

ASYMPTOTIC POWER OF SPHERICITY TESTS FOR HIGH-DIMENSIONAL DATA

BY ALEXEI ONATSKI, MARCELO J. MOREIRA¹ AND MARC HALLIN²

*University of Cambridge, FGV/EPGE, and Université libre de Bruxelles and
Princeton University*

This paper studies the asymptotic power of tests of sphericity against perturbations in a single unknown direction as both the dimensionality of the data and the number of observations go to infinity. We establish the convergence, under the null hypothesis and contiguous alternatives, of the log ratio of the joint densities of the sample covariance eigenvalues to a Gaussian process indexed by the norm of the perturbation. When the perturbation norm is larger than the *phase transition threshold* studied in Baik, Ben Arous and Pécché [*Ann. Probab.* **33** (2005) 1643–1697] the limiting process is degenerate, and discrimination between the null and the alternative is asymptotically certain. When the norm is below the threshold, the limiting process is nondegenerate, so that the joint eigenvalue densities under the null and alternative hypotheses are mutually contiguous. Using the asymptotic theory of statistical experiments, we obtain asymptotic power envelopes and derive the asymptotic power for various sphericity tests in the contiguity region. In particular, we show that the asymptotic power of the Tracy–Widom-type tests is trivial (i.e., equals the asymptotic size), whereas that of the eigenvalue-based likelihood ratio test is strictly larger than the size, and close to the power envelope.

1. Introduction. Recently, there has been much interest in testing sphericity in a high-dimensional setting. Various tests have been proposed and analyzed in Ledoit and Wolf (2002), Srivastava (2005), Birke and Dette (2005), Schott (2006), Bai et al. (2009), Fisher, Sun and Gallagher (2010), Chen, Zhang and Zhong (2010) and Berthet and Rigollet (2012). In many studies, a distinct interesting alternative to the null of sphericity is the existence of a low-dimensional structure or signal in the data. Detecting such a structure has been the focus of recent studies in various applied fields including population and medical genetics [Patterson, Price and Reich (2006)], econometrics [Onatski (2009, 2010)], wireless communication [Bianchi et al. (2010)], chemometrics [Kritchman and Nadler (2008)] and signal processing [Perry and Wolfe (2010)].

Received August 2011; revised June 2012.

¹Supported by CNPq and the NSF via Grant SES-0819761.

²Supported by the Sonderforschungsbereich “Statistical modelling of nonlinear dynamic processes” (SFB 823) of the Deutsche Forschungsgemeinschaft, and by a Discovery grant of the Australian Research Council.

MSC2010 subject classifications. Primary 62H15, 62B15; secondary 41A60.

Key words and phrases. Sphericity tests, large dimensionality, asymptotic power, spiked covariance, contiguity, power envelope, steepest descent, contour integral representation.

1 Most of the existing sphericity tests are based on the eigenvalues of the sample 1
2 covariance matrix, which constitute the maximal invariant statistic with respect to 2
3 orthogonal transformations of the data. The asymptotic power of such tests de- 3
4 pends on the asymptotic behavior of the sample covariance eigenvalues under the 4
5 alternative hypothesis. When the alternative is a rank- k perturbation of the null, the 5
6 corresponding population covariance matrix is proportional to a sum of the iden- 6
7 tity matrix and a matrix of rank k . Johnstone (2001) calls such a situation “spiked 7
8 covariance.” 8

9 The asymptotic behavior of the sample covariance eigenvalues in “spiked co- 9
10 variance” models of increasing dimension is well studied. Consider the simplest 10
11 case, when $k = 1$. If the largest population covariance eigenvalue is above the 11
12 “phase transition” threshold studied in Baik, Ben Arous and P  ch   (2005), then 12
13 the largest sample covariance eigenvalue remains separated from the rest of the 13
14 eigenvalues, which are asymptotically “packed together as in the support of the 14
15 Marchenko–Pastur density” [Baik and Silverstein (2006)]. Since the largest eigen- 15
16 value separates from the “bulk,” it is easy to detect a signal. 16

17 If the largest population covariance eigenvalue is at or below the threshold, the 17
18 empirical distribution of the sample covariance eigenvalues still converges to the 18
19 Marchenko–Pastur distribution, but the largest sample covariance eigenvalue now 19
20 converges to the upper boundary of its support, both under the null of sphericity 20
21 and the “spiked” alternative [Silverstein and Bai (1995) and Baik and Silverstein 21
22 (2006)]. Hence, the signal detection becomes problematic. At the threshold, the 22
23 null and the alternative hypotheses lead to different asymptotic distributions for the 23
24 centered and normalized largest sample covariance eigenvalue [Bloemendal and 24
25 Vir  g (2011) and Mo (2012)], which implies some asymptotic detection power. 25
26 However, below the threshold, the difference disappears with the joint distribu- 26
27 tion of any finite number of the centered and normalized largest sample covari- 27
28 ance eigenvalues converging to the multivariate Tracy–Widom law under both the 28
29 null and the alternative [Johnstone (2001), Baik, Ben Arous and P  ch   (2005), 29
30 El Karoui (2007) and F  ral and P  ch   (2009)]. 30

31 This similarity in the asymptotic behavior of covariance eigenvalues under the 31
32 null and the alternative prompts Nadakuditi and Edelman (2008) and Nadakuditi 32
33 and Silverstein (2010) to call the transition threshold “the fundamental asymptotic 33
34 limit of sample-eigenvalue-based detection.” They claim that no reliable signal de- 34
35 tection is possible below that limit in the asymptotic sense. This asymptotic impos- 35
36 sibility is also pointed out and discussed in several other recent studies, including 36
37 Patterson, Price and Reich (2006), Hoyle (2008), Nadler (2008), Kritchman and 37
38 Nadler (2009) and Perry and Wolfe (2010). 38

39 In this paper, we analyze the capacity of statistical tests to detect a one- 39
40 dimensional signal with the corresponding population covariance eigenvalue be- 40
41 low the “impossibility threshold,” showing that the terminology “impossibility 41
42 threshold” is overly pessimistic. We establish that the eigenvalue region below 42
43 the threshold actually is the region of mutual contiguity [in the sense of Le Cam 43

(1960)] of the joint distributions of the sample covariance eigenvalues under the null and under the alternative. We obtain the limit in distribution of the log likelihood ratio process inside this contiguity region and derive the asymptotic power envelope for sample-eigenvalue-based detection tests.

The power envelope is larger than size for local alternatives and monotonically tends to one as the signal's population eigenvalue approaches the threshold from below. Hence, the detection of a signal with high asymptotic probability is quite possible even in cases where the largest population covariance eigenvalue is smaller than the threshold, especially when the distance from the threshold remains small.

In the contiguity region, the log likelihood ratio is asymptotically equivalent to a simple statistic related to the Stieltjes transform of the empirical distribution of the sample covariance eigenvalues. The reason the asymptotic behavior of this statistic differs under the null and under the alternative despite the apparent similarity of eigenvalue behaviors just mentioned is that it is not based merely on a contrast between the largest and the rest of the eigenvalues. The information about the presence of the signal exploited by this statistic is hidden in the small deviations of the empirical distribution of the eigenvalues from its Marchenko–Pastur limit.

Let us examine our setting and our results in more detail. Suppose that data consist of n independent observations of p -dimensional real-valued vectors X_t distributed according to the Gaussian law with mean zero and covariance matrix $\sigma^2(I_p + hvv')$, where I_p is the p -dimensional identity matrix, σ and h are scalars and v is a p -dimensional vector with Euclidean norm one. We are interested in the asymptotic power of the tests of the null hypothesis $H_0 : h = 0$ against the alternative $H_1 : h > 0$ based on the eigenvalues of the sample covariance matrix of the data when both n and p go to infinity. The vector v is an unspecified nuisance parameter indicating the direction of the perturbation of sphericity. In contrast to Berthet and Rigollet (2012), who study signal detection in a similar setting where the vector v is sparse, we do not constrain v in any way except normalizing its Euclidean norm to one.

We consider the cases of known and unknown σ^2 . For the sake of brevity, in the rest of this Introduction, we discuss only the case of unknown σ^2 , which, in practice, is also more relevant. Let λ_j be the j th largest sample covariance eigenvalue, let $\mu_j = \lambda_j / (\lambda_1 + \dots + \lambda_p)$ be its normalized version and let $\mu = (\mu_1, \dots, \mu_{m-1})$, where $m = \min(n, p)$. We begin our analysis with a study of the asymptotic properties of the likelihood ratio process $L(h; \mu)$ defined as the ratio of the density of μ when $h \neq 0$ to that when $h = 0$. We represent $L(h; \mu)$ in the form of an integral over a contour in the complex plane and use the Laplace approximation method and recent results from the large random matrix theory to derive an asymptotic expansion of $L(h; \mu)$ as $p, n \rightarrow \infty$ so that $p/n \rightarrow c \in (0, \infty)$, which we throughout abbreviate into $p, n \rightarrow_c \infty$.

We show that, for any \bar{h} such that $0 < \bar{h} < \sqrt{c}$, $\ln L(h; \mu)$ converges in distribution under the null to a Gaussian process $\mathcal{L}(h; \mu)$ on $h \in [0, \bar{h}]$ with $E[\mathcal{L}(h; \mu)] =$

1 $\frac{1}{4}[\ln(1 - c^{-1}h^2) + c^{-1}h^2]$ and $\text{Cov}(\mathcal{L}(h_1; \mu), \mathcal{L}(h_2; \mu)) = -\frac{1}{2}[\ln(1 - c^{-1}h_1h_2) +$ 1
 2 $c^{-1}h_1h_2]$. By Le Cam's first lemma [see [van der Vaart \(1998\)](#), page 88], this implies 2
 3 that the joint distributions of the normalized sample covariance eigenvalues 3
 4 under the null and under the alternative are mutually contiguous for any $h \in [0, \bar{h}]$. 4
 5 We also show that these joint distributions are not mutually contiguous for any 5
 6 $h > \sqrt{c}$. 6

7 Since $\mathcal{L}(h; \mu)$, as a likelihood ratio process, is not of the Gaussian shift type, 7
 8 local asymptotic normality does not hold, and the asymptotic optimality analysis 8
 9 of tests of $H_0: h = 0$ against $H_1: h > 0$ is difficult. However, an asymptotic power 9
 10 envelope is easy to construct using the Neyman–Pearson lemma along with Le 10
 11 Cam's third lemma. We show that, for tests of asymptotic size α , the maximum 11
 12 achievable power against a specific alternative $h = h_1$ is $1 - \Phi[\Phi^{-1}(1 - \alpha) -$ 12
 13 $\sqrt{-\frac{1}{2}(\ln(1 - c^{-1}h_1^2) + c^{-1}h_1^2)}]$, where Φ , as usual, denotes the standard normal 13
 14 distribution function. 14

15 Using our result on the limiting distribution of $\ln L(h; \mu)$ and Le Cam's third 15
 16 lemma, we compute the asymptotic powers of several previously proposed tests of 16
 17 sphericity and of the likelihood ratio (LR) test based on μ . We find that the power 17
 18 of the LR test comes close to the asymptotic power envelope. The LR test outper- 18
 19 forms the test proposed by [John \(1971\)](#) and studied in [Ledoit and Wolf \(2002\)](#), as 19
 20 well as [Srivastava \(2005\)](#) and the test proposed by [Bai et al. \(2009\)](#). The asymp- 20
 21 totic powers of the tests based on the largest sample covariance eigenvalue, such as 21
 22 the tests proposed by [Bejan \(2005\)](#), [Patterson, Price and Reich \(2006\)](#), [Kritchman](#) 22
 23 [and Nadler \(2009\)](#), [Onatski \(2009\)](#), [Bianchi et al. \(2010\)](#) and [Nadakuditi and Sil-](#) 23
 24 [verstein \(2010\)](#), equals the tests' asymptotic size for alternatives in the contiguity 24
 25 region. 25

26 The rest of the paper is organized as follows. Section 2 provides a representation 26
 27 of the likelihood ratio in terms of a contour integral. Section 3 applies Laplace's 27
 28 method to obtain an asymptotic approximation to the contour integral. Section 4 28
 29 uses that approximation to establish the convergence of the log likelihood ratio 29
 30 process to a Gaussian process. Section 5 provides an analysis of the asymptotic 30
 31 power of various sphericity tests and derives the asymptotic power envelope. Sec- 31
 32 tion 6 concludes. Proofs are given in the [Appendix](#); the more technical ones are 32
 33 relegated to the [Supplementary Appendix](#) [[Onatski, Moreira and Hallin \(2013\)](#)]. 33
 34 34

35 **2. Likelihood ratios as contour integrals.** Let X be a $p \times n$ matrix with 35
 36 i.i.d. real Gaussian $N(0, \sigma^2(I_p + hvv'))$ columns. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ be the 36
 37 ordered eigenvalues of $\frac{1}{n}XX'$ and let $\lambda = (\lambda_1, \dots, \lambda_m)$, where $m = \min\{n, p\}$. 37
 38 Finally, let $\mu = (\mu_1, \dots, \mu_{m-1})$, where $\mu_j = \lambda_j/(\lambda_1 + \dots + \lambda_p)$. 38
 39 39

40 As explained in the [Introduction](#), our goal is to study the asymptotic power of 40
 41 the eigenvalue-based tests of $H_0: h = 0$ against $H_1: h > 0$. If σ^2 is known, the 41
 42 model is invariant with respect to orthogonal transformations, and the maximal 42
 43 invariant statistic is λ . Therefore, we consider tests based on λ . If σ^2 is unknown 43

(which, strictly speaking, is what is meant by “sphericity”), the model is invariant with respect to orthogonal transformations and multiplications by nonzero scalars, and the maximal invariant is μ . Hence, we consider tests based on μ . Note that the distribution of μ does not depend on σ^2 , whereas if σ^2 is known, we can always normalize λ dividing it by σ^2 . Therefore, in what follows, we will assume that $\sigma^2 = 1$ without loss of generality.

Let us denote the joint density of $\lambda_1, \dots, \lambda_m$ as $p(\lambda; h)$ and that of μ_1, \dots, μ_{m-1} as $p(\mu; h)$. The following proposition gives explicit formulas for $p(\lambda; h)$ and $p(\mu; h)$.

PROPOSITION 1. *Let $S(r)$ be the $(r - 1)$ -dimensional unit sphere, and let (dx_r) be the invariant measure on $S(r)$ normalized so that the total measure is one. Further, let $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ and $M = \text{diag}(\mu_1, \dots, \mu_p)$. Then*

$$(2.1) \quad p(\lambda; h) = \frac{\gamma(n, p, \lambda)}{(1+h)^{n/2}} \int_{S(p)} e^{(n/2)(h/(1+h))x'_p \Lambda x_p} (dx_p)$$

and

$$(2.2) \quad p(\mu; h) = \frac{\delta(n, p, \mu)}{(1+h)^{n/2}} \times \int_0^\infty y^{(np-2)/2} e^{-ny/2} \int_{S(p)} e^{(n/2)(yh/(1+h))x'_p M x_p} (dx_p) dy,$$

where $\gamma(n, p, \lambda)$ and $\delta(n, p, \mu)$ depend only on n and p , and on λ and μ , respectively.

The spherical integrals in (2.1) and (2.2) can be represented in the form of a confluent hypergeometric function ${}_1F_1$ of matrix argument [Hillier (2001), page 4]. For example, for the integral in (2.1),

$$\int_{S(p)} e^{(n/2)(h/(1+h))x'_p \Lambda x_p} (dx_p) = {}_1F_1\left(\frac{1}{2}, \frac{p}{2}; \frac{n}{2} \frac{h}{1+h} \Lambda\right).$$

Butler and Wood (2002) develop Laplace approximations to functions ${}_1F_1$ but do not analyze the asymptotic behavior of the approximation errors. The next lemma derives an alternative representation of the spherical integrals in Proposition 1. This representation has the form of a contour integral of a single complex variable, and our asymptotic analysis will be based on the Laplace approximation to such an integral.

LEMMA 2. *Let $D = \text{diag}(d_1, \dots, d_r)$, where d_j are arbitrary complex numbers. Further, let \mathcal{K} be a contour in the complex plane starting at $-\infty$, encircling counter-clockwise the points $0, d_1, \dots, d_r$, and going back to $-\infty$. Such a contour is shown in Figure 1. We have*

$$(2.3) \quad \int_{S(r)} e^{x'_r D x_r} (dx_r) = \frac{\Gamma(r/2)}{2\pi i} \oint_{\mathcal{K}} e^s \prod_{j=1}^r (s - d_j)^{-1/2} ds.$$

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

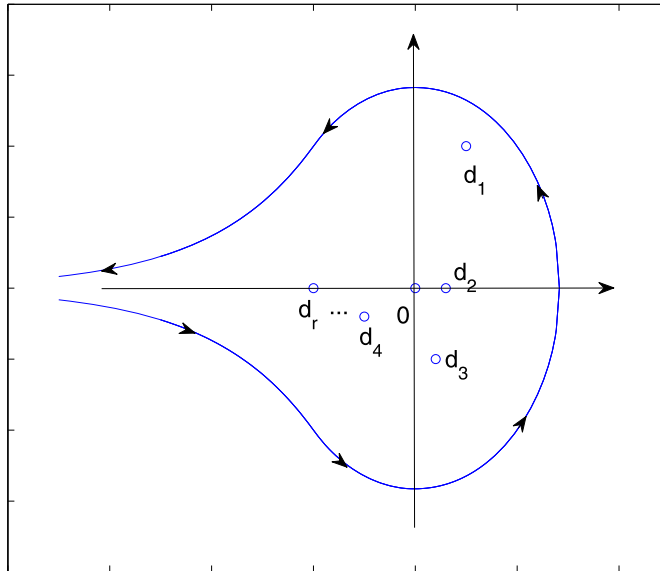


FIG. 1. Contour of integration K in (2.3).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

PROOF. The integral on the left-hand side of (2.3) is the expected value of $\exp(\frac{y_1^2 d_1 + \dots + y_r^2 d_r}{y_1^2 + \dots + y_r^2})$, where y_1, \dots, y_r are independent standard normal random variables. The variables $u_j = \frac{y_j^2}{y_1^2 + \dots + y_r^2}$, $j = 1, \dots, r$, have Dirichlet distribution $\mathcal{D}(k_1, \dots, k_r)$ with parameters $k_1 = \dots = k_r = \frac{1}{2}$. Denoting the expectation operator with respect to such a distribution as $E_{\mathcal{D}}$, we have

$$(2.4) \quad \int_{S(r)} e^{x_r' D x_r} (dx_r) = E_{\mathcal{D}} \exp(u_1 d_1 + \dots + u_r d_r).$$

Now, expanding the exponent in the latter expression into power series and taking expectations term by term yields

$$(2.5) \quad E_{\mathcal{D}} \exp(u_1 d_1 + \dots + u_r d_r) = \sum_{k=0}^{\infty} \frac{E_{\mathcal{D}}(u_1 d_1 + \dots + u_r d_r)^k}{k!}.$$

The Dirichlet average of $(u_1 d_1 + \dots + u_r d_r)^k$ is well studied. By Theorem 3.1 of Dickey (1983),

$$(2.6) \quad \begin{aligned} & E_{\mathcal{D}}[(u_1 d_1 + \dots + u_r d_r)^k] \\ &= \sum_{\substack{m_1, \dots, m_r \geq 0 \\ m_1 + \dots + m_r = k}} \frac{k!}{m_1! \dots m_r!} \frac{(1/2)_{m_1} \dots (1/2)_{m_r}}{(r/2)_k} d_1^{m_1} \dots d_r^{m_r}, \end{aligned}$$

1 where $(k)_s = k(k+1)\cdots(k+s-1)$ is Pochhammer's notation for the shifted
2 factorial.

3 Combining (2.6) with (2.5) and (2.4), we get

$$(2.7) \quad \int_{S(r)} e^{x_r' D x_r} (dx_r) = \sum_{m_1, \dots, m_r \geq 0} \frac{(1/2)_{m_1} \cdots (1/2)_{m_r} d_1^{m_1} \cdots d_r^{m_r}}{(r/2)_{m_1 + \dots + m_r} m_1! \cdots m_r!}$$

$$= {}_r\Phi(1/2, \dots, 1/2; r/2; d_1, \dots, d_r),$$

4 where the last equality is the definition of the confluent form of the Lauricella F_D
5 function, denoted as ${}_r\Phi(\cdot)$. The functions ${}_r\Phi(\cdot)$ were introduced by Erdelyi (1937)
6 and are discussed by Srivastava and Karlsson (1985). In probability and statistics,
7 they were recently used to study the mean of a Dirichlet process [see Lijoi and
8 Regazzini (2004) and references therein].

9 Erdelyi (1937), formula (8,6), establishes the following contour integral repre-
10 sentation of ${}_r\Phi(\cdot)$:

$$(2.8) \quad {}_r\Phi(k_1, \dots, k_r; t; d_1, \dots, d_r)$$

$$= \frac{\Gamma(t)}{2\pi i} \oint_{\mathcal{K}} e^{s^{-t+k_1+\dots+k_r}} \prod_{j=1}^r (s-d_j)^{-k_j} ds.$$

11 Lemma 2 follows from equalities (2.7) and (2.8). \square

12 The contour integral representation given in Lemma 2 has been derived inde-
13 pendently by Mo (2012) and Wang (2012), who use it to study the largest sample
14 covariance eigenvalue when the corresponding population eigenvalue equals the
15 critical threshold or lies above it. Our proof effectively takes advantage of old re-
16 sults of Dickey (1983) and Erdelyi (1937), and thus is different from the proofs in
17 the above mentioned papers.

18 Using Lemma 2 and Proposition 1, we derive contour integral representa-
19 tions for the likelihood ratios $L(h; \lambda) = p(\lambda; h)/p(\lambda; 0)$ and $L(h; \mu) = p(\mu; h)/$
20 $p(\mu; 0)$. The quantity $L(h; \lambda)$ is the likelihood ratio based on λ as opposed to the
21 entire data X . Similarly, $L(h; \mu)$ is the likelihood ratio based on μ .

22 LEMMA 3. *Let \mathcal{K} be a contour in the complex plane that starts at $-\infty$, then
23 encircles counter-clockwise the sample covariance eigenvalues $\lambda_1, \dots, \lambda_p$, and
24 goes back to $-\infty$. In addition, we require that for any $z \in \mathcal{K}$, $\operatorname{Re} z < \frac{1+h}{h}S$, where
25 $\operatorname{Re} z$ denotes the real part of $z \in \mathbb{C}$ and $S = \lambda_1 + \dots + \lambda_p$. Then,*

$$(2.9) \quad L(h; \lambda) = k_1 \left(\frac{2}{n}\right)^{(p-2)/2} \frac{1}{2\pi i} \oint_{\mathcal{K}} e^{(n/2)(h/(1+h))z} \prod_{j=1}^p (z - \lambda_j)^{-1/2} dz$$

1 and

$$\begin{aligned}
 & 2 \\
 & 3 \quad L(h; \mu) = k_2 \frac{S^{(p-2)/2}}{2\pi i} \\
 & 4 \quad (2.10) \\
 & 5 \quad \times \oint_{\mathcal{K}} e^{-((np-p+2)/2)\ln(1-(h/(1+h))(z/S))} \prod_{j=1}^p (z - \lambda_j)^{-1/2} dz, \\
 & 6 \\
 & 7
 \end{aligned}$$

8 where $k_1 = h^{-(p-2)/2}(1+h)^{(p-n-2)/2}\Gamma(p/2)$ and $k_2 = k_1 \frac{\Gamma((np-p+2)/2)}{\Gamma(np/2)}$.

9
10
11 Close inspection of the proof of Lemma 3 reveals that the right-hand side of
12 (2.10) depends on λ only through μ . Although it is possible to express $L(h; \mu)$ as
13 an explicit function of μ , the implicit form given in (2.10) is convenient because it
14 allows us to use similar methods for the asymptotic analysis of the two likelihood
15 ratios.

16 In the next two sections, we perform an asymptotic analysis of $L(h; \lambda)$ and
17 $L(h; \mu)$ that relies on the Laplace approximation of the contour integrals in Lem-
18 ma 3 after those contours have been suitably deformed without changing the value
19 of the integrals.

20
21 **3. Laplace approximation.** In this section, we derive the Laplace approxima-
22 tions to the contour integrals in Lemma 3. Laplace's method for contour integrals
23 is discussed, for example, in Chapter 4 of Olver (1997). The method describes
24 an asymptotic approximation to a contour integral $\oint_{\mathcal{K}} e^{-nf(z)} g(z) dz$ as $n \rightarrow \infty$,
25 where $f(z)$ and $g(z)$ are analytic functions of z . The approximation is usually
26 based on the part of the contour integral coming from a neighborhood of some
27 point $z_0 \in \mathcal{K}$, where z_0 is such that $\frac{d}{dz} f(z_0) = 0$ and $\operatorname{Re} f(z_0) = \min_{z \in \mathcal{K}} \operatorname{Re} f(z)$.
28 For such a point to exist, one might need to deform the contour so that, by Cauchy's
29 theorem, the value of the integral does not change. Typically, the deformation is
30 chosen so that $\operatorname{Re}(-f(z))$ declines in the fastest way possible as z goes away from
31 z_0 along the contour. For this reason, the method is called the *method of steepest*
32 *descent*.

33 The contour integrals in (2.9) and (2.10) can be represented in the Laplace form
34 with a deterministic function $f(z)$ and a random function $g(z)$ that converges to
35 a log-normal random process on the contour as $p, n \rightarrow_c \infty$. To see this, note that
36 the logarithm of the multiple product in (2.9) and (2.10) equals $-\frac{1}{2} \sum_{j=1}^p \ln(z -$
37 $\lambda_j)$. For each z , this expression is a special form of the linear spectral statistic
38 $\sum_{j=1}^p \varphi(\lambda_j)$ studied by Bai and Silverstein (2004). According to the central limit
39 theorem (Theorem 1.1) established in that paper, the random variable

$$\begin{aligned}
 & 40 \\
 & 41 \quad (3.1) \quad \Delta_p(z) = \sum_{j=1}^p \ln(z - \lambda_j) - p \int \ln(z - \lambda) d\mathcal{F}_p(\lambda) \\
 & 42 \\
 & 43
 \end{aligned}$$

converges in distribution to a normal random variable when $p, n \rightarrow_c \infty$. Here $\mathcal{F}_p(\lambda)$ is the cumulative distribution function of the Marchenko–Pastur distribution with a mass of $\max(0, 1 - c_p^{-1})$ at zero and density

$$(3.2) \quad \psi_p(x) = \frac{1}{2\pi c_p x} \sqrt{(b_p - x)(x - a_p)},$$

where $c_p = p/n$, $a_p = (1 - \sqrt{c_p})^2$ and $b_p = (1 + \sqrt{c_p})^2$.

Such a convergence suggests the following choices of $f(z)$ and $g(z)$ in the Laplace forms of the integrals in (2.9) and (2.10):

$$(3.3) \quad f(z) = -\frac{1}{2} \left(\frac{h}{1+h} z - c_p \int \ln(z - \lambda) d\mathcal{F}_p(\lambda) \right)$$

and

$$(3.4) \quad g(z) = \begin{cases} \exp\left\{-\frac{1}{2}\Delta_p(z)\right\}, & \text{for (2.9),} \\ \exp\left\{-\frac{np-p+2}{2} \ln\left(1 - \frac{h}{1+h} \frac{z}{S}\right) - \frac{n}{2} \frac{h}{1+h} z - \frac{1}{2}\Delta_p(z)\right\}, & \text{for (2.10).} \end{cases}$$

As mentioned above, a particularly useful deformation of \mathcal{K} passes through the point $z = z_0(h)$ where $\frac{d}{dz} f(z) = 0$. Taking the derivative of the right-hand side of (3.3), we see that $z_0(h)$ must satisfy

$$(3.5) \quad \frac{h}{1+h} + c_p m_p(z_0(h)) = 0,$$

where $m_p(z) = \int \frac{1}{\lambda - z} d\mathcal{F}_p(\lambda)$ is the Stieltjes transform of the Marchenko–Pastur distribution with parameter c_p . The properties of $m_p(z)$ are well studied. In particular, the analytic expression for $m_p(z)$ is known; see, for example, equation (2.3) in Bai (1993). For $z \neq 0$, which lies outside the support of $\mathcal{F}_p(\lambda)$, we have

$$(3.6) \quad m_p(z) = \frac{-z - c_p + 1 + \sqrt{(z - c_p - 1)^2 - 4c_p}}{2c_p z},$$

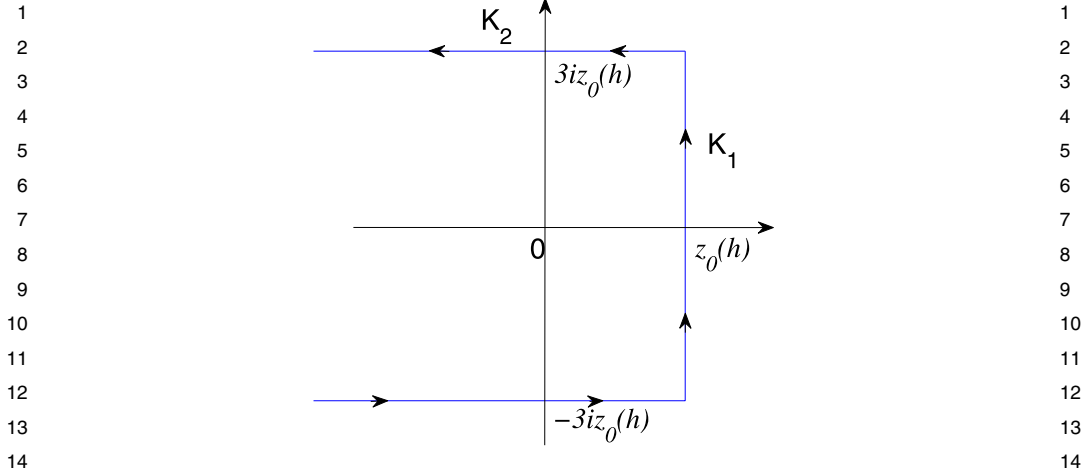
where the branch of the square root is chosen so that the real and the imaginary parts of $\sqrt{(z - c_p - 1)^2 - 4c_p}$ have the same signs as the real and the imaginary parts of $z - c_p - 1$, respectively.

Substituting (3.6) into (3.5) and solving for $z_0(h)$ when $h \in (0, \sqrt{c_p})$, we get

$$(3.7) \quad z_0(h) = \frac{(1+h)(c_p+h)}{h}.$$

When $h \geq \sqrt{c_p}$, there are no solutions to (3.5) that lie outside the support of $\mathcal{F}_p(\lambda)$.

When $h = \sqrt{c_p}$, the right-hand side of (3.7) equals $(1 + \sqrt{c_p})^2$, which lies exactly

FIG. 2. Deformation K of contour \mathcal{K} .

on the boundary of the support of $\mathcal{F}_p(\lambda)$. When $h > \sqrt{c_p}$, (3.7) provides a solution to (3.5) only when the branch of the square root in (3.6) is chosen differently. As can be verified using (3.3) and (3.6), in such a case, $\frac{d}{dz}f(z)$ is strictly negative at $z = z_0(h)$ given by (3.7).

As $c_p \rightarrow c$, any fixed h that is smaller than \sqrt{c} eventually satisfies the inequality $h < \sqrt{c_p}$, so that $\frac{d}{dz}f(z) = 0$ at $z = z_0(h)$. Therefore, for $h < \sqrt{c}$, we will deform the contour \mathcal{K} into a contour K that passes through $z_0(h)$. We define K as $K = K_+ \cup K_-$, where K_- is the complex conjugate of K_+ and $K_+ = K_1 \cup K_2$ with

$$(3.8) \quad K_1 = \{z_0(h) + it : 0 \leq t \leq 3z_0(h)\}$$

and

$$(3.9) \quad K_2 = \{x + 3iz_0(h) : -\infty < x \leq z_0(h)\}.$$

Figure 2 illustrates the choice of K .

A proof of the following technical lemma is relegated to the Supplementary Appendix [Onatski, Moreira and Hallin (2013)].

LEMMA 4. *Suppose that our null hypothesis is true, and let \bar{h} be any fixed number such that $0 < \bar{h} < \sqrt{c}$. Deforming contour \mathcal{K} into K leaves the value of the integrals (2.9) and (2.10) in Lemma 3 unchanged for all $h \in (0, \bar{h}]$ with probability approaching one as $p, n \rightarrow_c \infty$.*

We now derive, uniformly (over $h \in (0, \bar{h}]$), Laplace approximations to the integrals (2.9) and (2.10) in Lemma 3. First, we introduce additional notation. When $f(z)$ and $g(z)$ are analytic at $z_0 = z_0(h)$, let f_s and g_s with $s = 0, 1, \dots$ be the

coefficients in the power series representations

$$(3.10) \quad f(z) = \sum_{s=0}^{\infty} f_s(z - z_0)^s, \quad g(z) = \sum_{s=0}^{\infty} g_s(z - z_0)^s.$$

When $f(z)$ and $g(z)$ are not analytic at z_0 , let the coefficients f_s and g_s be arbitrary numbers for all s .

The following lemma is a generalization of the well-known Watson lemma for contour integrals; see [Olver \(1997\)](#), page 118. Theorem 7.1 in [Olver \(1997\)](#), page 127, derives a similar generalization for the case when $f(z)$ and $g(z)$ are fixed deterministic analytic functions. In contrast to Olver's theorem, our lemma allows $g(z)$ to be a random function, and $f(z)$ to depend on parameter h , and obtains a uniform approximation over $h \in (0, \bar{h}]$. The proof is relegated to the Supplementary Appendix [[Onatski, Moreira and Hallin \(2013\)](#)].

LEMMA 5. *Under the conditions of Lemma 4, for any $h \in (0, \bar{h}]$ and any positive integer m , as $p, n \rightarrow_c \infty$, we have*

$$(3.11) \quad \oint_K e^{-nf(z)} g(z) dz = 2e^{-nf_0} \left[\sum_{s=0}^{m-1} \Gamma\left(s + \frac{1}{2}\right) \frac{a_{2s}}{n^{s+1/2}} + \frac{O_p(1)}{hn^{m+1/2}} \right],$$

where $O_p(1)$ is uniform in $h \in (0, \bar{h}]$. The coefficients a_s in (3.11) can be expressed through f_s and g_s defined above. In particular, we have

$$(3.12) \quad a_0 = \frac{g_0}{2f_2^{1/2}} \quad \text{and} \quad a_2 = \left\{ 4g_2 - \frac{6f_3g_1}{f_2} + \left(\frac{15f_3^2}{2f_2^2} - \frac{6f_4}{f_2} \right) g_0 \right\} \frac{1}{8f_2^{3/2}}.$$

As we explained above, $z_0(h)$ is not a critical point of $f(z)$ when $h > \sqrt{c_p}$. This leads to a situation where the Laplace method for the integral $\oint_K e^{-nf(z)} g(z) dz$ delivers a rather crude approximation. Fortunately, our asymptotic analysis tolerates crude approximations when $h > \sqrt{c_p}$. The following lemma, which is proven in the Supplementary Appendix [[Onatski, Moreira and Hallin \(2013\)](#)], is sufficient for our purposes.

LEMMA 6. *Let $\tilde{h} > \sqrt{c}$, and denote by $K(\tilde{h})$ the corresponding contour, as defined in (3.8) and (3.9). Under the null hypothesis, deforming the contour K into $K(\tilde{h})$ leaves the value of the integrals in Lemma 3 unchanged for all $h \in [\tilde{h}, \infty)$ with probability approaching one as $p, n \rightarrow_c \infty$. Further, for any $h \in [\tilde{h}, \infty)$,*

$$(3.13) \quad \oint_{K(\tilde{h})} e^{-nf(z)} g(z) dz = e^{-nf(z_0(\tilde{h}))} O_p(1),$$

where $O_p(1)$ is uniform over $h \in [\tilde{h}, \infty)$.

Neither Lemma 5 nor Lemma 6 addresses interesting cases with h in a neighborhood of \sqrt{c} . In such cases, $z_0(h)$ would be close to the upper boundary of the support of the Marchenko–Pastur distribution. This may lead to the nonanalyticity of $f(z)$ and $g(z)$ on K and a more complicated asymptotic behavior of $g(z)$. We leave the analysis of cases where h may approach \sqrt{c} for future research.

Guionnet and Maïda (2005) study the asymptotic behavior of spherical integrals using large deviation techniques. Their Theorems 3 and 6 imply Lemma 6 and can be used to obtain the first term in the asymptotic expansion of Lemma 5.

4. Asymptotic behavior of the likelihood ratios. In this section, we discuss the asymptotic behavior of the likelihood ratios $L(h; \lambda)$ and $L(h; \mu)$. First, let us focus on the case where $h \leq \bar{h}$. In the Appendix, we use Lemmas 4 and 5 to derive the following theorem.

THEOREM 7. *Suppose that the null hypothesis is true ($h = 0$). Let \bar{h} be any fixed number such that $0 < \bar{h} < \sqrt{c}$ and let $C[0, \bar{h}]$ be the space of real-valued continuous functions on $[0, \bar{h}]$ equipped with the supremum norm. Then as $p, n \rightarrow_c \infty$, we have*

$$(4.1) \quad L(h; \lambda) = e^{-[\Delta_p(z_0(h)) - \ln(1 - h^2/c_p)]/2} + O_p(n^{-1})$$

and

$$(4.2) \quad L(h; \mu) = e^{-[\Delta_p(z_0(h)) - \ln(1 - h^2/c_p) - h^2/(2c_p) + (h/c_p)(S-p)]/2} + O_p(n^{-1}).$$

Furthermore, $\ln L(h; \lambda)$ and $\ln L(h; \mu)$, viewed as random elements of $C[0, \bar{h}]$, converge weakly to $\mathcal{L}(h; \lambda)$ and $\mathcal{L}(h; \mu)$ with Gaussian finite-dimensional distributions such that, for any $h_1, \dots, h_r \in [0, \bar{h}]$,

$$(4.3) \quad E(\mathcal{L}(h_j; \lambda)) = \frac{1}{4} \ln(1 - c^{-1}h_j^2),$$

$$(4.4) \quad \text{Cov}(\mathcal{L}(h_j; \lambda), \mathcal{L}(h_k; \lambda)) = -\frac{1}{2} \ln(1 - c^{-1}h_j h_k),$$

$$(4.5) \quad E(\mathcal{L}(h_j; \mu)) = \frac{1}{4} [\ln(1 - c^{-1}h_j^2) + c^{-1}h_j^2]$$

and

$$(4.6) \quad \text{Cov}(\mathcal{L}(h_j; \mu), \mathcal{L}(h_k; \mu)) = -\frac{1}{2} [\ln(1 - c^{-1}h_j h_k) + c^{-1}h_j h_k].$$

The log likelihood ratio processes studied in Theorem 7 are not of the standard locally asymptotically normal form. This is because they cannot be represented as $\varphi_1(h)W + \varphi_2(h)$, where $\varphi_1(h)$ and $\varphi_2(h)$ are some deterministic functions of h , and W is a standard normal random variable. Indeed, had the representation $\varphi_1(h)W + \varphi_2(h)$ been possible, the covariance of the limiting log likelihood process at h_1 and h_2 would have been $\varphi_1(h_1)\varphi_1(h_2)$. Hence, for $\mathcal{L}(h; \lambda)$, for instance, we would have had $\varphi_1(h) = \sqrt{-\frac{1}{2} \ln(1 - c^{-1}h^2)}$ and $\varphi_1(h_1)\varphi_1(h_2) = -\frac{1}{2} \ln(1 - c^{-1}h_1 h_2)$, which cannot be true for all $0 < h_1 < \sqrt{c}$ and $0 < h_2 < \sqrt{c}$.

1 The quantity $\Delta_p(z_0(h))$ plays an important role in the limits of experiments. 1
 2 The likelihood ratio processes are well approximated by simple functions of 2
 3 $\Delta_p(z_0(h))$ and S , which are easy to compute from the data and are asymptotically 3
 4 Gaussian by the central limit theorem of [Bai and Silverstein \(2004\)](#). Recalling the 4
 5 definition (3.1) of $\Delta_p(z_0(h))$, we see that asymptotically, all statistical information 5
 6 about parameter h is contained in the deviations of the sample covariance eigenval- 6
 7 ues $\lambda_1, \dots, \lambda_p$ from $\lim_{n,p \rightarrow \infty} z_0(h) = \frac{(1+h)(h+c)}{h}$. Although the latter limit does 7
 8 not have an obvious interpretation when $h < \sqrt{c}$, it is the probability limit of λ_1 8
 9 under alternatives with $h > \sqrt{c}$; see, for example, [Baik and Silverstein \(2006\)](#). 9

10 Let us now consider cases where $h > \sqrt{c}$. We prove the following theorem in 10
 11 the [Appendix](#). 11
 12 12

13 **THEOREM 8.** *Suppose that the null hypothesis is true ($h = 0$), and let H be 13
 14 any fixed number such that $\sqrt{c} < H < \infty$. Then as $p, n \rightarrow_c \infty$, the following 14
 15 holds. For any $h \in [H, \infty)$, the likelihood ratios $L(h; \lambda)$ and $L(h; \mu)$ converge to 15
 16 zero; more precisely, there exists $\delta > 0$ that depends only on H such that 16
 17*

$$18 \quad (4.7) \quad L(h; \lambda) = O_p(e^{-n\delta}) \quad \text{and} \quad L(h; \mu) = O_p(e^{-n\delta}). \quad 18$$

19 19
 20 Note that [Theorem 7](#) and [Le Cam's first lemma](#) [see [van der Vaart \(1998\)](#), 20
 21 page 88] imply that the joint distributions of $\lambda_1, \dots, \lambda_m$ (as well as those of 21
 22 μ_1, \dots, μ_{m-1}) under the null and under the alternative are mutually contiguous 22
 23 for any $h \in [0, \sqrt{c})$. In contrast, [Theorem 8](#) shows that mutual contiguity is lost 23
 24 for $h > \sqrt{c}$. For such h , consistent tests (as $p, n \rightarrow_c \infty$) exist at any probability 24
 25 level $\alpha > 0$. 25

26 In a similar setting, [Nadakuditi and Edelman \(2008\)](#) call the number of “signal 26
 27 eigenvalues” of the population covariance matrix that exceed $1 + \sqrt{c}$ the “effec- 27
 28 tive number of identifiable signals” [see also [Nadakuditi and Silverstein \(2010\)](#)]. 28
 29 [Theorems 7](#) and [8](#) shed light on the formal statistical content of this concept. The 29
 30 “identifiable signals” are detected with probability approaching one in large sam- 30
 31 ples (irrespective of the probability level $\alpha > 0$ at which identification tests are 31
 32 performed). Other signals still can be detected, but the probability of detecting 32
 33 them will never approach one (whatever the probability level $\alpha < 1$). 33
 34 34

35 **5. Asymptotic power analysis.** [Theorem 7](#) can be used to study “local” pow- 35
 36 ers of the tests for detecting signals in noise. The nonstandard form of the limit 36
 37 of log likelihood ratio processes in our setting makes it hard to develop tests with 37
 38 optimal local power properties. However, using the [Neyman–Pearson lemma](#) and 38
 39 [Le Cam's third lemma](#), we can analytically derive the local asymptotic power en- 39
 40 velope and compare local asymptotic powers of specific tests to this envelope. 40

41 It is convenient to reparametrize our problem to $\theta = \sqrt{-\ln(1 - h^2/c)}$. As h 41
 42 varies in the region of contiguity $[0, \sqrt{c})$, θ spans the entire half-line $[0, \infty)$. Note 42
 43 43

1 that the asymptotic mean and autocovariance functions of the log likelihood ra- 1
 2 tios derived in the previous section depend on h only through $h/\sqrt{c} = \sqrt{1 - e^{-\theta^2}}$. 2
 3 Therefore, under the new parametrization, they depend only on θ . Loosely speak- 3
 4 ing, θ and $\sqrt{p/n} \sim \sqrt{c}$ play the classical roles of a “local parameter” and a conti- 4
 5 guity rate, respectively. 5

6 Let $\beta(\theta_1; \lambda)$ and $\beta(\theta_1; \mu)$ be the asymptotic powers of the asymptotically most 6
 7 powerful λ - and μ -based tests of size α of the null $\theta = 0$ against the alternative 7
 8 $\theta = \theta_1$. The following proposition is proven in the [Appendix](#). 8
 9

10 PROPOSITION 9. Let Φ denote the standard normal distribution function. 10
 11 Then 11

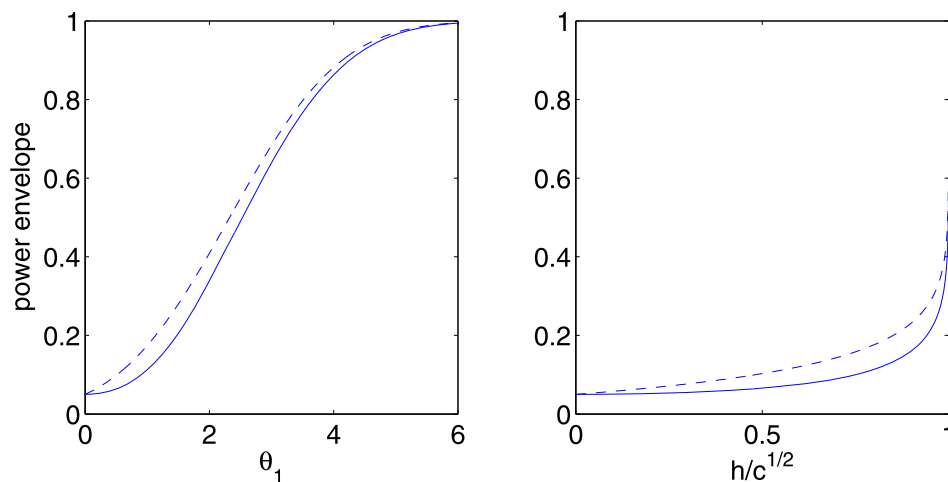
$$12 \quad (5.1) \quad \beta(\theta_1; \lambda) = 1 - \Phi\left[\Phi^{-1}(1 - \alpha) - \frac{\theta_1}{\sqrt{2}}\right] \quad 12$$

13 and 13

$$14 \quad (5.2) \quad \beta(\theta_1; \mu) = 1 - \Phi\left[\Phi^{-1}(1 - \alpha) - \sqrt{\frac{1}{2}(\theta_1^2 - 1 + e^{-\theta_1^2})}\right]. \quad 14$$

15 Plots of the asymptotic power envelopes $\beta(\theta_1; \lambda)$ and $\beta(\theta_1; \mu)$ against θ_1 for 15
 16 asymptotic size $\alpha = 0.05$ are shown in the left panel of Figure 3. The power loss 16
 17 of the μ -based tests relative to the λ -based tests is due to the nonspecification 17
 18 of σ^2 . In contrast to λ -based tests, μ -based tests may achieve the corresponding 18
 19 power envelope even when σ^2 is unknown. 19

20 The right panel of Figure 3 shows the envelopes as functions of the original 20
 21 parameter h normalized by \sqrt{c} . We see that the alternatives that can theoretically 21
 22 22
 23 23
 24 24
 25 25
 26 26



27 FIG. 3. The maximal asymptotic power of the λ -based tests (dashed lines) and μ -based tests (solid 27
 28 lines) of $\theta = 0$ against $\theta = \theta_1$. Left panel: θ -parametrization. Right panel: h -parametrization. 28
 29 29
 30 30
 31 31
 32 32
 33 33
 34 34
 35 35
 36 36
 37 37
 38 38
 39 39
 40 40
 41 41
 42 42
 43 43

1 be detected with high probability are concentrated near the threshold $h = \sqrt{c}$. The 1
 2 strong nonlinearity of the θ -parametrization should be kept in mind while inter- 2
 3 preting the figures that follow. 3

4 It is interesting to compare the power envelopes to the asymptotic powers 4
 5 of the likelihood ratio (LR) and weighted average power (WAP) tests. The λ - 5
 6 based LR and WAP tests of $\theta = 0$ against the alternative $\theta \in (0, M]$, where 6
 7 $M < \infty$, would reject the null if and only if, respectively, $2 \sup_{\theta \in (0, M]} \ln L(\theta; \lambda)$ 7
 8 and $\ln \int_0^M L(\theta; \lambda) W(d\theta)$ are sufficiently large. The power of a WAP test would, of 8
 9 course, depend on the choice of the weighting measure $W(d\theta)$. The μ -based LR 9
 10 and WAP tests are defined similarly. Theorem 7 and Le Cam's third lemma sug- 10
 11 gest a straightforward procedure for the numerical evaluation of the corresponding 11
 12 asymptotic power functions. 12

13 Consider, for example, the λ -based LR test statistic. According to Theorem 7, 13
 14 its asymptotic distribution under the null equals the distribution of $2 \sup_{\theta \in (0, M]} X_\theta$, 14
 15 where X_θ is a Gaussian process with $E(X_\theta) = -\theta^2/4$ and $\text{Cov}(X_{\theta_1}, X_{\theta_2}) =$ 15
 16 $-\frac{1}{2} \ln(1 - \sqrt{(1 - e^{-\theta_1^2})(1 - e^{-\theta_2^2})})$. According to Le Cam's third lemma, under 16
 17 a specific alternative $\theta = \theta_1 \leq M$, the asymptotic distribution of the LR statis- 17
 18 tic equals the distribution of $2 \sup_{\theta \in (0, M]} \tilde{X}_\theta$, where \tilde{X}_θ is a Gaussian process 18
 19 with the same covariance function as that of X_θ , but with a different mean: 19
 20 $E(\tilde{X}_\theta) = E(X_\theta) + \text{Cov}(X_\theta, X_{\theta_1})$. 20
 21

22 Therefore, to numerically evaluate the asymptotic power function of the λ -based 22
 23 LR test, we simulate 500,000 observations of X_θ on a grid of 1000 equally spaced 23
 24 points in $\theta \in [0, M = 6]$, where $M = 6$ is chosen as the upper limit of the grid 24
 25 because it is large enough for the power envelopes to reach the value of 99%. For 25
 26 each observation, we save its supremum on the grid, and use the empirical distri- 26
 27 bution of two times the suprema as the approximate asymptotic distribution of the 27
 28 likelihood ratio statistic under the null. We denote this distribution as \hat{F}_0 . Its 95% 28
 29 quantile equals 4.3982. 29

30 For each θ_1 on the grid, we repeat the simulation for process \tilde{X}_θ to obtain the 30
 31 approximate asymptotic distribution of the likelihood ratio statistic under the alter- 31
 32 native $\theta = \theta_1$, which we denote as \hat{F}_1 . We use the value of \hat{F}_1 at the 95% quantile 32
 33 of \hat{F}_0 as a numerical approximation to the asymptotic power at θ_1 of the λ -based 33
 34 LR test with asymptotic size 0.05. 34

35 Figure 4 shows the resulting asymptotic power curve of the LR test (solid line) 35
 36 along with the asymptotic power envelope (dotted line). It also shows the asymp- 36
 37 totic power of the WAP test with $W(d\theta)$ equal to the uniform measure on $[0, 6]$ 37
 38 (dashed line). The left and right panels correspond to λ - and μ -based tests, respec- 38
 39 tively. 39

40 The asymptotic powers of the LR and WAP tests both come close to the power 40
 41 envelope. The LR and WAP power functions are so close that they are difficult 41
 42 to distinguish clearly. The asymptotic power of the WAP test appears to be larger 42
 43 than that of the LR test for all θ_1 in the $[0, 6]$ range, except for relatively large θ_1 . 43

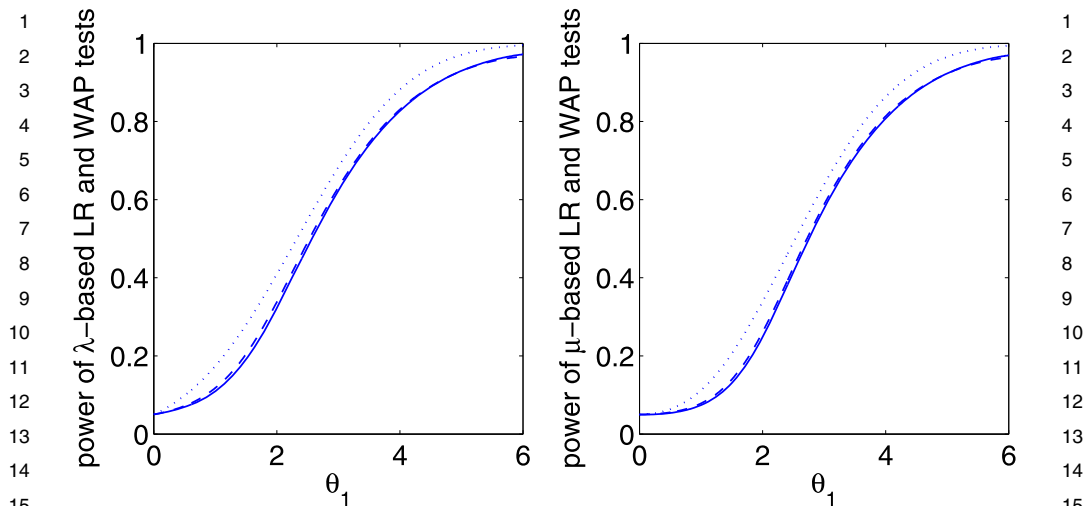


FIG. 4. The asymptotic power envelope (dotted line), the asymptotic power of the LR test (solid line), and the asymptotic power of the WAP test with uniform weighting measure on $\theta \in [0, 6]$ (dashed line). Left panel: λ -based tests and envelope. Right panel: μ -based tests and envelope.

Hence, the LR test still may be admissible. More accurate numerical analysis is needed to shed further light on this issue.

In the remaining part of this section, we consider some of the tests that have been proposed previously in the literature, and, in Proposition 10, derive their asymptotic power functions. We focus on four examples. Three of them are inspired by the “classical” fixed- p theory, while the fourth is more directly based on results from the large random matrix theory.

The problem of testing the hypothesis of sphericity has a long history, and has generated a considerable body of literature, which we only very briefly summarize here. The classical fixed- p Gaussian analysis of the various problems considered here goes back to Mauchly (1940), who first derived the Gaussian likelihood ratio test for sphericity. The (Gaussian) locally most powerful invariant (under shift, scale and orthogonal transformations) test was obtained by John (1971, 1972) and by Sugiura (1972), with adjusted versions resisting elliptical violations of the Gaussian assumptions proposed in Hallin and Paindaveine (2006), where a Le Cam approach is adopted under a general elliptical setting. Ledoit and Wolf (2002) propose two extensions (for the unknown and known scale problems, resp.) of John’s test, while Bai et al. (2009) adapt Mauchly’s (1940) likelihood ratio test.

EXAMPLE 1 [John’s (1971) test of sphericity]. John (1971) proposes testing the sphericity hypothesis $\theta = 0$ against general alternatives using the test statistic $U = \frac{1}{p} \text{tr}[(\frac{\hat{\Sigma}}{(1/p)\text{tr}(\hat{\Sigma})} - I_p)^2]$, where $\hat{\Sigma}$ is the sample covariance matrix of the data. He shows that, when $n > p$, such a test is locally most powerful invariant. Studying

1 John's test when $p/n \rightarrow c \in (0, \infty)$, Ledoit and Wolf (2002) prove that, under the 1
 2 null, $nU - p \xrightarrow{d} N(1, 4)$. Hence, the test with asymptotic size α rejects the null 2
 3 hypothesis of sphericity if $\frac{1}{2}(nU - p - 1) > \Phi^{-1}(1 - \alpha)$. 3
 4

5 EXAMPLE 2 [The Ledoit and Wolf (2002) test of $\Sigma = I$]. Ledoit and Wolf 5
 6 (2002) propose using $W = \frac{1}{p} \text{tr}[(\hat{\Sigma} - I)^2] - \frac{p}{n} [\frac{1}{p} \text{tr} \hat{\Sigma}]^2 + \frac{p}{n}$ as a test statistic for 6
 7 testing the hypothesis that the population covariance matrix is a unit matrix. Under 7
 8 the null, $nW - p \xrightarrow{d} N(1, 4)$. As in the previous example, the null hypothesis is 8
 9 rejected at asymptotic size α if $\frac{1}{2}(nW - p - 1) > \Phi^{-1}(1 - \alpha)$. 9
 10

11 EXAMPLE 3 [The "corrected" LRT of Bai et al. (2009)]. When $n > p$, 11
 12 Bai et al. (2009) propose a corrected version of the likelihood ratio statistic 12
 13 $\text{CLR} = \text{tr} \hat{\Sigma} - \ln \det \hat{\Sigma} - p - p(1 - (1 - \frac{p}{n}) \ln(1 - \frac{p}{n}))$ based on the entire 13
 14 data, as opposed to λ or μ only, to test the equality of the population covari- 14
 15 ance matrix to the identity matrix against general alternatives. Under the null, 15
 16 $\text{CLR} \xrightarrow{d} N(-\frac{1}{2} \ln(1 - c), -2 \ln(1 - c) - 2c)$. The null hypothesis is rejected when- 16
 17 ever $\text{CLR} + \frac{1}{2} \ln(1 - c) > \sqrt{-2 \ln(1 - c) - 2c} \Phi^{-1}(1 - \alpha)$. 17
 18
 19

20 More directly inspired by the asymptotic theory of random matrices, several 20
 21 authors have recently proposed and studied various tests based on λ_1 or μ_1 : see 21
 22 Bejan (2005), Patterson, Price and Reich (2006), Kritchman and Nadler (2009), 22
 23 Onatski (2009), Bianchi et al. (2010) and Nadakuditi and Silverstein (2010). We 23
 24 refer to these tests, which reject H_0 for large values of λ_1 or μ_1 , as Tracy–Widom- 24
 25 type tests. 25

26 EXAMPLE 4 (Tracy–Widom-type tests). Asymptotic critical values of such 26
 27 tests are obtained using the fact, established by Johnstone (2001), that under the 27
 28 null, 28
 29

$$29 \quad (5.3) \quad n^{2/3} c^{1/6} (1 + \sqrt{c})^{-4/3} (\lambda_1 - (1 + \sqrt{c})^2) \xrightarrow{d} \text{TW},$$

30 where TW denotes the Tracy–Widom law of the first kind. The null hypothesis is 30
 31 rejected when λ_1 or μ_1 exceeds the adequate Tracy–Widom quantile. 31
 32
 33

34 Consider the tests described in Examples 1, 2, 3 and 4, and denote by $\beta_J(\theta_1)$, 34
 35 $\beta_{\text{LW}}(\theta_1)$, $\beta_{\text{CLR}}(\theta_1)$, and $\beta_{\text{TW}}(\theta_1)$ their respective asymptotic powers at asymptotic 35
 36 level α . The following proposition is established in the Appendix. 36
 37

38 PROPOSITION 10. Denote $1 - e^{-\theta_1^2}$ as $\psi(\theta_1)$. The asymptotic power functions 38
 39 of the tests described in Examples 1–4 satisfy, for any $\theta_1 > 0$, 39
 40

$$40 \quad (5.4) \quad \beta_{\text{TW}}(\theta_1) = \alpha,$$

$$41 \quad (5.5) \quad \beta_J(\theta_1) = \beta_{\text{LW}}(\theta_1) = 1 - \Phi(\Phi^{-1}(1 - \alpha) - \frac{1}{2} \psi(\theta_1))$$

42

43

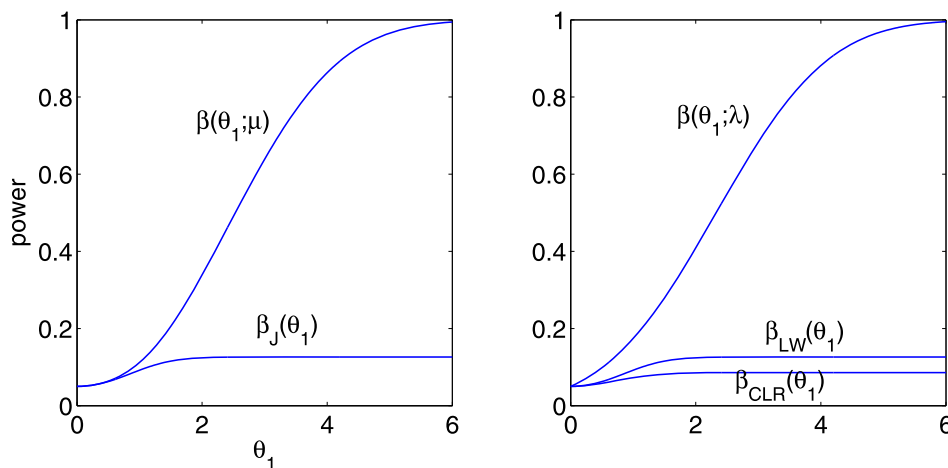
1 and

$$(5.6) \quad \beta_{\text{CLR}}(\theta_1) = 1 - \Phi\left(\Phi^{-1}(1 - \alpha) - \frac{\sqrt{c\psi(\theta_1)} - \ln(1 + \sqrt{c\psi(\theta_1)})}{\sqrt{-2\ln(1 - c) - 2c}}\right).$$

2
3
4
5 With the important exception of [Srivastava \(2005\)](#), (5.4)–(5.6) are the first
6 results on the asymptotic power of those tests against contiguous alternatives.
7 [Srivastava \(2005\)](#) analyzes the asymptotic power of tests similar to those in Ex-
8 amples 1 and 2. His Theorems 3.1 and 4.1 can be used to establish (5.5).
9

10 From Proposition 10, we see that the local asymptotic power of the Tracy–
11 Widom-type tests is trivial. As shown by [Baik, Ben Arous and P ech e \(2005\)](#) in the
12 complex data case and by [F eral and P ech e \(2009\)](#) in the real data case, the conver-
13 gence (5.3) holds not only under the null, but also under any alternative of the form
14 $h = h_0 < \sqrt{c}$. Under the “local” parametrization adopted in this section, such alter-
15 natives have the form $\theta = \theta_1 > 0$. It can be shown that the Tracy–Widom-type tests
16 are consistent against noncontiguous alternatives $h = h_1 > \sqrt{c}$. However, such a
17 consistency is likely to be also a property of the LR tests based on μ or on λ . If
18 this holds true, the LR tests asymptotically dominate the Tracy–Widom-type tests.
19 A more detailed analysis of the optimality properties of LR tests is the subject of
20 ongoing research.

21 The asymptotic power functions of the tests from Examples 1, 2 and 3 are non-
22 trivial. Figure 5 compares these power functions to the corresponding power en-
23 velopes. Since John’s test is invariant with respect to orthogonal transformations
24 and scalings, $\beta_J(\theta_1)$ is compared to the power envelope $\beta(\theta_1; \mu)$. The asymp-
25 totic power functions $\beta_{\text{LW}}(\theta_1)$ and $\beta_{\text{CLR}}(\theta_1)$ are compared to the power envelope
26



27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42 FIG. 5. Asymptotic powers ($\beta_J, \beta_{\text{LW}}, \beta_{\text{CLR}}$) of the tests described in Examples 1 (John), 2 (Ledoit
43 and Wolf) and 3 (Bai et al.).

1 $\beta(\theta_1; \lambda)$ because the Ledoit–Wolf test of $\Sigma = I$ and the “corrected” likelihood 1
 2 ratio test are invariant only with respect to orthogonal transformations. 2

3 Interestingly, whereas $\beta_J(\theta_1)$ and $\beta_{LW}(\theta_1)$ depend only on α and θ_1 , $\beta_{CLR}(\theta_1)$ 3
 4 depends also on c . As c converges to one, $\beta_{CLR}(\theta_1)$ converges to α , which corre- 4
 5 sponds to the case of trivial power. As c converges to zero, $\beta_{CLR}(\theta_1)$ converges to 5
 6 $\beta_J(\theta_1)$. In Figure 5, we provide the plot of $\beta_{CLR}(\theta_1)$ that corresponds to $c = 0.5$. 6

7 The left panel of Figure 5 shows that the power function of John’s test is very 7
 8 close to the power envelope $\beta(\theta_1; \mu)$ in the vicinity of $\theta_1 = 0$. Such behavior is 8
 9 consistent with the fact that John’s test is locally most powerful invariant. However, 9
 10 for large θ_1 , the asymptotic power functions of all the tests from Examples 1, 10
 11 2 and 3 are lower than the corresponding asymptotic power envelopes. We should 11
 12 stress here that these tests have power against general alternatives as opposed to the 12
 13 “spiked” alternatives that maintain the assumption that the population covariance 13
 14 matrix of data has the form $\sigma^2(I_p + hvv')$. 14

15 For the “spiked” alternatives, the λ - and μ -based LR tests may be more attrac- 15
 16 tive. However, implementing these tests requires some care. A “quick-and-dirty” 16
 17 approach would be to approximate $\ln L(\theta; \lambda)$ and $\ln L(\theta; \mu)$ by the simple but 17
 18 asymptotically equivalent expressions from (4.1) and (4.2), compute two times 18
 19 their maxima on a grid over $\theta \in (0, M]$, and compare them with critical values 19
 20 obtained by simulation as for the construction of Figure 4. Unfortunately, in fi- 20
 21 nite samples, this simple approach will lead to a numerical breakdown whenever 21
 22 $z_0(h(\theta))$ happens to be less than the largest sample covariance eigenvalue for some 22
 23 $\theta \leq M$. In addition, since the asymptotic approximation derived in Theorem 7 is 23
 24 not uniform over entire half-line $\theta \in [0, \infty)$, its quality will depend on the choice 24
 25 of M . For relatively large M , the asymptotic behavior of the LR test implemented 25
 26 as above may poorly match its finite sample behavior. 26

27 Instead, we recommend implementing the LR tests without using the asymptotic 27
 28 approximations. The finite sample log likelihood ratios $\ln L(\theta; \lambda)$ and $\ln L(\theta; \mu)$ 28
 29 can be computed using the contour integral representations (2.9) and (2.10). 29
 30 Choosing the contour of integration so that the sample covariance eigenvalues re- 30
 31 main to its left will eliminate the numerical breakdown problem associated with the 31
 32 asymptotic tests. Furthermore, under the Gaussianity assumption, the finite sample 32
 33 distributions of the log likelihood ratios are pivotal. Hence, the exact critical values 33
 34 can be computed via Monte Carlo simulations as follows: simulate many replica- 34
 35 tions of data under the null. For each replication, compute the log likelihood ratio 35
 36 and store two times its maximum. Use the 95% quantile of the empirical distribu- 36
 37 tion of the stored values as a numerical approximation for the exact critical value 37
 38 of the test. The finite sample properties of such a test are left as an important topic 38
 39 for future research. 39
 40

41 **6. Conclusion.** In this paper, we study the asymptotic power of tests for the 41
 42 existence of rank-one perturbations of sphericity as both the dimensionality of the 42
 43 43

1 data and the number of observations go to infinity. Focusing on tests that are in- 1
 2 variant with respect to orthogonal transformations and rescaling, we establish the 2
 3 convergence of the log ratio of the joint densities of the sample covariance eigen- 3
 4 values under the alternative and null hypotheses to a Gaussian process indexed by 4
 5 the norm of the perturbation. 5

6 When the perturbation norm is larger than the phase transition threshold stud- 6
 7 ied in Baik, Ben Arous and P ech e (2005), the limiting log-likelihood process is 7
 8 degenerate and the joint eigenvalue distributions under the null and alternative hy- 8
 9 potheses are asymptotically mutually singular, so that the discrimination between 9
 10 the null and the alternative is asymptotically certain. When the norm is below the 10
 11 threshold, the limiting log-likelihood process is nondegenerate and the joint eigen- 11
 12 value distributions under the null and alternative hypotheses are mutually con- 12
 13 tiguous. Using the asymptotic theory of statistical experiments, we obtain power 13
 14 envelopes and derive the asymptotic size and power for various eigenvalue-based 14
 15 tests in the region of contiguity. 15

16 Several questions are left for future research. First, we only considered rank- 16
 17 one perturbations of the spherical covariance matrices. It would be desirable to 17
 18 extend the analysis to finite-rank perturbations. Such an extension will require a 18
 19 more complicated technical analysis. Second, it would be interesting to extend our 19
 20 analysis to the asymptotic regime $p, n \rightarrow \infty$ with $p/n \rightarrow \infty$ or $p/n \rightarrow 0$. In the 20
 21 context of sphericity tests, such asymptotic regimes have been recently studied in 21
 22 Birke and Dette (2005). Third, a thorough analysis of the finite sample properties 22
 23 of the proposed LR tests would clarify the related practical implementation issues. 23
 24 Fourth, our Lemma 5 can be used to derive higher-order asymptotic approxima- 24
 25 tions to the likelihood ratios, which may improve finite-sample performances of 25
 26 asymptotic tests. Finally, it would be of considerable interest to relax the Gaussian 26
 27 assumptions, for example, into elliptical ones, preferably with unspecified radial 27
 28 densities, on the model (in a fixed- p context) of Hallin and Paindaveine (2006). 28

29 APPENDIX 29

30
 31 **A.1. Proof of Proposition 1.** For the joint density $p(\lambda; h)$ of $\lambda_1, \dots, \lambda_m$, we 31
 32 have 32

$$33 \quad p(\lambda; h) = \tilde{\gamma} \frac{\prod_{i=1}^m \lambda_i^{(|p-n|-1)/2} \prod_{i < j}^m (\lambda_i - \lambda_j)}{(1+h)^{n/2}} \quad 33$$

$$34 \quad \times \int_{\mathcal{O}(p)} e^{-(n/2) \text{tr}(\Pi Q' \Lambda Q)} (dQ), \quad 34$$

35 (A.1) 35

36 where $\tilde{\gamma}$ depends only on n and p , $\Pi = \text{diag}((1+h)^{-1}, 1, \dots, 1)$, $\mathcal{O}(p)$ is the 36
 37 set of all $p \times p$ orthogonal matrices and (dQ) is the invariant measure on the 37
 38 orthogonal group $\mathcal{O}(p)$ normalized to make the total measure unity. When $n \geq p$, 38
 39 (A.1) is a special case of the density given in James (1964), page 483. When $n < p$, 39
 40 (A.1) follows from Theorems 2 and 6 in Uhlig (1994). 40
 41
 42
 43 43

Let $\Psi = \text{diag}(\frac{h}{1+h}, 0, \dots, 0)$ be a $p \times p$ matrix. Since $\Pi = I_p - \Psi$, we have $\text{tr}(\Pi Q' \Lambda Q) = \text{tr} \Lambda - \text{tr}(\Psi Q' \Lambda Q)$, and we can rewrite (A.1) as

$$(A.2) \quad p(\lambda; h) = \tilde{\gamma} \frac{\prod_{i=1}^m \lambda_i^{(|p-n|-1)/2} \prod_{i < j}^m (\lambda_i - \lambda_j) e^{-(n/2) \text{tr} \Lambda}}{(1+h)^{n/2}} \times \int_{\mathcal{O}(p)} e^{(n/2) \text{tr}(\Psi Q' \Lambda Q)} (dQ).$$

Note that $\text{tr}(\Psi Q' \Lambda Q) = \text{tr}(Q \Psi Q' \Lambda) = \frac{h}{1+h} x_p' \Lambda x_p$, where x_p is the first column of Q . When Q is uniformly distributed over $\mathcal{O}(p)$, its first column x_p is uniformly distributed over $\mathcal{S}(p)$. Therefore, we have

$$(A.3) \quad p(\lambda; h) = \tilde{\gamma} \frac{\prod_{i=1}^m \lambda_i^{(|p-n|-1)/2} \prod_{i < j}^m (\lambda_i - \lambda_j) e^{-(n/2) \text{tr} \Lambda}}{(1+h)^{n/2}} \times \int_{\mathcal{S}(p)} e^{(n/2)(h/(1+h))x_p' \Lambda x_p} (dx_p),$$

which establishes (2.1). Now, let $y = \lambda_1 + \dots + \lambda_p$ so that $\mu_j = \lambda_j/y$. Note that $\text{tr} \Lambda = y$, $\text{tr} M = \mu_1 + \dots + \mu_p = 1$, and that the Jacobian of the coordinate change from $\lambda_1, \dots, \lambda_m$ to $\mu_1, \dots, \mu_{m-1}, y$ equals y^{m-1} . Changing variables in (A.3), and integrating y out, we obtain (2.2).

A.2. Proof of Lemma 3. Using (2.3) in the ratio of the right-hand side of (2.1) with $h > 0$ to that with $h = 0$, and changing the variable of integration from s to $z = \frac{1+h}{h} \frac{2}{n} s$, we get (2.9). Further, from (2.2), we have

$$(A.4) \quad \begin{aligned} p(\mu; 0) &= \delta(n, p, \mu) \int_0^\infty y^{np/2-1} e^{-ny/2} dy \\ &= \delta(n, p, \mu) \left(\frac{2}{n}\right)^{np/2} \Gamma\left(\frac{np}{2}\right). \end{aligned}$$

For $h > 0$, using (2.3) in (2.2), we get

$$(A.5) \quad \begin{aligned} p(\mu; h) &= \frac{\delta(n, p, \mu) \Gamma(p/2)}{(1+h)^{n/2} 2\pi i} \\ &\quad \times \int_0^\infty \oint_{\tilde{\mathcal{K}}} y^{(np-2)/2} e^{s-ny/2} \prod_{j=1}^p \left(s - \frac{n}{2} \frac{yh}{1+h} \mu_j\right)^{-1/2} ds dy, \end{aligned}$$

where $\tilde{\mathcal{K}}$ is a contour starting at $-\infty$, encircling counter-clockwise the points $0, \frac{ny}{2} \frac{h}{1+h} \mu_1, \dots, \frac{ny}{2} \frac{h}{1+h} \mu_m$ and going back to $-\infty$. Since $\frac{h}{1+h} \mu_j < 1$ by construction, we may and will choose $\tilde{\mathcal{K}}$ so that for any $s \in \tilde{\mathcal{K}}$, $\text{Re } s < \frac{ny}{2}$. Changing variables of integration from y and s to $w = \frac{ny}{2}$ and $z = s \frac{1+h}{hw}$, where S is any positive

constant, and dividing by the right-hand side of (A.4), we obtain

$$L(h; \mu) = \frac{S^{(p-2)/2}(1+h)^{(p-n-2)/2}\Gamma(p/2)}{h^{(p-2)/2}\Gamma(np/2)2\pi i} \times \int_0^\infty \oint_{\mathcal{K}} w^{np/2-p/2} e^{(wh/(1+h))(z/S)-w} \prod_{j=1}^p (z - S\mu_j)^{-1/2} dz dw,$$

where \mathcal{K} is a contour starting at $-\infty$, encircling counter-clockwise the points $0, S\mu_1, \dots, S\mu_m$, and going back to $-\infty$. In addition, for any $z \in \mathcal{K}$, $\operatorname{Re} z < \frac{1+h}{h}S$. Such a choice of \mathcal{K} guarantees that the integrand in the above double integral is absolutely integrable on $[0, \infty) \times \mathcal{K}$, so that Fubini's theorem can be used to justify the interchange of the order of the integrals. Changing the order of the integrals and setting $S = \lambda_1 + \dots + \lambda_p$, we obtain (2.10).

A.3. Proof of Theorem 7. First, let us formulate the following technical lemma. Its proof is in the Supplementary Appendix [Onatski, Moreira and Hallin (2013)].

LEMMA 11. (i) If $h < \sqrt{c_p}$, $f_0 = -\frac{1}{2}(c_p + (1 - c_p) \ln(1 + h) - c_p \ln \frac{c_p}{h})$.
(ii) If $h > \sqrt{c_p}$, $f_0 = -\frac{1}{2}(h + c_p + (1 - c_p) \ln(c_p + h) - \frac{c_p}{h} - \ln h)$.

Below, we prove Theorem 7 for $L(h; \mu)$. The proof for $L(h; \lambda)$ is similar but simpler, and we omit it to save space. As follows from Lemmas 4 and 5, the integral in (2.10) can be represented as $2e^{-nf_0}[\Gamma(\frac{1}{2})\frac{a_0}{n^{1/2}} + \frac{O_p(1)}{hn^{3/2}}]$ uniformly in $h \in (0, \bar{h}]$. Therefore, and since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, we can write

$$(A.5) \quad L(h; \mu) = \frac{k_2 S^{(p-2)/2}}{\sqrt{n\pi i}} e^{-nf_0} \left[a_0 + h^{-1} O_p\left(\frac{1}{n}\right) \right],$$

where $k_2 = h^{-(p-2)/2}(1+h)^{(p-n-2)/2}\frac{(n-1)p}{2}\Gamma(\frac{(n-1)p}{2})\Gamma(\frac{p}{2})\Gamma^{-1}(\frac{np}{2})$. Using Stirling's approximation $\Gamma(r) = e^{-r}r^r(\frac{2\pi}{r})^{1/2}(1 + O(r^{-1}))$ with $r = \frac{p}{2}, \frac{np}{2}$ and $\frac{(n-1)p}{2}$, and the fact that $\ln(n-1) = \ln n - n^{-1} - \frac{1}{2}n^{-2} + O(n^{-3})$, we find, after algebraic simplifications, that

$$(A.6) \quad \frac{k_2}{\sqrt{n\pi}} = h^{-(p-2)/2}(1+h)^{(p-n-2)/2} \times e^{-((p-2)/2)\ln n - p/2 + c_p/4 + \ln c_p/2}(1 + O(n^{-1})).$$

Using (A.6) and Lemma 11(i), we obtain

$$\frac{k_2 S^{(p-2)/2}}{\sqrt{n\pi i}} e^{-nf_0} h^{-1} O_p\left(\frac{1}{n}\right) = \frac{1}{1+h} \left(\frac{S}{p}\right)^{(p-2)/2} e^{c_p/4 - \ln c_p/2} O_p\left(\frac{1}{n}\right),$$

1 which, together with the fact that $S - p = O_p(1)$, implies that

$$2 \quad (A.7) \quad \frac{k_2 S^{(p-2)/2}}{\sqrt{n\pi i}} e^{-nf_0} h^{-1} O_p\left(\frac{1}{n}\right) = O_p\left(\frac{1}{n}\right) \quad 3$$

4 uniformly over $h \in (0, \bar{h}]$.

5 Now, as can be verified using (3.3) and (3.6), if $h < \sqrt{c_p}$, then

$$6 \quad (A.8) \quad f_2 = -\frac{h^2}{4(1+h)^2(c_p - h^2)}. \quad 7$$

8 Therefore, using (3.12), we obtain

$$9 \quad (A.9) \quad a_0 = i \frac{(1+h)(c_p - h^2)^{1/2}}{h} g_0. \quad 10$$

11 Using (3.4), (A.6), (A.9) and Lemma 11(i) in (A.5), after algebraic simplifications
12 and rearrangements of terms, we get

$$13 \quad \ln\left[\frac{k_2 S^{(p-2)/2} e^{-nf_0} a_0}{\sqrt{n\pi i}}\right] \quad 14$$

$$15 \quad (A.10) \quad = \frac{1}{2} \ln\left(1 - \frac{h^2}{c_p}\right) + \frac{c_p}{4} + \frac{p-2}{2} \ln\left(\frac{S}{p}\right) \quad 16$$

$$17 \quad - \frac{n}{2} \frac{hz_0(h)}{1+h} - \frac{np-p+2}{2} \ln\left(1 - \frac{h}{1+h} \frac{z_0(h)}{S}\right) - \frac{1}{2} \Delta_p(z_0(h)). \quad 18$$

19 Finally, using the fact that $S - p = O_p(1)$, we obtain $\ln(S/p) = (S - p)/p +$
20 $O_p(p^{-2})$ and

$$21 \quad \ln\left(1 - \frac{h}{1+h} \frac{z_0(h)}{S}\right) = -\frac{h}{1+h} \frac{z_0(h)}{p} - \frac{1}{2} \left(\frac{hz_0(h)}{(1+h)p}\right)^2 \quad 22$$

$$23 \quad + \frac{h}{1+h} \frac{z_0(h)}{p^2} (S - p) + O_p(p^{-3}). \quad 24$$

25 The latter two equalities, (A.10) and the fact that $\frac{h}{1+h} z_0(h) = h + c_p$ entail

$$26 \quad (A.11) \quad \frac{k_2 S^{(p-2)/2} e^{-nf_0} a_0}{\sqrt{n\pi i}} \quad 27$$

$$28 \quad = e^{-\{\Delta_p(z_0(h)) - \ln(1 - h^2/c_p) + (h/c_p)(S - p) - h^2/(2c_p) + O_p(p^{-1})\}/2}, \quad 29$$

30 which, together with (A.7), imply formula (4.2).

31 Now, let us prove the convergence of $\ln L(h; \mu)$ to $\mathcal{L}(h; \mu)$. By (4.2), the joint
32 convergence of $\ln L(h_j; \mu)$ with $j = 1, \dots, r$ to a Gaussian vector is equivalent
33 to the convergence of $(S - p, \Delta_p(z_0(h_1)), \dots, \Delta_p(z_0(h_r)))$ to a Gaussian vector.
34 A proof of the following technical lemma, based on Theorem 1.1 of Bai and Sil-
35 verstein (2004), is given in the Supplementary Appendix [Onatski, Moreira and
36 Hallin (2013)].

LEMMA 12. *Suppose that the null hypothesis holds. Then, as $p, n \rightarrow_c \infty$, the vector $(S - p, \Delta_p(z_0(h_1)), \dots, \Delta_p(z_0(h_r)))$ converges in distribution to a Gaussian vector $(\eta, \xi_1, \dots, \xi_r)$ with*

$$\begin{aligned} E\eta &= 0, & \text{Var}(\eta) &= 2c, & \text{Cov}(\eta, \xi_j) &= -2h_j, \\ \text{Cov}(\xi_j, \xi_k) &= -2\ln(1 - c^{-1}h_jh_k) & \text{and} & & E\xi_j &= \frac{1}{2}\ln(1 - c^{-1}h_j^2). \end{aligned}$$

Lemma 12 and (4.2) imply that $E[\mathcal{L}(h_j; \mu)] = -\frac{1}{2}E\xi_j + \frac{1}{2}\ln(1 - c^{-1}h_j^2) + \frac{1}{4}c^{-1}h_j^2 = \frac{1}{4}[\ln(1 - c^{-1}h_j^2) + c^{-1}h_j^2]$ and

$$\begin{aligned} \text{Cov}[\mathcal{L}(h_j; \mu), \mathcal{L}(h_k; \mu)] &= \frac{1}{4}\text{Cov}(\xi_j, \xi_k) + \frac{h_k}{4c}\text{Cov}(\xi_j, \eta) \\ &\quad + \frac{h_j}{4c}\text{Cov}(\xi_k, \eta) + \frac{h_jh_k}{4c^2}\text{Var}(\eta) \\ &= -\frac{1}{2}\ln(1 - c^{-1}h_jh_k) - \frac{h_jh_k}{2c}, \end{aligned}$$

which establishes (4.5) and (4.6).

To complete the proof of Theorem 7, we need to note that the tightness of $L(h; \mu)$, viewed as a random element of the space $C([0, \bar{h}])$, as $p, n \rightarrow_c \infty$, follows from formula (4.2) and the fact that $S - p$ and $\Delta_p(z_0(h))$, are $O_p(1)$, uniformly in $h \in (0, \bar{h}]$. This uniformity is a consequence of Lemma A2 proven in the Supplementary Appendix [Onatski, Moreira and Hallin (2013)].

A.4. Proof of Theorem 8. As in the proof of Theorem 7, we will focus on the case of the likelihood ratio based on μ . The proof for $L(h; \lambda)$ is similar. According to Lemma 6 and formula (2.10), for any $\tilde{h} > \sqrt{c}$, we have $L(h; \mu) = k_2 S^{(p-2)/2} e^{-nf(z_0(\tilde{h}))} O_p(1)$. Using (A.6) and the fact that $(\frac{S}{p})^p = (1 + \frac{S-p}{p})^p = (1 + \frac{O_p(1)}{p})^p = O_p(1)$, we can write

$$(A.12) \quad L(h; \mu) = e^{(n/2)(c_p \ln(c_p(1+h)/h) - \ln(1+h) - c_p - 2f(z_0(\tilde{h})))} O_p(n^{1/2}).$$

Noting that $\tilde{h} > \sqrt{c_p}$ for sufficiently large n and p , and using Lemma 11(ii) and the fact that $\frac{\tilde{h}}{1+\tilde{h}}z_0(\tilde{h}) = \tilde{h} + c_p$, we get $-2f(z_0(\tilde{h})) = (1 - c_p)\ln(c_p + \tilde{h}) - \frac{c_p}{\tilde{h}} - \ln \tilde{h} + \frac{h}{1+h}z_0(\tilde{h})$. Substituting the latter expression in (A.12) and simplifying, we obtain

$$(A.13) \quad L(h; \mu) = e^{(n/2)R(h, \tilde{h}, c_p)} O_p(n^{1/2}),$$

where $O_p(\cdot)$ is uniform in $h \in [\tilde{h}, \infty)$ and $R(h, \tilde{h}, c_p) = (1 - c_p)\ln(c_p + \tilde{h}) - \frac{c_p}{\tilde{h}} - \ln \tilde{h} + \frac{h}{1+h}z_0(\tilde{h}) - (1 - c_p)\ln(1+h) - c_p \ln h + c_p \ln c_p - c_p$.

As $n, p \rightarrow \infty$, $R(h, \tilde{h}, c_p) \rightarrow R(h, \tilde{h}, c)$ uniformly over $(h, \tilde{h}) \in [\sqrt{c}, H]^2$. On the other hand, $R(h, \tilde{h}, c)$ is continuous on $(h, \tilde{h}) \in [\sqrt{c}, H]^2$, $R(\sqrt{c}, \sqrt{c}, c) = 0$, and $\frac{d}{dh}R(h, \tilde{h}, c) = (1+h)^{-2} \left(\frac{(1+\tilde{h})(c+\tilde{h})}{\tilde{h}} - \frac{(1+h)(c+h)}{h} \right) < 0$ for all h and \tilde{h} such that $\sqrt{c} \leq \tilde{h} < h \leq H$. Therefore, for any $H > \sqrt{c}$, there exist \tilde{h} and δ such that $\sqrt{c} < \tilde{h} \leq H$, $\delta > 0$ and $R(H, \tilde{h}, c) < -3\delta$; and thus, for sufficiently large n and p , $R(H, \tilde{h}, c_p) < -3\delta$. Now, $\frac{d}{dh}R(h, \tilde{h}, c_p) = (1+h)^{-2}(z_0(\tilde{h}) - z_0(h)) < 0$ for all $h > \tilde{h}$, as long as $\tilde{h} \geq \sqrt{c_p}$. Hence, for sufficiently large n and p , $R(h, \tilde{h}, c_p) < -3\delta$ for all $h > \tilde{h}$. Using (A.13), we get $|L(h; \mu)| \leq e^{-3n\delta/2} O_p(n^{1/2}) = O_p(e^{-n\delta})$ uniformly over $h \in [H, \infty)$.

A.5. Proof of Proposition 9. For brevity, we derive only the asymptotic power envelope for the case of μ -based tests. According to the Neyman–Pearson lemma, the most powerful test of the null $\theta = 0$ against a particular alternative $\theta = \theta_1$ is the test which rejects the null when $\ln L(\theta_1; \mu)$ is larger than some critical value C . It follows from Theorem 7 that, for such a test to have asymptotic size α , C must be $C = \sqrt{V(\theta_1)}\Phi^{-1}(1 - \alpha) + m(\theta_1)$, where $m(\theta_1) = (-\theta_1^2 + 1 - e^{-\theta_1^2})/4$ and $V(\theta_1) = (\theta_1^2 - 1 + e^{-\theta_1^2})/2$ are obtained from (4.5) and (4.6) by the reparametrization $\theta = \sqrt{-\ln(1 - h^2/c)}$. Now, according to Le Cam’s third lemma and Theorem 7, under $\theta = \theta_1$, $\ln L(\theta_1; \mu) \xrightarrow{d} N(m(\theta_1) + V(\theta_1), V(\theta_1))$. Therefore, the asymptotic power $\beta(\theta_1; \mu)$ of the asymptotically most powerful test of $\theta = 0$ against $\theta = \theta_1$ is (5.2).

A.6. Proof of Proposition 10. As shown by Baik, Ben Arous and P ech e (2005) in the complex case and by F eral and P ech e (2009) in the real case, the convergence (5.3) takes place not only under the null, but also under alternatives $h = h_1$ with $h_1 < \sqrt{c}$, yielding $\theta = \theta_1 < \infty$ under the parametrization $\theta = \sqrt{-\ln(1 - h^2/c)}$. Hence, (5.4) follows.

Formulas (5.5) and (5.6) can be established using conceptually similar steps. To save space, below we only establish formula (5.6). The following technical lemma is proven in the Supplementary Appendix [Onatski, Moreira and Hallin (2013)].

LEMMA 13. *Let CLR be the “corrected” likelihood ratio statistic as defined in Example 3. Then, under the null, as $p, n \rightarrow_c \infty$, the vector $(\text{CLR}, \Delta_p(z_0(h)))$ converges in distribution to a Gaussian vector (ζ_1, ζ_2) with $\text{Cov}(\zeta_1, \zeta_2) = -2h + 2\ln(1 + h)$.*

Lemma 13 and (4.2) imply the convergence in distribution of the vector $(\text{CLR}, \ln L(h; \lambda))$ to a Gaussian vector $(\zeta_1, -\frac{1}{2}\zeta_2)$. From Bai et al. (2009), we know that, under the null, $\text{CLR} \xrightarrow{d} N(-\frac{1}{2}\ln(1 - c), -2\ln(1 - c) - 2c)$. By Le

1 Cam's third lemma, under the alternative $h = h_1$, CLR converges to a Gaus- 1
 2 sian random variable with the same variance but with mean equal to $-\frac{1}{2}\ln(1 - 2$
 3 $c) + \text{Cov}(\zeta_1, -\frac{1}{2}\zeta_2) = -\frac{1}{2}\ln(1 - c) + h - \ln(1 + h)$ evaluated at $h = h_1$. There- 3
 4 fore, the power of the "corrected" likelihood ratio test of asymptotic size α 4
 5 equals $1 - \Phi(\Phi^{-1}(1 - \alpha) - \frac{h_1 - \ln(1 + h_1)}{\sqrt{-2\ln(1 - c) - 2c}})$. Using the reparametrization $\theta_1 =$ 5
 6 $\sqrt{-\ln(1 - h_1^2/c)}$, we get (5.6). 6
 7

8
 9 **Acknowledgments.** This work started when the first two authors worked at 9
 10 and the third author visited Columbia University. We would like to thank Tony 10
 11 Cai, the Associate Editor, Nick Patterson and an anonymous referee for helpful 11
 12 and encouraging comments. 12

13 SUPPLEMENTARY MATERIAL 13

14
 15 **Supplementary Appendix** (DOI: 10.1214/13-AOS1100SUPP; .pdf). The Sup- 15 aos1100_supp.pdf
 16 plementary Appendix contains proofs of Lemmas 4, 5, 6, 11, 12 and 13. 16

17 REFERENCES 17

- 18
 19 BAI, Z. D. (1993). Convergence rate of expected spectral distributions of large random matrices. II. 19
 20 Sample covariance matrices. *Ann. Probab.* **21** 649–672. [MR1217560](#) 20 <mr>
 21 BAI, Z. D. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional 21
 22 sample covariance matrices. *Ann. Probab.* **32** 553–605. [MR2040792](#) 22 <mr>
 23 BAI, Z., JIANG, D., YAO, J.-F. and ZHENG, S. (2009). Corrections to LRT on large-dimensional 23
 24 covariance matrix by RMT. *Ann. Statist.* **37** 3822–3840. [MR2572444](#) 24 <mr>
 25 BAIK, J., BEN AROUS, G. and PÉCHÉ, S. (2005). Phase transition of the largest eigenvalue for 25
 26 nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. [MR2165575](#) 26 <mr>
 27 BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked 27
 28 population models. *J. Multivariate Anal.* **97** 1382–1408. [MR2279680](#) 28 <mr>
 29 BEJAN, A. I. (2005). Largest eigenvalues and sample covariance matrices. Tracy–Widom and 29
 30 Painleve II; computational aspects and realization in S-plus with applications. Unpublished 30
 31 manuscript, Univ. Warwick. 31 <unstr>
 32 BERTHET, Q. and RIGOLLET, P. (2012). Optimal detection of sparse principal components in high 32
 33 dimension. Available at arXiv:1202.5070. 33 <unstr>
 34 BIANCHI, P., DEBBAH, M., MAIDA, M. and NAJIM, J. (2010). Performance of statistical tests for 34
 35 single source detection using random matrix theory. Unpublished manuscript. 35 <unstr>
 36 BIRKE, M. and DETTE, H. (2005). A note on testing the covariance matrix for large dimension. 36
 37 *Statist. Probab. Lett.* **74** 281–289. [MR2189467](#) 37 <mr>
 38 BLOEMENDAL, A. and VIRÁG, B. (2011). Limits of spiked random matrices I. Available at 38
 39 arXiv:1011.1877v2. 39 <unstr>
 40 BUTLER, R. W. and WOOD, A. T. A. (2002). Laplace approximations for hypergeometric functions 40
 41 with matrix argument. *Ann. Statist.* **30** 1155–1177. [MR1926172](#) 41 <mr>
 42 CHEN, S. X., ZHANG, L.-X. and ZHONG, P.-S. (2010). Tests for high-dimensional covariance 42
 43 matrices. *J. Amer. Statist. Assoc.* **105** 810–819. [MR2724863](#) 43 <mr>
 44 DICKEY, J. M. (1983). Multiple hypergeometric functions: Probabilistic interpretations and statisti- 44
 45 cal uses. *J. Amer. Statist. Assoc.* **78** 628–637. [MR0721212](#) 45 <mr>
 46 EL KAROUI, N. (2007). Tracy–Widom limit for the largest eigenvalue of a large class of complex 46
 47 sample covariance matrices. *Ann. Probab.* **35** 663–714. [MR2308592](#) 47 <mr>

- 1 ERDELYI, A. (1937). Beitrag zur theorie der konfluenten hypergeometrischen funktionen von
2 mehreren veranderlichen. *Sitzungsberichte, Akademie der Wissenschaften in Wien, Abteilung IIa,*
3 *Mathematisch-Naturwissenschaftliche Klasse* **146** 431–467. 3 <auto>
- 4 FÉRAL, D. and PÉCHÉ, S. (2009). The largest eigenvalues of sample covariance matrices for a
5 spiked population: Diagonal case. *J. Math. Phys.* **50** 073302, 33. [MR2548630](#) 4 <mr>
- 6 FISHER, T. J., SUN, X. and GALLAGHER, C. M. (2010). A new test for sphericity of the covariance
7 matrix for high dimensional data. *J. Multivariate Anal.* **101** 2554–2570. [MR2719881](#) 5 <mr>
- 8 GEMAN, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 252–261.
9 [MR0566592](#) 6 <mr>
- 10 GUIONNET, A. and MAÏDA, M. (2005). A Fourier view on the R -transform and related asymptotics
11 of spherical integrals. *J. Funct. Anal.* **222** 435–490. [MR2132396](#) 7 <mr>
- 12 HALLIN, M. and PAINDAVEINE, D. (2006). Semiparametrically efficient rank-based inference for
13 shape. I. Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756. [MR2329465](#) 8 <mr>
- 14 HILLIER, G. (2001). The density of a quadratic form in a vector uniformly distributed on the n -
15 sphere. *Econometric Theory* **17** 1–28. [MR1863565](#) 9 <mr>
- 16 HOYLE, D. C. (2008). Automatic PCA dimension selection for high dimensional data and small
17 sample sizes. *J. Mach. Learn. Res.* **9** 2733–2759. 10 <auto>
- 18 JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples.
19 *Ann. Math. Statist.* **35** 475–501. [MR0181057](#) 11 <mr>
- 20 JOHN, S. (1971). Some optimal multivariate tests. *Biometrika* **58** 123–127. [MR0275568](#) 12 <mr>
- 21 JOHN, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions.
22 *Biometrika* **59** 169–173. [MR0312619](#) 13 <mr>
- 23 JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components
24 analysis. *Ann. Statist.* **29** 295–327. [MR1863961](#) 14 <mr>
- 25 KRITCHMAN, S. and NADLER, B. (2008). Determining the number of components in a factor model
26 from limited noisy data. *Chemometrics and Intelligent Laboratory Systems* **94** 19–32. 15 <auto>
- 27 KRITCHMAN, S. and NADLER, B. (2009). Non-parametric detection of the number of signals:
28 Hypothesis testing and random matrix theory. *IEEE Trans. Signal Process.* **57** 3930–3941.
29 [MR2683143](#) 16 <mr>
- 30 LE CAM, L. (1960). Locally asymptotically normal families of distributions. Certain approximations
31 to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ.*
32 *California Publ. Statist.* **3** 37–98. [MR0126903](#) 17 <mr>
- 33 LEDOIT, O. and WOLF, M. (2002). Some hypothesis tests for the covariance matrix when the di-
34 mension is large compared to the sample size. *Ann. Statist.* **30** 1081–1102. [MR1926169](#) 18 <mr>
- 35 LIJOI, A. and REGAZZINI, E. (2004). Means of a Dirichlet process and multiple hypergeometric
36 functions. *Ann. Probab.* **32** 1469–1495. [MR2060305](#) 19 <mr>
- 37 MAUCLY, J. W. (1940). Significance test for sphericity of a normal n -variate distribution. *Ann.*
38 *Math. Statist.* **11** 204–209. [MR0002084](#) 20 <mr>
- 39 MO, M. Y. (2012). Rank 1 real Wishart spiked model. *Comm. Pure Appl. Math.* **65** 1528–1638.
40 [MR2969495](#) 21 <mr>
- 41 NADAKUDITI, R. R. and EDELMAN, A. (2008). Sample eigenvalue based detection of high-
42 dimensional signals in white noise using relatively few samples. *IEEE Trans. Signal Process.*
43 **56** 2625–2638. [MR1500236](#) 22 <mr>
- 44 NADAKUDITI, R. R. and SILVERSTEIN, J. W. (2010). Fundamental limit of sample generalized
45 eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only
46 samples. *IEEE Journal of Selected Topics in Signal Processing* **4** 468–480. 23 <auto>
- 47 NADLER, B. (2008). Finite sample approximation results for principal component analysis: A matrix
48 perturbation approach. *Ann. Statist.* **36** 2791–2817. [MR2485013](#) 24 <mr>
- 49 OLVER, F. W. J. (1997). *Asymptotics and Special Functions*. AK Peters, Wellesley, MA.
50 [MR1429619](#) 25 <mr>

- 1 ONATSKI, A. (2009). Testing hypotheses about the numbers of factors in large factor models. *Econo-* 1
 2 *metrica* **77** 1447–1479. [MR2561070](#) 2 <mr>
- 3 ONATSKI, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. 3
 4 *Rev. Econom. Statist.* **92** 1004–1016. 4 <auto>
- 5 ONATSKI, A., MOREIRA, M. J. and HALLIN, M. (2013). Supplement to “Asymptotic power of 5
 6 sphericity tests for high-dimensional data.” DOI:[10.1214/13-AOS1100SUPP](#). 6 <unstr>
- 7 PATTERSON, N., PRICE, A. L. and REICH, D. (2006). Population structure and eigenanalysis. *PLoS* 7
 8 *Genetics* **2** 2074–2093. 8 <auto>
- 9 PERRY, P. O. and WOLFE, P. J. (2010). Minimax rank estimation for subspace tracking. *IEEE Jour-* 9
 10 *nal of Selected Topics in Signal Processing* **4** 504–513. 10 <auto>
- <uncited> 11 RATNARAJAH, T. and VAILLANCOURT, R. (2005). Complex singular Wishart matrices and appli- 11
 12 cations. *Comput. Math. Appl.* **50** 399–411. [MR2165429](#) 12 <mr>
- <uncited> 13 RUDIN, W. (1987). *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York. [MR0924157](#) 13
 14 SCHOTT, J. R. (2006). A high-dimensional test for the equality of the smallest eigenvalues of a 14
 15 covariance matrix. *J. Multivariate Anal.* **97** 827–843. [MR2256563](#) 15 <mr>
- 16 SILVERSTEIN, J. W. and BAI, Z. D. (1995). On the empirical distribution of eigenvalues of a class 16
 17 of large-dimensional random matrices. *J. Multivariate Anal.* **54** 175–192. [MR1345534](#) 17 <mr>
- 18 SRIVASTAVA, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. 18
 19 *J. Japan Statist. Soc.* **35** 251–272. [MR2328427](#) 19 <mr>
- 20 SRIVASTAVA, H. M. and KARLSSON, P. W. (1985). *Multiple Gaussian Hypergeometric Series*. Ellis 20
 21 Horwood, Chichester. [MR0834385](#) 21 <mr>
- 22 SUGIURA, N. (1972). Locally best invariant test for sphericity and the limiting distributions. *Ann.* 22
 23 *Math. Statist.* **43** 1312–1316. [MR0311032](#) 23 <mr>
- 24 UHLIG, H. (1994). On singular Wishart and singular multivariate beta distributions. *Ann. Statist.* **22** 24
 25 395–405. [MR1272090](#) 25 <mr>
- 26 VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Proba-* 26
 27 *bilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. [MR1652247](#) 27 <mr>
- 28 WANG, D. (2012). The largest eigenvalue of real symmetric, Hermitian and Hermitian self-dual ran- 28
 29 dom matrix models with rank one external source, Part I. *J. Stat. Phys.* **146** 719–761. [MR2916094](#) 29 <mr>
- 30 A. ONATSKI M. J. MOREIRA 30
- 31 FACULTY OF ECONOMICS ESCOLA DE PÓS-GRADUAÇÃO EM ECONOMIA 31
- 32 UNIVERSITY OF CAMBRIDGE FUNAÇÃO GETULIO VARGAS (FGV/EPGE) 32
- 33 SIDGWICK AVENUE PRAIA DE BOTAFOGO, 190-SALA 1100 33
- 34 CAMBRIDGE, CB3 9DD RIO DE JANEIRO-RJ 22250-900 34
- 35 UNITED KINGDOM BRAZIL 35
- 36 E-MAIL: ao319@cam.ac.uk E-MAIL: mjmoreira@fgv.br 36
- 37 M. HALLIN 37
- 38 ECARES 38
- 39 UNIVERSITÉ LIBRE DE BRUXELLES CP 114/04 39
- 40 50, AVENUE F.D. ROOSEVELT 40
- 41 B-1050 BRUXELLES 41
- 42 BELGIUM 42
- 43 E-MAIL: mhallin@ulb.ac.be 43

THE ORIGINAL REFERENCE LIST

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
- The list of entries below corresponds to the original Reference section of your article. The bibliography section on previous page was retrieved from MathSciNet applying an automated procedure.**
- Please check both lists and indicate those entries which lead to mistaken sources in automatically generated Reference list.**
- BAI, Z.D. (1993). Convergence rate of the expected spectral distributions of large random matrices. Part II. Sample covariance matrices. *Annals of Probability* **21** 649-672.
- 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.
- BAI, Z.D., JIANG, D., YAO, J.F., and ZHENG, S. (2009). Corrections to LRT on Large-dimensional Covariance Matrix by RMT. *Annals of Statistics* **37** 3822-3840.
- BAIK, J., BEN AROUS, G. and PECHE, S. (2005). Phase transition of the largest eigenvalue for non-null complex sample covariance matrices. *Annals of Probability* **33** 1643-1697.
- BAIK, J. and SILVERSTEIN, J.W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis* **97** 1382-1408.
- BEJAN, A.I.U. (2005). Largest Eigenvalues and Sample Covariance Matrices. Tracy-Widom and Painleve II; Computational Aspects and Realization In S-Plus with Applications. manuscript, University of Warwick.
- BERTHET, Q. and RIGOLLET, P. (2012). Optimal detection of sparse principal components in high dimension. *arXiv:1202.5070*.
- BIANCHI, P., DEBBAH, M., MAIDA, M. and NAJIM, J. (2010). Performance of Statistical Tests for Single Source Detection using Random Matrix Theory. manuscript.
- BIRKE, M. and DETTE, H. (2005). A note on testing the covariance matrix for large dimension. *Statistics and Probability Letters* **74** 281-289.
- BLOEMENDAL, A. and VIRÁG, B. (2011). Limits of spiked random matrices I. *arXiv:1011.1877v2*.
- BUTLER, R.W. and WOOD, A.T.A. (2002). Laplace approximations for hypergeometric functions with matrix argument. *Annals of Statistics* **30** 1155-1177.
- CHEN, S.X., ZHANG, L.X. and ZHONG, P.S. (2010). Tests for High-Dimensional Covariance Matrices. *Journal of the American Statistical Association* **105** 810-819.
- DICKEY, J.M. (1983). Multiple Hypergeometric Functions: Probabilistic Interpretations and Statistical Uses. *Journal of the American Statistical Association* **78** 628-637.
- EL KAROUI, N. (2007). Tracy-Widom Limit for the Largest Eigenvalue of a Large Class of Complex Sample Covariance Matrices. *Annals of Probability* **38** 663-714.
- ERDELYI, A. (1937). Beitrag zur Theorie der konfluenten hypergeometrischen Funktionen von mehreren Veränderlichen. *Sitzungsberichte, Akademie der Wissenschaften in Wien, Abteilung IIa, Mathematisch-naturwissenschaftliche Klasse* **146** 431-467.
- FERAL, D. and PECHE, S. (2009). The largest eigenvalues of sample covariance matrices for a spiked population: Diagonal case. *Journal of Mathematical Physics* **50** (7).
- FISHER, T.J., SUN, X. and GALLAGHER, C.M. (2010). A new test of sphericity of the covariance matrix for high-dimensional data. *Journal of Multivariate Analysis* **101** 2554-2570.
- GEMAN, S. (1980). A limit theorem for the norm of random matrices. *Annals of Probability* **8** 252-261.
- GUIONNET, A. and MAIDA, M. (2005) A Fourier view on the R-transform and related asymptotics of spherical integrals. *Journal of Functional Analysis* **222** 435-490.
- HALLIN, M. and PAINDAVEINE, D. (2006). Semiparametrically efficient rank-based inference for shape: I Optimal rank-based tests for sphericity. *Annals of Statistics* **34** 2707-2756.
- HILLIER, G. (2001). The Density of a Quadratic Form in a Vector Uniformly Distributed on the n -Sphere. *Econometric Theory* **17** 1-28.

- 1 HOYLE, D.C. (2008). Automatic PCA Dimension Selection for High Dimensional Data and Small 1
2 Sample Sizes. *Journal of Machine Learning Research* **9** 2733-2759. 2
- 3 JAMES, A.T. (1964). Distributions of matrix variates and latent roots derived from normal samples. 3
4 *Annals of Mathematical Statistics* **35** 475-501. 4
- 5 JOHN, S. (1971). Some optimal multivariate tests. *Biometrika* **58** 123–127. 5
- 6 JOHN, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. 6
7 *Biometrika* **59** 169-174. 6
- 8 JOHNSTONE, I.M. (2001). On the distribution of the largest eigenvalue in principal components 7
9 analysis. *Annals of Statistics* **29** 295–327. 8
- 10 KRITCHMAN, S. and NADLER, B. (2008). Determining the number of components in a factor model 9
11 from limited noisy data. *Chemometrics and Intelligent Laboratory Systems* **94** 19-32. 10
- 12 KRITCHMAN, S. and NADLER, B. (2009). Non-Parametric Detection of the Number of Signals: 11
13 Hypothesis Testing and Random Matrix Theory. *IEEE Transactions on Signal Processing* **57** 12
14 3930-3941. 12
- 15 LE CAM, L. (1960). Locally asymptotically normal families of distribution. *University of California* 13
16 *Publications in Statistics* **3** 37-98. 14
- 17 LEDOIT, O. and WOLF, M. (2002). Some Hypothesis Tests for the Covariance Matrix when the 15
18 Dimension is Large Compared to the Sample Size. *Annals of Statistics* **30** 1081-1102. 16
- 19 LUOI, A. and REGAZZINI, E. (2004). Means of a Dirichlet Process and Multiple Hypergeometric 17
20 Functions. *Annals of Probability* **32** 1469-1495. 17
- 21 MAUCHLY, J. W. (1940). Test for sphericity of a normal n-variate distribution. *Annals of Mathemat-* 18
22 *ical Statistics* **11** 204-209. 19
- 23 MO, M. Y. (2011). The rank 1 real Wishart spiked model. *arXiv:1101.5144v1*. To appear in *Com-* 20
24 *munications in Pure and Applied Mathematics*. 21
- 25 NADAKUDITI, R.R. and EDELMAN, A. (2008). Sample Eigenvalue Based Detection of High- 22
26 Dimensional Signals in White Noise Using Relatively Few Samples. *IEEE Transactions on Signal* 23
27 *Processing* **56** 2625-2638. 23
- 28 NADAKUDITI, R.R. and SILVERSTEIN, J.W. (2010). Fundamental Limit of Sample Generalized 24
29 Eigenvalue Based Detection of Signals in Noise Using Relatively Few Signal-Bearing and Noise- 25
30 Only Samples. *IEEE Journal of Selected Topics in Signal Processing* **4** 468-480. 26
- 31 NADLER, B. (2008). Finite Sample Approximation Results for Principal Components Analysis: a 27
32 Matrix Perturbation Approach. *Annals of Statistics* **36** 2791-2817. 27
- 33 OLVER, F.W.J. (1997). *Asymptotics and Special Functions*. A K Peters, Natick, MA. 28
- 34 ONATSKI, A. (2009). Testing Hypotheses About the Number of Factors in Large Factor Models. 29
35 *Econometrica* **77**, 1447-1479. 30
- 36 ONATSKI, A. (2010). Determining the Number of Factors from Empirical Distribution of Eigenval- 31
37 ues. *Review of Economics and Statistics* **92** 1004-1016. 32
- 38 Onatski, A., Moreira, M. J. and Hallin, M. (2013). Supplement to “Asymptotic power of sphericity 33
39 tests for high-dimensional data” 33
- 40 PATTERSON, N., PRICE, A.L. and REICH, D. (2006). Population Structure and Eigenanalysis. *PLoS* 34
41 *Genetics* **2** (12) 2074-2093. 35
- 42 PERRY, P.O. and WOLFE, P.J. (2010). Minimax Rank Estimation for Subspace Tracking. *IEEE* 36
43 *Journal of Selected Topics in Signal Processing* **4** 504-513. 37
- 44 RATNARAJAH, T. and VAILLANOURT, R. (2005). Complex Singular Wishart Matrices and Applica- 38
45 tions. *Computers & Mathematics with Applications* **50** 399–411. 38
- 46 RUDIN, W. (1987). *Real and Complex Analysis*. 3rd edition, McGraw-Hill, New York. 39
- 47 SCHOTT, J. (2006). A high-dimensional test for the equality of the smallest eigenvalues of a covari- 40
48 ance matrix. *Journal of Multivariate Analysis* **97** 827-843. 41
- 49 SILVERSTEIN, J.W. and BAI, Z.D. (1995). On the empirical distribution of eigenvalues of a class of 42
50 large dimensional random matrices. *Journal of Multivariate Analysis* **54** 175-192. 43

1	SRIVASTAVA, M.S. (2005). Some tests concerning the covariance matrix in high-dimensional data.	1
2	<i>Journal of the Japan Statistical Society</i> 35 251-272.	2
3	SRIVASTAVA, H.M. and KARLSSON, P.W. (1985). <i>Multiple Gaussian Hypergeometric Series</i> . Ellis	3
4	Horwood Limited, Chichester.	4
5	SUGIURA, N. (1972). Locally best invariant test for sphericity and the limiting distributions. <i>Annals</i>	5
6	<i>of Mathematical Statistics</i> 43 1312-1316.	6
7	UHLIG, H. (1994). On Singular Wishart and Singular Multivariate Beta Distributions. <i>Annals of</i>	7
8	<i>Statistics</i> 22 395-405.	8
9	VAN DER VAART, A.W. (1998). <i>Asymptotic Statistics</i> . Cambridge University Press.	9
10	WANG, D. (2012). The largest eigenvalue of real symmetric, Hermitian and Hermitian self-dual	10
11	random matrix models with rank one external source, part I. <i>Journal of Statistical Physics</i> 146	11
12	719-761.	12
13		13
14		14
15		15
16		16
17		17
18		18
19		19
20		20
21		21
22		22
23		23
24		24
25		25
26		26
27		27
28		28
29		29
30		30
31		31
32		32
33		33
34		34
35		35
36		36
37		37
38		38
39		39
40		40
41		41
42		42
43		43

1 META DATA IN THE PDF FILE 1

2 **Following information will be included as pdf file Document Properties:** 2

3
4 **Title** : Asymptotic power of sphericity tests for high-dimensional 4
data 4

5 **Author** : Alexei Onatski, Marcelo J. Moreira, Marc Hallin 5

6 **Subject** : The Annals of Statistics, 2013, Vol.0, No.00, 1-28 6

7 **Keywords**: 62H15, 62B15, 41A60, Sphericity tests, large dimensional- 7
ity, asymptotic power, spiked covariance, contiguity, power enve- 8
lope, steepest descent, contour integral representation 9

10
11 THE LIST OF URI ADRESSES 11

12
13
14 **Listed below are all uri addresses found in your paper. The non-active uri addresses, if any, are** 14
indicated as ERROR. Please check and update the list where necessary. The e-mail addresses 15
are not checked – they are listed just for your information. More information can be found in 16
the support page: 16

17 <http://www.e-publications.org/ims/support/urihelp.html>. 17

18
19 200 <http://www.imstat.org/aos/> [2:pp.1,1] OK 19

20 200 <http://www.imstat.org> [2:pp.1,1] OK 20

21 200 <http://www.ams.org/mathscinet/msc/msc2010.html> [2:pp.1,1] OK 21

22 404 <http://dx.doi.org/10.1214/13-AOS1100SUPP> [4:pp.26,26,33,33] Not Found 22

23 200 <http://arxiv.org/abs/1202.5070> [2:pp.27,27] OK 23

24 200 <http://arxiv.org/abs/1011.1877v2> [2:pp.28,28] OK 24

25 --- <mailto:ao319@cam.ac.uk> [2:pp.35,35] Check skip 25

26 --- <mailto:mjmoreira@fgv.br> [2:pp.35,35] Check skip 26

27 --- <mailto:mhallin@ulb.ac.be> [2:pp.35,35] Check skip 27

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43