UNIVERSITY OF CAMBRIDGE

Faculty of Economics

Cambridge Working Paper Economics

Cambridge Working Paper Economics: 1806

TESTING IN HIGH-DIMENSIONAL SPIKED MODELS

Iain M. Johnstone (Stanford University)

Alexei Onatski (University of Cambridge)

25 January 2018

We consider the five classes of multivariate statistical problems identified by James (1964), which together cover much of classical multivariate analysis, plus a simpler limiting case, symmetric matrix denoising. Each of James' problems involves the eigenvalues of $E^{-2}H$ 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 smaller than a case-specific phase transition threshold, none of the sample eigenvalues separate from the bulk, making the testing problem challenging. Using a unified strategy for the six 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 then derive asymptotic power envelopes for tests for the presence of a spike.

TESTING IN HIGH-DIMENSIONAL SPIKED MODELS

By Iain M. Johnstone* and Alexei Onatski†

Stanford University and University of Cambridge

We consider the five classes of multivariate statistical problems identified by James (1964), which together cover much of classical multivariate analysis, plus a simpler limiting case, symmetric matrix denoising. Each of James' problems involves 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 smaller than a case-specific phase transition threshold, none of the sample eigenvalues separate from the bulk, making the testing problem challenging. Using a unified strategy for the six 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 then derive asymptotic power envelopes for tests for the presence of a spike.

1. Introduction.

High-dimensional multivariate models and methods, such as regression, principal components, and canonical correlation analysis, repay study in frameworks where the dimensionality diverges to infinity together with the sample size. "Spiked" models that deviate from a reference model along a small fixed number of unknown directions have proven to be a fruitful abstraction and research tool in this context. A basic statistical question that arises in the analysis of such 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 remarkable classification, recalled in Table 1, relies on the five most common hypergeometric functions ${}_{\mathsf{p}}F_{\mathsf{q}}$. In this paper, we describe rank-one spiked models that represent each of James' classes in a high dimensional setting. We derive the asymptotic behavior of the corresponding likelihood ratios in a regime where the dimensionality p of the data and the degrees of freedom n_1, n_2 increase proportionally. Specifically, we study the ratios of the joint densities of the relevant data

*Supported in part by NSF DMS 0906812 and 1407813.

†Supported by the J.M. Keynes Fellowship Fund, University of Cambridge.

Primary 62E20, secondary 62H15

Keywords and phrases: likelihood ratio test, hypergeometric function, principal components analysis, canonical correlations, matrix denoising, multiple response regression

TABLE 1
The five cases of James (1964)

		Statistical method	$n_1 H$	$n_2 E$
${}_0F_0$	PCA	Principal components analysis [latent roots of covariance matrix]	$W_p(n_1, \Sigma + \Phi)$	$n_2\Sigma$
${}_1F_0$	SigD	Signal Detection [equality of covariance matrices]	$W_p(n_1, \Sigma + \Phi)$	$W_p(n_2, \Sigma)$
${}_0F_1$	REG$_0$	Multivariate regression, known error covariance [non-central means]	$W_p(n_1, \Sigma, n_1\Phi)$	$n_2\Sigma$
${}_1F_1$	REG	Multivariate regression, unknown error covariance [non-central latent roots]	$W_p(n_1, \Sigma, n_1\Phi)$	$W_p(n_2, \Sigma)$
${}_2F_1$	CCA	Canonical correlation analysis	$W_p(n_1, \Sigma, \Phi(Y))$	$W_p(n_2, \Sigma)$

James' names for the cases, when different from ours, are shown in brackets. Final two columns interpret H and E of (1) *for Gaussian data, so that W_p denotes a p-variate central or noncentral Wishart distribution, see Definitions. Matrix Φ has low rank, equal to one in this paper. For CCA, $\Phi(Y)$ is a random noncentrality matrix, see Supplementary Material (SM) 3.2 for definition. In cases 1 and 3, E is deterministic, Σ is known, and n_2 disappears. Otherwise E is assumed independent of H.*

under the alternative hypothesis, which assumes the presence of a spike, to that under the null of no spike. The relevant data consist, in each case, of the maximal invariant statistic represented by eigenvalues of a large random matrix.

We find that the joint distributions of the eigenvalues under the alternative hypothesis and under the null are mutually contiguous when the values of the spike is below a phase transition threshold. The value of the threshold depends on the problem type. Furthermore, we find that the log likelihood ratio processes parametrized by the value of the spike 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 classical results that assume Gaussian data. 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 the hypothesis H and error sums of squares E are proportional to Wishart matrices, as summarized for the various cases in Table 1. The five cases can be linked via sufficiency and invariance arguments to the statistical problems listed in the table. We briefly discuss these links in the next section.

James' classification suggests common features that call for a systematic approach. Thus 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, one of the arguments has low rank under our spiked model alternatives. Indeed, for tractability, we focus on the rank one setting. We can then 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, Lemma 1, using the recent result of [12]. Then we deform the contour of integration so that the integral becomes amenable to Laplace approximation analysis, extending [27, ch. 4].

Using the Laplace approximation technique, we show that the log likelihood ratios are asymptotically equivalent to simpler random functions of the spike parameters, Theorems 10 and 11. The randomness enters via a linear spectral statistic of a large random matrix of either sample covariance or F-ratio type. Using central limit theorems for the two cases, due to [6] and [38] respectively, we derive the asymptotic Gaussianity and obtain the mean and the autocovariance functions of the log likelihood ratio processes, Theorem 12.

These asymptotics of the log likelihood processes show that the corresponding statistical experiments do not converge to Gaussian shift models. In other words, the experiments that consist of observing the solutions to equation (1) parameterized by the value of the spike under the alternative hypothesis are not of 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 to find simple analytic expressions for the asymptotic power envelopes for the statistical tests of the null hypothesis of no spike in the data, Theorem 13. The form of the envelope depends only on whether both H and E in 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 more rapidly. 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 [30]. Our work here identifies common features in James' classification of multivariate statistical problems and uses them to extend the analysis to the full system. 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, Lemma 3.

The simple 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. Three extensively studied classical ensembles of random matrices are the Gaussian, Laguerre and Jacobi ensembles, e.g. [22]. However, only the Laguerre and Jacobi ensembles appear in high-dimensional analysis of James' five-fold classification. This prompts us to look 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 $\sqrt{n_1}\,(H-\Sigma)$ with $\Sigma = I_p$ as $n_1 \to \infty$, for p fixed. We call the corresponding statistical problem "symmetric matrix denoising"(SMD). 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, again of rank one for this paper. We add this "case zero"to James' classification and derive the asymptotics of the corresponding log likelihood ratio and power envelope.

To summarize, the contributions of this paper are as follows.

- We revisit James' classification, which covers a large part of classical multivariate analysis, now in the setting of high-dimensional data and show that the classification accommodates low rank structures as departures from the classical null hypotheses.
- We show that in such high dimensional settings with rank-one structure, random matrix theory allows tractable approximations to the joint eigenvalue density functions, in place of slowly converging zonal polynomial series.
- We show that the log likelihood ratio processes, when parametrized by spike magnitude, converge to Gaussian process limits in the subcritical interval.

- Hence, we show that informative tests are possible based on *all* the eigenvalues whereas tests based on the largest eigenvalue alone are uninformative.
- As a tool, we develop new uniform approximations to certain hypergeometric functions.
- We identify symmetric matrix denoising as a limiting case of each of James' models. It is the simplest model displaying all the phenomena seen in the paper. We clarify the manner in which the simpler cases are limits of the more complex ones.

The rest of the paper fleshes out this program and its conclusions. The proofs are largely deferred to the extensive Supplementary Material (SM). They reflect substantial effort to identify and exploit common structure in the six cases. Indeed some of this common structure appears remarkable and not yet fully explained.

Definitions and global assumptions. Let Z be an $n \times p$ data matrix with rows drawn i.i.d. from $N_p(0, \Sigma)$, a p-dimensional normal distribution with mean 0 and covariance Σ. Suppose that M is also $n \times p$, but deterministic. If $Y = M + Z$, then $H = Y'Y$ has a p dimensional Wishart distribution $W_p(n, \Sigma, \Psi)$ with n degrees of freedom, covariance matrix Σ and non-centrality matrix $\Psi = \Sigma^{-1}M'M$. The central Wishart distribution, corresponding to $M = 0$, is denoted $W_p(n, \Sigma)$.

Throughout the paper, we shall assume that

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

where p is the dimensionality of matrices H and E, and n_1, n_2 are the degrees of freedom of the corresponding Wishart distributions, as summarized in Table 1. The assumption $p \leq n_2$ ensures almost sure invertibility of matrix E in (1), whereas the assumption $p \leq n_1$ while not essential, is made for brevity, as it reduces the number of various situations which need to be considered.

2. Links to statistical problems. We briefly review examples of statistical problems, old and new, that lead to each of James' five cases, plus symmetric matrix denoising, and explain our choice of labels for those cases.

PCA. In the first case n_1 i.i.d. $N_p(0, \Omega)$ observations are used to test the null hypothesis that the population covariance Ω equals a given matrix Σ. The alternative of interest is

$$\Omega = \Sigma + \Phi \quad \text{with} \quad \Phi = \theta\psi\psi',$$

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

Without loss of generality (wlog), 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.

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 $n_1 H$ equal to the sample covariance matrix and $E = \Sigma$. We restrict attention to the invariant tests. Therefore, the relevant data are summarized by $\lambda_1, ..., \lambda_p$. For convenience, details of the invariance and sufficiency arguments for all cases are in SM 2.1.

SigD. Consider 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. The alternative hypothesis is the same as for case PCA. 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 'noise' sample covariance matrix. We again can assume that the population covariance of the 'noise' $\Sigma = I_p$, although this time it is unknown to the statistician (SM explains why such an assumption is wlog). Here, 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 second sample size $n_2 \to \infty$, while n_1 and p are held constant, equation (3) reduces to equation (1), E converges to Σ, and SigD reduces to PCA.

REG$_0$*, REG.* Now consider linear regression with multivariate response

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

when the goal is to test linear restrictions on the matrix of coefficients β. In case REG_0 the covariance matrix Σ of the i.i.d. Gaussian rows of the error matrix ε is assumed known. REG corresponds to unknown Σ.

As explained in [24, 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 linear restrictions (per each of the p columns of matrix β). 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 $\theta > 0$, $\|\Sigma^{-1/2}\psi\| = 1$, 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. SM contains a discussion of the relationship between alternative (4) and a corresponding constraint on the coefficients of the untransformed regression model.

In the important example of comparison of q group means, i.e. one-way MANOVA, the null hypothesis imposes equality of all means, while a rank one alternative would posit that the q mean vectors lie along a line, for example $\mu_k = \mu_1 + s_k\psi$ for scalar $s_k, k = 2, \ldots, q$ and $\psi \in \mathbb{R}^p$. This will be a plausible reduction of a global alternative hypothesis in some applications.

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

$$H = Y_1'Y_1/n_1 \text{ and } E = \Sigma.$$

These solutions represent a multivariate analog of the difference between the sum of squared residuals in the restricted and unrestricted regressions. Under the null hypothesis, n_1H is distributed as $W_p(n_1, \Sigma)$ whereas under the alternative, it is distributed as $W_p(n_1, \Sigma, n_1\Phi)$, where $\Phi = \theta\Sigma^{-1}\psi\psi'$. Without loss of generality, we may assume that $\Sigma = I_p$.

The canonical form of REG_0 is essentially equivalent to the recently studied setting of *matrix denoising*

$$Y_1 = M + Z.$$

References which point to a variety of applications include [11, 35, 25, 15]. Often M is assumed to have low rank, and the matrix valued noise Z to have i.i.d. Gaussian entries. Here we test $M = 0$ versus a rank one alternative.

For REG, similar arguments lead to tests based on the p solutions $\lambda_1, ..., \lambda_p$ of (3) with

$$H = Y_1'Y_1/n_1 \text{ and } E = Y_3'Y_3/n_2,$$

where the error d.f. $n_2 = T - q$. These solutions represent a multivariate analog of the F ratio: the difference between the sum of squared residuals in the restricted and unrestricted regressions to the sum of squared residuals in the restricted regression. Again, we may assume wlog that Σ, although unknown to the statistician, equals I_p. Note that, as $n_2 \to \infty$ while n_1 and p are held constant, REG reduces to REG$_0$.

CCA. Consider testing for 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 $(x_t', y_t')'$ into

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

respectively. Under $H_0 : \Sigma_{xy} = 0$, the alternative of interest is

$$\Sigma_{xy} = \sqrt{\frac{n_1\theta}{n_1\theta + n_1 + n_2}}\psi\varphi', \tag{5}$$

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

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

The peculiar parameterizations of the alternative $\theta \neq 0$ in (4) and (5) are chosen to allow unified treatments of PCA and REG$_0$ and of SigD, REG and CCA in our main results, Theorems 11 and 12 below.

The test can be based on the squared sample canonical correlations, which are solutions to (1) with

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

Remarkably, the squared sample canonical correlations scaled by n_2/n_1, which we denote as $\lambda_1, ..., \lambda_p$, 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 (see SM 3.2).

SMD. We observe a $p \times p$ matrix $X = \Phi + Z/\sqrt{p}$, where Z is a noise matrix from the Gaussian Orthogonal Ensemble (GOE), i.e. it is symmetric and

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

We seek to make inference about a symmetric rank-one "signal" matrix $\Phi = \theta\psi\psi'$. The null and the alternative hypotheses are given by (2). The nuisance vector $\psi \in \mathbb{R}^p$ 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 each of the above cases. For example, consider PCA with p held fixed and $n_1 \to \infty$. Take $\Sigma = I_p$ for convenience and set $\Omega = I_p + \sqrt{p/n_1}\Phi$ with $\Phi = \theta\psi\psi'$, so that the original value of the spike is rescaled to be a local perturbation. Now write H in the form $\Omega^{1/2}\check{H}\Omega^{1/2}$ where $\check{H} \sim W_p(n_1, I_p)$. A standard matrix central limit theorem for p fixed, e.g. [16, Th. 2.5.1], says that

$$\check{H} = I_p + Z/\sqrt{n_1} + o_{\mathrm{P}}(n_1^{-1/2}),$$

where Z belongs to GOE. Writing $\Omega^{1/2} = I_p + \frac{1}{2}\sqrt{p/n_1}\Phi + o(n_1^{-1/2})$, and introducing $\mu = \sqrt{n_1/p}\,(\lambda - 1)$, we can rewrite

$$\det(H - \lambda I_p) = (p/n_1)^{p/2} \det[\Phi + Z/\sqrt{p} - \mu I_p + o_{\mathrm{P}}(1)],$$

so that PCA degenerates to SMD. Compare also [4].

Indeed, each of the cases eventually degenerate to SMD via sequential asymptotic links (SM 2.2 has details). For convenience, we summarize links between the different cases and the definitions of the corresponding matrices H and E in Figure 1. We note that the SMD model has been studied recently, e.g. [9, 20] and references therein, though not with our techniques.

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.

3. The likelihood ratios. Our goal is to study the asymptotic behavior of likelihood ratios based on the observed eigenvalues

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

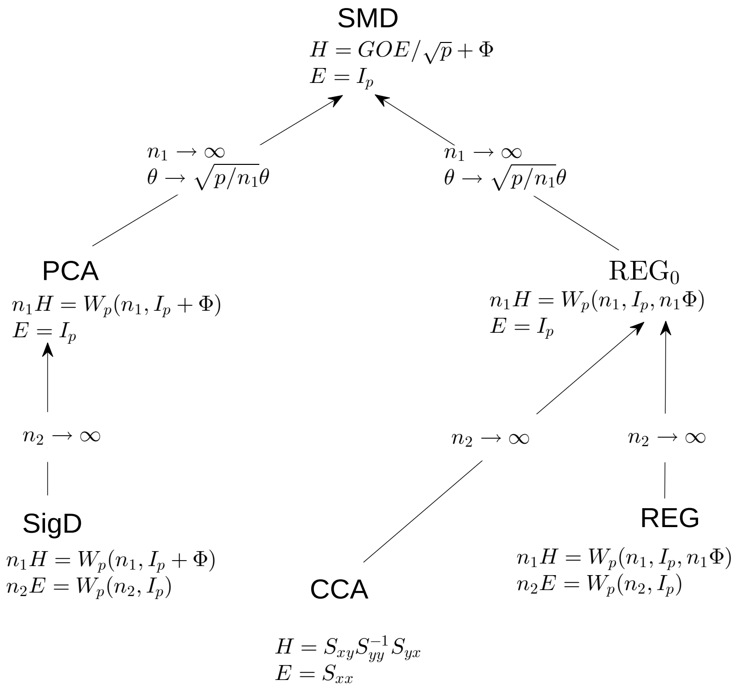

FIG 1. *Matrices H and E, and links between the different cases. Without loss of generality, matrix E or, in SigD, REG, and CCA cases, its population counterpart Σ is assumed to be equal to I_p. Matrix Φ has the form $\theta\psi\psi'$ with $\theta \geq 0$ and $\|\psi\| = 1$.*

Let $p(\Lambda;\theta)$ be the joint density of the eigenvalues under the alternative and $p(\Lambda;0)$ the corresponding density under the null. James' formulas for these joint densities lead to our starting point, which is a unified form for the likelihood ratio

$$L(\theta;\Lambda)=\frac{p(\Lambda;\theta)}{p(\Lambda;0)}=\alpha\left(\theta\right)\,{}_{\mathsf{p}}F_{\mathsf{q}}(a,b;\Psi,\Lambda)\;, \tag{6}$$

where $\Psi=\Psi(\theta)$ is a p-dimensional matrix $\operatorname{diag}\left\{\Psi_{11},0,...,0\right\}$, and the values of Ψ_{11}, $\alpha\left(\theta\right)$, p, q, a, and b are as given in Table 2.

For SMD, we prove that $L\left(\theta;\Lambda\right)$ is as in (6) in SM 3.1. For PCA, the explicit form of the likelihood ratio is derived in [30]. For SigD, REG_0, and REG, the expressions (6) follow, respectively, from equations (65), (68), and (73) of [18]. For CCA, the expression is a corollary of [24, Th. 11.3.2]. Further details appear in SM 3.2.

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

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

where $a=\left(a_1,...,a_{\mathsf{p}}\right)$ and $b=\left(b_1,....,b_{\mathsf{q}}\right)$ are parameters, κ are partitions of the integer k, $\left(a_j\right)_{\kappa}$ and $\left(b_i\right)_{\kappa}$ are the generalized Pochhammer symbols, and C_{κ} are the zonal polynomials, e.g. [24, Def. 7.3.2.]. Note that some links between the cases illustrated in Figure 1 can also be established via asymptotic relations between the hypergeometric functions. For example, the confluence relations

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

TABLE 2

Parameters of the explicit expression (6) for the likelihood ratios. Here $n\equiv n_1+n_2$.

Case	${}_{\mathsf{p}}F_{\mathsf{q}}$	$\alpha\left(\theta\right)$	a	b	Ψ_{11}
SMD	${}_0F_0$	$\exp\left(-p\theta^2/4\right)$	_	_	$\theta p/2$
PCA	${}_0F_0$	$\left(1+\theta\right)^{-n_1/2}$	_	_	$\theta n_1/(2\left(1+\theta\right))$
SigD	${}_1F_0$	$\left(1+\theta\right)^{-n_1/2}$	$n/2$	_	$\theta n_1/\left(n_2\left(1+\theta\right)\right)$
REG_0	${}_0F_1$	$\exp\left(-n_1\theta/2\right)$	_	$n_1/2$	$\theta n_1^2/4$
REG	${}_1F_1$	$\exp\left(-n_1\theta/2\right)$	$n/2$	$n_1/2$	$\theta n_1^2/\left(2n_2\right)$
CCA	${}_2F_1$	$\left(1+n_1\theta/n\right)^{-n/2}$	$(n/2,n/2)$	$n_1/2$	$\theta n_1^2/\left(n_2^2+n_2n_1\left(1+\theta\right)\right)$

e.g. [28, eq. 35.8.9], imply the links SigD $\mapsto$ PCA and REG $\mapsto$ REG$_0$ as $n_2 \to \infty$ for p and n_1 held constant.

In the next section, we shall study the asymptotic behavior of the likelihood ratios (6) as n_1, n_2, and p go to infinity so that

$$c_1 \equiv p/n_1 \to \gamma_1 \in (0,1) \text{ and } c_2 \equiv p/n_2 \to \gamma_2 \in (0,1]\,. \tag{7}$$

We denote this asymptotic regime by $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, where $\mathbf{n} = \{n_1, n_2\}$ and $\boldsymbol{\gamma} = \{\gamma_1, \gamma_2\}$. 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 **n**. We briefly discuss possible extensions of our analysis to the situations with $\gamma_1 \geq 1$ in Section 7.

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. First, some background on the eigenvalues. Under the null, $\lambda_1, ..., \lambda_p$ are the eigenvalues of $GOE/\sqrt{p}$ in the SMD case; of $W_p(n_1, I_p)/n_1$ for PCA and REG$_0$; and of a p-dimensional multivariate beta matrix, e.g. [23, p. 110], with parameters $n_1/2$ and $n_2/2$ and here scaled by a factor of n_2/n_1, in the SigD, REG, and CCA cases. The empirical distribution of $\lambda_1, ..., \lambda_p$

$$\hat{F} = \frac{1}{p}\sum_{j=1}^{p} I\{\lambda_j \leq \lambda\}$$

is well known, [3], to converge weakly almost surely (a.s.) in each case:

$$\hat{F} \Rightarrow F_{\boldsymbol{\gamma}} = \begin{cases} F^{\mathrm{SC}} & \text{for SMD} \\ F^{\mathrm{MP}} & \text{for PCA, REG}_0 \\ F^{\mathrm{W}} & \text{for SigD, REG, CCA,} \end{cases}$$

the semi-circle, Marchenko-Pastur and (scaled) Wachter distributions respectively. Table 3 recalls the explicit forms of these limiting distributions. The cumulative distribution functions $F_{\boldsymbol{\gamma}}^{\mathrm{lim}}(\lambda)$ are linked in the sense that

$$\begin{aligned} &F_{\boldsymbol{\gamma}}^{\mathrm{W}}(\lambda) &&\to F_{\gamma_1}^{\mathrm{MP}}(\lambda) &&\text{as } \gamma_2 \to 0, \\ &F_{\gamma_1}^{\mathrm{MP}}(\sqrt{\gamma_1}\lambda + 1) &&\to F^{\mathrm{SC}}(\lambda) &&\text{as } \gamma_1 \to 0. \end{aligned}$$

If φ is a 'well-behaved' function, the centered *linear spectral statistic*

$$\sum_{j=1}^{p} \varphi(\lambda_j) - p\int \varphi(\lambda)\, \mathrm{d}F_{\mathbf{c}}^{\mathrm{lim}}(\lambda)\,, \tag{8}$$

TABLE 3
Semi-circle, Marchenko-Pastur and scaled Wachter distributions

Case	$F_{\boldsymbol{\gamma}}^{\lim}$	Density, $\lambda \in [\beta_-, \beta_+]$	$\beta_\pm$	Threshold $\bar{\theta}$
SMD	SC	$\dfrac{R(\lambda)}{2\pi}$	± 2	1
PCA REG_0	MP	$\dfrac{R(\lambda)}{2\pi\gamma_1\lambda}$	$(1 \pm \sqrt{\gamma_1})^2$	$\sqrt{\gamma_1}$
SigD REG CCA	W	$\dfrac{(\gamma_1+\gamma_2)R(\lambda)}{2\pi\gamma_1\lambda(\gamma_1-\gamma_2\lambda)}$	$\gamma_1 \left(\dfrac{\rho \pm 1}{\rho \pm \gamma_2}\right)^2$	$\dfrac{\rho+\gamma_2}{1-\gamma_2}$

$$R(\lambda) = \sqrt{(\beta_+ - \lambda)(\lambda - \beta_-)} \qquad \rho = \sqrt{\gamma_1 + \gamma_2 - \gamma_1\gamma_2}$$

converges in distribution to a Gaussian random variable in each of the semi-circle [7], Marchenko-Pastur [6] and Wachter [38] cases. Note that the centering constant is defined in terms of $F_{\mathbf{c}}$, where $\mathbf{c} = \{c_1, c_2\}$. That is, the "correct centering" can be computed using the densities from Table 3, 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 recall the behavior of the largest eigenvalue λ_1 under the alternative hypothesis. As long as $\theta \leq \bar{\theta}$, the phase transition threshold reported in Table 3, the top eigenvalue $\lambda_1 \to \beta_+$, the upper boundary of support of $F_{\boldsymbol{\gamma}}$, almost surely. When $\theta > \bar{\theta}$, λ_1 separates from 'the bulk' of the other eigenvalues and a.s. converges to a point strictly above β_+. For details, we refer to [21, 8, 26, 29, 12, 10] for the respective cases SMD, PCA, SigD, REG_0, REG, and CCA.

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 $\bar{\theta}$. In such *super-critical* cases, 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 $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, as shown in [21] and [30] for SMD and PCA respectively. In contrast, as we show below, the sequences of measures corresponding to the distributions of Λ under the null and under *sub-critical* alternatives $\theta < \bar{\theta}$ are mutually contiguous, and the likelihood ratio converges to a Gaussian process. In the super-critical setting, an analysis of the likelihood ratios under local alternatives appears in [13].

4. Contour integral representation. The asymptotic behavior of the likelihood ratios (6) depends on that of ${}_{\mathsf{p}}F_{\mathsf{q}}\left(a,b;\Psi,\Lambda\right)$. When the dimension of the matrix arguments remains fixed, there is a large and well established literature on the asymptotics of ${}_{\mathsf{p}}F_{\mathsf{q}}\left(a,b;\Psi,\Lambda\right)$ for large parameters and norm of the matrix arguments, see [23] for a review. In contrast, relatively little is known about when the dimensionality of the matrix arguments Ψ,Λ diverge to infinity. It is this regime we study in this paper, noting that in single-spiked models, the matrix argument Ψ has rank one. This allows us to represent ${}_{\mathsf{p}}F_{\mathsf{q}}\left(a,b;\Psi,\Lambda\right)$ 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.

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*

(9)
$$L\left(\theta;\Lambda\right)=\frac{\Gamma\left(s+1\right)\alpha\left(\theta\right)q_s}{\Psi_{11}^s 2\pi\mathrm{i}}\int_{\mathcal{K}}{}_{\mathsf{p}}F_{\mathsf{q}}\left(a-s,b-s;\Psi_{11}z\right)\prod_{j=1}^{p}\left(z-\lambda_j\right)^{-1/2}\mathrm{d}z,$$

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

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

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 [12, Prop. 1] and from equation (6). Our next step is to apply the Laplace approximation to integrals (9). To this end, we shall transform the right hand side of (9) so that it has a "Laplace form"

$$L\left(\theta;\Lambda\right)=\sqrt{\pi p}\frac{1}{2\pi\mathrm{i}}\int_{\mathcal{K}}\exp\left\{-(p/2)f(z;\theta)\right\}g(z;\theta)\mathrm{d}z. \tag{10}$$

The dependence on θ will usually not be shown explicitly. 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. In order to apply the Laplace approximation, we shall deform the contour of

TABLE 4

Values of f_c and $\breve{g}_c = g_c/(1+o(1))$ for the different cases. The terms $o(1)$ do not depend on θ and converge to zero as $\mathbf{n}, p \to_\gamma \infty$. In the table, $l(\theta) = 1 + (1+\theta)c_2/c_1$ and $r^2 = c_1 + c_2 - c_1c_2$.

Case	f_c	$\breve{g}_c = g_c/(1+o(1))$
SMD	$1 + \theta^2/2 + \log\theta$	θ
PCA	$1 + \frac{1-c_1}{c_1}\log(1+\theta) + \log\frac{\theta}{c_1}$	$\theta(1+\theta)^{-1}c_1^{-1}$
SigD	$f_c^{\text{PCA}} + f_{10}$	$\breve{g}_c^{\text{PCA}}\breve{g}_{10}$
REG$_0$	$1 + \frac{\theta+c_1}{c_1} + \log\frac{\theta}{c_1} + \frac{1-c_1}{c_1}\log(1-c_1)$	$\theta c_1^{-1}(1-c_1)^{-1/2}$
REG	$f_c^{\text{REG}_0} + f_{10}$	$\breve{g}_c^{\text{REG}_0}\breve{g}_{10}$
CCA	$f_c^{\text{REG}} + f_{21}$	$\breve{g}_c^{\text{REG}}\breve{g}_{10}/l(\theta)$
	$f_{10} = -1 - \frac{r^2}{c_1c_2}\log\frac{r^2}{c_1+c_2} + \log\frac{c_1+c_2}{c_1}$	$\breve{g}_{10} = c_1^{-1}r(c_1+c_2)^{1/2}$
	$f_{21} = -1 - \frac{\theta}{c_1} - \frac{r^2}{c_1c_2}\log\frac{r^2}{c_1 l(\theta)}$	

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

$$f(z) = f_c + f_e(z) + f_h(z) \text{ and}$$
$$g(z) = g_c \times g_e(z) \times g_h(z), \tag{11}$$

where f_c and g_c do not depend on z. The subscripts (c,e,h) are mnemonic for 'coefficient', 'eigenvalues' and 'hypergeometric'.

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

$$\frac{\Gamma(s+1)\,\alpha(\theta)\,q_s}{\sqrt{\pi p}\Psi_{11}^s} = \exp\{-(p/2)f_c\}\, g_c, \tag{12}$$

where g_c remains bounded as $\mathbf{n}, p \to_\gamma \infty$. The values of f_c and g_c are given in Table 4. Details of the derivation are given in SM 4.1.

Second, we consider the decomposition

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

where

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

and

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

For $f_{\mathrm{e}}(z)$ and $g_{\mathrm{e}}(z)$ to be well-defined we need z not to belong to the support of $F_{\mathbf{c}}$, which we assume. In addition, $z \notin \mathrm{supp}(\hat{F})$ since by definition contour $\mathcal{K}$ encircles it. Note that $g_{\mathrm{e}}(z)$ is the exponent of a linear spectral statistic, which converges to a Gaussian random variable as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$ under the null hypothesis.

Third and finally, we describe a decomposition

$${}_{\mathsf{p}}F_{\mathsf{q}}\left(a-s, b-s; \Psi_{11}z\right)=\exp\left\{-(p/2)f_{\mathrm{h}}(z)\right\}g_{\mathrm{h}}(z). \tag{16}$$

For the $\mathsf{q}=0$ cases, the corresponding ${}_{\mathsf{p}}F_{\mathsf{q}}$ can be expressed in terms of elementary functions. Indeed, ${}_0F_0(z)=e^z$ and ${}_1F_0(a;z)=(1-z)^{-a}$. We set

$$f_{\mathrm{h}}(z)=\begin{cases} -z\theta & \text{for SMD} \\ -z\theta/\left(c_1(1+\theta)\right) & \text{for PCA} \\ \ln\left[1-c_2 z\theta/\left\{c_1(1+\theta)\right\}\right] r^2/\left(c_1c_2\right) & \text{for SigD,} \end{cases} \tag{17}$$

and

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

Unfortunately, for the $\mathsf{q}=1$ cases, the corresponding ${}_{\mathsf{p}}F_{\mathsf{q}}$ do not admit exact representations in terms of elementary functions. Therefore, we shall consider their asymptotic approximations instead. Let

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

Further, let

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

where

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

With this notation, we have

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

The function $F_0(z)$ can be expressed in terms of the modified Bessel function of the first kind $I_m(\cdot)$, see [2, eq. 9.6.47], as

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

This representation allows us to use a known uniform asymptotic approximation of the Bessel function [2, eq. 9.7.7] to obtain Lemma 2, proven in SM 4.2. To state it let

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

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

$$|\arg \eta_0| \le \pi - \delta, \text{ and } \eta_0 \ne 0.$$

LEMMA 2. *As $m \to \infty$, we have*

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

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

To foreshadow our results for $F_1(z)$ and $F_2(z)$, we note that the right hand side of (24) can be formally linked, via (22), to the saddle-point approximation of the integral representation, see [37, p. 181],

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

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

Turning now to functions $F_1(z)$ and $F_2(z)$, to obtain uniform asymptotic approximations, we use the contour integral representations, see [28, eqs. 13.4.9 and 15.6.2],

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

where

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

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

and

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

For $j=2$, the contour does not encircle $1/\eta_2$, and the representation is valid for η_2 such that $|\arg(1-\eta_2)|<\pi$. We derive a saddle-point approximation to the integral in (25) to be summarized in Lemma 3 below. The relevant saddle points are

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

We shall need the following additional notation. Let

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

where the branches of $\arg(\cdot)$ are chosen so that $|\omega_j + 2\omega_{0j}| \le \pi/2$.

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 $(\kappa,\eta)\in\Omega_{j\delta}$ for any $\delta>0$, where $\Omega_{j\delta}$ are as defined in Table 5.

Point-wise asymptotic approximation (31) was established in [34] for $j=1$, and in [32, 33] for $j=2$. However, those papers do not study the uniformity of the approximation error, which is important for our analysis. Lemma 3 is proved at length in SM 4.3. It is fair to say that the corresponding derivations constitute the technically most challenging part of our analysis. This further highlights the technical difficulties that occur when going from SMD, PCA, and SigD cases to REG_0, REG, and CCA.

Using Lemmas 2 and 3, and Stirling's approximation

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

we set the components of the "Laplace form" (16) of ${}_{\mathsf{p}}F_{\mathsf{q}}$ for the $\mathsf{q}=1$ cases as follows

$$f_{\mathrm{h}}(z) = \begin{cases} \frac{1-c_1}{c_1}\varphi_0(t_0) & \mathrm{REG}_0 \\ \frac{1-c_1}{c_1}\left(\varphi_j(t_j) + \kappa\ln\kappa - (\kappa-1)\ln(\kappa-1)\right) & \mathrm{REG,\ CCA} \end{cases} \tag{33}$$

and

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

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 the $o(1)$ 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 $(\kappa, \eta_j) \in \Omega_{j\delta}$ 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).

Confluences of functions f. As $c_2 \to 0$ with c_1 held fixed, we have

$$\begin{gathered} f^{\mathrm{SigD}}(z) \to f^{\mathrm{PCA}}(z), \\ f^{\mathrm{REG}}(z), f^{\mathrm{CCA}}(z) \to f^{\mathrm{REG}_0}(z). \end{gathered} \tag{35}$$

Also, as $c_1 \to 0$,

$$f^{\mathrm{PCA}}(z), f^{\mathrm{REG}_0}(z) \to f^{\mathrm{SMD}}(z), \tag{36}$$

after making the substitutions $\theta \to \sqrt{c_1}\theta$ and $z \to \sqrt{c_1}z + 1$ on the left hand side. Some details appear in SM 4.4.

TABLE 5
Definition of $\Omega_{j\delta}$ from Lemma 3.

$\Omega_{j\delta} = \Omega_\delta \cap \hat{\Omega}_{j\delta}$ with the following Ω_δ and $\hat{\Omega}_{j\delta}$

Set	Definition: pairs $(x,z) \in \mathbb{R}\times\mathbb{C}$ s.t.
Ω_δ	$\delta \le x-1 \le 1/\delta$, $\lvert z\rvert \le 1/\delta$, and $\inf_{y\in\mathbb{R}\setminus[0,\infty)} \lvert z-y\rvert \ge \delta$
$\hat{\Omega}_{1\delta}$	$\mathrm{Re}\, z \ge -2x+1$
$\hat{\Omega}_{2\delta}$	$\inf_{y\in\mathbb{R}\setminus(-\infty,1]} \lvert z-y\rvert \ge \delta$ and x is unconstrained.

4.2. *Saddlepoints and Contours of steep descent.* We shall now show how to deform contours $\mathcal{K}$ in (10) into the contours of steep descent. First, we find saddle points of functions $f(z)$ for each of the six cases. Note that

$$-\mathrm{d}f_{\mathrm{e}}(z)/\mathrm{d}z = \int (\lambda - z)^{-1} \mathrm{d}F_{\mathbf{c}}(\lambda) = m_{\mathbf{c}}(z),$$

the Stieltjes transform of $F_{\mathbf{c}}$. 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 $p \leq n_1$, $F_{\mathbf{c}}$ does not have any non-trivial mass at 0.

To find saddle points z_0 of $f(z)$ we therefore solve the equation

$$m_{\mathbf{c}}(z) = \mathrm{d}f_{\mathrm{h}}(z)/\mathrm{d}z. \tag{37}$$

A proof of the following lemma appears in SM 4.5.

Lemma 4. *The saddle points $z_0(\theta, \mathbf{c})$ of $f(z)$ satisfy*

$$z_0(\theta, \mathbf{c}) = \begin{cases} \theta + 1/\theta & \text{for SMD} \\ (1+\theta)(\theta + c_1)/\theta & \text{for PCA and REG}_0 \\ (1+\theta)(\theta + c_1)/[\theta l(\theta)] & \text{for SigD, REG, and CCA.} \end{cases} \tag{38}$$

For $\theta \in (0, \bar{\theta}_{\mathbf{c}})$, $z_0 > b_+$, where $\bar{\theta}_{\mathbf{c}}$ is the threshold corresponding to $F_{\mathbf{c}}$, which is an analogue of the threshold $\bar{\theta}_{\boldsymbol{\gamma}} \equiv \bar{\theta}$ corresponding to $F_{\boldsymbol{\gamma}}$ given in Table 3.

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 $z_0 = \theta + 1/\theta$.

Remark 5. For all the six cases that we study, $f(z_0)$ equals zero. SM 4.6 has a verification of this important fact.

Remark 6. As $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, $z_0(\theta, \mathbf{c}) \to z_0(\theta, \boldsymbol{\gamma}) > \beta_+$, where the latter inequality holds for any $\theta \in (0, \bar{\theta})$. Since $\lambda_1 \overset{a.s.}{\to} \beta_+$ the inequality $z_0(\theta, \mathbf{c}) > \lambda_1$ must hold with probability approaching one as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$.

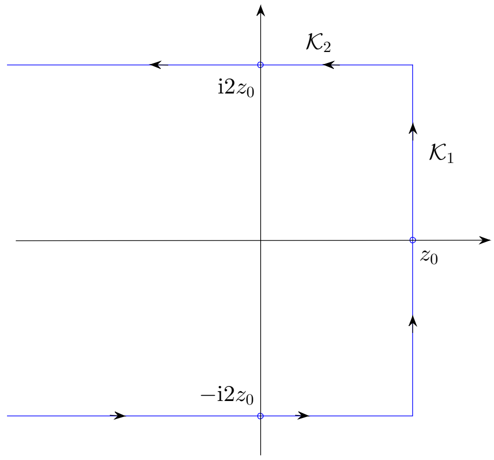

FIG 2. *Deformed contour $\mathcal{K}$ for SMD, PCA, and SigD.*

For the rest of the paper, assume that $\theta \in \left(0, \bar{\theta}\right)$. We deform contour $\mathcal{K}$ in (10) 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

$$\mathcal{K}_1 = \{z_0 + \mathrm{i}t : 0 \leq t \leq 2z_0\} \text{ and} \tag{39}$$

$$\mathcal{K}_2 = \{x + \mathrm{i}2z_0 : -\infty < x \leq z_0\}. \tag{40}$$

The deformed contour is shown on Figure 2.

Note that the singularities of the integrand in (10) are situated at $z = \lambda_j$ (plus an additional singularity at $z = c_1(1+\theta)/\left(\theta c_2\right) < z_0$ for SigD). Since $z_0 > \lambda_1$ holds with probability approaching one as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, Cauchy's theorem ensures that the deformation of the contour does not change the value of $L\left(\theta; \Lambda\right)$ with probability approaching one as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$.

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

REMARK 7. In the event of asymptotically negligible probability that the deformed contour $\mathcal{K}$ does not encircle all λ_j, we not only lose the equality

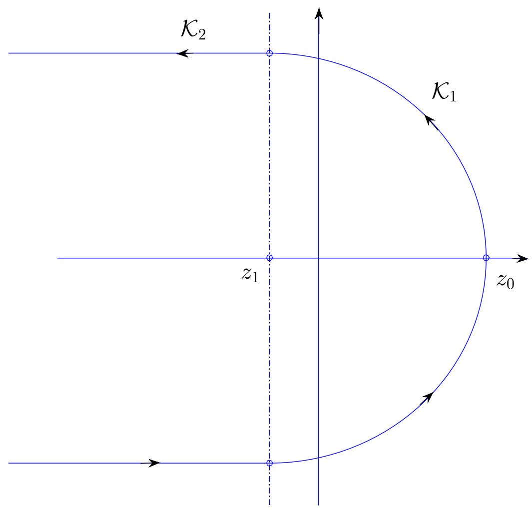

FIG 3. *Deformed contour $\mathcal{K}$ for REG_0 and CCA.*

(10) but also face the difficulty that function $g(z)$ ceases to be well defined as the definition of $g_e(z)$ contains a logarithm of a non-positive number. To eliminate any ambiguity, if such an event holds we shall redefine $g_e(z)$ as unity.

For REG_0 and CCA, let

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

and let

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

The corresponding contour $\mathcal{K}$ is shown on Figure 3. Similarly to the SMD, PCA and SigD cases, the deformation of the contour in (10) to $\mathcal{K}$ does not change the value of $L(\theta; \Lambda)$ with probability approaching one as $\mathbf{n}, p \to_\gamma \infty$.

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

$$\eta_1 = z\theta c_2 / [c_1(1-c_1)] \tag{41}$$

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

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

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

Using (41) and the identity

$$\eta_1 = \tau(\tau + 1) / (\tau + \kappa), \tag{42}$$

we obtain

$$z = \frac{c_1(1 - c_1)}{\theta c_2} \frac{\tau(\tau + 1)}{\tau + \kappa}. \tag{43}$$

We define the deformed contour $\mathcal{K}$ in z-plane as the image of $\mathcal{C}$ under the transformation $\tau \to z$ given by (43). 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 next lemma, proven in SM 4.7, 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 8. *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 integral (9) 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 is large, $\omega \in \Omega \subset \mathbb{R}^k$ is a k-dimensional parameter, and $\mathcal{K}_{p,\omega}$ is a path in $\mathbb{C}$ that starts at $a_{p,\omega}$ and ends at $b_{p,\omega}$. 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. Assume that there is a domain $T_{p,\omega} \supset \mathcal{K}_{p,\omega}$ on which for sufficiently large p, $\phi_{p,\omega}(z)$ and $\chi_{p,\omega}(z)$ are single-valued holomorphic functions of z, in the case of $\chi_{p,\omega}$ with probability increasing to 1.

We describe an extension of the Laplace approximation detailed by Olver [27, p. 127] to a situation in which functions ϕ, χ and contour $\mathcal{K}$ depend on p and ω and in addition χ is random. In Olver's original theorem, both functions and contour are fixed. In what follows, however, we omit subscripts p and ω from $\phi_{p,\omega}$, $\chi_{p,\omega}$, $\mathcal{K}_{p,\omega}$, etc. to lighten notation.

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 be the coefficients in the power series representations

$$\phi(z) = \sum_{s=0}^{\infty} \phi_s (z - z_0)^s, \qquad \chi(z) = \sum_{s=0}^{\infty} \chi_s (z - z_0)^s. \tag{44}$$

We assume that there exist positive constants $C_1, ..., C_4$ that do not depend on p or 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 $|z - z_0| \le C_1$

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

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

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

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

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

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 $\omega \in \Omega$.

Assumptions A0–A5 ensure that Olver's proof of the Laplace approximation theorem (Theorem 7.1 on p. 127 of Olver (1997)) can be extended to cases where functions $\phi(z)$ and $\chi(z)$, as well as the contour $\mathcal{K}$, depend on p and ω. Note that in Olver's notations, $\phi(z)$, $\chi(z)$, and p are, respectively $p(t)$, $q(t)$, and z.

The first part of A0, which requires the boundedness of $|\mathcal{K}|$, taken together with A5 and the assumption that $\mathcal{K}$ is a contour of steep descent guarantee the absolute convergence of the integral $\int_{\mathcal{K}} e^{-p(\phi(z)-\phi_0)}\chi(z)\mathrm{d}z$, in probability. The second part of A0 ensures that as $p\to\infty$, $\mathcal{K}$ does not collapse to a point.

Assumption A1 excludes situations where z_0 approaches singular points of $\phi(z)$ or $\chi(z)$ as $p\to\infty$. Assumption A2 guarantees that the second derivative of $\phi(z)$ at z_0 does not degenerate to 0 or infinity as $p\to\infty$. Assumption A3 implies that $|\phi(z)-\phi(z_0)|$ can be bounded from below by a fixed quadratic function of z in a vicinity of z_0 as $p\to\infty$. This ensures a regular behavior of function $(\phi(z)-\phi(z_0))^{1/2}$. Assumption A4 implies that $|\arg(\phi(z)-\phi(z_0))|<\pi/2$ is some neighborhood of z_0 as $p\to\infty$. We need this condition to be able to use an asymptotic expansion of an incomplete Gamma function in our proofs (Section 5.1 of SM). Assumption A5 ensures that $|\chi(z)|$ remains bounded in probability as $p\to\infty$.

LEMMA 9. *Under assumptions A0-A5, for any positive integer k, as $p\to\infty$, we have*

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

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

Lemma 9 is proved in SM 5.1. We use it to obtain the Laplace approximation to

$$L_1(\theta;\Lambda)=\sqrt{\pi p}\frac{1}{2\pi\mathrm{i}}\int_{\mathcal{K}_1\cup\bar{\mathcal{K}}_1}e^{-(p/2)f(z)}g(z)\mathrm{d}z. \tag{45}$$

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 $\mathrm{d}^2f(z_0)/\mathrm{d}z^2$. As was mentioned in Remark 5, $f(z_0)=0$ for all the six cases that we study. The values of $\mathrm{d}^2f(z_0)/\mathrm{d}z^2$ are derived in SM 5.2. All of them are negative. The explicit form of $D_2\equiv\theta^2\left(-\mathrm{d}^2f(z_0)/\mathrm{d}z^2\right)^{-1}$, which

TABLE 6
The values of $D_2 \equiv \theta^2(-\mathrm{d}^2 f(z_0)/\mathrm{d}z^2)^{-1}$ *for the different cases.*

Case	Value of D_2	Case	Value of D_2
SMD	$1-\theta^2$	REG$_0$	$c_1\,(1+c_1+2\theta)\left(c_1-\theta^2\right)$
PCA	$c_1\left(c_1-\theta^2\right)(1+\theta)^2$	REG	$c_1 h\,(c_1+\theta+(1+\theta)\,l)\,/l^4$
SigD	$r^2 h\,(1+\theta)^2\,/l^4$	CCA	$c_1^2 h\,(2\,(c_1+\theta)+l\,(1-c_1))\,/\left(l^3\,(c_1+c_2)\right)$
	$l\equiv l(\theta)=1+(1+\theta)c_2/c_1$		$h\equiv h(\theta)=c_1+c_2(1+\theta)^2-\theta^2$

is somewhat shorter than that for $\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$, is reported in Table 6. We formulate the main result of this section in the following theorem, proven in SM 5.3.

THEOREM 10. *Suppose that the null hypothesis holds, that is,* $\theta_0 = 0$*. Let* $\bar\theta$ *be the threshold corresponding to* $F_{\boldsymbol\gamma}$ *as given in Table 3, and let* ε *be an arbitrarily small fixed positive number. Then for any* $\theta \in \left(0, \bar\theta-\varepsilon\right]$*, as* $\mathbf{n}, p \to_{\boldsymbol\gamma} \infty$*, we have*

$$L\,(\theta;\Lambda) = \frac{g(z_0)}{\sqrt{-\mathrm{d}^2 f(z_0)/\mathrm{d}z^2}} + O_{\mathrm{P}}\left(p^{-1}\right), \tag{46}$$

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 10 with the definitions of $g(z)$ and the values of $-\mathrm{d}^2 f(z_0)/\mathrm{d}z^2$, given in Table 6, it is straightforward to establish the following theorem, details in SM 6.1. Let

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

In accordance with Remark 7, we define $\Delta_p(\theta)$ as zero in the event of asymptotically negligible probability that $z_0 \le \lambda_1$.

THEOREM 11. *Suppose that the null hypothesis holds, that is* $\theta_0 = 0$*. Let* $\bar\theta$ *be the threshold corresponding to* $F_{\boldsymbol\gamma}$ *as given in Table 3, and let* ε *be an arbitrarily small fixed positive number. Then for any* $\theta \in \left(0,\bar\theta-\varepsilon\right]$*, as* $\mathbf{n}, p \to_{\boldsymbol\gamma} \infty$*, we have*

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

where

$$\delta_p(\theta) = \begin{cases} \theta & \text{for SMD} \\ \theta/\sqrt{c_1} & \text{for PCA and } REG_0 \\ \theta r/(c_1 l(\theta)) & \text{for SigD, REG, and CCA,} \end{cases}$$

$r^2 = c_1 + c_2 - c_1 c_2$ *and* $o_{\mathrm{P}}(1)$ *is uniform in* $\theta \in (0, \bar{\theta} - \varepsilon]$.

Statistic $\Delta_p(\theta)$ is a linear spectral statistic. As follows from the CLT for such statistics derived by [7], [6], and [38] 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 $\theta \in (0, \bar{\theta} - \varepsilon]$. 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 [7, Th. 1.1] for SMD, in [6, Th. 1.1] for PCA and REG_0, and in [38, Th. 4.1 and Exmpl. 4.1] for the remaining cases. SM 6.2 provides details on the use of [7, 6, 38] to establish convergence of $\Delta_p(\theta)$, and the use of Theorem 11 to obtain the following theorem.

THEOREM 12. *Suppose that the null hypothesis holds, that is* $\theta_0 = 0$. *Let* $\bar{\theta}$ *be the threshold corresponding to* $F_{\boldsymbol{\gamma}}$ *as given in Table 3, and let* ε *be an arbitrarily small fixed positive number. Further, let* $C[0, \bar{\theta} - \varepsilon]$ *be the space of continuous functions on* $[0, \bar{\theta} - \varepsilon]$ *equipped with the supremum norm. Then* $\ln L(\theta; \Lambda)$ *viewed as random elements of* $C[0, \bar{\theta} - \varepsilon]$ *converge weakly to* $\mathcal{L}(\theta)$ *with Gaussian finite dimensional distributions such that*

$$\mathbb{E}\mathcal{L}(\theta) = \tfrac{1}{4}\ln(1 - \delta^2(\theta))$$

and

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

with

$$\delta(\theta) = \begin{cases} \theta & \text{for SMD} \\ \theta/\sqrt{\gamma_1} & \text{for PCA and } REG_0 \\ \theta\rho/(\gamma_1 + \gamma_2 + \theta\gamma_2) & \text{for SigD, REG, and CCA} \end{cases}.$$

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

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 12 implies, via Le Cam's first lemma, the mutual contiguity of $\{\mathbb{P}_{p,\theta}\}$ and $\{\mathbb{P}_{p,0}\}$ as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, for each $\theta < \bar{\theta}$. 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,[1] 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 should directly use Theorem 12.

Here we use it to derive the asymptotic power envelopes for tests of the null hypothesis $\theta_0 = 0$ against the point alternative $\theta_0 > 0$. By the Neyman-Pearson lemma, the most powerful test would reject the null when $\ln L(\theta; \Lambda)$ is above a critical value. By Theorem 12 and Le Cam's third lemma (see [36, Ch. 6]),

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

with the plus sign holding under the null, and the minus under the alternative. This implies the following theorem.

THEOREM 13. *Let $\bar{\theta}$ be the threshold corresponding to $F_{\boldsymbol{\gamma}}$ as given in Table 3. 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*

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

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 differ only for cases with different limiting spectral distributions: Semi-circle, Marchenko-Pastur, and Wachter, denoted $PE^{\mathrm{SC}}(\theta)$, $PE^{\mathrm{MP}}(\theta, \gamma_1)$ and $PE^{\mathrm{W}}(\theta, \boldsymbol{\gamma})$ respectively. Figure 4 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 degrees of freedom n_1 and n_2.

Envelope $PE^{\mathrm{MP}}(\theta, \gamma_1)$ can be obtained from $PE^{\mathrm{W}}(\theta, \boldsymbol{\gamma})$ by sending γ_2 to zero. Further, $PE^{\mathrm{SC}}(\theta)$ can be obtained from $PE^{\mathrm{MP}}(\theta, \gamma_1)$ by transformation $\theta \longmapsto \sqrt{\gamma_1}\theta$. Further, note the difference in the horizontal scale of the bottom panel of Figure 4 relative to the two other panels. For $\gamma_1 = \gamma_2 = 0.9$

[1] [17] has an interesting discussion of ubiquity of random processes with logarithmic covariance structure in physics and engineering applications. In that paper, such processes appear as limiting objects related to the behavior of the characteristic polynomials of large matrices from Gaussian Unitary Ensemble.

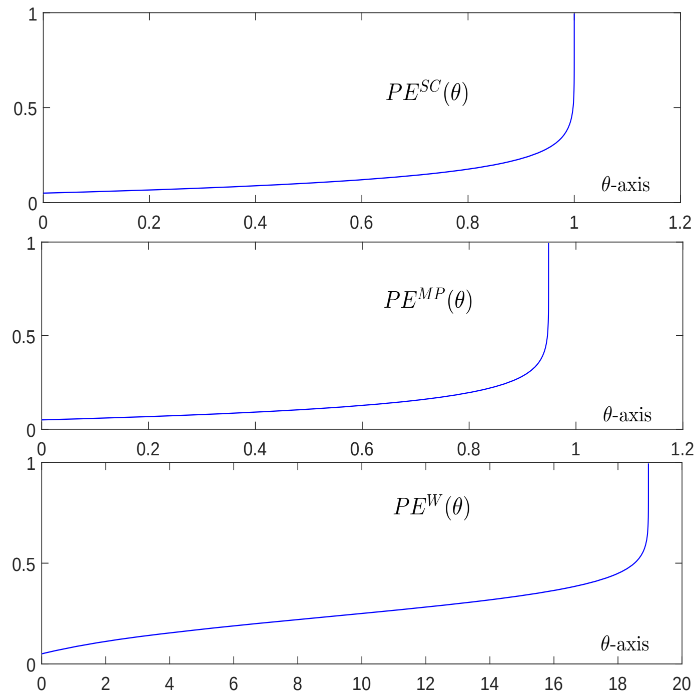

FIG 4. *The asymptotic power envelopes* $PE^{\mathrm{SC}}(\theta)$, $PE^{\mathrm{MP}}(\theta, \gamma_1)$, *and* $PE^{\mathrm{W}}(\theta, \gamma)$ *for* $\alpha = 0.05$, $\gamma_1 = \gamma_2 = 0.9$.

the phase transition threshold corresponding to the Wachter distribution is relatively large. It equals $(\gamma_2 + \rho)/(1 - \gamma_2) \approx 18.9$. Moreover, the value of $PE^{\mathrm{W}}(\theta)$ 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. Concluding remarks. Note that Theorem 12 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. It also can be used to construct asymptotic confidence intervals for a sub-critical spike as well as to describe the asymptotic properties of its maximum likelihood estimator. We do not pursue this line of research here, but provide a general outline.

Consider the log likelihood ratio $\ln L(\theta;\Lambda) - \ln L(\theta_0;\Lambda)$. According to Theorem 12, this ratio converges to $X(\theta) \equiv \mathcal{L}(\theta) - \mathcal{L}(\theta_0)$. By Le Cam's third lemma, under the null hypothesis that the true value of the spike equals θ_0, $X(\theta)$ is a Gaussian process with mean

$$\mathbb{E}X(\theta) = \frac{1}{4}\ln\frac{\left(1-\delta^2(\theta)\right)\left(1-\delta^2(\theta_0)\right)}{\left(1-\delta(\theta)\delta(\theta_0)\right)^2}$$

and covariance function

$$\mathbb{C}ov\left(X(\theta_1), X(\theta_2)\right) = -\frac{1}{2}\ln\frac{\left(1-\delta(\theta_1)\delta(\theta_2)\right)\left(1-\delta(\theta_0)\delta(\theta_0)\right)}{\left(1-\delta(\theta_1)\delta(\theta_0)\right)\left(1-\delta(\theta_2)\delta(\theta_0)\right)}.$$

An approximation to the distribution of the supremum of such a process over $\theta \in [0, \bar{\theta}-\varepsilon]$ can be obtained via simulation. Alternatively, it might be expressed analytically in the form of converging Rice series (see e.g. [1]). Quantiles of the distribution can be used as asymptotic critical values for the likelihood ratio test of the hypothesis $\theta = \theta_0$. Inverting the test, we obtain asymptotic confidence intervals for the true value of a sub-critical spike.

The maximum likelihood estimator for the spike, $\hat{\theta}_{ML}$, equals the arg max of $\ln L(\theta;\Lambda) - \ln L(\theta_0;\Lambda)$ over $\theta \in [0, \bar{\theta}-\varepsilon]$. By Lemma 2.6 of [19], the limiting process $X(\theta)$ achieves maximum at a unique point with probability one. Therefore by the argmax continuous mapping theorem, $\hat{\theta}_{ML}$ converges in distribution to the arg max of $X(\theta)$. The distribution of such an arg max can be approximated using simulations.

Unfortunately, the quality of the estimator $\hat{\theta}_{ML}$ cannot be "good". For PCA, we were able to prove that no estimator of θ has root mean squared error better than the order of magnitude of the sub-critical parameter θ. This result will appear in another work.

Our asymptotic discussion of James' framework can likely be extended to a fixed number of sub-critical spikes. Such an extension would require developing Laplace approximations to multiple contour integrals, and uniform approximations to hypergeometric functions of two matrix arguments in terms of elementary functions. Alternatively, one may employ large deviation analysis of spherical integrals as in [31], which covers the PCA case. As this paper is already long, the extension will appear separately.

Addressing the case of slowly increasing number of spikes may require new techniques, perhaps, similar to those developed in [14]. In such a case, relatively little is known even about the phase transition phenomenon. For sample covariance matrices, Theorem 1.1 of [5] can be used to show that the phase transition still happens at the usual threshold $\bar{\theta} = \sqrt{\gamma_1}$. However, it is not clear whether the experiments of observing sample covariance eigenvalues corresponding to the null case and an alternative with a growing number of sub-critical eigenvalues remain mutually contiguous.

Note that, intuitively, the asymptotic power of the likelihood ratio test of the null hypothesis of no spikes against the alternative of one spike should not decrease if the rank-one assumption on the alternative is wrong and there are additional spikes. In SM, we confirm this intuition for SMD and PCA cases. A confirmation or refutation of the intuition for the other James' cases requires further analysis and is left for future research.

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 probably 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. For PCA the case $p > n_1$ was explicitly covered in [30]. For REG_0 the assumption can be relaxed using the symmetry of the problem.

Specifically, the canonical REG_0 problem tests restriction $M = 0$ in the model $Y = M + \varepsilon$, where all matrices are $n_1 \times p$ and ε has i.i.d. standard normal components. Clearly, interchanging roles of n_1 and p yields essentially the same problem.

For CCA, the sample canonical correlations are not well defined for $p > n_1$, so we are not interested in such a case. This leaves us with SigD and REG cases, which we mark as more difficult from the point of view of relaxing $n_1 \geq p$ assumption.

For SigD, our derivations (not reported here) show that the equivalent of (6) for $p > n_1$ involves the hypergeometric function ${}_2F_1$. Therefore, SigD with $p > n_1$ represents the fifth, rather than the second, group of multivariate statistical problems according to James' (1964) classification. For REG, an equivalent of (6) for $p > n_1$ can be obtained using [18, eq. (74)]. However, further analysis of SigD and REG in the situation where $p > n_1$ needs more substantial changes to our derivations. We leave such an analysis for future research.

Finally, 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 Gaussian assumption. We hope that the results of this paper might provide a benchmark for such future studies.

Acknowledgements. We thank two anonymous referees, an Associate Editor and the Editor for careful reading of the manuscript and helpful

comments.

SUPPLEMENTARY MATERIAL

Supplement A:
(link TBA). The supplement has proofs for all results in the paper, organized by section for easier cross-reference.

References.

[1] Azais, J-M. and Wschebor, M. (2002). The distribution of the maximum of a Gaussian process: Rice method revisited, in Sidoravicius, V. (ed.) *In and Out of Equilibrium: Probability with a Physical Flavour, Progress in Probability.* Birkhauser, Basel.

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

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

[4] Bai, Z. D. and Yin, Y. Q. (1988). Convergence to the semicircle law. *Annals of Probability* **2** 863–875.

[5] Bai, Z. D. and Silverstein, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Annals of Probability* **26** 316–345.

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

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

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

[9] Baik, J. and Wang, D. (2013). On the largest eigenvalue of a Hermitian random matrix model with spiked external source II: Higher rank cases. Int. Math. Res. Not. IMRN, **14** 3304–3370.

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

[11] Cai, J-F., Candés, E. J. and Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization* **20** 1956–1982.

[12] Dharmawansa, P. and Johnstone, I. M. (2014). Joint density of eigenvalues in spiked multivariate models. *Stat* **3** 240–249.

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

[14] Dobriban, E. (2017). Sharp detection in PCA under correlations: All eigenvalues matter. *Annals of Statistics* **45** 1810–1833.

[15] Donoho, D. and Gavish, M. (2014). Minimax risk of matrix denoising by singular value thresholding. *Annals of Statistics* **42** 2413–2440.

[16] Fujikoshi, Y., Ulyanov, V. V. and Shimizu, R. (2010). *Multivariate statistics: High-dimensional and large-sample approximations.* John Wiley & Sons Inc., Hoboken, NJ.

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

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

[19] Kim, J. and Pollard, D. (1990) Cube root asymptotics. *Annals of Statistics* **18** 191–219.

[20] Lelarge, M. and Miolane, L. (2016) Fundamental limits of symmetric low-rank matrix estimation. *arXiv 1611.03888*.

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

[22] Mehta, L. M. (2004). *Random Matrices*, 3rd ed. Elsevier Academic Press.

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

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

[25] Nadakuditi, R. R. (2014). OptShrink: an algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage. *IEEE Trans. Inform. Theory* **60** 3002–3018.

[26] Nadakuditi, R. R. and Silverstein, J. W. (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** 468–480.

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

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

[29] 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.

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

[31] Onatski, A., Moreira, M. J. and Hallin, M. (2014). Signal detection in high dimension: the multispiked case. *Annals of Statistics* **42** 225–254.

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

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

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

[35] Shabalin, A. A. and Nobel, A. B. (2013). Reconstruction of a low-rank matrix in the presence of Gaussian noise. *J. Multivariate Anal.* **118** 67–76.

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

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

[38] 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.

E-mail: imj@stanford.edu

E-mail: ao319@cam.ac.uk

SUPPLEMENTARY MATERIAL FOR "TESTING IN HIGH-DIMENSIONAL SPIKED MODELS."

BY IAIN M. JOHNSTONE AND ALEXEI ONATSKI

This note contains supplementary material for Johnstone and Onatski (2016) (JO in what follows). It is lined up with sections in the main text to make it relatively easy to see how and where the proof details fit in.

CONTENTS

1. Introduction. There is no supplementary material for the Introduction section of JO.

2. Links to classical statistical problems.

2.1. *Sufficiency and invariance considerations.* In this subsection, we clarify which sufficiency and invariance arguments lead us to consider tests based on the solutions of

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

and

$$\det \left(H - \lambda \left(E + \frac{n_1}{n_2} H \right) \right) = 0 \tag{2}$$

for SMD, PCA, SigD, RED_0, REG, and CCA problems. Most of this discussion is standard and can be found, for example, in Muirhead (1982).

SMD: Consider the group of transformations

$$G = \{U : U \in \mathcal{O}(p)\}, \tag{3}$$

where $\mathcal{O}(p)$ is the group of $p \times p$ orthogonal matrices, acting on the space of $p \times p$ symmetric matrices $X = \theta\psi\psi' + GOE/\sqrt{p}$ by

$$U \circ X = UXU'.$$

The corresponding induced group of transformations on the parameter space of points (ψ, θ) is given by

$$U \circ (\psi, \theta) = (U\psi, \theta).$$

A maximal invariant in the parameter space is θ, whereas that in the sample space is given by the ordered eigenvalues $\lambda_1 \geq ... \geq \lambda_p$ of X. Since neither the null nor the alternative hypothesis,

$$H_0 : \theta_0 = 0 \text{ and } H_1 : \theta_0 > 0, \tag{4}$$

is affected by the transformations, it is natural to base the test on the maximal invariant in the sample space.

PCA: In this case, the data are given by $X \sim N(0, \Omega \otimes I_{n_1})$, where $\Omega = \Sigma + \theta\psi\psi'$, where Σ is a known positive definite symmetric matrix and $\left\|\Sigma^{-1/2}\psi\right\| = 1$. Without loss of generality, we can set $\Sigma = I_p$. A sufficient statistic is $H = XX'/n_1$. Consider the group of transformations (3) that acts on the sample space of the sufficient statistic by

$$U \circ H = UHU',$$

and on the parameter space by

$$U \circ (\psi, \theta) = (U\psi, \theta).$$

The maximal invariant in the parameter space is θ, and we base the test of (4) on a maximal invariant in the sample space of the sufficient statistic, which is given by the ordered eigenvalues of the sample covariance matrix H.

SigD: The data are given by independent matrices

$$X \sim N(0, \Omega \otimes I_{n_1}) \text{ and } Y \sim N(0, \Sigma \otimes I_{n_2}),$$

where $\Omega = \Sigma + \theta\psi\psi'$, Σ is an unknown positive definite symmetric matrix, and $\left\|\Sigma^{-1/2}\psi\right\| = 1$. A sufficient statistic consists of the sample covariance matrices $H = XX'/n_1$ and $E = YY'/n_2$. Let $\mathcal{GL}(p)$ be the group of non-singular $p \times p$ matrices. Consider the group of transformations $G = \{B : B \in \mathcal{GL}(p)\}$ that acts on the space of points $(H, E) \in \mathcal{S}_p \times \mathcal{S}_p$, where $\mathcal{S}_p$ is the space of positive definite symmetric $p \times p$ matrices, by

$$B \circ (H, E) = (BHB', BEB')$$

and on the parameter space by

$$B \circ (\Sigma, \psi, \theta) = (B\Sigma B', B\psi, \theta).$$

Note that we restrict the sample space to $\mathcal{S}_p \times \mathcal{S}_p$, that is we exclude from consideration zero-probability event where the matrix HE is singular. The maximal invariant in the parameter space is θ and we base the test of (4) on a maximal invariant in the sample space of the sufficient statistic, which is given by the ordered solutions to (1) or to (2) (see Theorem 8.2.2 of Muirhead (1982)). The links between SigD and PCA become particularly clear when we work with the solutions to (2).

Note that we can assume that $\Sigma = I_p$ wlog. It is because $\lambda_1 \geq ... \geq \lambda_p$ that solve equation (2) are invariant with respect to the transformation

$$(H, E) \mapsto (\Sigma^{-1/2} H \Sigma^{-1/2}, \Sigma^{-1/2} E \Sigma^{-1/2}).$$

In particular, the joint distribution of $\lambda_1 \geq ... \geq \lambda_p$ under the null hypothesis $H_0 : \Omega = \Sigma$ is the same as in the case where $\Omega = \Sigma = I_p$. Similarly, the joint distribution of $\lambda_1 \geq ... \geq \lambda_p$ under the alternative $H_1 : \Omega = \Sigma + \theta\psi\psi'$ with $\left\|\Sigma^{-1/2}\psi\right\| = 1$ is the same as in the situation where $\Omega = I_p + \theta\psi\psi'$ with $\|\psi\| = 1$ and $\Sigma = I_p$.

REG$_0$: Consider linear regression $Y = X\beta + \varepsilon$, where Y is $T \times p$, X is $T \times q$, β is $q \times p$, and ε has i.i.d. $N(0, \Sigma)$ rows. For REG$_0$, Σ is a know symmetric positive definite matrix, which can be set to I_p wlog. We would like to test a general linear hypothesis $C\beta = 0$, where C is a known $n_1 \times q$ matrix of rank n_1.

As explained in Muirhead (1982, pp 433-434), the problem can be cast in the canonical form, where the matrix of transformed response variables is split into three parts: an $n_1 \times p$ matrix Y_1, a $(q - n_1) \times p$ matrix Y_2, and an $n_2 \times p$ matrix Y_3 with $n_2 = T - q$. Under the null hypothesis, $M \equiv \mathbb{E}Y_1 = 0$, whereas under the alternative,

$$M = \sqrt{n_1\theta}\varphi\psi', \tag{5}$$

where $\theta > 0$, $\|\Sigma^{-1/2}\psi\| = 1$, and $\|\varphi\| = 1$. Matrices Y_2 and Y_3 have, respectively, unrestricted and zero means under both the null and the alternative.

In terms of the original regression model, matrix M can be expressed as the product of an invertible matrix, which depends only on C and X, and matrix $C\beta$. In particular, $M = 0$ if and only if $C\beta = 0$. Alternative (5) corresponds to a rank-one alternative $C\beta = \sqrt{n_1\theta}\tilde{\varphi}\psi'$ in the original model, where vector $\tilde{\varphi}$ is obtained from vector φ via a linear transformation that depends on matrices C and X.

A sufficient statistic for $\sqrt{n_1\theta}\varphi\psi'$ is Y_1. Consider a group of transformations

$$G = \{(U, V) : U \in \mathcal{O}(p), V \in \mathcal{O}(n_1)\}$$

that acts on the points Y_1 of the sample space $\mathbb{R}^{n_1 \times p}$ by

$$(U, V) \circ Y_1 = VY_1U'$$

and on the parameter space by

$$(U, V) \circ (\varphi, \psi, \theta) = (V\varphi, U\psi, \theta).$$

A maximal invariant in the parameter space is θ, whereas the maximal invariant statistic consists of the ordered eigenvalues of $H = Y_1Y_1'/n_1$.

REG: The difference between the cases REG and REG$_0$ is that in REG Σ is assumed to be an unknown matrix from $\mathcal{S}_p$. The sufficient statistic now is (Y_1, Y_2, E), where $E = Y_3'Y_3/n_2$. Consider a group of transformations

$$G = \left\{(B, V, A) : B \in \mathcal{GL}(p), V \in \mathcal{O}(n_1), A \in \mathbb{R}^{(T - n_1 - n_2) \times p}\right\}$$

that acts on the points (Y_1, Y_2, E) of the sample space $\mathbb{R}^{n_1 \times p} \times \mathbb{R}^{T-n_1-n_2 \times p} \times \mathcal{S}_p$ by

$$(B, V, A) \circ (Y_1, Y_2, E) = (VY_1B', Y_2B' + A, BEB')$$

and on the parameter space by

$$(B, V, A) \circ (\varphi, \psi, \theta, M, \Sigma) = (V\varphi, B\psi, \theta, MB' + A, B\Sigma B')\,.$$

A maximal invariant in the parameter space is θ, whereas the maximal invariant in the sample space consists of the ordered roots of equation (2), where $H = Y_1{\prime}Y_1/n_1$ and $E = Y_3'Y_3/n_2$ (see Theorem 10.2.1 on page 437 of Muirhead (1982)).

CCA: For this case, the sufficient statistic is $S = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix}$. Consider a group of transformations

$$G = \{B : B = diag\,\{B_1, B_2\}\,, B_1 \in \mathcal{GL}\,(p)\,,\ B_2 \in \mathcal{GL}\,(n_1)\}$$

acting on the sample space, restricted so that S_{xx} and S_{yy} are invertible, by

$$B \circ S = BSB'.$$

On the parameter space, the group acts by

$$B \circ (\Sigma_{xx}, \Sigma_{yy}, \psi, \varphi, \theta) = (B_1\Sigma_{xx}B_1', B_2\Sigma_{yy}B_2', B_1\psi, B_2\varphi, \theta)\,.$$

As follows from Muirhead's (1982) Theorem 11.2.2, a maximal invariant in the parameter space is θ and that in the sample space consists of the solutions to (1) with

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

2.2. *Sequential asymptotic links between the cases.* **PCA→SMD:** Recall that the relevant data for PCA case are represented by the solutions to equation (1) with $E = \Sigma$ and $n_1 H \sim W_p\,(n_1, \Omega)$. Let

$$\Omega = \Sigma + \sqrt{p/n_1}\theta\psi\psi' \tag{6}$$

with $\left\|\Sigma^{-1/2}\psi\right\| = 1$. That is, let the value of the spike in the original version of PCA be scaled by $\sqrt{p/n_1}$. Equation (6) implies that

$$\Sigma^{-1} = \Omega^{-1} + \frac{\sqrt{p/n_1}\theta}{1 + \sqrt{p/n_1}\theta}\Sigma^{-1}\psi\psi'\Sigma^{-1},$$

and therefore, equation (1) is equivalent to

$$\det\left(\left(\Omega^{-1} + \frac{\sqrt{p/n_1}\theta}{1 + \sqrt{p/n_1}\theta}\Sigma^{-1}\psi\psi'\Sigma^{-1}\right)H - \lambda I_p\right) = 0,$$

which, in its turn, is equivalent to

$$\det\left(\Omega^{-1/2}H\Omega^{-1/2} + \sqrt{p/n_1}\theta\eta\eta'\Omega^{-1/2}H\Omega^{-1/2} - \lambda I_p\right) = 0, \tag{7}$$

where

$$\eta = \Omega^{1/2}\Sigma^{-1}\psi / \left(1+\sqrt{p/n_1}\theta\right)^{1/2}$$

is such that $\|\eta\| = 1$. The latter equality follows from the fact $\psi'\Sigma^{-1}\Omega\Sigma^{-1}\psi = 1 + \sqrt{p/n_1}\theta$, which is a consequence of (6) and of the normalization $\left\|\Sigma^{-1/2}\psi\right\| = 1$.

Now assume that n_1 diverges to infinity while p is held constant. Then, by a CLT

$$\Omega^{-1/2}H\Omega^{-1/2} = I_p + Z/\sqrt{n_1} + o_{\mathrm{P}}\left(n_1^{-1/2}\right), \tag{8}$$

where Z belongs to GOE. Multiplying (7) by $(n_1/p)^{p/2}$ and using (8), we see that, as $n_1 \to \infty$, equation (7) degenerates to

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

Hence, PCA degenerates to SMD.

SigD→PCA: As shown in JO, SigD degenerates to PCA as $n_2 \to \infty$ while n_1 and p are held constant. Therefore, SigD can be linked to SMD via PCA.

REG$_0$ →SMD: 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 (JO4)) is scaled by $(p/n_1)^{1/2}$. Suppose now that n_1 diverges to infinity while p is held constant. Then, by a CLT,

$$\Sigma^{-1/2}H\Sigma^{-1/2} - I_p = Z/\sqrt{n_1} + \sqrt{p/n_1}\theta\eta\eta' + o_{\mathrm{P}}\left(n_1^{-1/2}\right), \tag{9}$$

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

$$\det\left(\Sigma^{-1/2}H\Sigma^{-1/2} - \lambda I_p\right) = 0. \tag{10}$$

Multiplying it by $(n_1/p)^{p/2}$ and using (9), we see that equation (10) degenerates to

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

Hence, REG$_0$ degenerates to SMD.

REG→REG$_0$: The REG case degenerates to REG$_0$ as $n_2 \to \infty$ while n_1 and p are held constant. Therefore, REG can be linked to SMD via REG$_0$.

CCA→REG$_0$: Recall that the CCA case is based on the solutions to equation (1) with

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

where S_{xx} and S_{yy} are sample covariance matrices corresponding to i.i.d. $N(0, \Sigma_{xx})$ sample $x_t \in \mathbb{R}^p$, $t = 1, ..., n_1 + n_2$, and i.i.d. $N(0, \Sigma_{yy})$ sample $y_t \in \mathbb{R}^{n_1}$, $t = 1, ..., n_1 + n_2$, respectively. Matrices S_{xy} and S_{yx} are the corresponding sample cross-covariance matrices. Since the transformations

$x_t \mapsto \Sigma_{xx}^{-1/2} x_t$ and $y_t \mapsto \Sigma_{yy}^{-1/2} y_t$ do not affect the roots of (1), we shall assume without loss of generality that $\Sigma_{xx} = I_p$ and $\Sigma_{yy} = I_{n_1}$. Recall that, by assumption,

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

Suppose that n_2 diverges to infinity while n_1 and p are held constant. Then, by a CLT,

$$S_{xx} = I_p + o_{\mathrm{P}}(1)\,,\ S_{yy} = I_{n_1} + o_{\mathrm{P}}(1)\,,$$

whereas

$$S_{xy} = \Sigma_{xy} + Z_{xy}/\sqrt{n_1 + n_2} + o_{\mathrm{P}}\left((n_1+n_2)^{-1/2}\right),$$

where Z_{xy} is a $p \times n_1$ matrix with i.i.d. $N(0,1)$ entries. Therefore, equation (1) degenerates to

$$\det\left(\frac{1}{n_1}\left(\tilde{\Sigma}_{xy} + Z_{xy}\right)\left(\tilde{\Sigma}_{xy} + Z_{xy}\right)' - \nu I_p\right) = 0 \tag{11}$$

with

$$\tilde{\Sigma}_{xy} = \sqrt{n_1\theta}\psi\varphi'$$

and

$$\nu = (1 + n_2/n_1)\,\lambda.$$

Hence, CCA degenerates to REG_0. It can further be linked to SMD via REG_0.

3. The likelihood ratios.

3.1. *SMD entry of Table JO2.* The explicit expression for $L^{(\mathrm{SMD})}(\theta;\Lambda)$ given in Table JO2 follows from the following lemma.

LEMMA 1. *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*

$$c_p(x)\exp\left\{-p\theta^2/4\right\}{}_0F_0(\Psi, x)\,, \tag{12}$$

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

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

where $\mathrm{etr}\{\cdot\}$ denotes the exponential trace function, and $(\mathrm{d}u)$ 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. Using the strategy of the proof of Muirhead's (1982) Theorem 3.2.17, integrate H out to obtain (12) with

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

where $\Gamma_p(p/2)$ is the multivariate Gamma function. □

3.2. *Identification of the parameters of expression (JO6).* For the reader's convenience, we provide some extra detail on the identification of the parameters of expression (JO6) for the likelihood ratio $L(\theta;\Lambda)$ summarized in Table JO2. To have a self-contained source for derivations, we refer below to Muirhead (1982), henceforth [M], in addition to James (1964), [J] below.

Some Notational conventions. $|A| = \det(A)$, and c_{pn} for a constant depending only on p, n. The hypergeometric function

$${}_{\mathsf{p}}F_{\mathsf{q}}(a,b;A,B) = \int_{\mathcal{O}(p)} {}_{\mathsf{p}}F_{\mathsf{q}}(a,b;AHBH')(\mathrm{d}H),$$

We sometimes drop explicit mention of the parameter vectors a, b, and write ${}_{\mathsf{p}}F_{\mathsf{q}}[A;B]$. In particular, we have

$$ {}_{\mathsf{p}}F_{\mathsf{q}}[A;B] = {}_{\mathsf{p}}F_{\mathsf{q}}[B;A] \qquad \text{and} \qquad {}_{\mathsf{p}}F_{\mathsf{q}}[cA;B] = {}_{\mathsf{p}}F_{\mathsf{q}}[A;cB], \tag{13}$$

and

$${}_{\mathsf{p}}F_{\mathsf{q}}[A;0] = {}_{\mathsf{p}}F_{\mathsf{q}}[0] = 1$$

For ${}_0F_0(A) = \operatorname{etr}(A)$ we also have

$${}_0F_0(A, I+C) = \operatorname{etr}(A)\,{}_0F_0(A,C). \tag{14}$$

To indicate the extension to rank r perturbations, we write $\psi = [\psi_1 \cdots \psi_r]$ for a $p \times r$ matrix with $\psi'\Sigma^{-1}\psi = I_r$, $\theta = \operatorname{diag}(\theta_1,\ldots,\theta_r)$, $1+\theta$ for $I_r + \theta$ and $\sqrt{\theta}$ for $\operatorname{diag}(\sqrt{\theta}_1,\ldots,\sqrt{\theta}_r)$.

PCA. [J, eq. (58)], [M, Th. 9.4.1]. We assume a $p \times n_1$ matrix $X \sim N(0, \Omega \otimes I_{n_1})$ with $\Omega = \Sigma + \psi\theta\psi'$ for $\Sigma > 0$ and $\psi'\Sigma^{-1}\psi = I_r$. Without loss of generality we can set $\Sigma = I_p$. The matrix $n_1 H = XX'$ has eigenvalues $\Lambda = \operatorname{diag}(\lambda_i)$. Using the dictionary

M:	S	m	n	Σ	L
JO:	H	p	n_1	Ω	Λ

[M, Th. 9.4.1] gives the joint density of Λ as

$$p(\Lambda|\Omega) = c_{pn_1}|\Omega|^{-n_1/2}|\Lambda|^{(n_1-p-1)/2}v(\Lambda)\,{}_0F_0(-\tfrac{1}{2}n_1\Lambda, \Omega^{-1}),$$

where

$$v(\Lambda) = \prod_{i<j}^{p}(\lambda_i - \lambda_j).$$

Since ${}_0F_0(A,I) = \operatorname{etr}(A)$, the likelihood ratio

$$L(\theta;\Lambda) = \frac{p(\Lambda|\Omega)}{p(\Lambda|I)} = |\Omega|^{-n_1/2}\operatorname{etr}(\tfrac{1}{2}n_1\Lambda)\,{}_0F_0(-\tfrac{1}{2}n_1\Lambda, \Omega^{-1}).$$

We have $|\Omega| = |I+\theta|$, and $\Omega^{-1} = I - \psi\theta(1+\theta)^{-1}\psi'$, and referring to (14), we obtain

$$ {}_0F_0(-\tfrac{1}{2}n_1\Lambda, \Omega^{-1}) = \mathrm{etr}(-\tfrac{1}{2}n_1\Lambda)\, {}_0F_0(\tfrac{1}{2}n_1\Lambda, \psi\theta(1+\theta)^{-1}\psi'), $$

and arrive at

$$ L(\theta;\Lambda) = |1+\theta|^{-n_1/2}\, {}_0F_0(\tfrac{1}{2}n_1\psi\theta(1+\theta)^{-1}\psi', \Lambda). $$

SigD. [J, eq. (65), citing Constantine (unpublished)], [M, Th. 8.2.8]. Now assume independent matrices

$$ X \sim N(0, \Omega\otimes I_{n_1}) \qquad \text{and} \qquad Y \sim N(0, \Sigma\otimes I_{n_2}), $$

with dimensions $p\times n_1$ and $p\times n_2$, and $\Omega = \Sigma + \psi\theta\psi'$ for $\Sigma > 0$ unknown and $\psi'\Sigma^{-1}\psi = I_r$. Without loss of generality (wlog) we can again set $\Sigma = I_p$. The sample covariance matrices are given by

$$ H = XX'/n_1 \qquad \text{and} \qquad E = YY'/n_2. $$

Using now the dictionary

M:	A_1	A_2	m	n_1	n_2	Σ_1	Σ_2	Δ
JO:	n_1H	n_2E	p	n_1	n_2	Ω	I_p	Ω

,

[M, Th. 8.2.8] gives the joint density of the eigenvalues $F = \mathrm{diag}(f_1,\ldots,f_p)$ of

$$ \det(n_1H - f_in_2E) = 0 \tag{15} $$

as

$$ p(F|\Omega) = c_{pn_1n_2}|\Omega|^{-n_1/2}|F|^{(n_1-p-1)/2}v(F)\, {}_1F_0(\tfrac{1}{2}n; -\Omega^{-1}, F), $$

where $n = n_1+n_2$. It is helpful to transform the hypergeometric function using [M, Lemma 8.2.10], due to Khatri (1967), which says here that

$$ {}_1F_0[-\Omega^{-1}, F] = |I+F|^{-n/2}\, {}_1F_0[I-\Omega^{-1}, F(I+F)^{-1}] $$

Note that, as for PCA, $I-\Omega^{-1} = \psi\theta(1+\theta)^{-1}\psi'$. The (generalized) eigenvalues $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ of (JO3) are seen to be related to those of (15) via the transformation $\Lambda = (n_2/n_1)F(I+F)^{-1}$. In forming the likelihood ratio, terms not depending on θ cancel, including the Jacobian of this transformation. Hence we arrive at

$$ L(\theta;\Lambda) = \frac{p(\Lambda|\Omega)}{p(\Lambda|I)} = \frac{p(F|\Omega)}{p(F|I)} = |1+\theta|^{-n_1/2}\, {}_1F_0[(n_1/n_2)\psi\theta(1+\theta)^{-1}\psi', \Lambda]. $$

REG$_0$. [J, eq. (68)], [M, Exer. 10.9]. After reduction to canonical form, we assume that we observe an $n_1\times p$ matrix $Y_1 \sim N(M, I_{n_1}\otimes\Sigma)$. The unnormalized sample covariance matrix $n_1H = Y_1'Y_1$ has a non-central Wishart distribution with non-centrality matrix $\Omega = \Sigma^{-1}M'M$. Without loss of generality we can set $\Sigma = I_p$. Using the dictionary

M:	A	m	n
JO:	n_1H	p	n_1

,

[M, Exer. 10.9] gives the joint density of the eigenvalues $W = \text{diag}(w_i)$ of $n_1 H$ as

$$p(W|\Omega) = c_{pn_1} \operatorname{etr}(-\tfrac{1}{2}\Omega) \operatorname{etr}(-\tfrac{1}{2}W)|W|^{(n_1-p-1)/2} v(W)\, {}_0F_1(\tfrac{1}{2}n_1; \tfrac{1}{4}\Omega, W).$$

The low rank assumption (JO4) posits $M = \sqrt{n_1}\varphi\sqrt{\theta}\psi'$ with $\varphi'\varphi = \psi'\Sigma^{-1}\psi = I_r$, so that with $\Sigma = I_p$, we have $\Omega = \Omega_\theta = n_1\psi\theta\psi'$. Note that $\mathbb{E}H = I + \psi\theta\psi'$, which explains the normalization chosen for M.

The eigenvalues Λ of H are related to the eigenvalues W of $n_1 H$ by $\Lambda = W/n_1$ and so

$$\begin{aligned} L(\theta;\Lambda) = \frac{p(\Lambda|\Omega_\theta)}{p(\Lambda|\Omega_0)} = \frac{p(W|\Omega_\theta)}{p(W|\Omega_0)} &= \operatorname{etr}\{-\tfrac{1}{2}n_1\theta\}\ {}_0F_1[\tfrac{1}{4}n_1\psi\theta\psi', n_1\Lambda] \\ &= \operatorname{etr}\{-\tfrac{1}{2}n_1\theta\}\ {}_0F_1[\tfrac{1}{4}n_1^2\psi\theta\psi', \Lambda], \end{aligned}$$

where we used (13).

REG. [J, eq. (73), citing Constantine (1963)], [M, Th. 10.4.2]. We are in the situation of REG_0, but with Σ unknown and estimated by an independent Wishart matrix $n_2 E \sim W_p(n_2, \Sigma)$. [M, Th. 10.4.2] gives the joint density of the eigenvalues F of equation (15). Using the dictionary

M:	A	B	m	r	$n-p$	Σ	Ω
JO:	n_1H	n_2E	p	n_1	n_2	I_p	Ω

,

this may be written as

$$p(F|\Omega) = c_{pn_1n_2} \operatorname{etr}(-\tfrac{1}{2}\Omega) w(F)\, {}_1F_1(\tfrac{1}{2}n, \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega, F(I+F)^{-1}).$$

where $w(F) = |F|^{(n_1-p-1)/2}|I+F|^{-(n_1+n_2)/2} v(F)$ does not depend on θ.

As for SigD, we make the transformation $\Lambda = (n_2/n_1)F(I+F)^{-1}$ to the generalized eigenvalues of (JO3). So, as in previous cases,

$$\begin{aligned} L(\theta;\Lambda) = \frac{p(\Lambda|\Omega_\theta)}{p(\Lambda|\Omega_0)} = \frac{p(F|\Omega_\theta)}{p(F|\Omega_0)} &= \operatorname{etr}\{-\tfrac{1}{2}n_1\theta\}\ {}_1F_1[\tfrac{1}{2}n_1\psi\theta\psi', (n_1/n_2)\Lambda] \\ &= \operatorname{etr}\{-\tfrac{1}{2}n_1\theta\}\ {}_1F_1[\tfrac{1}{2}(n_1^2/n_2)\psi\theta\psi', \Lambda]. \end{aligned}$$

CCA. [J, eq. (76), citing Constantine (1963)], [M, Th. 11.3.2]. We recall some of the steps from [M, Th. 11.2.6], borrowing some text from Johnstone and Nadler (2015). The canonical correlation problem is invariant under change of basis for each of the two sets of variables, e.g. [M, Th. 11.2.2]. We may therefore assume that the matrix Σ takes the canonical form

$$\Sigma = \begin{pmatrix} I_p & \tilde{P} \\ \tilde{P}' & I_{n_1} \end{pmatrix}, \quad \tilde{P} = [P\ 0], \quad P = \text{diag}(\rho_1, \ldots, \rho_r, 0, \ldots, 0)$$

where $\tilde{P}$ is $p \times n_1$ and the matrix P is of size $p \times p$ with r non-zero population canonical correlations $\rho_1, ...\rho_r$. Furthermore, in this new basis, we decompose the sample covariance matrix as follows,

$$nS = \begin{pmatrix} X'X & X'Y \\ Y'X & Y'Y \end{pmatrix} \tag{16}$$

where the columns of the $n \times p$ matrix X contain the first p variables of the $n \equiv n_1 + n_2$ samples, now assumed to have mean 0, represented in the transformed basis. Similarly, the columns of

$n \times n_1$ matrix Y contain the remaining n_1 variables. For future use, we note that the matrix $Y'Y \sim W_{n_1}(n, I_{n_1})$.

The squared canonical correlations $\{r_i^2\}$ are the eigenvalues of $S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx}$. According to [M, Th. 11.3.2], the joint density of $R^2 = \text{diag}(r_1^2, \ldots, r_p^2)$ is given by

$$p(R^2|P^2) = c_{pn_1n_2} |I_p - P^2|^{n/2} w(R^2) \, {}_2F_1(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}n_1; P^2, R^2),$$

where $w(R^2) = |R^2|^{(n_1-p-1)/2} |I_p - R^2|^{(n_2-p-1)/2} v(R^2)$ does not depend on P^2. Below, we abbreviate the hypergeometric function as ${}_2F_1[P^2, R^2]$ since the parameters $(\frac{1}{2}n, \frac{1}{2}n; \frac{1}{2}n_1)$ don't change.

If we set $P_Y = Y(Y'Y)^{-1}Y'$ the canonical correlations r_i^2 can be rewritten as the roots of $\det(r^2 X'X - X'P_Y X) = 0$. Now set $n_1 H = X'P_Y X$ and $n_2 E = X'(I - P_Y)X$: the previous equation becomes

$$\det(n_1 H - r^2(n_1 H + n_2 E)) = 0. \tag{17}$$

We now recall a standard partitioned Wishart argument. Conditional on Y, matrix X is Gaussian with independent rows, and mean and covariance matrices

$$\begin{aligned} M(Y) &= Y \Sigma_{yy}^{-1} \Sigma_{yx} = Y \tilde{P}' \\ \Sigma_{xx \cdot y} &= \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} = I - P^2. \end{aligned}$$

Conditional on Y, and using Cochran's theorem, the matrices

$$n_1 H \sim W_p(n, \Sigma_{xx \cdot y}, \Phi(Y)) \quad \text{and} \quad n_2 E \sim W_p(n_2, \Sigma_{xx \cdot y})$$

are independent, where the noncentrality matrix

$$\Phi(Y) = \Sigma_{xx \cdot y}^{-1} M(Y)' M(Y).$$

The generalized eigenvalues λ_i of (JO3) are related to the canonical correlations r_i^2, the generalized eigenvalues of (17), by $\lambda_i = (n_2/n_1) r_i^2$. Thus we obtain the interpretation of the roots of (JO3) in terms of a pair of matrices $n_1 H$ and $n_2 E$ which are conditionally independently Wishart given (part of the data) Y. Further, as for the previous cases, we can write the likelihood ratio as

$$\begin{aligned} \frac{p(\Lambda|P^2)}{p(\Lambda|0)} = \frac{p(R^2|P^2)}{p(R^2|0)} &= |I_p - P^2|^{n/2} \, {}_2F_1[P^2, R^2] \\ &= |I_p - P^2|^{n/2} \, {}_2F_1[(n_1/n_2)P^2, \Lambda]. \end{aligned}$$

Now in our rank r setting, $P^2 = \sum_1^r \rho_i^2 e_i e_i'$ with $\rho_i^2 = n_1\theta_i/(n_1\theta_i + n_1 + n_2)$. From the previous display we obtain, after setting $\psi = [I_r \; 0_{r \times (p-r)}]'$,

$$L(\theta, \Lambda) = \frac{p(\Lambda|P^2)}{p(\Lambda|0)} = |I_r + n_1\theta/n|^{-n/2} \, {}_2F_1[n_1^2 \psi \theta (n_2^2 I_r + n_2 n_1 (I_r + \theta))^{-1} \psi', \Lambda].$$

4. Contour integral representation.

4.1. *Derivations for Table JO4.* In this subsection, we obtain decomposition (JO12)

$$\mathcal{A} \equiv \frac{\Gamma(s+1)\,\alpha(\theta)\,q_s}{\sqrt{\pi p}\Psi_{11}^s} = \exp\{-(p/2)\,f_{\rm c}\}\,g_{\rm c},$$

where $s = p/2 - 1$, and $g_{\rm c}$ and $f_{\rm c}$ remain bounded as $\mathbf{n}, p \to_\gamma \infty$, for SMD, PCA, SigD, RED_0, REG, and CCA. The values of $g_{\rm c}$ and $f_{\rm c}$ for the different cases are given in Table JO4.

Structure of the prefactor $\mathcal{A}$. Let us rewrite

$$\mathcal{A} = \alpha(\theta)\frac{\Gamma(s+1)}{(p/2)^s\sqrt{\pi p}}\left[\frac{p/2}{\Psi_{11}}\right]^s q_s \tag{18}$$

as a product of terms $\mathcal{A}_k = g_k e^{-(p/2)f_k}(1+o(1))$ where f_k, g_k depend only on (c_1, c_2, θ), p, and q, see (24) below. The idea is to show the dependence on p, q. Referring to Table JO2, we have $a_j = n/2$ and $b_j = n_1/2$ whenever they are present, and so

$$q_s = \left[\frac{\Gamma(n/2-s)}{\Gamma(n/2)}\right]^{\mathsf{p}}\left[\frac{\Gamma(n_1/2)}{\Gamma(n_1/2-s)}\right]^{\mathsf{q}}. \tag{19}$$

Table JO2 also shows that

$$\alpha(\theta) = A(\theta)^{-n_1/2}, \qquad \frac{p/2}{\Psi_{11}} = \frac{p}{n_1}\frac{1}{B(\theta)}\frac{(n_2/2)^{\mathsf{p}}}{(n_1/2)^{\mathsf{q}}} \tag{20}$$

where $A(\theta)$ and $B(\theta)$ depend on the particular case in James' classification. This dependence is shown in Table 1 below.

TABLE 1
Terms $A(\theta), B(\theta)$ in the prefactor ${}_{\mathsf{p}}\mathcal{A}_{\mathsf{q}}$, formula (24).

Case	${}_{\mathsf{p}}F_{\mathsf{q}}$	$A(\theta)$	$B(\theta)$
SMD*	${}_0F_0$	$e^{\theta^2/2}$	θ
PCA	${}_0F_0$	$1+\theta$	$\theta/(1+\theta)$
SigD	${}_1F_0$	$1+\theta$	$\theta/(1+\theta)$
REG_0	${}_0F_1$	e^θ	θ
REG	${}_1F_1$	e^θ	θ
CCA	${}_2F_1$	$(1+n_1\theta/n)^{n/n_1}$	$\theta/l(\theta)$

(*) replace n_1 by p, $\quad l(\theta) = 1 + \frac{c_2}{c_1}(1+\theta)$

Combine like terms in (19) and (20) to get

$$\frac{(n_2/2)^{\mathsf{p}s}}{(n_1/2)^{\mathsf{q}s}}q_s = \left(\frac{n_2}{n}\right)^{\mathsf{p}s}\frac{\Theta^{\mathsf{p}}(n/2,p/2)}{\Theta^{\mathsf{q}}(n_1/2,p/2)}, \tag{21}$$

where we define

$$\Theta(N,M) = \frac{N^{M-1}\Gamma(N-M+1)}{\Gamma(N)} \sim e^M\left(1-\frac{M}{N}\right)^{N-M+1/2}. \tag{22}$$

[This is verified at the end of this section.] Finally, define

$$E(M) = \frac{\Gamma(M)}{M^{M-1}\sqrt{2\pi M}} \sim e^{-M}. \tag{23}$$

Combining (18), (20)–(22), we obtain the desired form

$$\mathcal{A} = {}_{\mathsf{p}}\mathcal{A}_{\mathsf{q}} = E(p/2)A(\theta)^{-n_1/2}B(\theta)^{-s}\left(\frac{p}{n_1}\right)^s\left(\frac{n_2}{n}\right)^{\mathsf{p}s}\frac{\Theta^{\mathsf{p}}(n/2,p/2)}{\Theta^{\mathsf{q}}(n_1/2,p/2)}. \tag{24}$$

Each factor in this product is easily cast in the form $g_k e^{-(p/2)f_k}(1+o(1))$, with the resulting values of f_k and g_k shown in Table 2. When needed, we factorize $g_k = \check{g}_k\tilde{g}_k$ to show the leading term $\check{g}_k$ and the term $\tilde{g}_k = 1+o(1)$, with the specific dependence of the $o(1)$ term (which comes from the error bound in Stirling's formula) shown in the final column of Table 2.

TABLE 2

Form of each term in (24), when expressed as $g_k e^{-(p/2)f_k}$, with $g_k = \check{g}_k\tilde{g}_k$. Here ϑ_m denotes a term that is $O(m^{-1})$.

	f_k	$\check{g}_k$	$\tilde{g}_k$
$E(p/2)$	1	1	$1+\vartheta_p$
$A(\theta)^{-n_1/2}$	$c_1^{-1}\log A(\theta)$	1	1
$B(\theta)^{-s}$	$\log B(\theta)$	$B(\theta)$	1
$(p/n_1)^s$	$-\log c_1$	$1/c_1$	1
$\left(\frac{n_2}{n}\right)^s$	$\log\left(1+\frac{c_2}{c_1}\right)$	$1+\frac{c_2}{c_1}$	1
$\Theta(n/2,p/2)$	$-1-\frac{r^2}{c_1c_2}\log\frac{r^2}{c_1+c_2}$	$\frac{r}{(c_1+c_2)^{1/2}}$	$1+\vartheta_n+\vartheta_{n-p}$
$\Theta^{-1}(n_1/2,p/2)$	$1+\frac{1-c_1}{c_1}\log(1-c_1)$	$(1-c_1)^{-1/2}$	$1+\vartheta_{n_1}+\vartheta_{n_1-p}$

TABLE 3

Table JO4. Values of f_c and $\check{g}_c = g_c/(1+o(1))$ for the different cases. The terms $o(1)$ do not depend on θ and converge to zero as $\mathbf{n}, p \to_\gamma \infty$. In the table, $l(\theta) = 1+(1+\theta)c_2/c_1$ and $r^2 = c_1+c_2-c_1c_2$.

Case	f_c	$\check{g}_c = g_c/(1+o(1))$
SMD	$1+\theta^2/2+\log\theta$	θ
PCA	$1+\frac{1-c_1}{c_1}\log(1+\theta)+\log\frac{\theta}{c_1}$	$\theta(1+\theta)^{-1}c_1^{-1}$
SigD	$f_c^{\mathrm{PCA}}+f_{10}$	$\check{g}_c^{\mathrm{PCA}}\check{g}_{10}$
REG$_0$	$1+\frac{\theta+c_1}{c_1}+\log\frac{\theta}{c_1}+\frac{1-c_1}{c_1}\log(1-c_1)$	$\theta c_1^{-1}(1-c_1)^{-1/2}$
REG	$f_c^{\mathrm{REG}_0}+f_{10}$	$\check{g}_c^{\mathrm{REG}_0}\check{g}_{10}$
CCA	$f_c^{\mathrm{REG}}+f_{21}$	$\check{g}_c^{\mathrm{REG}}\check{g}_{10}/l(\theta)$

$$f_{10} = -1-\frac{r^2}{c_1c_2}\log\frac{r^2}{c_1+c_2}+\log\frac{c_1+c_2}{c_1} \qquad \check{g}_{10} = c_1^{-1}r(c_1+c_2)^{1/2}$$

$$f_{21} = -1-\frac{\theta}{c_1}-\frac{r^2}{c_1c_2}\log\frac{r^2}{c_1l(\theta)}$$

Verification of Table JO4. We write $g_c = \breve{g}_c(1+o(1))$ and $g_{10} = \breve{g}_{10}(1+o(1))$ when we seek to be explicit about the leading term. The $o(1)$ term differs from row to row, but depends only on p, n_1, n_2 (and not θ). The explicit dependence can be constructed from the rows of Table 2, from which it is seen in fact always to be $O(m^{-1})$, where $m = \min(p, n_1 - p)$.

The lines for SMD, PCA and REG_0 in Table JO4 – reproduced here as Table 3 below – are immediately verified from Table 2. Next, we consider ratios in which the p index decreases by one from numerator to denominator. We then have from (24)

$$\frac{\mathcal{A}^{\text{SigD}}}{\mathcal{A}^{\text{PCA}}} = \frac{{}_1\mathcal{A}_0}{{}_0\mathcal{A}_0} = \frac{\mathcal{A}^{\text{REG}}}{\mathcal{A}^{\text{REG}_0}} = \frac{{}_1\mathcal{A}_1}{{}_0\mathcal{A}_1} = \left(\frac{n_2}{n}\right)^s \Theta(n/2, p/2) = g_{10} e^{-(p/2) f_{10}}.$$

Referring to Table 2, we recover the terms f_{10} and g_{10} and hence the lines for SigD and REG in Table JO4. For future reference, it is useful to decompose

$$\begin{aligned} &f_{10} = k_1 + k_0, \\ (25) \qquad &k_1 = -1 - \frac{r^2}{c_1 c_2}\log r^2, \qquad k_0 = \frac{r^2}{c_1 c_2}\log(c_1+c_2) + \log\frac{c_1+c_2}{c_1} \end{aligned}$$

Using (24) we have, in an obvious notation,

$$\begin{aligned} \frac{\mathcal{A}^{\text{CCA}}}{\mathcal{A}^{\text{REG}}} = \frac{{}_2\mathcal{A}_1}{{}_1\mathcal{A}_1} &= \left(\frac{A^C}{A^R}\right)^{-n_1/2}\left(\frac{B^C}{B^R}\right)^{-s} \cdot \left(\frac{n_2}{n}\right)^s \Theta(n/2, p/2) = \mathcal{R} \cdot \frac{\mathcal{A}^{\text{REG}}}{\mathcal{A}^{\text{REG}_0}}, \\ &= g_{21} e^{-(p/2) f_{21}}. \end{aligned}$$

and referring to Table 2, $\mathcal{R} = l^{-1}(\theta)\exp\{-(p/2) f_{20}\}$, where

$$\begin{aligned} (26) \qquad f_{20} = \frac{n_1}{p}\log\frac{A^C}{A^R} + \log\frac{B^C}{B^R} &= \frac{c_1+c_2}{c_1 c_2}\log\frac{c_1 l(\theta)}{c_1+c_2} - \frac{\theta}{c_1} - \log l(\theta) \\ &= k_2 - k_0, \end{aligned}$$

and

$$(27) \qquad k_2 = -\frac{\theta}{c_1} + \frac{r^2}{c_1 c_2}\log c_1 l(\theta).$$

This establishes the CCA line of Table JO4 after we note that

$$(28) \qquad f_{21} = f_{20} + f_{10} = k_2 - k_0 + k_1 + k_0 = k_2 + k_1.$$

Verification of (22). Use Stirling's formula (23) twice:

$$\begin{aligned} \Gamma(N) &\sim \sqrt{2\pi N}\; N^{N-1} e^{-N}, \qquad \text{and} \\ \Gamma(N-M+1) &= (N-M)\Gamma(N-M) \\ &\sim \sqrt{2\pi(N-M)}(N-M)^{N-M} e^{-N+M} \end{aligned}$$

to arrive at

$$\frac{N^{-1}\Gamma(N-M+1)}{\Gamma(N)} \sim \left(\frac{N-M}{N}\right)^{1/2} \frac{(N-M)^{N-M}}{N^N} e^M$$

and hence formula (22).

4.2. *Proof of Lemma JO2 (approximation to ${}_0F_1$).* By equation 9.6.47 in Abramowitz and Stegun (1964), we have

$$F_0 = \Gamma(m+1)\left(m^2\eta_0\right)^{-m/2} I_m\left(2m\eta_0^{1/2}\right), \tag{29}$$

where the principal branches of the fractional powers are taken, and $I_m(\cdot)$ is the modified Bessel function of the first kind. Using equation 9.7.7 in Abramowitz and Stegun (1964), we obtain

$$I_m\left(2m\eta_0^{1/2}\right) = \frac{\eta_0^{m/2}}{(1+4\eta_0)^{1/4}\sqrt{2\pi m}} e^{m(2t_0-1-\ln t_0)}(1+o(1)), \tag{30}$$

where $o(1) \to 0$ as $m \to \infty$ uniformly with respect to $\eta_0 \in \Omega_{0\delta}$ for any $\delta > 0$. Using (30) in (29), and invoking Stirling's approximation

$$\Gamma(m+1) = m^m e^{-m}\sqrt{2\pi m}(1+o(1)),$$

we obtain

$$F_0 = (1+4\eta_0)^{-1/4} e^{-m(-2t_0+2+\ln t_0)}(1+o(1)).$$

Since $1 - t_0 = -\eta_0/t_0$, we obtain $-2t_0 + 2 + \ln t_0 = \varphi_0(t_0)$ and thus,

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

4.3. *Proof of Lemma JO3 (approximation to ${}_1F_1$, ${}_2F_1$).* First, let us change variable of integration in

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

from t to $\tau = t\eta_j$. We obtain

$$F_j = \frac{C_m\eta_j^{-m}}{2\pi\mathrm{i}}\int_0^{(\eta_j+)} \exp\{-m\phi_j(\tau)\}\chi_j(\tau)\,\mathrm{d}\tau, \tag{31}$$

where

$$\phi_j(\tau) = \begin{cases} -\tau - \kappa\ln\tau + (\kappa-1)\ln(\tau-\eta_j) & \text{for } j=1 \\ -\kappa\ln(\tau/(1-\tau)) + (\kappa-1)\ln(\tau-\eta_j) & \text{for } j=2 \end{cases} \tag{32}$$

and

$$\chi_j(\tau) = \begin{cases} (\tau-\eta_j)^{-1} & \text{for } j=1 \\ (\tau-\eta_j)^{-1}(1-\tau)^{-1} & \text{for } j=2 \end{cases}.$$

Note that, for $j = 2$, the contour in (31) does not encircle 1.

To obtain point-wise asymptotic approximation to (31), the method of the steepest descent (ascent) is very convenient. However, establishing the uniformity of the approximation requires the knowledge of details of the structure of the steepest descent paths. For example, this knowledge becomes essential when some of the steepest descent paths contain two saddle points. Unfortunately, for our problem, the steepest descent paths are relatively complicated. Therefore, we will consider very simple paths that are steep (but not the steepest) in a neighborhood of a saddle point. This strategy allows us to rigorously establish the uniformity for relatively large sets of parameters κ and η_j. A downside of this approach is that we need to explicitly characterize the behavior of $\phi_j(\tau)$ on the simple paths, which requires some relatively lengthy but elementary calculus.

We shall prove Lemma JO3 separately for $j=1$ (REG) and for $j=2$ (CCA). Therefore, we shall omit subscripts j from the notation below.

Proof of Lemma JO3 for REG.

Saddle points, REG. The saddle points satisfy

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\phi(\tau) = -1 - \frac{\kappa}{\tau} + \frac{\kappa-1}{\tau-\eta} = -\frac{\tau^2+(1-\eta)\tau-\kappa\eta}{\tau(\tau-\eta)} = 0.$$

There are two solutions to this equation

$$\tau_{\pm} = \tfrac{1}{2}\left\{\eta+2\kappa-1 \pm \sqrt{(\eta+2\kappa-1)^2 - 4\kappa(\kappa-1)}\right\} - \kappa, \tag{33}$$

where we choose the principal branch of the square root (cut along $(-\infty,0]$) when $\operatorname{Re}\eta \geq -2\kappa+1$, and the other branch when $\operatorname{Re}\eta < -2\kappa+1$. The following lemma collects facts about the behavior of τ_+ for various (κ,η). Suppose that $\kappa>1$ (which is certainly true if $0<p<\min\{n_1,n_2\}$). Let $\beta=\arg\eta$. Here and in what follows the principal branch of arg (cut along $(-\infty,0]$) is considered, unless stated otherwise.

LEMMA 2. *(i) If $\operatorname{Im}\eta>0$, then $0<\arg(\tau_+-\eta)<\beta$; if $\operatorname{Im}\eta<0$, then $\beta<\arg(\tau_+-\eta)<0$. For real $\eta>0$, τ_+ is real and $\tau_+>\eta$.*

(ii) For $\eta\in\mathbb{C}\setminus(-\infty,0]$, function $\operatorname{Re}\phi(\tau)$ is strictly increasing as τ moves away from τ_+ (in any direction) along the circle with center η and radius $|\tau_+-\eta|$ until it reaches a point B on the circle. If $\operatorname{Im}\eta>0$, then $-\pi\leq\arg(B-\eta)\leq\beta-\pi$. If $\operatorname{Im}\eta<0$, then $\pi+\beta\leq\arg(B-\eta)\leq\pi$. If $\eta>0$, then $B=2\eta-\tau_+$.

Proof: *(i)* For $\operatorname{Im}\eta>0$ and the branch of the square root chosen as described above, we have

$$\operatorname{Im}\sqrt{(\eta+2\kappa-1)^2-4\kappa(\kappa-1)} > \operatorname{Im}(\eta+2\kappa-1) = \operatorname{Im}\eta.$$

Since

$$2\operatorname{Im}(\tau_+-\eta) = -\operatorname{Im}\eta + \operatorname{Im}\sqrt{(\eta+2\kappa-1)^2-4\kappa(\kappa-1)},$$

we have $\operatorname{Im}(\tau_+-\eta)>0$. Therefore,

$$\text{if } \operatorname{Im}\eta>0, \text{ then } 0<\arg(\tau_+-\eta)<\pi. \tag{34}$$

Similarly, we can show that if $\operatorname{Im}\eta<0$, then $-\pi<\arg(\tau_+-\eta)<0$.

Now let $\rho=|\tau_+-\eta|$. Then, for $\tau=\eta+\rho e^{\mathrm{i}x}$ we have

$$\begin{aligned}\operatorname{Re}\phi(\tau) &= (\kappa-1)\ln\rho - \operatorname{Re}\eta - \rho\cos x \\ &\quad -\frac{\kappa}{2}\ln\left(\rho^2+|\eta|^2+2\rho|\eta|\cos(x-\beta)\right),\end{aligned}$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) = \rho\left\{\sin x + k_\eta(x)\sin(x-\beta)\right\}, \tag{35}$$

where

$$k_\eta(x) = \frac{\kappa|\eta|}{\rho^2+|\eta|^2+2\rho|\eta|\cos(x-\beta)} > 0, \tag{36}$$

unless $\cos(x-\beta) = -1$ and $\rho = |\eta|$, in which case $\tau = \eta + \rho e^{ix} = 0$ and $\frac{d}{dx}\operatorname{Re}\phi(\tau) \to -\infty$ as $x \downarrow \beta - \pi$ and $\frac{d}{dx}\operatorname{Re}\phi(\tau) \to +\infty$ as $x \uparrow \beta + \pi$.

For $\operatorname{Im}\eta > 0$, (35) implies that

$$\frac{d}{dx}\operatorname{Re}\phi(\tau) > 0 \text{ for } x \in [\beta, \pi], \text{ and} \tag{37}$$

$$\frac{d}{dx}\operatorname{Re}\phi(\tau) < 0 \text{ for } x \in [\beta-\pi, 0]. \tag{38}$$

But, since τ_+ is a saddle point of $\phi(\tau)$,

$$\frac{d}{dx}\operatorname{Re}\phi(\tau) = 0 \text{ for } x = \arg(\tau_+ - \eta).$$

Therefore, inequalities (34) and (37) guarantee that $\arg(\tau_+ - \eta) \in (0, \beta)$. Similarly, we can show that $\operatorname{Im}\eta < 0$ implies that $\arg(\tau_+ - \eta) \in (\beta, 0)$. The part of *(i)* that deals with real $\eta > 0$ holds by inspection.

(ii) Consider the case $\operatorname{Im}\eta > 0$. Let us show that there are no zeros of $\frac{d}{dx}\operatorname{Re}\phi(\tau)$ on $(0,\beta)$ other than $\arg(\tau_+ - \eta)$. First, suppose that $k_\eta(\beta/2) < 1$, where $k_\eta(\cdot)$ is as defined in (36). Then, since $k_\eta(x)$ is a decreasing function of $x \in (0,\beta)$, equation (35) implies that all zeros of $\frac{d}{dx}\operatorname{Re}\phi(\tau)$ on $x \in (0,\beta)$ must belong to $(0,\beta/2)$. Furthermore, at any zero x of $\frac{d}{dx}\operatorname{Re}\phi(\tau)$ we must have $k_\eta(x) < 1$.

Indeed, let $x = \beta/2 + y$. Then,

$$\sin x + k_\eta(x)\sin(x-\beta) = \sin(\beta/2 + y) - k_\eta(x)\sin(\beta/2 - y)$$

On the other hand,

$$\sin(\beta/2 + y) - \sin(\beta/2 - y) = 2\sin y \cos(\beta/2),$$

which is positive for $0 < y < \beta/2$ and negative for $0 > y > -\beta/2$. Therefore, $\sin x + k_\eta(x)\sin(x-\beta)$ (and $\frac{d}{dx}\operatorname{Re}\phi(\tau)$) is positive for $x \in (\beta/2, \beta)$, and it may equal zero for some $x \in (0, \beta/2)$ only if $k_\eta(x) < 1$.

If there are more than one zero for $x \in (0,\beta/2)$, then by the mean-value theorem there must exist $x_1 \in (0,\beta/2)$ such that, at $x = x_1$, $\frac{d^2}{dx^2}\operatorname{Re}\phi(\tau) \le 0$ and $\frac{d}{dx}\operatorname{Re}\phi(\tau) \ge 0$. The latter inequality and the fact that $\frac{d}{dx}\operatorname{Re}\phi(\tau) < 0$ at $x = 0$ implies that some zeros of $\frac{d}{dx}\operatorname{Re}\phi(\tau)$ must be less than or equal to x_1, and hence, $k_\eta(x_1) < 1$.

To summarize, if there are more than one zero of $\frac{d}{dx}\operatorname{Re}\phi(\tau)$ on $(0,\beta/2)$, we must have

$$\rho\left\{\cos x_1 + k_\eta(x_1)\cos(x_1-\beta) + k'_\eta(x_1)\sin(x_1-\beta)\right\} \le 0 \tag{39}$$

for some $x_1 \in (0,\beta/2)$ with $k_\eta(x_1) < 1$. Since

$$k'_\eta(x_1)\sin(x_1-\beta) > 0 \text{and } (1 - k_\eta(x_1))\cos x_1 > 0,$$

we must have

$$\cos x_1 + \cos(x_1 - \beta) < 0.$$

Therefore,

$$2\cos(x_1 - \beta/2)\cos(\beta/2) < 0, \tag{40}$$

which is impossible for $x_1 \in (0,\beta)$ and $0 < \beta < \pi$.

Now, suppose that $k_\eta(\beta/2) > 1$. Then, all zeros of $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ on $x \in (0,\beta)$ belong to $(\beta/2,\beta)$. If there are more than one zero, there must exist $x_1 \in (\beta/2,\beta)$ such that $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\operatorname{Re}\phi(\tau) \leq 0$ at $x = x_1$ with $k_\eta(x_1) > 1$. That is, (39) holds. Since

$$k_\eta'(x_1)\sin(x_1-\beta) > 0,\ \ \cos(x_1-\beta) > 0,\ \text{and}\ k_\eta(x_1) > 1,$$

we still must have (40), which is impossible.

Finally, if $k_\eta(\beta/2) = 1$ then, since $k_\eta(x)$ is decreasing, (35) implies that there is only one zero of $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ on $x \in (0,\beta)$, which is $x = \beta/2$. To summarize, we have shown that

$$x_+ \equiv \arg(\tau_+ - \eta)$$

is the only zero of $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ on $(0,\beta)$. Similar arguments show that there exists only one zero, say x_-, of $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ on $(-\pi,\beta-\pi)$. (If $|\eta| = \rho$ so that $\operatorname{Re}\phi(\tau)$ is singular at $x = \beta-\pi$ with $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) \to -\infty$ as $x \downarrow \beta-\pi$ and $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) \to +\infty$ as $x \uparrow \beta+\pi$, we formally define $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ at $x = \beta-\pi$ as zero).

We will set

$$B = \eta + \rho\exp\{\mathrm{i}x_-\}.$$

The uniqueness of the zeros of $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ on $(0,\beta)$ and on $(-\pi,\beta-\pi)$, and inequalities (37,38) imply *(ii)* for the situation where $\operatorname{Im}\eta > 0$. The analysis for the cases with $\operatorname{Im}\eta < 0$ is similar to the above, and we omit it.

It remains to note that for real η such that $\eta > 0$, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) = \rho\{1 + k_\eta(x)\}\sin x$$

which implies the validity of *(ii)* for $\eta > 0$. □

Contours of steep descent, REG. We shall choose the contour of integration in (31) so that it passes through τ_+, and $\operatorname{Re}\phi(\tau)$ increases as τ moves away from τ_+ along the contour, at least in a neighborhood of τ_+. Such contours are called contours of steep descent (of $-\operatorname{Re}\phi(\tau)$). The contour consists of a circle with center η and radius $\rho = |\tau_+ - \eta|$ (which, in what follows, we refer to as the circle) and two overlapping straight segments of opposite orientations.

We consider four situations. The first and the second ones correspond to $\operatorname{Re}\eta > 0$ and to $\rho < |\eta|$ and $\rho \geq |\eta|$, respectively. The third and the fourth ones correspond to $\operatorname{Re}\eta \leq 0$ and to $\rho < |\eta|$ and $\rho \geq |\eta|$, respectively. In situations 1, 3, and 4, the two straight segments of opposite orientation connect zero and the point A where the circle is intersected by a half-line that starts at η and passes through zero. In situation 2, the point A is the intersection point of the circle and a half-line that starts at τ_- and passes through zero. Figure 1 illustrates the choice of the contour. The points B on the circles are as defined in Lemma 2.

Let us show that in situation 2, that is when $\operatorname{Re}\eta > 0$ and $\rho > |\eta|$, the circle intersects the straight segment $[\tau_-, 0)$, as shown in Figure 1. Indeed, by definition of $\tau_\pm$ we have

$$-(\tau_- - \eta) = (\tau_+ - \eta) + \eta + 1. \tag{41}$$

Since, by Lemma 2 (i), $\operatorname{Im}(\tau_+ - \eta)$ has the same sign as $\operatorname{Im}\eta$ and $\operatorname{Re}(\tau_+ - \eta) \geq 0$, and since $\operatorname{Re}(\eta+1) > 0$ and $\operatorname{Im}(\eta+1) = \operatorname{Im}\eta$, we have

$$|\operatorname{Re}\{-(\tau_- - \eta)\}| > |\operatorname{Re}(\tau_+ - \eta)|\ \text{and}\ |\operatorname{Im}\{-(\tau_- - \eta)\}| \geq |\operatorname{Im}(\tau_+ - \eta)|,$$

which implies that the circle must intersect the straight segment $[\tau_-, 0)$.

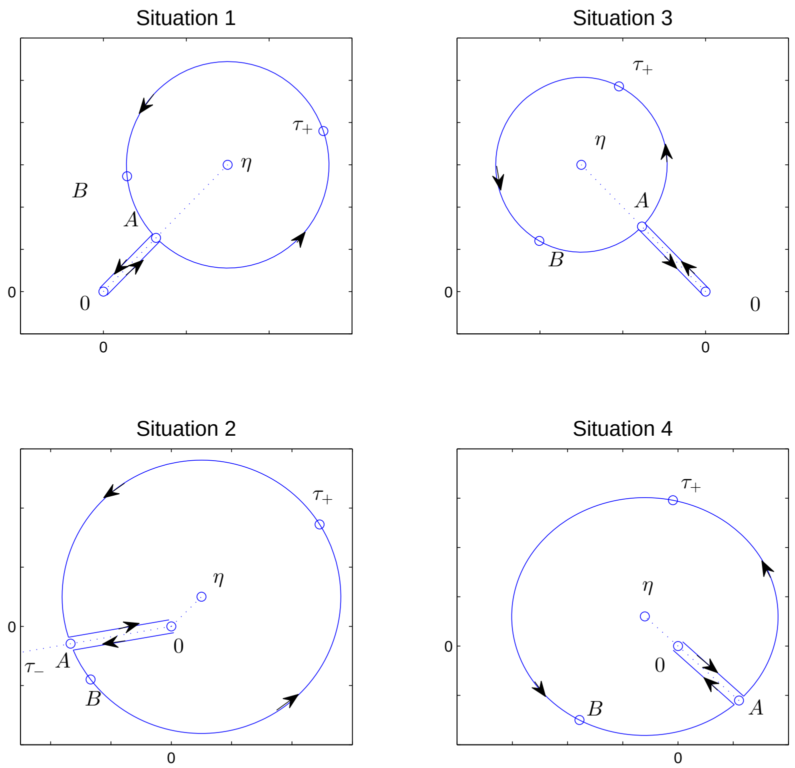

FIG 1. *Contours of steep descent, $j = 1$.*

We shall split the contour, which we shall call $\mathcal{K}$, in three parts. In situations 1, 3, and 4, the splitting is

$$\mathcal{K} = \mathcal{K}_{[0,A]} + \mathcal{K}_{[A,\tau_+,B]} + \mathcal{K}_{[B,A,0]}. \tag{42}$$

This decomposition assumes that $\operatorname{Im}\eta \geq 0$. If the sign of $\operatorname{Im}\eta$ changes to the negative, so that $\eta \longmapsto \bar{\eta}$, then $\mathcal{K}$ is transformed to its complex conjugate, and the orientation of such a complex conjugate must be changed to the opposite one. The decomposition then becomes

$$\mathcal{K} = \mathcal{K}_{[0,A,B]} + \mathcal{K}_{[B,\tau_+,A]} + \mathcal{K}_{[A,0]}. \tag{43}$$

In situation 2, when $\operatorname{Im}\eta \geq 0$, the splitting is (43) because $\arg(B - \eta) \geq \arg(A - \eta)$. (We will verify the latter inequality shortly.) In what follows, we consider only the case $\operatorname{Im}\eta \geq 0$. The complex conjugate case is analyzed similarly, and we omit details of such an analysis.

As follows from the proof of Lemma 2, $\operatorname{Re}\phi(\tau)$ is strictly increasing as τ is going along $\mathcal{K}_{[A,\tau_+,B]}$ away from τ_+. In other words, $\mathcal{K}_{[A,\tau_+,B]}$ is a contour of steep descent. Below, we shall use Lemma JO9 to analyze

$$\mathcal{I}_{[A,\tau_+,B]} = \int_{\mathcal{K}_{[A,\tau_+,B]}} e^{-m\phi(\tau)}\chi(\tau)\,\mathrm{d}\tau.$$

We shall then show that $\mathcal{I}_{[0,A]}$ and $\mathcal{I}_{[B,A,0]}$, which are defined similarly to $\mathcal{I}_{[A,\tau_+,B]}$, are asymptotically dominated by $\mathcal{I}_{[A,\tau_+,B]}$. However, before we embark on this agenda, let us show that $\arg(B - \eta) \geq \arg(A - \eta)$ as was claimed above.

As follows from Lemma 2, to see that the latter inequality holds, it is sufficient to verify that $\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Re}\,\phi\,(\tau)$ is positive at $\tau = A$. For such a verification, we will refer to Figure 2.

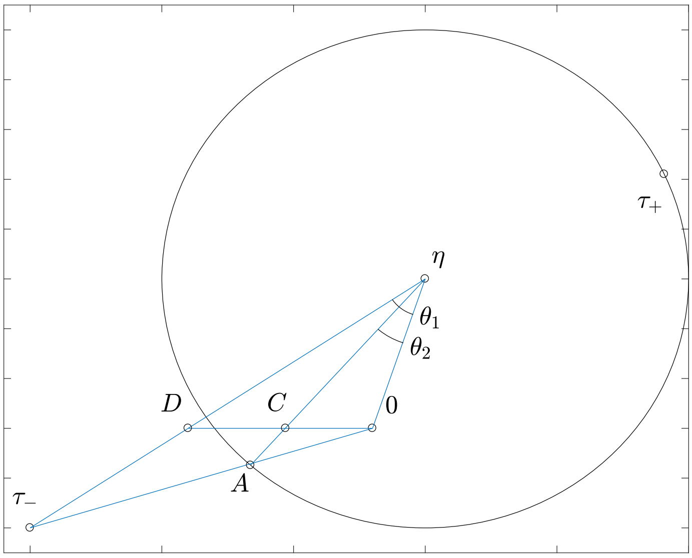

FIG 2. *An illustration to the argument that* $\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Re}\,\phi\,(\tau)$ *is positive at* $\tau = A$.

First, note that $\tau_+\tau_- = -\kappa\eta$ and $(\tau_+ - \eta)(\tau_- - \eta) = (1-\kappa)\eta$, where by assumption, $\kappa > 1$. The first of these equalities implies that $\arg\tau_- = -\pi + \arg\eta - \arg\tau_+ > -\pi$, so that point C on Figure 2 rightly belongs to $[A, \eta]$ (the line passing through $D, C, 0$ is a horizontal line). The second of the equalities implies that the angle $\angle D\eta 0 \equiv \theta_1$ equals $\arg(\tau_+ - \eta)$. Furthermore, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Re}\,\phi\,(\tau) = \rho\{\sin x + \frac{|\tau_+||\tau_-|}{|\tau|^2}\sin(x-\beta)\}.$$

For $\tau = A$, we have $x = \arg(A - \eta)$ and $\beta - x - \pi$ equals $\angle C\eta 0 \equiv \theta_2$. This implies

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Re}\,\phi\,(\tau) = \rho\{-\sin(\beta-\theta_2) + \frac{|\tau_+||\tau_-|}{|A|^2}\sin\theta_2\}. \tag{44}$$

For $\tau = \tau_+$, the derivative is zero, and hence

$$0 = \sin\theta_1 + \frac{|\tau_-|}{|\tau_+|}\sin(\theta_1 - \beta). \tag{45}$$

Now, by the law of sines applied to the triangle $\eta C 0$, we have

$$\frac{\sin\theta_2}{|C|} = \frac{\sin(\beta-\theta_2)}{|\eta|}. \tag{46}$$

Similarly, for the triangle $\eta D 0$, we have

$$\frac{\sin\theta_1}{|D|} = \frac{\sin(\beta-\theta_1)}{|\eta|}. \tag{47}$$

Combining (46) with (44), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)=\rho\sin\theta_2\left\{\frac{|\tau_+||\tau_-|}{|A|^2}-\frac{|\eta|}{|C|}\right\}. \tag{48}$$

Combining (47) with (45), we obtain

$$|\eta|=\frac{|\tau_+||D|}{|\tau_-|} \tag{49}$$

Using (49) in (48), we get

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)=\rho\sin\theta_2\frac{|\tau_+|}{|\tau_-|}\left\{\frac{|\tau_-|^2}{|A|^2}-\frac{|D|}{|C|}\right\}>0.\square$$

Saddle point approximation for $\mathcal{I}_{[A,\tau_+,B]}$, REG. We shall now derive a saddle point approximation to the integral $\mathcal{I}_{[A,\tau_+,B]}$ which is uniform with respect $(\kappa,\eta)\in\Omega_\delta$, where

$$\Omega_\delta=\left\{(\kappa,\eta)\in\mathbb{R}\times\mathbb{C}:\delta\le\kappa-1\le\delta^{-1},\ \operatorname{dist}(\eta,\mathbb{R}^-)\ge\delta,\ \text{and}\ |\eta|\le\delta^{-1}\right\}, \tag{50}$$

δ is an arbitrary fixed number that satisfies inequalities $0<\delta<1$, $\mathbb{R}^-=(-\infty,0)$ and, for any $A\subseteq\mathbb{C}$ and $B\subseteq\mathbb{C}$,

$$\operatorname{dist}(A,B)=\inf_{a\in A,b\in B}|a-b|.$$

Let us show that assumptions A0-A5 of Lemma JO9 hold. For this, we shall need the following lemma.

LEMMA 3. *The quantities $|\tau_+-\eta|$ and $|\tau_+|$ are bounded away from zero and infinity, uniformly with respect to $(\kappa,\eta)\in\Omega_\delta$.*

Proof: Note that $|\tau_+-\eta|$ and $|\tau_+|$ are continuous functions of $(\kappa,\eta)\in\Omega_\delta$. On the other hand, the definitions (33, 50) of τ_+ and Ω_δ together with Lemma 2 imply that $|\tau_+-\eta|\neq0$ and $|\tau_+|\neq0$ for any $(\kappa,\eta)\in\Omega_\delta$. The lemma follows from these observations and the compactness of Ω_δ.$\square$

Lemma 3 implies that the length of $\mathcal{K}_{[A,\tau_+,B]}$ is bounded uniformly with respect to $(\kappa,\eta)\in\Omega_\delta$. Further,

$$\sup_{\tau\in\mathcal{K}_{[A,\tau_+]}}|\tau-\tau_+|\ge|A-\tau_+|\ \text{and}\ \sup_{\tau\in\mathcal{K}_{[\tau_+,B]}}|\tau-\tau_+|\ge|B-\tau_+|$$

with $|A-\tau_+|$ and $|B-\tau_+|$ being continuous functions of $(\kappa,\eta)\in\Omega_\delta$, which are not equal to zero for any $(\kappa,\eta)\in\Omega_\delta$. Therefore, $|A-\tau_+|$ and $|B-\tau_+|$ are bounded away from zero, uniformly with respect to $(\kappa,\eta)\in\Omega_\delta$ and assumption A0 holds.

Assumptions A1, A2, A3 and A5 follow from Lemma 3. Finally, let τ_1 and τ_2 be the points of intersection of $\mathcal{K}$ with a circle with center at τ_+ and a sufficiently small fixed radius ε_1. The validity of Assumption A4 follows from the fact that $\operatorname{Re}(\phi(\tau_s)-\phi(\tau_+))$, $s=1,2$, are positive continuous functions of $(\kappa,\eta)\in\Omega_\delta$ (the positivity being a consequence of Lemma 2 (ii)) and $\operatorname{Im}(\phi(\tau_s)-\phi(\tau_+))$, $s=1,2$, are continuous functions of $(\kappa,\eta)\in\Omega_\delta$.

Since assumptions A0-A5 hold, we have by Lemma JO9

$$\mathcal{I}_{[A,\tau_+,B]}=2e^{-m\phi_0}\left[\sqrt{\pi}\frac{a_0}{m^{1/2}}+\frac{O(1)}{m^{3/2}}\right],$$

where $O(1)$ is uniform with respect to $(\kappa,\eta)\in\Omega_\delta$,

$$\phi_0 = -\tau_+ - \kappa\ln\tau_+ + (\kappa-1)\ln(\tau_+-\eta) \tag{51}$$

and

$$a_0 = \frac{(\tau_+-\eta)^{-1}}{\sqrt{2\kappa/\tau_+^2 - 2(\kappa-1)/(\tau_+-\eta)^2}} \tag{52}$$

with the branch of the square root chosen as described in Lemma JO9.

Precisely, let

$$\alpha = \pi/2 + \arg(\tau_+-\eta),$$

where the principal branch of $\arg(\cdot)$ is taken, and let

$$w = \arg\left(2\kappa/\tau_+^2 - 2(\kappa-1)/(\tau_+-\eta)^2\right), \tag{53}$$

where the branch of $\arg(\cdot)$ is chosen so that $|w+2\alpha|\le\pi/2$. Then

$$a_0 = \frac{e^{-\mathrm{i}w/2}(\tau_+-\eta)^{-1}}{\sqrt{\left|2\kappa/\tau_+^2 - 2(\kappa-1)/(\tau_+-\eta)^2\right|}}. \tag{54}$$

Analysis of $\mathcal{I}_{[0,A]}$ and $\mathcal{I}_{[B,A,0]}$, REG. Let us show that $\mathcal{I}_{[A,\tau_+,B]}$ asymptotically dominates $\mathcal{I}_{[0,A]}$ and $\mathcal{I}_{[B,A,0]}$ uniformly with respect to $(\kappa,\eta)\in\Omega_{1\delta}$, where

$$\Omega_{1\delta} = \Omega_\delta \cap \{(\kappa,\eta)\in\mathbb{R}\times\mathbb{C} : \operatorname{Re}\eta \ge -2\kappa+1\}.$$

It is sufficient to prove that there exists a positive constant S, such that, for τ on $\mathcal{K}_{[0,A]}$ and $\mathcal{K}_{[B,A,0]}$, we have $\operatorname{Re}\phi(\tau)\ge\operatorname{Re}\phi_0+S$, uniformly with respect to $(\kappa,\eta)\in\Omega_{1\delta}$. For concreteness, we again assume that $\operatorname{Im}\eta\ge0$. The complex conjugate case is very similar, and we omit its analysis.

Note that, by Lemma 2 (ii), for any $\tau\in\mathcal{K}_{[B,A]}$, $\operatorname{Re}\phi(\tau)\ge\operatorname{Re}\phi(A)$. Hence, it is sufficient to prove that $\operatorname{Re}\phi(\tau)\ge\operatorname{Re}\phi_0+S$ for τ from $\mathcal{K}_{[0,A]}$. Moreover, for situations 1, 2 and 4, shown on Figure 1, it is sufficient to establish the fact that $\operatorname{Re}\phi(A)\ge\operatorname{Re}\phi_0+S$. Indeed, let $\tau\in\mathcal{K}_{[0,A]}$, and let $x=|\tau|$. For situations 1 and 4, using the definition of $\phi(\tau)$ we have, respectively,

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) = -\cos\beta - \kappa/x - (\kappa-1)/(|\eta|-x) < 0,$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) = -\cos(\pi-\beta) - \kappa/x + (\kappa-1)/(|\eta|+x) < 0,$$

where $\beta=\arg\eta$. Therefore,

$$\operatorname{Re}\phi(A) = \inf_{\tau\in\mathcal{K}_{[0,A]}}\operatorname{Re}\phi(\tau). \tag{55}$$

For situation 2, we have

$$\operatorname{Re}\phi(\tau) = -x\cos(\arg\tau_-) - \kappa\ln x + \frac{\kappa-1}{2}\ln\left(x^2+|\eta|^2-2x|\eta|\cos\gamma\right),$$

where $\gamma = 2\pi + \arg\tau_- - \beta$, and thus,

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\operatorname{Re}\phi(\tau) &= \kappa/x^2 + (\kappa-1)\frac{-x^2 - |\eta|^2\cos(2\gamma) + 2x|\eta|\cos\gamma}{\left(x^2 + |\eta|^2 - 2x|\eta|\cos\gamma\right)^2} \\ &\geq \kappa/|\tau|^2 - (\kappa-1)/|\tau-\eta|^2.\end{aligned}$$

On the other hand, using the fact that $\tau_+\tau_- = -\kappa\eta$ and Lemma 2 (i), it is straightforward to verify that $\gamma > \pi/2$ and therefore, $|\tau|^2 < |\tau-\eta|^2$ for any $\tau \in \mathcal{K}_{[0,A]}$. Hence, $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\operatorname{Re}\phi(\tau) > 0$. But the first derivative of $\operatorname{Re}\phi(\tau)$ with respect to x must become positive for $x \to \infty$, negative for $x \to 0$, and zero for $x = |\tau_-|$, where x is any point on the ray connecting 0 with τ_-. Hence, $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau)$ must be negative for $\tau \in \mathcal{K}_{[0,A]}$, and (55) again holds.

For situation 3, let $\tau \in \mathcal{K}_{[0,A]}$. There are two possibilities. First, there exists τ_1 on the circle, such that $\operatorname{Re}\tau = \operatorname{Re}\tau_1$ and $|\operatorname{Im}\tau_1| \leq |\operatorname{Im}\eta|$. In such a case, $\operatorname{Re}\phi(\tau) \geq \operatorname{Re}\phi(\tau_1)$. Furthermore, by Lemma 2 (ii), $\operatorname{Re}\phi(\tau_1) > \operatorname{Re}\phi(C)$, where $C = \eta + |\tau_+ - \eta|$. Hence,

$$\operatorname{Re}\phi(\tau) > \operatorname{Re}\phi(C). \tag{56}$$

Second, $\operatorname{Re}\tau > \operatorname{Re}\eta + |\tau_+ - \eta|$. Assuming that $(\kappa,\eta) \in \Omega_{1\delta}$, the latter inequality implies that $\operatorname{Re}\tau \geq -\kappa$. Indeed, for $(\kappa,\eta) \in \Omega_{1\delta}$, the definition (33) of τ_+ implies that $\operatorname{Re}\tau_+ \geq -\kappa$. Therefore,

$$\operatorname{Re}\tau > \operatorname{Re}\eta + |\tau_+ - \eta| \geq \operatorname{Re}\tau_+ \geq -\kappa.$$

Let $x = |\tau|$, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) = -\cos\beta - \kappa/x - (\kappa-1)/(|\eta| - x),$$

and

$$\frac{1}{\cos\beta}\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) = -1 - \frac{\kappa}{\operatorname{Re}\tau} - \frac{\kappa-1}{\operatorname{Re}\eta - \operatorname{Re}\tau}.$$

But

$$-1 - \frac{\kappa}{\operatorname{Re}\tau} \geq 0 \text{ and } -\frac{\kappa-1}{\operatorname{Re}\eta - \operatorname{Re}\tau} > 0.$$

Therefore $\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Re}\phi(\tau) < 0$, and (56) holds. Note that in the analysis of situation 3 we used the assumption $(\kappa,\eta) \in \Omega_{1\delta}$, and in particular that $\operatorname{Re}\eta \geq -2\kappa + 1$. If the latter inequality is not satisfied, the minimum of $\operatorname{Re}\phi(\tau)$ on $\mathcal{K}$ can be achieved at some point on $\mathcal{K}_{[0,A]}$. This fact will be used later, in our proof of Theorem JO10.

It remains to show that, for some positive S,

$$\operatorname{Re}\phi(A) \geq \operatorname{Re}\phi_0 + S \tag{57}$$

uniformly with respect to $(\kappa,\eta) \in \Omega_{1\delta}$, and

$$\operatorname{Re}\phi(C) \geq \operatorname{Re}\phi_0 + S \tag{58}$$

uniformly with respect to $(\kappa,\eta) \in \tilde{\Omega}_{1\delta}$, where

$$\tilde{\Omega}_{1\delta} = \Omega_{1\delta} \cap \{\kappa \geq 1, \operatorname{Re}\eta \leq 0\}.$$

Inequality (58) follows from the fact that function $\operatorname{Re}\phi(C) - \operatorname{Re}\phi_0$ is continuous and positive for $(\kappa,\eta) \in \tilde{\Omega}_{1\delta}$ and from the compactness of $\tilde{\Omega}_{1\delta}$.

We cannot use a similar argument to establish inequality (57) because $\operatorname{Re}\phi(A) - \operatorname{Re}\phi_0$ is not a continuous function of $(\kappa,\eta) \in \Omega_{1\delta}$, as we may have $A = B = 0$ and $\operatorname{Re}\phi(A) = +\infty$ for some $(\kappa,\eta) \in \Omega_{1\delta}$. However, we can bound $\operatorname{Re}\phi(A) - \operatorname{Re}\phi_0$ from below by the minimum of two positive continuous functions $\operatorname{Re}\phi(\tau_1) - \operatorname{Re}\phi_0$ and $\operatorname{Re}\phi(\tau_2) - \operatorname{Re}\phi_0$, where τ_1 and τ_2 are the points of the intersection of the circle with center η and radius $|\tau_+ - \eta|$ and a circle with center τ_+ and a fixed radius, which is smaller than $|A - \tau_+|$, uniformly with respect to $(\kappa,\eta) \in \Omega_{1\delta}$. Therefore, there exists $S > 0$ such that (57) holds uniformly with respect to $(\kappa,\eta) \in \Omega_{1\delta}$.

Asymptotics in terms of φ and ψ, REG. The above analysis implies the following asymptotic representation

$$F_1 = \frac{C_m \eta^{-m}}{2\pi \mathrm{i}} 2e^{-m\phi_0} \left[\sqrt{\pi} \frac{e^{-\mathrm{i}w/2} (\tau_+ - \eta)^{-1}}{m^{1/2} \sqrt{\left|2\kappa/\tau_+^2 - 2(\kappa-1)/(\tau_+ - \eta)^2\right|}} + \frac{O(1)}{m^{3/2}} \right],$$

where $O(1)$ is uniform with respect to $(\kappa,\eta) \in \Omega_{1\delta}$. We would like to express this formula in terms of $t_1, \varphi_1(\cdot)$, and $\psi_1(\cdot)$. As follows from the definition of φ_1 (see equation (JO27)) and the fact that $\tau_+ = t_1\eta$,

$$\left|2\kappa/\tau_+^2 - 2(\kappa-1)/(\tau_+ - \eta)^2\right| = \left|2\varphi_1''(t_1)/\eta^2\right|,$$

Furthermore,

$$\varphi_1(t_1) = \phi_0 - \ln\eta,$$

and by (JO28)

$$(\tau_+ - \eta)^{-1} = \psi_1(t_1)\,\eta^{-1}.$$

Therefore, we have

$$F_1 = C_m e^{-m\varphi_1(t_1)} \left[\frac{e^{-\mathrm{i}w/2} e^{-\mathrm{i}\arg\eta}}{\mathrm{i}} \frac{\psi_1(t_1)}{\sqrt{|2\pi m \varphi_1''(t_1)|}} + \frac{O(1)}{m^{3/2}} \right].$$

On the other hand, by definition (53),

$$w = \arg(\varphi_1''(t_1)) - 2\arg\eta = \omega_1 - \pi - 2\arg\eta,$$

where ω_1 is as defined in equation (JO30). Hence,

$$\begin{aligned} F_1 &= C_m e^{-m\varphi_1(t_1)} e^{-\mathrm{i}\omega_1/2} \left[\frac{\psi_1(t_1)}{\sqrt{|2\pi m \varphi_1''(t_1)|}} + \frac{O(1)}{m^{3/2}} \right] \\ &= C_m \psi_1(t_1)\, e^{-\mathrm{i}\omega_1/2} \left|2\pi m \varphi_1''(t_1)\right|^{-1/2} \exp\{-m\varphi_1(t_1)\}\,(1 + o(1)). \end{aligned}$$

Proof of Lemma JO3 for CCA.

Saddle points, CCA. From equation (32) with $j = 2$, we see that the saddle points satisfy

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\phi(\tau) = -\frac{\kappa}{\tau} - \frac{\kappa}{1-\tau} + \frac{\kappa-1}{\tau-\eta} = \frac{\tau^2(\kappa-1) + \tau - \kappa\eta}{\tau(\tau-1)(\tau-\eta)} = 0.$$

There are two solutions to this equation

$$\tau_{\pm} = \frac{-1 \pm \sqrt{1 + 4\kappa(\kappa-1)\eta}}{2(\kappa-1)}, \tag{59}$$

where we choose the principal branch of the square root cut along $(-\infty, 0]$.

The following lemma collects facts about the behavior of τ_+ for various (κ, η). As usual, we assume that $\kappa > 1$. In addition, we assume that $\eta \notin (-\infty, \eta_*) \cup [1, \infty)$, where

$$\eta_* = -\frac{1}{4\kappa(\kappa-1)}.$$

Note that set $(-\infty, \eta_*) \cup [1, \infty)$ does not intersect with $\Omega_{2\delta}$ for any $\delta > 0$.

LEMMA 4. *(i)* $|\tau_+ - \eta| < |1 - \eta|$, *and* $\tau_+ = 0$ *if and only if* $\eta = 0$.

(ii) If $\operatorname{Im}\eta > 0$ *and* $\operatorname{Re}\tau_+ > 1/2$, *then* $0 < \operatorname{Im}\tau_+ < \operatorname{Im}\eta$. *If* $\operatorname{Im}\eta > 0$ *and* $\operatorname{Re}\tau_+ < 1/2$, *then* $\operatorname{Im}\tau_+ > \operatorname{Im}\eta$. *Similarly, if* $\operatorname{Im}\eta < 0$ *and* $\operatorname{Re}\tau_+ > 1/2$, *then* $0 > \operatorname{Im}\tau_+ > \operatorname{Im}\eta$. *If* $\operatorname{Im}\eta < 0$ *and* $\operatorname{Re}\tau_+ < 1/2$, *then* $\operatorname{Im}\tau_+ < \operatorname{Im}\eta$.

(iii) For $\eta \notin (-\infty, \eta_*) \cup [1, \infty)$, *function* $\operatorname{Re}\phi(\tau)$ *is strictly increasing as* τ *moves away from* τ_+ *(in any direction) along the circle with center* η *and radius* $|\tau_+ - \eta|$ *until it reaches a point* B *on the circle.*

Proof: *(i)* Let

$$-\eta_*^{-1}(\eta - \eta_*) = \rho^2 \exp\{\mathrm{i}2\theta\} \tag{60}$$

with $\theta \in (-\pi/2, \pi/2)$. Then

$$\tau_+ = \frac{-1 + \rho\exp\{\mathrm{i}\theta\}}{2(\kappa-1)} \tag{61}$$

and a direct calculation (we perform it using Maple's symbolic algebra software) shows that

$$\eta_*^{-2}\left(|\tau_+ - \eta|^2 - |1-\eta|^2\right) = -4\kappa(\kappa + \rho\cos\theta - 1)\left((2\kappa-1)^2 + \rho^2 - 2\rho(2\kappa-1)\cos\theta\right).$$

Since $\theta \in (-\pi/2, \pi/2)$ and $\kappa > 1$, the latter expression is less than zero. Further, equation (61) implies that $\tau_+ = 0$ if and only if $\theta = 0$ and $\rho = 1$. The latter two equalities are equivalent to $\eta = 0$.

(ii) From (61), we see that $\operatorname{Re}\tau_+ > 1/2$ if and only if

$$\rho\cos\theta > \kappa. \tag{62}$$

On the other hand,

$$-\eta_*^{-1}\operatorname{Im}(\tau_+ - \eta) = 2\rho\sin\theta(\kappa - \rho\cos\theta). \tag{63}$$

Combining (62) and (63), we obtain *(ii)*.

(iii) Recall that

$$\phi(\tau) = -\kappa\ln\frac{\tau}{1-\tau} + (\kappa-1)\ln(\tau-\eta).$$

Therefore, on the circle with center η and radius $|\tau_+ - \eta|$, $\operatorname{Re}\phi(\tau)$ equals $-\kappa\ln|\tau/(1-\tau)|$ plus a constant. Further, for $c > 0$ such that $c \neq 1$, the set of τ that satisfy equality $|\tau/(1-\tau)| = c$ is

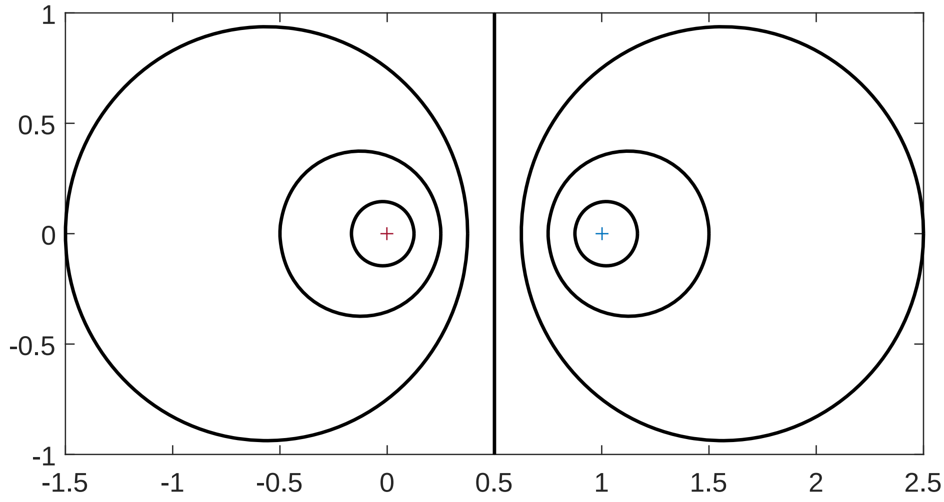

FIG 3. *Isolines of the function* $|\tau/(1-\tau)|$.

a circle with center $c^2/(c^2-1)$ and radius $c/|c^2-1|$. For $c=1$, $|\tau/(1-\tau)|=c$ along the line $\mathrm{Re}\,\tau=1/2$. Figure 3 shows the iso-lines of $|\tau/(1-\tau)|$. For $c<1$, the isolines are encircling 0, for $c>1$, they are encircling 1.

Since τ_+ is a critical point of $\mathrm{Re}\,\phi(\tau)$, the circle with center η and radius $|\tau_+-\eta|$ must have a common tangent with one of the isolines at $\tau=\tau_+$. Therefore, $\mathrm{Re}\,\phi(\tau)$ must be strictly monotone as τ moves away from τ_+ along the circle with center η and radius $|\tau_+-\eta|$ until it reaches a point B on the circle. Part *(ii)* of the lemma implies that $\mathrm{Re}\,\phi(\tau)$ is strictly increasing. □

Contours of steep descent, CCA. We shall choose the contour of integration in (31), which we shall call $\mathcal{K}$, so that it passes through τ_+, and $\mathrm{Re}\,\phi(\tau)$ increases as τ moves away from τ_+ along the contour, at least in a neighborhood of τ_+. The contour consists of a circle with center η and radius $r=|\tau_+-\eta|$, which, in what follows, we refer to as C_1, and two overlapping circular segments of opposite orientations, which we will refer to as C_2.

We consider four situations. The first and the second ones correspond to $r<|\eta|$ and to $\mathrm{Re}\,\eta<0$ and $\mathrm{Re}\,\eta\ge 0$, respectively. The third and the fourth ones correspond to $r\ge|\eta|$ and to $\mathrm{Re}\,\eta<0$ and $\mathrm{Re}\,\eta\ge 0$, respectively. Using (61), we obtain

$$\eta_*^{-2}\,|\tau_+-\eta|^2-\eta_*^{-2}\,|\eta|^2=4\kappa\left(\rho^2-2\rho\cos\theta+1\right)(\kappa-\rho\cos\theta-1)\,.$$

Therefore, situations 3 or 4 are realized whenever

$$\rho\cos\theta\le\kappa-1. \tag{64}$$

In particular, the corresponding τ_+ must be such that $\mathrm{Re}\,\tau_+<1/2$ (compare to (62)).

For situation 1 and 2, C_2 consists of a segment of the circle that passes through 0, 1, and η. The segment starts at the closest to 0 intersection of the latter circle with C_1 and ends at 0. It does not pass through 1 or η. For situation 3 and 4, C_2 consists of the segment of the circle with center at 1 and radius 1 that connects 0 with the point A of the intersection of this circle with C_1, and lies inside C_1. Out of the two intersection points we choose the one with the imaginary part of the opposite sign to that of $\mathrm{Im}\,\eta$. Figures 4, 5, 6, and 7 illustrate the choice of $\mathcal{K}$ for situations 1, 2, 3, and 4, respectively.

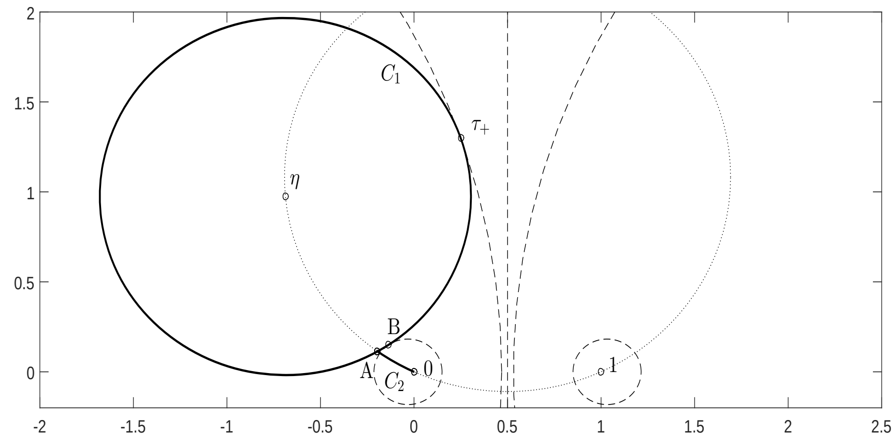

FIG 4. *Choice of contour $\mathcal{K}$ in situation 1. The contour is represented by the dark black circle and the circle segment ending at* 0*. The dashed lines are iso-lines of function* $|\tau/(\tau-1)|$.

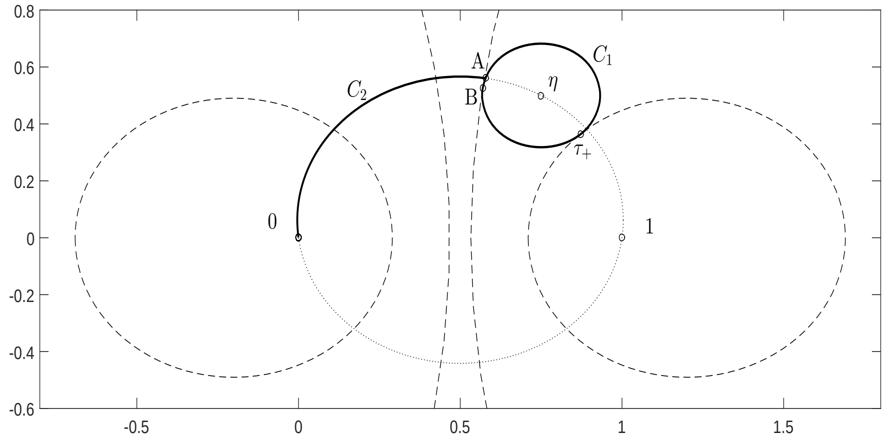

FIG 5. *Choice of contour $\mathcal{K}$ in situation 2. The contour is represented by the dark black circle and the circle segment ending at* 0*. The dashed lines are iso-lines of function* $|\tau/(\tau-1)|$.

We split the contour in three parts

$$\mathcal{K} = \mathcal{K}_{[0,A]} + \mathcal{K}_{[A,\tau_+,B]} + \mathcal{K}_{[B,A,0]}, \tag{65}$$

or

$$\mathcal{K} = \mathcal{K}_{[0,A,B]} + \mathcal{K}_{[B,\tau_+,A]} + \mathcal{K}_{[A,0]} \tag{66}$$

depending on whether moving counter-clockwise along C_1 from A to B reaches τ_+ or not. In the rest of this note, we shall refer to (65) for concreteness. Our arguments do not depend on the specific form of the splitting.

As follows from Lemma 4, $\operatorname{Re}\phi(\tau)$ is strictly increasing as τ is going along $\mathcal{K}_{[A,\tau_+,B]}$ away from τ_+. In other words, $\mathcal{K}_{[A,\tau_+,B]}$ is a contour of steep descent. Below, we shall use Lemma JO9 to analyze

$$\mathcal{I}_{[A,\tau_+,B]} = \int_{\mathcal{K}_{[A,\tau_+,B]}} e^{-m\phi(\tau)}\chi(\tau)\,\mathrm{d}\tau.$$

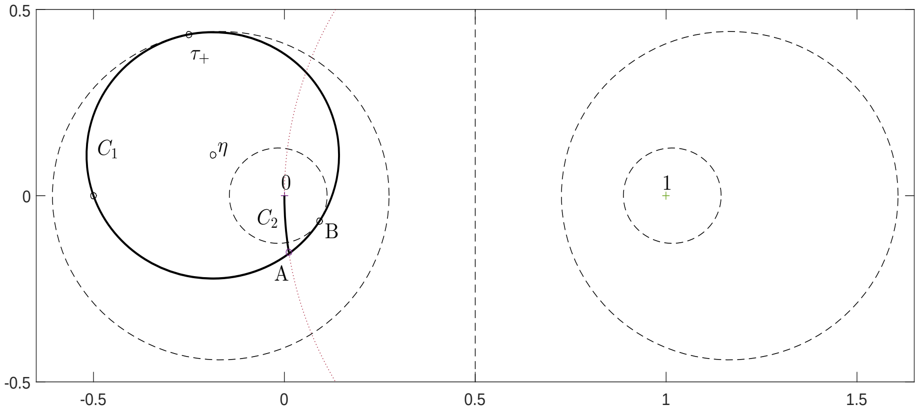

Fig 6. *Choice of contour $\mathcal{K}$ in situation 3. The contour is represented by the dark black circle and the circle segment ending at* 0*. The dashed lines are iso-lines of function* $|\tau/(\tau-1)|$.

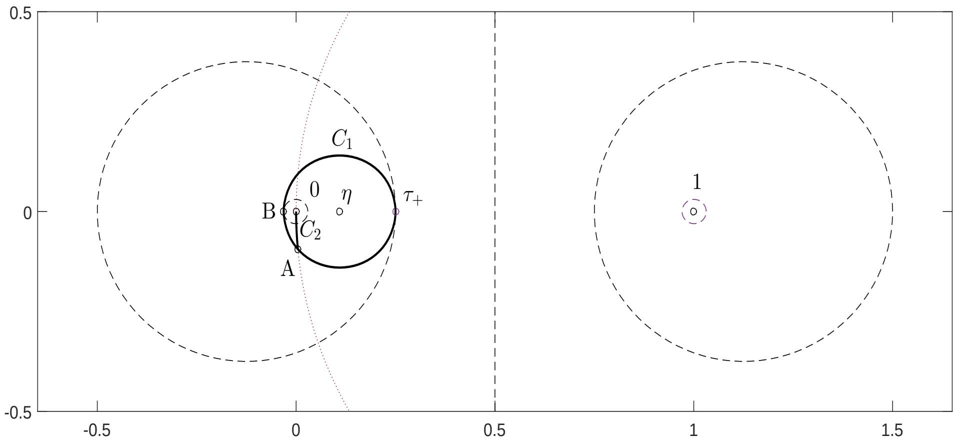

Fig 7. *Choice of contour $\mathcal{K}$ in situation 4. The contour is represented by the dark black circle and the circle segment ending at* 0*. The dashed lines are iso-lines of function* $|\tau/(\tau-1)|$.

We shall then show that $\mathcal{I}_{[0,A]}$ and $\mathcal{I}_{[B,A,0]}$, which are defined similarly to $\mathcal{I}_{[A,\tau_+,B]}$, are asymptotically dominated by $\mathcal{I}_{[A,\tau_+,B]}$.

Saddle point approximation for $\mathcal{I}_{[A,\tau_+,B]}$, CCA. We now derive a saddle point approximation to the integral $\mathcal{I}_{[A,\tau_+,B]}$ which is uniform with respect $(\kappa,\eta)\in\Omega_{2\delta}$, where

$$\Omega_{2\delta}=\left\{(\kappa,\eta):\delta\le\kappa-1\le\delta^{-1},\ \mathrm{dist}\,(\eta,\mathbb{R}\backslash[0,1])\ge\delta,\ \text{and }\ |\eta|\le\delta^{-1}\right\}, \tag{67}$$

and δ is an arbitrary fixed number that satisfies inequalities $0<\delta<1$. Let us verify assumptions A0-A5 of Lemma JO9. For this verification, we need the following lemma.

Lemma 5. *The quantities $|\tau_+-\eta|$ and $|\tau_+|$ are bounded away from zero and infinity, uniformly with respect to $(\kappa,\eta)\in\Omega_{2\delta}$.*

Proof: The lemma follows from Lemma 4 (i,ii), the fact that $\tau_+\ne\eta$ for $(\kappa,\eta)\in\Omega_{2\delta}$, and the compactness of $\Omega_{2\delta}$. □

Lemma 5 implies that the length of $\mathcal{I}_{[A,\tau_+,B]}$ is bounded uniformly with respect to $(\kappa,\eta)\in\Omega_{2\delta}$. Further,

$$\sup_{\tau\in\mathcal{K}_{[A,\tau_+]}}|\tau-\tau_+|\geq|A-\tau_+| \text{ and } \sup_{\tau\in\mathcal{K}_{[\tau_+,B]}}|\tau-\tau_+|\geq|B-\tau_+|,$$

where $|A-\tau_+|$ and $|B-\tau_+|$ are continuous functions of $(\kappa,\eta)\in\Omega_{2\delta}$, which are not equal to zero for any $(\kappa,\eta)\in\Omega_{2\delta}$. Therefore, $|A-\tau_+|$ and $|B-\tau_+|$ are bounded away from zero, uniformly with respect to $(\kappa,\eta)\in\Omega_{2\delta}$ and assumption A0 holds.

Assumptions A1, A2, A3 and A5 follow from Lemma 5. Finally, let τ_1 and τ_2 be the points of intersection of $\mathcal{K}$ with a circle with center at τ_+ and a sufficiently small fixed radius ε_1. The validity of Assumption A4 follows from the fact that $\operatorname{Re}(\phi(\tau_s)-\phi(\tau_+))$, $s=1,2$, are positive continuous functions of $(\kappa,\eta)\in\Omega_{2\delta}$ (the positivity being a consequence of Lemma 4 (iii)) and $\operatorname{Im}(\phi(\tau_s)-\phi(\tau_+))$, $s=1,2$, are continuous functions of $(\kappa,\eta)\in\Omega_{2\delta}$.

Since assumptions A0-A5 hold, by Lemma JO9, we have

$$\mathcal{I}_{[A,\tau_+,B]}=2e^{-m\phi_0}\left[\sqrt{\pi}\frac{a_0}{m^{1/2}}+\frac{O(1)}{m^{3/2}}\right],$$

where $O(1)$ is uniform with respect to $(\kappa,\eta)\in\Omega_{2\delta}$,

$$\phi_0=-\kappa\ln\frac{\tau_+}{1-\tau_+}+(\kappa-1)\ln(\tau_+-\eta) \tag{68}$$

and

$$a_0=\frac{(\tau_+-\eta)^{-1}(1-\tau_+)^{-1}}{\sqrt{2\kappa(1-2\tau_+)/\left((1-\tau_+)^2\tau_+^2\right)-2(\kappa-1)/(\tau_+-\eta)^2}} \tag{69}$$

with the branch of the square root chosen as described in Lemma JO9.

Precisely, let

$$\alpha=\pi/2+\arg(\tau_+-\eta),$$

where the principal branch of $\arg(\cdot)$ is taken, and let

$$w=\arg\left(2\kappa(1-2\tau_+)/\left((1-\tau_+)^2\tau_+^2\right)-2(\kappa-1)/(\tau_+-\eta)^2\right),$$

where the branch of $\arg(\cdot)$ is chosen so that

$$|w+2\alpha|\leq\pi/2.$$

Then

$$a_0=\frac{e^{-\mathrm{i}w/2}(\tau_+-\eta)^{-1}(1-\tau_+)^{-1}}{\sqrt{\left|2\kappa(1-2\tau_+)/\left((1-\tau_+)^2\tau_+^2\right)-2(\kappa-1)/(\tau_+-\eta)^2\right|}}. \tag{70}$$

Analysis of $\mathcal{I}_{[0,A]}$ and $\mathcal{I}_{[B,A,0]}$, CCA. Let us show that $\mathcal{I}_{[A,\tau_+,B]}$ asymptotically dominates $\mathcal{I}_{[0,A]}$ and $\mathcal{I}_{[B,A,0]}$ uniformly with respect to $(\kappa,\eta)\in\Omega_{2\delta}$. It is sufficient to prove that there exists a positive constant S, such that, for τ on $\mathcal{I}_{[0,A]}$ or on $\mathcal{I}_{[B,A,0]}$, we have $\operatorname{Re}\phi(\tau)\geq\operatorname{Re}\phi_0+S$, uniformly with respect to $(\kappa,\eta)\in\Omega_{2\delta}$.

Note that, by Lemma 4 (iii), for any $\tau \in \mathcal{K}_{[A,B]}$, $\operatorname{Re}\phi(\tau) \geq \operatorname{Re}\phi(A)$. Hence, it is sufficient to prove that $\operatorname{Re}\phi(\tau) \geq \operatorname{Re}\phi_0 + S$ for τ from $\mathcal{K}_{[0,A]}$. Moreover, it is sufficient to establish the fact that

$$\operatorname{Re}\phi(A) \geq \operatorname{Re}\phi_0 + S \tag{71}$$

uniformly with respect to $(\kappa, \eta) \in \Omega_{2\delta}$. It is because for any $\tau \in \mathcal{K}_{[0,A]}$, $\operatorname{Re}\phi(\tau) \geq \operatorname{Re}\phi(A)$.

Indeed, for situations 1 and 2 this property of $\operatorname{Re}\phi(\tau)$ follows from the fact that $|\tau/(1-\tau)|$ is strictly decreasing and $|\tau - \eta|$ is strictly increasing as τ moves along $\mathcal{K}_{[0,A]}$ away from A. For situation 3, we have

$$\begin{aligned} \operatorname{Re}\phi(\tau) - \operatorname{Re}\phi(A) &= -\kappa \log \frac{|\tau|}{|A|} + (\kappa - 1) \log \frac{|\tau - \eta|}{|A - \eta|} \\ &> -\kappa \log \frac{|\tau|}{|\tau - \eta|} + \kappa \log \frac{|A|}{|A - \eta|}, \end{aligned}$$

where the latter inequality holds because $|\tau - \eta| < |A - \eta|$. The iso-lines of function $|\tau| / |\tau - \eta|$ are similar to those shown on Figure 3 with the concentration points 0 and η instead of 0 and 1. As τ moves along $\mathcal{K}_{[0,A]}$ away from A, the isolines are crossed so that $|\tau| / |\tau - \eta|$ is decreasing. Therefore,

$$\operatorname{Re}\phi(\tau) - \operatorname{Re}\phi(A) \geq 0. \tag{72}$$

For situation 4, the analysis is more involved. We have the following lemma.

Lemma 6. *Inequality (72) holds for situation 4.*

Proof: The analysis is similar to that of situation 3. However, in contrast to situation 3, we cannot immediately claim that as τ moves along $\mathcal{K}_{[0,A]}$ away from A, the isolines of the function $|\tau| / |\tau - \eta|$ are crossed so that the function is decreasing. For this claim to be valid, we must verify that

$$|A| / |A - \eta| < 1, \tag{73}$$

so that A and 0 lie on the same side of the iso-line $|\tau| / |\tau - \eta| = 1$.

Let x be the point on C_1 where $|x| / |x - \eta| = 1$, such that $\operatorname{Im}(\eta - x)$ has the same sign as $\operatorname{Im}\eta$. To establish (73), it is sufficient to show that x lies inside the circle with center at 1 and radius 1 (circumference of which contains $\mathcal{K}_{[0,A]}$). That is, it is sufficient to show that $|1 - x|^2 < 1$.

For concreteness, let us focus on the case $\operatorname{Im}\eta > 0$. Then, we have

$$x = \eta/2 - \mathrm{i}\eta a, \text{ where } a = \sqrt{\left(r^2 - |\eta/2|^2\right) / |\eta|^2}$$

and r is the radius of C_1. A straightforward algebra shows that

$$|1 - x|^2 = r^2 + 1 - \operatorname{Re}\eta - 2a \operatorname{Im}\eta.$$

Furthermore, since $r^2 \geq |\eta|^2$, we have $a > \sqrt{3}/2 > 1$. Therefore, the inequality $|1 - x|^2 < 1$ would follow from the inequality $r^2 \leq |\eta| < \operatorname{Re}\eta + \operatorname{Im}\eta$. Let us now show that in situation 4, $r^2 \leq |\eta|$.

Let $z = \rho \exp\{\mathrm{i}\theta\}$, where ρ and θ are as in (60). Situation 4 imposes the following constraints on z: 1) $\operatorname{Re} z \geq 0$, 2) $\operatorname{Re} z \leq \kappa - 1$, 3) $(\operatorname{Re} z)^2 - (\operatorname{Im} z)^2 \geq 1$. The first one is equivalent to $\theta \in [-\pi/2, \pi/2]$, which must be true by definition of θ. The second one is equivalent to (64), and the last one ensures that $\operatorname{Re}\eta \geq 0$. We have

$$\tau_+ = \frac{-1+z}{2(\kappa-1)} \text{ and } \eta = \frac{-1+z^2}{4\kappa(\kappa-1)}.$$

Therefore,

$$r^2 \equiv |\tau_+ - \eta|^2 = \frac{|z-1|^2 |z-(2\kappa-1)|^2}{16\kappa^2(\kappa-1)^2}$$

and

$$|\eta| = \frac{|z-1||z+1|}{4\kappa(\kappa-1)}.$$

For z that satisfies the above three constraints, we must have $|z+1| > |z-1|$. Therefore, to establish inequality $r^2 \leq |\eta|$, it is sufficient to show that

$$\frac{|z-(2\kappa-1)|^2}{4\kappa(\kappa-1)} \leq 1.$$

The latter inequality is equivalent to

$$(\operatorname{Im} z)^2 + 1 \leq 2\operatorname{Re} z(2\kappa-1) - (\operatorname{Re} z)^2.$$

In view of the third constraint, it is sufficient to show that

$$2\operatorname{Re} z(2\kappa-1) - (\operatorname{Re} z)^2 \geq (\operatorname{Re} z)^2.$$

But the second constraint implies this inequality. The situation where $\operatorname{Im}\eta \leq 0$ is analyzed similarly. □

To summarize, in all the four situations we only need to show that (71) holds. Note that $\operatorname{Re}\phi(A) - \operatorname{Re}\phi_0$ is not a continuous function of $(\kappa,\eta) \in \Omega_{2\delta}$ because we may have $A = B = 0$ and $\operatorname{Re}\phi(A) = +\infty$ for some $(\kappa,\eta) \in \Omega_{2\delta}$. However, we can bound $\operatorname{Re}\phi(A) - \operatorname{Re}\phi_0$ from below by the minimum of two positive continuous functions $\operatorname{Re}\phi(\tau_1) - \operatorname{Re}\phi_0$ and $\operatorname{Re}\phi(\tau_2) - \operatorname{Re}\phi_0$, where τ_1 and τ_2 are points of the intersection of the circle with center η and radius $|\tau_+ - \eta|$ and a circle with center τ_+ and a fixed radius, which is smaller than $|A - \tau_+|$, uniformly with respect to $(\kappa,\eta) \in \Omega_{2\delta}$. Therefore, there exists $S > 0$ such that (71) holds uniformly with respect to $(\kappa,\eta) \in \Omega_{2\delta}$.

Asymptotics in terms of φ and ψ, CCA.

The above analysis implies the following asymptotic representation

$$F_2 = \frac{C_m\eta^{-m}}{2\pi\mathrm{i}} 2e^{-m\phi_0}\left[\frac{\sqrt{\pi}e^{-\mathrm{i}w/2}(\tau_+-\eta)^{-1}(1-\tau_+)^{-1}}{m^{1/2}\sqrt{\left|2\kappa(1-2\tau_+)/\left((1-\tau_+)^2\tau_+^2\right) - 2(\kappa-1)/(\tau_+-\eta)^2\right|}} + \frac{O(1)}{m^{3/2}}\right],$$

where $O(1)$ is uniform with respect to $(\kappa,\eta) \in \Omega_{2\delta}$. We would like to express this formula in terms of $t_2, \varphi_2(\cdot)$, and $\psi_2(\cdot)$. Since

$$\left|2\kappa(1-2\tau_+)/\left((1-\tau_+)^2\tau_+^2\right) - 2(\kappa-1)/(\tau_+-\eta)^2\right| = \left|2\varphi_2''(t_2)/\eta^2\right|,$$

$$\varphi_2(t_2) = \phi_0 - \ln\eta,$$

and

$$(\tau_+-\eta)^{-1}(1-\tau_+)^{-1} = \psi_2(t_2)\eta^{-1},$$

we have

$$F_2 = C_m e^{-m\varphi_2(t_2)} \left[\frac{e^{-\mathrm{i}w/2} e^{-\mathrm{i}\arg\eta}}{\mathrm{i}} \frac{\psi_2(t_2)}{\sqrt{|2\pi m \varphi_2''(t_2)|}} + \frac{O(1)}{m^{3/2}} \right].$$

On the other hand, by definition,

$$w = \arg(\varphi_2''(t_2)) - 2\arg\eta = \omega_2 - \pi - 2\arg\eta,$$

where ω_2 is as defined in equation (JO30). Therefore,

$$\begin{aligned} F_2 &= C_m e^{-m\varphi_2(t_2)} e^{-\mathrm{i}\omega_2/2} \left[\frac{\psi_2(t_2)}{\sqrt{|2\pi m \varphi_2''(t_2)|}} + \frac{O(1)}{m^{3/2}} \right] \\ &= C_m \psi_2(t_2) e^{-\mathrm{i}\omega_2/2} |2\pi m \varphi_2''(t_2)|^{-1/2} \exp\{-m\varphi_2(t_2)\}(1 + o(1)). \end{aligned}$$

4.4. *Proof of Confluences.* The confluences (JO35) are established by showing convergence of each of the components in (JO11). For the f_c and f_e components, this follows from inspection of Tables JO4 and JO3 respectively, while for $f_\mathrm{h}^{\mathrm{SigD}}(z)$, this follows from (JO17). For $f_\mathrm{h}^{\mathrm{REG}}(z)$ and $f_\mathrm{h}^{\mathrm{CCA}}(z)$, one uses the definitions of φ_j and t_j and calculation, though one can also appeal to the confluences

$$\begin{aligned} {}_1F_1(a/\epsilon; b; \epsilon x) &\to {}_0F_1(b; ax) \\ {}_2F_1(a_1/\epsilon, a_2/\epsilon; b; \epsilon^2 x) &\to {}_0F_1(b; a_1 a_2 x) \end{aligned}$$

as $\epsilon \to 0$, and observe in (JO21) that with $\kappa \sim c_1/((1-c_1)c_2)$ and as $c_2 \to 0$, we have

$$\begin{aligned} (m\kappa + 1) m\eta_1 &\to m^2 \eta_0 \\ (m\kappa + 1)^2 \eta_2 &\to m^2 \eta_0. \end{aligned}$$

For the confluences (JO36), there is some crosstalk between the components. We write $f_\mathrm{c}[\theta]$ to show the dependence on θ explicitly. Writing $\theta = \sqrt{c_1}\xi$, it is direct to verify that

$$\begin{aligned} f_\mathrm{c}^{\mathrm{PCA}}[\sqrt{c_1}\xi] &= f_\mathrm{c}^{\mathrm{SMD}}[\xi] + \xi/\sqrt{c_1} - \xi^2 - \log\sqrt{c_1} + O(\sqrt{c_1}) \\ f_\mathrm{c}^{\mathrm{REG}_0}[\sqrt{c_1}\xi] &= f_\mathrm{c}^{\mathrm{SMD}}[\xi] + \xi/\sqrt{c_1} - \xi^2/2 - \log\sqrt{c_1} + O(\sqrt{c_1}). \end{aligned}$$

From (JO14) and the MP entry in Table JO3, and writing $z = 1 + \sqrt{c_1}w$, we have

$$f_\mathrm{e}^{\mathrm{PCA}}(1 + \sqrt{c_1}w) = f_\mathrm{e}^{\mathrm{REG}_0}(1 + \sqrt{c_1}w) = f_\mathrm{e}^{\mathrm{SMD}}(w) + \log\sqrt{c_1} + o(1).$$

For the h term, we write $f_\mathrm{h}(z;\theta)$ to show the dependence on θ explicitly. From (JO17), one quickly has

$$f_\mathrm{h}^{\mathrm{PCA}}(1 + \sqrt{c_1}w; \sqrt{c_1}\xi) = f_\mathrm{h}^{\mathrm{SMD}}(w;\xi) - \xi/\sqrt{c_1} + \xi^2 + O(\sqrt{c_1}).$$

For $f_\mathrm{h}^{\mathrm{REG}_0}$, we have $\varphi_0(t_0) = \log t_0 - 2(t_0 - 1)$ and that $t_0 = 1 + \eta_0 - \eta_0^2 + O(\eta_0^3)$ for small η_0. This leads to $f_\mathrm{h}(z;\theta) = (1 - c_1)c_1^{-1}[-\eta_0 + \frac{1}{2}\eta_0^2 + O(\eta_0^3)]$ and thence by elementary evaluation to

$$f_\mathrm{h}^{\mathrm{REG}_0}(1 + \sqrt{c_1}w; \sqrt{c_1}\xi) = f_\mathrm{h}^{\mathrm{SMD}}(w;\xi) - \xi/\sqrt{c_1} + \xi^2/2 + O(\sqrt{c_1}).$$

Combining terms from the preceding displays yields the confluences (JO36).

4.5. *Proof of Lemma JO4 (saddle points z_0).* $\mathsf{q} = 0$ *cases: (SMD, PCA, SigD).*
First, note that

$$f'(z) = f_{\rm e}'(z) + f_{\rm h}'(z) = -m_{\mathbf{c}}(z) + f_{\rm h}'(z), \tag{74}$$

where $m_{\mathbf{c}}(z)$ is the appropriate Stieltjes transform. We proceed, then, by solving for z in the equation $f_{\rm h}'(z) = m_{\mathbf{c}}(z)$.

SMD. We have $f_{\rm h}'(z) = -\theta$, so substituting $m_{\mathbf{c}}(z_0) = -\theta$ into the quadratic equation

$$m^2 + zm + 1 = 0 \tag{75}$$

satisfied by $m = m_{\mathbf{c}}^{\rm SC}(z)$, we get

$$z_0(\theta) = -\frac{m^2+1}{m} = \frac{\theta^2+1}{\theta} = \theta + 1/\theta.$$

Obviously, for any $\theta \in \left(0, \bar{\theta}^{\rm SMD}\right) \equiv (0,1)$, $z_0(\theta)$ is larger than $b_+^{\rm SMD} = \beta_+^{\rm SMD} \equiv 2$.

PCA. Now $f_{\rm h}'(z) = -\theta/[c_1(1+\theta)]$, so we substitute $m_{\mathbf{c}}(z_0) = -\theta/[c_1(1+\theta)]$ into the quadratic equation

$$c_1 z m^2 + (z + c_1 - 1)m + 1 = 0 \tag{76}$$

satisfied by $m = m_{\mathbf{c}}^{\rm MP}(z)$. This is a linear equation for z whose solution is

$$z_0(\theta) = (\theta+1)(\theta+c_1)/\theta.$$

Note that the minimum of $z_0(\theta)$ over $\theta > 0$ equals $b_+^{\rm PCA} \equiv \left(1+\sqrt{c_1}\right)^2$ and is achieved at

$$\theta = \bar{\theta}_{\mathbf{c}} \equiv \sqrt{c_1}.$$

Therefore, since $m_{\mathbf{c}}^{\rm MP}(z)$ is well defined for $z > b_+^{\rm PCA}$, $m_{\mathbf{c}}^{\rm MP}(z_0)$ must be well defined for any $\theta \in (0, \bar{\theta}_{\mathbf{c}})$.

SigD. The Stieltjes transform $m = m_{\mathbf{c}}^{\rm W}(z)$ of the Wachter distribution, as normalized here, satisfies the quadratic equation

$$c_1 z(c_1 - c_2 z)m^2 + [c_1(1-c_2)z - (1-c_1)(c_1 - c_2 z)]m + r^2 = 0, \tag{77}$$

while

$$f_{\rm h}'(z) = \frac{ab}{1-az}, \qquad a = \frac{c_2\theta}{c_1(1+\theta)}, \qquad b = -\frac{r^2}{c_1 c_2}. \tag{78}$$

To solve $m(z) = f_{\rm h}'(z)$, insert $m = ab/(1-az)$ into (77) to obtain an apparently quadratic equation. However the coefficient of z^2 vanishes, so that as with SMD and PCA, z_0 is the solution of a linear equation $\beta z + \gamma = 0$, where in this case

$$\beta = abc_1 l(\theta)/(1+\theta)$$
$$\gamma = -abc_1(c_1+\theta)/\theta,$$

so that

$$z_0(\theta) = \frac{(c_1+\theta)(1+\theta)}{\theta l(\theta)}. \tag{79}$$

It also follows that

$$az_0 = \frac{c_2(c_1+\theta)}{c_1 l(\theta)}, \qquad 1 - az_0 = \frac{r^2}{c_1 l(\theta)}$$

$$f'_{\mathrm{h,SigD}}(z_0) = m(z_0) = -\frac{\theta l(\theta)}{c_1(1+\theta)}. \tag{80}$$

Recall that $l(\theta) = 1 + (1+\theta)c_2/c_1$. Therefore, (79) implies that that the minimum of $z_0(\theta)$ over $\theta > 0$ equals

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

and is achieved at

$$\theta = \bar{\theta}_{\mathbf{c}} \equiv \frac{c_2+r}{1-c_2}.$$

Therefore, $m_{\mathbf{c}}^{\mathrm{W}}(z)$ is well defined for $z > b_+^{\mathrm{SigD}}$, as $\mathbf{n}, p \to_\gamma \infty$, and $m_{\mathbf{c}}^{\mathrm{W}}(z_0)$ must be well defined for any $\theta \in (0, \bar{\theta}_{\mathbf{c}})$.

$\mathsf{q} = 1$ *cases: (REG$_0$, REG, CCA).*

We find the critical points $z_0(\theta)$ for the $\mathsf{q} = 1$ cases by showing that they are the same as for the corresponding $\mathsf{q} = 0$ cases. This is cast as a verification rather than a derivation as we still lack a good explanation for this curious fact.

We have seen, based on (74), that

$$f'_{\mathrm{h,PCA}}(z_0) = m_{\mathbf{c}}^{\mathrm{MP}}(z_0), \qquad f'_{\mathrm{h,SigD}}(z_0) = m_{\mathbf{c}}^{\mathrm{W}}(z_0),$$

for $z_0 = z_0^{\mathrm{PCA}}$ and z_0^{SigD} respectively. We now show that

$$f'_{\mathrm{h,REG_0}}(z_0) = f'_{\mathrm{h,PCA}}(z_0), \qquad f'_{\mathrm{h,REG}}(z_0) = f'_{\mathrm{h,CCA}}(z_0) = f'_{\mathrm{h,SigD}}(z_0)$$

for $z_0 = z_0^{\mathrm{PCA}}$ and z_0^{SigD} respectively. In combination with (74), this verifies that z_0^{PCA} and z_0^{SigD} are critical points for the $\mathsf{q} = 1$ cases as well.

The functions defined in (JO23) and (JO27) will sometimes be written in the form $\varphi_j(t, \eta_j)$ to show the dependence on η_j explicitly. We have

$$f_{\mathrm{h}}(z) = \frac{1-c_1}{c_1}[\varphi_j(t_j, \eta_j) + \gamma_j],$$

where $\gamma_j = \gamma_j(\mathbf{n}, p)$ and $t_j = t_j(\eta_j)$ satisfies

$$\frac{\partial}{\partial t}\varphi_j(t, \eta_j) = 0, \tag{81}$$

a quadratic equation for t_j with coefficients depending on η_j and κ. We therefore have, dropping the subscript j temporarily,

$$f'_{\mathrm{h}}(z) = \frac{1-c_1}{c_1}\frac{\mathrm{d}}{\mathrm{d}\eta}\varphi(t(\eta), \eta)\frac{\mathrm{d}\eta}{\mathrm{d}z} = \frac{1-c_1}{c_1}\frac{\partial}{\partial\eta}\varphi(t(\eta), \eta)\frac{\mathrm{d}\eta}{\mathrm{d}z} \tag{82}$$

From definitions (JO23) and (JO27), again with $t_j = t_j(\eta_j)$, and $\kappa = r^2/[c_2(1-c_1)]$,

$$\frac{\partial}{\partial \eta_j}\varphi(t_j,\eta_j) = \begin{cases} -1/t_0 \\ -t_1 \\ -\kappa t_2/(1-\eta_2 t_2). \end{cases} \tag{83}$$

We now turn to the specifics of the three cases.

REG$_0$. We show that $z = z_0^{\mathrm{PCA}} = (\theta+1)(\theta+c_1)/\theta$ solves

$$f'_{\mathrm{h,REG}_0}(z) = m(z_0) = -\frac{\theta}{c_1(\theta+1)}.$$

From (82) and (83),

$$f'_{\mathrm{h,REG}_0}(z) = -\frac{\theta}{c_1(1-c_1)}\frac{1}{t_0(\eta)},$$

so that we should solve $t_0(\eta_0(z)) = (\theta+1)/(1-c_1)$ for z. Since t_0 satisfies a quadratic equation, the equation for z becomes

$$\eta_0(z) = t_0^2 - t_0 = \frac{(\theta+1)(\theta+c_1)}{(1-c_1)^2},$$

which implies that $z_0^{\mathrm{REG}_0} = \theta^{-1}(1-c_1)^2\eta_0 = (\theta+1)(\theta+c_1)/\theta = z_0^{\mathrm{PCA}}(\theta)$.

REG. This time we solve for z in

$$f'_{\mathrm{h,REG}}(z) = f'_{\mathrm{h,SigD}}(z_0) = -\frac{\theta l(\theta)}{c_1(1+\theta)},$$

where the second equality uses (80). From (82) and (83) we have $f'_{\mathrm{h,REG}}(z) = -c_2\theta t_1(\eta)/c_1^2$, and so

$$t_1(\eta) = \frac{c_1 l(\theta)}{c_2(1+\theta)}, \qquad t_1(\eta) - 1 = \frac{c_1}{c_2(1+\theta)}. \tag{84}$$

The quadratic equation for t_1 is $\eta_1 t_1^2 + (1-\eta_1)t_1 - \kappa = 0$, so that

$$\eta_1 = \frac{\kappa - t_1}{t_1(t_1-1)} = \frac{c_2(\theta+1)(\theta+c_1)}{(1-c_1)c_1 l(\theta)} \tag{85}$$

which implies that $z_0^{\mathrm{REG}} = c_1(1-c_1)\eta_1/(\theta c_2) = z_0^{\mathrm{SigD}}(\theta)$.

CCA. Treat this as a modification of REG. Thus

$$\begin{aligned} \varphi_2(t) &= \kappa\log(1-\eta_2 t) + \eta_1 t + \varphi_1(t), \qquad \text{and} \\ \varphi_2'(t) &= -\frac{\kappa\eta_2}{1-\eta_2 t} + \eta_1 + \varphi_1'(t). \end{aligned} \tag{86}$$

We verify that at $z = z_0^{\mathrm{SigD}}$,

$$t_2 = \frac{c_1 l(\theta)}{c_2(1+\theta)} = t_1(\eta_1(z_0)) \tag{87}$$

satisfies $\varphi_2'(t_2) = 0$ for $\eta_2 = \eta_2(z_0)$. Indeed, writing $L(\theta) = c_1 l(\theta)$, we have

$$\eta_2 = \frac{c_2^2(c_1+\theta)(1+\theta)}{L^2(\theta)}, \qquad t_2\eta_2 = \frac{c_2(c_1+\theta)}{L(\theta)}, \qquad 1 - t_2\eta_2 = \frac{r^2}{L(\theta)}, \tag{88}$$

and

$$\frac{\kappa}{1-t_2\eta_2}=\frac{L(\theta)}{(1-c_1)c_2}=\frac{\eta_1}{\eta_2},$$

so that from (86), $\varphi_2'(t_2)=\varphi_1'(t_1)=0$. But now we can see that, at $z=z_0$,

$$\begin{aligned}f'_{\mathrm{h,CCA}}(z_0)&=-\frac{1-c_1}{c_1}\frac{\kappa}{1-\eta_2 t_2}\cdot t_2\cdot\frac{\mathrm{d}\eta_2}{\mathrm{d}z}\\&=-\frac{1-c_1}{c_1}\frac{\eta_1}{\eta_2}\cdot t_1\cdot\frac{\mathrm{d}\eta_2}{\mathrm{d}z}=-\frac{1-c_1}{c_1}\cdot t_1\cdot\frac{\mathrm{d}\eta_1}{\mathrm{d}z}=f'_{\mathrm{h,REG}}(z_0),\end{aligned}$$

so that z_0 also satisfies $f'_{\mathrm{CCA}}(z_0)=0$.

4.6. *Verification of Remark JO5: that $f(z_0)=0$.* Recall that $f(z_0)=f_{\mathrm{c}}+f_{\mathrm{e}}(z_0)+f_{\mathrm{h}}(z_0)$. The term f_{c} is given in Table 3. The next term,

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

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

LEMMA 7. *For SigD, REG, and CCA, for any $\theta\in\left(0,\bar\theta\right)$ and for sufficiently large $\mathbf{n},p$, we have*

$$f_{\mathrm{e}}\left(z_0\right)=2\ln c_1-\ln\theta-\frac{1-c_1}{c_1}\ln\left(1+\theta\right)-\frac{c_1+c_2}{c_1c_2}\ln\left(c_1+c_2\right)+\frac{r^2}{c_1c_2}\ln\left[c_1 l\left(\theta\right)\right]. \tag{89}$$

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

$$\alpha=\frac{b_++b_-}{2}=\frac{c_1\left(r^2+c_1^2\right)}{\left(c_1+c_2\right)^2},\qquad\beta=\frac{b_+-b_-}{2}=\frac{2rc_1^2}{\left(c_1+c_2\right)^2}. \tag{90}$$

We obtain

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

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 (90) the factorizations

$$\begin{aligned}c_1\left(\alpha-\beta\cos\varphi\right)&=\frac{\beta}{2r}\left(r-c_1z\right)\left(r-c_1z^{-1}\right),\\c_1-c_2\alpha+c_2\beta\cos\varphi&=\frac{\beta}{2r}\left(r+c_2z\right)\left(r+c_2z^{-1}\right),\\z_0-\alpha+\beta\cos\varphi&=q(z)q\left(z^{-1}\right)\ \text{ with}\\q\left(z\right)&=\frac{c_1}{c_1+c_2}\left(\sqrt{c_1l\left(\theta\right)/\theta}+rz\sqrt{\theta/\left[c_1l\left(\theta\right)\right]}\right).\end{aligned}$$

Our integral becomes

$$f_{\mathrm{e}}\left(z_0\right)=\frac{-\left(c_1+c_2\right)r^2}{4\pi\mathrm{i}}\int_{\mathcal{C}}\frac{\left(z-z^{-1}\right)^2\ln\left(q(z)q\left(z^{-1}\right)\right)}{\left(r-c_1z\right)\left(r-c_1z^{-1}\right)\left(r+c_2z\right)\left(r+c_2z^{-1}\right)}\frac{\mathrm{d}z}{z}.$$

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

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

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 (89). □

COROLLARY 8. *For PCA and REG*$_0$*, for any* $\theta \in \left(0, \bar{\theta}\right)$ *and for sufficiently large* $\mathbf{n}, p$*, we have*

$$f_{\mathrm{e}}\left(z_0\right) = \ln c_1 - \ln \theta - \frac{1-c_1}{c_1} \ln\left(1+\theta\right) + \theta/c_1. \tag{91}$$

Proof: The corollary is obtained from Lemma 7 by taking the limit as $c_2 \to 0$. □

COROLLARY 9. *For SMD, for any* $\theta \in \left(0, \bar{\theta}\right)$ *and for sufficiently large* $\mathbf{n}, p$*, we have*

$$f_{\mathrm{e}}\left(z_0\right) = -\ln \theta + \theta^2/2. \tag{92}$$

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,

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

These equations imply that

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

Using this relationship together with Corollary 8 yields $f_{\mathrm{e}}\left(z_0\right) = -\ln\theta + \theta^2/2$ for SMD. □

Observe from Lemma 7 and Corollary 8 that

$$\begin{aligned} f_{\mathrm{e}}^{\mathrm{SigD}} &= f_{\mathrm{e}}^{\mathrm{PCA}} + \log c_1 - \frac{c_1+c_2}{c_1 c_2}\log(c_1+c_2) + \frac{r^2}{c_1 c_2}\log[c_1 l(\theta)] - \theta/c_1 \\ &= f_{\mathrm{e}}^{\mathrm{PCA}} + f_{20}, \qquad \text{and} \\ f_{\mathrm{e}}^{\mathrm{REG}} &= f_{\mathrm{e}}^{\mathrm{REG}_0} + f_{20}, \end{aligned}$$

where f_{20} is defined at (26). Combining Table 3 for f_{c} with this display and Corollaries 8 and 9 for $f_{\mathrm{e}}(z_0)$, we can summarize the results for $f_{\mathrm{c}} + f_{\mathrm{e}}(z_0)$ by case in Table 4 below. For the SigD and REG lines we use (28), namely $f_{21} = f_{10} + f_{20}$, while for CCA we recall that $f_{\mathrm{e}}^{\mathrm{REG}} = f_{\mathrm{e}}^{\mathrm{CCA}}$.

Case	$F = f_c + f_e(z_0)$
SMD	$1+\theta^2$
PCA	$1+\theta/c_1$
SigD	$F^{\mathrm{PCA}} + f_{21}$
REG_0	$2(1+\theta/c_1) + \frac{1-c_1}{c_1}\log\frac{1-c_1}{1+\theta}$
REG	$F^{\mathrm{REG}_0} + f_{21}$
CCA	$F^{\mathrm{REG}} + f_{21}$

TABLE 4

Explicit form of $f_c + f_e(z_0)$ for the different cases.

We turn to the evaluation of $f_h(z_0)$: in each case it will turn out to equal $-F = -f_c - f_e(z_0)$ as shown in Table 4. Again we start with the $\mathsf{q} = 0$ cases, in which $f_h(z)$ is an elementary function.

SMD. We immediately have $f_h(z_0) = -z_0\theta = -\theta^2 - 1$.

PCA. Now $f_h(z_0) = -z_0\theta/[c_1(1+\theta)] = -1 - \theta/c_1$.

SigD. This time, referring to the definition of f_{21} in Table 3,

$$f_h^{\mathrm{SigD}}(z_0) = \frac{r^2}{c_1c_2}\log\left[1 - \frac{c_2}{c_1}\frac{c_1+\theta}{l(\theta)}\right] = \frac{r^2}{c_1c_2}\log\left[\frac{r^2}{c_1 l(\theta)}\right]$$
$$= -f_{21} - F^{\mathrm{PCA}}.$$

REG_0. Since $t_0 = \frac{1}{2}(1+\sqrt{1+4\eta_0})$ satisfies $t_0^2 - t_0 - \eta_0 = 0$, we have

$$\varphi_0(t_0) = \log t_0 - t_0 - \eta_0/t_0 + 1 = \log t_0 - 2\eta_0/t_0.$$

Since $\eta_0(z_0) = (1+\theta)(c_1+\theta)/(1-c_1)^2$, we find after algebra that

$$\sqrt{1+4\eta_0} = \frac{1+c_1+2\theta}{1-c_1}, \tag{93}$$

so that

$$t_0 = \frac{1+\theta}{1-c_1}, \qquad \frac{\eta_0}{t_0} = \frac{c_1+\theta}{1-c_1},$$

and

$$f_h(z_0) = \frac{1-c_1}{c_1}\varphi_0(t_0) = \frac{1-c_1}{c_1}\log\frac{1+\theta}{1-c_1} - \frac{2(c_1+\theta)}{c_1} = -F^{\mathrm{REG}_0}.$$

REG. Combining the definitions of f_h and $\varphi_1(t_1)$ we have

$$f_h^{\mathrm{REG}}(z) = \frac{1-c_1}{c_1}\left\{-\eta_1 t_1 + \kappa\log\frac{\kappa}{t_1} + (\kappa-1)\log\left(\frac{t_1-1}{\kappa-1}\right)\right\}.$$

Combining (84) and (85) gives

$$\eta_1 t_1 = \frac{c_1+\theta}{1-c_1}, \qquad \frac{\kappa}{t_1} = \frac{r^2}{1-c_1}\frac{1+\theta}{L(\theta)}, \qquad \frac{t_1-1}{\kappa-1} = \frac{1-c_1}{1+\theta}, \tag{94}$$

so that

$$f_{\rm h}^{\rm REG}(z_0) = -\frac{c_1+\theta}{c_1} + \frac{r^2}{c_1c_2}\log\frac{r^2}{L(\theta)} + \frac{1-c_1}{c_1}\log\frac{1+\theta}{1-c_1}$$

We can now compare REG with REG_0 just as SigD was compared with PCA: thus

$$f_{\rm h}^{\rm REG}(z_0) - f_{\rm h}^{\rm REG_0}(z_0) = \frac{r^2}{c_1c_2}\log\frac{r^2}{c_1 l(\theta)} + \frac{c_1+\theta}{c_1} = -f_{21},$$

and so

$$f_{\rm h}^{\rm REG}(z_0) = -F^{\rm REG_0} - f_{21} = -F^{\rm REG}. \tag{95}$$

CCA. Combining the definitions of $f_{\rm h}$ and $\varphi_2(t_2)$ we have

$$f_{\rm h}^{\rm CCA}(z) = \frac{1-c_1}{c_1}\left\{\kappa\log(1-\eta_2t_2) + \kappa\log\frac{\kappa}{t_2} + (\kappa-1)\log\left(\frac{t_2-1}{\kappa-1}\right)\right\}.$$

In particular, recalling that $t_2 = t_1$,

$$\begin{aligned} f_{\rm h}^{\rm CCA}(z_0) - f_{\rm h}^{\rm REG}(z_0) &= \frac{1-c_1}{c_1}\left[\kappa\log(1-\eta_2t_2) + \eta_1t_1\right] \\ &= \frac{r^2}{c_1c_2}\log\frac{r^2}{L(\theta)} + \frac{c_1+\theta}{c_1} = -f_{21}, \end{aligned}$$

after substitution from (88) and (94). In combination with (95), we get

$$f_{\rm h}^{\rm CCA}(z_0) = -F^{\rm REG} - f_{21} = -F^{\rm CCA}.$$

4.7. *Proof of Lemma JO8 (contours of steep descent).* 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,

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

is strictly increasing. On the other hand, the definition (JO17) of $f_{\rm h}(z)$ implies that $\mathrm{Re}\, f_{\rm h}(z)$ 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, $\mathrm{Re}\, f_{\rm e}(z)$ is strictly increasing. To show that $\mathrm{Re}\, f_{\rm h}(z)$ is strictly increasing too, it is sufficient to prove that $\mathrm{Re}\,\varphi_j(t_j)$ is strictly increasing for $j=0,2$.

Proof of the monotonicity of $\mathrm{Re}\,\varphi_j(t_j)$ for $j=0,2$. Let us show that $\mathrm{Re}\,\varphi_j(t_j)$ is strictly increasing for $j=0,2$ as z moves away from z_0 along $\mathcal{K}_1$. Recall that for $z\in\mathcal{K}_1$, we have

$$z = z_1 + |z_0-z_1|\exp\{\mathrm{i}\gamma\}, \gamma\in[0,\pi/2].$$

Let

$$R_j = \begin{cases} |z_0-z_1|\,\theta/(1-c_1)^2 & \text{for } j=0 \\ |z_0-z_1|\,\theta c_2^2/\left[c_1^2 l(\theta)\right] & \text{for } j=2 \end{cases}. \tag{96}$$

For REG$_0$, using

$$\eta_j = \begin{cases} z\theta/(1-c_1)^2 & \text{for } j=0 \\ z\theta c_2^2/\left[c_1^2 l(\theta)\right] & \text{for } j=2 \end{cases} . \tag{97}$$

and the definition of φ_0 and t_0, we obtain

$$\operatorname{Re}\varphi_0(t_0) = \tfrac{1}{2}\ln\left(1+4R_0^{1/2}\cos(\gamma/2)+4R_0\right) - 2R_0^{1/2}\cos(\gamma/2)+1-\ln 2.$$

Since the derivative of the above expression with respect to $\gamma \in [0,\pi/2)$ is positive, $\operatorname{Re}\varphi_0(t_0)$ does strictly increase as z moves away from z_0 along $\mathcal{K}_1$.

For CCA, using the identity

$$1-\eta_2 t_2 = \frac{\kappa}{\kappa-1}\frac{t_2-1}{t_2}$$

we obtain

$$\operatorname{Re}\varphi_2(t_2) = -2\kappa\ln|t_2| + (2\kappa-1)\ln|t_2-1| + \kappa\ln\frac{\kappa}{\kappa-1} \tag{98}$$

Further, we have

$$\eta_2 = -\frac{1}{4\kappa(\kappa-1)} + R_2\exp\{\mathrm{i}\gamma\}$$

and

$$t_2 = \frac{2\kappa}{(k_2\exp\{\mathrm{i}\gamma/2\}+1)}, \tag{99}$$

where $k_2 = \sqrt{4R_2\kappa(\kappa-1)}$. Taking the derivative of $\operatorname{Re}\varphi_2(t_2)$ with respect to γ, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\operatorname{Re}\varphi_2(t_2) = \frac{-k_2\sin\gamma/2}{2\left|k_2\exp\{\mathrm{i}\gamma/2\}+1\right|^2} + \frac{k_2\sin\gamma/2}{2\left|1-\frac{k_2}{2\kappa-1}\exp\{\mathrm{i}\gamma/2\}\right|^2}.$$

For $\gamma \in [0,\pi/2]$, the above derivative is positive if

$$\left|k_2\exp\{\mathrm{i}\gamma/2\}+1\right| > \left|1-\frac{k_2}{2\kappa-1}\exp\{\mathrm{i}\gamma/2\}\right|.$$

The latter inequality does hold because $k_2/(2\kappa-1) < k_2$. Hence, $\frac{\mathrm{d}}{\mathrm{d}\gamma}\operatorname{Re}\varphi_2(t_2) > 0$ for $\gamma \in [0,\pi/2]$. □

It remains to prove Lemma JO8 for REG. In the REG case, z moves away from z_0 along $\mathcal{K}_1$ when τ moves away from τ_0 along $\mathcal{C}_1$. Using the definition of φ_j (JO27), the formula (JO33) for $f_{\mathrm{h}}(z)$, and the expression (JO42) for η_1, we obtain

$$\operatorname{Re} f_{\mathrm{h}}(\tau) = \frac{1-c_1}{2c_1}\left(-\operatorname{Re}\tau + \ln|\tau+1| + \kappa\ln|\tau+\kappa| + \kappa\ln\kappa\right).$$

On the other hand, $|\tau+\kappa|$ remains constant on $\mathcal{C}_1$ whereas both $-\operatorname{Re}\tau$ 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 $-\kappa$ of the circumference that represents $\mathcal{C}_1$ is to the left of the point -1. Hence, $\operatorname{Re} f_{\mathrm{h}}(\tau)$ is strictly increasing.

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

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

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

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

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

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

$$\begin{aligned}\gamma(\tau,x)&=\rho^4+(2\kappa-1+x)^2\rho^2-2\rho^3(2\kappa-1+x)\cos\alpha\\&+\kappa^2(\kappa-1)^2+2\left(\rho^2\cos 2\alpha-(2\kappa-1+x)\rho\cos\alpha\right)\kappa(\kappa-1)\end{aligned}$$

so that

$$\frac{\mathrm{d}\gamma(\tau,x)}{\mathrm{d}\cos\alpha}=2\rho\left\{-(2\kappa-1+x)\left[\rho^2+\kappa(\kappa-1)\right]+4\rho\kappa(\kappa-1)\cos\alpha\right\}. \tag{100}$$

We would like to prove that the derivative $\mathrm{d}\gamma(\tau,x)/\mathrm{d}\cos\alpha$ is negative. As is seen from (100), 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

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

This is negative because the expression in the figure brackets is positive. The positivity follows from the observation that

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

To summarize, both $\operatorname{Re}f_{\mathrm{e}}(\tau)$ and $\operatorname{Re}f_{\mathrm{h}}(\tau)$ 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. □

5. Laplace approximation.

5.1. *Proof of Lemma JO9 (extends Olver's asy. expansion).* We closely follow Olver's (1997, pp. 121-125) derivation of an approximation to a similar integral, augmenting Olver's proof by explicit uniform bounds on the approximation errors. First, focus on the integral

$$I^+=\int_{[z_0,b]_{\mathcal{K}}}e^{-p\phi(z)}\chi(z)\mathrm{d}z.$$

Let us introduce new variables v and w by the equations

$$w^2=v=\phi(z)-\phi_0, \tag{101}$$

where the branch of w is determined by $\lim\{\arg v\}=\arg\phi_2+2\beta$ as $z\to z_0$ along $(z_0,b)_{\mathcal{K}}$, and by continuity elsewhere. Here $\beta=\lim\arg(z-z_0)$ as $z\to z_0$ along $(z_0,b)_{\mathcal{K}}$.

Consider w as a function of z. A proof of the following auxiliary lemma is given in the next subsection of this note.

LEMMA 10. *Let $B(\alpha,R)$ and $\overline{B}(\alpha,R)$ denote, respectively, the open and closed balls in $\mathbb{C}$ with center at α and radius R. Suppose that assumptions A0-A4 hold. Then, there exist $\rho_1,\rho_2>0$ with $\rho_2<\rho_1$, which do not depend on p and ω, such that, for sufficiently large p,* *(i) $w(z)$ is holomorphic in $\overline{B}(z_0,\rho_1)$. Furthermore, for any ζ_1,ζ_2 from $\overline{B}(z_0,\rho_1)$, we have* $|w(\zeta_2)-w(\zeta_1)|\geq\frac{1}{2}\left|\phi_2^{1/2}\right||\zeta_2-\zeta_1|$.

(ii) $w(z)$ maps $B(z_0,\rho_1)$ on an open set W that contains 0. The inverse function $z(w)$ is holomorphic in W.

(iii) For any $z_1\in[z_0,b]_{\mathcal{K}}$ such that $|z_1-z_0|=\rho_2$, $\overline{B}(0,2|w(z_1)|)$ is contained in W.

Let z_1 be a point of $[z_0,b]_{\mathcal{K}}$ satisfying Lemma 10 (iii). Then the portion $[z_0,z_1]_{\mathcal{K}}$ of $\mathcal{K}$ can be deformed, without changing the value of the integral

$$\overline{I^+}=\int_{[z_0,z_1]_{\mathcal{K}}}e^{-p\phi(z)}\chi(z)\mathrm{d}z,$$

to make its $w(z)$ map a straight line. Since $\chi(z)$ may be random, the latter statement is only true under qualification: "with probability arbitrarily close to one (w.p.a.c.1) for sufficiently large p." Transformation to the variable v gives

$$\overline{I^+}=e^{-p\phi_0}\int_{[0,\tau]}e^{-pv}\varphi(v)\mathrm{d}v, \tag{102}$$

where

$$\tau=\phi(z_1)-\phi_0,\ \varphi(v)=\chi(z)/\phi'(z), \tag{103}$$

and the path for the integral on the right-hand side of (102) is also a straight line.

For $|v|\leq\tau$ with $|v|\neq 0$, $\varphi(v)$ has a convergent expansion of the form

$$\varphi(v)=\sum_{s=0}^{\infty}a_s v^{(s-1)/2}, \tag{104}$$

w.p.a.c.1 for sufficiently large p. Indeed, it is sufficient to show that expansion

$$w\varphi(v)\equiv w\chi(z)/\phi'(z)=\sum_{s=0}^{\infty}a_s w^s \tag{105}$$

converges for $w\in W$, w.p.a.c.1 for sufficiently large p. But by Lemma 10, $w\chi(z)$ and $\phi'(z)$ viewed as functions of w, are holomorphic in W, w.p.a.c.1 for sufficiently large p. Furthermore, since

$$\phi'(z)\frac{\mathrm{d}}{\mathrm{d}w}z(w)=2w,$$

$\phi'(z)$ is not equal to zero for $w\in\overline{B}\left(0,2\tau^{1/2}\right)\setminus\{0\}$, and, since $\phi_2\neq 0$, $\phi'(z)$ has a simple zero at $w=0$. Therefore, the desired convergence holds, w.p.a.c.1 for sufficiently large p.

The coefficients a_s in (104) can be computed from the coefficients ϕ_j and χ_j defined by equation (JO44). The formulae for a_0,a_1, and a_2 are given, for example, on p. 86 of Olver (1997). We use the formula for a_0 in the statement of Lemma JO9.

Define $\varphi_k(v)$, $k=0,1,2,...$ by the relations $\varphi_k(0)=a_k$ and

$$\varphi(v)=\sum_{s=0}^{k-1}a_s v^{(s-1)/2}+v^{(k-1)/2}\varphi_k(v)\ \text{ for } v\neq 0. \tag{106}$$

Then the integral on the right-hand side of (102) can be rearranged in the form

$$\int_{[0,\tau]} e^{-pv}\varphi(v)\mathrm{d}v=\sum_{s=0}^{k-1}\Gamma\left(\frac{s+1}{2}\right)\frac{a_s}{p^{(s+1)/2}}-\varepsilon_{k,1}\left(p,\omega\right)+\varepsilon_{k,2}\left(p,\omega\right), \tag{107}$$

where

$$\varepsilon_{k,1}\left(p,\omega\right)=\sum_{s=0}^{k-1}\Gamma\left(\frac{s+1}{2},\tau p\right)\frac{a_s}{p^{(s+1)/2}}, \tag{108}$$

$$\varepsilon_{k,2}\left(p,\omega\right)=\int_{[0,\tau]} e^{-pv}v^{(k-1)/2}\varphi_k\left(v\right)\mathrm{d}v, \tag{109}$$

and $\Gamma\left(\alpha,x\right)=\int_x^\infty e^{-t}t^{\alpha-1}\mathrm{d}t$ is the incomplete Gamma function. Keep in mind that τ, a_s, and φ_k depend on p and ω.

Note that $\arg v$ is a continuous function of z, and as mentioned above, $\lim|\arg v|=|\arg\phi_2+2\beta|$ as $z\to z_0$ along $(z_0,b)_{\mathcal{K}}$. On the other hand, Lemma JO9 requires that $|\arg\phi_2+2\beta|\leq\pi/2$. Therefore, $\lim|\arg v|\leq\pi/2$ as $z\to z_0$ along $(z_0,b)_{\mathcal{K}}$. But since $\mathcal{K}$ is a path of steep descent (of $-\operatorname{Re}\phi(z)$), $\operatorname{Re}(v)$ must be positive for $z\in(z_0,b]_{\mathcal{K}}$. Hence, by continuity, $|\arg v|<\pi/2$ for $z\in(z_0,b]_{\mathcal{K}}$. In particular, $|\arg\tau|=|\arg\left(\phi\left(z_1\right)-\phi_0\right)|<\pi/2$. Therefore, each incomplete Gamma function in (108) takes its principal value.

Consider $\varphi(v)w$ as a function of w. Since $\phi'\left(z\right)=2w(z)w'(z)$, we have

$$\varphi(v)w=\chi\left(z\right)/(2w'(z)). \tag{110}$$

By Lemma 10 (i),

$$\left|w'(z)\right|>\tfrac{1}{2}\left|\phi_2^{1/2}\right| \tag{111}$$

for $z\in\overline{B}\left(z_0,\rho_1\right)$. Equation (110), inequality (111), and Assumptions A2, A5 imply that

$$\sup_{w\in W}|\varphi(v)w|=\sup_{z\in\overline{B}(z_0,\rho_1)}|\chi(z)/(2w'(z))|=O_{\mathrm{P}}(1) \tag{112}$$

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

Further, by Assumption A4, there exist positive constants τ_1 and τ_2 (that may depend on $\rho_2\equiv|z_1-z_0|$) such that for all $\omega\in\Omega$ and sufficiently large p, $\operatorname{Re}\tau>\tau_1$ and $|\operatorname{Im}\tau|<\tau_2$. Since $|\tau|\geq|\operatorname{Re}\tau|>\tau_1$, $B\left(0,|\tau_1|^{1/2}\right)$ is contained in W, where $\varphi(v)w$ is analytic. Using Cauchy's estimates for the derivatives of an analytic function (see Theorem 10.26 in Rudin (1987)), (105) and (112), we get

$$|a_s|\leq|\tau_1|^{-s/2}\sup_{w\in B\left(0,|\tau_1|^{1/2}\right)}|\varphi w|=O_{\mathrm{P}}(1). \tag{113}$$

Next, Olver (1997, ch. 4, pp.109-110) shows that $\Gamma\left(\alpha,\zeta\right)=O\left(e^{-\zeta}\zeta^{\alpha-1}\right)$ as $|\zeta|\to\infty$, uniformly in the sector $|\arg\left(\zeta\right)|\leq\pi/2-\delta$ for an arbitrary positive δ. Let us take $\alpha=(s+1)/2$ and $\zeta=\tau p$. Since $\operatorname{Re}\tau>\tau_1$ and $|\operatorname{Im}\tau|<\tau_2$, we have

$$|\tau p|>\tau_1 p\to\infty$$

and

$$|\arg\left(\tau p\right)|=|\arctan(\operatorname{Im}\tau/\operatorname{Re}\tau)|<\arctan(\tau_2/\tau_1)<\pi/2,$$

uniformly in $\omega \in \Omega$ for sufficiently large p. Therefore,

$$\Gamma\left(\frac{s+1}{2}, \tau p\right) = O\left(e^{-\tau p} (\tau p)^{\frac{s-1}{2}}\right) = O\left(e^{-\frac{1}{2}\tau_1 p}\right) \tag{114}$$

for any integer s, uniformly in $\omega \in \Omega$. Equality (114), the definition (108) of $\varepsilon_{k,1}(p, \omega)$, and inequality (113) imply that

$$\varepsilon_{k,1}(p, \omega) = O_{\mathrm{P}}\left(e^{-\frac{1}{2}\tau_1 p}\right), \tag{115}$$

where O_{P} is uniform in $\omega \in \Omega$.

Now consider $w^k \varphi_k(v)$ as a function of w. Since, by definition,

$$w^k \varphi_k(v) = \varphi(v)\, w - \sum_{s=0}^{k-1} a_s w^s,$$

it can be interpreted as a remainder in the Taylor expansion of $\varphi(v)\, w$. As explained above, such an expansion is valid in W, which includes the ball $B\left(0, 2|\tau|^{1/2}\right)$ by Lemma 10 (iii). By a general formula for remainders in Taylor expansions, for any $w \in B\left(0, |\tau|^{1/2}\right)$,

$$\left|w^k \varphi_k(v)\right| \leq \frac{|w|^k}{k!} \max_{w \in B\left(0, |\tau|^{1/2}\right)} \left|\frac{\mathrm{d}^k}{\mathrm{d}w^k}(w\varphi(v))\right|. \tag{116}$$

Further, for any $w \in B\left(0, |\tau|^{1/2}\right)$, a ball with radius $|\tau_1|^{1/2}$ centered in w is contained in the ball $B\left(0, 2|\tau|^{1/2}\right) \subset W$. Therefore, using (112) and Cauchy's estimates for the derivatives of an analytic function (see Theorem 10.26 in Rudin (1987)), we get

$$\max_{w \in B\left(0, |\tau|^{1/2}\right)} \left|\frac{\mathrm{d}^k}{\mathrm{d}w^k}(w\varphi(v))\right| \leq k!\, |\tau_1|^{-k/2} \sup_{w \in W} |w\varphi(v)| = O_{\mathrm{P}}(1). \tag{117}$$

Combining (116) and (117), we have

$$\sup_{v \in (0, \tau]} |\varphi_k(v)| = O_{\mathrm{P}}(1).$$

This equality together with (113) and the fact that, by definition, $\varphi_k(0) = a_k$ imply that

$$\max_{v \in [0, \tau]} |\varphi_k(v)| = O_{\mathrm{P}}(1), \tag{118}$$

where $O_{\mathrm{P}}(1)$ is uniform in $\omega \in \Omega$.

For $\varepsilon_{k,2}(p, \omega)$, the substitution of variable $v = \tau x/p$ in the integral (109) yields

$$\varepsilon_{k,2}(p, \omega) = p^{-(k+1)/2} \int_0^p e^{-\tau x} x^{\frac{k-1}{2}} \tau^{\frac{k+1}{2}} \varphi_k(v)\, \mathrm{d}x.$$

Therefore,

$$\begin{aligned} \left|\varepsilon_{k,2}(p, \omega)\, p^{(k+1)/2}\right| &< \max_{v \in [0, \tau]} |\varphi_k(v)| \int_0^p e^{-\operatorname{Re}\tau x} x^{\frac{k-1}{2}} |\tau|^{\frac{k+1}{2}}\, \mathrm{d}x \\ &< \max_{v \in [0, \tau]} |\varphi_k(v)| \int_0^\infty e^{-\frac{\operatorname{Re}\tau}{|\tau|} y} y^{\frac{k-1}{2}}\, \mathrm{d}y. \end{aligned} \tag{119}$$

Since $\operatorname{Re}\tau > \tau_1$ and $|\operatorname{Im}\tau| < \tau_2$, we have

$$\frac{\operatorname{Re}\tau}{|\tau|} \geq \frac{\operatorname{Re}\tau}{|\operatorname{Re}\tau| + |\operatorname{Im}\tau|} > \frac{\tau_1}{\tau_1+\tau_2}$$

for all $\omega \in \Omega$ and sufficiently large p. Therefore, the integral in (119) is bounded uniformly in $\omega \in \Omega$. Using (118), we conclude that

$$\varepsilon_{k,2}(p,\omega) = O_{\mathrm{P}}\left(p^{-(k+1)/2}\right). \tag{120}$$

Combining (102), (107), (115), and (120), we obtain

$$\overline{I^+} = e^{-p\phi_0}\left(\sum_{s=0}^{k-1}\Gamma\left(\frac{s+1}{2}\right)\frac{a_s}{p^{(s+1)/2}} + \frac{O_{\mathrm{P}}(1)}{p^{(k+1)/2}}\right), \tag{121}$$

where $O_{\mathrm{P}}(1)$ is uniform in $\omega \in \Omega$.

Let us now consider the contribution of $[z_1,b]_{\mathcal{K}}$ to the contour integral

$$I^+ = \int_{[z_0,b]_{\mathcal{K}}} e^{-p\phi(z)}\chi(z)\mathrm{d}z.$$

Since $\mathcal{K}$ is a contour of steep descent,

$$\inf_{z\in[z_1,b]_{\mathcal{K}}} \operatorname{Re}\left(\phi(z)-\phi_0\right) \geq \operatorname{Re}\tau > \tau_1.$$

Therefore, by assumptions A5 and A0, we have

$$\begin{aligned}\left|I^+ - \overline{I^+}\right| &\leq e^{-p\phi_0}e^{-p\tau_1}\int_{[z_1,b]_{\mathcal{K}}}|\chi(z)\mathrm{d}z| \\ &\leq e^{-p\phi_0}e^{-p\tau_1}|\mathcal{K}|\,O_{\mathrm{P}}(1) = e^{-p\phi_0}e^{-p\tau_1}O_{\mathrm{P}}(1),\end{aligned} \tag{122}$$

where $O_{\mathrm{P}}(1)$ is uniform in $\omega \in \Omega$.

Combining (121) and (122), we obtain

$$I^+ = e^{-p\phi_0}\left(\sum_{s=0}^{k-1}\Gamma\left(\frac{s+1}{2}\right)\frac{a_s}{p^{(s+1)/2}} + \frac{O_{\mathrm{P}}(1)}{p^{(k+1)/2}}\right). \tag{123}$$

Finally, note that

$$I_{p,\omega} = I^+ - I^-,$$

where

$$I^- = \int_{[z_0,a]_{\mathcal{K}}} e^{-p\phi(z)}\chi(z)\mathrm{d}z$$

where $[z_0,a]_{\mathcal{K}}$ is a contour that coincides with $[a,z_0]_{\mathcal{K}}$ but has the opposite orientation. The integral I^- can be analyzed similarly to I^+. As explained in Olver (1997, pp.121–122), a_s with odd s in the asymptotic expansion for I^- coincides with the corresponding a_s in the asymptotic expansion for I^-. However, a_s with even s in the two expansions differ by the sign. Therefore, coefficients a_s with odd s cancel out, but those with even s double in the difference of the two expansions. Setting $k = 2m$, we have

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

which establishes the lemma. □

Proof of Lemma 10.

First, we show that there exists ρ_1 such that $w(z)$ is holomorphic in $\overline{B}(z_0,\rho_1)$ and that $\frac{\mathrm{d}}{\mathrm{d}z}w(z_0) = \phi_2^{1/2}$. Let $\phi^{(j)}(z)$ denote the j-th order derivative of $\phi(z)$. Consider a Taylor expansion of $\phi^{(j)}(z)$ at z_0

$$\phi^{(j)}(z) = \sum_{s=0}^{k} \frac{1}{s!}\phi^{(j+s)}(z_0)(z-z_0)^s + R_{j,k+1}.$$

In general, for any $z \in \overline{B}(z_0,x)$, the remainder $R_{j,k+1}$ satisfies

$$|R_{j,k+1}| \leq \frac{|z-z_0|^{k+1}}{(k+1)!} \max_{|t-z_0|\leq x} \left|\phi^{(j+k+1)}(t)\right|. \tag{124}$$

By assumptions A1–A3, there exist constants C_1, C_2, and C_4, such that

$$\left|\phi^{(3)}(t)\right| \leq \frac{C_4}{C_2}\left|\phi^{(2)}(z_0)\right| \tag{125}$$

for any $t \in \overline{B}(z_0,C_1)$. Let $\rho_1 = \min\left\{C_1, \frac{C_2}{2C_4}\right\}$. Then, combining (125) with (124) and recalling that $\frac{1}{j!}\phi^{(j)}(z_0) = \phi_j$, we obtain for $z \in \overline{B}(z_0,\rho_1)$,

$$|R_{0,3}| \leq \frac{|z-z_0|^2}{6}|\phi_2|, \text{ and } |R_{1,2}| \leq \frac{|z-z_0|}{2}|\phi_2|. \tag{126}$$

Further, since

$$R_{0,2} = \phi_2(z-z_0)^2 + R_{0,3},$$

the first of the inequalities in (126) implies that, for $z \in \overline{B}(z_0,\rho_1)$,

$$\frac{5}{6}|\phi_2||z-z_0|^2 \leq |R_{0,2}| \leq \frac{7}{6}|\phi_2||z-z_0|^2. \tag{127}$$

Next, since $\phi_1 = 0$, inequalities (127) imply that

$$|\phi(z)-\phi_0| = |R_{0,2}| \geq \frac{5}{6}|\phi_2||z-z_0|^2 \tag{128}$$

for any $z \in \overline{B}(z_0,\rho_1)$. Since $\phi_2 \neq 0$, inequality (128) implies that $\phi(z)-\phi_0$ does not have zeros in $\overline{B}(z_0,\rho_1)$ except a zero of the second order at $z=z_0$. Therefore,

$$\sqrt{\frac{\phi(z)-\phi_0}{(z-z_0)^2}} = \frac{w(z)}{z-z_0}$$

is holomorphic inside $\overline{B}(z_0,\rho_1)$, and converges to $\phi_2^{1/2}$ as $z \to z_0$. This implies that $w(z)$ is holomorphic in $\overline{B}(z_0,\rho_1)$ and $\frac{\mathrm{d}}{\mathrm{d}z}w(z_0) = \phi_2^{1/2}$.

Now let us show that, for any $z \in \overline{B}(z_0,\rho_1)$,

$$\left|\frac{\mathrm{d}}{\mathrm{d}z}w(z) - \frac{\mathrm{d}}{\mathrm{d}z}w(z_0)\right| \leq \tfrac{1}{2}\left|\frac{\mathrm{d}}{\mathrm{d}z}w(z_0)\right|. \tag{129}$$

Indeed, since

$$\frac{\mathrm{d}}{\mathrm{d}z}w(z) = \frac{\phi^{(1)}(z)}{2w(z)} = \tfrac{1}{2}(\phi(z)-\phi_0)^{-1/2}\phi^{(1)}(z)$$

and $\frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right) = \phi_2^{1/2} \neq 0,$

$$\frac{\frac{\mathrm{d}}{\mathrm{d}z}w\left(z\right)}{\frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right)} = \left(1 + \frac{R_{0,3}}{\phi_2\left(z - z_0\right)^2}\right)^{-\frac{1}{2}}\left(1 + \frac{R_{1,2}}{2\phi_2\left(z - z_0\right)}\right). \tag{130}$$

Note that for any y_1 and y_2 such that $|y_2| < 1$,

$$\left|\frac{1 + y_1}{\sqrt{1 + y_2}} - 1\right| \leq \frac{|y_1| + |y_2|}{1 - |y_2|}, \tag{131}$$

where the principal branch of the square root is used. This follows from the facts that, for $|y_2| < 1$, $\left|\sqrt{1 + y_2}\right| \geq 1 - |y_2|$ and $\left|1 + y_1 - \sqrt{1 + y_2}\right| \leq |y_1| + |y_2|$. Both of these inequalities follow from $\left|1 - \sqrt{1 + y_2}\right| \leq |y_2|$, which can be established by denoting $\sqrt{1 + y_2}$ as x so that the inequality becomes $|1 - x| \leq \left|x^2 - 1\right|$ and using the fact that $1 \leq |x + 1|$ (because $\mathrm{Re}\, x \geq 0$ when $|y_2| < 1$). Setting

$$y_1 = \frac{R_{1,2}}{2\phi_2\left(z - z_0\right)} \text{ and } y_2 = \frac{R_{0,3}}{\phi_2\left(z - z_0\right)^2}$$

and using (126) and (130), we obtain

$$\left|\frac{\frac{\mathrm{d}}{\mathrm{d}z}w\left(z\right)}{\frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right)} - 1\right| \leq \tfrac{1}{2}.$$

Hence, (129) holds.

Finally, let ζ_1 and ζ_2 be any two points in $\overline{B}\left(z_0, \rho_1\right)$, and let $\gamma(t) = (1 - t)\,\zeta_1 + t\zeta_2$, where $t \in [0, 1]$. We have

$$\int_0^1 \left(\frac{\mathrm{d}}{\mathrm{d}z}w\left(\gamma(t)\right) - \frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right)\right)\mathrm{d}t = \frac{w\left(\zeta_2\right) - w\left(\zeta_1\right)}{\zeta_2 - \zeta_1} - \frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right).$$

Therefore, using (129), we obtain

$$\left|\frac{w\left(\zeta_2\right) - w\left(\zeta_1\right)}{\zeta_2 - \zeta_1} - \frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right)\right| \leq \tfrac{1}{2}\left|\frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right)\right|.$$

This inequality and the fact that $\frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right) = \phi_2^{1/2}$ imply part (i) of the lemma.

Part (ii) of the lemma is a simple consequence of part (i). Indeed, by the open mapping theorem, W is an open set. Next, by (i), $w(z)$ is one-to-one mapping of $B\left(z_0, \rho_1\right)$ on W and has a non-zero derivative in $B\left(z_0, \rho_1\right)$. Further, let $\psi\left(w\right)$ be defined on W by $\psi\left(w\left(z\right)\right) = z$. Fix $\tilde{w} \in W$. Then $\psi\left(\tilde{w}\right) = \tilde{z}$ for a unique $\tilde{z}$ in $B\left(z_0, \rho_1\right)$. If $w \in W$ and $\psi\left(w\right) = z$, we have

$$\frac{\psi\left(w\right) - \psi\left(\tilde{w}\right)}{w - \tilde{w}} = \frac{z - \tilde{z}}{w\left(z\right) - w\left(\tilde{z}\right)}.$$

By (i), $w \to \tilde{w}$ as $z \to \tilde{z}$, and the latter equality implies $\frac{\mathrm{d}}{\mathrm{d}w}\psi\left(\tilde{w}\right) = \frac{1}{\frac{\mathrm{d}}{\mathrm{d}z}w(\tilde{z})}$. Therefore, $z(w) \equiv \psi\left(w\right)$ is an analytic inverse of $w(z)$ on W.

Finally, part (iii) of the lemma can be established as follows. Note that by part (i),

$$\left|w\left(z_0 + \rho_1 e^{i\varphi}\right) - w\left(z_0\right)\right| \geq \frac{\rho_1}{2}\left|\frac{\mathrm{d}}{\mathrm{d}z}w\left(z_0\right)\right|$$

for any $\varphi \in [0, 2\pi]$. Therefore, for any w_1 such that $|w_1 - w(z_0)| \leq \frac{\rho_1}{4} \left|\frac{\mathrm{d}}{\mathrm{d}z} w(z_0)\right|$, we have

$$\min_{\varphi \in [0,2\pi]} \left| w_1 - w\left(z_0 + \rho_1 e^{i\varphi}\right) \right| \geq \frac{\rho_1}{4} \left| \frac{\mathrm{d}}{\mathrm{d}z} w(z_0) \right|.$$

By a corollary to the maximum modulus theorem (see Rudin (1987), p. 212), the latter inequality implies that the function $w(z) - w_1$ has a zero in $B(z_0, \rho_1)$. Thus, region W includes $\overline{B}(0, \frac{\rho_1}{4} \left|\frac{\mathrm{d}}{\mathrm{d}z} w(z_0)\right|)$. On the other hand,

$$|w(z_1)| \leq 2\rho_2 \left| \frac{\mathrm{d}}{\mathrm{d}z} w(z_0) \right|. \tag{132}$$

Indeed, consider the identity

$$w^2(z_1) = \phi_1(z_1 - z_0) + R_{0,2}.$$

Since $\phi_1 = 0$, (127) imply

$$|w(z_1)|^2 \leq \frac{7}{6} |\phi_2| |z_1 - z_0|^2. \tag{133}$$

But, by definition,

$$|z_1 - z_0| = \rho_2. \tag{134}$$

Since $\frac{\mathrm{d}}{\mathrm{d}z} w(z_0) = \phi_2^{1/2}$, (133) and (134) imply (132). Setting $\rho_2 = \rho_1/16$, we obtain that W includes $\overline{B}(0, 2|w(z_1)|)$.

5.2. *Evaluation of $d^2 f(z_0)/dz^2$.* Note that $-\mathrm{d}^2 f_{\mathrm{e}}(z_0)/\mathrm{d}z^2 = \mathrm{d}m_{\mathbf{c}}(z_0)/\mathrm{d}z$. Therefore $\mathrm{d}^2 f_{\mathrm{e}}(z_0)/\mathrm{d}z^2$ can be directly evaluated using explicit expressions for the Stieltjes transforms of the semicircle, Marchenko-Pastur and Wachter distributions. Further, using the definition of $f_{\mathrm{h}}(z)$, we directly evaluate $\mathrm{d}^2 f_{\mathrm{h}}(z_0)/\mathrm{d}z^2$. Combining the expressions for the second derivatives of f_{e} and f_{h}, we obtain values of the second derivative of f reported in Table JO6.

Evaluation of $\mathrm{d}m_{\mathbf{c}}(z_0)/\mathrm{d}z$. For each of the three cases, it is a little easier to evaluate

$$a(\theta) = \frac{m'(z_0)}{m^2(z_0)} = -\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{m}\right)\bigg|_{z=z_0}. \tag{135}$$

In each case $v = -1/m$ satisfies a quadratic equation in $v = v(z)$. Differentiation with respect to z yields an equation for v' which we write in the form

$$(C + \Delta)v' = C. \tag{136}$$

SMD. From (75), $v = -1/m$ satisfies $1 - zv + v^2 = 0$, and so, differentiating w.r.t. z,

$$(2v - z)v' = v.$$

At $z = z_0 = \theta + 1/\theta$, with $m(z_0) = -\theta$, we get $C = v_0 = 1/\theta$ and $\Delta = v_0 - z_0 = -\theta$, and

$$a(\theta) = v'(z_0) = \frac{C}{C + \Delta} = \frac{1}{1 - \theta^2}.$$

PCA. From (76), $v = -1/m$ satisfies $c_1 z - (z + c_1 - 1)v + v^2 = 0$, and so, differentiating,

$$(2v - z - c_1 + 1)v' = v - c_1.$$

At $z_0 = (1+\theta)(c_1+\theta)/\theta$ and $v_0 = c_1(1+\theta)/\theta$, we have $C = v_0 - c_1 = c_1/\theta$ and $\Delta = v_0 - z_0 + 1 = -\theta$, so that

$$a(\theta) = v'(z_0) = \frac{C}{C+\Delta} = \frac{c_1}{c_1-\theta^2}.$$

SigD. From (77), $v = -1/m$ satisfies

$$c_1 z(c_1 - c_2 z) - [c_1^2 - c_1 + (c_1 + c_2 - 2c_1c_2)z]v + r^2v^2 = 0,$$

and so v' satisfies (136) with

$$C = c_1^2 - 2c_1c_2(z-v) - (c_1+c_2)v, \qquad \Delta = -c_1 + (c_1+c_2)(z-v).$$

At $z_0 = (1+\theta)(c_1+\theta)/\theta l(\theta)$ and $v_0 = c_1(1+\theta)/\theta l(\theta)$, we find $z_0 - v_0 = (1+\theta)/l(\theta)$, and eventually

$$C = -\frac{c_1}{\theta l(\theta)}[h(\theta)+\theta^2], \qquad \Delta = \frac{c_1}{\theta l(\theta)}\theta^2,$$

with $h(\theta) = c_1 + c_2(1+\theta)^2 - \theta^2$, and hence

$$a(\theta) = v'(z_0) = \frac{h(\theta)+\theta^2}{h(\theta)}.$$

The results are summarized for later reference in Table 5.

	$m(z_0)$	$a(\theta)$	$m'(z_0)$
SMD	$-\theta$	$\frac{1}{1-\theta^2}$	$\frac{\theta^2}{1-\theta^2}$
PCA, REG_0	$-\frac{\theta}{c_1(1+\theta)}$	$\frac{c_1}{c_1-\theta^2}$	$\frac{\theta^2}{c_1(1+\theta)^2(c_1-\theta^2)}$
SigD, REG, CCA	$-\frac{\theta l(\theta)}{c_1(1+\theta)}$	$\frac{h(\theta)+\theta^2}{h(\theta)}$	$\frac{\theta^2 l^2(\theta)}{c_1^2(1+\theta)^2}\frac{h(\theta)+\theta^2}{h(\theta)}$

TABLE 5

Summary of Stieltjes transform quantities. $a(\theta)$ *is defined at* (135), $h(\theta) = c_1 + c_2(1+\theta)^2 - \theta^2$.

Computation of $\mathrm{d}^2 f_{\mathrm{h}}(z_0)/\mathrm{d}z^2$. Since $f''(z) = f_{\mathrm{e}}''(z) + f_{\mathrm{h}}''(z)$ and $f_{\mathrm{e}}''(z) = -m'(z)$, we have

$$-f''(z_0) = m'(z_0) - f_{\mathrm{h}}''(z_0).$$

We will see that in each case there is a factorization

$$m'(z_0) = m^2(z_0)a(\theta)$$
$$f_{\mathrm{h}}''(z_0) = m^2(z_0)b(\theta).$$

Note that the functions $a(\theta), b(\theta)$ are distinct from the constants a, b in (78). Thus

$$-f''(z_0) = m^2(z_0)[a(\theta) - b(\theta)],$$

and the entries of Table JO6 are

(137) $$D_2 = \frac{\theta^2}{-f''(z_0)} = \frac{\theta^2}{m^2(z_0)}\frac{1}{a(\theta)-b(\theta)}.$$

	$b(\theta)$	$\frac{\theta^2}{m^2(z_0)}$	$\frac{1}{a(\theta)-b(\theta)}$
SMD	0	1	$1-\theta^2$
PCA	0	$c_1^2(1+\theta)^2$	$\frac{h_0}{c_1}$
SigD	$-\frac{c_1c_2}{r^2}$	$\frac{c_1^2(1+\theta)^2}{l^2}$	$\frac{hr^2}{c_1^2l^2}$
REG_0	$\frac{c_1}{K_0}$	$c_1^2(1+\theta)^2$	$\frac{h_0K_0}{c_1(1+\theta)^2}$
REG	$\frac{c_1}{K_1}$	$\frac{c_1^2(1+\theta)^2}{l^2}$	$\frac{hK_1}{c_1(1+\theta)^2l^2}$
CCA	$\frac{c_1-c_2(1+\theta)}{K_2}$	$\frac{c_1^2(1+\theta)^2}{l^2}$	$\frac{hK_2}{(c_1+c_2)(1+\theta)^2l}$

TABLE 6

Remaining quantities needed for Table JO6: as shown at (137), the entries there are obtained by multiplying the last two columns of this table. In the last three cases, some algebra is required to verify that $a(\theta)-b(\theta)$ factorizes as shown in the last column. Here $h_0=c_1-\theta^2$, $K_0=1+c_1+2\theta$, $K_1=c_1+\theta+(1+\theta)l$ and $K_2=2(c_1+\theta)+(1-c_1)l$. As $c_2\to 0$, we have $h\to h_0, l\to 1, r^2\to c_1$ and $K_1,K_2\to K_0$.

Evaluation of $b(\theta)$. For **SMD** and **PCA**, $f_{\rm h}(z)$ is linear in z so $b(\theta)=0$.
For **SigD**, from (78) and (80), we find that

$$f_{\rm h}''(z_0)=\frac{1}{b}\left(\frac{ab}{1-az_0}\right)^2=-\frac{c_1c_2}{r^2}m^2(z_0).$$

For the $\mathsf{q}=1$ cases, we have from (82) that

$$f_{\rm h}''(z)=\frac{1-c_1}{c_1}\frac{\mathrm{d}}{\mathrm{d}\eta}\chi(\eta)\left(\frac{\mathrm{d}\eta}{\mathrm{d}z}\right)^2. \tag{138}$$

where $\chi(\eta)=\chi_j(\eta_j)=(\partial/\partial\eta_j)\varphi_j(t_j,\eta_j)$ is given by (83).
REG_0. Recall that $t_0=\frac{1}{2}(1+\sqrt{1+4\eta_0})=(1+\theta)/(1-c_1)$, so that from (93)

$$\dot t_0=\frac{\mathrm{d}}{\mathrm{d}\eta}t_0=(1+4\eta_0)^{-1/2}=\frac{1-c_1}{1+c_1+2\theta}.$$

We have both

$$\frac{\mathrm{d}}{\mathrm{d}\eta}\chi_0(\eta)=\frac{\mathrm{d}}{\mathrm{d}\eta}\left(-\frac{1}{t_0(\eta)}\right)=\frac{\dot t_0}{t_0^2},\qquad\text{and}\qquad\frac{\mathrm{d}\eta}{\mathrm{d}z}=\frac{\theta}{(1-c_1)^2},$$

and so

$$f_{\rm h}''(z_0)=\frac{\theta^2}{c_1(1+c_1+2\theta)(1+\theta)^2}=m^2(z_0)\frac{c_1}{1+c_1+2\theta}.$$

REG. We have $\chi_1(\eta_1)=-t_1(\eta_1)$ and recall that t_1 satisfies a quadratic equation $\eta_1t^2+(1-\eta_1)t-\kappa=0$, so that $\dot t_1=\mathrm{d}t_1/\mathrm{d}\eta$ solves

$$[2\eta_1t_1+1-\eta_1]\dot t_1=t_1(1-t_1).$$

Using (84), we can evaluate

$$t_1(1-t_1)=-\frac{c_1^2l(\theta)}{c_2^2(1+\theta)^2},$$

and setting

$$K_1(\theta) = c_1 + \theta + (1+\theta)l(\theta),$$

we also have from (94) and (85)

$$2\eta_1 t_1 + 1 - \eta_1 = \frac{c_1 K_1(\theta)}{(1-c_1)L(\theta)}.$$

We then find from (138), the previous displays and $\mathrm{d}\eta_1/\mathrm{d}z = \theta c_2/[c_1(1-c_1)]$ that

$$f_{\mathrm{h}}''(z_0) = \frac{\theta^2 l^2(\theta)}{(1+\theta)^2}\frac{1}{c_1 K_1(\theta)} = m^2(z_0)\frac{c_1}{K_1(\theta)}.$$

CCA. Recall that $t_2(\eta_2)$ satisfies $\eta_2(\kappa-1)t^2 + t - \kappa = 0$, and hence $\dot{t}_2 = \mathrm{d}t_2(\eta)/\mathrm{d}\eta$ is given by

$$\dot{t}_2 = \frac{-(\kappa-1)t_2^2}{1+2\eta_2(\kappa-1)t_2}.$$

Since $\chi_2(t_2) = -\kappa t_2/(1-\eta_2 t_2)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\eta}\chi_2(\eta) = \frac{-\kappa}{(1-\eta_2 t_2)^2}(t_2^2 + \dot{t}_2).$$

We have $\kappa = r^2/c_2(1-c_1)$ and $\kappa - 1 = c_1/c_2(1-c_1)$, and so from (88),

$$(\kappa-1)\eta_2 t_2 = \frac{c_1+\theta}{(1-c_1)l(\theta)}$$

and if we define

$$K_2(\theta) = (1-c_1)l(\theta) + 2(c_1+\theta),$$

we arrive at

$$t_2^2 + \dot{t} = \frac{t_2^2}{c_2}\left[c_2 - \frac{L(\theta)}{K_2(\theta)}\right].$$

Some algebra shows that

$$c_1[c_2K_2(\theta) - L(\theta)] = r^2[c_2(1+\theta) - c_1].$$

From (138) and the preceding displays,

$$f_{\mathrm{h}}''(z_0) = -\frac{1-c_1}{\kappa c_1}\left[\frac{\kappa t_2}{1-\eta_2 t_2}\frac{\mathrm{d}\eta_2}{\mathrm{d}z}\right]^2 \frac{r^2}{c_1 c_2}\frac{c_2(1+\theta)-c_1}{K_2(\theta)}.$$

Now from (87) and (88),

$$\frac{\kappa t_2}{1-\eta_2 t_2}\frac{\mathrm{d}\eta}{\mathrm{d}z} = \frac{\theta l(\theta)}{(1+\theta)(1-c_1)} = -\frac{c_1}{1-c_1}m(z_0)$$

and so finally

$$f_{\mathrm{h}}''(z_0) = m^2(z_0)\frac{c_1 - c_2(1+\theta)}{K_2(\theta)}.$$

5.3. *Proof of Theorem JO10.* First, let us show that

$$L_1(\theta;\Lambda) = \frac{g(z_0)}{\sqrt{-\mathrm{d}^2 f(z_0)/\mathrm{d}z^2}} + O_{\mathrm{P}}\left(p^{-1}\right), \tag{139}$$

where $O_{\mathrm{P}}(1)$ *is uniform with respect to* $\theta \in \left(0, \bar{\theta} - \varepsilon\right]$. Changing the variable of integration in (JO45) from z to $\zeta = \theta z$, we obtain

$$L_1(\theta;\Lambda) = \sqrt{\pi p}\frac{1}{2\pi \mathrm{i}}\int_{\tilde{\mathcal{K}}} e^{-p\phi(\zeta)}\chi(\zeta)\mathrm{d}\zeta, \tag{140}$$

where

$$\phi(\zeta) = f(\zeta/\theta)/2,\ \chi(\zeta) = g(\zeta/\theta)/\theta,$$

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 JO6 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 JO9 hold for the integral in (140) for all the six cases that we consider. The validity of A5 follows from Lemma 11 given below and from the definitions of $g(z)$. Let

$$\Delta(\zeta) = p\int \ln(\zeta/\theta - \lambda)\,\mathrm{d}\left(\hat{F}(\lambda) - F_{\mathbf{c}}(\lambda)\right), \tag{141}$$

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

LEMMA 11. *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 $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, where $O_{\mathrm{P}}(1)$ is uniform with respect to $\theta \in \Omega \equiv \left(0, \bar{\theta} - \varepsilon\right]$.

Proof: Let us rewrite (141) 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 $\varphi(\xi)$ is analytic, that encloses these supports. Theorem

2.1 and equation (2.3) of Bai and Yao (2005) for SMD case, and Lemma 1.1 of Bai and Silverstein (2004) for the rest of the cases, imply 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

$$\int_{\mathcal{P}} |p(\hat{m}(\xi) - m_{\mathbf{c}}(\xi)) \mathrm{d}\xi| = O_{\mathrm{P}}(1).$$

(Throughout these notes, notation $\int_{\mathcal{P}} |f(\xi)\mathrm{d}\xi|$ should be interpreted as $\int_{\alpha}^{\beta} |f(\mathcal{P}(t))\mathcal{P}'(t)|\,\mathrm{d}t$, where $\mathcal{P}$ is parameterized as a continuously differentiable complex function on $[\alpha, \beta] \subseteq \mathbb{R}^1$. For piecewise continuously differential pathes, $[\alpha, \beta]$ should be split into a finite number of sub-intervals where $\mathcal{P}$ is continuously differentiable.) Therefore, for any $\delta > 0$, there exists $B > 0$, such that

$$\Pr\left(|\Delta(\varphi)| \le B \sup_{\xi \in \mathcal{P}} |\varphi(\xi)|\right) > 1 - \delta \tag{142}$$

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

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

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

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

Since $\Delta(\varphi_{\zeta,\theta}) = \Delta(\zeta)$, we obtain from (142) that for any $\delta > 0$, there exists $B_2 > 0$ such that for sufficiently large $\mathbf{n}$ and p,

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

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

Applying Lemma JO9 to the integral in (140) and using the fact that $f(z_0) = 0$, we obtain (139). 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

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

Explicitly evaluating the latter integral and using the exact form of g_{c}, available from Table JO4, we obtain

$$|L_2(\theta; \Lambda)| \le \frac{2C}{\sqrt{\pi p}} e^{-(p/2)f_{\mathrm{c}}} (2z_0)^{-p/2} e^{-(p/2)f_{\mathrm{h}}(z_0)} (1 + o(1)),$$

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,

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

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 $\theta \in (0, \bar{\theta} - \varepsilon]$ 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 (139), we establish Theorem JO10 for SMD, PCA, and SigD.

For $\mathbf{REG}_0$, we shall need the following lemma.

LEMMA 12. *For sufficiently large* $\mathbf{n}$ *and* p, *we have*

$$|{}_0F_1(b-s;\Psi_{11}z)| < 4\sqrt{\pi m}\,|\exp\{-m\varphi_0(t_0)\}| \tag{143}$$

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 (JO22) imply that

$${}_0F_1(b-s;\Psi_{11}z) = \Gamma(m+1)\left(m^2\eta_0\right)^{-m/2} e^{-m\pi\mathrm{i}/2} J_m\left(\mathrm{i}2m\eta_0^{1/2}\right). \tag{144}$$

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

$$|J_K(K\zeta)| \leq \left\{1 + \left|\frac{\sin K\pi}{K\pi}\right|\right\}\left|\left\{\frac{\zeta\exp\left\{\sqrt{1-\zeta^2}\right\}}{1+\sqrt{1-\zeta^2}}\right\}^K\right|, \tag{145}$$

(see Watson (1944), p. 270). The latter inequality, equation (144), and the Stirling formula for $\Gamma(m+1)$ imply that (143) holds for sufficiently large m, for any z and $\theta > 0$. The constant 4 on the right hand side of (143) is not the smallest possible one, but it is sufficient for our purposes. □

Using inequality (143), we obtain for REG_0

$$|L_2(\theta;\Lambda)| \leq 4e^{-(p/2)f_{\mathrm{c}}} g_{\mathrm{c}}\sqrt{pm}\int_{\mathcal{K}_2}\left|\exp\{-m\varphi_0(t_0)\}\prod_{j=1}^p (z-\lambda_j)^{-1/2}\,\mathrm{d}z\right|. \tag{146}$$

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

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

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 (146) yields

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

Since, for some constant τ_1, $\operatorname{Re} f(\bar{z}) > f(z_0) + \tau_1 = \tau_1$ and since, by Lemma 11, $4g_{\mathrm{e}}(\bar{z}) = O_{\mathrm{P}}(1)$ uniformly over $\theta \in (0, \bar{\theta} - \varepsilon]$, we obtain

$$|L_2(\theta;\Lambda)| \leq e^{-(p/2)\tau_1} g_{\mathrm{c}} \sqrt{pm} \int_{\mathcal{K}_2} \prod_{j=1}^{p} \left| \frac{\bar{z}-\lambda_j}{z-\lambda_j} \right|^{1/2} |\mathrm{d}z| \, O_{\mathrm{P}}(1) . \tag{147}$$

Note that for any $z \in \mathcal{K}_2$ and any $j = 1, ..., p$, $|(\bar{z}-\lambda_j)/(z-\lambda_j)| \leq 1$ 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

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

Combining this with (147) and noting that $g_{\mathrm{c}} |\bar{z}| = O(1)$ uniformly over $\theta \in (0, \bar{\theta} - \varepsilon]$, we obtain

$$|L_2(\theta;\Lambda)| \leq \sqrt{pm} e^{-(p/2)\tau_1} O_{\mathrm{P}}(1) , \tag{148}$$

where $O_{\mathrm{P}}(1)$ is uniform with respect to $\theta \in (0, \bar{\theta} - \varepsilon]$. Theorem JO10 for REG$_0$ follows from the latter equality and (139).

For **REG** and **CCA**, the Theorem follows from (139) and inequalities

$$|L_2(\theta;\Lambda)| \leq p e^{-p\tau_2} O_{\mathrm{P}}(1) , \tag{149}$$

where τ_2 is a positive constant. We obtain (149) by combining the method used to derive (148) with upper bounds on ${}_1F_1$ and ${}_2F_1$, which we establish using the integral representations (JO25). □

A proof of the domination of $L_2(\theta;\Lambda)$ by $L_1(\theta;\Lambda)$ (via establishing (149)). By definition, we have

$$L_2(\theta;\Lambda) = \sqrt{\pi p} \exp\left\{ -\frac{p}{2} \left(f_{\mathrm{c}} + f_{\mathrm{e}}(z_0) \right) \right\} \frac{g_{\mathrm{c}} g_{\mathrm{e}}(z_0)}{2\pi \mathrm{i}} \int_{\mathcal{K}_2 \cup \bar{\mathcal{K}}_2} F_j \prod_{j=1}^{p} \left(\frac{z_0-\lambda_j}{z-\lambda_j} \right)^{1/2} \mathrm{d}z \tag{150}$$

with $j = 1$ for REG and $j = 2$ for CCA. The idea of the proof is to use the integral representations (31), that is

$$F_j = \frac{C_m \eta_j^{-m}}{2\pi \mathrm{i}} \int_0^{(\eta_j +)} \exp\{-m\phi_j(\tau)\} \chi_j(\tau) \, \mathrm{d}\tau,$$

to find simple upper bounds for $|F_j|$ corresponding to $z \in \mathcal{K}_2 \cup \bar{\mathcal{K}}_2$. Note that since $F_j(\bar{z}) = \overline{F_j(z)}$, it is sufficient to establish the bounds for $z \in \mathcal{K}_2$. These upper bounds will then be used to estimate the integral in (150) from above, and eventually to establish the domination of $L_2(\theta;\Lambda)$ by $L_1(\theta;\Lambda)$.

REG.

LEMMA 13. *Let $\tau_+ \in \mathcal{C}_2$ and z be the corresponding point of $\mathcal{K}_2$. Then $\operatorname{Re} f_{\mathrm{h}}(z) > f_{\mathrm{h}}(z_0) + \alpha$, where $\alpha > 0$ does not depend on $\tau_+ \in \mathcal{C}_2$ and does not depend on θ.*

Proof: Parametrize points $\tau_+ \in \mathcal{C}_2$ as

$$\tau_+ = -\kappa - x + |\tau_0 + \kappa| \exp\{\mathrm{i}\pi/2\}, \tag{151}$$

$x \geq 0$. As x goes from 0 to ∞, the corresponding z tracks contour $\mathcal{K}_2$ from the point ζ, where $\mathcal{K}_2$ and $\mathcal{K}_1$ meet, to $-\infty$. Recall that

$$-\frac{p}{2} f_{\mathrm{h}}(z) = -m\left(\varphi_1(t_1) + k\right) = -m\left(\phi_1\left(\tau_+\right) + \ln \eta_1 + k\right), \tag{152}$$

where $k = \kappa \ln \kappa - (\kappa - 1) \ln(\kappa - 1)$. Using the definition (32) of ϕ_1 and the identity

$$\eta_1 = \frac{\tau_+\left(\tau_+ + 1\right)}{\tau_+ + \kappa}, \tag{153}$$

we obtain

$$\frac{p}{2m} \operatorname{Re} f_{\mathrm{h}}(z) = -\operatorname{Re} \tau_+ + \ln|\tau_+ + 1| - \kappa \ln|\tau_+ + \kappa| + \kappa \ln \kappa.$$

Taking the derivative of both sides of the latter equality with respect to x, we obtain

$$\frac{p}{2m} \frac{\mathrm{d}}{\mathrm{d}x} \operatorname{Re} f_{\mathrm{h}}(z) = 1 + \frac{x + \kappa - 1}{|\tau_+ + 1|^2} - \frac{\kappa x}{|\tau_+ + \kappa|^2}.$$

For $x \geq 0$, we have

$$\begin{aligned} |\tau_+ + \kappa| &\equiv |-x + |\tau_0 + \kappa| \exp\{\mathrm{i}\pi/2\}| > x \text{ and} \\ |\tau_+ + \kappa| &\equiv |-x + |\tau_0 + \kappa| \exp\{\mathrm{i}\pi/2\}| > \kappa. \end{aligned}$$

Therefore, $\kappa x / |\tau_+ + \kappa|^2 < 1$ and $\frac{\mathrm{d}}{\mathrm{d}x} \operatorname{Re} f_{\mathrm{h}}(z) > 0$. This implies that

$$\operatorname{Re} f_{\mathrm{h}}(z) > \operatorname{Re} f_{\mathrm{h}}(\zeta).$$

On the other hand, as shown in subsection 4.7 (pp 38-40 of these notes), $\operatorname{Re} f_{\mathrm{h}}(z)$ strictly increases as z moves along $\mathcal{K}_1$ from z_0 to ζ. Hence, there exists $\alpha > 0$ that does not depend on $\tau_+ \in \mathcal{C}_2$, such that

$$\operatorname{Re} f_{\mathrm{h}}(z) > \operatorname{Re} f_{\mathrm{h}}(z_0) + \alpha = f_{\mathrm{h}}(z_0) + \alpha.$$

From the definitions of $\mathcal{C}_1$ (the image of which under $\tau \mapsto z$ transformation is $\mathcal{K}_1$) and of $f_{\mathrm{h}}(z)$, it is easy to see that α can be chosen so that it does not depend on θ as well. □

LEMMA 14. *There exist positive constants α and α_1 that do not depend on θ such that, for any $\tau_+ \in \mathcal{C}_2$*

$$|F_1| \leq \alpha_1 \sqrt{p}\, |\eta_1| \exp\left\{-\frac{p}{2}\left(f_{\mathrm{h}}(z_0) + \alpha\right)\right\}. \tag{154}$$

Proof: Let $\tau_+ \in \mathcal{C}_2$ and z be the corresponding point of $\mathcal{K}_2$. Choose the contour in the integral representation (31) of F_1 as in subsection 4.3 (that contains a proof of Lemma JO3) of this note. We shall call such a contour $\mathcal{K}^*$. As explained in subsection 4.3, the minimum of $\operatorname{Re} \phi_1(\tau)$ over $\tau \in \mathcal{K}^*$ is achieved either at τ_+ or, in some cases corresponding to situation 3, at τ^* that belongs to

$[0, A]$ and is such that $\operatorname{Re}\tau^* \leq -\kappa$ (see a discussion around equation (56), which shows that points $\tau^* \in [0, A]$ with $\operatorname{Re}\tau^* > -\kappa$ cannot correspond to the minimum of $\operatorname{Re}\phi_1(\tau)$ over $\tau \in \mathcal{K}^*$).

If the minimum of $\operatorname{Re}\phi_1(\tau)$ over $\tau \in \mathcal{K}^*$ is achieved at τ_+ then using (31), (152), and the Stirling's approximation

$$C_m = \frac{\sqrt{\pi p\,(1-c_1)}}{r} \exp\{m(\kappa-1)\ln(\kappa-1) - m\kappa\ln\kappa\}(1+o(1)), \tag{155}$$

we obtain, for some $\tilde{\alpha} > 0$ that does not depend on τ_+ and on θ,

$$|F_1| \leq \tilde{\alpha}\sqrt{p}\exp\left\{-\frac{p}{2}\operatorname{Re} f_{\mathrm{h}}(z)\right\}\int_{\mathcal{K}^*}|\chi_1(\tau)\,\mathrm{d}\tau|. \tag{156}$$

Recall that $\chi_1(\tau) = (\tau - \eta_1)^{-1}$. By definition of $\mathcal{K}^*$,

$$\sup_{\tau\in\mathcal{K}^*}|\chi_1(\tau)| \leq \max\left\{|\tau_+ - \eta_1|^{-1}, |\eta_1|^{-1}\right\} \text{ and } |\mathcal{K}^*| \leq |\eta_1| + 2\pi|\tau_+ - \eta_1|. \tag{157}$$

Identity (153) implies that $|\tau_+ - \eta_1| = (\kappa-1)|\tau_+/(\tau_+ + \kappa)|$ is bounded away from zero uniformly with respect to $\tau_+ \in \mathcal{C}_2$. Therefore, (156) and (157) imply that there exists $\alpha_1 > 0$ that does not depend on τ_+ and on θ such that

$$|F_1| \leq \alpha_1\sqrt{p}\,|\eta_1|\exp\left\{-\frac{p}{2}\operatorname{Re} f_{\mathrm{h}}(z)\right\}.$$

Combining this with Lemma 13, we obtain (154).

If the minimum of $\operatorname{Re}\phi_1(\tau)$ over $\tau \in \mathcal{K}^*$ is achieved at τ^* then we must be in situation 3 so that $|\tau^* - \eta_1| > |\tau_+ - \eta_1|$ and

$$\begin{aligned}\operatorname{Re}\phi_1(\tau^*) &= -\operatorname{Re}\tau^* - \kappa\ln|\tau^*| + (\kappa-1)\ln|\tau^* - \eta_1| \\ &> -\operatorname{Re}\tau^* - \kappa\ln|\tau^*| + (\kappa-1)\ln|\tau_+ - \eta_1|.\end{aligned}$$

Let τ be any point on the ray starting at 0 and passing through η_1, let $\arg\eta_1 = \beta$ (note that $\beta > \pi/2$ so that $\cos\beta < 0$), and let $x = |\tau|$. Then

$$-\operatorname{Re}\tau^* - \kappa\ln|\tau^*| \geq -\max_{x\geq 0}\{x\cos\beta + \kappa\ln x\} = \kappa - \kappa\ln(-\kappa/\cos\beta).$$

Therefore,

$$\operatorname{Re}\phi_1(\tau^*) > \kappa - \kappa\ln(-\kappa/\cos\beta) + (\kappa-1)\ln|\tau_+ - \eta_1|.$$

This inequality implies

$$\operatorname{Re}\phi_1(\tau^*) + \ln|\eta_1| > \kappa - \kappa\ln(-\kappa/\cos\beta) + (\kappa-1)\ln|\tau_+ - \eta_1| + \ln|\eta_1|.$$

Using (153) and the fact that $\tau_+ \in \mathcal{C}_2$, we obtain

$$\begin{aligned}\operatorname{Re}\phi_1(\tau^*) + \ln|\eta_1| &> \kappa - \kappa\ln(-\kappa/\cos\beta) + \kappa\ln\left|\frac{\tau_+}{\tau_+ + \kappa}\right| \\ &\quad + \ln|\tau_+ + 1| + (\kappa-1)\ln(\kappa-1) \\ &> \kappa - \kappa\ln(-\kappa/\cos\beta) + \ln|\bar{\tau}_+ + 1| + (\kappa-1)\ln(\kappa-1),\end{aligned} \tag{158}$$

where $\bar{\tau}_+ = -\kappa + |\tau_0 + \kappa|\exp\{\mathrm{i}\pi/2\}$ is the point where $\mathcal{C}_2$ and $\mathcal{C}_1$ meet. On the other hand,

$$\cos\beta \leq \cos\arg\bar{\tau}_+ = -\kappa/|\bar{\tau}_+|$$

and thus

$$\begin{aligned}\kappa-\kappa\ln\left(-\kappa/\cos\beta\right) &> \kappa-\kappa\ln\left|\bar{\tau}_{+}\right|=\kappa+\kappa\ln\left(\left|\bar{\tau}_{+}+\kappa\right|/\left|\bar{\tau}_{+}\right|\right)-\kappa\ln\left|\bar{\tau}_{+}+\kappa\right| \\ &> \kappa+\kappa\ln\left(1/\sqrt{2}\right)-\kappa\ln\left|\bar{\tau}_{+}+\kappa\right|>-\kappa\ln\left|\bar{\tau}_{+}+\kappa\right|.\end{aligned}$$

Using this inequality and (158), we obtain

$$\operatorname{Re}\phi_1\left(\tau^*\right)+\ln\left|\eta_1\right|>-\kappa\ln\left|\bar{\tau}_{+}+\kappa\right|+\ln\left|\bar{\tau}_{+}+1\right|+(\kappa-1)\ln(\kappa-1).$$

Since $|\tau+\kappa|$ stays constant for $\tau\in\mathcal{C}_1$ whereas $|\tau+1|$ is strictly decreasing as τ moves along $\mathcal{C}_1$ from τ_0 to $\bar{\tau}_+$, there exists $\alpha_2>0$ which is independent of τ_+ and θ, such that

$$\begin{aligned}\operatorname{Re}\phi_1\left(\tau^*\right)+\ln\left|\eta_1\right| &> -\kappa\ln\left(\tau_0+\kappa\right)+\ln\left(\tau_0+1\right)+(\kappa-1)\ln(\kappa-1)+\alpha_2 \\ &> -\tau_0-\kappa\ln\left(\tau_0+\kappa\right)+\ln\left(\tau_0+1\right)+(\kappa-1)\ln(\kappa-1)+\alpha_2 \\ &= \operatorname{Re}\phi_1\left(\tau_0\right)+\ln\left|\eta_{10}\right|+\alpha_2,\end{aligned}$$

where η_{10} is the value of η_1 that corresponds to z_0. Therefore, by (152), we have

$$-m\left(\operatorname{Re}\phi_1\left(\tau^*\right)+\ln\left|\eta_1\right|\right)<-m\left(\frac{p}{2m}f_{\mathrm{h}}(z_0)+\alpha_2-k\right).$$

Using this inequality together with (31) and (155), we obtain that, for some $\tilde{\alpha}>0$ that does not depend on τ_+ and θ,

$$\left|F_1\right|\leq\tilde{\alpha}\sqrt{p}\exp\left\{-\frac{p}{2}\left(f_{\mathrm{h}}(z_0)+\frac{2m}{p}\alpha_2\right)\right\}\int_{\mathcal{K}^*}\left|\chi_1(\tau)\,\mathrm{d}\tau\right|.$$

Analysing the integral $\int_{\mathcal{K}^*}\left|\chi_1(\tau)\,\mathrm{d}\tau\right|$ as above, we conclude that there exist $\alpha,\alpha_1>0$ that do not depend on τ_+ and θ, such that (154) holds. □

Using Lemma 14 and equation (150), we obtain the following bound on $\left|L_2(\theta;\Lambda)\right|$

$$\left|L_2(\theta;\Lambda)\right|\leq\alpha_1 p\exp\left\{-\frac{p}{2}\alpha\right\}\left|g_{\mathrm{c}}g_{\mathrm{e}}(z_0)\right|\int_{\mathcal{K}_2}\left|\eta_1\prod_{j=1}^{p}\left(\frac{z_0-\lambda_j}{z-\lambda_j}\right)^{1/2}\mathrm{d}z\right|. \tag{159}$$

On the other hand, for any λ_j from the support of $F_{\mathbf{c}}$, we have

$$\left|\frac{z_0-\lambda_j}{z-\lambda_j}\right|<\left|\frac{z_0}{z}\right|=\left|\frac{\eta_{10}}{\eta_1}\right|=\left|\frac{\tau_0\left(\tau_0+1\right)\left(\tau_{+}+\kappa\right)}{\left(\tau_0+\kappa\right)\tau_{+}\left(\tau_{+}+1\right)}\right| \tag{160}$$

and

$$\mathrm{d}z=\frac{c_1\left(1-c_1\right)}{\theta c_2}\mathrm{d}\eta_1=\frac{c_1\left(1-c_1\right)}{\theta c_2}\left(1-\frac{\kappa(\kappa-1)}{\left(\tau_{+}+\kappa\right)^2}\right)\mathrm{d}\tau_{+}. \tag{161}$$

Note that, for any $\tau_+\in\mathcal{C}_2$

$$\left|\frac{\kappa(\kappa-1)}{\left(\tau_{+}+\kappa\right)^2}\right|<\frac{\kappa(\kappa-1)}{\left(\tau_0+\kappa\right)^2}.$$

A direct calculation based on the definitions

$$\begin{aligned}\tau_0 &= \tfrac{1}{2}\left\{\eta_{10}-1+\sqrt{\left(\eta_{10}-1\right)^2+4\kappa\eta_{10}}\right\}, \\ \eta_{10} &= \frac{z_0\theta c_2}{c_1\left(1-c_1\right)},\ \kappa=\frac{c_1+c_2-c_1c_2}{c_2\left(1-c_1\right)},\ \text{and} \\ z_0 &= \frac{(1+\theta)\left(\theta+c_1\right)}{\theta\left(1+(1+\theta)c_2/c_1\right)}\end{aligned}$$

yields

$$\tau_0 = \frac{\theta + c_1}{1 - c_1}, \ \eta_{10} = \frac{c_2 (\theta + 1)(\theta + c_1)}{(c_1 + c_2 + \theta c_2)(1 - c_1)}, \text{ and}$$
$$\frac{\kappa(\kappa - 1)}{(\tau_0 + \kappa)^2} = c_1 \frac{c_1 + c_2 - c_1 c_2}{(c_1 + c_2 + \theta c_2)^2}.$$

The latter two equalities together with (161) imply that there exists a constant $\alpha_2 > 0$ that does not depend on $\theta \in (0, \bar{\theta} - \varepsilon]$ such that (for sufficiently large $\mathbf{n}, p$ as $\mathbf{n}, p \to_\gamma \infty$)

$$|\eta_{10} \mathrm{d}z| < \frac{\alpha_2}{\theta} |\mathrm{d}\tau_+| .$$

Using this and (160) in (159), we obtain

$$|L_2(\theta; \Lambda)| \le \alpha_1 \alpha_2 p \exp\left\{-\frac{p}{2}\alpha\right\} \left|\frac{g_c}{\theta} g_e(z_0)\right| \int_{\mathcal{C}_2} \left|\frac{\tau_0 (\tau_0 + 1)(\tau_+ + \kappa)}{(\tau_0 + \kappa)\tau_+ (\tau_+ + 1)}\right|^{p/2-1} |\mathrm{d}\tau_+| .$$

Note that for any $\tau_+ \in \mathcal{C}_2$ we have $|\tau_+ + 1| > |\tau_+ + \kappa|$. On the other hand, $\tau_0 + 1 < \tau_0 + \kappa$. Therefore,

$$\left|\frac{(\tau_0 + 1)(\tau_+ + \kappa)}{(\tau_0 + \kappa)(\tau_+ + 1)}\right| < 1$$

and

$$|L_2(\theta; \Lambda)| \le \alpha_1 \alpha_2 p \exp\left\{-\frac{p}{2}\alpha\right\} \left|\frac{g_c}{\theta} g_e(z_0)\right| \int_{\mathcal{C}_2} \left|\frac{\tau_0}{\tau_+}\right|^{p/2-1} |\mathrm{d}\tau_+| .$$

Using parameterization (151), we obtain

$$\begin{aligned} |L_2(\theta; \Lambda)| &\le \alpha_1 \alpha_2 p \exp\left\{-\frac{p}{2}\alpha\right\} \left|\frac{g_c}{\theta} g_e(z_0)\right| \int_0^\infty \left|\frac{\tau_0}{x + \kappa - \mathrm{i}|\tau_0 + \kappa|}\right|^{p/2-1} \mathrm{d}x \\ &\le \alpha_1 \alpha_2 \alpha_3 p \exp\left\{-\frac{p}{2}\alpha\right\} \left|\frac{g_c}{\theta} g_e(z_0)\right| \end{aligned}$$

for some $\alpha_3 > 0$ that does not depend on θ. Finally, note that $g_c/\theta = O(1)$ and $g_e(z_0) = O_P(1)$, so that the above display implies equation (149). Since

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

we see that $L_2(\theta; \Lambda)$ is asymptotically dominated by $L_1(\theta; \Lambda)$.

CCA. Let A be an arbitrarily large positive constant. Split the contour $\mathcal{K}_2$ into $\mathcal{K}_{21}$ and $\mathcal{K}_{22}$, where

$$\mathcal{K}_{21} = \{z : z \in \mathcal{K}_2, \operatorname{Re} z > -A\} .$$

Note that the approximation

$$F_2 = C_m \psi_2(t_2) e^{-\mathrm{i}\omega_2/2} \left|2\pi m \varphi_2''(t_2)\right|^{-1/2} \exp\{-m\varphi_2(t_2)\} (1 + o(1))$$

derived in Lemma JO3 remains valid for $z \in \mathcal{K}_{21}$. Therefore, the representation

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

is valid for $z \in \mathcal{K}_{21} \cup \mathcal{K}_1$. Hence, if we show that $\mathcal{K}_{21} \cup \mathcal{K}_1$ is a contour of steep descent for $-\operatorname{Re} f(z)$, then

$$L_{21}(\theta;\Lambda) + L_1(\theta;\Lambda)$$

must be asymptotically equivalent to $L_1(\theta;\Lambda)$, where

$$L_{21}(\theta;\Lambda) = \sqrt{\pi p}\frac{1}{2\pi \mathrm{i}}\int_{\mathcal{K}_{21}\cup\bar{\mathcal{K}}_{21}} \exp\left\{-\frac{p}{2}f(z)\right\} g(z)\mathrm{d}z,$$

and thus, $L_{21}(\theta;\Lambda)$ must be asymptotically dominated by $L_1(\theta;\Lambda)$.

Obviously, $-\operatorname{Re} f_{\mathrm{e}}(z)$ is decreasing as z moves along $\mathcal{K}_{21}$ so that $\operatorname{Re} z$ becomes more and more negative. Let us consider the behavior of

$$-\operatorname{Re} f_{\mathrm{h}}(z) = \frac{1-c_1}{c_1}\left(-\varphi_2(t_2) - \kappa\ln\kappa + (\kappa-1)\ln(\kappa-1)\right). \tag{162}$$

Recall (98), that states

$$\operatorname{Re}\varphi_2(t_2) = -2\kappa\ln|t_2| + (2\kappa-1)\ln|t_2-1| + \kappa\ln\frac{\kappa}{\kappa-1}.$$

Parametrize $z \in \mathcal{K}_{21}$ as

$$z = z_1 - x|z_0-z_1| + |z_0-z_1|\mathrm{i},\ x \in [0, (A+z_1)/|z_0-z_1|]$$

where

$$z_1 = -\frac{c_1(1-c_1)^2 l(\theta)}{4\theta r^2}.$$

For the corresponding $\eta_2 = z\theta c_2^2/\left[c_1^2 l(\theta)\right]$ we have

$$\eta_2 = R_0 - xR_1 + R_1\mathrm{i},\ x \in [0, (A+z_1)/|z_0-z_1|],$$

where

$$R_0 = -\frac{1}{4\kappa(\kappa-1)} \text{ and } R_1 = |z_0-z_1|\frac{\theta c_2^2}{c_1^2 l(\theta)}.$$

From the definition of t_2 we obtain

$$t_2 = \frac{2\kappa}{1+\sqrt{1+4\kappa(\kappa-1)(R_0-xR_1+R_1\mathrm{i})}},$$

which implies that

$$t_2 = \frac{2\kappa}{1+\rho\sqrt{-x+\mathrm{i}}}, \tag{163}$$

where

$$\rho = \sqrt{4\kappa(\kappa-1)R_1}.$$

LEMMA 15. *Let (163) hold. Then* $\frac{\mathrm{d}}{\mathrm{d}x}\left(-\operatorname{Re}\varphi_2(t_2)\right) < 0$ *for* $x \geq 0$.

Proof: Since

$$\mathrm{Re}\sqrt{-x+\mathrm{i}}=\sqrt{\frac{\sqrt{x^2+1}-x}{2}}\text{ and }\mathrm{Im}\sqrt{-x+\mathrm{i}}=\sqrt{\frac{\sqrt{x^2+1}+x}{2}},$$

we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(-\mathrm{Re}\,\varphi_2\left(t_2\right)\right) &= -\frac{1}{2\sqrt{x^2+1}}\frac{\rho^2x-\rho\,\mathrm{Re}\sqrt{-x+\mathrm{i}}}{\left|1+\rho\sqrt{-x+\mathrm{i}}\right|^2}\\ &\quad -\frac{2\kappa-1}{2\sqrt{x^2+1}}\frac{\rho^2x+(2\kappa-1)\,\rho\,\mathrm{Re}\sqrt{-x+\mathrm{i}}}{\left|2\kappa-1-\rho\sqrt{-x+\mathrm{i}}\right|^2}.\end{aligned}$$

For $x\geq 0$ this is no larger than

$$-\frac{\rho\,\mathrm{Re}\sqrt{-x+\mathrm{i}}}{2\sqrt{x^2+1}}\left(\frac{-1}{\left|1+\rho\sqrt{-x+\mathrm{i}}\right|^2}+\frac{(2\kappa-1)^2}{\left|2\kappa-1-\rho\sqrt{-x+\mathrm{i}}\right|^2}\right),$$

which is negative because

$$\left|1+\rho\sqrt{-x+\mathrm{i}}\right|>\left|1-\frac{\rho}{2\kappa-1}\sqrt{-x+\mathrm{i}}\right|.\square$$

Lemma 15 and identity (162) imply that $-\mathrm{Re}\,f_{\mathrm{e}}(z)$ is decreasing as z moves along $\mathcal{K}_{21}$. Hence $\mathcal{K}_{21}\cup\mathcal{K}_1$ is indeed a contour of steep descent for $-\mathrm{Re}\,f(z)$, and therefore $L_{21}(\theta;\Lambda)$ is asymptotically dominated by $L_1(\theta;\Lambda)$. It remains to be shown that $L_{22}(\theta;\Lambda)=L_2(\theta;\Lambda)-L_1(\theta;\Lambda)$ is asymptotically dominated by $L_1(\theta;\Lambda)$.

For any $z\in\mathcal{K}_{22}$ and the corresponding $\eta_2=z\theta c_2^2/\left[c_1^2 l(\theta)\right]$, consider the integral representation

$$F_2=\frac{C_m}{2\pi\mathrm{i}}\int_0^{(1+)}\exp\left\{-m\varphi_2(t)\right\}\psi_2(t)\,\mathrm{d}t, \tag{164}$$

where

$$\begin{aligned}\varphi_2(t)&=-\kappa\ln(t)+(\kappa-1)\ln(t-1)+\kappa\ln(1-\eta_2 t)\\ \psi_2(t)&=(t-1)^{-1}(1-\eta_2 t)^{-1}.\end{aligned}$$

For a fixed contour $\mathcal{K}_*$ in (164), it is clearly possible to make $\mathrm{Re}\,\varphi_2(t)$ arbitrarily large and $|\psi_2(t)|$ arbitrarily close to zero, uniformly with respect to $t\in\mathcal{K}_*$ by choosing A sufficiently large (so that $|\eta_2|$ is sufficiently large). Therefore, by choosing A sufficiently large, we shall have inequality

$$|F_2|\leq\tilde{\alpha}\sqrt{p}\exp\left\{-\frac{p}{2}\left(\mathrm{Re}\,f_{\mathrm{h}}(z_0)+\alpha\right)\right\}$$

for some $\tilde{\alpha},\alpha>0$ (that do not depend on θ) and any $z\in\mathcal{K}_{22}$. Using this upper bound in (150), we obtain

$$L_{22}(\theta;\Lambda)\leq\alpha_1 p\exp\left\{-\frac{p}{2}\alpha\right\}|g_{\mathrm{c}}g_{\mathrm{e}}(z_0)|\int_{\mathcal{K}_{22}}\left|\prod_{j=1}^{p}\left(\frac{z_0-\lambda_j}{z-\lambda_j}\right)^{1/2}\mathrm{d}z\right|$$

for some $\alpha_1>0$ that does not depend on θ.

Clearly for any $z \in \mathcal{K}_{22}$ and any λ_j from the support of $F_{\mathbf{c}}$ we have

$$\left|\frac{z_0 - \lambda_j}{z - \lambda_j}\right| \leq \left|\frac{z_0}{z}\right| = \left|\frac{\eta_{20}}{\eta_2}\right|,$$

where η_{20} is the value of η_2 that correspond to $z = z_0$. Therefore, we have for some $\alpha_2 > 0$ that does not depend on θ

$$L_{22}\left(\theta;\Lambda\right) \leq \alpha_2 p \exp\left\{-\frac{p}{2}\alpha\right\} \left|\frac{g_{\mathrm{c}}}{\theta} g_{\mathrm{e}}(z_0)\right| \left|\int_{\mathcal{K}_{22}} \left|\frac{\eta_{20}}{\eta_2}\right|^{p/2} \mathrm{d}\eta_2\right|,$$

and thus, for some $\alpha_3 > 0$ that does not depend on θ,

$$L_{22}\left(\theta;\Lambda\right) \leq \alpha_3 p \exp\left\{-\frac{p}{2}\alpha\right\} \left|\frac{g_{\mathrm{c}}}{\theta} g_{\mathrm{e}}(z_0)\right|.$$

Finally, note that $g_{\mathrm{c}}/\theta = O(1)$ and $g_{\mathrm{e}}(z_0) = O_{\mathrm{P}}(1)$, so that the above display implies (149) with L_{22} replacing L_2. Since

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

we see that $L_{22}\left(\theta;\Lambda\right)$ is asymptotically dominated by $L_1\left(\theta;\Lambda\right)$.

6. Asymptotics of LR.

6.1. *Derivations for Theorem JO11 (limiting LR).* We record details to verify that

$$\frac{g(z_0)}{\sqrt{-f''(z_0)}} = \exp\left\{-\tfrac{1}{2}\Delta_p(\theta) + \tfrac{1}{2}\log[1-\delta_p^2(\theta)]\right\}(1+o(1)),$$

where, perhaps surprisingly, our six cases reduce to the three values for $\delta_p(\theta)$ given in Theorem JO11. Recall the decomposition $g = g_{\mathrm{c}} g_{\mathrm{e}} g_{\mathrm{h}}$ and note from the definitions (JO15) that $g_{\mathrm{e}}(z_0) = \exp\{-\frac{1}{2}\Delta_p(\theta)\}$. Consequently, from the definition of D_2 in Table JO6, the left side of the previous display may be written as

$$\theta^{-1} g_{\mathrm{c}} g_{\mathrm{h}} \sqrt{D_2} \exp\{-\tfrac{1}{2}\Delta_p(\theta)\},$$

so our task is to verify that

$$P = \theta^{-1} g_{\mathrm{c}} g_{\mathrm{h}} \sqrt{D_2} = (1-\delta_p^2(\theta))^{1/2}(1+o(1)). \tag{165}$$

To this end, Table 7 collects values for $\theta^{-1}\breve{g}_{\mathrm{c}}$, g_{h}, and $\sqrt{D_2}$ from Table JO4, Section JO4.1 and Table JO6 respectively. Cases SMD and PCA require no further comment. For the remaining cases, we add remarks on the evaluation of $g_{\mathrm{h}}(z_0)$ and then the product (165).

SigD. First observe that since $z_0\theta = (1+\theta)(c_1+\theta)/l(\theta)$,

$$g_{\mathrm{h}}(z_0) = \left(1 - \frac{c_2 z_0 \theta}{c_1(1+\theta)}\right)^{-1} = \frac{c_1 l(\theta)}{r^2},$$

and we get the claimed expression for P,

$$P^2 = \frac{(c_1+c_2)h}{c_1^2 l^2} = 1 - \frac{\theta^2 r^2}{c_1^2 l^2}, \tag{166}$$

	$\theta^{-1}\breve{g}_c$	g_h	$\sqrt{D_2}$
SMD	1	1	$\sqrt{1-\theta^2}$
PCA	$\dfrac{1}{c_1(1+\theta)}$	1	$c_1(1+\theta)\sqrt{h_0/c_1}$
SigD	$\dfrac{r\sqrt{c_1+c_2}}{c_1^2(1+\theta)}$	$\dfrac{c_1 l(\theta)}{r^2}$	$\dfrac{r(1+\theta)\sqrt{h}}{l^2(\theta)}$
REG_0	$\dfrac{1}{c_1\sqrt{1-c_1}}$	$\dfrac{\sqrt{1-c_1}}{\sqrt{K_0}}$	$c_1\sqrt{K_0 h_0/c_1}$
REG	$\dfrac{r\sqrt{c_1+c_2}}{c_1^2\sqrt{1-c_1}}$	$\dfrac{\sqrt{c_1(1-c_1)}l(\theta)}{r\sqrt{K_1}}$	$\dfrac{\sqrt{c_1K_1h}}{l^2(\theta)}$
CCA	$\dfrac{r^2(c_1+c_2)}{c_1^3\sqrt{1-c_1}l(\theta)}$	$\dfrac{c_1\sqrt{1-c_1}l^{3/2}(\theta)}{r^2\sqrt{K_2}}$	$\dfrac{c_1\sqrt{K_2h}}{\sqrt{c_1+c_2}l^{3/2}(\theta)}$

TABLE 7

Components of the product $P = \theta^{-1}g_c g_h\sqrt{D_2}$. The CCA entry for g_h is shown for completeness – it is derived, post facto, from the calculations above.

after using the identity

$$(c_1+c_2)h = c_1^2 l^2 - \theta^2 r^2 \tag{167}$$

REG$_0$. From (JO34) and (93), we have

$$g_h(z_0) \sim (1+4\eta_0)^{-1/4} \sim \sqrt{1-c_1}/\sqrt{K_0}.$$

REG. We use (JO34) to evaluate $g_h(z_0)$. Using (84) to evaluate $t_1(z_0)$, we have

$$\begin{aligned}\varphi_1''(t_1) &= \frac{\kappa}{t_1^2} - \frac{\kappa-1}{(t_1-1)^2} = \frac{c_2^2(1+\theta)^2}{c_2(1-c_1)}\left[\frac{r^2}{L^2(\theta)} - \frac{1}{c_1}\right] \\ &= -\frac{c_2^2(1+\theta)^2}{c_1^2(1-c_1)}\frac{K_1(\theta)}{l^2(\theta)},\end{aligned} \tag{168}$$

using the identity

$$L^2(\theta) - c_1 r^2 = c_1c_2K_1(\theta).$$

Since $t_1-1>0$ and $\varphi_1''(t_1)<0$, we can take $\omega_1 = 0$. Together with $\psi_1(t_1) = (t_1-1)^{-1} = c_2(1+\theta)/c_1$, we obtain from (JO34) and (84)

$$g_h(z_0) \sim \sqrt{\frac{c_1}{r^2}}\,|\varphi_1''(t_1)|^{-1/2}\psi_1(t_1) = \frac{\sqrt{c_1(1-c_1)}l(\theta)}{r\sqrt{K_1(\theta)}}$$

The product P then reduces to the first expression in (166).

CCA. We show that $P_{\mathrm{CCA}} = P_{\mathrm{REG}}(1+o(1))$. From Table 7 and (JO28), we have

$$\frac{\theta^{-1}g_{c,C}}{\theta^{-1}g_{c,R}} = \frac{r\sqrt{c_1+c_2}}{c_1 l(\theta)}, \qquad \sqrt{\frac{D_{2,C}}{D_{2,R}}} = \sqrt{\frac{c_1l(\theta)}{c_1+c_2}\frac{K_2}{K_1}}, \qquad \frac{\psi_2(t_2)}{\psi_1(t_1)} = \frac{1}{1-\eta_2 t_2} = \frac{c_1 l(\theta)}{r^2}.$$

Multiplying these ratios and referring to (JO34), we obtain

$$\frac{P_{CCA}}{P_{REG}} \sim \left|\frac{\varphi_1''(t_1)}{\varphi_2''(t_2)}\right|^{1/2} \left[\frac{c_1 l}{r^2}\frac{K_2(\theta)}{K_1(\theta)}\right]^{1/2} e^{i\omega_2/2}. \tag{169}$$

We now compare $\varphi_2''(t_2)$ to $\varphi_1''(t_1)$, recalling that $t_2 = t_1$. First, from (86),

$$\varphi_2''(t) = -\frac{\kappa\eta_2^2}{(1-\eta_2 t_2)^2} + \varphi_1''(t).$$

In particular, $\varphi_2''(t_2) < 0$ and, as with ω_1, also $\omega_2 = 0$. From (88), we evaluate

$$-\frac{\kappa\eta_2^2}{(1-\eta_2 t_2)^2} = -\frac{c_2^3(1+\theta)^2}{c_1^2(1-c_1)}\frac{(c_1+\theta)^2}{l^2 r^2},$$

so that from (168),

$$\frac{\varphi_2''(t_2)}{\varphi_1''(t_2)} = 1 + \frac{c_2(c_1+\theta)^2}{r^2 K_1(\theta)} = \frac{c_1 l(\theta) K_2(\theta)}{r^2 K_1(\theta)},$$

where the second identity follows after some algebra. The latter display and (169) show that $P_{\mathrm{CCA}} = P_{\mathrm{REG}}(1+o(1))$.

6.2. *Proof of Theorem JO12 (Gaussian process limit). Some general considerations*

Almost sure continuity of $\ln L(\theta;\Lambda)$**.** Let $\varepsilon > 0$ be a fixed small number. First, let us show that $\ln L(\theta;\Lambda)$ are continuous functions of $\theta \in [0, \bar\theta - \varepsilon]$ for each of the six cases under study. Recall equation (JO6)

$$L^{(Case)}(\theta;\Lambda) = \alpha(\theta)\, {}_{\mathsf{p}}F_{\mathsf{q}}(a, b; \Psi, \Lambda), \tag{170}$$

where Ψ is a p-dimensional matrix $\operatorname{diag}\{\Psi_{11}, 0, ..., 0\}$, and the values of Ψ_{11}, $\alpha(\theta)$, p, q, a, and b are as given in Table 8. Consider the series representation

$$\begin{aligned} {}_{\mathsf{p}}F_{\mathsf{q}}(a, b; \Psi, \Lambda) &= \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{\kappa\vdash k}\frac{(a_1)_\kappa \dots (a_{\mathsf{p}})_\kappa}{(b_1)_\kappa \dots (b_{\mathsf{q}})_\kappa}\frac{C_\kappa(\Psi)C_\kappa(\Lambda)}{C_\kappa(I_p)} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\frac{(a_1)_k \dots (a_{\mathsf{p}})_k}{(b_1)_k \dots (b_{\mathsf{q}})_k}\frac{\Psi_{11}^k C_k(\Lambda)}{C_k(I_p)}, \end{aligned}$$

where the second equality follows from the fact that $C_\kappa(\Psi) = 0$ unless partition $\kappa \vdash k$ is trivial, that is $\kappa = k$, in which case $C_\kappa(\Psi) = \Psi_{11}^k$ (see definition 7.2.1 iii in Muirhead (1982)). James (1968) shows that the coefficients of zonal polynomials are positive. Therefore, for non-negative Ψ_{11} and λ_j, $j = 1, ..., p$, we have

$$0 \le \frac{\Psi_{11}^k C_k(\Lambda)}{C_k(I_p)} \le (\Psi_{11}\lambda_1)^k.$$

This implies that ${}_{\mathsf{p}}F_{\mathsf{q}}(a, b; \Psi, \Lambda)$ is an analytic function of $\theta \in [0, \bar\theta - \varepsilon]$ and ${}_{\mathsf{p}}F_{\mathsf{q}}(a, b; \Psi, \Lambda) \ge 1$ (the first term in the expansion of ${}_{\mathsf{p}}F_{\mathsf{q}}(a, b; \Psi, \Lambda)$ is 1) when $\mathsf{p} \le \mathsf{q}$, that is for SMD, PCA, REG_0, and REG cases. For SigD and CCA, ${}_{\mathsf{p}}F_{\mathsf{q}}(a, b; \Psi, \Lambda)$ is an analytic function of θ in the domain

$$\Psi_{11}\lambda_1 < 1.$$

Case	${}_{\mathsf{p}}F_{\mathsf{q}}$	$\alpha(\theta)$	a	b	Ψ_{11}
SMD	${}_0F_0$	$\exp\left(-p\theta^2/4\right)$	_	_	$\theta p/2$
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	${}_2F_1$	$(1+n_1\theta/n)^{-n/2}$	$(n/2,n/2)$	$n_1/2$	$\theta n_1^2/\left(n_2^2+n_2n_1(1+\theta)\right)$

TABLE 8

Parameters of the JO's explicit expression (JO6) for the likelihood ratios. Here $n \equiv n_1+n_2$.

But for SigD and CCA λ_j are solutions to

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

and hence, with probability 1, $\lambda_1 \leq n_2/n_1$ because H and E are positive definite. Therefore, for SigD we have

$$\Psi_{11}\lambda_1=\frac{\theta n_1}{n_2(1+\theta)}\lambda_1 \leq \frac{\theta}{1+\theta}<1$$

for any $\theta \in [0,\bar{\theta}-\varepsilon]$, and for CCA we have

$$\Psi_{11}\lambda_1=\frac{\theta n_1^2}{n_2^2+n_2n_1(1+\theta)}\lambda_1 \leq \frac{\theta n_1}{n_2+n_1(1+\theta)}<1$$

for any $\theta \in [0,\bar{\theta}-\varepsilon]$. Thus, ${}_{\mathsf{p}}F_{\mathsf{q}}(a,b;\Psi,\Lambda)$ is an analytic function of $\theta \in [0,\bar{\theta}-\varepsilon]$ and ${}_{\mathsf{p}}F_{\mathsf{q}}(a,b;\Psi,\Lambda) \geq 1$ for all six cases that we consider. Using (170) we conclude that $\ln L(\theta;\Lambda)$ are continuous functions of $\theta \in [0,\bar{\theta}-\varepsilon]$ with probability one. In particular (see Bosq (2000) p. 22) $\ln L(\theta;\Lambda)$ can be interpreted as random element of the space $C_{[0,1-\varepsilon]}$ of continuous functions on $[0,1-\varepsilon]$ equipped with the supremum norm.

Reduction to a linear spectral statistic. By Theorem JO11 we have

$$\ln L(\theta;\Lambda)=-\tfrac{1}{2}\Delta_p(\theta)+\frac{1}{2}\ln\left(1-[\delta_p(\theta)]^2\right)+o_{\mathrm{P}}(1), \tag{171}$$

where

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

and

$$\Delta_p(\theta)=p\int \ln(z_0-\lambda)\,\mathrm{d}\left(\hat{F}(\lambda)-F_{\mathbf{c}}(\lambda)\right) \tag{172}$$

with

$$z_0 = \begin{cases} \theta + 1/\theta & \text{for SMD} \\ (1+\theta)(\theta + c_1)/\theta & \text{for PCA and REG}_0 \\ (1+\theta)(\theta+c_1)/[\theta l(\theta)] & \text{for SigD, REG, and CCA} \end{cases} \tag{173}$$

and $F_{\mathbf{c}}$ equals the semicircle distribution for SMD, the Marchenko-Pastur distribution for PCA and REG_0, and the scaled Wachter distribution for SigD, REG, and CCA. As explained in JO, the statistic $\Delta_p(\theta)$ should be interpreted as zero whenever $z_0 \leq \lambda_1$.

Since both $\ln L(\theta; \Lambda)$ and $\Delta_p(\theta)$ are random element of $C_{[0,1-\varepsilon]}$, $o_{\mathrm{P}}(1)$ is also a random element of $C_{[0,1-\varepsilon]}$, and $\|o_{\mathrm{P}}(1)\| \xrightarrow{\mathrm{P}} 0$. Therefore by the standard argument, see for example Theorem 3.1 of Billingsley (1999), p. 27, the weak limits of $\ln L(\theta;\Lambda)$ and of $-\frac{1}{2}\Delta_p(\theta) + \frac{1}{2}\ln\left(1 - [\delta_p(\theta)]^2\right)$ coincide. Note that $\frac{1}{2}\ln\left(1-[\delta_p(\theta)]^2\right)$ is converging in the space $C_{[0,1-\varepsilon]}$ to

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

Therefore, we only need to establish the weak convergence of $\Delta_p(\theta)$. There are two facts to be established. First, the tightness of $\Delta_p(\theta)$, and second, the convergence of its finite dimensional distributions.

Tightness of $\Delta_p(\theta)$. There are three cases to consider: $F_{\mathbf{c}}$ is the semicircle, the Marchenko-Pastur, and the Wachter distribution. Whether the Marchenko-Pastur $F_{\mathbf{c}}$ corresponds to PCA or REG_0 cases is of no importance because we consider the tightness under the null hypothesis so that $\hat{F}$ is the same for PCA and REG_0. Similarly, the differences between SigD, REG and CCA cases are of no importance here.

Tightness, Semi-circle $F_{\mathbf{c}}$. The tightness of $\Delta_p(\theta)$ in this case is a direct consequence of Theorem 1.1 of Bai and Yao (2005).

Tightness, Marchenko-Pastur $F_{\mathbf{c}}$. Following Bai and Silverstein (2004), let us represent the linear spectral statistic $\Delta_p(\theta)$ in the following form

$$\Delta_p(\theta) = -\frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{R}} \ln(z_0 - z)\, p\left[\hat{s}(z) - s_{\mathbf{c}}(z)\right] \mathrm{d}z,$$

where $\mathcal{R}$ is contour that does not intersect the supports of $\hat{F}$ and $F_{\mathbf{c}}$ and does not encircle z_0. Here

$$\hat{s}(z) = \int (\lambda - z)^{-1} \mathrm{d}\hat{F}(\lambda) \text{ and } s_{\mathbf{c}}(z) = \int (\lambda - z)^{-1} \mathrm{d}F_{\mathbf{c}}(\lambda).$$

With asymptotically negligible probability the above requirements for $\mathcal{R}$ are impossible to satisfy. We will therefore condition our arguments on the high probability event that ensures the existence of required $\mathcal{R}$.

Precisely, recall that the supports of $F_{\boldsymbol{\gamma}}$ and $F_{\mathbf{c}}$ are given by

$$\begin{aligned} [\beta_-, \beta_+] &= \left[(1-\sqrt{\gamma_1})^2, (1+\sqrt{\gamma_1})^2\right] \text{ and} \\ [b_-, b_+] &= \left[(1-\sqrt{c_1})^2, (1+\sqrt{c_1})^2\right], \end{aligned}$$

respectively, and the threshold $\bar{\theta}$ equals $\sqrt{\gamma_1}$. Furthermore, $\mathbf{c} \to \boldsymbol{\gamma}$. Using these facts and the definition of z_0, it is straightforward to verify that there exists $\eta > 0$ that depends on ε such that

$$\min_{\theta \in [0, \bar{\theta} - \varepsilon]} (z_0 - \beta_+ - \eta) > 0$$

for all sufficiently large n_1, n_2, p along the sequence $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$. Further, note that $\lambda_1 \overset{a.s.}{\to} \beta_+$ and $\lambda_p \overset{a.s.}{\to} \beta_-$ when $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$.

Consider the event

$$Q_p = \{\max\{\lambda_1, b_+\} \leq \beta_+ + \eta/2 < z_0 - \eta/2, \ \min\{\lambda_p, b_-\} \geq \beta_- - \eta/2\}. \tag{174}$$

The discussion above implies that

$$\lim_{p \to \infty} \Pr\{Q_p\} = 1. \tag{175}$$

Let $\mathcal{R}$ be the rectangular contour with the vertices at $(\beta_+ + \eta) \pm \mathrm{i}v$ and $(\beta_- - \eta) \pm \mathrm{i}v$ for an arbitrary fixed positive v. Conditional on the event Q_p, $\mathcal{R}$ does not intersect the supports of $\hat{F}$ and $F_{\mathbf{c}}$ and does not encircle z_0 as required. Since $\Pr\{Q_p\} \to \infty$, it is sufficient to establish the tightness of $\Delta_p(\theta)$ conditional on Q_p. Therefore, in what follows we shall assume that Q_p holds.

Let $\mathcal{C}$ be the part of $\mathcal{R}$ that lies in the upper half complex plane. Then

$$\Delta_p(\theta) = -\frac{1}{\pi} \operatorname{Im} \int_{\mathcal{C}} \ln(z_0 - z)\, p\, [\hat{s}(z) - s_{\mathbf{c}}(z)]\, \mathrm{d}z.$$

Since the mapping

$$f(z) \mapsto g(\theta) = -\frac{1}{\pi} \operatorname{Im} \int_{\mathcal{C}} \ln(z_0 - z)\, f(z) \mathrm{d}z$$

is a continuous mapping from the space $C_{\mathcal{C}}$ of the complex-valued continuous functions on $\mathcal{C}$ (with the supremum norm) to the space $C_{[0,1-\varepsilon]}$, the tightness of $\Delta_p(\theta)$ would follow from that of

$$M_p(z) \equiv p\, [\hat{s}(z) - s_{\mathbf{c}}(z)].$$

As in Bai and Silverstein (2004) p. 561, choose sequence $\{\varepsilon_p\}$ such that $\varepsilon_p \to 0$ as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$ and

$$\varepsilon_p \geq p^{-\alpha}$$

for some $\alpha \in (0, 1)$. Further, let

$$\begin{aligned} \mathcal{C}_u &= \{x + \mathrm{i}v : x \in [\beta_- - \eta, \beta_+ + \eta]\}, \\ \mathcal{C}_l &= \left\{(\beta_- - \eta) + \mathrm{i}y : y \in \left[p^{-1}\varepsilon_p, v\right]\right\}, \\ \mathcal{C}_r &= \left\{(\beta_+ + \eta) + \mathrm{i}y : y \in \left[p^{-1}\varepsilon_p, v\right]\right\}, \end{aligned}$$

and let $\mathcal{C}_p = \mathcal{C}_l \cup \mathcal{C}_u \cup \mathcal{C}_r$. Define the process $\hat{M}_p(z)$ on $\mathcal{C}$ as follows

$$\hat{M}_p(z) = \begin{cases} M_p(z) & \text{for } z \in \mathcal{C}_p \\ M_p(\beta_+ + \eta + \mathrm{i}p^{-1}\varepsilon_p) & \text{for } z = \beta_+ + \eta + \mathrm{i}y, \ y \in [0, p^{-1}\varepsilon_p] \\ M_p(\beta_- - \eta + \mathrm{i}p^{-1}\varepsilon_p) & \text{for } z = \beta_- - \eta + \mathrm{i}y, \ y \in [0, p^{-1}\varepsilon_p] \end{cases}.$$

Note that

$$\hat{M}_p(z) = p\, [\hat{s}(z) - s_{\mathbf{c}}(z)] + o_{\mathrm{P}}(1),$$

where $o_{\mathrm{P}}(1)$ is uniform over $z \in \mathcal{C}$. Indeed, for any $z \in \mathcal{C}_p$ we have

$$\hat{M}_p(z) = p\left[\hat{s}(z) - s_{\mathbf{c}}(z)\right],$$

whereas by the definition of $\hat{s}(z)$ and (174)

$$\sup_{y\in[0,p^{-1}\varepsilon_p]} p\left|\hat{s}(\beta_\pm \pm \eta + \mathrm{i}y) - \hat{s}(\beta_\pm \pm \eta + \mathrm{i}p^{-1}\varepsilon_p)\right| \leq p\frac{p^{-1}\varepsilon_p}{(\eta/2)^2} \to 0,$$

and similarly

$$\sup_{y\in[0,p^{-1}\varepsilon_p]} p\left|s_{\mathbf{c}}(\beta_\pm \pm \eta + \mathrm{i}y) - s_{\mathbf{c}}(\beta_\pm \pm \eta + \mathrm{i}p^{-1}\varepsilon_p)\right| \mathbf{1}\left\{Q_p\right\} \leq p\frac{p^{-1}\varepsilon_p}{(\eta/2)^2} \to 0.$$

Therefore, it is sufficient to prove the tightness of $\hat{M}_p(\cdot)$ as a sequence of random elements of $C_{\mathcal{C}}$. Lemma 1.1 of Bai and Silverstein (2004) establishes this result along with the weak convergence of $\hat{M}_p(\cdot)$ to a Gaussian process.

Tightness, Wachter $F_{\mathbf{c}}$. We shall base our arguments on the results established in Zheng (2012). He establishes a CLT for linear spectral statistics of multivariate F and *Beta* matrices via representing those statistics in the form of a contour integral that involves a process related to $M_p(z)$ (see the previous section). The CLT follows from his proving the convergence of the process to a Gaussian process.

In contrast to JO, whose attention is focused on the eigenvalues of $H\left(E + \frac{n_1}{n_2}H\right)^{-1}$, Zheng's (2012) primary focus is on the eigenvalues of HE^{-1}. Let $\hat{F}$ and $\hat{G}$ be the empirical distributions of the eigenvalues of $H\left(E + \frac{n_1}{n_2}H\right)^{-1}$ and HE^{-1}, respectively. If x is an eigenvalue of HE^{-1}, then $x\left(1 + c_2x/c_1\right)^{-1}$ is an eigenvalue of $H\left(E + \frac{n_1}{n_2}H\right)^{-1}$, and thus

$$\hat{G}(x) = \hat{F}\left(\frac{x}{1 + c_2x/c_1}\right).$$

A similar equality holds for the corresponding limiting distributions $G_{\mathbf{c}}$ and $F_{\mathbf{c}}$. Therefore,

$$\begin{aligned}\Delta_p(\theta) &\equiv p\int \ln(z_0 - \lambda)\,\mathrm{d}\left(\hat{F}(\lambda) - F_{\mathbf{c}}(\lambda)\right)\\ &= p\int \ln\left(z_0 - \frac{x}{1 + c_2x/c_1}\right)\mathrm{d}\left(\hat{G}(x) - G_{\mathbf{c}}(x)\right).\end{aligned}$$

Denote the Stieltjes transform of $\hat{G}$ as $\hat{m}(z)$ and that of $G_{\mathbf{c}}$ as $m_{\mathbf{c}}(z)$. Then, similarly to the Marchenko-Pastur case, we have

$$\Delta_p(\theta) = -\frac{1}{2\pi\mathrm{i}}\oint_{\mathcal{R}} \ln\left(z_0 - \frac{z}{1 + c_2z/c_1}\right) p\left[\hat{m}(z) - m_{\mathbf{c}}(z)\right]\mathrm{d}z,$$

where $\mathcal{R}$ is contour that does not intersect the supports of $\hat{G}$ and $G_{\mathbf{c}}$ and does not encircle $z_0/\left(1 - c_2z_0/c_1\right)$. As above, the existence of such a contour requires conditioning on a large probability event, which we shall assume.

Zheng (2012) pp. 467–470 sketches a proof of the weak convergence of $p\left[\hat{m}(z) - m_{\mathbf{c}}(z)\right]$. Such a weak convergence implies the tightness, which in its turn implies the tightness of $\Delta_p(\theta)$.

For the reader's convenience, we provide here a brief description of the main steps in Zheng's proof. The proof is based on the decomposition

$$p\left[\hat{m}(z)-m_{\mathbf{c}}(z)\right]=p\left[\hat{m}(z)-m_{\mathbf{c}}^{(E)}(z)\right]+p\left[m_{\mathbf{c}}^{(E)}(z)-m_{\mathbf{c}}(z)\right],$$

where $m_{\mathbf{c}}^{(E)}(z)$ is the Stieltjes transform of $G_{\mathbf{c}}^{(E)}$, the limiting spectral distribution (as $\mathbf{n},p\to_{\mathbf{c}}\infty$) of HA_p^{-1} where the empirical spectral distribution of symmetric positive definite matrix A_p converges to that of E as $\mathbf{n},p\to_{\mathbf{c}}\infty$. First, Zheng establishes the weak convergence of $p\left[\hat{m}(z)-m_{\mathbf{c}}^{(E)}(z)\right]$ conditional on $\{E,p=1,2...\}$ by appealing to Lemma 1.1 of Bai and Silverstein (2004). Since the limiting process does not depend on $\{E,p=1,2...\}$, the unconditional convergence also follows. Next, Zheng represents $p\left[m_{\mathbf{c}}^{(E)}(z)-m_{\mathbf{c}}(z)\right]$ as a product of a continuous function of z that converges in $C_{\mathcal{R}}$ and the term $p\left[\hat{m}_E\left(-\underline{m}_{\mathbf{c}}(z)\right)-m_{c_2}\left(-\underline{m}_{\mathbf{c}}(z)\right)\right]$, where $\hat{m}_E$ is the Stieltjes transform of the empirical spectral distribution of E, m_{c_2} is that of the corresponding limiting distribution as $\mathbf{n},p\to_{\mathbf{c}}\infty$, and $\underline{m}_{\mathbf{c}}$ is defined via the Stieltjes transform $m_{\mathbf{c}}$ of $G_{\mathbf{c}}$ by

$$\underline{m}_{\mathbf{c}}(z)=-\frac{1-c_1}{z}+c_1m_{\mathbf{c}}(z).$$

Then, she points out that $-\underline{m}_{\mathbf{c}}(z)$ converges to $-\underline{m}_{\boldsymbol{\gamma}}(z)$, which is defined analogously with $\mathbf{c}$ replaced by $\boldsymbol{\gamma}$. Function $-\underline{m}_{\boldsymbol{\gamma}}(z)$ transforms $\mathcal{R}$ to a contour encircling the support of the limiting spectral distribution of E. Zheng appeals to Lemma 1.1 of Bai and Silverstein (2004) to establish the weak convergence of $p\left[\hat{m}_E(z)-m_{c_2}(z)\right]$ as a random continuous function on such a contour. Zheng's proof omits some details, probably for the sake of saving the space. For example, she does not mention that to be able to view $p\left[\hat{m}(z)-m_{\mathbf{c}}^{(E)}(z)\right]$ and $p\left[m_{\mathbf{c}}^{(E)}(z)-m_{\mathbf{c}}(z)\right]$ as continuous random functions on $\mathcal{R}$, a conditioning on some event of increasing probability in needed. Having a detailed proof would be useful, but requires a separate research effort.

Finite dimensional convergence. The convergence of the finite dimensional distributions to Gaussian distributions follow from Theorem 1.1 of Bai and Yao (2005) for the semicircle $F_{\mathbf{c}}$, from Theorem 1.1 of Bai and Silverstein (2004) for the Marchenko-Pastur $F_{\mathbf{c}}$, and from Theorem 4.1 of Zheng (2012) for the Wachter $F_{\mathbf{c}}$. We now use the results in the above mentioned papers to compute the means and covariance matrices of the asymptotic finite-dimensional distributions of $\Delta_p(\theta)$.

Finite dimensional asymptotics, Semi-circle $F_{\mathbf{c}}$. Recall that $z_0=\theta+1/\theta$, we obtain

$$\Delta_p(\theta)=p\int\ln\left(\theta^2-\lambda\theta+1\right)\mathrm{d}\left(\hat{F}(\lambda)-F_{\mathbf{c}}(\lambda)\right).$$

Theorem 1.1 Bai and Yao (2005) implies that the random vector $(\Delta_p(\theta_1),...,\Delta_p(\theta_k))$ with $\theta_i\in[0,\bar{\theta}-\varepsilon]$ converges in distribution to a Gaussian vector $(\mathcal{D}(\theta_1),...,\mathcal{D}(\theta_k))$ with

$$\mathbb{E}\mathcal{D}(\theta_i)=\frac{1}{4}\left[\ln\left[(1-\theta_i)^2\right]+\ln\left[(1+\theta_i)^2\right]\right]-\tfrac{1}{2}\tau_0(\theta) \tag{176}$$

and

$$Cov\left(\mathcal{D}(\theta_i),\mathcal{D}(\theta_j)\right)=2\sum_{l=1}^{\infty}l\tau_l(\theta_i)\tau_l(\theta_j), \tag{177}$$

where

$$\tau_l(\theta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\left(1+\theta^2-2\theta\cos\varphi\right)\cos(l\varphi)\,\mathrm{d}\varphi.$$

LEMMA 16. *For any θ, such that $|\theta| < 1$, and any integer $l > 0$, we have $\tau_l(\theta) = -\theta^l/l$ and $\tau_0(\theta) = 0$.*

Proof: Changing the variable of integration from φ to $z = e^{\mathrm{i}\varphi}$, we obtain

$$\tau_l(\theta) = \frac{1}{2\pi \mathrm{i}} \oint \ln\left[(1-\theta z)\left(1-\theta z^{-1}\right)\right] z^{l-1}\mathrm{d}z,$$

where the contour integral is taken over the counter-clockwise oriented unit circle. Representing the logarithm of a product as the sum of logarithms, we obtain

$$\tau_l(\theta) = \frac{1}{2\pi \mathrm{i}} \oint \ln(1-\theta z) z^{l-1}\mathrm{d}z + \frac{1}{2\pi i}\oint \ln\left[1-\theta z^{-1}\right] z^{l-1}\mathrm{d}z.$$

Since, for $|\theta| < 1$, $\ln(1-\theta z)$ is analytic in the unit circle and equal to zero at $z = 0$, we have

$$\frac{1}{2\pi \mathrm{i}} \oint \ln(1-\theta z) z^{l-1}\mathrm{d}z = 0$$

for any integer $l \geq 0$. Hence,

$$\tau_l(\theta) = \frac{1}{2\pi \mathrm{i}} \oint \ln\left[1-\theta z^{-1}\right] z^{l-1}\mathrm{d}z.$$

Changing the variable of integration from z to $\zeta = z^{-1}$, and noting that $\mathrm{d}z/z = -\mathrm{d}\zeta/\zeta$, we get

$$\tau_l(\theta) = \frac{1}{2\pi \mathrm{i}} \oint \ln[1-\theta\zeta]\,\zeta^{-l-1}\mathrm{d}\zeta.$$

On the other hand, for $|\zeta| \leq 1$, we have the following power series expansion

$$\ln[1-\theta\zeta] = -\sum_{j=1}^{\infty}\frac{\theta^j}{j}\zeta^j.$$

Thus, by Cauchy's residue theorem, $\tau_l(\theta) = -\theta^l/l$ for $l > 0$ and $\tau_0(\theta) = 0$. □

Lemma 16 together with (176) and (177) yield

$$\mathbb{E}\mathcal{D}(\theta_i) = \tfrac{1}{2}\ln\left(1-\theta_i^2\right)$$

and

$$Cov\left(\mathcal{D}(\theta_i), \mathcal{D}(\theta_j)\right) = -2\ln\left(1-\theta_i\theta_j\right).$$

Finite dimensional asymptotics, Marchenko-Pastur $F_{\mathbf{c}}$. For PCA and REG_0 the finite dimensional distributions of $\Delta_p(\theta)$ are derived in Lemma 12 of Onatski et al (2013). They show that the random vector $(\Delta_p(\theta_1), ..., \Delta_p(\theta_k))$ with $\theta_i \in [0, \bar{\theta}-\varepsilon]$ converges in distribution to a Gaussian vector $(\mathcal{D}(\theta_1), ..., \mathcal{D}(\theta_k))$ with

$$\mathbb{E}\mathcal{D}(\theta_i) = \tfrac{1}{2}\ln\left(1-\theta_i^2/\gamma_1\right)$$

and

$$Cov\left(\mathcal{D}(\theta_i), \mathcal{D}(\theta_j)\right) = -2\ln\left(1-\theta_i\theta_j/\gamma_1\right).$$

Finite dimensional asymptotics, Wachter $F_{\mathbf{c}}$. Let

$$\hat{G}(x) = \hat{F}\left(\frac{x}{1 + c_2 x/c_1}\right) \text{ and } G_{\mathbf{c}}(x) = F_{\mathbf{c}}\left(\frac{x}{1 + c_2 x/c_1}\right).$$

Then

$$\Delta_p(\theta) = p \int \ln\left(z_0 - \frac{x}{1 + c_2 x/c_1}\right) \mathrm{d}\left(\hat{G}(x) - G_{\mathbf{c}}(x)\right). \tag{178}$$

Recall that

$$z_0 = \frac{(1+\theta)(\theta + c_1)}{\theta(1 + (1+\theta) c_2/c_1)}.$$

Let

$$z_{\gamma 0} = \frac{(1+\theta)(\theta + \gamma_1)}{\theta(1 + (1+\theta)\gamma_2/\gamma_1)}. \tag{179}$$

Since $z_0 \to z_{\gamma 0}$ and $\mathbf{c} \to \gamma$ as $\mathbf{n}, p \to_{\boldsymbol{\gamma}} \infty$, the asymptotic distribution of the random vector $(\Delta_p(\theta_1), ..., \Delta_p(\theta_k))$ must be the same as that of $(\Delta_{\gamma p}(\theta_1), ..., \Delta_{\gamma p}(\theta_k))$, where

$$\Delta_{\gamma p}(\theta) = p \int \ln\left(z_{\gamma 0} - \frac{x}{1 + \gamma_2 x/\gamma_1}\right) \mathrm{d}\left(\hat{G}(x) - G_{\mathbf{c}}(x)\right). \tag{180}$$

This can be formally shown by considering the representation

$$\Delta_p(\theta) - \Delta_{\gamma p}(\theta) = -\frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{R}} \ln\left[\frac{z_0 - \frac{z}{1+c_2 z/c_1}}{z_{\gamma 0} - \frac{z}{1+\gamma_2 z/\gamma_1}}\right] p\left[\hat{m}(z) - m_{\mathbf{c}}(z)\right] \mathrm{d}z$$

(see subsection "Tightness, Wachter $F_{\mathbf{c}}$") and using the convergence of $p\left[\hat{m}(z) - m_{\mathbf{c}}(z)\right]$ established by Zheng (2012) to demonstrate that

$$\Delta_p(\theta) - \Delta_{\gamma p}(\theta) = o_{\mathrm{P}}(1). \tag{181}$$

Theorem 3.1 of Zheng (2012) implies that the random vector $(\Delta_{\gamma p}(\theta_1), ..., \Delta_{\gamma p}(\theta_k))$ with $\theta_i \in [0, \bar{\theta} - \varepsilon]$ converges in distribution to a Gaussian vector $(\mathcal{D}(\theta_1), ..., \mathcal{D}(\theta_k))$. Let us find the asymptotic mean and covariances.

Equations (179), (180) and some elementary algebra yield

$$\Delta_{\gamma p}(\theta) = \Delta_{\gamma p}^{(1)}(\theta) - \Delta_{\gamma p}^{(2)}(\theta), \tag{182}$$

where

$$\Delta_{\gamma p}^{(1)}(\theta) = p \int \ln\left(\theta + \gamma_1 + \left(\gamma_2 - \frac{\theta}{1+\theta}\right) x\right) \mathrm{d}\left(\hat{G}(x) - G_{\mathbf{c}}(x)\right),$$

and

$$\Delta_{\gamma p}^{(2)}(\theta) = p \int \ln(\gamma_1/\gamma_2 + x) \, \mathrm{d}\left(\hat{G}(x) - G_{\mathbf{c}}(x)\right).$$

Note that both $\Delta_{\gamma p}^{(1)}(\theta)$ and $\Delta_{\gamma p}^{(2)}(\theta)$ have form

$$Y_{ab} = p \int \ln(a + bx) \, \mathrm{d}\left(\hat{G}(x) - G_{\mathbf{c}}(x)\right).$$

For $a, b, a', b' > 0$, Zheng (2012), Example 4.1 proves that $(Y_{ab}, Y_{a'b'})$ converge to a Gaussian vector $(X_{ab}, X_{a'b'})$ with

$$\mathbb{E}X_{ab} = \tfrac{1}{2}\log\frac{(c^2-d^2)\,\rho^2}{(c\rho-\gamma_2 d)^2} \tag{183}$$

and

$$Cov\,(X_{ab}, X_{a'b'}) = 2\log\frac{cc'}{cc'-dd'}, \tag{184}$$

where $c > d > 0$ satisfy

$$c^2+d^2 = a + b\frac{1+\rho^2}{(1-\gamma_2)^2} \text{ and } cd = \frac{b\rho}{(1-\gamma_2)^2} \tag{185}$$

and $c' > d' > 0$ satisfy

$$c'^2+d'^2 = a' + b'\frac{1+\rho^2}{(1-\gamma_2)^2} \text{ and } c'd' = \frac{b'\rho}{(1-\gamma_2)^2}. \tag{186}$$

A direct inspection reveals that Zheng's proof of (183) and (184) remains valid for any real a, b, a', and b' such that $\log(a+bz)$ and $\log(a'+b'x)$ are analytic in an open domain containing the support of $G_{\boldsymbol{\gamma}}$ as long as there exist real c and d satisfying (185) and real c' and d' satisfying (186) such that $|c| > |d|$ and $|c'| > |d'|$. Such c, d, c' and d' do exist for $Y_{ab} = \Delta^{(1)}_{\gamma p}(\theta)$ and $Y_{a'b'} = \Delta^{(2)}_{\gamma p}(\theta)$. Indeed, the values of a and b for $Y_{ab} = \Delta^{(1)}_{\gamma p}(\theta)$ are

$$a = \theta + \gamma_1 \text{ and } b = \gamma_2 - \frac{\theta}{1+\theta}.$$

The corresponding c and d that satisfy (185) are

$$c = \frac{\rho}{\sqrt{\theta+1}\,(1-\gamma_2)} \text{ and } d = \frac{\gamma_2 - \theta\,(1-\gamma_2)}{\sqrt{\theta+1}\,(1-\gamma_2)}. \tag{187}$$

Since $\gamma_2 < \rho$, $|c|$ is clearly larger than $|d|$ for positive d. For non-positive d, $|c| > |d|$ if and only if

$$\theta\,(1-\gamma_2) - \gamma_2 < \rho,$$

But this inequality folds for any $\theta \in [0, \bar{\theta} - \varepsilon]$ because $\bar{\theta} = (\gamma_2 + \rho)\,/\,(1-\gamma_2)$ (see Table JO3).

Further, the values of a' and b' for $Y_{a'b'} = \Delta^{(2)}_{\gamma p}(\theta)$ are

$$a' = \gamma_1/\gamma_2 \text{ and } b' = 1.$$

The corresponding c' and d' that satisfy (186) are

$$c' = \frac{\rho}{(1-\gamma_2)\sqrt{\gamma_2}} \text{ and } d' = \frac{\sqrt{\gamma_2}}{1-\gamma_2}. \tag{188}$$

Since $\gamma_2 < \rho$, we have $c' > d' > 0$.

Using (182), (183), (187), and (188), we find that

$$\begin{aligned}\mathbb{E}\mathcal{D}(\theta) &= \tfrac{1}{2}\log\frac{\left(c^2-d^2\right)\left(c'\rho-\gamma_2 d'\right)^2}{\left(c\rho-\gamma_2 d\right)^2\left(c'^2-d'^2\right)}\\ &= \tfrac{1}{2}\log\left(1-\frac{\rho^2\theta^2}{\left(\gamma_1+\gamma_2\left(1+\theta\right)\right)^2}\right)\end{aligned}$$

and

$$\begin{aligned}Cov\left(\mathcal{D}\left(\theta_i\right),\mathcal{D}\left(\theta_j\right)\right) &= 2\log\frac{\rho^2}{\rho^2-\left(\gamma_2-\theta_i\left(1-\gamma_2\right)\right)\left(\gamma_2-\theta_j\left(1-\gamma_2\right)\right)}\\ &\quad -2\log\frac{\rho^2}{\rho^2-\left(\gamma_2-\theta_j\left(1-\gamma_2\right)\right)\gamma_2}\\ &\quad -2\log\frac{\rho^2}{\rho^2-\left(\gamma_2-\theta_i\left(1-\gamma_2\right)\right)\gamma_2}\\ &\quad +2\log\frac{\rho^2}{\rho^2-\gamma_2^2}\\ &= -2\log\left(1-\frac{\rho^2\theta_i\theta_j}{\left(\gamma_1+\gamma_2\left(1+\theta_i\right)\right)\left(\gamma_1+\gamma_2\left(1+\theta_j\right)\right)}\right).\end{aligned}$$

7. Concluding remarks.

7.1. *Power of the LR test under multi-spike alternatives.* Consider the likelihood ratio test that rejects the null hypothesis of no spikes when the supremum of $\ln L\left(\theta;\Lambda\right)$ over $\theta\in\left[0,\bar{\theta}-\varepsilon\right]$ is above an asymptotic critical value. In this section, we study the power of such a test in the situation where the rank-one assumption on the alternative is wrong and there are multiple spikes, the highest of which is at least as high as the spike under our rank-one setting.

Intuitively, the power should increase under such a multi-spike alternative because it is "further away" from the null than the one-spike alternative. Below, we confirm this intuition for SMD and PCA cases.

First let us show that, in any of James' cases, the corresponding likelihood ratio test has a monotone acceptance region. That is, the null is accepted if and only if $g\left(\lambda_1,...,\lambda_p\right)<\mathsf{const}$ for a function g which is non-decreasing in each argument. Recall that the likelihood ratio has the following form

$$L\left(\theta;\Lambda\right)=\alpha(\theta)\ {}_{\mathsf{p}}F_{\mathsf{q}}\left(a,b;\Psi,\Lambda\right), \tag{189}$$

where $\Psi=\mathsf{diag}\left\{\Psi_{11},0,...,0\right\}$, $\Lambda=\mathsf{diag}\left\{\lambda_1,...,\lambda_p\right\}$, and the values of $\Psi_{11},\alpha(\theta),a,b,\mathsf{p}$, and q for the different cases are given in Table JO2. As explained in Section 6.2, we have the following expansion

$$L\left(\theta;\Lambda\right)=\alpha(\theta)\sum_{k=0}^{\infty}\frac{1}{k!}\frac{\left(a_1\right)_k...\left(a_{\mathsf{p}}\right)_k}{\left(b_1\right)_k...\left(b_{\mathsf{q}}\right)_k}\frac{\Psi_{11}^kC_k\left(\Lambda\right)}{C_k\left(I_p\right)},$$

where C_k are zonal polynomials. James (1968) shows that zonal polynomials have positive coefficients. Therefore, $C_k\left(\Lambda\right)$ and $L\left(\theta;\Lambda\right)$ are nondecreasing in each λ_j for any fixed $\theta\in\left[0,\bar{\theta}-\varepsilon\right]$. As a consequence, the supremum of $\ln L\left(\theta;\Lambda\right)$ over $\theta\in\left[0,\bar{\theta}-\varepsilon\right]$ is a non-decreasing function in each λ_j.

Next, recall that SMD refers to the problem of testing $H_0:\Phi=0$ against $H_1:\Phi=\theta_1\psi_1\psi_1'$ using the eigenvalues λ_j, $j=1,...,p$ of matrix $X=\Phi+Z/\sqrt{p}$, where Z is a noise matrix from the Gaussian Orthogonal Ensemble. Now suppose that the actual situation corresponds to the alternative

$$H_{\text{mult}} : \Phi = \sum_{j=1}^{r} \theta_j \psi_j \psi_j',$$

where $\theta_1 \geq ... \geq \theta_r > 0$ and $\psi_1, ..., \psi_r$ is a set of orthonormal nuisance vectors. Since Φ under H_{mult} is no smaller than under H_1, the j-th largest eigenvalue of X under H_{mult} is no smaller than under H_1. But as shown above, the likelihood ratio test has a monotone acceptance region. Hence, its power to reject H_0 in favour of H_{mult} is at least as high as its power to reject H_0 in favour of H_1.

Similarly, recall that PCA refers to the problem of testing $H_0 : \Omega = I_p$ against $H_1 : \Omega = I_p + \theta_1 \psi_1 \psi_1'$ using the eigenvalues of YY'/n_1, where $Y = \Omega^{1/2}\varepsilon$ and ε is a $p \times n_1$ matrix with i.i.d. standard normal entries. Suppose that the actual situation corresponds to the alternative

$$H_{\mathsf{mult}} : \Omega = I_p + \sum_{j=1}^{r} \theta_j \psi_j \psi_j'.$$

Note that the non-zero eigenvalues of YY'/n_1 coincide with those of $\varepsilon'\Omega\varepsilon/n_1$. Since Ω under H_{mult} is no smaller than under H_1, the j-th largest eigenvalue of $\varepsilon'\Omega\varepsilon/n_1$ under H_{mult} is no smaller than under H_1. Therefore, using the monotonicity of the acceptance region of the test, we conclude that the power corresponding to H_{mult} is no smaller than that corresponding to H_1.

Unfortunately, for the remaining cases, the above logic does not go through. For example, for SigD, we test $H_0 : \Omega = I_p$ against $H_1 : \Omega = I_p + \theta_1 \psi_1 \psi_1'$ using eigenvalues of $(XX'/n_2)^{-1} (YY'/n_1)$, where $Y = \Omega^{1/2}\varepsilon$ is as above, and X is a $p \times n_2$ matrix with i.i.d. standard normal entries independent from Y. It is conceivable that

$$\varepsilon'\Omega^{1/2} (XX'/n_2)^{-1} \Omega^{1/2}\varepsilon/n_1,$$

as opposed to $\varepsilon'\Omega\varepsilon/n_1$ (cf. the PCA case above), has some of its eigenvalues under H_{mult} smaller than the corresponding eigenvalues under H_1. Of course, on average over the distribution of X, the situation will be exactly the same as in the PCA case. Therefore, although we cannot prove the increase in power, it remains intuitively plausible.

Perlman and Olkin (1980) study the unbiasedness and power monotonicity of tests with monotone acceptance regions in cases that correspond to our REG_0, REG, and CCA. Although they prove the unbiasedness of such tests, the power monotonicity remains a "strong conjecture" (see p. 1329 of their paper). Their Proposition 2.6 (ii) formulates conditions on the likelihood ratio (corresponding to general alternatives) that guarantee the power monotonicity. However, as shown in Richards (2004), these conditions do not hold for likelihoods of form (189), in general. Of course, this does not mean that Perlman and Olkin's conjecture is wrong, it just cannot be established directly via Proposition 2.6.

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 Silverstein, J. W. (2004). CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices. *Annals of Probability* **32** 553–605.

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

[5] Baik, J. and Silverstein, J. W. (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** 1059–1092.

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

[8] Billingsley, P. (1999), *Convergence of Probability Measures.* John Wiley & Sons, Inc.

[9] Bosq, D. (2000). *Linear Processes in Function Spaces.* Springer-Verlag.

[10] Constantine, A. G. (1963). Some noncentral distribution problems in multivariate analysis. *Annals of Mathematical Statistics* **34** 1270–1285.

[11] Dharmawansa, P. and Johnstone, I. M. (2014). Joint density of eigenvalues in spiked multivariate models. *Stat* **3** 240–249.

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

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

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

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

[16] James, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Annals of Mathematical Statistics* **39** 1711–1718.

[17] Johnstone, I. M. and Nadler, B. (2015). Roy's largest root test under rank-one alternatives. *ArXiv:* 1310.6581v3.

[18] Johnstone, I. M. and Onatski, A. (2016). Testing in high-dimensional spiked models. *Manuscript*

[19] Khatri, C. G. (1967). Some distribution problems associated with the characteristic roots of $S_1 S_2^{-1}$. *Annals of Mathematical Statistics* **38** 944–948.

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

[21] Mehta, L. M. (2004). *Random Matrices*, 3rd ed. Elsevier Academic Press.

[22] 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*

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

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

[25] Nadakuditi, R. R. and Silverstein, J. W. (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** 468–480.

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

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

[28] 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.

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

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

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

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

[33] Perlman, M. D. and Olkin, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Annals of Statistics* **8** 1326–1341.

[34] Richards, D. St. P. (2004). Total positivity properties of generalized hypergeometric functions of matrix argument. *Journal of Statistical Physics* **116** 907–922.

[35] Rudin, W. (1987). *Real and Complex Analysis.* WCB/McGraw-Hill.

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

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

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

[39] Watson, G. N. (1944). *A Treatise on the Theory of Bessel Functions*. Cambridge University Press.
[40] 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** 508–516.
[41] 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.

E-mail: imj@stanford.edu

E-mail: ao319@cam.ac.uk