theta c_1^{-2}(1+\theta)^{-1}r(c_1+c_2)^{1/2}$
REG$_0$	$1+\frac{\theta+c_1}{c_1}+\ln\frac{\theta}{c_1}+\frac{1-c_1}{c_1}\ln(1-c_1)$	$\theta c_1^{-1}(1-c_1)^{-1/2}$
REG	$2f_I^{(REG_0)}-1+\ln\frac{c_1+c_2}{c_1}-\frac{r^2}{c_1c_2}\ln\frac{r^2}{c_1+c_2}$	$\theta c_1^{-2}(1-c_1)^{-1/2}r(c_1+c_2)^{1/2}$
CCA	$2f_I^{(REG)}-1-\frac{\theta}{c_1}-\frac{r^2}{c_1c_2}\ln\frac{r^2}{c_1 l}$	$\theta c_1^{-3}(1-c_1)^{-1/2}r^2(c_1+c_2)l^{-1}$

Table 3: Values of $2f_I$ and $g_I/(1+o(1))$ for the different cases. The terms $o(1)$ do not depend on θ and converge to zero as $\mathbf{n}, p \to_\gamma \infty$. The term r^2 is defined as $r^2 = c_1 + c_2 - c_1c_2$. The term $l \equiv l(\theta)$ is defined as $l(\theta) = 1 + (1+\theta)c_2/c_1$.

In order to apply the Laplace approximation, we shall deform the contour of integration so that it passes through a critical point z_0 of $f(z)$ and is such that $\operatorname{Re} f(z)$ is strictly increasing as z moves away from z_0 along the contour, at least in a vicinity of z_0.

4.1 The Laplace form

We shall transform (10) to (11) in three steps. As a result, functions f and g will have the forms of a sum and a product,

$$f(z) = f_I + f_{II}(z) + f_{III}(z) \text{ and}$$
$$g(z) = g_I \times g_{II}(z) \times g_{III}(z),$$

where f_I and g_I do not depend on z.

First, using the definitions of $\alpha(\theta)$, q_s, Ψ_{11} and employing Stirling's approximation, we obtain a decomposition

$$\frac{\Gamma(s+1)\,\alpha(\theta)\,q_s}{\sqrt{\pi p}\Psi_{11}^s} = \exp\{-pf_I\}\,g_I, \tag{12}$$

where g_I remains boudned as $\mathbf{n}, p \to_\gamma \infty$. The values of $2f_I$ and g_I are given in Table 3. It should be noted that $f_I^{(REG)}, f_I^{(CCA)} \to f_I^{(REG_0)}$ and $f_I^{(SigD)} \to f_I^{(PCA)}$ as $c_2 \to 0$.

Next, we consider the decomposition

$$\prod_{j=1}^{p}(z-\lambda_j)^{-1/2} = \exp\{-pf_{II}(z)\}\, g_{II}(z), \tag{13}$$

where

$$2f_{II}(z) = \int \ln(z-\lambda)\, \mathrm{d}F_{\mathbf{c}}(\lambda), \tag{14}$$

and

$$g_{II}(z) = \exp\left\{-\frac{p}{2}\int \ln(z-\lambda)\, \mathrm{d}\left(\hat{F}(\lambda) - F_{\mathbf{c}}(\lambda)\right)\right\}. \tag{15}$$

For $f_{II}(z)$ and $g_{II}(z)$ to be well-defined we need z not to belong to the support of $F_{\mathbf{c}}$, which we assume.[1] Note that $g_{II}(z)$ is the exponent of a linear spectral statistic, which converges to a Gaussian random variable as $\mathbf{n}, p \to_\gamma \infty$ under the null hypothesis. Since $F_{\mathbf{c}}^{W}(\lambda) \to F_{c_1}^{MP}(\lambda)$ as $c_2 \to 0$, we have $f_{II}^{(SigD)}(z) = f_{II}^{(REG)}(z) = f_{II}^{(CCA)}(z)$ converging to $f_{II}^{(PCA)}(z) = f_{II}^{(REG_0)}(z)$.

Finally, we obtain a decomposition

$${}_{\bar{p}}F_{\bar{q}}\left(a-s, b-s; \Psi_{11} z\right) = \exp\{-pf_{III}(z)\}\, g_{III}(z). \tag{16}$$

For SMD, PCA, and SigD, the corresponding ${}_{\bar{p}}F_{\bar{q}}$ can be expressed in terms of elementary functions, and we set

$$2f_{III}(z) = \begin{cases} -z\theta & \text{for SMD} \\ -z\theta/\left(c_1(1+\theta)\right) & \text{for PCA} \\ \ln\left[1 - c_2 z\theta/\{c_1(1+\theta)\}\right] r^2/(c_1 c_2) & \text{for SigD} \end{cases}, \tag{17}$$

and

$$g_{III}(z) = \begin{cases} 1 & \text{for SMD and PCA} \\ \left[1 - c_2 z\theta/\{c_1(1+\theta)\}\right]^{-1} & \text{for SigD} \end{cases}. \tag{18}$$

As $c_2 \to 0$, $f_{III}^{(SigD)}(z)$ converges to $f_{III}^{(PCA)}(z)$. Since, as has been shown above, a similar convergence holds for f_I and f_{II}, we have $f^{(SigD)}(z) \to f^{(PCA)}(z)$ as $c_2 \to 0$. Combining (14) and (17) with the information supplied by Table 3, we also see that $f^{(PCA)}(z) \to f^{(SMD)}(z)$ as $c_1 \to 0$ after the transformations $\theta \mapsto \sqrt{c_1}\theta$ and $z \mapsto \sqrt{c_1} z + 1$.

Unfortunately, for REG_0, REG, and CCA, the corresponding ${}_{\bar{p}}F_{\bar{q}}$ do not admit

[1] By definition, contour $\mathcal{K}$ encircles the support of $\hat{F}$, and hence $z \in \mathcal{K}$ does not belong to such support.

exact representations in terms of elementary functions. Therefore, we shall consider their asymptotic approximations instead. Let

$$m = (n_1 - p)/2 \text{ and } \varepsilon = (n - p)/(n_1 - p).$$

Further, let

$$\eta_j = \begin{cases} z\theta/(1-c_1)^2 & \text{for } j = 0 \\ z\theta c_2/[c_1(1-c_1)] & \text{for } j = 1 \\ z\theta c_2^2/[c_1^2 l(\theta)] & \text{for } j = 2 \end{cases}, \tag{19}$$

where

$$l(\theta) = 1 + (1+\theta)\, c_2/c_1. \tag{20}$$

With this notation, we have

$${}_{\bar{p}}F_{\bar{q}} = \begin{cases} {}_0F_1\left(m+1; m^2\eta_0\right) \equiv F_0 & \text{for REG}_0 \\ {}_1F_1\left(m\varepsilon+1, m+1; m\eta_1\right) \equiv F_1 & \text{for REG} \\ {}_2F_1\left(m\varepsilon+1, m\varepsilon+1; m+1; \eta_2\right) \equiv F_2 & \text{for CCA} \end{cases}. \tag{21}$$

The function F_0 can be expressed in terms of the modified Bessel function of the first kind $I_m(\cdot)$ as (see Abramowitz and Stegun (1964), equation 9.6.47)

$$F_0 = \Gamma(m+1)\left(m^2\eta_0\right)^{-m/2} I_m\left(2m\eta_0^{1/2}\right). \tag{22}$$

This representation allows us to use a known uniform asymptotic approximation of the Bessel function (see Abramowitz and Stegun (1964), equation 9.7.7) to obtain the following lemma. Let

$$\varphi_0(t) = \ln t - t - \eta_0/t + 1 \text{ and } t_0 = \left(1 + \sqrt{1+4\eta_0}\right)/2. \tag{23}$$

Further, for any $\delta > 0$, let $\Omega_{0\delta}$ be the set of $\eta_0 \in \mathbb{C}$ such that

$$|\arg \eta_0| \leq \pi - \delta, \text{ and } \eta_0 \neq 0.$$

Lemma 2 *As $m \to \infty$, we have*

$$F_0 = (1+4\eta_0)^{-1/4} \exp\{-m\varphi_0(t_0)\}(1+o(1)). \tag{24}$$

The convergence $o(1) \to 0$ holds uniformly with respect to $\eta_0 \in \Omega_{0\delta}$ for any $\delta > 0$.

We would like to point out that the right hand side of (24) can be formally linked, via (22), to the saddle-point approximation of the integral representation (see Watson (1944), p. 181)

$$I_m\left(2m\eta_0^{1/2}\right)=\frac{\eta_0^{m/2}e^m}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)}\exp\left\{-m\varphi_0\left(t\right)\right\}t^{-1}\mathrm{d}t.$$

Point t_0 can be interpreted as a saddle point of $\varphi_0\left(t\right)$, and the term $\left(1+4\eta_0\right)^{-1/4}$ in (24) can be interpreted as a factor of $\left(\varphi_0''\left(t_0\right)\right)^{-1/2}$.

To obtain uniform asymptotic approximations to functions F_1 and F_2, we use the contour integral representations (see Olver et al (2010), equations 13.4.9 and 15.6.2)

$$F_j=\frac{C_m}{2\pi\mathrm{i}}\int_0^{(1+)}\exp\left\{-m\varphi_j\left(t\right)\right\}\psi_j\left(t\right)\mathrm{d}t, \tag{25}$$

where

$$C_m=\frac{\Gamma\left(m+1\right)\Gamma\left(m\left(\varepsilon-1\right)+1\right)}{\Gamma\left(m\varepsilon+1\right)}, \tag{26}$$

$$\varphi_j(t)=\begin{cases}-\eta_j t-\varepsilon\ln t+\left(\varepsilon-1\right)\ln\left(t-1\right) & \text{for } j=1\\ -\varepsilon\ln\left(t/\left(1-\eta_j t\right)\right)+\left(\varepsilon-1\right)\ln\left(t-1\right) & \text{for } j=2\end{cases}, \tag{27}$$

and

$$\psi_j\left(t\right)=\begin{cases}\left(t-1\right)^{-1} & \text{for } j=1\\ \left(t-1\right)^{-1}\left(1-\eta_j t\right)^{-1} & \text{for } j=2\end{cases}. \tag{28}$$

For $j=2$, the contour does not encircle $1/\eta_2$, and the representation is valid for η_2 such that $\left|\arg\left(1-\eta_2\right)\right|<\pi$. We obtain the following lemma by deriving a saddle-point approximation to the integral in (25). The relevant saddle points are

$$t_j=\begin{cases}\frac{1}{2\eta_j}\left\{\eta_j-1+\sqrt{\left(\eta_j-1\right)^2+4\varepsilon\eta_j}\right\} & \text{for } j=1\\ \frac{1}{2\eta_j\left(\varepsilon-1\right)}\left\{-1+\sqrt{1+4\varepsilon\left(\varepsilon-1\right)\eta_j}\right\} & \text{for } j=2\end{cases}. \tag{29}$$

We shall need the following additional notation. Let

$$\omega_j=\arg\varphi_j''\left(t_j\right)+\pi \text{ and } \omega_{0j}=\arg\left(t_j-1\right), \tag{30}$$

where the branches of $\arg\left(\cdot\right)$ are chosen so that $\left|\omega_j+2\omega_{0j}\right|\leq\pi/2$. Further, for any small $\delta>0$ let $\Omega_{1\delta}$ be the set of $\left(\varepsilon,\eta_1\right)\in\mathbb{R}\times\mathbb{C}$ such that $\delta\leq\varepsilon-1\leq 1/\delta$, and

$$\operatorname{Re}\eta_1\geq-2\varepsilon+1,\ \operatorname{dist}\left(\eta_1,\mathbb{R}\backslash\left[0,\infty\right)\right)\geq\delta,\ \left|\eta_1\right|\leq 1/\delta.$$

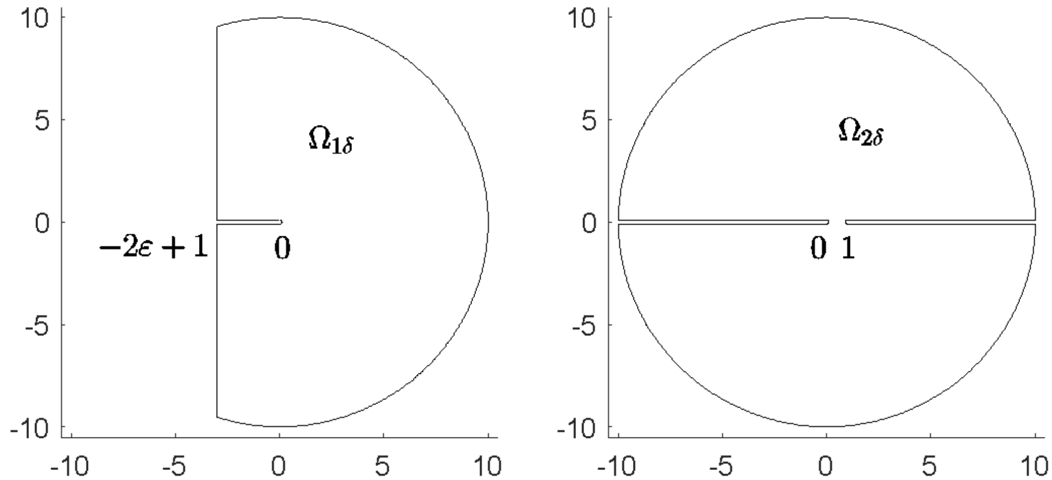

Figure 2: Cross-sections of the sets $\Omega_{j\delta}$ for $\varepsilon = 2$ and $\delta = 0.1$. The horizontal and vertical axes correspond to the real and purely imaginary numbers, respectively.

Similarly, let $\Omega_{2\delta}$ be the set of $(\varepsilon, \eta_2) \in \mathbb{R} \times \mathbb{C}$ such that $\delta \leq \varepsilon - 1 \leq 1/\delta$, and

$$\operatorname{dist}(\eta_2, \mathbb{R} \backslash [0,1]) \geq \delta, \ \ |\eta_2| \leq 1/\delta.$$

Here, for any $A \subseteq \mathbb{C}$ and $B \subseteq \mathbb{C}$, $\operatorname{dist}(A, B) = \inf_{a \in A, b \in B} |a - b|$. Figure 2 shows cross-sections of $\Omega_{1\delta}$ and $\Omega_{2\delta}$ for fixed ε.

Lemma 3 *As $m \to \infty$, we have for $j = 1, 2$*

$$F_j = C_m \psi_j(t_j) e^{-\mathrm{i}\omega_j/2} \left|2\pi m \varphi_j''(t_j)\right|^{-1/2} \exp\{-m\varphi_j(t_j)\}(1 + o(1)). \tag{31}$$

The convergence $o(1) \to 0$ holds uniformly with respect to $(\varepsilon, \eta) \in \Omega_{j\delta}$ for any $\delta > 0$.

Point-wise asymptotic approximation (31) was established in Passemier et al (2014) for $j = 1$, and in Paris (2013a,b) for $j = 2$. However, those papers do not study the uniformity of the approximation error, which is important for our analysis. A proof of Lemma 3 is available from the authors upon request. We shall report it elsewhere.

Using Lemmas 2, 3, and Stirling's approximation

$$C_m = \frac{\sqrt{\pi p(1 - c_1)}}{r} \exp\{m(\varepsilon - 1)\ln(\varepsilon - 1) - m\varepsilon \ln \varepsilon\}(1 + o(1)) \tag{32}$$

we set the components of the "Laplace form" (16) of ${}_{\bar{p}}F_{\bar{q}}$ for cases REG_0, REG, and CCA as follows

$$2f_{III}(z) = \begin{cases} \frac{1-c_1}{c_1}\varphi_0(t_0) & \text{for REG}_0 \\ \frac{1-c_1}{c_1}\left(\varphi_j(t_j) + \varepsilon\ln\varepsilon - (\varepsilon-1)\ln(\varepsilon-1)\right) & \text{for REG and CCA} \end{cases} \tag{33}$$

and

$$g_{III}(z) = \begin{cases} (1+4\eta_0)^{-1/4}(1+o(1)) & \text{for REG}_0 \\ \sqrt{c_1/r^2}e^{-\mathrm{i}\omega_j/2}\left|\varphi_j''(t_j)\right|^{-1/2}\psi_j(t_j)(1+o(1)) & \text{for REG and CCA} \end{cases}. \tag{34}$$

To express t_j and η_j in terms of z, one should use (29) and (19). We do not need to know how exactly $o(1)$ in (34) depend on z. For our purposes, the knowledge of the fact that $o(1)$ are analytic functions of η_j that converge to zero uniformly with respect to $(\varepsilon, \eta_j) \in \Omega_{j\delta}$ is sufficient. The analyticity of $o(1)$ follows from the analyticity of the functions on the left hand sides, and of the factors of $1+o(1)$ on the right hand sides of the equations (24) and (31).

Using the definitions of φ_j and t_j, it is straightforward to verify that $f_{III}^{(REG)}(z)$ and $f_{III}^{(CCA)}(z)$ converge to $f_{III}^{(REG_0)}(z)$ as $c_2 \to 0$. Since, as has been shown above, a similar convergence holds for f_I and f_{II}, we have $f^{(REG)}(z), f^{(CCA)}(z) \to f^{(REG_0)}(z)$ as $c_2 \to 0$. Elementary calculations that use equations (14), (23), (33) together with the explicit forms of $f_I^{(REG_0)}$ and $f_I^{(SMD)}$ given in Table 3 show that $f^{(REG_0)}(z) \to f^{(SMD)}(z)$ as $c_1 \to 0$ after transformations $\theta \mapsto \sqrt{c_1}\theta$ and $z \mapsto \sqrt{c_1}z + 1$.

4.2 Contours of steep descent

We shall now show how to deform contours $\mathcal{K}$ in (11) into the contours of steep descent. First, we find saddle points of functions $f(z)$ for each of the six cases. Note that the derivative of $f_{II}(z)$ equals minus half of the Stieltjes transform $m_{\mathbf{c}}(z)$ of the corresponding limiting spectral distribution $F_{\mathbf{c}}$. Although the Stieltjes transform is formally defined on $\mathbb{C}^+$, the definition remains valid on the part of the real line outside the support $[b_-, b_+]$ of $F_{\mathbf{c}}$. Since we assume that $\gamma_1 \leq 1$, $F_{\mathbf{c}}$ does not have any non-trivial mass at 0 for sufficiently large $\mathbf{n}$ and p.

To find saddle points of $f(z)$ we solve equation

$$m_{\mathbf{c}}(z) = 2\mathrm{d}f_{III}(z)/\mathrm{d}z. \tag{35}$$

In the Appendix, we find real solutions to (35), z_0, that satisfy inequality $z_0 > b_+$. These solutions are reported in the following lemma.

Lemma 4 *Let b_+ be the upper boundary of support of $F_{\mathbf{c}}$, and $\bar{\theta}$ be the threshold corresponding to F_γ as given in Table 2. Then, for $\theta \in \left(0, \bar{\theta}\right)$ and sufficiently large $\mathbf{n}$ and p as $\mathbf{n}, p \to_\gamma \infty$,*

$$z_0 = \begin{cases} \theta + 1/\theta & \text{for SMD} \\ (1+\theta)(\theta + c_1)/\theta & \text{for PCA and REG}_0 \\ (1+\theta)(\theta + c_1)/[\theta l(\theta)] & \text{for SigD, REG, and CCA} \end{cases} \tag{36}$$

satisfy inequality $z_0 > b_+$ and solve equation (35).

As $c_2 \to 0$ while c_1 stays constant, the value of z_0 for SigD, REG, and CCA converges to that for PCA and REG_0. The latter value in its turn converges to the value of z_0 for SMD when $c_1 \to 0$, after the transformations $\theta \mapsto \sqrt{c_1}\theta$ and $z_0 \mapsto \sqrt{c_1} z_0 + 1$. Precisely, solving equation

$$\sqrt{c_1} z_0 + 1 = (1 + \sqrt{c_1}\theta)(\sqrt{c_1}\theta + c_1)/(\sqrt{c_1}\theta)$$

for z_0 and taking limit as $c_1 \to 0$ yields $z_0 = \theta + 1/\theta$.

For the rest of the paper, assume that $\theta \in \left(0, \bar{\theta}\right)$. We deform contour $\mathcal{K}$ in (11) so that it passes through the saddle point z_0 as follows. Let $\mathcal{K} = \mathcal{K}_+ \cup \mathcal{K}_-$, where $\mathcal{K}_-$ is the complex conjugate of $\mathcal{K}_+$ and $\mathcal{K}_+ = \mathcal{K}_1 \cup \mathcal{K}_2$. For SMD, PCA, and SigD, let

$$\mathcal{K}_1 = \{z_0 + \mathrm{i}t : 0 \le t \le 2z_0\} \text{ and} \tag{37}$$

$$\mathcal{K}_2 = \{x + \mathrm{i}2z_0 : -\infty < x \le z_0\}. \tag{38}$$

The deformed contour is shown on Figure 3.

Note that the singularities of the integrand in (11) are situated at $z = \lambda_j$ (plus an additional singularity at $z = c_1(1+\theta)/(\theta c_2) < z_0$ for SigD). Since $\lambda_1 \overset{a.s.}{\to} \beta_+$ and $z_0 > b_+$, inequality $z_0 > \lambda_1$ must hold with probability approaching one as $\mathbf{n}, p \to_\gamma \infty$. Therefore by Cauchy's theorem, the deformation of the contour does not change the value of $L(\theta; \Lambda)$ with probability approaching one as $\mathbf{n}, p \to_\gamma \infty$.

Strictly speaking, the deformation of the contour is not continuous because $\mathcal{K}_+$ does not approach $\mathcal{K}_-$ at $-\infty$. In particular, in contrast to the original contour, the deformed one is not "closed" at $-\infty$. Nevertheless, such an "opening up" at

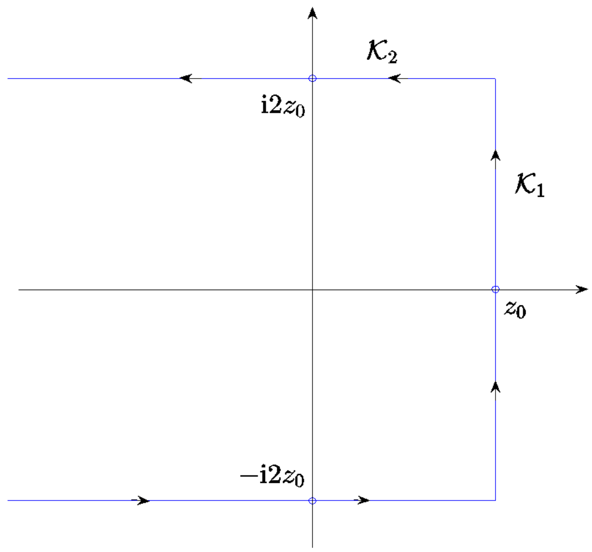

Figure 3: Deformed contour $\mathcal{K}$ for SMD, PCA, and SigD.

$-\infty$ does not lead to the change of the value of the integral because the integrand converges fast to zero by absolute value as $\operatorname{Re} z \to -\infty$.

Remark 5 *In the event of asymptotically negligeable probability that the deformed contour $\mathcal{K}$ does not encircle all λ_j, we not only loose the equality (11) but also face the difficulty that function $g(z)$ ceases to be well defined as the definition of $g_{II}(z)$ contains a logarithm of a non-positive number. To eliminate any ambiguity, if such an event holds we shall redefine $g_{II}(z)$ as unity.*

For REG_0 and CCA, let

$$z_1 = \begin{cases} -\left(1-c_1\right)^2 / \left[4\theta\right] & \text{for REG}_0 \\ -c_1\left(1-c_1\right)^2 l\left(\theta\right) / \left[4\theta r^2\right] & \text{for CCA} \end{cases},$$

and let

$$\mathcal{K}_1 = \left\{z_1 + \left|z_0 - z_1\right| \exp\left\{\mathrm{i}\gamma\right\} : \gamma \in \left[0, \pi/2\right]\right\} \text{ and}$$
$$\mathcal{K}_2 = \left\{z_1 - x + \left|z_0 - z_1\right| \exp\left\{\mathrm{i}\pi/2\right\} : x \geq 0\right\}.$$

The corresponding contour $\mathcal{K}$ is shown on Figure 4. Similarly to the SMD, PCA and SigD cases, the deformation of the contour in (11) to $\mathcal{K}$ does not change the value of $L\left(\theta; \Lambda\right)$ with probability approaching one as $\mathbf{n}, p \to_\gamma \infty$.

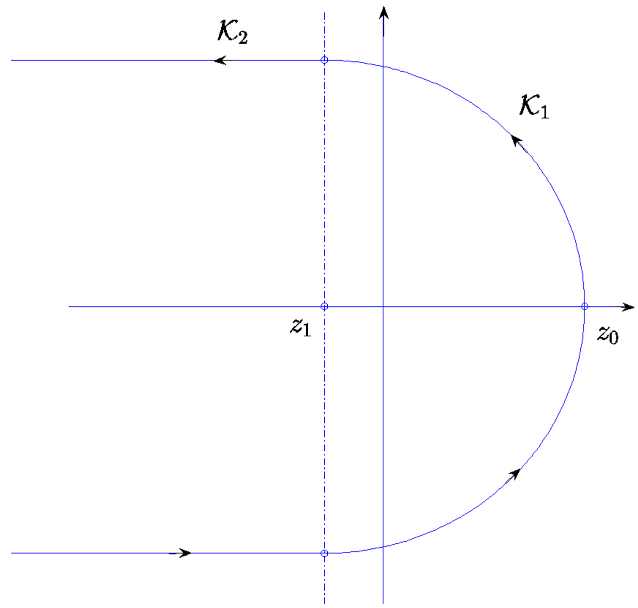

Figure 4: Deformed contour $\mathcal{K}$ for REG_0 and CCA.

For REG, deformed contour $\mathcal{K}$ in z-plane is simpler to describe as an image of a contour $\mathcal{C}$ in τ-plane, where $\tau = \eta_1 t_1$ with

$$\eta_1 = z\theta c_2 / \left[c_1 \left(1 - c_1\right)\right] \tag{39}$$

and t_1 as defined in (29). Let $\mathcal{C} = \mathcal{C}_+ \cup \mathcal{C}_-$, where $\mathcal{C}_-$ is the complex conjugate of $\mathcal{C}_+$ and $\mathcal{C}_+ = \mathcal{C}_1 \cup \mathcal{C}_2$, and let

$$\mathcal{C}_1 = \left\{-\varepsilon + |\tau_0 + \varepsilon| \exp\left\{\mathrm{i}\gamma\right\} : \gamma \in [0, \pi/2]\right\} \text{ and}$$
$$\mathcal{C}_2 = \left\{-\varepsilon - x + |\tau_0 + \varepsilon| \exp\left\{\mathrm{i}\pi/2\right\} : x \geq 0\right\},$$

where $\tau_0 = \left(\theta + c_1\right) / \left(1 - c_1\right)$.

Using (39) and the identity

$$\eta_1 = \tau\left(\tau + 1\right) / (\tau + \varepsilon), \tag{40}$$

we obtain

$$z = \frac{c_1\left(1 - c_1\right)}{\theta c_2} \frac{\tau\left(\tau + 1\right)}{\tau + \varepsilon}. \tag{41}$$

We define the deformed contour $\mathcal{K}$ in z-plane as the image of $\mathcal{C}$ under the trasformation $\tau \to z$ given by (41). The parts $\mathcal{K}_+, \mathcal{K}_-, \mathcal{K}_1$ and $\mathcal{K}_2$ of $\mathcal{K}$ are defined as the images of the corresponding parts of $\mathcal{C}$. Note that τ_0 is transformed to z_0 so that $\mathcal{K}$ passes through the saddle point z_0.

The following lemma is proven in the Appendix. It shows that $\mathcal{K}_1$ are contours of steep descent of $-\operatorname{Re} f(z)$ for all the six cases, SMD, PCA, SigD, REG_0, REG, and CCA.

Lemma 6 *For any of the six cases that we study, as z moves along the corresponding $\mathcal{K}_1$ away from z_0, $-\operatorname{Re} f(z)$ is strictly decreasing.*

5 Laplace approximation

The goal of this section is to derive Laplace approximations to the integrals

$$L(\theta; \Lambda) = \sqrt{\pi p} \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{K}} \exp\{-pf(z)\} g(z) \mathrm{d}z$$

for the six cases that we study. First, consider a general integral

$$I_{p,\omega} = \int_{\mathcal{K}_{p,\omega}} e^{-p\phi_{p,\omega}(z)} \chi_{p,\omega}(z) \mathrm{d}z,$$

where $p \to \infty$, ω is a k-dimensional parameter that belongs to a subset Ω of $\mathbb{R}^k$, $\mathcal{K}_{p,\omega}$ is a path in $\mathbb{C}$ that starts at $a_{p,\omega}$ and ends at $b_{p,\omega}$, and for sufficiently large p, $\phi_{p,\omega}(z)$ is a single-valued holomorphic function of z in a domain $T_{p,\omega}$ that contains $\mathcal{K}_{p,\omega}$.

We allow $\chi_{p,\omega}(z)$ to be a random element of the normed space of continuous functions on $\mathcal{K}_{p,\omega}$ with the supremum norm. Furthermore, we suppose that for any $\delta > 0$, there exists $\bar{p}$ such that for any $p > \bar{p}$, $\chi_{p,\omega}(z)$ is a single-valued holomorphic function of z in the domain $T_{p,\omega}$ with probability larger than $1-\delta$. In what follows, we shall omit subscripts p and ω from the notation $\phi_{p,\omega}$, $\chi_{p,\omega}$, $\mathcal{K}_{p,\omega}$, etc. to make it lighter.

Suppose that $\phi'(z) = 0$ at z_0 which is an interior point of $\mathcal{K}$, and suppose that $\operatorname{Re} \phi(z)$ is strictly increasing as z moves away from z_0 along the path. In other words, the path $\mathcal{K}$ is a contour of steep descent of $-\operatorname{Re} \phi(z)$. Denote a closed segment of $\mathcal{K}$ contained between z_1 and z_2 as $[z_1, z_2]_{\mathcal{K}}$. Similarly denote the segments that exclude one or both endpoints as $[z_1, z_2)_{\mathcal{K}}$, $(z_1, z_2]_{\mathcal{K}}$, and $(z_1, z_2)_{\mathcal{K}}$. Let β be the limiting value of $\arg(z - z_0)$ on the principal branch as $z \to z_0$ along

$(z_0, b)_{\mathcal{K}}$. Finally, let ϕ_s and χ_s with $s = 0, 1, ...$ be the coefficients in the power series representations

$$\phi(z) = \sum_{s=0}^{\infty} \phi_s (z - z_0)^s, \;\; \chi(z) = \sum_{s=0}^{\infty} \chi_s (z - z_0)^s. \tag{42}$$

We assume that there exist positive constants $C_1, ..., C_4$ that do not depend on p and on ω, such that for all $\omega \in \Omega$, for sufficiently large p :

A0 The length of the path $\mathcal{K}$ is bounded, uniformly over $\omega \in \Omega$ and all sufficiently large p. Furthermore,

$$\sup_{z \in (z_0, b)_{\mathcal{K}}} |z - z_0| > C_1, \text{ and } \sup_{z \in (a, z_0)_{\mathcal{K}}} |z - z_0| > C_1$$

A1 Functions $\phi(z)$ and $\chi(z)$ are holomorphic in the ball $|z - z_0| \leq C_1$

A2 The coefficient ϕ_2 satisfies $C_2 \leq |\phi_2| \leq C_3$

A3 The third derivative of $\phi(z)$ satisfies inequality

$$\sup_{|z - z_0| \leq C_1} \left| \mathrm{d}^3 \phi(z) / \mathrm{d}z^3 \right| \leq C_4$$

A4 For any positive $\varepsilon < C_1$, which does not depend on p and ω, and for all $z_1 \in \mathcal{K}$ such that $|z_1 - z_0| = \varepsilon$, there exist positive constants C_5, C_6, such that

$$\mathrm{Re}\,(\phi(z_1) - \phi_0) > C_5 \text{ and } |\mathrm{Im}\,(\phi(z_1) - \phi_0)| < C_6$$

A5 For a subset Θ of $\mathbb{C}$ that consists of all points whose Euclidean distance from $\mathcal{K}$ is no larger than C_1,

$$\sup_{z \in \Theta} |\chi(z)| = O_{\mathrm{P}}(1)$$

as $p \to \infty$, where $O_{\mathrm{P}}(1)$ is uniform in $\omega \in \Omega$.

The following lemma is a fairly straightforward extension of Theorem 7.1 of Olver (1997), p. 127 to the situation where functions $\phi(z)$, $\chi(z)$ and the contour $\mathcal{K}$ depend on p and ω. In Olver's original theorem, which uses different notation, both the functions and the contour are fixed. A proof of the extension is available from the authors upon request.

Case	Value of D_2	Case	Value of D_2
SMD	$1-\theta^2$	REG_0	$c_1\left(1+c_1+2\theta\right)\left(c_1-\theta^2\right)$
PCA	$c_1\left(c_1-\theta^2\right)\left(1+\theta\right)^2$	REG	$c_1 h\left(c_1+\theta+\left(1+\theta\right)l\right)/l^4$
SigD	$r^2 h\left(1+\theta\right)^2/l^4$	CCA	$c_1^2 h\left(2\left(c_1+\theta\right)+l\left(1-c_1\right)\right)/\left(l^3\left(c_1+c_2\right)\right)$

Table 4: The values of $D_2 \equiv \theta^2(-2\mathrm{d}^2 f(z_0)/\mathrm{d}z^2)^{-1}$. Here $l \equiv l(\theta)$ is as defined in (20) and $h \equiv h(\theta) = c_1 + c_2(1+\theta)^2 - \theta^2$.

Lemma 7 *Under assumptions A0-A5, for any positive integer κ, as $p \to \infty$, we have*

$$I_{p,\omega} = 2e^{-p\phi_0}\left[\sum_{s=0}^{\kappa-1}\Gamma\left(s+\frac{1}{2}\right)\frac{a_{2s}}{p^{s+1/2}} + \frac{O_{\mathrm{P}}(1)}{p^{\kappa+1/2}}\right],$$

where $O_{\mathrm{P}}(1)$ is uniform in $\omega \in \Omega$ and the coefficients a_j can be expressed through ϕ_i and χ_i defined above. In particular we have $a_0 = \chi_0/[2\phi_2^{1/2}]$, where $\phi_2^{1/2} =$ $\exp\left\{\left(\log|\phi_2| + \arg\phi_2\right)/2\right\}$ *with the branch of* $\arg\phi_2$ *chosen so that* $|\arg\phi_2 + 2\beta| \leq \pi/2$.

We use Lemma 7 to obtain the Laplace approximation to

$$L_1(\theta;\Lambda) = \sqrt{\pi p}\frac{1}{2\pi\mathrm{i}}\int_{\mathcal{K}_1\cup\bar{\mathcal{K}}_1} e^{-pf(z)}g(z)\mathrm{d}z. \tag{43}$$

Then we show that $L_1(\theta;\Lambda)$ asymptotically dominates the "residual" $L(\theta;\Lambda) - L_1(\theta;\Lambda)$. For this analysis, it is important to know the values of $f(z_0)$ and $\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$. We derive them in the Appendix. It turns out that as long as $\theta \in \left[0,\bar{\theta}\right)$, $f(z_0) = 0$ for all the six cases that we study. The values of $\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$ are all negative. The explicit form of $D_2 \equiv \theta^2\left(-2\mathrm{d}^2 f(z_0)/\mathrm{d}z^2\right)^{-1}$, which is somewhat shorter than that for $\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$ is reported in Table 4. We formulate the main result of this section in the following theorem. Its proof is given in the Appendix.

Theorem 8 *Suppose that the null hypothesis holds, that is, $\theta_0 = 0$. Let $\bar{\theta}$ be the threshold corresponding to F_γ as given in Table 2, and let ε be an arbitrarily small fixed positive number. Then for any $\theta \in \left(0,\bar{\theta}-\varepsilon\right]$, as $\mathbf{n}, p \to_\gamma \infty$, we have*

$$L(\theta;\Lambda) = \frac{g(z_0)}{\sqrt{-2\mathrm{d}^2 f(z_0)/\mathrm{d}z^2}} + O_{\mathrm{P}}\left(p^{-1}\right), \tag{44}$$

where $O_{\mathrm{P}}\left(p^{-1}\right)$ *is uniform in* $\theta \in \left(0, \bar{\theta}-\varepsilon\right]$ *and the principal branch of the square root is taken.*

6 Asymptotics of LR

Combining the results of Theorem 8 with the definitions of $g(z)$ and the values of $-2\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$ (given in Table 4), it is straightforward to establish the following theorem. Let

$$\Delta_p(\theta) = p \int \ln\left(z_0 - \lambda\right) \mathrm{d}\left(\hat{F}\left(\lambda\right) - F_{\mathbf{c}}\left(\lambda\right)\right).$$

In accordance with the remark made above, we define $\Delta_p(\theta)$ as zero in the event of asymptotically negligeable probability that $z_0 \leq \lambda_1$.

Theorem 9 *Suppose that the null hypothesis holds, that is* $\theta_0 = 0$*. Let* $\bar{\theta}$ *be the threshold corresponding to* F_γ *as given in Table 2, and let* ε *be an arbitrarily small fixed positive number. Then for any* $\theta \in \left(0, \bar{\theta}-\varepsilon\right]$, *as* $\mathbf{n}, p \to_\gamma \infty$, *we have*

$$L\left(\theta; \Lambda\right) = \exp\left\{-\frac{1}{2}\Delta_p(\theta) + \frac{1}{2}\ln\left(1 - \left[\delta_p\left(\theta\right)\right]^2\right)\right\}\left(1 + o_{\mathrm{P}}(1)\right),$$

where

$$\delta_p\left(\theta\right) = \begin{cases} \theta & \text{for SMD} \\ \theta/\sqrt{c_1} & \text{for PCA and REG}_0 \\ \theta r/\left(c_1 l\left(\theta\right)\right) & \text{for SigD, REG, and CCA} \end{cases}$$

and $o_{\mathrm{P}}(1)$ *is uniform in* $\theta \in \left(0, \bar{\theta}-\varepsilon\right]$.

Statistic $\Delta_p(\theta)$ is a linear spectral statistic. As follows from the CLT derived by Bai and Yao (2005), Bai and Silverstein (2004), and Zheng (2012) for the semi-circle, Marchenko-Pastur, and Wachter limiting distributions $F_{\mathbf{c}}$, respectively, statistic $\Delta_p(\theta)$ weakly converges to a Gaussian process indexed by $\theta \in \left(0, \bar{\theta}-\varepsilon\right]$. The explicit form of the mean and the covariance structure can be obtained from the general formulae for the asymptotic mean and covariance of linear spectral statistics given in Theorem 1.1 of Bai and Yao (2005) for SMD, in Theorem 1.1 of Bai and Silverstein (2004) for PCA and REG_0, and in Theorem 4.1 and Example 4.1 of Zheng (2012) for the remaining cases. For PCA, the corresponding calculations have been done in Onatski et al (2013). We omit details of the similar calculations for the remaining cases to save space. The convergence of $\Delta_p(\theta)$ and Theorem 9 imply the following theorem.

Theorem 10 *Suppose that the null hypothesis holds, that is $\theta_0 = 0$. Let $\bar{\theta}$ be the threshold corresponding to F_γ as given in Table 2, and let ε be an arbitrarily small fixed positive number. Further, let $C\left[0, \bar{\theta} - \varepsilon\right]$ be the space of continuous functions on $\left[0, \bar{\theta} - \varepsilon\right]$ equipped with the supremum norm. Then $\ln L(\theta; \Lambda)$ viewed as random elements of $C\left[0, \bar{\theta} - \varepsilon\right]$ converge weakly to $\mathcal{L}(\theta)$ with Gaussian finite dimensional distributions such that*

$$\mathbb{E}\mathcal{L}(\theta) = \frac{1}{4}\ln\left(1 - [\delta(\theta)]^2\right)$$

and

$$\mathbb{C}ov\left(\mathcal{L}(\theta_1), \mathcal{L}(\theta_2)\right) = -\frac{1}{2}\ln\left(1 - \delta(\theta_1)\delta(\theta_2)\right)$$

with

$$\delta(\theta) = \begin{cases} \theta & \text{for SMD} \\ \theta/\sqrt{\gamma_1} & \text{for PCA and REG}_0 \\ \theta\rho/\left(\gamma_1 + \gamma_2 + \theta\gamma_2\right) & \text{for SigD, REG, and CCA} \end{cases}.$$

Here ρ, γ_1, γ_2 are the limits of r, c_1, c_2 as $\mathbf{n}, p \to_\gamma \infty$.

Note that the theorem establishes the weak convergence of the log likelihood ratio viewed as a random element of the space of continuous functions. This is much stronger than simply the convergence of the finite dimensional distributions of the log likelihood process. In particular, the theorem can be used to find the asymptotic distribution of the supremum of the likelihood ratio, and thus, to find the asymptotic critical values of the likelihood ratio test. We do not pursue this line of research here.

Let $\{\mathbb{P}_{p,\theta}\}$ and $\{\mathbb{P}_{p,0}\}$ be the sequences of measures corresponding to the joint distributions of $\lambda_1, ..., \lambda_p$ when $\theta_0 = \theta$ and when $\theta_0 = 0$ respectively. Then Theorem 10 implies, via Le Cam's first lemma, the mutual contiguity of $\{\mathbb{P}_{p,\theta}\}$ and $\{\mathbb{P}_{p,0}\}$ as $\mathbf{n}, p \to_\gamma \infty$. This reveals the statistical meaning of the phase transition thresholds as the upper boundaries of the contiguity regions for spiked models.

The precise form of the autocovariance of $\mathcal{L}(\theta)$ shows that,[2] although the experiment of observing $\lambda_1, ..., \lambda_p$ is asymptotically normal, it does not converge to a Gaussian shift experiment. In particular, the optimality results available for Gaussian shifts cannot be used in our framework. To analyze asymptotic risks of various statistical problems related to the experiment of observing $\lambda_1, ..., \lambda_p$, one

[2] Fyodorov et al (2013) have an interesting discussion of ubiquity of random processes with logarithmic covariance structure in physics and engineering applications. In their paper, such processes appear as limiting objects related to the behavior of the characteristic polynomials of large matrices from Gaussian Unitary Ensemble.

should directly use Theorem 10.

In this paper, we use Theorem 10 to derive the asymptotic power envelopes for tests of the null hypothesis $\theta_0 = 0$ against the alternative $\theta_0 > 0$. Such a power envelope has been derived by Onatski et al (2013) for the case of PCA. By the Neyman-Pearson lemma, the most powerful test of the null against a point alternative $\theta_0 = \theta$ would reject the null when $\ln L(\theta; \Lambda)$ is above a critical value. By Theorem 10 and Le Cam's third lemma (see van der Vaart (1998), chapter 6),

$$\ln L(\theta; \Lambda) \xrightarrow{d} N\left(\frac{1}{4}\ln\left(1 - [\delta(\theta)]^2\right), -\frac{1}{2}\ln\left(1 - [\delta(\theta)]^2\right)\right)$$

under the null, and

$$\ln L(\theta; \Lambda) \xrightarrow{d} N\left(-\frac{1}{4}\ln\left(1 - [\delta(\theta)]^2\right), -\frac{1}{2}\ln\left(1 - [\delta(\theta)]^2\right)\right)$$

under the alternative. This implies the following theorem.

Theorem 11 *Let $\bar{\theta}$ be the threshold corresponding to F_γ as given in Table 2. For any $\theta \in [0, \bar{\theta})$, the value of the asymptotic power envelope for the tests of the null $\theta_0 = 0$ against the alternative $\theta_0 > 0$ which are based on $\lambda_1, ..., \lambda_p$ and have asymptotic size α is given by*

$$PE(\theta) = 1 - \Phi\left[\Phi^{-1}(1-\alpha) - \sqrt{-\frac{1}{2}\ln\left(1 - [\delta(\theta)]^2\right)}\right].$$

Here Φ denotes the standard normal cumulative distribution function. For $\theta \geq \bar{\theta}$ the value of the asymptotic power envelope equals one.

The envelopes are different only for the cases that correspond to different limiting spectral distributions: the semi-circle, the Marchenko-Pastur, and the Wachter distribution. Therefore, we can denote $PE(\theta)$ as $PE^{SC}(\theta)$ for SMD, as $PE^{MP}(\theta)$ for PCA and REG_0, and as $PE^{W}(\theta)$ for the remaining cases. Figure 5 shows the graphs of the envelopes for $\alpha = 0.05$ and $\gamma_1 = \gamma_2 = 0.9$. Such large values of γ_1 and γ_2 correspond to situations where the dimensionality p is not very different from the "sample sizes" n_1 and n_2. Of course, the values of γ_1 and γ_2 are irrelevant for $PE^{SC}(\theta)$, and the value of γ_2 is irrelevant for $PE^{MP}(\theta)$.

Note that the asymptotic power envelope $PE^{MP}(\theta)$ can be obtained from $PE^{W}(\theta)$ by sending γ_2 to zero. Further, $PE^{SC}(\theta)$ can be obtained from $PE^{MP}(\theta)$

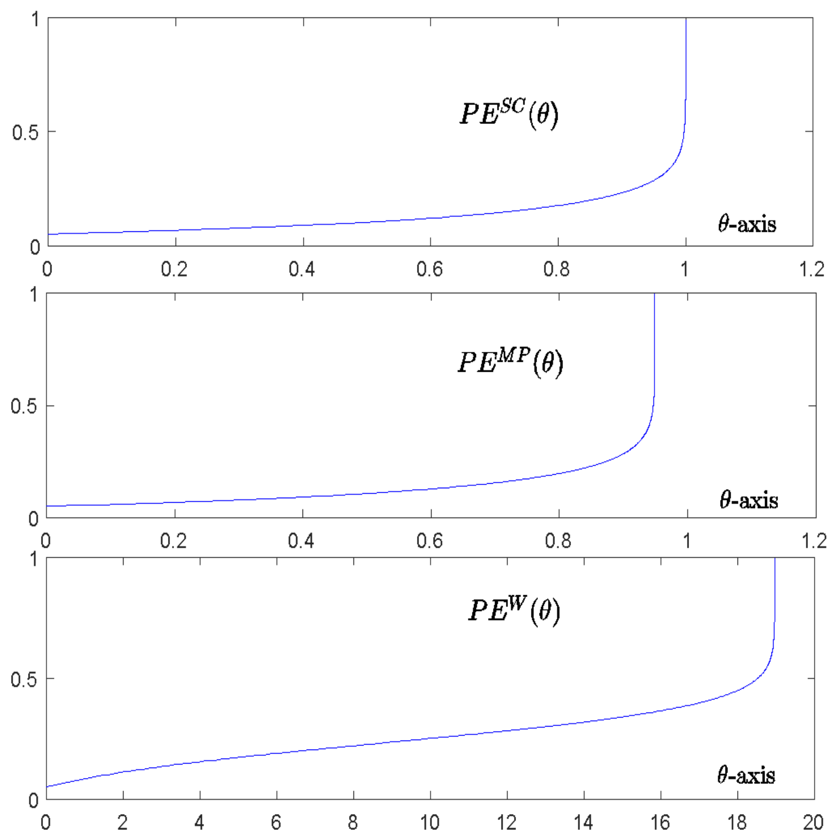

Figure 5: The asymptotic power envelopes $PE^{SC}(\theta)$, $PE^{MP}(\theta)$, and $PE^{W}(\theta)$ for $\alpha = 0.05$, $\gamma_1 = \gamma_2 = 0.9$.

by transformation $\theta \longmapsto \sqrt{\gamma_1}\theta$. Further, note the difference in the horizontal scale of the bottom panel of Figure 5 relative to the two other panels. For $\gamma_1 = \gamma_2 = 0.9$ the phase transition threshold corresponding to Wachter distribution is relatively large. It equals $(\gamma_2 + \rho)/(1 - \gamma_2) \approx 18.9$. Moreover, the value of $PE^W(\theta)$ becomes substantially larger than the nominal size $\alpha = 0.05$ for θ that are situated far below this threshold. This suggests that the information in all the eigenvalues $\lambda_1, ..., \lambda_p$ might be effectively used to detect spikes that are small relative to the phase transition threshold in two sample problems. We leave a confirmation or rejection of this speculation for future research.

7 Conclusion

This paper derives the asymptotics of the likelihood ratio processes corresponding to the null hypothesis of no spikes and the alternative of a single spike in various high-dimensional multivariate models. We cover all the five classes of multivariate statistical problems identified by James (1964). In addition, we consider a symmetric matrix denoising problem that does not fit in James' classification. We find that, as the dimensionality and the number of observations go to infinity proportionally, the log likelihood processes converge to Gaussian limits as long as the value of the spike parameter is below corresponding phase transition thresholds. We derive explicit formulae for the autocovariance and the mean of the limiting processes and use them to obtain asymptotic power envelopes for tests for the presence of a spike.

In this paper, we make the assumption that $n_2 \geq p$ to ensure the invertibility of matrix E in (1) with probability one. However, we also make the assumption $n_1 \geq p$, which can be lifted without a substantial reformulation of the problem. We make the latter assumption mostly to simplify our exposition. The assumption is irrelevant for SMD. The PCA results are obtained in Onatski et al (2013) without using this assumption. For SigD, our derivations (not reported here) show that the equivalent of (7) for $n_1 < p$ involves the hypergeometric function ${}_2F_1$. Therefore, SigD with $n_1 < p$ represents the fifth, rather than the second, group of multivariate statistical problems according to James' (1964) classification. The REG_0 problem is symmetric with respect to n_1 and p after a simple reparametrization. For REG, an equivalent of (7) for $n_1 < p$ can be obtained using equation (74) of James (1964). However, further analysis of REG in this situation needs more substantial changes to our analysis. In CCA case, the sample canonical correlations are only well defined if $n_1 \geq p$. To summarize, when $n_1 < p$, the only interesting untreated cases are SigD and REG. We leave their study for future research.

8 Appendix

8.1 SMD entry of Table 1

The explicit expression for $L^{(SMD)}(\theta;\Lambda)$ given in Table 1 follows from the following lemma.

Lemma 12 *For SMD case, the joint density of the diagonal elements of Λ evaluated at the diagonal elements of $x = \mathrm{diag}\{x_1, ..., x_p\}$ with $x_1 \geq ... \geq x_p$ equals*

$$c_p(x) \exp\left\{-p\theta^2/4\right\} {}_0F_0(\Psi, x), \tag{45}$$

where $c_p(x)$ is a quantity that depends on p and x, but not on θ, and $\Psi = \mathrm{diag}\{\theta p/2, 0, ..., 0\}$. The density under the null hypothesis is obtained from the above expression by setting $\theta = 0$.

Proof: The proof is based on the "symmetrization trick" used by James (1955) to derive the density of non-central Wishart distribution. Let $Y = U'XU$, where U is a random matrix from $\mathcal{O}(p)$ and $X = Z/\sqrt{p} + \eta\theta\eta'$ with Z from GOE, $\theta \geq 0$, and $\|\eta\| = 1$. Note that the eigenvalues of X and Y are the same. The joint density of the functionally independent elements of Y evaluated at y is

$$(2\pi/p)^{-p(p+1)/4} 2^{-p/2} \int_{\mathcal{O}(p)} \mathrm{etr}\left\{-\frac{p}{4}(uyu' - \eta\theta\eta')^2\right\} (\mathrm{d}u),$$

where $(\mathrm{d}u)$ is the normalized uniform measure over $\mathcal{O}(p)$. Taking the square under etr and factorizing, we obtain an equivalent expression

$$(2\pi/p)^{-p(p+1)/4} 2^{-p/2} \exp\left\{-\frac{p}{4}\theta^2\right\} \mathrm{etr}\left\{-\frac{p}{4}y^2\right\} \int_{\mathcal{O}(p)} \mathrm{etr}\left\{\frac{p\theta}{2}uyu'\eta\eta'\right\} (\mathrm{d}u).$$

Now change the variables from y to (H, x), where $y = HxH'$ is the spectral decomposition of y, and integrate H out to obtain (45) with

$$c_p(x) = \frac{p^{p(p+1)/4}\pi^{p(p-1)/4}}{2^{p(p-1)/4+p}\Gamma_p(p/2)} \mathrm{etr}\left(-\frac{p}{4}x^2\right) \prod_{i<j}^{p} (x_i - x_j).$$

Here $\Gamma_p(p/2)$ is the multivariate Gamma function (see Muirhead (1982), pp 61-63). $\square$

8.2 Proof of Lemma 4

It is sufficient to prove the lemma for SigD, REG and CCA. For PCA and REG_0, the lemma follows by taking the limits of SigD and REG cases as $c_2 \to 0$. For SMD, the lemma then follows by taking the limit of PCA case as $c_1 \to 0$, after the transformations $\theta \mapsto \sqrt{c_1}\theta$ and $z \mapsto \sqrt{c_1}z + 1$.

Our proofs of the lemma are very similar for SigD, REG and CCA. Here we show only the proof for SigD. First, note that the minimum of z_0 over $\theta > 0$ equals

$$b_+ \equiv c_1 \left(\frac{r+1}{r+c_2}\right)^2$$

and is achieved at

$$\theta = \bar{\theta}_p \equiv (c_2 + r) / (1 - c_2).$$

Therefore, since $m_{\mathbf{c}}^W(z)$ is well defined for $z > b_+$ and since $\bar{\theta}_p \to \bar{\theta}$ as $\mathbf{n}, p \to_\gamma \infty$, $m_{\mathbf{c}}^W(z_0)$ must be well-defined for any $\theta \in \left(0, \bar{\theta}\right)$ and for sufficiently large $\mathbf{n}, p$. Using an explicit expression for the Stieltjes transform of the limiting spectral distribution of the multivariate F matrix, which is given by Bai and Silverstein (2006) p.71, we obtain

$$\begin{aligned} m_{\mathbf{c}}^W(z) = & -\frac{(c_1 - 1)(c_1 - c_2 z) + (1 - c_2) c_1 z}{2c_1 z (c_1 - c_2 z)} \\ & + \frac{\sqrt{((c_1 - c_2) z + c_1 (1 - c_1))^2 - 4c_1 z (c_1 - c_2 z)}}{2c_1 z (c_1 - c_2 z)}. \end{aligned} \tag{46}$$

Further,

$$2\frac{\mathrm{d}}{\mathrm{d}z} f_{III}^{(SigD)}(z) = -\frac{\theta r^2}{c_1^2 (1+\theta) - c_1 c_2 \theta z}.$$

It now takes a direct algebra, which we perform using Maple's symbolic algebra software, to verify that z_0 solves equation (35).$\square$

8.3 Proof of Lemma 6

For SMD, PCA, and SigD, $|z-\lambda|$ is obviously strictly increasing for any $\lambda \in \mathbb{R}$ and as z moves away from z_0 along $\mathcal{K}_1$. Therefore,

$$2\,\mathrm{Re}\, f_{II}(z) \equiv \int \ln|z-\lambda|\, \mathrm{d}F_{\mathbf{c}}(\lambda)$$

is strictly increasing. On the other hand, by (17), $\mathrm{Re}\, f_{III}(z)$ is non-decreasing. Hence $\mathrm{Re}\, f(z)$ is strictly increasing.

For REG_0 and CCA, $|z-\lambda|$ is strictly increasing for any $\lambda \geq 0$ as z moves away from z_0 along $\mathcal{K}_1$ because the center of the circumference that includes $\mathcal{K}_1$ is a negative real number. Therefore, $\mathrm{Re}\, f_{II}(z)$ is strictly increasing. To show that $\mathrm{Re}\, f_{III}(z)$ is strictly increasing too, it is sufficient to prove that $\mathrm{Re}\,\varphi_j(t_j)$ is strictly increasing for $j=0,2$. A proof of this fact relies on elementary calculus. It is available from the authors upon request.

For REG, z moves away from z_0 along $\mathcal{K}_1$ when τ moves away from τ_0 along $\mathcal{C}_1$. Using (27), (33), and (40), we obtain

$$\mathrm{Re}\, f_{III}(\tau) = \frac{1-c_1}{c_1}\left(-\mathrm{Re}\,\tau + \ln|\tau+1| + \varepsilon\ln|\tau+\varepsilon| + \varepsilon\ln\varepsilon\right).$$

On the other hand, $|\tau+\varepsilon|$ remains constant on $\mathcal{C}_1$ whereas both $-\mathrm{Re}\,\tau$ and $|\tau+1|$ increase as τ moves away from τ_0 along $\mathcal{C}_1$. To see that $|\tau+1|$ indeed increases recall that the center $-\varepsilon$ of the circumference that represents $\mathcal{C}_1$ is to the left of the point -1. Hence, $\mathrm{Re}\, f_{III}(\tau)$ is strictly increasing.

To show that $\mathrm{Re}\, f_{II}(\tau)$ is strictly increasing too it is sufficient to verify that

$$|z-\lambda| \equiv \left|\frac{c_1(1-c_1)}{\theta c_2}\frac{\tau(\tau+1)}{\tau+\varepsilon} - \lambda\right|$$

is strictly increasing for any λ from the support of $F_{\mathbf{c}}$. Since $|\tau+\varepsilon|$ remains constant, it is sufficient to show that

$$\gamma(\tau, x) \equiv |\tau(\tau+1) - x(\tau+\varepsilon)|^2$$

increases as τ moves away from τ_0 along $\mathcal{C}_1$ for any $x = \lambda\theta c_2/[c_1(1-c_1)]$.

Parameterize $\tau \in \mathcal{C}_1$ as $-\varepsilon + \rho e^{\mathrm{i}\alpha}, \alpha \in [0, \pi/2]$. Then elementary calculations yield

$$\gamma(\tau, x) = \rho^4 + (2\varepsilon - 1 + x)^2 \rho^2 - 2\rho^3 (2\varepsilon - 1 + x) \cos\alpha \\ + \varepsilon^2 (\varepsilon - 1)^2 + 2\left(\rho^2 \cos 2\alpha - (2\varepsilon - 1 + x) \rho \cos\alpha\right) \varepsilon (\varepsilon - 1)$$

so that

$$\frac{\mathrm{d}\gamma(\tau, x)}{\mathrm{d}\cos\alpha} = 2\rho\left\{-(2\varepsilon - 1 + x)\left[\rho^2 + \varepsilon(\varepsilon - 1)\right] + 4\rho\varepsilon(\varepsilon - 1)\cos\alpha\right\}. \tag{47}$$

We would like to prove that the derivative $\mathrm{d}\gamma(\tau, x)/\mathrm{d}\cos\alpha$ is negative. As is seen from (47), the derivative is decreasing in x and increasing in $\cos\alpha$. Since $x \geq 0$ and $\cos\alpha \leq 1$, it is sufficient to show that $\mathrm{d}\gamma(\tau, 0)/\mathrm{d}\cos\alpha$ is negative for $\cos\alpha = 1$. We have

$$\left.\frac{\mathrm{d}\gamma(\tau, 0)}{\mathrm{d}\cos\alpha}\right|_{\cos\alpha = 1} = -2\rho(2\varepsilon - 1)\left\{\left(\rho - \frac{2\varepsilon(\varepsilon - 1)}{2\varepsilon - 1}\right)^2 + \right. \\ \left. \varepsilon(\varepsilon - 1) - \left(\frac{2\varepsilon(\varepsilon - 1)}{2\varepsilon - 1}\right)^2\right\}.$$

This is negative because the expression in the second line of the above display is positive. Indeed,

$$\varepsilon(\varepsilon - 1)(2\varepsilon - 1)^2 - 4\varepsilon^2(\varepsilon - 1)^2 = \varepsilon(\varepsilon - 1) > 0.$$

To summarize, both $\operatorname{Re} f_{II}(\tau)$ and $\operatorname{Re} f_{III}(\tau)$ are strictly increasing as τ moves away from τ_0 along $\mathcal{C}_1$. Hence, the image of $\mathcal{C}_1$, $\mathcal{K}_1$, is a contour of steep descent of $-\operatorname{Re} f(z)$ in z-plane.

8.4 Values of $f(z_0)$ and $\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$

Let us first show that $f(z_0) = 0$. Recall that $f(z) = f_I + f_{II}(z) + f_{III}(z)$. The value of f_I is given in Table 3. The value of $f_{III}(z_0)$ is straightforward to compute using the definitions of f_{III} and z_0. The

$$2f_{II}(z_0) = \int_{b_-}^{b_+} \ln(z_0 - \lambda)\,\mathrm{d}F_{\mathbf{c}}(\lambda)$$

takes on three different values: one for SMD, another for PCA and REG_0, and the third one for SigD, REG, and CCA.

Lemma 13 *For SigD, REG, and CCA, for any* $\theta \in \left(0, \bar{\theta}\right)$ *and for sufficiently large* $\mathbf{n}, p$, *we have*

$$2 f_{II}\left(z_0\right) = 2 \ln c_1 - \ln \theta - \frac{1 - c_1}{c_1} \ln\left(1 + \theta\right) - \frac{c_1 + c_2}{c_1 c_2} \ln\left(c_1 + c_2\right) + \frac{r^2}{c_1 c_2} \ln\left[c_1 l\left(\theta\right)\right]. \tag{48}$$

Proof: We follow the usual strategy of reduction to a contour integral. First make the change of variables $\lambda = \alpha - \beta \cos \varphi$. In order to arrange that $\lambda = b_-$ and b_+ at $\varphi = 0$ and π respectively, we set

$$\alpha = \frac{b_+ + b_-}{2} = \frac{c_1\left(r^2 + c_1^2\right)}{\left(c_1 + c_2\right)^2}, \qquad \beta = \frac{b_+ - b_-}{2} = \frac{2 r c_1^2}{\left(c_1 + c_2\right)^2}. \tag{49}$$

We obtain

$$2 f_{II}\left(z_0\right) = \frac{c_1 + c_2}{4 \pi c_1} \int_0^{2\pi} \frac{\beta^2 \sin^2 \varphi \ln\left(z_0 - \alpha + \beta \cos \varphi\right)}{\left(\alpha - \beta \cos \varphi\right)\left(c_1 - c_2 \alpha + c_2 \beta \cos \varphi\right)} \mathrm{d}\varphi$$

after extending the integral from $[0, \pi]$ to $[0, 2\pi]$ using the symmetry of the integrand about $\varphi = \pi$. Now introduce $z = e^{\mathrm{i}\varphi}$. Since $\cos \varphi = \left(z + z^{-1}\right)/2$, we have from (49) the factorizations

$$\begin{aligned}
c_1\left(\alpha - \beta \cos \varphi\right) &= \frac{\beta}{2r}\left(r - c_1 z\right)\left(r - c_1 z^{-1}\right), \\
c_1 - c_2 \alpha + c_2 \beta \cos \varphi &= \frac{\beta}{2r}\left(r + c_2 z\right)\left(r + c_2 z^{-1}\right), \\
z_0 - \alpha + \beta \cos \varphi &= q(z) q\left(z^{-1}\right) \text{ with} \\
q\left(z\right) &= \frac{c_1}{c_1 + c_2}\left(\sqrt{c_1 l\left(\theta\right)/\theta} + r z \sqrt{\theta / \left[c_1 l\left(\theta\right)\right]}\right).
\end{aligned}$$

Our integral becomes

$$2 f_{II}\left(z_0\right) = \frac{-\left(c_1 + c_2\right) r^2}{4 \pi \mathrm{i}} \int_{\mathcal{C}} \frac{\left(z - z^{-1}\right)^2 \ln\left(q(z) q\left(z^{-1}\right)\right)}{\left(r - c_1 z\right)\left(r - c_1 z^{-1}\right)\left(r + c_2 z\right)\left(r + c_2 z^{-1}\right)} \frac{\mathrm{d}z}{z}.$$

The integral has form $I = \oint \ln\left(q(z)q\left(z^{-1}\right)\right) H\left(z\right) z^{-1}\mathrm{d}z$ with $H(z) = H\left(z^{-1}\right)$. Hence, expanding the logarithm yields two identical terms, so that

$$2f_{II}\left(z_0\right) = \frac{-\left(c_1+c_2\right)}{2\pi\mathrm{i}} \int_{\mathcal{C}} \frac{\left(z^2-1\right)^2 \ln q(z)}{\left(r-c_1 z\right)\left(z-c_1/r\right)\left(r+c_2 z\right)\left(z+c_2/r\right)} \frac{\mathrm{d}z}{z}.$$

For $\theta \in \left(0, \bar{\theta}\right)$ and sufficiently large $\mathbf{n}, p$, we have $\theta \in \left(0, \bar{\theta}_p\right)$ with $\bar{\theta}_p = \left(c_2+r\right)/\left(1-c_2\right).$ On the other hand, for $\theta \in \left(0, \bar{\theta}_p\right)$, the function $\ln q\left(z\right)$ is analytic inside the circle $|z| = 1$, and so the whole integrand is analytic inside the circle except for simple poles at $z = 0, c_1/r$ and $-c_2/r$. The residues at these poles are respectively

$$\frac{c_1+c_2}{c_1 c_2} \ln \frac{c_1\sqrt{c_1 l/\theta}}{c_1+c_2}, -\frac{1-c_1}{c_1} \ln \frac{c_1\left(1+\theta\right)}{\sqrt{\theta c_1 l}}, \text{ and } -\frac{1-c_2}{c_2} \ln \frac{c_1}{\sqrt{\theta c_1 l}}$$

and their sum, after collecting terms, yields formula (48).□

Corollary 14 *For PCA and REG*$_0$*, for any* $\theta \in \left(0, \bar{\theta}\right)$ *and for sufficiently large* $\mathbf{n}, p$*, we have*

$$2f_{II}\left(z_0\right) = \ln c_1 - \ln\theta - \frac{1-c_1}{c_1} \ln\left(1+\theta\right) + \theta/c_1. \tag{50}$$

Proof: The corollary is obtained from Lemma 13 by taking the limit as $c_2 \to 0$.□

Corollary 15 *For SMD, for any* $\theta \in \left(0, \bar{\theta}\right)$ *and for sufficiently large* $\mathbf{n}, p$*, we have*

$$2f_{II}\left(z_0\right) = -\ln\theta + \theta^2/2. \tag{51}$$

Proof: We remarked earlier that SMD is a limit of PCA and REG$_0$ as $c_1 \to 0$ after the transformations $\theta \mapsto \sqrt{c_1}\theta$ and $z \mapsto \sqrt{c_1}z + 1$. In particular,

$$z_0^{(SMD)} = \lim_{c_1\to 0, \theta\mapsto\sqrt{c_1}\theta} \left(z_0^{(PCA)} - 1\right)/\sqrt{c_1} \text{ and } F^{SC}\left(\lambda\right) = \lim_{c_1\to 0} F_{\mathbf{c}}^{MP}\left(\sqrt{c_1}\lambda + 1\right).$$

These equations imply that

$$2f_{II}^{(SMD)}\left(z_0^{(SMD)}\right) = \lim_{c_1\to 0, \theta\mapsto\sqrt{c_1}\theta} \left[2f_{II}^{(PCA)}\left(z_0^{(PCA)}\right) - \ln\sqrt{c_1}\right].$$

Using this relationship together with Corollary 14 yields $2f_{II}\left(z_0\right) = -\ln\theta + \theta^2/2$ for SMD.□

Combining equations (48), (50), and (51) with the explicit expressions for f_I and $f_{III}(z_0)$, we obtain the desired result: $f(z_0) = 0$ for all the six cases we consider.

To compute $\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$, note that $-2\mathrm{d}^2 f_{II}(z_0)/\mathrm{d}z^2 = \mathrm{d}m_{\mathbf{c}}(z_0)/\mathrm{d}z$. Therefore $\mathrm{d}^2 f_{II}(z_0)/\mathrm{d}z^2$ can be directly evaluated using explicit expressions for the Stieltjes transforms of the semicircle, Marchenko-Pastur and Wachter distributions. Formula (46) gives such an explicit expression for $m_{\mathbf{c}}^W(z)$. The explicit expressions for $m_{\mathbf{c}}^{MP}(z)$ and $m^{SC}(z)$ are well known. To perform the necessary computations, we use Maple's symbolic algebra software. Further, using the definition of $f_{III}(z)$, we directly evaluate $\mathrm{d}^2 f_{III}(z_0)/\mathrm{d}z^2$. Combining the expressions for the second derivatives of f_{II} and f_{III}, we obtain values of the second derivative of f reported in Table 4.□

8.5 Proof of Theorem 8

First, let us show that

$$L_1(\theta;\Lambda) = \frac{g(z_0)}{\sqrt{-2\mathrm{d}^2 f(z_0)/\mathrm{d}z^2}} + O_{\mathrm{P}}\left(p^{-1}\right), \tag{52}$$

where $O_{\mathrm{P}}(1)$ *is uniform with rspect to* $\theta \in \left(0, \bar{\theta} - \varepsilon\right]$. Changing the variable of integration in (43) from z to $\zeta = \theta z$, we obtain

$$L_1(\theta;\Lambda) = \sqrt{\pi p}\frac{1}{2\pi\mathrm{i}}\int_{\tilde{\mathcal{K}}} e^{-p\phi(\zeta)}\chi(\zeta)\mathrm{d}\zeta, \tag{53}$$

where

$$\phi(\zeta) = f(\zeta/\theta),\ \chi(\zeta) = g(\zeta/\theta)/\theta,$$

and $\tilde{\mathcal{K}}$ is the image of $\mathcal{K}_1 \cup \bar{\mathcal{K}}_1$ under the transformation $z \mapsto \zeta$. The set of possible values of θ is $\Omega \equiv \left(0, \bar{\theta} - \varepsilon\right]$.

Using Table 4 and the definitions of $\mathcal{K}_1$, z_0, $f(z)$, and $g(z)$, it is straightforward to verify that the assumptions A0-A4 of Lemma 7 hold for the integral in (53) for all the six cases that we consider. The validity of A5 follows from Lemma 16 given below and from the definitions of $g(z)$. Let

$$\Delta(\zeta) = p\int \ln(\zeta/\theta - \lambda)\,\mathrm{d}\left(\hat{F}(\lambda) - F_{\mathbf{c}}(\lambda)\right), \tag{54}$$

so that $\Delta(\zeta) = -2\ln g_{II}(\zeta/\theta)$.

Lemma 16 *Suppose that the null hypothesis holds, that is $\theta_0 = 0$. Then there exists a positive constant C_1, such that for a subset Θ of $\mathbb{C}$ that consists of all points whose Euclidean distance from $\tilde{\mathcal{K}}$ is no larger than C_1, we have*

$$\sup_{\zeta\in\Theta} |\Delta(\zeta)| = O_{\mathrm{P}}(1)$$

as $\mathbf{n}, p \to_\gamma \infty$, where $O_{\mathrm{P}}(1)$ is uniform with respect to $\theta \in \Omega \equiv \left(0, \bar{\theta} - \varepsilon\right]$.

Proof: Let us rewrite (54) in the following equivalent form

$$\Delta(\zeta) = p \int \ln(1 - \lambda\theta/\zeta) \, \mathrm{d}\left(\hat{F}(\lambda) - F_{\mathbf{c}}(\lambda)\right).$$

Statistic $\Delta(\zeta)$ is a special form of a linear spectral statistic

$$\Delta(\varphi) = p \int \varphi(\lambda) \, \mathrm{d}\left(\hat{F}(\lambda) - F_{\mathbf{c}}(\lambda)\right)$$

studied by Bai and Yao (2005), Bai and Silverstein (2004), and Zheng (2012) for the cases of the Semi-circle, Marchenko-Pastur, and Wachter limiting distributions, respectively. These papers note that

$$\Delta(\varphi) = -\frac{p}{2\pi \mathrm{i}} \int_{\mathcal{P}} \varphi(\xi) \left(\hat{m}(\xi) - m_{\mathbf{c}}(\xi)\right) \mathrm{d}\xi,$$

where

$$\hat{m}(\xi) = \int \frac{1}{\lambda - \xi} \mathrm{d}\hat{F}(\xi), \; m_{\mathbf{c}}(\xi) = \int \frac{1}{\lambda - \xi} \mathrm{d}F_{\mathbf{c}}(\lambda)$$

are the Stieltjes transforms of $\hat{F}$ and $F_{\mathbf{c}}$, and $\mathcal{P}$ is a positively oriented contour in an open neighborhood of the supports of $\hat{F}$ and $F_{\mathbf{c}}$, where $\varphi(\zeta)$ is analytic, that encloses these supports. Further, the papers show that if the distance from $\mathcal{P}$ to the supports of $\hat{F}$ and $F_{\mathbf{c}}$ stays away from zero with probability approaching one as $\mathbf{n}, p \to_\gamma \infty$, then

$$\Delta(\varphi) = -\frac{p}{2\pi \mathrm{i}} \int_{\mathcal{P}} \varphi(\xi) \, \hat{M}(\xi) \mathrm{d}\xi + O_{\mathrm{P}}(1), \tag{55}$$

where $p\hat{M}(\xi)$ is a truncated version of $p\left[\hat{m}(\xi) - m_{\mathbf{c}}(\xi)\right]$ that weakly converges to a random continuous function on $\mathcal{P}$ with Gaussian finite dimensional distributions. Furthermore, $O_{\mathrm{P}}(1)$ in (55) is uniform in φ that are analytic in the open neighborhood of the supports of $\hat{F}$ and $F_{\mathbf{c}}$ and such that $\sup_{\xi\in\mathcal{P}} |\varphi(\xi)| < K$ for some constant K. Therefore, for any $\delta > 0$, there exists $B > 0$, such that

$$\Pr\left(|\Delta(\varphi)| \leq B \sup_{\xi \in \mathcal{P}} |\varphi(\xi)|\right) > 1 - \delta \tag{56}$$

for all $\mathbf{n}$ and p. Moreover, constant B does not depend on φ. Now, consider a family of functions $\varphi_{\zeta,\theta}(\xi)$

$$\left\{\varphi_{\zeta,\theta}(\xi) = \ln(1 - \xi\theta/\zeta) : \zeta \in \Theta \text{ and } \theta \in \Omega\right\}.$$

By the definitions of Θ and Ω, there exists an open neighborhood $\mathcal{N}$ of the supports of $\hat{F}$ and $F_{\mathbf{c}}$ and a constant B_1, such that, with probability arbitrarily close to one, for sufficiently large $\mathbf{n}$ and p, $\varphi_{\zeta,\theta}(\xi)$ are analytic in $\mathcal{N}$ for all $\zeta \in \Theta$ and $\theta \in \Omega$ and

$$\sup_{\theta \in \Omega} \sup_{\zeta \in \Theta} \sup_{\xi \in \mathcal{N}} |\varphi_{\zeta,\theta}(\xi)| \leq B_1.$$

Since $\Delta(\varphi_{\zeta,\theta}) = \Delta(\zeta)$, we obtain from (56) that for any $\delta > 0$, there exists $B_2 > 0$ such that for sufficiently large $\mathbf{n}$ and p,

$$\Pr\left(\sup_{\theta \in \Omega} \sup_{\zeta \in \Theta} |\Delta(\zeta)| \leq B_2\right) > 1 - \delta.$$

In other words, $\sup_{\zeta \in \Theta} |\Delta(\zeta)| = O_{\mathrm{P}}(1)$ uniformly over $\theta \in \Omega$.□

Applying Lemma 7 to the integral in (53) and using the fact that $f(z_0) = 0$, we obtain (52). It remains to show that $L_2(\theta; \Lambda)$ is asymptotically dominated by $L_1(\theta; \Lambda)$, where

$$L_2(\theta; \Lambda) = L(\theta; \Lambda) - L_1(\theta; \Lambda).$$

For SMD, PCA, and SigD we have

$$\begin{aligned} |L_2(\theta; \Lambda)| &= \left| \frac{\sqrt{\pi p}}{2\pi \mathrm{i}} \int_{\mathcal{K}_2 \cup \bar{\mathcal{K}}_2} e^{-p(f_I + f_{III}(z))} g_I g_{III}(z) \prod_{j=1}^{p} (z - \lambda_j)^{-1/2} \, \mathrm{d}z \right| \\ &\leq \sqrt{\frac{p}{\pi}} e^{-pf_I} g_I (2z_0)^{-p/2} \int_{\mathcal{K}_2} \left| e^{-pf_{III}(z)} g_{III}(z) \mathrm{d}z \right| \\ &\leq \sqrt{\frac{p}{\pi}} e^{-pf_I} g_I (2z_0)^{-p/2} \int_{-\infty}^{z_0} e^{-pf_{III}(x)} g_{III}(x) \mathrm{d}x. \end{aligned}$$

Explicitly evaluating the latter integral and using the exact form of g_I, available from Table 3, we obtain

$$|L_2(\theta;\Lambda)| \leq \frac{2C}{\sqrt{\pi p}} e^{-pf_I} (2z_0)^{-p/2} e^{-pf_{III}(z_0)} (1+o(1)),$$

where $o(1)$ does not depend on θ, $C=1$ for SMD and PCA, and $C=\sqrt{c_1+c_2}/r$ for SigD. Therefore,

$$\begin{aligned} |L_2(\theta;\Lambda)| &\leq \frac{2C}{\sqrt{\pi p}} e^{-pf(z_0)} \exp\left\{-p\left(\ln(2z_0)/2 - f_{II}(z_0)\right)\right\}(1+o(1)) \\ &= \frac{2C}{\sqrt{\pi p}} \exp\left\{-\frac{p}{2}\int \ln\left(\frac{2z_0}{z_0-\lambda}\right) \mathrm{d}F_{\mathbf{c}}(\lambda)\right\}(1+o(1)), \end{aligned}$$

where we used the fact that $f(z_0)=0$. But $\ln(2z_0/(z_0-\lambda))$ is positive and bounded away from zero uniformly over $\theta \in \left(0, \bar{\theta}-\varepsilon\right]$ with probability arbitrarily close to one, for sufficiently large $\mathbf{n}, p$. Hence, there exists a positive constant K such that

$$|L_2(\theta;\Lambda)| \leq \frac{2C}{\sqrt{\pi p}} e^{-pK}(1+o(1))$$

with probability arbitrarily close to one for sufficiently large $\mathbf{n}, p$. Combining this inequality with (52), we establish Theorem 8 for SMD, PCA, and SigD.

For REG_0, we shall need the following lemma.

Lemma 17 *For sufficiently large* $\mathbf{n}$ *and* p*, we have*

$$\left|{}_0F_1(b-s;\Psi_{11}z)\right| < 4\sqrt{\pi m}\left|\exp\left\{-m\varphi_0(t_0)\right\}\right| \tag{57}$$

for any z *and any* $\theta > 0$*.*

Proof: We use the identity (see formula 9.6.3 in Abramowitz and Stegun (1964))

$$I_m(\zeta) = e^{-m\pi \mathrm{i}/2} J_m(\mathrm{i}\zeta) \text{ for } -\pi < \arg\zeta \leq \pi/2,$$

where $J_m(\cdot)$ is the Bessel function. The identity and (22) imply that

$$ {}_0F_1(b-s;\Psi_{11}z) = \Gamma(m+1)\left(m^2\eta_0\right)^{-m/2} e^{-m\pi\mathrm{i}/2} J_m\left(\mathrm{i}2m\eta_0^{1/2}\right). \tag{58}$$

On the other hand, for any ζ and any positive K,

$$\left|J_K(K\zeta)\right| \leq \left\{1+\left|\frac{\sin K\pi}{K\pi}\right|\right\}\left|\left\{\frac{\zeta\exp\left\{\sqrt{1-\zeta^2}\right\}}{1+\sqrt{1-\zeta^2}}\right\}^K\right|, \tag{59}$$

(see Watson (1944), p.270). The latter inequality, equation (58), and the Stirling formula for $\Gamma(m+1)$ imply that (57) holds for sufficiently large m, for any z and $\theta > 0$. The constant 4 on the right hand side of (57) is not the smallest possible one, but it is sufficient for our purposes.□

Using inequality (57), we obtain for REG_0

$$\left|L_2(\theta;\Lambda)\right| \leq 4e^{-pf_I}g_I\sqrt{pm}\int_{\mathcal{K}_2}\left|\exp\{-m\varphi_0(t_0)\}\prod_{j=1}^{p}(z-\lambda_j)^{-1/2}\,\mathrm{d}z\right|. \tag{60}$$

It is straightforward to verify that $\mathrm{Re}\,\varphi_0(t_0)$ is strictly increasing as z is moving along $\mathcal{K}_2$ towards $-\infty$. Therefore, for any $z \in \mathcal{K}_2$,

$$\mathrm{Re}\,\varphi_0(t_0(z)) > \mathrm{Re}\,\varphi_0(t_0(\bar{z})),$$

where $\bar{z} = z_1 + \mathrm{i}(z_0 - z_1)$ is the point of $\mathcal{K}_2$ where $\mathcal{K}_2$ meets $\mathcal{K}_1$. The latter inequality together with (60) yields

$$\left|L_2(\theta;\Lambda)\right| \leq 4e^{-p\,\mathrm{Re}\,f(\bar{z})}g_I\left|g_{II}(\bar{z})\right|\sqrt{pm}\int_{\mathcal{K}_2}\prod_{j=1}^{p}\left|\frac{\bar{z}-\lambda_j}{z-\lambda_j}\right|^{1/2}|\mathrm{d}z|.$$

Since, for some constant τ_1, $\mathrm{Re}\,f(\bar{z}) > f(z_0) + \tau_1 = \tau_1$ and since, by Lemma 16, $4g_{II}(\bar{z}) = O_{\mathrm{P}}(1)$ uniformly over $\theta \in \left(0, \bar{\theta}-\varepsilon\right]$, we obtain

$$\left|L_2(\theta;\Lambda)\right| \leq e^{-p\tau_1}g_I\sqrt{pm}\int_{\mathcal{K}_2}\prod_{j=1}^{p}\left|\frac{\bar{z}-\lambda_j}{z-\lambda_j}\right|^{1/2}|\mathrm{d}z|\,O_{\mathrm{P}}(1). \tag{61}$$

Note that for any $z \in \mathcal{K}_2$ and any $j = 1,...,p$, $\left|(\bar{z}-\lambda_j)/(z-\lambda_j)\right| \leq 1$ and $|z-\lambda_j| > |z|$. Further, since $z_0 < |\bar{z}|$ and with probability arbitrary close to one, for sufficiently large $\mathbf{n}$ and p, $\lambda_1 < z_0$, we have $|\bar{z}-\lambda_j| < |\bar{z}-z_0| < 2|\bar{z}|$.

Thus, for $p \geq 4$, we have

$$\int_{\mathcal{K}_2} \prod_{j=1}^{p} \left| \frac{\bar{z} - \lambda_j}{z - \lambda_j} \right|^{1/2} |\mathrm{d}z| \leq \int_{\mathcal{K}_2} 4 \left| z/\bar{z} \right|^{-2} |\mathrm{d}z| = |\bar{z}| \, O(1)$$

Combining this with (61) and noting that $g_I |\bar{z}| = O(1)$ uniformly over $\theta \in \left(0, \bar{\theta} - \varepsilon\right]$, we obtain

$$|L_2(\theta; \Lambda)| \leq \sqrt{pm} e^{-p\tau_1} O_{\mathrm{P}}(1), \tag{62}$$

where $O_{\mathrm{P}}(1)$ is uniform with respect to $\theta \in \left(0, \bar{\theta} - \varepsilon\right]$. Theorem 8 for REG_0 follows from the latter equality and (52).

For REG and CCA, the Theorem follows from (52) and inequalities

$$|L_2(\theta; \Lambda)| \leq p e^{-p\tau_2} O_{\mathrm{P}}(1), \tag{63}$$

where τ_2 is a positive constant. We obtain (63) by combining the method used to derive (62) with upper bounds on ${}_1F_1$ and ${}_2F_1$, which we establish using the integral representations (25). We omit details to save space.□

References

[1] Abramowitz, M., and Stegun, I. A. (eds.) (1964) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards.

[2] Bai, Z.D. (1999) "Methodologies in Spectral Analysis of Large Dimensional Random Matrices, a Review," *Statistica Sinica* 9, 611-677.

[3] Bai, Z.D. and J.W. Silverstein (2004) "CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices", *Annals of Probability* 32, 553-605.

[4] Bai, Z.D., and J.W. Silverstein (2006) *Spectral Analysis of Large Dimensional Random Matrices*, Science Press, Beijing.

[5] Baik, J. and J.W. Silverstein. (2006) "Eigenvalues of large sample covariance matrices of spiked population models". Journal of Multivariate Analysis 97, 1382–1408.

[6] Bai, Z.D., and Yao, J. (2005) "On the Convergence of the Spectral Empirical Process of Wigner matrices", Bernoulli 11(6), 1059-1092.

[7] Bao, Z., J. Hu, G. Pan, and W. Zhou (2014) "Canonical correlation coefficients of high-dimensional normal vectors: finite rank case," arXiv 1407.7194

[8] Dharmawansa, P. and I. M. Johnstone (2014) "Joint density of eigenvalues in spiked multivariate models," Stat 3, no. 1, 240–249.

[9] Dharmawansa, P., I.M. Johnstone, and A. Onatski (2014a) "Local Asymptotic Normality of the spectrum of high-dimensional spiked F-ratios," arXiv 1411.3875

[10] Fyodorov, Y. V., B.A. Khoruzhenko, and N.J. Simm (2013) "Fractional Brownian motion with Hurst index H=0 and the Gaussian Unitary Ensemble," arXiv 1312.0212v1.

[11] James, A.T. (1955) "The Non-central Wishart Distribution," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 229, pp. 364-366.

[12] James, A. T. (1964) "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics* 35, 475-501.

[13] Luke, Y.L. (1969) *The Special Functions and Their Approximations*, Volume I. Academic Press.

[14] Maïda, M. (2007) "Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles," Electronic Journal of Probability 12, 1131–1150.

[15] Mehta, L.M. (2004) *Random Matrices*. Third edition. Elsevier Academic Press.

[16] Montanari, A., Reichman, D., and Zeitouni, O. (2014) "On the limitation of spectral methods: From the Gaussian hidden clique problem to rank one perturbations of Gaussian tensors," arXiv: 1411.6149v1.

[17] Muirhead, R.J. (1978) "Latent roots and matrix variates: a review of some asymptotic results," *Annals of Statistics* 6, 5-33.

[18] Muirhead, R.J. (1982) Aspects of Multivariate Statistical Theory. John Wiley & Son, Hoboken, New Jersey.

[19] Nadakuditi, R.R. and J. W. Silverstein (2010) "Fundamental Limit of Sample Generalized Eigenvalue Based Detection of Signals in Noise Using Relatively Few Signal-Bearing and Noise-Only Samples," IEEE Journal of Selected Topics in Signal Processing 4 (3), 468-480.

[20] Olver, F.W.J. (1997) *Asymptotics and Special Functions*, A K Peters, Natick, Massachusetts.

[21] Olver, F.W.J., D. W. Lozier, R.F.Boisvert, and C.W. Clark (eds.) (2010) *NIST Handbook of Mathematical Functions*, Cambridge University Press, Cambridge.

[22] Onatski, A. (2007) "Asymptotics of the principal components estimator of large factor models with weak factors and i.i.d. Gaussian noise," manuscript, University of Cambridge. Available at http://www.econ.cam.ac.uk /people /faculty /ao319 /pubs /inference45a.pdf.

[23] Onatski, A., Moreira, M.J., and Hallin, M. (2013) "Asymptotic power of sphericity tests for high-dimensional data", *Annals of Statistics* 41, 1204-1231.

[24] Paris, R.B. (2013a) "Asymptotics of the Gauss Hypergeometric Function with Large Parameters, I," Journal of Classical Analysis 2, 183-203.

[25] Paris, R.B. (2013b) "Asymptotics of the Gauss Hypergeometric Function with Large Parameters, II," Journal of Classical Analysis 3, 1-15.

[26] Passemier, D., McKay, M.R., and Y. Chen (2014) "Asymptotic Linear Spectral Statistics for Spiked Hermitian Random Matrix Models", arXiv:1402.6419v1.

[27] Silverstein, J.W. (1985) "The Limiting Eigenvalue Distribution of a Multivariate F Matrix," *SIAM Journal of Mathematical Analysis* 16 (3), 641-646.

[28] van der Vaart, A.W. (1998) *Asymptotic Statistics*, Cambridge University Press.

[29] Wachter, K. (1980) "The limiting empirical measure of multiple discriminant ratios," *The Annals of Statistics* 8, 937-957.

[30] Watson, G.N. (1944) *A Treatise on the Theory of Bessel Functions*, Cambridge University Press.

[31] Yin, Y.Q, Bai, Z.D., and Krishnaiah, P.R. (1983) "Limiting Behavior of the Eigenvalues of a Multivariate F Matrix," *Journal of Multivariate Analysis* 13 (4), 508-516.

[32] Zheng, S. (2012) "Central limit theorems for linear spectral statistics of large dimensional F-matrices", *Annales de l'Institut Henry Poincaré - Probabilités et Statistiques* 48, 444-476.