

## SHARP DETECTION IN PCA UNDER CORRELATIONS: ALL EIGENVALUES MATTER

BY EDGAR DOBRIBAN<sup>1</sup>

*Stanford University*

Principal component analysis (PCA) is a widely used method for dimension reduction. In high-dimensional data, the “signal” eigenvalues corresponding to weak principal components (PCs) do not necessarily separate from the bulk of the “noise” eigenvalues. Therefore, popular tests based on the largest eigenvalue have little power to detect weak PCs. In the special case of the spiked model, certain tests asymptotically equivalent to linear spectral statistics (LSS)—averaging effects over *all* eigenvalues—were recently shown to achieve some power.

We consider a “local alternatives” model for the spectrum of covariance matrices that allows a general correlation structure. We develop new tests to detect PCs in this model. While the top eigenvalue contains little information, due to the strong correlations between the eigenvalues we can detect weak PCs by averaging over all eigenvalues using LSS. We show that it is possible to find the optimal LSS, by solving a certain integral equation. To solve this equation, we develop efficient algorithms that build on our recent method for computing the limit empirical spectrum [Dobriban (2015)]. The solvability of this equation also presents a new perspective on phase transitions in spiked models.

**1. Introduction.** Introduced by Pearson and Hotelling in the early 1900s, Principal Component Analysis (PCA) is a widely used method for dimension reduction. Inference in PCA is classically based on the asymptotic distribution of the top eigenvalues of the sample covariance matrix, which are consistent estimators of the top population eigenvalues under low-dimensional asymptotics—when the sample size grows while the dimension is fixed [Anderson (1963, 2003)].

In contrast, in high dimensions—when the dimension is proportional to the sample size—the behavior of the eigenvalues is different. Below a critical value of the top eigenvalue in the population, the top sample eigenvalue has the same behavior as if there were only null eigenvalues; see, for example, Baik, Ben Arous and Pécché (2005), Benaych-Georges and Nadakuditi (2011) for results in this direction, and Hachem, Hardy and Najim (2015) for a survey. Tests based on the top eigenvalue alone have small power to detect weak PCs in high dimensions.

---

Received February 2016; revised August 2016.

<sup>1</sup>Supported in part by NSF Grants DMS-14-18362 and DMS-14-07813, and by an HHMI International Student Research Fellowship.

*MSC2010 subject classifications.* Primary 62H25; secondary 62H15, 45B05.

*Key words and phrases.* Principal component analysis, linear spectral statistic, random matrix theory, linear integral equation, optimal testing.

Despite this work, our current methods for detecting weak PCs are limited to a small number of covariance matrix models solved explicitly. These all center on the special case of the “spiked model,” where the covariance matrix is a low rank perturbation of the identity [Johnstone (2001)]. For instance, Onatski, Moreira and Hallin (2013, 2014) study likelihood ratio tests in Gaussian spiked models. This is extended to  $F$ -matrices and a few other explicitly solved examples [Dharmawansa, Johnstone and Onatski (2014), Johnstone and Onatski (2015)]. Is it possible to detect weak PCs under general covariance matrix models, such as those of Marchenko and Pastur (1967)? This question is relevant for settings where the spiked model is not a good description of empirical data.

A second key challenge is that the pre-existing theoretical approach does not generalize directly. The likelihood approach of [Onatski, Moreira and Hallin \(2013, 2014\)](#) is hard to extend to general covariance matrices, and a new approach is needed.

As a consequence of our results, below certain problem-specific signal strengths quite generally *all eigenvalues matter* to achieve optimal detection of PCs in high-dimensional data. This is in contrast to the low-dimensional case, as well as to the high-dimensional case with *strong* PCs separating from the bulk [e.g., Baïk, Ben Arous and P  ch   (2005), Paul (2007), etc.]. Thus, our results identify a special but broad regime where optimal inference must be based on all eigenvalues.

1.1. *Our contributions.* Suppose we have an  $n \times p$  data matrix  $X_{n \times p}$ , with  $n$  rows sampled from a  $p$ -dimensional population. The samples are allowed to have



**2. Sharp detection in PCA.** More rigorously, we observe an  $n \times p$  data matrix  $X_{n \times p}$ , where  $n$  is the sample size and  $p$  is the dimensionality. If the samples are drawn independently from a population with covariance matrix  $\Sigma_p$ , then one can model  $X_{n \times p} = Z_{n \times p} \Sigma_p^{1/2}$ , where the  $n \times p$  matrix  $Z_{n \times p}$  has i.i.d. standardized entries, and  $\Sigma_p$  is a  $p \times p$  deterministic positive semidefinite population covariance matrix. Let  $H_p$  be the spectral distribution of  $\Sigma_p$ , that is, the discrete uniform distribution on its eigenvalues  $l_i$ , sorted so that  $l_1 \geq l_2 \geq \dots \geq l_p$ . Its cumulative distribution function is defined as  $H_p(x) = p^{-1} \sum_{i=1}^p I(l_i \leq x)$ . The  $l_i$  are the population variances of the principal components.

The null hypothesis of identity  $\Sigma_p = I_p$  is equivalent to  $H_p = \delta_1$ , where  $\delta_c$  is the point mass at  $c$ . The spiked alternative  $\Sigma_p = I_p + \sum_{j=1}^k h_j v_j v_j^\top$ , for orthonormal  $v_j$ , is equivalent to  $H_p = (1 - k/p)\delta_1 + p^{-1} \sum_{j=1}^k \delta_{1+h_j}$ . This expresses the null and spiked alternative in terms of the spectrum of  $\Sigma_p$ .

We consider a more general *local alternatives model*. Let  $H = d^{-1} \sum_{i=1}^d \delta_{t_i}$  and  $G_j = h^{-1} \sum_{i=1}^h \delta_{s_{ij}}$ ,  $j = 0, 1$  be fixed probability distributions on  $[0, \infty)$ . Under the null, we take the eigenvalues to be  $t_1, t_2, \dots, t_d$  each with multiplicity  $m$ , and  $s_1^0, s_2^0, \dots, s_h^0$ . Under the alternative, the eigenvalues are  $t_i$  with the same multiplicity, and  $s_1^1, s_2^1, \dots, s_h^1$ .

In this paper,  $d$  and  $h$  are fixed constants, so the total number of eigenvalues is  $p = dm + h$ , and  $h$  of them differ between null and alternative. Without loss of generality, we can take  $p \rightarrow \infty$  along such a subsequence (as  $m \rightarrow \infty$ ).

We can write this sequence of null hypotheses  $H_{p,0}$  and alternatives  $H_{p,1}$  as

$$(1) \quad H_{p,0} : H_p = (1 - hp^{-1})H + hp^{-1}G_0,$$

$$(2) \quad H_{p,1} : H_p = (1 - hp^{-1})H + hp^{-1}G_1.$$

Taking  $H = \delta_1$ ,  $G_0 = \delta_1$  and  $G_1 = h^{-1} \sum_{j=1}^h \delta_{1+h_j}$ , the above null generalizes the hypothesis of identity  $H_p = \delta_1$  against spiked alternatives.

In our spiked paper,  $H$  is a finite mixture of point masses. The CLT of [Bai and Silverstein \(2004\)](#) requires a fixed sequence of covariance matrices, implying that  $H_p$  are finite mixture of point masses. Further, we need the form  $H_p = (1 - hp^{-1})H + hp^{-1}G_0$  for our argument via the weak derivative. This explains why  $H$  must be a finite mixture of point masses in our paper.

However, we insist that this is not a significant limitation of the proposed methodology. The reason is twofold. First,  $d$  and  $h$  can be arbitrarily large constants. Therefore, our model is still quite flexible. Second, it is more natural to test hypotheses about the spectrum  $H_p$  than about the limiting spectrum  $H$ . The limiting spectrum  $H$  is a purely theoretical quantity that helps with the analysis. However, it seems more natural to test hypotheses about the finite-sample quantity  $H_p$ . Now,  $H_p$  is usually a discrete mixture of point masses, so we can use the methods proposed in this paper.

As explained later in Section 2.4, there are some analogies between our model and optimal testing under local alternatives in classical statistics. Building on this analogy, we call  $h$  the *local parameter*.

We will construct tests based on linear spectral statistics (LSS)  $T_p(\varphi) = \text{tr}(\varphi(\widehat{\Sigma})) = \sum_{i=1}^p \varphi(\lambda_i)$ , where  $\widehat{\Sigma} = n^{-1} X_{n \times p}^\top X_{n \times p}$  is the sample covariance matrix, and  $\lambda_i$  are its eigenvalues. We will find the *optimal LSS* for the hypothesis testing problem (1) versus (2), when the sample size  $n$  and dimension  $p$  grow such that  $\gamma_p = p/n \rightarrow \gamma > 0$ . In fact, we will assume that  $\gamma_p = \gamma$ , so that  $\gamma$  must be rational. This is not a limitation, because in practice we always have *finite*  $n, p$ , and we can set  $\gamma := p/n$  to use our methods.

In this model, the Marchenko–Pastur *forward map*—or simply *Marchenko–Pastur map*—describes the weak limit of the spectral distribution  $F_p$  of  $\widehat{\Sigma}$ . If the entries of  $Z_{n \times p}$  come from an infinite array of i.i.d. variables with mean zero and variance 1, and  $H_p \Rightarrow H$  weakly, then with probability 1,  $F_p \Rightarrow \mathcal{F}_\gamma(H)$  for a probability measure  $\mathcal{F}_\gamma(H)$  [Marchenko and Pastur (1967), Bai and Silverstein (2009)]. We will assume  $H \neq \delta_0$ . An example is the autoregressive covariance matrix of order 1, where the entries of  $\Sigma_p$  are  $\Sigma_p[i, j] = \rho^{|i-j|}$ ,  $\rho \in (0, 1)$ ; for other examples, see Dobriban and Wager (2015).

The Marchenko–Pastur map  $\mathcal{F}_\gamma$  has a smoothing effect: for any  $H$ ,  $\mathcal{F}_\gamma(H)$  has a continuous density for all  $x \neq 0$ , and also for  $x = 0$  if  $\gamma < 1$  [Silverstein and Choi (1995)]. If  $\gamma > 1$ , the so-called companion limit empirical spectral distribution (ESD)  $\underline{F}$ , defined by  $\underline{F} = \gamma \mathcal{F}_\gamma(H) + (1 - \gamma)I_{[0, \infty)}$  has a density at zero. We will find it convenient to work with this distribution. The companion limit ESD is the limit of the spectral distribution of the matrix  $\widehat{\underline{\Sigma}} = n^{-1} X_{n \times p} X_{n \times p}^\top$ .

The asymptotic distribution of the LSS is known for smooth functions  $\varphi$ . Let  $\mathcal{I} = [a, b]$  be a compact interval whose interior includes  $[\liminf l_p(\Sigma_p)I(\gamma \in (0, 1))(1 - \sqrt{\gamma})^2, \limsup l_1(\Sigma_p)(1 + \sqrt{\gamma})^2]$  for both null and alternative  $\Sigma_p$  sequences, where we assume  $l_1(\Sigma_p)$  is uniformly bounded above. This interval includes the support of the limiting ESD  $\mathcal{F}_\gamma(H)$  [Bai and Silverstein (2009)]. Let  $\mathcal{H}(\mathcal{I})$  be the set of complex analytic functions on some open domain of  $\mathbb{C}$  containing  $\mathcal{I}$ , and let  $\varphi \in \mathcal{H}(\mathcal{I})$ . Suppose that the i.i.d. real standardized random variables  $Z_{n \times p}[i, j] = Z[i, j]$  come from an infinite array, with  $\mathbb{E}[Z[i, j]^4] = 3$ .

The CLT for linear spectral statistics of Bai and Silverstein (2004) implies that the centered test statistics converge weakly:  $T_p(\varphi) - p \int \varphi(x) d\mathcal{F}_{\gamma_p}(H_p) \Rightarrow \mathcal{N}(m_\varphi, \sigma_\varphi^2)$  under the null and alternative, for a certain mean  $m_\varphi$  and variance  $\sigma_\varphi^2$ . The limit parameters depend on  $H$  and  $\gamma$ . We focus on variables whose fourth moment matches the Gaussian distribution, but a similar approach should work for more generally, using the CLT of Pan and Zhou (2008) for diagonal  $\Sigma_p$  or the CLT of Zheng, Bai and Yao (2015).

Recall that the Stieltjes transform of a signed measure  $\mu$  on  $[0, \infty)$  is defined as the map  $m : \mathbb{C} \setminus [0, \infty) \rightarrow \mathbb{C}$ ,  $m(z) = \int (x - z)^{-1} d\mu(x)$ . Let  $v(z) = v_\gamma(z; H)$  be the Stieltjes transform of the limiting companion ESD  $\underline{F}$ . The limit  $v(x) =$

$\lim_{z \rightarrow x} v(z)$  exists for all  $x \in \mathbb{R} \setminus \{0\}$  [Silverstein and Choi (1995)]. We will also need the kernel (well defined a.s. with respect to Lebesgue measure on  $\mathbb{R}$ ):

$$(3) \quad k(x, y) = k_{\gamma}(x, y; H) = \frac{1}{2\pi^2} \log \left( 1 + 4 \frac{\Im(v(x))\Im(v(y))}{|v(x) - v(y)|^2} \right).$$

Note that  $k \neq 0$  only within the support of  $\mathcal{F}_{\gamma}(H)$ .

**2.1. Main results.** In the above model, the optimal LSS depends on the *weak derivative*  $\delta\mathcal{F}_{\gamma}$  of the Marchenko–Pastur map. For two probability measures  $H, G$ , we define this as the signed measure arising in the weak limit

$$(4) \quad \delta\mathcal{F}_{\gamma}(H, G) = \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{F}_{\gamma}((1 - \varepsilon)H + \varepsilon G) - \mathcal{F}_{\gamma}(H)}{\varepsilon}.$$

We will show in Theorem 4.1 that the limit exists. To find the optimal LSS, we will first give an asymptotically equivalent normal test for *fixed* LSS.

**THEOREM 2.1** (Asymptotically equivalent normal test). *Consider the problem of testing for weak PCs in the local alternatives model (1) versus (2). For each  $\varphi \in \mathcal{H}(\mathcal{I})$ , there is a sequence of constants  $c_p$  such that under the null  $H_{p,0}$ , one has  $T_p(\varphi) - c_p \Rightarrow \mathcal{N}(0, \sigma_{\varphi}^2)$ , while under the alternative  $H_{p,1}$ , one has  $T_p(\varphi) - c_p \Rightarrow \mathcal{N}(\mu_{\varphi}, \sigma_{\varphi}^2)$ .*

*The mean and variance are*

$$(5) \quad \mu_{\varphi} = -h \int_{\mathcal{I}} \varphi'(x) \Delta(x) dx \quad \text{and}$$

$$(6) \quad \sigma_{\varphi}^2 = \int_{\mathcal{I}} \int_{\mathcal{I}} \varphi'(x) \varphi'(y) k(x, y) dx dy.$$

*Here,  $\Delta = \delta\mathcal{F}_{\gamma}(H, G_1) - \delta\mathcal{F}_{\gamma}(H, G_0)$  is the difference of the weak derivatives  $\delta\mathcal{F}_{\gamma}(H, G_i)$ , and  $k$  denotes the kernel defined in (3).*

The proofs of the results in this section are outlined in Section 4.1. Therefore, using the linear spectral statistic  $T_p(\varphi)$  is asymptotically equivalent to a hypothesis test of a distribution  $\mathcal{N}(0, \sigma_{\varphi}^2)$  against  $\mathcal{N}(\mu_{\varphi}, \sigma_{\varphi}^2)$ . The next step is to optimize over LSS  $\varphi$ . In analogy to the asymptotic theory of optimal testing in i.i.d. models, we will call  $\theta(\varphi) = \mu_{\varphi}/\sigma_{\varphi}$  the *efficacy* of a test sequence  $T_p(\varphi)$  [Lehmann and Romano (2005), page 536]. If  $\sigma_{\varphi} = 0$  while  $\mu_{\varphi} \neq 0$ , we define  $\theta(\varphi) = +\infty$ , because the efficacy in distinguishing  $\mathcal{N}(0, \sigma_{\varphi}^2)$  from  $\mathcal{N}(\mu_{\varphi}, \sigma_{\varphi}^2)$  is infinite. Similarly, if  $\sigma_{\varphi} = 0$  while  $\mu_{\varphi} = 0$ , define  $\theta(\varphi) = 0$ . With these definitions, one does not have to worry about dividing by 0.

We will maximize the efficacy over certain function classes  $\mathcal{X}$ :

$$(7) \quad \sup_{\varphi \in \mathcal{X}} \frac{\mu_{\varphi}}{\sigma_{\varphi}}.$$

The value of the optimization problem will be called the *efficacy over  $\mathcal{X}$* , and will be denoted  $\theta^*(\mathcal{X})$ . A function  $\varphi \in \mathcal{X}$  achieving this value will be called an *optimal LSS* over  $\mathcal{X}$ . Due to the quadratic nature of the the objective, it will be easier to optimize first over the space  $\mathcal{W}(\mathcal{I}) = \{\varphi : \mathcal{I} \rightarrow \mathbb{R} : \varphi'(x) \text{ exists for almost every } x \in \mathcal{I}; \text{ and } \varphi' \in L^2[\mathcal{I}]\}$ , using Hilbert space techniques.

Accordingly, we define the linear integral operator  $K = K_{\gamma, H} : L^2[\mathcal{I}] \rightarrow L^2[\mathcal{I}]$  induced by  $k$  in the usual way:  $K(\varphi)(x) = \int_{\mathcal{I}} k(x, y)\varphi(y) dy$ . Since  $k$  is a logarithmically weakly singular kernel [Bai and Silverstein (2004), page 564],  $K$  is compact [see Kress (2014), pages 29 and 62, for this property]. We write  $\langle \cdot, \cdot \rangle$  for the inner product on  $L^2[\mathcal{I}]$ ,  $\text{Im}(K) = \{Kl : l \in L^2(\mathcal{I})\}$  for the image of  $K$ , and  $\overline{\text{Im}}(K)$  for the closure of  $\text{Im}(K)$ . We now find the optimal LSS.

**THEOREM 2.2** [Optimal linear spectral statistics over  $\mathcal{W}(\mathcal{I})$ ]. *Consider the optimization of the efficacy over  $\mathcal{W}(\mathcal{I})$ . The following dichotomy arises:*

1. *If  $\Delta \in \text{Im}(K)$ , then the efficacy over  $\mathcal{W}(\mathcal{I})$  equals  $h \cdot \langle \Delta, K^+ \Delta \rangle^{1/2} < \infty$ . The optimal linear spectral statistics over  $\mathcal{W}(\mathcal{I})$  are given by a Fredholm integral equation of the first kind for their derivatives:*

$$(8) \quad K(\varphi') = -\eta \Delta,$$

where  $\eta > 0$  is any constant.

2. *On the other hand, if  $\Delta \notin \text{Im}(K)$ , then the efficacy over  $\mathcal{W}(\mathcal{I})$  equals  $+\infty$ . If in addition  $\Delta \notin \overline{\text{Im}}(K)$ , the optimal LSS are all functions  $\varphi \in \mathcal{W}(\mathcal{I})$  with  $K(\varphi') = 0$  and  $\langle \Delta, \varphi' \rangle < 0$ .*

This gives an equation for the optimal LSS, which we call the *optimal LSS equation*. Since the equation does not depend on  $h$ , the optimal LSS is *uniformly optimal* against all  $h > 0$ . If the equation is not solvable in  $L^2(\mathcal{I})$ , we will construct a sequence  $\varphi_n \in \mathcal{W}(\mathcal{I})$  with efficacies  $\theta(\varphi_n) \rightarrow \infty$ , showing that the supremum of asymptotic power over  $\mathcal{W}(\mathcal{I})$  is unity.

We now return to smooth LSS. While the solution of the optimal LSS may not be an analytic function, we will show that analytic functions in  $\mathcal{H}(\mathcal{I})$  have the same maximum power as functions in  $\mathcal{W}(\mathcal{I})$ . Denoting the centered test statistics  $\tilde{T}_p(\varphi) = T_p(\varphi) - p \int \varphi(x) d\mathcal{F}_{\gamma_p}(H_p) - m_\varphi$ , we consider two-sided testing procedures that reject  $H_{p,0}$  if  $\tilde{T}_p(\varphi) \notin [t_\varphi^-, t_\varphi^+]$  for some constants  $t_\varphi^- < t_\varphi^+$ . Our goal is to optimize over smooth functions  $\varphi \in \mathcal{H}(\mathcal{I})$  and the critical values  $t_\varphi^- < t_\varphi^+$ . The maximal asymptotic power is defined as

$$\beta = \sup_{\varphi \in \mathcal{H}(\mathcal{I}), t_\varphi^- < t_\varphi^+} \lim_{p \rightarrow \infty} \mathbb{P}_{H_{p,1}}(\tilde{T}_p(\varphi) \notin [t_\varphi^-, t_\varphi^+]).$$

We find an expression for the power, depending on the null, the spikes, and the local parameter. This requires the generalized inverse  $K^+$  of  $K$ , the linear operator assigning to each  $\Delta \in \text{Im}(K)$  the minimum norm solution to the equation  $Kl = \Delta$  [see, e.g., Groetsch (1977), page 115].



**THEOREM 2.3 (Asymptotic power).** *Among tests based on linear spectral statistics  $T_p(\varphi)$  for  $\varphi \in \mathcal{H}(\mathcal{I})$  with asymptotic level  $\alpha \in (0, 1)$ , the maximal asymptotic power is*

$$\beta = \begin{cases} \Phi(z_\alpha + h\langle \Delta, K^+ \Delta \rangle^{1/2}), & \text{if } \Delta \in \text{Im}(K), \\ 1, & \text{if } \Delta \notin \text{Im}(K). \end{cases}$$

Here,  $\Delta = \delta\mathcal{F}_\gamma(H, G_1) - \delta\mathcal{F}_\gamma(H, G_0)$  is the difference of the weak derivatives,  $K$  is the compact operator induced by the kernel (3), and  $K^+$  is the pseudoinverse of  $K$ .

This shows that there are two possibilities, depending on the relation between the null and the alternative. If  $\Delta \in \text{Im}(K)$ , the asymptotic power depends on the norm of  $\Delta$  via  $\langle \Delta, K^+ \Delta \rangle^{1/2}$ . This is reasonable, as a “larger” derivative  $\Delta$  perturbs the null more, and should be easier to detect. A larger local parameter  $h > 0$  also leads to more power, as there are more spikes.

The second case,  $\Delta \notin \text{Im}(K)$ , can occur—for instance—if the alternative sample spikes separate from the bulk. In certain spiked models, the existence of a threshold beyond which the top eigenvalue separates from the bulk was shown for complex-valued Gaussian white noise in Baik, Ben Arous and P\'ech\'e (2005), and for correlated noise in Benaych-Georges and Nadakuditi (2011), Bai and Yao (2012) [see also Yao, Zheng and Bai (2015), Chapter 11]. While the models differ slightly between the authors, the location of the phase transition is the same.

For large spikes, we will show in Section 2.5 that the weak derivative  $\delta\mathcal{F}_\gamma$  has mass outside of the support  $S$  of  $\mathcal{F}_\gamma(H)$ . Hence, the distribution function  $\Delta$  is not in the image of  $K$ , which is supported on  $S$ . Thus, there is full power above the phase transition.

Intuitively,  $\Delta \in \text{Im}(K)$  should correspond to spikes below the phase transition. Indeed, in this case  $\Delta$  is supported within  $S$ . However, it is not clear that  $\Delta$  belongs to the image of the compact operator  $K$ . Showing this requires a more detailed operator-analytic study of  $K$ . We leave this for future research.

## 2.2. Examples of optimal LSS; Numerical results.

**2.2.1. Standard spiked model.** We take a detour to illustrate the optimal LSS in two simple cases. First, in the “standard spiked model” introduced in Johnstone (2001), the null is specified by  $H = \delta_1$  and  $G_0 = \delta_1$ , while the alternative has  $G_1 = \delta_t$ . We take the aspect ratio  $\gamma = 1/2$ . The well known BBP phase transition [Baik, Ben Arous and P\'ech\'e (2005)] states that for a “subcritical” spike  $t$  below the “phase transition” (PT) threshold  $1 + \sqrt{\gamma} \approx 1.7$ , the corresponding “sample spike” moves to the top of the bulk spectrum. For a “supercritical” spike  $t$  above the PT threshold, the sample spike moves to a value  $z(t) = t[1 + \gamma/(t - 1)]$  above the bulk edge.





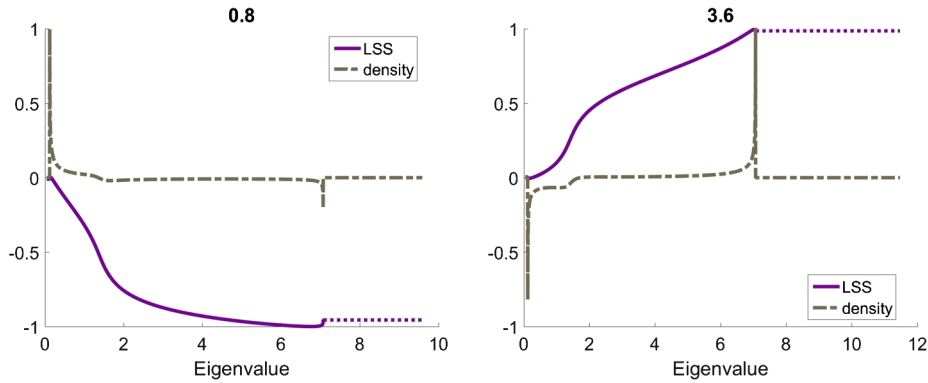


FIG. 2. Density of  $\delta\mathcal{F}_\gamma(H, G_1)$  and optimal LSS with  $H = 2^{-1}(\delta_1 + \delta_3)$ ,  $G_1 = \delta_t$ ,  $\gamma = 1/2$ . On the left plot, the spike  $t = 0.8$ ; while on the right plot  $t = 3.6$ ; both are subcritical.

Our theory only specifies the optimal LSS within the support of the Marchenko–Pastur map—and we extend it as a constant to the complement; see Section 3. This is illustrated by the dotted line.

For a supercritical spike, there is more latitude in the choice of the optimal LSS. Here, we set it to 0 on the support of the bulk and to unity above the location of the sample spike  $z(t)$ , interpolating by an Epanechnikov kernel (see Section 3).

**2.2.2. Local alternatives.** Next, we consider a general instance of local alternatives, with a mixture  $H = 2^{-1}(\delta_1 + \delta_3)$ , and  $G_0 = H$ . In this background noise, we want to test for the presence of a PC with magnitude  $t$ , corresponding to  $G_1 = \delta_t$ .

We show the density of  $\delta\mathcal{F}_\gamma(H, G_1)$ , and the optimal LSS for  $\gamma = 1/2$  (Figure 2) and  $\gamma = 1/10$  (Figure 3). We consider two values for  $t$ , 0.8 and 3.6, both of which turn out to be subcritical.

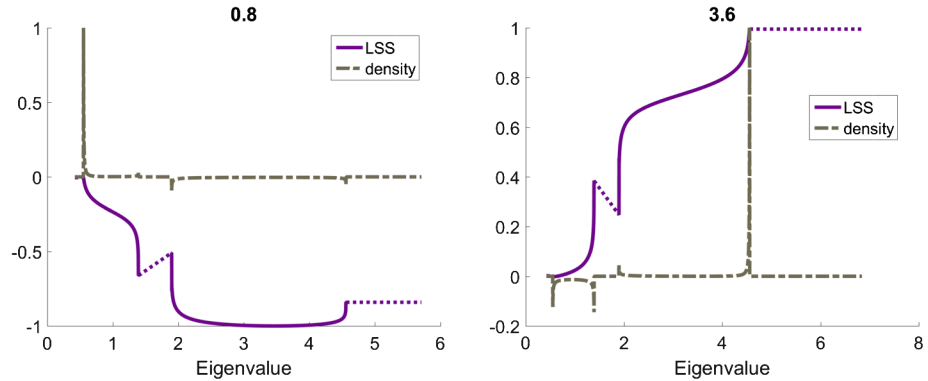


FIG. 3. The same plot as Figure 2, except with  $\gamma = 1/10$ .

We observe the following:

1. *The density of  $\delta\mathcal{F}_\gamma$* : For  $\gamma = 1/10$ , the bulk of sample eigenvalues has two components; for  $\gamma = 1/2$ , it has only one. This affects both the weak derivative and the optimal LSS. For  $\gamma = 1/2$ , the singularities of  $\delta\mathcal{F}_\gamma$  are similar to the standard case. For  $\gamma = 1/10$ , the spike seems to perturb positively the component of the bulk containing it, and perturb negatively the other component.

2. *The LSS*: The optimal LSS are highly nonlinear, and differ a great deal between the four settings ( $\gamma \in \{1/10, 1/2\}$ ,  $t \in \{0.8, 3.6\}$ ). Our theorem only specifies the LSS within the support of the bulk  $\mathcal{S}$ . We extend them by linear interpolation outside; see Section 3. This is indicated by the dotted lines.

In general, the optimal LSS are “large” where the density of  $\delta\mathcal{F}_\gamma$  is positive. However, they have nontrivial shapes; in particular, they show sharp “peaks” at the edges. It appears that the test statistics have qualitatively novel properties.

**2.3. Simulation results.** To illustrate the finite-sample performance of our methods, we present the results of a Monte Carlo (MC) simulation (Figure 4). The eigenvalues of an autoregressive covariance matrix of order 1 (AR-1) with  $\Sigma_{ij} = \rho^{|i-j|}$ , and  $\rho = 0.5$  make up the null  $H$ . The sample size is  $n = 500$  while  $\gamma = 1/2$ , so the dimension is  $p = 250$ . Experiments reported in Appendix C [see Dobriban (2017)] show parallel results for  $\gamma = 2$ . For large  $p$ , it is well known that the largest eigenvalue of  $\Sigma$  is approximately  $(1 + \rho)/(1 - \rho)$ , which equals three (3) in our case. The null spike  $s^0 = 1$  is buried within the population bulk, while the alternative spike  $s^1 = 3.5$  sticks out. The spike  $s^1$  can be seen clearly in the histogram in the top left plot of Figure 4.

We generate a random Gaussian matrix  $X = Z\Sigma^{1/2}$  with this covariance structure. The histograms of the sample eigenvalues—for both null and alternative—are in the top right plot of Figure 4. The top sample spike does not separate from the sample bulk. This is reinforced by the scree plots of the top 10 eigenvalues under null and alternative, shown in the middle row, left plot of Figure 4. The two look nearly indistinguishable.

Is it possible to distinguish the two distributions? Our approach is to use the optimal LSS, plotted in the middle row, right plot of Figure 4. This LSS puts a large weight on the top eigenvalues, while also putting a smaller weight on the middle eigenvalues; and it is extended as a constant outside the bulk. This can indeed distinguish between the two distributions—in the bottom left plot of Figure 4 we show the histogram of the LSS over 200 MC samples; using the empirical mean and standard error under the null to standardize both histograms. The distributions look approximately normal. Under the alternative, the distribution has mean approximately equal to two (2), which is encouraging.

**2.3.1. Increasing the spike.** To examine the power more thoroughly, we perform a broader MC simulation, increasing the alternative spike  $s^1$  from 1 to 5.

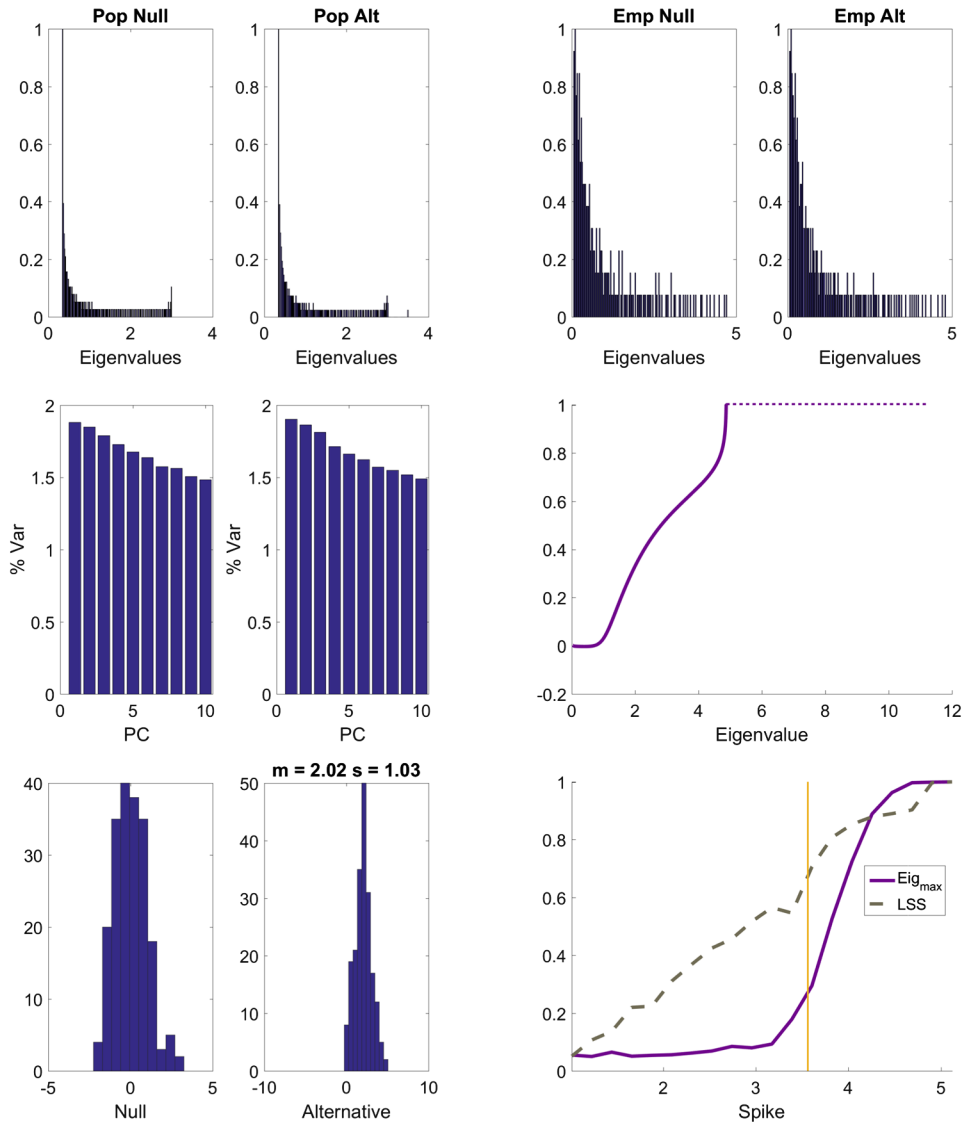


FIG. 4. Simulation results. Top row, left: Histogram of population eigenvalues under null and alternative. Top row, right: Histogram of sample eigenvalues under null and alternative, for one MC instance. Middle row, left: Scree plot of top 10 sample eigenvalues under null and alternative, for one MC instance. Middle row, right: pointwise plot of optimal LSS. Bottom row, left: Histogram LSS under null and alternative, over 200 MC instances. Bottom row, right: Power of optimal LSS and top eigenvalue-based tests as a function of the position of the spike under the alternative.

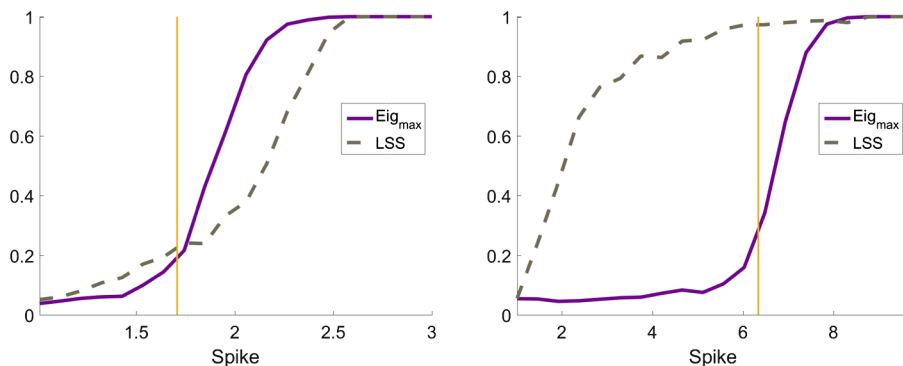


FIG. 5. Power of optimal LSS and top eigenvalue based tests for increasing alternative spike. Left:  $\rho = 0$  (identity matrix). Right:  $\rho = 0.7$ .

We compare the test which rejects if the top eigenvalue is large which rejects if the LSS is large. We set the critical values based on the empirical distribution under the null, to ensure type I error control at level  $\alpha = 0.05$ . We record 1000 MC iterates with sample size  $n = 2000$  and other parameters kept the same as before.

The results—in the bottom right plot of Figure 4—show that the LSS-based test has power even below the PT threshold, while the top eigenvalue test does not. The vertical line shows the location of the asymptotic PT.

To get a broader view of the achievable power in various scenarios, we repeat the last experiment for two additional values of  $\rho$ . We use  $\rho = 0$ —corresponding to an identity covariance matrix—and  $\rho = 0.7$ , which allows for higher correlations. In Figure 5, we show the results recorded over 1000 MC iterates with sample size  $n = 500$  and  $\gamma = 1/2$ .

For the identity case, the optimal LSS has weak finite sample power. The top eigenvalue test surpasses it above the PT. In contrast, for  $\rho = 0.7$ , the LSS has a lot of power below the PT. The broad conclusion of these experiments is that for eigenvalue distributions that are “widely spread,” one has indeed the power to detect spikes below the PT. The larger power is not due only to the larger phase transition location. In simulations reported in Appendix C, we show that the power does not increase perceptibly by scaling the sample covariance matrix  $\Sigma = \sigma^2 I_p$  so that the two phase transitions agree.

**2.4. Comparison to classical optimal testing.** There are connections between our work and the large literature on optimal testing under local alternatives. In that work, dating back to the pioneering results of Le Cam, Hájek, and others in the 1960s, the limiting power of tests under sequences of local alternatives is calculated using the contiguity of the null and alternative hypotheses. In finite-dimensional smooth parametric models, the local asymptotic normality of the log-likelihood ratio process holds, leading to strong optimality results. See [Lehmann and Romano \(2005\)](#), [van der Vaart \(1998\)](#) for graduate-level introductions.

It is worth comparing our theorems to some classical results. For instance, Theorem 13.6.1 from [Lehmann and Romano \(2005\)](#) considers nonparametric testing of a smooth statistical functional  $\theta(P)$  based on  $n$  i.i.d. samples from  $P \in \mathbb{P}$ , where  $\mathbb{P}$  is the set of all probabilities on a space. The power at  $P = P_{u,0}$  along a one-dimensional quadratic mean differentiable (q.m.d.) submodel  $P_{u,t}$ , with score function  $u$ , against local alternatives  $t = hn^{-1/2}$  is bounded by  $\Phi(z_\alpha + h\langle \tilde{\theta}_P, u \rangle_P / |\tilde{\theta}_P|_P)$ , where  $\tilde{\theta}_P$  is a type of weak derivative of  $\theta$  along  $P_{u,t}$  and  $\langle u, v \rangle_P = \int uv dP$ . This is similar to our Theorem 2.1: the efficacy of a LSS is  $-h\langle \varphi', \Delta \rangle / \langle \varphi', K\varphi' \rangle^{1/2}$  for the natural inner product on  $L^2(\mathcal{I})$ . The similarity is due to the asymptotic normality of both problems; however, the probabilistic reasons for normality are of course different.

**2.5. Full power above the phase transition.** Here, we show that the optimal LSS have full power when the spikes are above the known phase transition threshold from classical spiked models. This relies on studying the weak derivative of the Marchenko–Pastur map. For simplicity, we will let  $G_0 = H$ , so that  $\Delta = \delta\mathcal{F}_\gamma(H, G_1)$ , but similar results hold for general  $G_0$ .

We want to find out if the weak derivative has mass outside of the support  $S = \text{Supp}(\mathcal{F}_\gamma(H))$ . In such a case,  $\Delta(x) \neq 0$  must occur on a set of positive measure outside  $S$ . Since the kernel is supported on  $S$ , the optimal LSS equation cannot have a solution. This argument shows that the asymptotic power is unity.

We say that a spike  $s_j$  is *above the phase transition* if  $s_j \in -1/v(S^c)$ , where  $v$  is the companion Stieltjes transform of  $\mathcal{F}_\gamma(H)$ . This is consistent with the previous definitions for the “generalized” spiked model in [Benaych-Georges and Nadakuditi \(2011\)](#), [Bai and Yao \(2012\)](#) [see also [Yao, Zheng and Bai \(2015\)](#), Chapter 11]. Under certain orthogonal invariance assumptions, the phase transition is derived in [Benaych-Georges and Nadakuditi \(2011\)](#) by first reducing it to the case  $\Sigma_p = I_p + M_p$  for diagonal  $M_p$ , then working with the determinantal equation  $\det(\lambda I_p - n^{-1}X^\top X \cdot (I_p + M_p)) = 0$  for the eigenvalues of  $n^{-1}X^\top X \cdot \Sigma = \Sigma^{-1/2}\hat{\Sigma}_p\Sigma^{1/2}$ . The PT locations can also be computed numerically (Section 3).

Our goal is to prove the following result.

**THEOREM 2.4 (Full power above phase transition).** *Suppose that in the local alternatives model we have  $H = d^{-1} \sum_{i=1}^d \delta_{t_i}$ ,  $G_0 = H$ , and  $G_1 = h^{-1} \sum_{i=1}^h \delta_{s_i}$ . If there are spikes above the phase transition—so that  $s_j \in -1/v(S^c)$  for some  $j$ —then the asymptotic power of the optimal LSS is unity.*

**PROOF.** If there is a spike  $s_j$ —with mass  $u_j$  in  $G_1$ —above the phase transition, then  $\delta\mathcal{F}_\gamma(H, G_1)$  has a point mass of weight  $\gamma u_j > 0$  for some  $x \in S^c$  by Proposition 2.5 (to be proved next). Therefore, the distribution function  $\Delta$  has a

discontinuity at  $x$ , so it is nonzero on a subset of  $S^c$  with positive Lebesgue measure. Since the kernel  $k$  is zero on  $S^c$ ,  $\Delta$  is not in the image of  $K$ . By Theorem 2.1, the asymptotic power is unity.  $\square$

It remains to prove the following key proposition, which establishes properties of the weak derivative  $\delta\mathcal{F}_\gamma$ . It will be convenient to define the *spike forward map*  $\psi(s)$ , which for a population spike  $s$  and bulk  $H$ , gives the location of the sample spike under the effect of the bulk  $H$ . This is defined through its functional inverse, which is expressed as  $\psi^{-1}(x) = -1/v(x)$  [see Yao, Zheng and Bai (2015), Chapter 11]; and one can verify that  $\psi$  is well defined outside of the support of  $H$ . The values  $x \in S^c$  in the image of the spike forward map, that is, for which  $x = \psi(s_j)$  for some  $j$ , will be called the *sample spikes*. We study the weak derivative somewhat more generally than the setting of our main results, for arbitrary weighted mixtures of point masses.

**PROPOSITION 2.5** (Properties of the weak derivative). *Suppose the population bulk is  $H = \sum_{i=1}^k w_i \delta_{t_i}$ , with  $w_i > 0$  such that  $\sum_i w_i = 1$ . Suppose the spikes have distribution  $G = \sum_{j=1}^l u_j \delta_{s_j}$  with distinct  $s_j > 0$  and weights  $u_j > 0$  summing to one. Let the support of the forward map be  $S = \text{Supp}(\mathcal{F}_\gamma(H))$ , and consider the weak derivative  $\delta\mathcal{F}_\gamma(H, G)$ . Then:*

1.  $\delta\mathcal{F}_\gamma$  has a density at all  $x$  in the interior of  $S$ ,  $x \in \text{int}(S)$ .
2.  $\delta\mathcal{F}_\gamma$  has a point mass  $\gamma u_j$  at sample spikes  $x = \psi(s_j)$ , that is, for the values  $x \in S^c$  such that  $s_j = -1/v(x)$  for some  $j$ .
3.  $\delta\mathcal{F}_\gamma$  has zero density at all  $x$  outside  $\text{int}(S)$  that are not sample spikes.

The proof is postponed to Section A.3. The result sheds new light on the phase transition phenomenon. It shows that the population spikes  $s_j$  are “above the phase transition,” if and only if they create an isolated point mass in the weak derivative. We find this explanation illuminating.

**2.6. Sphericity tests—PCA with unknown scale.** Our entire framework can be extended to sphericity tests, which allow for an unknown scale parameter in PCA. Classically this corresponds to the composite null hypothesis  $\Sigma_p = \sigma^2 I_p$ , for some unknown  $\sigma^2 > 0$ . When studying PCA, the alternative hypothesis of interest is  $\Sigma_p = \sigma^2(I_p + \sum_{j=1}^k h_j v_j v_j^\top)$ , for orthonormal  $v_j$ . We will study the natural generalization of the local alternatives model where the  $p$ th problem is

$$(9) \quad H_{p,0} : H_p = \sigma^2[(1 - hp^{-1})H + hp^{-1}G_0] \quad \text{for some } \sigma^2 > 0,$$

$$(10) \quad H_{p,1} : H_p = \sigma^2[(1 - hp^{-1})H + hp^{-1}G_1] \quad \text{for some } \sigma^2 > 0.$$

Here,  $H, G_0$  and  $G_1$  are probability measures and the integer  $h > 0$  is the local parameter, with same properties as in the previous sections. When  $H = \delta_1$ ,  $G_0 = \delta_1$ ,  $h = k$  and  $G_1 = k^{-1} \sum_{j=1}^k \delta_{h_j+1}$ , this recovers the classical setup.



The null and alternative are both invariant with respect to orthogonal rotations and scaling. It is reasonable to consider tests based on the set of standardized eigenvalues  $\lambda_i/\hat{\sigma}^2$  of the sample covariance matrix, with  $\hat{\sigma}^2 = \hat{\sigma}_p^2 = p^{-1} \text{tr} \hat{\Sigma}$ . With Gaussian data, and when  $H = \delta_1$ ,  $G_0 = \delta_1$ , they form a set of maximal invariants with respect to rotations and scaling. Moreover, the standardized eigenvalues are distribution-free—or pivotal—under the null. Therefore, we consider linear *standardized* spectral statistics ( $\text{LS}^3$ ), which we define as  $S_p(\varphi) = \text{tr}(\varphi(\hat{\Sigma}/\hat{\sigma}^2)) = \sum_{i=1}^p \varphi(\lambda_i/\hat{\sigma}^2)$ .

Our goal will be to find the optimal  $\text{LS}^3$ . We first establish their asymptotic distribution. We assume the same model as in Section 2.1. We consider smooth functions  $\varphi \in \mathcal{H}(\mathcal{I}/m_1)$ , where  $m_1 = \int x dH(x)$ . This is because the eigenvalues  $\lambda_i$  still belong to the compact interval  $\mathcal{I}$  almost surely and—as we will see in the proofs— $\hat{\sigma}^2 \rightarrow m_1 > 0$  almost surely. We will use the notation  $\mathcal{F}_{\gamma_p}(g(x)) = \int g(x) d\mathcal{F}_{\gamma_p}(x)$  for the integral of a function  $g$  under  $\mathcal{F}_{\gamma_p}(H)$ .

**LEMMA 2.6 (CLT for  $\text{LS}^3$ ).** *For  $\varphi \in \mathcal{H}(\mathcal{I}/m_1)$ , under the null and alternative (9), (10) the linear standardized spectral statistics  $S_p(\varphi)$  are asymptotically normal. There is a sequence of constants  $c_p$  such that under  $H_{p,0}$ ,  $S_p(\varphi) - c_p \Rightarrow \mathcal{N}(0, \sigma_{\varphi,s}^2)$ , while under  $H_{p,1}$ ,  $S_p(\varphi) - c_p \Rightarrow \mathcal{N}(\mu_{\varphi,s}, \sigma_{\varphi,s}^2)$ , for a mean shift  $\mu_{\varphi,s}$  and variance  $\sigma_{\varphi,s}^2$ . The mean shift and variance are the same as those in the asymptotic distribution of the LSS  $T_p(j)$ , where  $j \in \mathcal{H}(\mathcal{I})$  is defined by*

$$(11) \quad j(x) = \varphi\left(\frac{x}{m_1}\right) - \frac{x}{m_1} \mathcal{F}_{\gamma}\left(\frac{x}{m_1} \varphi'\left(\frac{x}{m_1}\right)\right).$$

The lemma, proved in Section A.8, states that the  $\text{LS}^3$  for  $\varphi$  and the LSS for  $j$  are asymptotically equivalent. Hence, we will find the optimal  $\text{LS}^3$  by optimizing over LSS of the form (11). By scale invariance, we can restrict to working with  $\sigma^2 = 1$ , which implies  $m_1 = 1$ .

First, we characterize the LSS that are of the required form  $j(x) = \varphi(x) - x\mathcal{F}_{\gamma}(x\varphi'(x))$ . We claim that a function  $j \in \mathcal{H}(\mathcal{I})$  is of this form if and only if  $\mathcal{F}_{\gamma}(xj'(x)) = 0$ . Indeed, if  $j$  has this form, then  $\mathcal{F}_{\gamma}(xj'(x)) = \mathcal{F}_{\gamma}[x\varphi'(x) - \mathcal{F}_{\gamma}(x\varphi'(x))] = 0$ . On the other hand, if  $\mathcal{F}_{\gamma}(xj'(x)) = 0$ , then by taking  $f = j$ , clearly  $j$  is of the required form, as the second term cancels.

Therefore, we optimize the efficacy from (7) over the function class  $\mathcal{H}_0(\mathcal{I}) = \{\varphi \in \mathcal{H}(\mathcal{I}) : \mathcal{F}_{\gamma}(x\varphi'(x)) = 0\}$ . The constraint  $\mathcal{F}_{\gamma}(x\varphi'(x)) = 0$  is a linear equation  $\langle g, D \rangle = 0$  for the derivative  $g = \varphi'$ , with  $D = x d\mathcal{F}_{\gamma}(x) \in L^2(\mathcal{I})$ .  $D$  is an  $L^2$  function, because  $\mathcal{F}_{\gamma}$  has a continuous density except at 0, while the  $x$  term is null at 0. From the previous sections, it follows that the efficacy optimization over a space  $\mathcal{X}$  can be written in terms of  $g = \varphi'$  as

$$\sup_{g \in \mathcal{X}} -h \frac{\langle g, \Delta \rangle}{\langle g, Kg \rangle^{1/2}} \quad \text{s.t. } \langle g, D \rangle = 0.$$

As in the previous section, at first we will optimize over  $g \in L^2(\mathcal{I})$ , and then extend to analytic functions. Consider the projection operator  $P$  into the orthogonal complement of the one-dimensional space spanned by  $D$ :  $Pg = g - D\langle g, D \rangle / \|D\|^2$ . Optimizing subject to the linear constraint is equivalent to optimizing over the set  $g \in \text{Im}(P)$ —or with  $g = PL$  to solving the problem

$$\sup_{l \in L^2(\mathcal{I})} -h \frac{\langle PL, \Delta \rangle}{\langle PL, KPL \rangle^{1/2}}.$$

Denoting  $\Delta_1 = PL$  and  $K_1 = PKP$ , this reduces to the type of optimization problem solved previously [see (7)]. Putting this together with Lemma 2.6 and the analogue of Theorem 4.3 for  $\text{LS}^3$ —whose statement and proof is omitted due to its similarity to Theorem 4.3—we obtain the power of  $\text{LS}^3$ .

We consider tests that reject the null if  $S_p(\varphi) - c_p \notin [t_\varphi^-, t_\varphi^+]$  for some function-dependent constants  $t_\varphi^-, t_\varphi^+$ . By scale-invariance it is enough to consider  $\sigma^2 = 1$ . In this case, we denote the  $p$ th null and alternative distribution as  $H_{p,0}$  and  $H_{p,1}$ , respectively. The maximal asymptotic power of  $\text{LS}^3$  is

$$\beta_s = \sup_{\varphi \in \mathcal{H}(\mathcal{I}), t_\varphi^- < t_\varphi^+} \lim_{p \rightarrow \infty} \mathbb{P}_{H_{p,1}}(S_p(\varphi) \notin [t_\varphi^-, t_\varphi^+]).$$

**THEOREM 2.7 (Asymptotic power of  $\text{LS}^3$ ).** *Consider scale-invariant tests for detecting weak PCs based on linear standardized spectral statistics  $S_p(\varphi)$ . Suppose  $\varphi \in \mathcal{H}(\mathcal{I})$  and the tests have asymptotic level  $\alpha \in (0, 1)$ . The maximal asymptotic power is*

$$\beta_s = \begin{cases} \Phi(z_\alpha + h\langle K_1^+ \Delta_1, \Delta_1 \rangle^{1/2}), & \text{if } \Delta_1 \in \text{Im}(K_1), \\ 1, & \text{if } \Delta_1 \notin \text{Im}(K_1). \end{cases}$$

This result quantifies the loss of power due to restricting to scale-invariant  $\text{LS}^3$  from LSS. If  $\Delta, D \in \text{Im}(K)$ , it can be checked that  $\Delta_1 \in \text{Im}(K_1)$ . Moreover, the efficacy is  $\theta_s = h[\langle K^+L, \Delta \rangle - \langle K^+L, D \rangle^2 / \langle K^+D, D \rangle]^{1/2}$ . This shows that the efficacy is reduced from  $\theta = h\langle K^+L, \Delta \rangle^{1/2}$ , and the power loss depends on the “correlation” between  $\Delta$  and  $D$  with respect to  $K$ .

**3. Computation.** We now explain the computational details of our method. A MATLAB implementation, and the code to reproduce our computational experiments, is available at [github.com/dobriban/eigenedge](https://github.com/dobriban/eigenedge). Appendix B has additional details.

The computational problem is the following: Given a null distribution  $H = d^{-1} \sum_{i=1}^d \delta_{t_i}$ , spikes  $G_i = h^{-1} \sum_{j=1}^h \delta_{s_j^i}$ ,  $i = 0, 1$ , and an aspect ratio  $\gamma$ , compute the optimal LSS. For simplicity, we take all spikes in  $G_0$  subcritical—which is the only case we need in simulations—but the general case is similar. We will outline the needed steps, giving parameter choices (Table 2) and pseudocode in Appendix B.

**3.1. Computing  $v$  and the support.** First, we compute the companion Stieltjes transform  $v(x)$  of the limit ESD  $\mathcal{F}_\gamma(H)$  on a dense grid  $\{x_m\}$  on the real line (see Algorithm 1). We use the SPECTRODE method [Dobriban (2015)], which produces an approximation  $\tilde{v}(x)$  that depends on a user-specified accuracy parameter  $\varepsilon > 0$ , and converges to  $v(x)$  as  $\varepsilon \rightarrow 0$ . In Dobriban (2015), we showed that  $\pi^{-1}\Im(\tilde{v}(x))$  converges to the density  $\pi^{-1}\Im(v(x))$  of the limit ESD. An analogous argument shows that  $\tilde{v}(x)$  converges to  $v(x)$ . SPECTRODE also produces a converging approximation to the support  $S$  of  $\mathcal{F}_\gamma(H)$  as a union of closed intervals  $\tilde{S} = \bigcup_j [\tilde{l}_j, \tilde{u}_j]$ ,  $j = 1, \dots, J$ .

There are two cases—below and above the PT—which depend on whether or not  $\Delta \in \text{Im}(K)$ . As a proxy to this abstract statement, we check if the sample spikes corresponding to  $s_j^1$  belong to the support, as in Section 2.5. We have shown that  $\Delta \notin \text{Im}(K)$  if some sample spikes are outside the limit ESD. This is the first case that we handle (Algorithm 2). Second, if all sample spikes are in the support, we directly attempt to solve a discretized version of the optimal LSS equation (Algorithm 3).

**3.2. Above the PT.** From  $\tilde{v}(x)$  and the support, we check if there are any spikes above the phase transition by verifying if any *sample spike*  $\psi(s_j^1)$  falls outside the support:  $\psi(s_j^1) \notin \tilde{S}$  for any  $j$ . Recall here that  $\psi$  is the spike forward map from Section 2.5, and equals  $\psi(s) = s[1 + \gamma d^{-1} \sum_{i=1}^d t_i/(s - t_i)]$ , see Yao, Zheng and Bai (2015), Chapter 11, equation 11.15]. If there are spikes above the PT, by Theorem 2.2, the asymptotic power is unity. As candidate optimal LSS, inspired by the second part of Theorem 2.2, we consider smooth functions  $\varphi$  that equal unity in a small neighborhood of the sample spike, and zero on  $\tilde{S}$ .

Specifically, we take an LSS that has a small Epanechnikov kernel centered at each sample spike  $\psi(s_j^1)$ , and zero elsewhere (see Algorithm 2). Since the fluctuations of the spikes are asymptotically normal above the phase transition, we choose the width of the kernel as  $n_{SD} \cdot n^{-1/2} \hat{\sigma}_j$ . Here,  $n_{SD}$  is a constant given in Table 2,  $n$  is the sample size (provided as an optional input), and  $\hat{\sigma}_j$  is the asymptotic standard deviation of the sample spike,  $\hat{\sigma}_j^2 = 2[s_j^1]^2 \psi'(s_j^1)$ ; see Yao, Zheng and Bai (2015), Theorem 11.11, and also Onatski (2012), Theorem 2 for closely related earlier results. Moreover, we extend the LSS as a constant equal to unity in the direction pointing away from the support  $S$ , for any *extremal* spikes that fall above  $\max(S)$ , or below  $\min(S)$ . If the optional input  $n$  is not provided, we set  $n = (d + h)/\gamma$ , which is equivalent to assuming that  $p = d + h$ . A small refinement of this approach is described in the Appendix B.1.

**3.3. Below the PT.** If there are no spikes above the PT, we proceed to solve the optimal LSS equation (see Algorithm 3). The LSS is well defined only within the support  $S$  of the bulk  $\mathcal{F}_\gamma(H)$ , so we restrict to that subset of the grid. First, the kernel  $k$  is evaluated pointwise using  $\tilde{v}$ .

Next, we compute the difference  $\Delta$  of the weak derivatives (Algorithm 4). As explained in Dobriban (2015),  $v'(z)$  can be expressed in closed form as a function of  $v(z)$ . Hence, using equation (13) we can approximate the Stieltjes transforms of  $\delta\mathcal{F}_\gamma(H, G_i)$ . We find their density from the inversion formula for Stieltjes transforms, and their distribution by integrating the density numerically.

Finally, we need to solve the optimal LSS equation  $Kg = -\eta\Delta$  (where we set the constant  $\eta$  to 1 without loss of generality); see Algorithm 5. This is a Fredholm integral equation of the first kind with a logarithmically weakly singular kernel. There are many methods for solving such equations numerically [see Kress (2014)]. We implement two methods: A fast heuristic diagonal regularization method, and a slower but potentially more accurate collocation method. The details are given in Appendix B.2. Finally, Appendix B.4 reports the results of unit tests to verify the accuracy of the methods.

## 4. Proofs.

### 4.1. Main steps of the proofs.

4.1.1. *Weak derivative of the Marchenko–Pastur map.* We start by explaining the main steps in proving Theorems 2.1 (asymptotically normal equivalent) and 2.2 (optimal LSS equation). These lead to the proof of Theorem 2.3 (asymptotic power). Starting with Theorem 2.1, our assumptions imply that the Bai–Silverstein CLT for linear spectral statistics [Bai and Silverstein (2004), Theorem 1.1] applies both under the sequences of null and alternative hypotheses. Denoting—perhaps with a slight abuse of notation—by  $H_{p,i}$  the spectral distributions under null ( $i = 0$ ) and alternative ( $i = 1$ ), this shows that

$$\begin{aligned} \text{under } H_0 : T_p(\varphi) - p \int_{\mathcal{I}} \varphi(x) d\mathcal{F}_{\gamma_p}(H_{p,0}) &\Rightarrow \mathcal{N}(m_\varphi, \sigma_\varphi^2), & \text{while} \\ \text{under } H_1 : T_p(\varphi) - p \int_{\mathcal{I}} \varphi(x) d\mathcal{F}_{\gamma_p}(H_{p,1}) &\Rightarrow \mathcal{N}(m_\varphi, \sigma_\varphi^2). \end{aligned}$$

Here,  $m_\varphi, \sigma_\varphi^2$  are certain constants that are the same under the null and the alternative. Indeed, in Theorem 1.1 of Bai and Silverstein (2004), these limiting parameters are given by certain contour integrals that *only depend on the weak limit of the PSD*, and in our case these weak limits are both equal to  $H$ . The explicit form of these constants will only matter later. The important part is the *difference* in the centering terms, that is, the change from the argument of  $\mathcal{F}_{\gamma_p}$  from  $H_{p,0}$  to  $H_{p,1}$ . Therefore, the mean shift between the two hypotheses ought to equal

$$\mu_\varphi = \lim_{p \rightarrow \infty} p \int_{\mathcal{I}} \varphi(x) d[\mathcal{F}_{\gamma_p}(H_{p,1}) - \mathcal{F}_{\gamma_p}(H_{p,0})],$$

provided this limit is well defined. It is natural to conjecture that the signed measures  $D_p = p[\mathcal{F}_{\gamma_p}(H_{p,1}) - \mathcal{F}_{\gamma_p}(H_{p,0})]$  have a weak limit—and we will in fact

prove this. We can write

$$\begin{aligned} D_p &= p[\mathcal{F}_{\gamma_p}(H_{p,1}) - \mathcal{F}_\gamma(H_{p,1})] - p[\mathcal{F}_{\gamma_p}(H_{p,0}) - \mathcal{F}_\gamma(H_{p,0})] \\ &\quad + p[\mathcal{F}_\gamma(H_{p,1}) - \mathcal{F}_\gamma(H)] - p[\mathcal{F}_\gamma(H_{p,0}) - \mathcal{F}_\gamma(H)]. \end{aligned}$$

Since  $\gamma_p = \gamma$ , the first two terms are 0; if we relaxed the assumptions to  $\gamma_p \rightarrow \gamma$ , these limits would need to be evaluated. Therefore, by the definition of the weak derivative of the Marchenko–Pastur map (4), and by the definition of  $H_{p,i} = (1 - hp^{-1})H + hp^{-1}G_i$  the limit of  $D_p$  will be  $h \cdot [\delta\mathcal{F}_\gamma(H, G_1) - \delta\mathcal{F}_\gamma(H, G_0)]$ . Further,  $\varphi$  is continuous and bounded on  $\mathcal{I}$ , since by assumption  $\varphi'$  exists on  $\mathcal{I}$ . Therefore, by the definition of weak convergence of signed measures [see, e.g., Bogachev (2007), Chapter 8], the mean shift will be

$$(12) \quad \mu_\varphi = h \int_{\mathcal{I}} \varphi(x) d[\delta\mathcal{F}_\gamma(H, G_1) - \delta\mathcal{F}_\gamma(H, G_0)](x).$$

We are therefore naturally lead to the study of the weak derivative. We will study it in a slightly more general setting than above, allowing for *arbitrary compactly supported* probability distributions  $H$  and  $G$ .

For any signed measure  $\mu$ , it will be convenient to define the *companion measure*  $\underline{\mu} = (1 - \gamma)\mu + \gamma\delta_0$ . The companion Stieltjes transform of a measure  $\mu$  is then the Stieltjes transform of its companion  $\underline{\mu}$ . This terminology is consistent with the limit companion ESD, which we already used. Let  $\mathcal{P}_c$  be the set of compactly supported probability measures on  $([0, \infty), \mathcal{B}([0, \infty)))$ . It is known that for  $H \in \mathcal{P}_c$ , one has  $\mathcal{F}_\gamma(H) \in \mathcal{P}_c$  [Bai and Silverstein (2009)]. Our main result on the derivative of the Marchenko–Pastur map is the following.

**THEOREM 4.1** (Weak derivative of the Marchenko–Pastur map). *Let  $\mathcal{F}_\gamma : \mathcal{P}_c \rightarrow \mathcal{P}_c$  be the forward Marchenko–Pastur map, which takes the population spectral distribution  $H$  to the limit empirical spectral distribution  $\mathcal{F}_\gamma(H)$ . Then  $\mathcal{F}_\gamma$  has a well-defined weak derivative  $\delta\mathcal{F}_\gamma(\cdot, \cdot)$ , that is, for any  $H, G \in \mathcal{P}_c$ , the following weak limit exists as  $\varepsilon \rightarrow 0$ :*

$$\frac{\mathcal{F}_\gamma((1 - \varepsilon)H + \varepsilon G) - \mathcal{F}_\gamma(H)}{\varepsilon} \Rightarrow \delta\mathcal{F}_\gamma(H, G).$$

*The limit  $\delta\mathcal{F}_\gamma$  is a compactly supported signed measure with finite total variation, and has zero total mass:  $\delta\mathcal{F}_\gamma(\mathbb{R}) = 0$ . Furthermore:*

1. *The companion Stieltjes transform  $s(z)$  of the weak derivative can be expressed as*

$$(13) \quad s(z) = -\gamma v'(z) \int \frac{t}{1 + tv(z)} dv(t),$$

*where  $v = G - H$ , and  $v(z)$  is the companion Stieltjes transform of the limit empirical spectral distribution  $\mathcal{F}_\gamma(H)$ .*

2. Therefore, the weak derivative is linear in the second argument:  $\delta\mathcal{F}_\gamma(H, aP + bQ) = a\delta\mathcal{F}_\gamma(H, P) + b\delta\mathcal{F}_\gamma(H, Q)$  for all  $P, Q \in \mathcal{P}_c$ , and  $a, b > 0$  with  $a + b = 1$ .

3. The distribution function of the weak derivative belongs to  $L^2(\mathcal{I})$ .

The proof of this result is given later in Section A.1.

4.1.2. *Completing the proof.* We now have the tools to complete the proof of Theorems 2.1 and 2.2.

PROOF OF THEOREM 2.1 (CONTINUED). In Section 4.1.1, we showed that under the null  $T_p(\varphi) - c_p \Rightarrow \mathcal{N}(0, \sigma_\varphi^2)$ , while under the alternative  $T_p(\varphi) - c_p \Rightarrow \mathcal{N}(\mu_\varphi, \sigma_\varphi^2)$ , for constants  $c_p$ . It follows from equation (1.17) on page 564 of [Bai and Silverstein \(2004\)](#) that the variance  $\sigma_\varphi^2$  has the form stated in Theorem 2.1 [see (6)], while we showed that  $\mu_\varphi$  has the form in (12).

Recall that the distribution function of the weak derivative was defined by  $\Delta(x) = D((-\infty, x])$ , where  $D = \delta\mathcal{F}_\gamma(H, G_1) - \delta\mathcal{F}_\gamma(H, G_0)$ . Since  $H$  and  $G_i$  are compactly supported, from Theorem 4.1 it follows that the  $\delta\mathcal{F}_\gamma$ —and  $D$ —are also compactly supported. The compact interval  $\mathcal{I} = [a, b]$  is such that it includes this support. Since  $D$  has zero total mass,  $\Delta(x) = 0$  for  $x \leq a$  and for  $x \geq b$ . Using the integration by parts formula for the Lebesgue–Stieltjes integral, which is valid since  $\varphi$  is absolutely continuous, and  $D$  is a bounded Borel measure on  $\mathcal{I} = [a, b]$  with  $\Delta(a) = \Delta(b) = 0$  [see, e.g., [Bogachev \(2007\)](#), Example 5.8.112], we thus have

$$\mu_\varphi = h \int_{\mathcal{I}} \varphi(x) d[\delta\mathcal{F}_\gamma(H, G_1) - \delta\mathcal{F}_\gamma(H, G_0)](x) = -h \int_{\mathcal{I}} \varphi'(x) \Delta(x) dx.$$

This shows the asymptotic equivalence to the normal problem stated in Theorem 2.1 and completes its proof.  $\square$

We will now proceed to prove Theorem 2.2.

PROOF OF THEOREM 2.2. To optimize over  $\varphi$ , we will use properties of the Hilbert space  $L^2(\mathcal{I})$  and its inner product  $\langle g, j \rangle = \int_{\mathcal{I}} g(x) j(x) dx$ . Let us write  $g = \varphi' \in L^2(\mathcal{I})$ . We are optimizing over  $\varphi \in \mathcal{W}(\mathcal{I})$ , which by the definition of  $\mathcal{W}(\mathcal{I})$  is equivalent to optimizing over  $\varphi' = g \in L^2(\mathcal{I})$ . The mean and variance are  $\mu = -h\langle g, \Delta \rangle$ , and  $\sigma^2 = \langle g, Kg \rangle$ . The expression  $\mu = -h\langle g, \Delta \rangle$  is valid because  $\Delta \in L^2(\mathcal{I})$  by Theorem 4.1.

Therefore, the efficacy optimization is equivalent to the problem of maximizing  $\theta(g) = -h\langle g, \Delta \rangle / \langle g, Kg \rangle^{1/2}$  over  $g \in L^2(\mathcal{I})$ . The following lemma, proved in Section A.5, gives the desired answer.

LEMMA 4.2. *Consider maximizing  $\theta(g)$  over  $g \in L^2(\mathcal{I})$ . If  $\Delta \notin \text{Im}(K)$ , the supremum is  $+\infty$ . Moreover, if  $\Delta \notin \overline{\text{Im}}(K)$ , the supremum is achieved for  $g$  such that  $Kg = 0$  and  $\langle g, \Delta \rangle < 0$ . If  $\Delta \in \text{Im}(K)$ , the maximum is  $h\langle \Delta, K^+ \Delta \rangle^{1/2}$ , and is achieved for  $g$  such that  $Kg = -\eta \Delta$ , for some  $\eta > 0$ .*

The conclusion of Theorem 2.2 follows immediately from the above lemma and completes the proof.  $\square$

Finally, we can prove Theorem 2.3.

PROOF OF THEOREM 2.3. Consider first the choice of the critical values  $t_\varphi^-, t_\varphi^+$  for a fixed  $\varphi$ . From Theorem 2.1, under the null  $T_p(\varphi) - c_p(\varphi) \Rightarrow \mathcal{N}(0, \sigma_\varphi^2)$ , while under the alternative  $T_p(\varphi) - c_p(\varphi) \Rightarrow \mathcal{N}(\mu_\varphi, \sigma_\varphi^2)$ . If the effect size of  $\varphi$  is 0,  $\mu_\varphi = 0$ , then using  $\varphi$  leads to trivial power, so we will examine  $\mu_\varphi < 0$  in the remainder; the case  $\mu_\varphi > 0$  is analogous.

If  $\sigma_\varphi > 0$ , the asymptotically optimal choices are  $t_\varphi^- = m_\varphi + \sigma_\varphi z_\alpha$  and  $t_\varphi^+ = +\infty$ ; while the asymptotic power equals  $\Phi(z_\alpha + |\mu_\varphi|/\sigma_\varphi)$ . If  $\sigma_\varphi = 0$ , then we can take  $t_\varphi^\pm = m_\varphi \pm \varepsilon$  for any  $\varepsilon > 0$ , and still have asymptotic level  $\alpha$ . Moreover, the test statistic converges in probability to two different values—0 and  $\mu_\varphi$ —under the null and the alternative. Therefore, the power of such a test is asymptotically unity for small  $\varepsilon$ . We conclude that the maximal power over analytic functions is  $\beta = \sup_{\varphi \in \mathcal{H}(\mathcal{I})} \Phi(z_\alpha + |\mu_\varphi|/\sigma_\varphi) = \Phi(z_\alpha + \theta^*(\mathcal{H}(\mathcal{I})))$ . Here, we used the convention of defining  $\mu_\varphi/\sigma_\varphi$  as 0 or  $+\infty$  in corner cases.

We now show that the efficacy over the set of analytic functions  $\mathcal{H}(\mathcal{I})$  equals the efficacy over  $\mathcal{W}(\mathcal{I})$ , because the optimal LSS can be approximated arbitrarily well—in an  $L^2$  sense—by analytic functions.

LEMMA 4.3 [Optimal Linear Spectral Statistics over  $\mathcal{H}(\mathcal{I})$ ]. *The efficacy over the set of analytic functions  $\mathcal{H}(\mathcal{I})$  equals that over  $\mathcal{W}(\mathcal{I})$ :  $\theta^*(\mathcal{H}(\mathcal{I})) = \theta^*(\mathcal{W}(\mathcal{I}))$ . There is a sequence  $\varphi_n \in \mathcal{H}(\mathcal{I})$  such that  $\theta(\varphi_n) \uparrow \theta^*(\mathcal{W}(\mathcal{I}))$ .*

The proof is in Section A.6. From Lemma 4.3, we conclude that  $\beta = \Phi(z_\alpha + \theta^*(\mathcal{W}(\mathcal{I})))$ . Now Theorem 2.2 shows that  $\theta^*(\mathcal{W}(\mathcal{I}))$  has the desired form, completing the proof.  $\square$

**5. Discussion.** An interesting set of questions would be to extend the results of the spiked model, and in particular our investigation, to the case of “mesoscopic perturbations” where the number of spikes can grow with  $n$ . The probabilistic foundations of these models are only beginning to be worked out. For instance, Huang (2014) shows concentration inequalities for the individual spikes above the phase transition. Using this and extensions of other results, such as the CLT for the individual spikes, in future work it may be possible to answer many questions for



growing number of spikes, such as convergence rates and simultaneous confidence intervals.

Another question of interest is to develop a better understanding of the optimal LSS functions. For instance, are they unique up affine scaling? Can we find explicit expressions for them in certain cases? Is it possible to show that they are smooth—and thus optimal over function classes possessing more derivatives? Answering these questions will lead to a more complete picture about optimal testing in high-dimensional PCA.

**Acknowledgments.** We are grateful to David Donoho and Iain Johnstone for helpful discussion. We thank Marc Hallin and Alexei Onatski for comments on an earlier version of the manuscript. We thank the Associate Editor and the referees for remarks, and for suggestions for future research, that helped improve the paper.

## SUPPLEMENTARY MATERIAL

**Supplement to “Sharp detection in PCA under correlations: All eigenvalues matter”** (DOI: [10.1214/16-AOS1514SUPP](https://doi.org/10.1214/16-AOS1514SUPP); .pdf). In the supplementary material, we give the remaining details of the proofs, algorithms implementing our method and further simulations.

## REFERENCES

- ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Stat.* **34** 122–148. [MR0145620](#)
- ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley-Interscience, Hoboken, NJ. [MR1990662](#)
- BAI, Z. D. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. [MR2040792](#)
- BAI, Z. and SILVERSTEIN, J. W. (2009). *Spectral Analysis of Large Dimensional Random Matrices*. Springer.
- BAI, Z. and YAO, J. (2012). On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.* **106** 167–177. [MR2887686](#)
- BAIK, J., BEN AROUS, G. and PÉCHÉ, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. [MR2165575](#)
- BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. [MR2279680](#)
- BENAYCH-GEORGES, F. and NADAKUDITI, R. R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** 494–521. [MR2782201](#)
- BOGACHEV, V. I. (2007). *Measure Theory*, Vol. 1. Springer, Berlin.
- COUILLET, R. and DEBBAH, M. (2011). *Random Matrix Methods for Wireless Communications*. Cambridge Univ. Press, Cambridge. [MR2884783](#)
- DHARMAWANSA, P., JOHNSTONE, I. M. and ONATSKI, A. (2014). Local asymptotic normality of the spectrum of high-dimensional spiked F-ratios. ArXiv preprint [arXiv:1411.3875](https://arxiv.org/abs/1411.3875).
- DOBRIBAN, E. (2015). Efficient computation of limit spectra of sample covariance matrices. *Random Matrices Theory Appl.* **4** 1550019, 36. [MR3418848](#)
- DOBRIBAN, E. (2017). Supplement to “Sharp detection in PCA under correlations: All eigenvalues matter.” DOI:[10.1214/16-AOS1514SUPP](https://doi.org/10.1214/16-AOS1514SUPP).

- DOBRIAN, E. and WAGER, S. (2015). High-dimensional asymptotics of prediction: Ridge regression and classification. ArXiv preprint [arXiv:1507.03003](#).
- GROETSCH, C. W. (1977). *Generalized Inverses of Linear Operators: Representation and Approximation*. *Monographs and Textbooks in Pure and Applied Mathematics* **37**. Dekker, New York. [MR0458859](#)
- HACHEM, W., HARDY, A. and NAJIM, J. (2015). A survey on the eigenvalues local behavior of large complex correlated Wishart matrices. In *Modélisation Aléatoire et Statistique—Journées MAS 2014. ESAIM Proc. Surveys* **51** 150–174. EDP Sci., Les Ulis. [MR3440796](#)
- HUANG, J. (2014). Mesoscopic perturbations of large random matrices. ArXiv preprint [arXiv:1412.4193](#).
- JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. [MR1863961](#)
- JOHNSTONE, I. M. and ONATSKI, A. (2015). Testing in high-dimensional spiked models. ArXiv preprint [arXiv:1509.07269](#).
- KRESS, R. (2014). *Linear Integral Equations*, 3rd ed. Springer, New York. [MR3184286](#)
- LEHMANN, E. L. and ROMANO, J. P. (2005). *Testing Statistical Hypotheses*, 3rd ed. Springer, New York. [MR2135927](#)
- MARCHENKO, V. A. and PASTUR, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mat. Sb.* **114** 507–536.
- ONATSKI, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *J. Econometrics* **168** 244–258. [MR2923766](#)
- ONATSKI, A., MOREIRA, M. J. and HALLIN, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Ann. Statist.* **41** 1204–1231. [MR3113808](#)
- ONATSKI, A., MOREIRA, M. J. and HALLIN, M. (2014). Signal detection in high dimension: The multispiked case. *Ann. Statist.* **42** 225–254. [MR3189485](#)
- PAN, G. M. and ZHOU, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.* **18** 1232–1270. [MR2418244](#)
- PAUL, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. [MR2399865](#)
- PAUL, D. and AUE, A. (2014). Random matrix theory in statistics: A review. *J. Statist. Plann. Inference* **150** 1–29. [MR3206718](#)
- SILVERSTEIN, J. W. and CHOI, S.-I. (1995). Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal.* **54** 295–309. [MR1345541](#)
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. [MR1652247](#)
- YAO, J., ZHENG, S. and BAI, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. Cambridge Univ. Press, New York. [MR3468554](#)
- ZHENG, S., BAI, Z. and YAO, J. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Ann. Statist.* **43** 546–591. [MR3316190](#)

DEPARTMENT OF STATISTICS  
 STANFORD UNIVERSITY  
 STANFORD, CALIFORNIA 94305  
 USA  
 E-MAIL: [dobriban@stanford.edu](mailto:dobriban@stanford.edu)