RANK TESTS FOR PCA UNDER WEAK IDENTIFIABILITY

BY DAVY PAINDAVEINE[*], LAURA PERALVO MAROTO[†] AND THOMAS VERDEBOUT[‡]

ECARES and Department of Mathematics, Université libre de Bruxelles, [*]*Davy.Paindaveine@ulb.be* [†]*Laura.Peralvo@ulb.be;* [‡]*Thomas.Verdebout@ulb.be*

In a triangular array framework where n observations are randomly sampled from a p-dimensional elliptical distribution with shape matrix $\mathbf{V}_n$, we consider the problem of testing the null hypothesis $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against the alternative hypothesis $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}$ is the (fixed) leading unit eigenvector of $\mathbf{V}_n$ and $\boldsymbol{\theta}_0$ is a given unit p-vector. The dependence of the shape matrix on the sample size allows us to consider challenging asymptotic scenarios in which the parameter of interest $\boldsymbol{\theta}$ is unidentified in the limit, because the ratio between both leading eigenvalues of $\mathbf{V}_n$ converges to one. We carefully study the corresponding limiting experiments under such *weak identifiability*, and we show that these may be LAN or non-LAN. While earlier work in the framework was strictly limited to Gaussian distributions, where the study of local log-likelihood ratios could simply rely on explicit expressions, our asymptotic investigation allows for essentially arbitrary elliptical distributions. This requires original results on quadratic mean differentiable families for triangular arrays of observations, that are likely to be of interest in other models, too. Even in non-LAN experiments, our results enable us to investigate, through Le Cam's first and third lemmas, the asymptotic null and non-null properties of multivariate rank tests. These nonparametric tests are shown to exhibit an excellent behavior under weak identifiability: not only do they maintain the target nominal size irrespective of the amount of weak identifiability, but they also keep their outstanding uniform efficiency properties under such non-standard scenarios. In particular, Gaussian-score rank tests, under arbitrarily weak identifiability, still *uniformly* dominate their parametric pseudo-Gaussian competitor in terms of asymptotic relative efficiencies. Our theoretical results, that are the first ones to study rank tests in the triangular array framework allowing for weak identifiability, are supported by several Monte-Carlo exercises.

1. Introduction. Dimension reduction is nowadays a very classical topic in Statistics. Among the various dimension reduction techniques, Principal Component Analysis (PCA) remains by far the most commonly used. For a random p-vector $\mathbf{X}$ with arbitrary mean vector and covariance matrix $\boldsymbol{\Sigma}$, it is well-known that the first principal component is obtained by computing the projections $\boldsymbol{\theta}'\mathbf{X}$ of $\mathbf{X}$ onto the unit eigenvector $\boldsymbol{\theta}$ associated with the largest eigenvalue of $\boldsymbol{\Sigma}$. The remaining principal components are then obtained by projecting $\mathbf{X}$ onto the other eigenvectors, ordered decreasingly according to their companion eigenvalues. Obviously, in practice, the covariance matrix $\boldsymbol{\Sigma}$ is usually unknown, so that one of the key issues in PCA is to perform inference on the corresponding eigenvectors and eigenvalues.

The seminal paper Anderson (1963) provided asymptotic results for estimated eigenvectors and eigenvalues of $\boldsymbol{\Sigma}$ in the Gaussian case. Later, Tyler (1981, 1983) extended these results to the broader elliptical case. Inference on the eigenvalues of $\boldsymbol{\Sigma}$ is still a very important topic; see, among others, the recent papers Donoho, Gavish and Johnstone (2018), Dörnemann and Dette (2025), Bernard and Verdebout (2024a,b), and the references therein.

MSC2020 subject classifications: Primary 62H25,62F05; secondary 62G10,62G35.

Keywords and phrases: Elliptical densities, limiting experiments, multivariate signs and ranks, principal component analysis, spiked scatter matrices, triangular arrays of observations, weak identifiability.

Inference on eigenvectors of covariance/scatter matrices is also a very popular topic: Hallin, Paindaveine and Verdebout (2014) proposed efficient R-estimators for the eigenvectors of scatter matrices, while Croux and Haesbroeck (2000), Hubert, Rousseeuw and Vanden Branden (2005), and He et al. (2011) developed robust PCA methods. Recently, Johnstone and Lu (2009), Han and Liu (2014), and Fan et al. (2022) considered inference on eigenvectors of $\mathbf{\Sigma}$ in the high-dimensional framework.

In the present work, we consider, in a general elliptical framework, inference on the leading eigenvector $\boldsymbol{\theta}$, that is, the one associated with the largest eigenvalue of the corresponding scatter matrix $\mathbf{\Sigma}$ (all our results would actually apply to inference on any other eigenvector of $\mathbf{\Sigma}$). More precisely, we consider the problem of testing $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}_0$ is a given unit p-vector. While the emphasis in PCA is usually more on point estimation, the testing problem above is of high practical relevance, too. For instance, it is of paramount importance in "confirmatory PCA", that is, when one wants to test that $\boldsymbol{\theta}$ (or any other eigenvector) coincides with an eigenvector obtained from an earlier real data analysis ("historical data") or with an eigenvector resulting from a theory or a model. In line with this, tests for the null hypothesis $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ have been used, for instance, in Jackson (2005) to analyze the concentration of a chemical component in a solution and in Sylvester, Kramer and Jungers (2008) to study the geometric similarity among modern humans.

Denoting the sample covariance matrix as $\mathbf{S}^{(n)} := (1/n)\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}}^{(n)})(\mathbf{X}_i - \bar{\mathbf{X}}^{(n)})'$ (with $\bar{\mathbf{X}}^{(n)} := (1/n)\sum_{i=1}^{n}\mathbf{X}_i$, as usual) and its largest eigenvalue as $\hat{\lambda}_1$, the textbook test for this problem is the Anderson (1963) Gaussian likelihood ratio test, ϕ_{A} say, that rejects the null hypothesis at asymptotic level α when

$$Q_{\mathrm{A}} := n\big(\hat{\lambda}_1\boldsymbol{\theta}_0'(\mathbf{S}^{(n)})^{-1}\boldsymbol{\theta}_0 + \hat{\lambda}_1^{-1}\boldsymbol{\theta}_0'\mathbf{S}^{(n)}\boldsymbol{\theta}_0 - 2\big) > \chi^2_{p-1,1-\alpha},$$

where $\chi^2_{p-1,1-\alpha}$ stands for the upper α-quantile of the chi-square distribution with $p-1$ degrees of freedom. Various extensions of this test have been proposed in the literature. To mention only a few, Jolicoeur (1984) considered a small-sample version of this test, Flury (1988) proposed an extension to a larger number of eigenvectors, Tyler (1981, 1983) robustified the test to possible (elliptical) departures from multinormality, while Schwartzman, Mascarenhas and Taylor (2008) considered extensions to the case of Gaussian random matrices. Le Cam optimal tests for the same problem were obtained in Hallin, Paindaveine and Verdebout (2010a). More precisely, Hallin, Paindaveine and Verdebout (2010a) first defined a pseudo-Gaussian test ϕ_{G} that achieves asymptotically the target nominal size under any null elliptical distribution with finite fourth-order moments, while keeping the same asymptotic power as the Gaussian likelihood ratio test ϕ_{A} under multinormality. Motivated by invariance arguments, Hallin, Paindaveine and Verdebout (2010a) further developed multivariate rank tests, that in particular avoid any moment assumption. When based on a score function K, this rank test, ϕ_K say, rejects the null hypothesis at asymptotic level α when

$$Q_K := \frac{np(p+2)}{\mathcal{J}_p(K)}\sum_{j=2}^{p}\big(\tilde{\boldsymbol{\theta}}_j'\mathbf{S}_K^{(n)}\boldsymbol{\theta}_0\big)^2 > \chi^2_{p-1,1-\alpha},$$

where $\mathbf{S}_K^{(n)}$ is a signed-rank scatter matrix involving the score function K and where $\hat{\lambda}_j$ and $\tilde{\boldsymbol{\theta}}_j$, $j = 2,\ldots,p$ are suitable estimators of the $p-1$ smallest eigenvalues of the underlying shape matrix and of the corresponding unit eigenvectors, respectively; see Section 2 for details. When the corresponding population eigenvalues $\lambda_1,\ldots,\lambda_p$ are fixed and satisfy $\lambda_1 > \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_p$ (the minimal condition under which $\boldsymbol{\theta}$ is identifiable—up to an unimportant sign), the above tests compare as follows:

(a_1) in the Gaussian case, the test statistics of $\phi_{\rm A}$ and $\phi_{\rm G}$ are asymptotically equivalent in probability under the null hypothesis, hence also under sequences of contiguous alternatives. The advantage of $\phi_{\rm G}$ with respect to $\phi_{\rm A}$ is that it is robust to non-normality under the null hypothesis: it shows the target nominal size asymptotically under any null elliptical distribution with finite fourth-order moments;

(b_1) robustness of the rank test ϕ_K under the null hypothesis is actually even better: unlike the pseudo-Gaussian test, the rank test still achieves asymptotically the target nominal size even under infinite fourth-order moments;

(c_1) rank tests combine the above "validity-robustness" (robustness in terms of Type I risk) with "efficiency-robustness" (robustness in terms of Type II risk): in particular, denoting as $\mathrm{ARE}_g(\phi_a/\phi_b)$ the asymptotic relative efficiency of ϕ_a with respect to ϕ_b under the elliptical density g, the Chernoff–Savage result in Paindaveine (2006) entails that the Gaussian-score rank test $\phi_{K_{\phi_1}}$ (the one that achieves Le Cam optimality under Gaussian distributions) satisfies

$$\mathrm{ARE}_g(\phi_{K_{\phi_1}}/\phi_{\rm G}) \geq 1 \tag{1.1}$$

under any g with finite fourth-order moments, and that equality in (1.1) holds in the Gaussian case only (see (2.6) below for details on $\phi_{K_{\phi_1}}$). As soon as the underlying elliptical distribution is not Gaussian, thus, the Gaussian-score rank test strictly improves over its pseudo-Gaussian competitor in terms of asymptotic relative efficiencies.

In the classical asymptotic scenario where eigenvalues remain fixed as the sample size diverges to infinity, thus, the pseudo-Gaussian test $\phi_{\rm G}$ should be favoured over the Anderson test $\phi_{\rm A}$ in terms of validity-robustness, whereas rank tests, and in particular the Gaussian-score rank test, should be favoured over $\phi_{\rm G}$ in terms of both validity- and efficiency-robustness.

As recently shown in Paindaveine, Remy and Verdebout (2020a) (hereafter, PRV20), however, another type of robustness is of high practical relevance, too. PRV20 proved that the asymptotic equivalence between $\phi_{\rm G}$ and $\phi_{\rm A}$ in the Gaussian case (see (a_1) above) does not resist situations under which $\boldsymbol{\theta}$ is *weakly identified*: more precisely, under triangular arrays of Gaussian observations for which the ratio between both leading eigenvalues of the underlying scatter matrix satisfies $\lambda_{n1}/\lambda_{n2} = 1 + O(n^{-1/2})$, the Anderson test $\phi_{\rm A}$ becomes extremely liberal (we also refer to Dümbgen (1995) for the behavior of likelihood ratio tests under weak identifiability) whereas the pseudo-Gaussian test $\phi_{\rm G}$ still shows asymptotically the target null size (the terminology *weak identifiability* here reflects the fact that, in this non-standard asymptotic scenario, the parameter of interest $\boldsymbol{\theta}$ is not identified in the limit). In other words, when λ_{n1} and λ_{n2} get "too close to each other", $\phi_{\rm G}$ remains an asymptotically valid test while $\phi_{\rm A}$ does not. In this sense, $\phi_{\rm G}$ is robust to weak identifiability while $\phi_{\rm A}$ is not. In view of the dominance of rank tests over pseudo-Gaussian tests in standard asymptotic scenarios, an urgent question is then the following: "do points (b_1)–(c_1) above extend to situations involving weak identifiability?" Answering this question is the main objective of the present paper. More precisely, the questions we consider are the following ones:

(b_2) Under general elliptical distributions, does the validity-robustness of the rank test ϕ_K survive asymptotic scenarios involving weak identifiability?

(c_2) If so, does the efficiency-robustness of the rank test ϕ_K survive such non-standard asymptotic scenarios, still under general elliptical distributions? In particular, does the Chernoff–Savage result in (1.1) still hold under weak identifiability?

Answering these questions raises important technical challenges that go much further what was considered in earlier works on the topic, including PRV20. While, as mentioned above, the asymptotic behavior of the pseudo-Gaussian test ϕ_{G} was already studied in PRV20, all non-null results obtained there on this test were strictly limited to multinormal distributions. The reason is clear: it is only in that particular distributional framework that PRV20 investigated the asymptotic structure of local log-likelihood ratios. However, deriving the non-null asymptotic behaviors of the pseudo-Gaussian and rank tests under general elliptical distributions, which is obviously needed to explore whether or not the above Chernoff–Savage result extends to scenarios involving weakly identifiability, requires studying local log-likelihood ratios in the framework of non-Gaussian triangular arrays. *To this end, the methodology used in PRV20, that relies on explicit expressions of the corresponding multinormal likelihoods, cannot be used.* In the present work, we therefore adopt a much more general approach that allows us to study the asymptotic behavior of local log-likelihood ratios under double-asymptotic scenarios involving arbitrarily weak identifiability. This is done in particular by applying an original result on quadratic mean differentiable families in a general triangular array framework; see Proposition A.1 below. This allows us to characterize, under a general elliptical framework, the resulting limiting experiments; these are LAN (Locally Asymptotically Normal) or non-LAN, depending on the severity of weak identifiability. Even in non-LAN cases, Le Cam's first and third lemmas will allow us to derive the non-null asymptotic behaviors of the pseudo-Gaussian and rank tests above.

Other challenging points that will be tackled for the first time relate to rank tests themselves. Due to the mutual dependence of ranks, studying the asymptotic behavior of rank tests under weak identifiability requires suitable *asymptotic representation results* in triangular arrays of observations—this was not required when considering sign tests in Paindaveine, Remy and Verdebout (2020b), as multivariate signs are mutually independent random vectors. An even more delicate issue for rank tests relates to the estimation of nuisance parameters. Typically, controlling the resulting "aligned" ranks is done by establishing an *asymptotic linearity property*. Establishing such a property in the double-asymptotic scenarios we consider here, however, is another key difficulty we will need to address. We will actually do so by deriving a general asymptotic linearity result under triangular arrays of observations; see Proposition A.2 below. The general results in Propositions A.1 and A.2 are likely to be of interest in other contexts, too.

The paper is organized as follows. In Section 2, we define the pseudo-Gaussian test ϕ_{G} and the rank tests ϕ_K. In Section 3, we first investigate the asymptotic behavior of local log-likelihood ratios under single-spiked elliptical distributions and possible weak identifiability. Then, we study the corresponding null and non-null asymptotic behaviors of both the pseudo-Gaussian and rank tests. In Section 4, we perform Monte-Carlo exercises in order to explore the finite-sample relevance (and check the correctness) of our asymptotic results. In Section 5, we provide power-enhanced versions of our tests that, unlike their original version, are consistent under any fixed alternative. In Section 6, we extend our null and non-null asymptotic results from single-spiked spectra to much more general spectra (this is another major improvement over PRV20, where all results were strictly restricted to single-spiked spectra). Finally, we conclude the paper in Section 7, where we also briefly discuss perspectives for future research. In an appendix, we provide both aforementioned general results, that allow us to study the asymptotic behavior of local log-likelihood ratios in a triangular array framework (Proposition A.1) and to control aligned ranks under weak identifiability (Proposition A.2). All proofs are deferred to the supplement Paindaveine, Peralvo Maroto and Verdebout (2025).

For the sake of convenience, we introduce here some notation that will be used throughout the paper. We will denote as $\mathcal{S}^{p-1} := \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\|^2 = \mathbf{x}'\mathbf{x} = 1\}$ the unit hypersphere of $\mathbb{R}^p$. Since unit eigenvectors are at best defined up to a sign, we will consider the "positive" hemisphere $\mathcal{S}^{p-1}_+$ that collects the vectors in $\mathcal{S}^{p-1}$ whose first non-zero component is positive. Throughout, $\mathbf{I}_p$ will stand for the $p \times p$ identity matrix, and $\mathrm{diag}(\ell_1, \ldots, \ell_k)$ will stand for the $k \times k$ diagonal matrix with diagonal entries $\ell_1, \ldots, \ell_k$. For a $p \times p$ matrix $\mathbf{A}$, the quantities $\mathrm{vec}\,(\mathbf{A})$ and $\mathrm{vech}\,(\mathbf{A})$ will respectively denote the vector obtained by stacking the columns of $\mathbf{A}$ on top of each other and the vector obtained by doing the same operation but keeping only the diagonal and upper-diagonal entries of $\mathbf{A}$. We will write $\mathbf{J}_p := \mathrm{vec}\,(\mathbf{I}_p)\mathrm{vec}'(\mathbf{I}_p)$. Also, $\mathbf{K}_p$ will be the $p^2 \times p^2$ commutation matrix, that is such that $\mathbf{K}_p\mathrm{vec}\,(\mathbf{A}) = \mathrm{vec}\,(\mathbf{A}')$. For a symmetric positive semi-definite matrix $\mathbf{A}$, the matrix $\mathbf{A}^{1/2}$ will stand for its symmetric square root and the matrix $\mathbf{A}^{-1/2}$ will then be the inverse of this symmetric square root. The determinant of $\mathbf{A}$ will be denoted as $|\mathbf{A}|$. For a $p \times p$ scatter matrix $\boldsymbol{\Sigma}$ (that is, for a $p \times p$ symmetric positive definite matrix $\boldsymbol{\Sigma}$), the corresponding scale is $\sigma = |\boldsymbol{\Sigma}|^{1/(2p)}$ and the corresponding shape matrix is $\mathbf{V} = \boldsymbol{\Sigma}/|\boldsymbol{\Sigma}|^{1/p} = \boldsymbol{\Sigma}/\sigma^2$. As already mentioned, $\chi^2_{p-1,1-\alpha}$ is the quantile of order $1-\alpha$ of the chi-square distribution with $p-1$ degrees of freedom. We write $\mathbb{I}[A]$ for the indicator function of A. All deterministic and stochastic convergences will be as the sample size n diverges to infinity (this will include o and O statements, as well as o_{P} and O_{P} ones).

2. The pseudo-Gaussian and rank tests. In the rest of the paper, we focus on the problem of testing $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}$ is the eigenvector in $\mathcal{S}^{p-1}_+$ associated with the largest eigenvalue of a scatter matrix $\boldsymbol{\Sigma}_n$, or equivalently of the corresponding shape matrix $\mathbf{V}_n$, and where $\boldsymbol{\theta}_0 \in \mathcal{S}^{p-1}_+$ is fixed. We consider this problem within sequences of elliptical models indexed by a location parameter $\boldsymbol{\mu} \in \mathbb{R}^p$, a sequence of scatter matrices $(\boldsymbol{\Sigma}_n)$, and a radial density g_1; see (3.1) below for details. We now precisely define the pseudo-Gaussian and rank tests we will study in Section 3.

We start with rank tests. Based on a random sample $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$, the rank test ϕ_K from Hallin, Paindaveine and Verdebout (2010a), which results from the general methodology we will describe around (A.9), rejects the null hypothesis at asymptotic level α when

$$Q_K := \frac{np(p+2)}{\mathcal{J}_p(K)} \sum_{j=2}^{p} \big(\tilde{\boldsymbol{\theta}}_j' \mathbf{S}_K^{(n)} \boldsymbol{\theta}_0\big)^2 > \chi^2_{p-1,1-\alpha}, \tag{2.1}$$

where $\mathcal{J}_p(K)$ and $\mathbf{S}_K^{(n)}$ are respectively a normalizing constant and a signed-rank scatter matrix defined properly below and where $\tilde{\boldsymbol{\theta}}_2, \ldots, \tilde{\boldsymbol{\theta}}_p$ are defined recursively through

$$\tilde{\boldsymbol{\theta}}_j := s_j \frac{(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0' - \sum_{h=2}^{j-1} \tilde{\boldsymbol{\theta}}_h \tilde{\boldsymbol{\theta}}_h')\hat{\boldsymbol{\theta}}_{j,\mathrm{Tyler}}}{\|(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0' - \sum_{h=2}^{j-1} \tilde{\boldsymbol{\theta}}_h \tilde{\boldsymbol{\theta}}_h')\hat{\boldsymbol{\theta}}_{j,\mathrm{Tyler}}\|}; \tag{2.2}$$

here, $s_j \in \{-1, 1\}$ is such that $\tilde{\boldsymbol{\theta}}_j \in \mathcal{S}^{p-1}_+$, and summation over an empty collection of indices being defined as zero). These $\tilde{\boldsymbol{\theta}}_j$'s are thus obtained from a Gram–Schmidt orthogonalization of $\boldsymbol{\theta}_0, \hat{\boldsymbol{\theta}}_{2,\mathrm{Tyler}}, \ldots, \hat{\boldsymbol{\theta}}_{p,\mathrm{Tyler}} \in \mathcal{S}^{p-1}_+$, where

$$\begin{aligned}\hat{\mathbf{V}}_{\mathrm{Tyler}} &= (\hat{\boldsymbol{\theta}}_{1,\mathrm{Tyler}} \;\ldots\; \hat{\boldsymbol{\theta}}_{p,\mathrm{Tyler}})\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}(\hat{\boldsymbol{\theta}}_{1,\mathrm{Tyler}} \;\ldots\; \hat{\boldsymbol{\theta}}_{p,\mathrm{Tyler}})' \\ &= (\hat{\boldsymbol{\theta}}_{1,\mathrm{Tyler}} \;\ldots\; \hat{\boldsymbol{\theta}}_{p,\mathrm{Tyler}})\mathrm{diag}(\hat{\lambda}_{1,\mathrm{Tyler}}, \ldots, \hat{\lambda}_{p,\mathrm{Tyler}})(\hat{\boldsymbol{\theta}}_{1,\mathrm{Tyler}} \;\ldots\; \hat{\boldsymbol{\theta}}_{p,\mathrm{Tyler}})',\end{aligned}$$

with $\hat{\lambda}_{1,\mathrm{Tyler}} > \ldots > \hat{\lambda}_{p,\mathrm{Tyler}}$, is the spectral decomposition of the Tyler (1987) estimator of shape $\hat{\mathbf{V}}_{\mathrm{Tyler}}$ computed with respect to the Hettmansperger and Randles (2002) affine-equivariant median $\hat{\boldsymbol{\mu}}_n$ of $\mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}$; to be more specific, $\hat{\mathbf{V}}_{\text{Tyler}}$ is the symmetric positive definite $p\times p$ matrix with unit determinant satisfying

$$\frac{p}{n}\sum_{i=1}^{n}\frac{(\mathbf{X}_{ni}-\hat{\boldsymbol{\mu}}_n)(\mathbf{X}_{ni}-\hat{\boldsymbol{\mu}}_n)'}{(\mathbf{X}_{ni}-\hat{\boldsymbol{\mu}}_n)\hat{\mathbf{V}}_{\text{Tyler}}^{-1}(\mathbf{X}_{ni}-\hat{\boldsymbol{\mu}}_n)}=\hat{\mathbf{V}}_{\text{Tyler}}.$$

With this notation, a natural estimator of the shape matrix $\mathbf{V}_n$ under the null hypothesis is $\tilde{\mathbf{V}}_n := \tilde{\boldsymbol{\theta}}_0\hat{\boldsymbol{\Lambda}}_{\text{Tyler}}\tilde{\boldsymbol{\theta}}_0'$, with $\tilde{\boldsymbol{\theta}}_0 := (\boldsymbol{\theta}_0\,\tilde{\boldsymbol{\theta}}_2\,\ldots\,\tilde{\boldsymbol{\theta}}_p)$. The signed-rank covariance matrix $\mathbf{S}_K^{(n)}$ in (2.1) is then

$$\mathbf{S}_K^{(n)} := \frac{1}{n}\sum_{i=1}^{n}K\left(\frac{R_{ni}(\hat{\boldsymbol{\mu}}_n,\tilde{\mathbf{V}}_n)}{n+1}\right)\mathbf{U}_{ni}(\hat{\boldsymbol{\mu}}_n,\tilde{\mathbf{V}}_n)\mathbf{U}_{ni}'(\hat{\boldsymbol{\mu}}_n,\tilde{\mathbf{V}}_n), \tag{2.3}$$

where we wrote $R_{ni}(\boldsymbol{\mu},\mathbf{V})$ for the rank of $d_{ni}(\boldsymbol{\mu},\mathbf{V}) := \|\mathbf{V}^{-1/2}(\mathbf{X}_{ni}-\boldsymbol{\mu})\|$ among $d_{n1}(\boldsymbol{\mu},\mathbf{V}),\ldots,d_{nn}(\boldsymbol{\mu},\mathbf{V})$ and $\mathbf{U}_{ni}(\boldsymbol{\mu},\mathbf{V}) := \mathbf{V}^{-1/2}(\mathbf{X}_{ni}-\boldsymbol{\mu})/d_{ni}(\boldsymbol{\mu},\mathbf{V})$ for the "standardized spatial sign" of the observation $\mathbf{X}_{ni}$ with respect to $\boldsymbol{\mu}$; see, e.g., Oja (2010). These ranks and signs together form the maximal invariant associated with a group of monotone radial transformations from $\boldsymbol{\mu}$ (more precisely, they are asymptotically equivalent in probability to this maximal invariant); see Hallin and Paindaveine (2006) for details on this group. When deriving asymptotic results, we will assume that K satisfies the following mild assumption.

ASSUMPTION (A). The score function $K:(0,1)\to\mathbb{R}^+$

(A1) is continuous and square-integrable,
(A2) can be expressed as the difference of two monotone increasing functions, and
(A3) is normalized so that $\int_0^1 K(u)du = p$.

Optimality at a target radial density f_1 is achieved by choosing a score function $K=K_{f_1}$ defined by $K_{f_1}(u)=\tilde{F}_{1p}^{-1}(u)\varphi_{f_1}(\tilde{F}_{1p}^{-1}(u))$; see the beginning of Section 3 for the definition of the functions φ_{f_1} and $\tilde{F}_{1p}$. For score functions K, K_1 and K_2 satisfying Assumption (A), we let

$$\mathcal{J}_p(K_1,K_2) := \mathrm{E}[K_1(U)K_2(U)] \quad\text{and}\quad \mathcal{J}_p(K) := \mathcal{J}_p(K,K), \tag{2.4}$$

where U stands for a random variable that is uniformly distributed over $(0,1)$. We also let

$$\mathcal{J}_p(K,f_1) := \mathcal{J}_p(K,K_{f_1}) \quad\text{and}\quad \mathcal{J}_p(f_1) = \mathcal{J}_p(K_{f_1},K_{f_1}) \tag{2.5}$$

for score functions achieving optimality at f_1 (it is only when using score functions associated to a radial density f_1 that the dependence on p needs to be stressed in the notation, but we will still use the notation in (2.4) throughout). Classical score functions satisfying Assumption (A) include the power score functions $K(u)=p(a+1)u^a$ $(a\geq 0)$, with $\mathcal{J}_p(K)=p^2(a+1)^2/(2a+1)$; $a=0$, $a=1$ and $a=2$ provide the *Laplace* (or *sign*), *Wilcoxon*, and *Spearman* scores, respectively. An important example of score functions K_{f_1} is that of normal scores (also called *van der Waerden* scores) obtained when $f_1=\phi_1$ is the radial density associated with p-variate multinormal distributions, for which

$$K_{\phi_1}(u)=\Psi_p^{-1}(u) \quad\text{and}\quad \mathcal{J}_p(\phi_1)=p(p+2), \tag{2.6}$$

where Ψ_p denotes the chi-square distribution function with p degrees of freedom. If f_1 is rather the radial density associated with p-variate elliptical t distributions with ν degrees of freedom, then

$$K_{f_1}(u)=\frac{p(p+\nu)G_{p,\nu}^{-1}(u)}{\nu+pG_{p,\nu}^{-1}(u)} \quad\text{and}\quad \mathcal{J}_p(f_1)=\frac{p(p+2)(p+\nu)}{p+\nu+2},$$

where $G_{p,\nu}$ stands for the Fisher–Snedecor distribution function with p and ν degrees of freedom. In standard asymptotic scenarios where eigenvalues are fixed (i.e., scenarios that do not involve weak identifiability), this test is known to be locally asymptotically optimal (more precisely, locally asymptotically most stringent) under radial density f_1; see Hallin, Paindaveine and Verdebout (2010a).

The second test we will consider is the pseudo-Gaussian test ϕ_{G} that results from the general methodology we will describe around (A.8). This test rejects the null hypothesis at asymptotic level α when

$$Q_{\mathrm{G}} := \frac{n}{(1+\hat{\kappa}_n)\hat{\lambda}_{n1}} \sum_{j=2}^{p} \hat{\lambda}_{nj}^{-1} (\tilde{\boldsymbol{\theta}}_j' \mathbf{S}^{(n)} \boldsymbol{\theta}_0)^2 > \chi^2_{p-1,1-\alpha}, \tag{2.7}$$

where $\hat{\lambda}_{n1}, \ldots, \hat{\lambda}_{np}$ are the eigenvalues of $\mathbf{S}^{(n)} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_{ni} - \bar{\mathbf{X}}^{(n)})(\mathbf{X}_{ni} - \bar{\mathbf{X}}^{(n)})'$ and where $\tilde{\boldsymbol{\theta}}_2, \ldots, \tilde{\boldsymbol{\theta}}_p$ are obtained via the Gram–Schmidt procedure in (2.2) applied here to $\boldsymbol{\theta}_0, \hat{\boldsymbol{\theta}}_2, \ldots, \hat{\boldsymbol{\theta}}_p$, where $\hat{\boldsymbol{\theta}}_j$ is the eigenvector of $\mathbf{S}^{(n)}$ in $\mathcal{S}_+^{p-1}$ associated with the eigenvalue $\hat{\lambda}_{nj}$. The quantity $\hat{\kappa}_n$ in (2.7) is defined as

$$\hat{\kappa}_n := \frac{p\{\frac{1}{n}\sum_{i=1}^{n} d_{ni}^4(\bar{\mathbf{X}}^{(n)}, \mathbf{S}^{(n)})\}}{(p+2)\{\frac{1}{n}\sum_{i=1}^{n} d_{ni}^2(\bar{\mathbf{X}}^{(n)}, \mathbf{S}^{(n)})\}^2} - 1;$$

it is a consistent estimator of the kurtosis coefficient $\kappa_p(g_1)$ of the elliptical distribution with location centre $\boldsymbol{\mu}$, scatter matrix $\boldsymbol{\Sigma}$ and radial density g_1, defined as

$$\kappa_p(g_1) := \frac{pE_p(g_1)}{(p+2)D_p^2(g_1)} - 1,$$

with $D_p(g_1) := \mathrm{E}[d_{n1}^2(\boldsymbol{\mu}, \boldsymbol{\Sigma})]$ and $E_p(g_1) := \mathrm{E}[d_{n1}^4(\boldsymbol{\mu}, \boldsymbol{\Sigma})]$ (the value of $\kappa_p(g_1)$ does not depend on $\boldsymbol{\mu}$, nor on $\boldsymbol{\Sigma}$, which justifies the notation). Obviously, existence of $\kappa_p(g_1)$ requires finite fourth-order moments.

The asymptotic behavior of ϕ_{G} under weak identifiability was already studied in PRV20, but all non-null results there focused on the simple Gaussian case only. It is of course of interest, however, to know what is the non-null behavior of this test away from the Gaussian case, particularly so under heavy tails. An investigation of the asymptotic properties of the rank test ϕ_K is even more urgent; indeed, no result on the asymptotic behavior of this test is available, neither under the null hypothesis nor under local alternatives, even in the Gaussian case. In the next section, we therefore conduct a systematic study, under weak identifiability and in the general elliptic case, of the asymptotic behaviors of ϕ_K and ϕ_{G}, both under the null hypothesis and under suitable local alternatives.

3. Asymptotics in the single-spiked case. In the present section, we state our main asymptotic results under the single-spiked spectra that were considered in PRV20 (the extension to more general spectra will be conducted in Section 6 below). We start by properly providing the assumptions under which the various results will be obtained. Consider triangular arrays of elliptically symmetric observations $\mathbf{X}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$, where $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ form a random sample from the distribution that admits the density (with respect to the Lebesgue measure on $\mathbb{R}^p$)

$$f_{\boldsymbol{\mu}, \boldsymbol{\Sigma}_n, g_1}(\mathbf{x}) := \frac{c_{p,g_1}}{|\boldsymbol{\Sigma}_n|^{1/2}} g_1\left(\sqrt{(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}_n^{-1}(\mathbf{x}-\boldsymbol{\mu})}\right), \quad \mathbf{x} \in \mathbb{R}^p, \tag{3.1}$$

where $\boldsymbol{\mu} \in \mathbb{R}^p$ is a *location parameter*, $\boldsymbol{\Sigma}_n := \xi(\mathbf{I}_p + r_n v \boldsymbol{\theta}\boldsymbol{\theta}')$, with $\xi > 0$ and $\boldsymbol{\theta} \in \mathcal{S}_+^{p-1}$, is a sequence of *scatter matrix parameters*, g_1 is a *radial density*, and $c_{p,g_1} (> 0)$ is a normalizing constant. Letting $\mu_{\ell,h} := \int_0^\infty r^\ell h(r)dr$, we assume throughout that g_1 belongs to the class

$$\mathcal{F}_1 := \left\{ g_1 : \mathbb{R}_0^+ \to \mathbb{R}^+ \text{ such that } \mu_{p-1,g_1} < \infty \text{ and } \frac{1}{\mu_{p-1,g_1}} \int_0^1 r^{p-1} g_1(r)dr = \frac{1}{2} \right\}$$

of *standardized radial densities* (we will explain the nature of the standardization below). Above, (r_n) is a positive real sequence and v is a positive real number (we see r_n as a rate and v as an intensity parameter that governs the signal strength for any fixed rate r_n; see the next paragraph). The parameter ξ is an overall scale factor. The resulting hypothesis will be denoted as $\mathrm{P}^{(n)}_{\boldsymbol{\mu},\xi,\boldsymbol{\theta},r_n,v,g_1}$.

We factorize the scatter parameter $\boldsymbol{\Sigma}_n$ into $\sigma_n^2 \mathbf{V}_n$, where

$$\sigma_n := |\boldsymbol{\Sigma}_n|^{1/(2p)} = \sqrt{\xi}(1 + r_n v)^{1/(2p)}$$

is a *scale parameter* (with values in $\mathbb{R}_0^+$) and

$$\mathbf{V}_n := \boldsymbol{\Sigma}_n / \sigma_n^2 := (1 + r_n v)^{-1/p} (\mathbf{I}_p + r_n v \boldsymbol{\theta}\boldsymbol{\theta}')$$

is a *shape parameter* (with values in the collection of symmetric positive definite $p \times p$ matrices with unit determinant). Such a factorization is natural in the PCA context we consider since $\boldsymbol{\Sigma}_n$ and $\mathbf{V}_n$ share the same eigenvectors. The shape matrix $\mathbf{V}_n$ has eigenvalues

$$\text{(a)}\ \ \lambda_{n1,\mathbf{V}_n} = (1 + r_n v)^{(p-1)/p} \quad \text{and} \quad \text{(b)}\ \ \lambda_{n2,\mathbf{V}_n} = \cdots = \lambda_{np,\mathbf{V}_n} = (1 + r_n v)^{-1/p},$$

and the corresponding eigenspaces are (a) the vectorial line spanned by $\boldsymbol{\theta}$ and (b) its orthogonal complement in $\mathbb{R}^p$. Consequently, the considered triangular array of elliptical random vectors has a single-spike structure, in the sense of Donoho, Gavish and Johnstone (2018) or Banerjee and Ma (2022). The direction of the spike is $\boldsymbol{\theta}$ and its strength is driven by $r_n v$. When $r_n \equiv 1$ for any n, the ratio between the two largest eigenvalues of $\mathbf{V}_n$ (or equivalently, of $\boldsymbol{\Sigma}_n$) is $1 + v$ for any n, so that the leading eigenvector $\boldsymbol{\theta}$ remains identified in the limit as n diverges to infinity. In contrast, if (r_n) is $o(1)$, then the ratio between the two largest eigenvalues of $\mathbf{V}_n$ converges to one, and $\boldsymbol{\theta}$ is no longer identified in the limit; we then say that $\boldsymbol{\theta}$ is *weakly identified*.

Under $\mathrm{P}^{(n)}_{\boldsymbol{\mu},\xi,\boldsymbol{\theta},r_n,v,g_1}$, the Mahalanobis distances $d_{ni}(\boldsymbol{\mu}, \mathbf{V}_n)$, $i = 1, \ldots, n$, are i.i.d, with density and distribution functions given by

$$r \mapsto \frac{1}{\sigma_n} \tilde{g}_{1p}\left(\frac{r}{\sigma_n}\right) := \frac{1}{\sigma_n} (\mu_{p-1,g_1})^{-1} \left(\frac{r}{\sigma_n}\right)^{p-1} g_1\left(\frac{r}{\sigma_n}\right) \mathbb{I}[r > 0] \tag{3.2}$$

and

$$r \mapsto \tilde{G}_{1p}\left(\frac{r}{\sigma_n}\right) := \int_0^{r/\sigma_n} \tilde{g}_{1p}(s)ds, \tag{3.3}$$

respectively. The standardization of g_1 above is such that the density in (3.2) has median σ_n, a constraint that properly identifies ξ and g_1 without requiring any moment assumption; see Hallin and Paindaveine (2006) for a discussion. Under the same hypothesis, the multivariate signs $\mathbf{U}_{ni}(\boldsymbol{\mu}, \mathbf{V}_n)$, $i = 1, \ldots, n$, are i.i.d. uniform over the unit hypersphere $\mathcal{S}^{p-1}$ in $\mathbb{R}^p$, and they are independent of the Mahalanobis distances $d_{ni}(\boldsymbol{\mu}, \mathbf{V}_n)$, $i = 1, \ldots, n$.

Achieving optimality at a fixed value f_1 of g_1 will require some mild regularity conditions on f_1: we need to impose that f_1 belongs to the collection $\mathcal{F}_1^a$ of absolutely continuous densities in $\mathcal{F}_1$ such that

$$\mathcal{J}_p(f_1) = \int_0^1 r^2 \varphi_{f_1}^2(r) \tilde{f}_{1p}(r)dr < \infty, \tag{3.4}$$

where $\varphi_{f_1} := -\dot{f}_1/f_1$ involves the a.e. derivative $\dot{f}_1$ of f_1 (note that the integral expression in (3.4) agrees with the definition of $\mathcal{J}_p(f_1)$ in (2.5)). Finally, define the collection

$$\mathcal{F}_1^4 := \left\{ g_1 \in \mathcal{F}_1 : \int_0^\infty r^4 \tilde{g}_{1p}(r)dr < \infty \right\} = \left\{ g_1 \in \mathcal{F}_1 : \mu_{p+3,g_1} < \infty \right\}$$

of radial densities for which the densities defined in (3.2), hence also the corresponding elliptical densities in (3.1), have finite fourth-order moments. Since both the pseudo-Gaussian and rank tests are invariant under translations and homothetic transformations, that is, under transformations of the form $(\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}) \mapsto (\mathbf{X}_{n1} + \mathbf{t}, \ldots, \mathbf{X}_{nn} + \mathbf{t})$ (with $\mathbf{t} \in \mathbb{R}^p$) and $(\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}) \mapsto (\boldsymbol{\mu} + a(\mathbf{X}_{n1} - \boldsymbol{\mu}), \ldots, \boldsymbol{\mu} + a(\mathbf{X}_{nn} - \boldsymbol{\mu}))$ (with $a > 0$), we may assume without loss of generality that $\boldsymbol{\mu} = \mathbf{0}$ and $\xi = 1$. Consequently, we focus in the sequel on $\mathrm{P}^{(n)}_{\boldsymbol{\theta},r_n,v,f_1} := \mathrm{P}^{(n)}_{\mathbf{0},1,\boldsymbol{\theta},r_n,v,f_1}$. We will then simply write $d_{ni}(\mathbf{V}_n)$, $\mathbf{U}_{ni}(\mathbf{V}_n)$ and $R_{ni}(\mathbf{V}_n)$ rather than $d_{ni}(\mathbf{0}, \mathbf{V}_n)$, $\mathbf{U}_{ni}(\mathbf{0}, \mathbf{V}_n)$ and $R_{ni}(\mathbf{0}, \mathbf{V}_n)$.

Our first aim is to study the asymptotic behavior of local log-likelihood ratios of the form

$$\Lambda_n := \log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+\nu_n\boldsymbol{\tau}_n,r_n,v,f_1}}{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,r_n,v,f_1}},$$

where (ν_n) is a suitable positive real sequence and where $(\boldsymbol{\tau}_n)$ is a bounded sequence in $\mathbb{R}^p$ such that

$$\boldsymbol{\tau}_n'\boldsymbol{\theta}_0 = -\frac{\nu_n}{2}\|\boldsymbol{\tau}_n\|^2 \quad \text{for any } n; \tag{3.5}$$

the constraint (3.5) ensures that $\boldsymbol{\theta}_0 + \nu_n\boldsymbol{\tau}_n$ is a unit p-vector, hence is an admissible perturbation of $\boldsymbol{\theta}_0$. The asymptotic behavior of the log-likelihood ratio Λ_n is described in the following result (see Section S.1 of the supplement for a proof that crucially relies on our general result in Proposition A.1).

THEOREM 3.1. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $v > 0$, a positive sequence (r_n), and $f_1 \in \mathcal{F}_1^a$. Then, writing $d_{ni} := d_{ni}(\mathbf{V}_n)$ and $\mathbf{U}_{ni} := \mathbf{U}_{ni}(\mathbf{V}_n)$, we have the following (where, in each case, $(\boldsymbol{\tau}_n)$ is a bounded sequence in $\mathbb{R}^p$ such that (3.5) holds):*

(i) if $r_n \equiv 1$, then, with $\nu_n = 1/\sqrt{n}$,

$$\boldsymbol{\Delta}^{(n)}_{f_1} := \frac{v}{\sqrt{1+v}}\sqrt{n}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\left(\frac{1}{n}\sum_{i=1}^n \frac{d_{ni}}{\sigma_n}\varphi_{f_1}\left(\frac{d_{ni}}{\sigma_n}\right)\mathbf{U}_{ni}\mathbf{U}_{ni}' - \mathbf{I}_p\right)\boldsymbol{\theta}_0$$

and

$$\boldsymbol{\Gamma}_{f_1} = \frac{\mathcal{J}_p(f_1)v^2}{p(p+2)(1+v)}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0'),$$

we have that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,r_n,v,f_1}$, $\Lambda_n = \boldsymbol{\tau}_n'\boldsymbol{\Delta}^{(n)}_{f_1} - \frac{1}{2}\boldsymbol{\tau}_n'\boldsymbol{\Gamma}_{f_1}\boldsymbol{\tau}_n + o_{\mathrm{P}}(1)$ and that $\boldsymbol{\Delta}^{(n)}_{f_1}$ is asymptotically normal with mean zero and covariance matrix $\boldsymbol{\Gamma}_{f_1}$;

(ii) if r_n is $o(1)$ with $\sqrt{n}r_n \to \infty$, then, with $\nu_n = 1/(\sqrt{n}r_n)$,

$$\boldsymbol{\Delta}^{(n)}_{f_1} := v\sqrt{n}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\left(\frac{1}{n}\sum_{i=1}^n \frac{d_{ni}}{\sigma_n}\varphi_{f_1}\left(\frac{d_{ni}}{\sigma_n}\right)\mathbf{U}_{ni}\mathbf{U}_{ni}' - \mathbf{I}_p\right)\boldsymbol{\theta}_0$$

and

$$\boldsymbol{\Gamma}_{f_1} = \frac{\mathcal{J}_p(f_1)v^2}{p(p+2)}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0'),$$

we still have that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,r_n,v,f_1}$, $\Lambda_n = \boldsymbol{\tau}_n'\boldsymbol{\Delta}^{(n)}_{f_1} - \frac{1}{2}\boldsymbol{\tau}_n'\boldsymbol{\Gamma}_{f_1}\boldsymbol{\tau}_n + o_{\mathrm{P}}(1)$ *and that* $\boldsymbol{\Delta}^{(n)}_{f_1}$ *is asymptotically normal with mean zero and covariance matrix* $\boldsymbol{\Gamma}_{f_1}$*;*

(iii) if $r_n = 1/\sqrt{n}$*, then, with* $\nu_n \equiv 1$*,*

$$\Lambda_n = \boldsymbol{\tau}_n'\Bigg[v\sqrt{n}\bigg(\frac{1}{n}\sum_{i=1}^n \frac{d_{ni}}{\sigma_n}\varphi_{f_1}\Big(\frac{d_{ni}}{\sigma_n}\Big)\mathbf{U}_{ni}\mathbf{U}_{ni}' - \mathbf{I}_p\bigg)\Big(\boldsymbol{\theta}_0 + \frac{1}{2}\boldsymbol{\tau}_n\Big)\Bigg] - \frac{\mathcal{J}_p(f_1)v^2}{2p(p+2)}\bigg(\|\boldsymbol{\tau}_n\|^2 - \frac{\|\boldsymbol{\tau}_n\|^4}{4}\bigg) + o_{\mathrm{P}}(1) \tag{3.6}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,r_n,v,f_1}$*, where, if* $\boldsymbol{\tau}_n = \boldsymbol{\tau} + o(1)$*, then*

$$\boldsymbol{\tau}_n'\sqrt{n}\bigg(\frac{1}{n}\sum_{i=1}^n \frac{d_{ni}}{\sigma_n}\varphi_{f_1}\Big(\frac{d_{ni}}{\sigma_n}\Big)\mathbf{U}_{ni}\mathbf{U}_{ni}' - \mathbf{I}_p\bigg)\Big(\boldsymbol{\theta}_0 + \frac{1}{2}\boldsymbol{\tau}_n\Big)$$

is asymptotically normal with mean zero and covariance matrix

$$\frac{\mathcal{J}_p(f_1)}{p(p+2)}\bigg(\|\boldsymbol{\tau}\|^2 - \frac{\|\boldsymbol{\tau}\|^4}{4}\bigg);$$

(iv) if $r_n = o(1/\sqrt{n})$*, then, even with* $\nu_n \equiv 1$*, we have that* $\Lambda_n = o_{\mathrm{P}}(1)$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,r_n,v,f_1}$*.*

It follows from Theorem 3.1 that in regime (i) ($r_n \equiv 1$) and in regime (ii) (r_n is $o(1)$ with $\sqrt{n}r_n \to \infty$), the considered sequence of models is LAN, with central sequence

$$\boldsymbol{\Delta}^{(n)}_{f_1} = \frac{\sqrt{n}v}{\sqrt{1+\delta v}}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\bigg(\frac{1}{n}\sum_{i=1}^n \frac{d_{ni}}{\sigma_n}\varphi_{f_1}\Big(\frac{d_{ni}}{\sigma_n}\Big)\mathbf{U}_{ni}\mathbf{U}_{ni}' - \mathbf{I}_p\bigg)\boldsymbol{\theta}_0 \tag{3.7}$$

and Fisher information matrix

$$\boldsymbol{\Gamma}_{f_1} = \frac{\mathcal{J}_p(f_1)v^2}{p(p+2)(1+\delta v)}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0'), \tag{3.8}$$

where $\delta = 1$ in regime (i) and $\delta = 0$ in regime (ii). The situation in regime (iii) ($r_n = 1/\sqrt{n}$) is much less standard: indeed, the sequence of experiments there is neither LAN nor LAMN (Locally Asymptotically Mixed Normal), and the stochastic expansion of the local log-likelihood ratio in (3.6) is not a LAQ (Locally Asymptotically Quadratic) one. Nevertheless, Theorem 3.1(iii) and Le Cam's first lemma (see, e.g., Lemma 6.4 in van der Vaart (1998)) entail that $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,1/\sqrt{n},v,f_1}$ and $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+\boldsymbol{\tau}_n,1/\sqrt{n},v,f_1}$ are mutually contiguous. It is remarkable that fixed alternatives, that is alternatives associated with $\nu_n \equiv 1$, are contiguous to the null hypothesis (the severe weak identifiability in regime (iii) explains that this is possible). Finally, in regime (iv), local log-likelihood ratios converge in probability to zero even for the most severe alternatives that can be considered (that is, the fixed alternatives associated with $\nu_n \equiv 1$), so that no test will be able to detect even fixed alternatives in this case. Intuitively, the weak identifiability is so severe in regime (iv) that the considered inference problem is too difficult to allow for a non-trivial solution.

Now, we turn our attention to the pseudo-Gaussian and rank tests described in Section 2 and aim at deriving their asymptotic behavior, both under the null hypothesis and under suitable local alternatives. Theorem 3.1 will obviously be a key result to derive the asymptotic powers of these tests under the contiguous alternatives identified in regimes (i)–(iv). The following result, which was proved in PRV20 provides the asymptotic null behavior of the pseudo-Gaussian test.

PROPOSITION 3.1. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $\upsilon > 0$, a bounded positive sequence (r_n), and $g_1 \in \mathcal{F}_1^4$. Then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0, r_n, \upsilon, g_1}$, the test statistic Q_{G} in (2.7) is aymptotically chi-square with $p-1$ degrees of freedom.*

It follows that ϕ_{G} is robust to weak identifiability under any elliptical distribution that has finite fourth-order moments. The next result describes the asymptotic behavior of this test under the contiguous alternatives associated with the various regimes of Theorem 3.1 (see Section S.1 of the supplement for a proof).

THEOREM 3.2. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $\upsilon > 0$ and $g_1 \in \mathcal{F}_1^4 \cap \mathcal{F}_1^a$. Then, we have the following (in each case, $(\boldsymbol{\tau}_n)$ is a sequence in $\mathbb{R}^p$ that satisfies (3.5) and that converges to $\boldsymbol{\tau}$):*

(i) *if $r_n \equiv 1$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0 + \boldsymbol{\tau}_n/\sqrt{n}, r_n, \upsilon, g_1}$, Q_{G} is asymptotically chi-square with $p-1$ degrees of freedom and non-centrality parameter*

$$\frac{\upsilon^2}{(1+\kappa_p(g_1))(1+\upsilon)} \|\boldsymbol{\tau}\|^2;$$

(ii) *if $r_n = o(1)$ with $\sqrt{n} r_n \to \infty$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0 + \boldsymbol{\tau}_n/(\sqrt{n} r_n), r_n, \upsilon, g_1}$, Q_{G} is asymptotically chi-square with $p-1$ degrees of freedom and non-centrality parameter*

$$\frac{\upsilon^2}{1+\kappa_p(g_1)} \|\boldsymbol{\tau}\|^2;$$

(iii) *if $r_n = 1/\sqrt{n}$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0 + \boldsymbol{\tau}_n, r_n, \upsilon, g_1}$, Q_{G} is asymptotically chi-square with $p-1$ degrees of freedom and non-centrality parameter*

$$\frac{\upsilon^2}{16(1+\kappa_p(g_1))} \|\boldsymbol{\tau}\|^2 (4 - \|\boldsymbol{\tau}\|^2)(2 - \|\boldsymbol{\tau}\|^2)^2. \tag{3.9}$$

(iv) *if $r_n = o(1/\sqrt{n})$, then, under under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0 + \boldsymbol{\tau}_n, r_n, \upsilon, g_1}$, Q_{G} is asymptotically chi-square with $p-1$ degrees of freedom.*

While the asymptotic behavior of the pseudo-Gaussian test statistic Q_{G} is rather standard in regimes (i) and (ii), it is highly non-standard in regime (iii), with a non-centrality parameter that is not monotonic in $\|\boldsymbol{\tau}\|$. In regime (iv), Q_{G} remains (central) chi-square with $p-1$ degrees of freedom under fixed alternatives, which is in line with the fact that, from Theorem 3.1(iv), no test can detect even the most severe alternatives in this challenging regime.

We now turn to the asymptotic behavior of the rank test ϕ_K, which is much more delicate since it involves the asymptotic representation of signed-rank statistics in the context of weak identifiability. First rewrite the test statistic Q_K in (2.1) as

$$\begin{aligned} Q_K &= \frac{np(p+2)}{\mathcal{J}_p(K)} \sum_{j=2}^{p} \big(\tilde{\boldsymbol{\theta}}_j' \mathbf{S}_K^{(n)} \boldsymbol{\theta}_0\big)^2 \\ &= \frac{np(p+2)}{\mathcal{J}_p(K)} \boldsymbol{\theta}_0' \mathbf{S}_K^{(n)} (\mathbf{I}_p - \boldsymbol{\theta}_0 \boldsymbol{\theta}_0') \mathbf{S}_K^{(n)} \boldsymbol{\theta}_0 \\ &= \frac{p(p+2)}{\mathcal{J}_p(K)} \|\mathbf{T}_K^{(n)}\|^2, \end{aligned}$$

where we let $\mathbf{T}_K^{(n)} := \sqrt{n}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\mathbf{S}_K^{(n)}\boldsymbol{\theta}_0$. An important step to study the asymptotic behavior of the rank test ϕ_K is to obtain an *asymptotic representation result* for $\mathbf{T}_K^{(n)}$ that holds for any sequence (r_n). This actually requires modifying the estimator $(\hat{\boldsymbol{\mu}}_n, \tilde{\mathbf{V}}_n = \tilde{\boldsymbol{\theta}}_0\hat{\boldsymbol{\Lambda}}_{\text{Tyler}}\tilde{\boldsymbol{\theta}}_0')$ of $(\boldsymbol{\mu}, \mathbf{V}_n)$ introduced in Section 2 so that it is (or more precisely, so that its vectorized form is) *locally and asymptotically discrete* in the following sense; see, e.g., Definition 4.3 in Kreiss (1987) or Assumption (B3) in Hallin, Paindaveine and Verdebout (2010a).

DEFINITION 3.1. *A sequence of statistics $\mathbf{T}_n$ with values in $\mathbb{R}^k$ is* locally and asymptotically discrete *(with respect to the standard root-n rate) if and only if, for all $c > 0$, there exists $M = M(c) > 0$ such that the number of possible values of $\mathbf{T}_n$ in balls of the form $\{\mathbf{t} \in \mathbb{R}^k : \sqrt{n}\|\mathbf{t} - \mathbf{t}_0\| \leq c\}$ is bounded by M, uniformly as $n \to \infty$.*

In other words, the estimator $(\hat{\boldsymbol{\mu}}_n, \tilde{\mathbf{V}}_n)$ should be discretized so that it only takes a bounded number of distinct values in balls with $O(n^{-1/2})$ radius centered at $(\boldsymbol{\mu}, \mathbf{V}_n)$. This discretization has no practical consequences for fixed n since the discretization radius can be taken arbitrarily large; see page 2467 in Ilmonen and Paindaveine (2011) for more details. Such a discretization is thus a purely technical requirement that is needed to establish asymptotic results, and we will tacitly assume in the sequel that $(\hat{\boldsymbol{\mu}}_n, \tilde{\mathbf{V}}_n)$ has indeed been discretized. We then have the following asymptotic representation result (see Section S.1 of the supplement for a proof).

PROPOSITION 3.2. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $v > 0$, a bounded positive sequence (r_n), and $g_1 \in \mathcal{F}_1^a$, and let Assumption (A) hold. Let $\mathbf{T}_{K,g_1}^{(n)} := \sqrt{n}(\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\mathbf{S}_{K,g_1}^{(n)}\boldsymbol{\theta}_0$, with*

$$\mathbf{S}_{K,g_1}^{(n)} := \frac{1}{n}\sum_{i=1}^n K\left(\tilde{G}_{1p}\left(\frac{d_{ni}(\mathbf{V}_n)}{\sigma_n}\right)\right)\mathbf{U}_{ni}(\mathbf{V}_n)\mathbf{U}_{ni}'(\mathbf{V}_n), \tag{3.10}$$

where $\tilde{G}_{1p}$ was defined in (3.3). Then, $\mathbf{T}_K^{(n)} = \mathbf{T}_{K,g_1}^{(n)} + o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\theta}_0,r_n,v,g_1}^{(n)}$.

This result directly entails that, for any bounded sequence (r_n),

$$Q_K = \frac{p(p+2)}{\mathcal{J}_p(K)}\|\mathbf{T}_{K,g_1}^{(n)}\|^2 + o_{\mathrm{P}}(1)$$

under $\mathrm{P}_{\boldsymbol{\theta}_0,r_n,v,g_1}^{(n)}$. This asymptotic equivalence in probability, that thus holds in all regimes (i)–(iv) from Theorem 3.1, greatly simplifies the study of the asymptotic behavior of the test ϕ_K under weak identifiability. The following result provides the null asymptotic behavior of this test (see Section S.1 of the supplement for a proof).

THEOREM 3.3. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $v > 0$, a bounded positive sequence (r_n), and $g_1 \in \mathcal{F}_1^a$, and let Assumption (A) hold. Then, under $\mathrm{P}_{\boldsymbol{\theta}_0,r_n,v,g_1}^{(n)}$, the test statistic Q_K is asymptotically chi-square with $p-1$ degrees of freedom.*

It follows from Theorem 3.3 that the rank test ϕ_K is asymptotically robust to weak identifiability, without any moment assumption on the underlying elliptical distribution. To check that such a strong validity-robustness is not obtained at the expense of power, we derive next the asymptotic behavior of ϕ_K under the same local alternatives as in Theorem 3.2. We have the following result (see again Section S.1 of the supplement for a proof).

THEOREM 3.4. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $v>0$ and $g_1 \in \mathcal{F}_1^a$, and let Assumption (A) hold. Then, we have the following (in each case, $(\boldsymbol{\tau}_n)$ is a sequence in $\mathbb{R}^p$ that satisfies (3.5) and that converges to $\boldsymbol{\tau}$):*

(i) *if $r_n \equiv 1$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+\boldsymbol{\tau}_n/\sqrt{n},r_n,v,g_1}$, the test statistic Q_K is asymptotically non-central chi-square with $p-1$ degrees of freedom and non-centrality parameter*

$$\frac{\mathcal{J}_p^2(K,g_1)v^2}{\mathcal{J}_p(K)p(p+2)(1+v)}\|\boldsymbol{\tau}\|^2;$$

(ii) *if $r_n = o(1)$ with $\sqrt{n}r_n \to \infty$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+\boldsymbol{\tau}_n/(\sqrt{n}r_n),r_n,v,g_1}$, Q_K is asymptotically non-central chi-square with $p-1$ degrees of freedom and non-centrality parameter*

$$\frac{\mathcal{J}_p^2(K,g_1)v^2}{\mathcal{J}_p(K)p(p+2)}\|\boldsymbol{\tau}\|^2;$$

(iii) *if $r_n = 1/\sqrt{n}$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+\boldsymbol{\tau}_n,r_n,v,g_1}$, Q_K is asymptotically non-central chi-square with $p-1$ degrees of freedom and non-centrality parameter*

$$\frac{v^2\mathcal{J}_p^2(K,g_1)}{16\mathcal{J}_p(K)p(p+2)}\|\boldsymbol{\tau}\|^2(4-\|\boldsymbol{\tau}\|^2)(2-\|\boldsymbol{\tau}\|^2)^2; \tag{3.11}$$

(iv) *if $r_n = o(1/\sqrt{n})$, then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+\boldsymbol{\tau}_n,r_n,v,g_1}$, Q_K is asymptotically chi-square with $p-1$ degrees of freedom.*

By combining Theorems 3.2 and 3.4, we can obtain the Pitman asymptotic relative efficiencies (AREs) of the rank test ϕ_K with respect to the pseudo-Gaussian test ϕ_G. Recall that the ARE of a test ϕ_a with respect to a test ϕ_b is defined as the limit, when it exists, as n diverges to infinity, of the ratio $N(n)/n$ of the sample size $N(n)$ that is needed by ϕ_b to match the local power ϕ_a will show based on a sample of size n. As usual, such AREs in the present situation are obtained as the ratio of the non-centrality parameters in the corresponding non-null asymptotic distributions of these test statistics. For any sequence (r_n) in regimes (i)–(iii) (regime (iv) is irrelevant here since no test can show non-trivial asymptotic powers in this regime), the resulting AREs thus are given by

$$\mathrm{ARE}_{\boldsymbol{\theta}_0,r_n,v,g_1}(\phi_K/\phi_G) := \frac{(1+\kappa_p(g_1))\mathcal{J}_p^2(K,g_1)}{p(p+2)\mathcal{J}_p(K)}. \tag{3.12}$$

Since this expression does not depend on the underlying sequence (r_n), we conclude that the AREs of the rank tests with respect to their pseudo-Gaussian competitor are the same in the weak identifiability regimes (ii)–(iii) as in the standard regime (i). As a direct corollary, the Chernoff–Savage result of Paindaveine (2006) is robust to weak identifiability: irrespective of the severity of the possible weak identifiability (with the exception, again, of the extreme regime (iv)), the ARE of the Gaussian-score rank test with respect to the pseudo-Gaussian test is larger than or equal to one at any g_1 and is equal to one only when g_1 is the Gaussian radial density. In Table 1, we provide some numerical values of the ARE in (3.12) for various rank tests and for various underlying distributions. This illustrates in particular the aforementioned Chernoff–Savage result.

4. Simulations. In this section, we perform several Monte-Carlo exercises (a) to check that the various tests considered in this work are robust to weak identifiability under the null hypothesis (i.e., that they show asymptotically the target Type I risk in each asymptotic

		Underlying density						
K	p	t_5	t_8	t_{12}	$\mathcal{N}$	e_2	e_3	e_5
vdW	2	2.204	1.215	1.078	1.000	1.129	1.308	1.637
	3	2.270	1.233	1.086	1.000	1.108	1.259	1.536
	4	2.326	1.249	1.093	1.000	1.093	1.223	1.462
	6	2.413	1.275	1.106	1.000	1.072	1.174	1.363
	10	2.531	1.312	1.126	1.000	1.050	1.121	1.254
	∞	3.000	1.500	1.250	1.000	1.000	1.000	1.000
L	2	1.500	0.750	0.625	0.500	0.392	0.365	0.347
	3	1.800	0.900	0.750	0.600	0.493	0.464	0.444
	4	2.000	1.000	0.833	0.667	0.565	0.537	0.517
	6	2.250	1.125	0.938	0.750	0.662	0.636	0.617
	10	2.500	1.250	1.041	0.833	0.766	0.746	0.730
	∞	3.000	1.500	1.250	1.000	1.000	1.000	1.000
W	2	2.258	1.174	1.067	0.844	0.789	0.804	0.842
	3	2.386	1.246	1.070	0.913	0.897	0.933	1.001
	4	2.432	1.273	1.094	0.945	0.955	1.006	1.095
	6	2.451	1.283	1.105	0.969	1.008	1.075	1.188
	10	2.426	1.264	1.088	0.970	1.032	1.106	1.233
	∞	2.250	1.125	0.938	0.750	0.750	0.750	0.750
SP	2	2.301	1.230	1.067	0.934	0.965	1.042	1.168
	3	2.277	1.225	1.070	0.957	1.033	1.141	1.317
	4	2.225	1.200	1.051	0.956	1.057	1.179	1.383
	6	2.128	1.146	1.007	0.936	1.057	1.189	1.414
	10	2.001	1.068	0.936	0.891	1.017	1.144	1.365
	∞	1.667	0.833	0.694	0.556	0.556	0.556	0.556

TABLE 1

AREs of the van der Waerden (vdW), Laplace (L), Wilcoxon (W), and Spearman (SP) rank tests with respect to the pseudo-Gaussian test, under p-dimensional t (with 5, 8, and 12 degrees of freedom), normal, and power-exponential densities (with parameter $\eta \in \{2,3,5\}$), for $p \in \{2,3,4,6,10\}$ and $p \to \infty$.

regime), (b) to compare the small-sample powers of the rank and pseudo-Gaussian tests under weak identifiability, and (c) to check correctness of the highly non-standard asymptotic result obtained in Theorem 3.4(iii).

(a) In a first simulation exercise, we generated for any $b = 0, 1, \ldots, 5$, $M = 2{,}500$ mutually independent random samples $\mathbf{X}_i^{(b,s)}$, $i = 1, \ldots, n = 200$, from several trivariate elliptical distributions with location zero and scatter matrix

$$\boldsymbol{\Sigma}_n^{(b)} := \mathbf{I}_3 + n^{-b/6}\boldsymbol{\theta}_0\boldsymbol{\theta}_0', \qquad \text{with } \boldsymbol{\theta}_0 = (1,0,0)'. \tag{4.1}$$

For $s = 1, 2, 3$, the $\mathbf{X}_i^{(b,s)}$'s are sampled from a t distribution with one degree of freedom, from a t distribution with 5 degrees of freedom, and from the multinormal distribution, respectively. For each sample, we performed the following tests for $\mathcal{H}_0^{(n)} : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ at asymptotic level 5%: the pseudo-Gaussian test, the Wilcoxon rank test, and the van der Waerden rank test. The value of b allows us to consider the various regimes considered in the paper, namely regime (i) ($b = 0$), regime (ii) ($b = 1, 2$), regime (iii) ($b = 3$), and regime (iv) ($b = 4, 5$). Increasing values of b therefore provide weaker and weaker identifiability. Figure 1, that reports the resulting rejection frequencies, indicates that the pseudo-Gaussian test asymptotically shows the target Type I risk only when moments of order four are finite (which is not the case for the t distribution with one degree of freedom). In contrast, both considered rank tests achieve the target Type 1 risk in all cases, hence are validity-robust to both heavy tails and weak identifiability.

(b) In a second simulation exercise, we generated $M = 2{,}500$ mutually independent random samples $\mathbf{X}_i^{(b,s,\ell)}$, $i = 1, \ldots, n = 500$, $b = 0, 1, 2$, $\ell = 0, 1, \ldots, 5$, from several trivariate

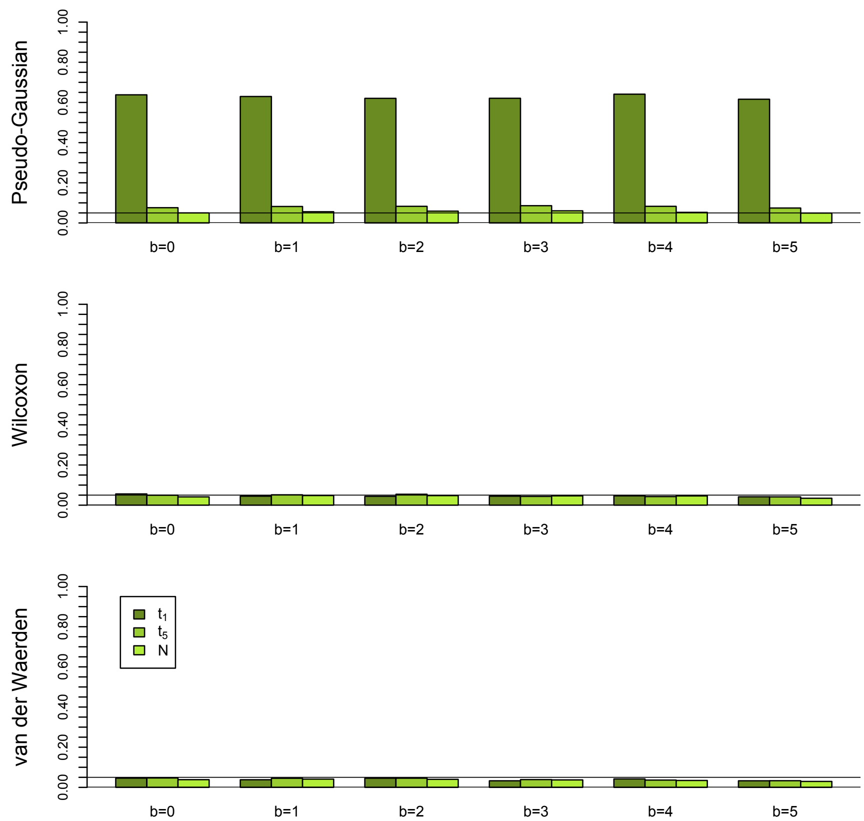

FIG 1. *Empirical rejection frequencies, under the null hypothesis, of the pseudo-Gaussian test, the Wilcoxon rank test, and the van der Waerden rank test, all performed at asymptotic level* 5%. *Results are based on* $M = 2{,}500$ *mutually independent three-dimensional random samples of size* $n = 200$, *obtained from the* t *distribution with one degree of freedom* (t_1), *the* t *distribution with 5 degrees of freedom* (t_5), *and the multinormal distribution* (N). *Increasing values of* b *bring the underlying spiked shape matrix closer and closer to the identity matrix, hence provide weaker and weaker identifiability.*

elliptical distributions with location zero and scatter matrix

$$\boldsymbol{\Sigma}_n^{(b,\ell)} := \mathbf{I}_3 + n^{-b/6}(\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell)(\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell)', \tag{4.2}$$

with $\boldsymbol{\theta}_0 = (1,0,0)'$ and $\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell = (\cos(\ell\pi/25), \sin(\ell\pi/25)), 0)'$. For $s = 1, 2, 3$, the $\mathbf{X}_i^{(b,s,\ell)}$'s are sampled from a t distribution with one degree of freedom, from a t distribution with 5 degrees of freedom, and from the multinormal distribution, respectively. The value $\ell = 0$ is associated with the null hypothesis, whereas the values $\ell = 1, \ldots, 5$ provide increasingly severe alternatives. The larger b, the more severe the weak identifiability (only regimes (i) and (ii) are considered here), hence the more challenging the considered testing problem. For each sample, we performed the same three tests for $\mathcal{H}_0^{(n)} : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ as in the first simulation exercise, still at asymptotic level 5%. The resulting rejection frequencies are plotted in Figure 2. Inspection of this figure reveals that, as expected from the AREs in Table 1, the Wilcoxon rank test is doing very well globally in terms of power. The pseudo-Gaussian test and the van der Waerden rank test share the same performances in the Gaussian case, but, in line with the theory, the van der Waerden test dominates its parametric competitor away from the Gaussian case. Clearly, the pseudo-Gaussian test does not meet the nominal level constraint in the Cauchy case, where fourth-order moments are infinite.

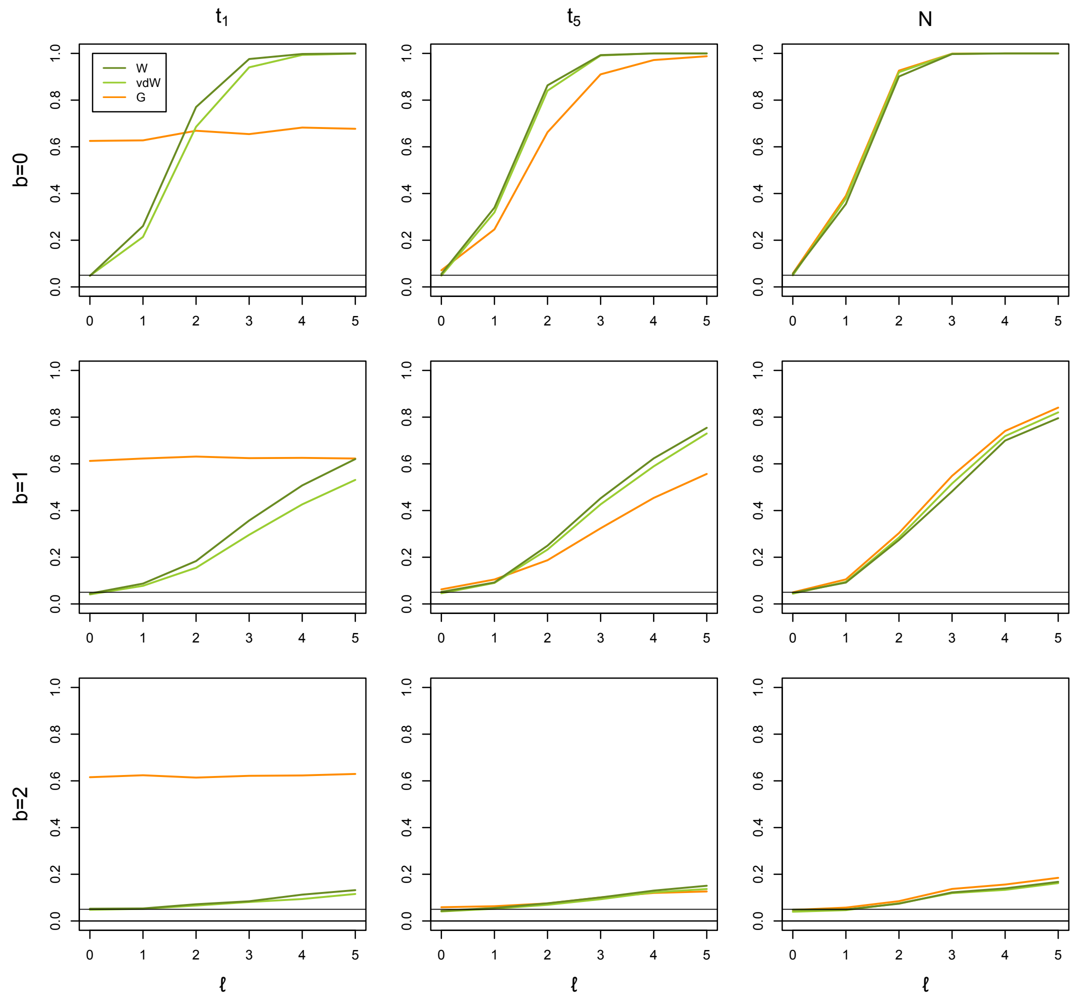

FIG 2. *Empirical rejection frequencies, under the null hypothesis and under some fixed alternatives, of the pseudo-Gaussian test (G), the Wilcoxon rank test (W), and the van der Waerden rank test (vdW), all performed at asymptotic level* 5%. *Results are based on* $M = 2{,}500$ *mutually independent three-dimensional random samples of size* $n = 500$, *obtained from the* t *distribution with one degree of freedom* (t_1), *the* t *distribution with* 5 *degrees of freedom* (t_5), *and the multinormal distribution* (N). *Increasing values of* b *bring the underlying spiked shape matrix closer and closer to the identity matrix, while increasing values of* ℓ *provide more and more severe alternatives (the null hypothesis is obtained for* $\ell = 0$).

(c) In a last simulation exercise, we generated $M = 100{,}000$ mutually independent random samples $\mathbf{X}_i^{(\ell)}$, $i = 1, \ldots, n = 10,000$, $\ell = 0, 1, \ldots, 20$, from the bivariate t distribution with one degree of freedom, location zero, and scatter matrix

$$\boldsymbol{\Sigma}_n^{(\ell)} := \mathbf{I}_p + n^{-1/2}(\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell)(\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell)', \tag{4.3}$$

with $\boldsymbol{\theta}_0 = (1, 0)'$ and $\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell = (\cos(\ell\pi/40), \sin(\ell\pi/40))'$. For each sample, we performed the van der Waerden rank test for $\mathcal{H}_0^{(n)} : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ at asymptotic level 5%. The resulting rejection frequencies are plotted in Figure 3, along with the asymptotic powers obtained through Theorem 3.4(iii). The results clearly support correctness of Theorem 3.4(iii).

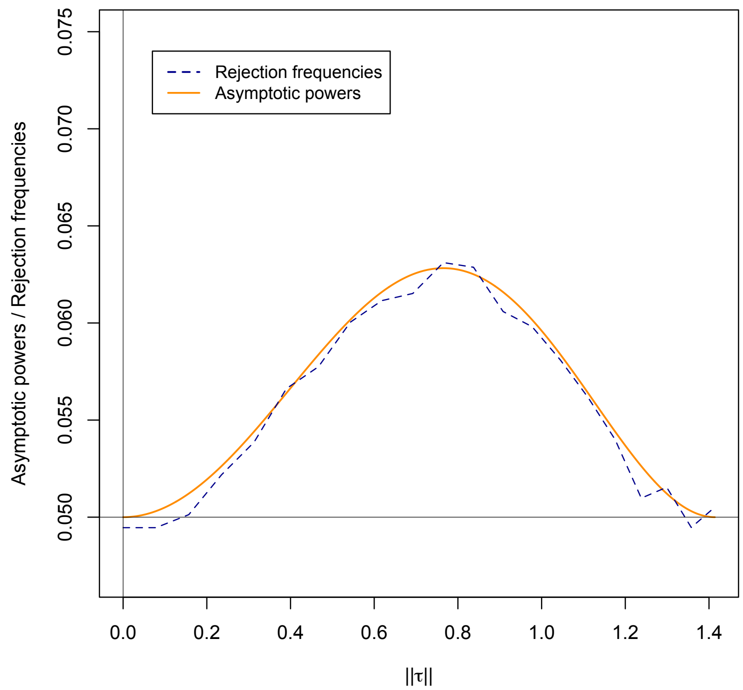

FIG 3. *Empirical rejection frequencies (dashed line), under the null hypothesis* ($\|\boldsymbol{\tau}\| = 0$) *and local alternatives* ($\|\boldsymbol{\tau}\| > 0$) *in regime (iii), of the van der Waerden test performed at asymptotic level* 5%. *The corresponding asymptotic powers obtained from Theorem 3.4(iii) are also provided (solid line).*

5. Power-enhanced tests. While the tests considered in the previous sections behave well under local alternatives[1], they actually suffer consistency issues under specific fixed alternatives, even in cases that do not involve weak identifiability. To show this, assume that the observations form a random sample from an elliptical distribution with a radial density g_1 ensuring finite fourth-order moments (so that the population kurtosis $\kappa_p(g_1)$ is well-defined) and with covariance matrix $\boldsymbol{\Sigma}_{\rm cov} := \xi(\mathbf{I}_p + v\boldsymbol{\theta}_1\boldsymbol{\theta}_1')$, where $\xi, v > 0$ and $\boldsymbol{\theta}_1 \in \mathcal{S}_+^{p-1}$ are fixed. Then, we readily have that the pseudo-Gaussian test statistic $Q_{\rm G}$ in (2.7) is such that

$$
\begin{aligned}
n^{-1}Q_{\rm G} &= \frac{1}{(1+\hat{\kappa}_n)\hat{\lambda}_{n1}} \sum_{j=2}^{p} \hat{\lambda}_{nj}^{-1} \big(\tilde{\boldsymbol{\theta}}_j' \mathbf{S}^{(n)} \boldsymbol{\theta}_0\big)^2 \\
&\xrightarrow{\rm P} \frac{1}{\xi^2(1+\kappa_p(g_1))(1+v)} \boldsymbol{\theta}_0' \boldsymbol{\Sigma}_{\rm cov} (\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0') \boldsymbol{\Sigma}_{\rm cov} \boldsymbol{\theta}_0 \\
&= \frac{v^2}{(1+\kappa_p(g_1))(1+v)} (1 - (\boldsymbol{\theta}_1'\boldsymbol{\theta}_0)^2)(\boldsymbol{\theta}_0'\boldsymbol{\theta}_1)^2 =: \xi(\boldsymbol{\theta}_1),
\end{aligned}
$$

where $\xrightarrow{\rm P}$ denotes convergence in probability. If $\boldsymbol{\theta}_1 \in \mathcal{S}_+^{p-1} \setminus \{\boldsymbol{\theta}_0\}$ is such that $\boldsymbol{\theta}_1'\boldsymbol{\theta}_0 \neq 0$, then we have $\xi(\boldsymbol{\theta}_1) > 0$, so that the pseudo-Gaussian test $\phi_{\rm G}$ is consistent (note that $\boldsymbol{\theta}_1'\boldsymbol{\theta}_0$

[1]In particular, irrespective of the possible weak identifiability, the AREs of the Gaussian-score rank test with respect to the pseudo-Gaussian test are larger than or equal to one at any radial density g_1 and are equal to one only when g_1 is the Gaussian radial density.

cannot be equal to -1 since $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1 \in \mathcal{S}_+^{p-1}$). However, if $\boldsymbol{\theta}_1'\boldsymbol{\theta}_0 = 0$, then $\tilde{\boldsymbol{\theta}}_j'\boldsymbol{\Sigma}_{\rm cov}\boldsymbol{\theta}_0 = \xi\tilde{\boldsymbol{\theta}}_j'\boldsymbol{\theta}_0 = 0$ for $j = 2, \ldots, p$, so that Lemma 5.1 from PRV20 entails that

$$Q_{\rm G} = \frac{1}{(1+\hat{\kappa}_n)\hat{\lambda}_{n1}} \sum_{j=2}^{p} \hat{\lambda}_{nj}^{-1} \big(\tilde{\boldsymbol{\theta}}_j' \sqrt{n}(\mathbf{S}^{(n)} - \boldsymbol{\Sigma}_{\rm cov})\boldsymbol{\theta}_0\big)^2$$

$$= \frac{1}{\xi^2(1+\kappa_p(g_1))(1+v)} \|(\boldsymbol{\theta}_0' \otimes (\mathbf{I}_p - \boldsymbol{\theta}_0\boldsymbol{\theta}_0'))\sqrt{n}\,\mathrm{vec}(\mathbf{S}^{(n)} - \boldsymbol{\Sigma}_{\rm cov})\|^2 + o_{\rm P}(1)$$

is asymptotically χ^2_{p-1}, which implies that the asymptotic power of the pseudo-Gaussian test $\phi_{\rm G}$ is then equal to the nominal level α. It is easy to show that, similarly, the rank test ϕ_K is consistent if and only if the fixed alternative $\boldsymbol{\theta}_1 \in \mathcal{S}_+^{p-1} \setminus \{\boldsymbol{\theta}_0\}$ satisfies $\boldsymbol{\theta}_1'\boldsymbol{\theta}_0 \neq 0$. In the present single-spiked scenario, the tests studied in the paper thus actually address the larger null hypothesis that $\boldsymbol{\theta}_0$ is an eigenvector of the underlying scatter matrix (as opposed to the null hypothesis that it is the eigenvector associated to the leading eigenvalue).

In order to correct the lack of consistency obtained when $\boldsymbol{\theta}_1'\boldsymbol{\theta}_0 = 0$, we propose the *power-enhanced* pseudo-Gaussian and rank tests rejecting the null hypothesis at asymptotic level α whenever

$$Q_{\rm G}^{\rm PE} := Q_{\rm G} + n(\hat{\lambda}_{n1} - \hat{\lambda}_{n2})^2 \big(\hat{\boldsymbol{\theta}}_1'\boldsymbol{\theta}_0 - 1\big)^2 > \chi^2_{p-1,1-\alpha} \tag{5.1}$$

and

$$Q_K^{\rm PE} := Q_K + n(\hat{\lambda}_{1,\rm Tyler} - \hat{\lambda}_{2,\rm Tyler})^2 \big(\hat{\boldsymbol{\theta}}_{1,\rm Tyler}'\boldsymbol{\theta}_0 - 1\big)^2 > \chi^2_{p-1,1-\alpha}, \tag{5.2}$$

respectively; here, $\hat{\lambda}_{n1}$, $\hat{\lambda}_{n2}$ and $\hat{\boldsymbol{\theta}}_1$ are respectively the two largest eigenvalues and the leading eigenvector of the empirical covariance matrix $\mathbf{S}^{(n)}$, whereas $\hat{\lambda}_{1,\rm Tyler}$, $\hat{\lambda}_{2,\rm Tyler}$ and $\hat{\boldsymbol{\theta}}_{1,\rm Tyler}$ are the corresponding quantities obtained from the estimator of shape $\hat{\mathbf{V}}_{\rm Tyler}$ considered in Section 2. We then have the following result (see Section S.2 of the supplement for a proof).

PROPOSITION 5.1. *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}_+^{p-1}$, $v > 0$, a bounded positive sequence (r_n) such that $\sqrt{n}r_n \to \infty$, and a radial density g_1 that, for the results on $Q_{\rm G}^{\rm PE}$ below, should ensure finite fourth-order moments (such moment condition is not needed for the results on $Q_K^{\rm PE}$). Then,*

(i) *under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,r_n,v,g_1}$, $Q_{\rm G}^{\rm PE} = Q_{\rm G} + o_{\rm P}(1)$ and $Q_K^{\rm PE} = Q_K + o_{\rm P}(1)$ as $n \to \infty$;*
(ii) *for any $\boldsymbol{\theta}_1 \in \mathcal{S}_+^{p-1} \setminus \{\boldsymbol{\theta}_0\}$,*

$$\mathrm{P}^{(n)}_{\boldsymbol{\theta}_1,r_n,v,g_1}[Q_{\rm G}^{\rm PE} > t] \to 1 \quad \text{and} \quad \mathrm{P}^{(n)}_{\boldsymbol{\theta}_1,r_n,v,g_1}[Q_K^{\rm PE} > t] \to 1$$

as $n \to \infty$ for any $t > 0$.

Proposition 5.1(i) entails that, in regimes (i)–(ii), the power-enhanced tests in (5.1)–(5.2) have asymptotic level α under the null hypothesis and, from contiguity, share the same local asymptotic powers as their original versions (in particular, the power-enhanced Gaussian-score rank test still uniformly dominates the power-enhanced pseudo-Gaussian test in terms of AREs). Proposition 5.1(ii), however, implies that the power-enhanced tests are consistent under *any* fixed alternative in regimes (i)–(ii). Note that this is as good as one may expect, as such universal consistency cannot be achieved in regime (iii): since any fixed alternative $\boldsymbol{\theta}_1$ is contiguous to the null hypothesis in regime (iii), the asymptotic equivalences in Proposition 5.1(i) that would be obtained from any power-enhancement mechanism would indeed automatically extend to the alternatives with $\boldsymbol{\theta}_1'\boldsymbol{\theta}_0 = 0$, so that the corresponding tests would suffer the same consistency issue as the original pseudo-Gaussian and rank tests.

We conclude this section with a simulation exercise to illustrate the results obtained above. We generated $M = 10{,}000$ mutually independent random samples $\mathbf{X}_i^{(b,\ell,s)}$, $i = 1, \ldots, n = 500$, $b = 0, 1$, $\ell = 0, 1, \ldots, L = 50$, $s = 1, 2, 3$, from several bivariate elliptical distributions with location zero and scatter matrix

$$\boldsymbol{\Sigma}_n^{(b,\ell)} := \mathbf{I}_2 + n^{-b/6}(\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell)(\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell)', \tag{5.3}$$

where $\boldsymbol{\theta}_0 = (1, 0)'$ and $\boldsymbol{\theta}_0 + \boldsymbol{\tau}_\ell = (\cos(\ell\pi/(2L)), \sin(\ell\pi/(2L)))'$. For $s = 1, 2, 3$, the $\mathbf{X}_i^{(b,\ell,s)}$'s are sampled from a t distribution with one degree of freedom, from a t distribution with 5 degrees of freedom, and from a multinormal distribution, respectively. The value $\ell = 0$ is associated with the null hypothesis $\mathcal{H}_0^{(n)} : \boldsymbol{\theta} = \boldsymbol{\theta}_0$, whereas the values $\ell = 1, \ldots, L$ provide increasingly severe alternatives; note that $\boldsymbol{\theta}_0'\boldsymbol{\theta}$ decreases from 1 to 0 as ℓ ranges from 0 to L, so that $\ell = L$ provides a value of $\boldsymbol{\theta}$ that is orthogonal to $\boldsymbol{\theta}_0$. For each sample, we performed the pseudo-Gaussian test ϕ_{G} and the van der Waerden rank test ϕ_K, as well as their power-enhanced versions in (5.1)–(5.2), all at asymptotic level 5%. The resulting rejection frequencies are plotted in Figure 4. The results confirm that the original pseudo-Gaussian and van der Waerden tests are blind to alternatives that are orthogonal to $\boldsymbol{\theta}_0$ and that such alternatives are detected by the power-enhanced tests.

6. Extension to more general spectra. The asymptotic analysis provided in Section 3 focused on scatter matrices of the form $\boldsymbol{\Sigma}_n := \xi(\mathbf{I}_p + r_n v\boldsymbol{\theta}\boldsymbol{\theta}')$, with $\xi, v > 0$ and $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$, involving a single-spiked spectrum for which, when excluding the leading eigenvalue, all other eigenvalues are equal. In the present section, our objective is to show that the main results obtained in Section 3 can be extended to much more general spectra. More precisely, we will consider triangular arrays of elliptically symmetric observations $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ with location parameter $\boldsymbol{\mu}$, scatter matrix

$$\boldsymbol{\Sigma}_n := \xi\left\{(1 + r_n v)\boldsymbol{\theta}\boldsymbol{\theta}' + \left(\mathbf{I}_p - \boldsymbol{\theta}\boldsymbol{\theta}' - \sum_{j=q+1}^{p} \boldsymbol{\theta}_j\boldsymbol{\theta}_j'\right) + \sum_{j=q+1}^{p} \lambda_j \boldsymbol{\theta}_j\boldsymbol{\theta}_j'\right\},$$

and radial density $g_1 \in \mathcal{F}_1$; here, $\xi, v > 0$, $1 > \lambda_{q+1} > \ldots > \lambda_p > 0$, $\boldsymbol{\theta}, \boldsymbol{\theta}_{q+1}, \ldots, \boldsymbol{\theta}_p$ are pairwise orthogonal vectors in $\mathcal{S}_+^{p-1}$, and (r_n) is a positive real sequence. The same invariance arguments as in Section 3 imply that we may assume without loss of generality that $\boldsymbol{\mu} = \mathbf{0}$ and $\xi = 1$. The scatter parameter $\boldsymbol{\Sigma}_n$ has then eigenvalues

$$\lambda_{n1} = 1 + r_n v, \quad \lambda_2 = \cdots = \lambda_q = 1 > \lambda_{q+1} > \ldots > \lambda_p; \tag{6.1}$$

obviously, $\boldsymbol{\theta}$ is the eigenvector in $\mathcal{S}_+^{p-1}$ associated with the leading eigenvalue λ_{n1}, $\boldsymbol{\theta}_{q+1}, \ldots, \boldsymbol{\theta}_p$ are the eigenvectors in $\mathcal{S}_+^{p-1}$ associated with the eigenvalues $\lambda_{q+1}, \ldots, \lambda_p$, respectively, whereas the eigenvectors associated with the eigenvalue 1 span the orthogonal complement to the vector subspace that is spanned by $\boldsymbol{\theta}, \boldsymbol{\theta}_{q+1}, \ldots, \boldsymbol{\theta}_p$. The eigenvectors structure is thus fully determined by the $p \times (p - q + 1)$ matrix $\boldsymbol{\Gamma} := (\boldsymbol{\theta}\ \boldsymbol{\theta}_{q+1} \ldots \boldsymbol{\theta}_p)$. We will denote the resulting sequence of hypotheses as $\mathrm{P}^{(n)}_{\boldsymbol{\Gamma}, r_n, v, \boldsymbol{\lambda}, g_1}$, with $\boldsymbol{\lambda} = (\lambda_{q+1}, \ldots, \lambda_p)'$.

We aim at studying the asymptotic behavior of the rank test ϕ_K and of the pseudo-Gaussian test ϕ_{G} under local alternatives involving the general spectra above. Since the eigenvectors $\boldsymbol{\theta}_{q+1}, \ldots, \boldsymbol{\theta}_p$ must belong to the orthogonal complement to $\boldsymbol{\theta}$, perturbing $\boldsymbol{\theta}$ induces a perturbation of $\boldsymbol{\theta}_{q+1}, \ldots, \boldsymbol{\theta}_p$. Accordingly, we will consider local perturbations of the form

$$\boldsymbol{\Gamma} + \boldsymbol{\tau}_n \boldsymbol{\nu}_n := \boldsymbol{\Gamma} + (\boldsymbol{\tau}_{n1}\ \boldsymbol{\tau}_{n,q+1}\ \ldots\ \boldsymbol{\tau}_{np}) \begin{pmatrix} \nu_n & 0 \\ 0 & n^{-1/2}\mathbf{I}_{p-q} \end{pmatrix}$$

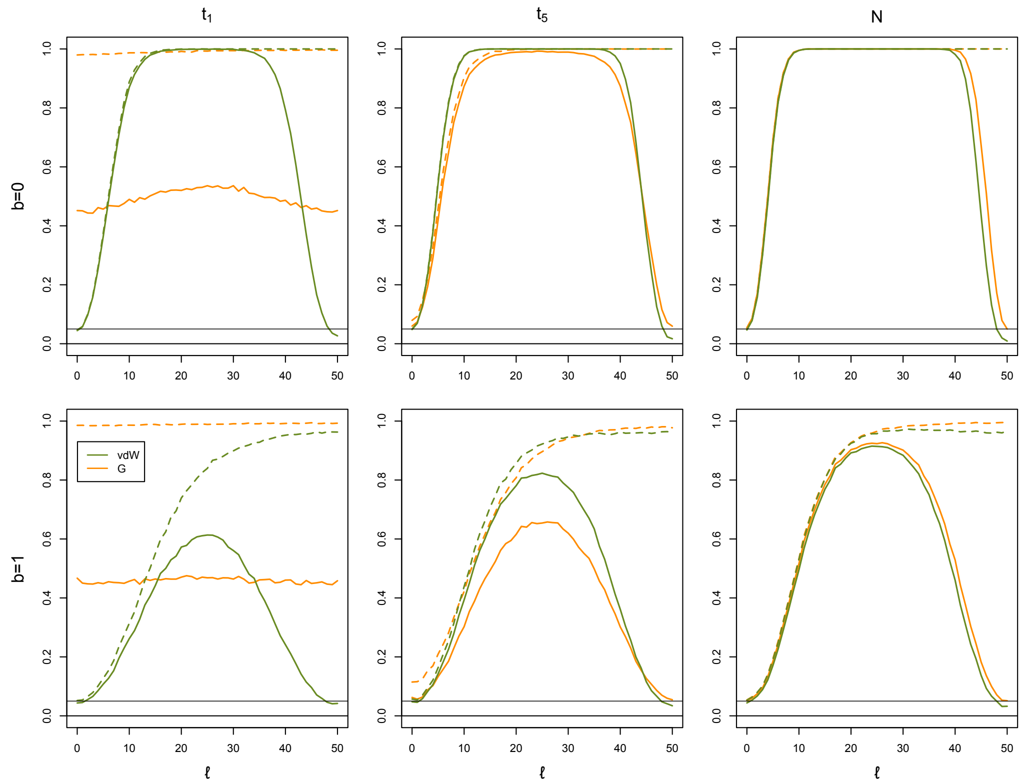

FIG 4. *Empirical rejection frequencies, under the null hypothesis and under some fixed alternatives, of the pseudo-Gaussian test (G, solid line) and the van der Waerden rank test (vdW, solid line), and of their power-enhanced version in (5.1)–(5.2) (G and vdW, dashed lines), all performed at asymptotic level* 5%*. Results are based on* $M = 10{,}000$ *mutually independent two-dimensional random samples of size* $n = 500$*, obtained from the* t *distribution with one degree of freedom* (t_1)*, the* t *distribution with* 5 *degrees of freedom* (t_5)*, and the multinormal distribution* (N)*. The values* $b = 0$ *and* $b = 1$ *correspond to regimes (i) and (ii), respectively. Increasing values of* ℓ *provide more and more severe alternatives* (*the null hypothesis is obtained for* $\ell = 0$ *and the most severe alternatives, obtained for* $\ell = 50$*, are associated with a leading eigenvector that is orthogonal to the null leading eigenvector* $\boldsymbol{\theta}_0$)*; see Section 5 for details.*

$$= (\boldsymbol{\theta} + \nu_n \boldsymbol{\tau}_{n1} \;\; \boldsymbol{\theta}_{q+1} + n^{-1/2}\boldsymbol{\tau}_{n,q+1} \;\; \ldots \;\; \boldsymbol{\theta}_p + n^{-1/2}\boldsymbol{\tau}_{np}),$$

for some suitable positive real sequence (ν_n). Since the column vectors of $\boldsymbol{\Gamma} + \boldsymbol{\tau}_n \boldsymbol{\nu}_n$ should be pairwise orthogonal vectors in $\mathcal{S}_+^{p-1}$, we must have

$$\boldsymbol{\Gamma}'\boldsymbol{\tau}_n\boldsymbol{\nu}_n + \boldsymbol{\nu}_n'\boldsymbol{\tau}_n'\boldsymbol{\Gamma} + \boldsymbol{\nu}_n'\boldsymbol{\tau}_n'\boldsymbol{\tau}_n\boldsymbol{\nu}_n = \mathbf{0} \quad \text{for any } n. \tag{6.2}$$

The following result then describes the corresponding null and non-null asymptotic behaviors of the rank test ϕ_K, hence extends to the general spectra above the results in Theorems 3.3–3.4 focusing on single-spiked spectra (see Section S.3 of the supplement for a proof).

THEOREM 6.1. *Fix a matrix* $\boldsymbol{\Gamma}_0 := (\boldsymbol{\theta}_0\, \boldsymbol{\theta}_{q+1} \ldots \boldsymbol{\theta}_p)$ *with pairwise orthogonal column vectors in* $\mathcal{S}_+^{p-1}$, $v > 0$, $\boldsymbol{\lambda} = (\lambda_{q+1}, \ldots, \lambda_p)'$ *with* $0 < \lambda_p < \ldots < \lambda_{q+1} < 1$*, and* $g_1 \in \mathcal{F}_1^a$*, and let Assumption (A) hold. Let* $\boldsymbol{\tau}_n = (\boldsymbol{\tau}_{n1}\, \boldsymbol{\tau}_{n,q+1} \ldots \boldsymbol{\tau}_{np})$ *be a sequence of bounded matrices such that (6.2) holds and such that* $\boldsymbol{\tau} = (\boldsymbol{\tau}_1\, \boldsymbol{\tau}_{q+1} \ldots \boldsymbol{\tau}_p) := \lim_{n\to\infty} \boldsymbol{\tau}_n$ *exists. Then, we have the following:*

(i) *if* $r_n \equiv 1$, *then, under* $\mathrm{P}^{(n)}_{\boldsymbol{\Gamma}_0+\boldsymbol{\tau}_n\boldsymbol{\nu}_n,r_n,v,\boldsymbol{\lambda},g_1}$ *with* $\nu_n = 1/\sqrt{n}$, Q_K *is asymptotically non-central chi-square with* $p-1$ *degrees of freedom and non-centrality parameter*

$$\frac{\mathcal{J}_p^2(K,g_1)}{\mathcal{J}_p(K)p(p+2)(1+v)} \times\left\{ v^2\|\boldsymbol{\tau}_1\|^2 + \sum_{j=q+1}^{p} \frac{1}{\lambda_j}\left(v^2(1-\sqrt{\lambda_j})^2 + 2v(1-\lambda_j) + (1-\lambda_j)^2 \right)(\boldsymbol{\tau}_j'\boldsymbol{\theta}_0)^2 \right\};$$

(ii) *if* $r_n = o(1)$ *with* $\sqrt{n}r_n \to \infty$, *then, under* $\mathrm{P}^{(n)}_{\boldsymbol{\Gamma}_0+\boldsymbol{\tau}_n\boldsymbol{\nu}_n,r_n,v,\boldsymbol{\lambda},g_1}$ *with* $\nu_n = 1/(\sqrt{n}r_n)$, Q_K *is asymptotically non-central chi-square with* $p-1$ *degrees of freedom and non-centrality parameter*

$$\frac{\mathcal{J}_p^2(K,g_1)}{\mathcal{J}_p(K)p(p+2)}\left\{ v^2\|\boldsymbol{\tau}_1\|^2 + \sum_{j=q+1}^{p} \frac{(1-\lambda_j)^2}{\lambda_j}(\boldsymbol{\tau}_j'\boldsymbol{\theta}_0)^2 \right\};$$

(iii) *if* $r_n = 1/\sqrt{n}$, *then, under* $\mathrm{P}^{(n)}_{\boldsymbol{\Gamma}_0+\boldsymbol{\tau}_n\boldsymbol{\nu}_n,r_n,v,\boldsymbol{\lambda},g_1}$ *with* $\nu_n \equiv 1$, Q_K *is asymptotically non-central chi-square with* $p-1$ *degrees of freedom and non-centrality parameter*

$$\frac{\mathcal{J}_p^2(K,g_1)}{\mathcal{J}_p(K)p(p+2)}\left\{ \frac{v^2}{16}\|\boldsymbol{\tau}_1\|^2\big(4-\|\boldsymbol{\tau}_1\|^2\big)\big(2-\|\boldsymbol{\tau}_1\|^2\big)^2 + \sum_{j=q+1}^{p} \frac{(1-\lambda_j)^2}{\lambda_j}(\boldsymbol{\tau}_j'\boldsymbol{\theta}_0)^2 \right\};$$

(iv) *if* $r_n = o(1/\sqrt{n})$, *then, under* $\mathrm{P}^{(n)}_{\boldsymbol{\Gamma}_0+\boldsymbol{\tau}_n\boldsymbol{\nu}_n,r_n,v,\boldsymbol{\lambda},g_1}$ *with* $\nu_n \equiv 1$, Q_K *is asymptotically non-central chi-square with* $p-1$ *degrees of freedom and non-centrality parameter*

$$\frac{\mathcal{J}_p^2(K,g_1)}{\mathcal{J}_p(K)p(p+2)} \sum_{j=q+1}^{p} \frac{(1-\lambda_j)^2}{\lambda_j}(\boldsymbol{\tau}_j'\boldsymbol{\theta}_0)^2.$$

Some comments are in order. First, taking $\boldsymbol{\tau}_n \equiv \mathbf{0}$ in this result shows that, under the general spectra above, Q_K remains asymptotically chi-square with $p-1$ degrees of freedom in regimes (i)–(iv), so that the rank test ϕ_K is still robust to weak identifiability under the null hypothesis. Second, compared to the non-centrality parameters in regimes (i)–(iv) from Theorem 3.4, those in Theorem 6.1 include additional positive contributions involving $(\boldsymbol{\tau}_j'\boldsymbol{\theta}_0)^2$, $j=q+1,\ldots,p$. This can be interpreted as follows: (a) when $\boldsymbol{\tau}_j'\boldsymbol{\theta}_0 \neq 0$, the perturbation/rotation of $\boldsymbol{\theta}_j$ into $\boldsymbol{\theta}_{nj} = \boldsymbol{\theta}_j + n^{-1/2}\boldsymbol{\tau}_{nj}$ induces a perturbation/rotation of $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, which, since $\boldsymbol{\theta}_j$ remains identified in the limit, brings extra asymptotic power at the standard root-n rate. (b) When $\boldsymbol{\tau}_j'\boldsymbol{\theta}_0 = 0$, however, the corresponding rotation of $\boldsymbol{\theta}_j$ fixes $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, so that it is only natural that no extra power is then obtained. Except for this new feature that can only be achieved in the general spectra considered in the present section, the non-null behavior of the rank test ϕ_K is the same for general spectra as for single-spiked ones. Third, while we do not state the result explicitly here in order to save space, working exactly along the same lines as in the proof of Theorem 6.1 allows one to show that the corresponding asymptotic non-null behavior of the pseudo-Gaussian test ϕ_{G} is simply obtained by replacing the factors

$$\frac{\mathcal{J}_p^2(K,g_1)}{\mathcal{J}_p(K)p(p+2)} \quad \text{with} \quad \frac{1}{1+\kappa_p(g_1)}$$

in the statement of Theorem 6.1(i)–(iv), which extends Theorem 3.2 to the more general spectra considered here. An important corollary is that, while the local asymptotic powers of ϕ_K and ϕ_{G} depend on the structure of the underlying spectrum, the AREs of ϕ_K with respect to ϕ_{G} do not (in particular, the uniform dominance of the Gaussian-score rank test over its pseudo-Gaussian competitor not only hold for single-spiked spectra but also for the general spectra we considered in this section). Last, we would like to mention that correctness of Theorem 6.1 was checked through Monte Carlo exercises; for the sake of brevity, we do not provide the corresponding results here, but they are available on a simple request.

7. Conclusions. This work tackled, in a general elliptical framework, the one-sample testing problem on the leading principal direction, under double asymptotic scenarios where this direction is weakly identified. For the first time, the corresponding limiting experiments were studied away from the multinormal case, which revealed that, depending on the severity of weak identifiability, both LAN and non-LAN structures can be obtained. This was a key prerequisite to study the asymptotic null and non-null behaviors of multivariate rank tests under weak identifiability. As it was showed, such rank tests exhibit extremely good robustness to weak identifiability, both in terms of Type 1 risk and in terms of Type 2 risk: irrespective of the severity of the weak identifiability, rank tests show the target size under the null hypothesis as soon as the underlying elliptical density satisfies some mild regularity conditions (that in particular do not impose any moment assumption). Moreover, it remains so under arbitrarily weak identifiability that, when based on normal scores, rank tests uniformly dominate their pseudo-Gaussian competitor in terms of asymptotic relative efficiencies. Power-enhanced versions of our tests allow us to combine this excellent local behavior with consistency under any alternative. Our null and non-null asymptotic findings required deriving general results (i) on the asymptotic behavior of local log-likelihood ratios for triangular arrays and (ii) asymptotic linearity results in a similar framework. The tests considered in the present work achieve semiparametric efficiency bounds at a target radial density f_1 only. Perspectives for future research could thus aim at defining tests that are uniformly semiparametrically efficient, which can be obtained by using scores $K_{\hat{g}_1}$ associated to a suitable kernel density estimator $\hat{g}_1$ of the underlying radial density; we refer to Hallin and Werker (2003) for the general principle and to Section 6.2 of Hallin and Paindaveine (2004) for an application in the elliptical framework. Of course, the high-dimensional case is another most interesting (and most challenging) venue for future research work.

APPENDIX: GENERAL ASYMPTOTIC RESULTS UNDER TRIANGULAR ARRAYS

This appendix provides the general asymptotic results that allowed us to study the asymptotic behavior of local log-likelihood ratios in a triangular array framework and to control aligned ranks under weak identifiability. Since these results are likely to be of interest for other problems, we present them in a generic semiparametric model indexed by a finite-dimensional parameter of interest $\boldsymbol{\theta}$, a finite-dimensional nuisance parameter $\boldsymbol{\eta}$, and an infinite-dimensional nuisance g in some collection $\mathcal{F}$ of functions (for the elliptical testing problem considered in the paper, $\boldsymbol{\theta}$ would be the leading eigenvector of the underlying scatter matrix, $\boldsymbol{\eta}$ would collect the location parameter and the remaining parameters in the scatter matrix, and g would be the “radial” density that makes the difference between, e.g., a multinormal distribution and a multivariate t distribution).

As in the body of the paper, we actually consider triangular arrays of observations where the nth row is made of a random sample from a distribution $\mathrm{P}_{\boldsymbol{\theta},\boldsymbol{\eta}_n,g}$ such that, at the limit as n diverges to infinity, the parameter of interest $\boldsymbol{\theta}$ is not identified: in other words, the sequences $\boldsymbol{\eta}_n$ we consider are such that the weak limits, as n diverges to infinity, of the distributions $\mathrm{P}_{\boldsymbol{\theta},\boldsymbol{\eta}_n,g}$ form a model in which injectivity of the parametrization in $\boldsymbol{\theta}$ is violated. This is precisely what we call *weak identifiability* in this generic context (in the PCA problem we considered, for instance, the leading eigenvector $\boldsymbol{\theta}$ is not identified in the limit if the ratio between both leading eigenvalues (that are part of the nuisance parameter $\boldsymbol{\eta}_n$) converges to one. Under such weak identifiability, local log-likelihood ratios may show non-standard asymptotic behaviors and in particular may provide limiting experiments that are not LAN.

We now provide a general result that allows one to determine the corresponding limiting experiments. In line with the above discussion, we phrase the result in a semiparametric framework, although the result itself is intrinsically parametric: it will be associated to a given value f of the infinite-dimensional nuisance g, namely the value f at which one aims to achieve optimality. Our general result is the following (see Section S.4 of the supplement for a proof).

PROPOSITION A.1. *Fix a sequence $(\boldsymbol{\vartheta}_n)$ in $\boldsymbol{\Theta}(\subset \mathbb{R}^k)$ and $f \in \mathcal{F}$. Let $\mathbf{X}_{ni}$, $i=1,\ldots,n$, $n=1,2,\ldots$ be a triangular array of observations such that, for any n, $\mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}$ form a random sample from the distribution $\mathrm{P}_{\boldsymbol{\vartheta}_n,f}$. Assume that there exist a positive real sequence (ν_n) and real-valued functions $\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}$ such that, for any sequence $(\boldsymbol{\tau}_n)$ in $\mathbb{R}^k$ such that $\boldsymbol{\vartheta}_n+\nu_n\boldsymbol{\tau}_n \in \boldsymbol{\Theta}$ for any n, we have*

$$\int \left\{ f^{1/2}_{\boldsymbol{\vartheta}_n+\nu_n\boldsymbol{\tau}_n}(\mathbf{x}) - f^{1/2}_{\boldsymbol{\vartheta}_n}(\mathbf{x}) - \frac{1}{2}\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{x}) f^{1/2}_{\boldsymbol{\vartheta}_n}(\mathbf{x}) \right\}^2 d\mu(\mathbf{x}) = o(n^{-1}), \tag{A.1}$$

where we let $f_{\boldsymbol{\vartheta}} := \mathrm{dP}_{\boldsymbol{\vartheta},f}/\mathrm{d}\mu$ for some suitable dominating measure μ. Denoting as $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}$ the joint distribution of $\mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}$, further assume that

$$\mathrm{E}_{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}}\big[n\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{n1})\big] = o(1), \tag{A.2}$$

$$\mathcal{I}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n} := n\mathrm{E}_{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}}\Big[\big(\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{n1})\big)^2\Big] = O(1), \tag{A.3}$$

$$\sum_{i=1}^n \big(\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{ni})\big)^2 = \mathcal{I}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n} + o_{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}}(1), \tag{A.4}$$

and that

$$\mathrm{E}_{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}}\Big[n(\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{n1}))^2 \mathbb{I}\big[n(\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{n1}))^2 \geq n\varepsilon^2\big]\Big] = o(1) \tag{A.5}$$

for any $\varepsilon > 0$. Then,

$$\log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n+\nu_n\boldsymbol{\tau}_n,f}}{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}} = \sum_{i=1}^n \dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{ni}) - \frac{1}{2}\mathcal{I}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n} + o_{\mathrm{P}}(1) \tag{A.6}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}$.

By taking $\boldsymbol{\vartheta}_n := (\boldsymbol{\theta},\boldsymbol{\eta}_n)$, the result in (A.6) provides an asymptotic representation of local log-likelihood ratios in a general triangular array framework. In regular situations, Proposition A.1 extends to triangular arrays the well-known result stating that quadratic mean differentiability of an absolutely continuous distribution $\mathrm{P}_{\boldsymbol{\vartheta},f}$ with respect to some dominating measure μ implies LAN of the corresponding sequence of experiments; see, e.g., Theorem 7.2 in van der Vaart (1998). In such cases, (A.6) is a standard *locally asymptotically quadratic (LAQ)* expansion, in which the random term $\dot{\ell}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}(\mathbf{X}_{ni})$ is linear in $\boldsymbol{\tau}_n$ and the deterministic term $\mathcal{I}^{(n)}_{\boldsymbol{\vartheta}_n,\boldsymbol{\tau}_n}$ is quadratic in $\boldsymbol{\tau}_n$. Proposition A.1 may lead to non-standard expansions, that are incompatible with LAN: applying the result in the PCA context described in the introduction led to both LAN and non-LAN structures, depending on the severity of the involved weak identifiability; see Theorem 3.1.

Now, in a semiparametric sequence of probability measures $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta},f}$, indexed by $\boldsymbol{\vartheta} = (\boldsymbol{\theta}, \boldsymbol{\eta}) \in \boldsymbol{\Theta}$ and $g \in \mathcal{F}$, the testing problem

$$\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \qquad \text{against} \qquad \mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0, \tag{A.7}$$

in the standard asymptotic scenario where the finite-dimensional nuisance parameter $\boldsymbol{\eta}$ does not depend on n, can classically be tackled in the following two ways. A first strategy is to use pseudo- or quasi-likelihood methods, that robustify to the present semiparametric framework parametric procedures associated with some fixed $f \in \mathcal{F}$; usually, f is Gaussian, and one then speaks of *pseudo-Gaussian procedures*. Such pseudo- or quasi-likelihood methods have been used, e.g., in Muirhead and Waternaux (1980), Shapiro and Browne (1987), Hallin and Paindaveine (2009), and Hallin, Paindaveine and Verdebout (2010b) to cite only a few. If the parametric submodel obtained by taking $g = f$ is LAN, with central sequence $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\eta},f}$ say (LAN is the rule in the above standard asymptotic scenario where $\boldsymbol{\eta}$ does not depend on n), the resulting tests reject $\mathcal{H}_0$ for large values of

$$(\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,\hat{\boldsymbol{\eta}},f})' \hat{\boldsymbol{\Gamma}}^{-} \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,\hat{\boldsymbol{\eta}},f} \tag{A.8}$$

where $\hat{\boldsymbol{\eta}}$ is a suitable estimator of $\boldsymbol{\eta}$ and $\hat{\boldsymbol{\Gamma}}$ is a consistent estimator of the covariance matrix in the Gaussian weak limit of $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\eta},f}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\eta},g}$; here, $\mathbf{A}^{-}$ stands for the Moore–Penrose generalized inverse of $\mathbf{A}$. Usually, the "studentization" in (A.8) makes such tests valid under a broad class of nuisances g, yet asymptotic parametric efficiency is achieved at f only. Such tests can then be said to be "somewhere parametrically efficient".

A second strategy to define tests in a large class of semiparametric models consists in eliminating the infinite-dimensional nuisance g through invariance arguments. If the testing problem is invariant with respect to some group $\mathcal{G}$ of transformations and if this group generates the nonparametric submodels obtained by fixing a value $(\boldsymbol{\theta}, \boldsymbol{\eta})$ of the finite-dimensional parameter, then an important quantity is the maximal invariant $\mathcal{M}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\eta})$ associated with $\mathcal{G}$. Indeed, tests that are measurable with respect to $\mathcal{M}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\eta})$ will be automatically distribution-free with respect to g. Moreover, Hallin and Werker (2003) showed that, in LAN experiments, tests based on the invariant central sequence at f, namely

$$\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\eta},f} := \mathrm{E}_{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,g}}[\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\eta},f} | \mathcal{M}^{(n)}(\boldsymbol{\theta}_0, \boldsymbol{\eta})],$$

are semiparametrically efficient at f (from invariance, this new central sequence does not depend on the value of g under which the expectation is computed). For the testing problem above, these invariant tests reject $\mathcal{H}_0$ for large values of

$$(\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta}_0,\hat{\boldsymbol{\eta}},f})' \underset{\sim}{\boldsymbol{\Gamma}}^{-} \underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta}_0,\hat{\boldsymbol{\eta}},f}, \tag{A.9}$$

where $\underset{\sim}{\boldsymbol{\Gamma}}$ is a consistent estimator of the covariance matrix in the Gaussian weak limit of $\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta}_0,\hat{\boldsymbol{\eta}},f}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,\boldsymbol{\eta},g}$ (from invariance, this covariance matrix again does not depend on g). Such tests are "somewhere semiparametrically efficient", in the sense that they achieve semiparametric efficiency at f. From invariance, they are (at least asymptotically) distribution-free with respect to g, which ensures validity under any g. For many problems, including the PCA testing problem considered in this paper, the maximal invariant is a vector of ranks (or, more precisely, of signed ranks), and the resulting invariant tests (that is, the resulting rank tests, or, more precisely, signed-rank tests), compared to pseudo-Gaussian tests, are valid under a broader class of nuisances g (see (b$_1$) in the introduction) and they also may be more powerful away from the target density f (see (c$_1$) in the introduction); in particular, for many problems, van der Waerden—that is, Gaussian score—rank tests are strictly more powerful than pseudo-Gaussian tests away from the Gaussian target densities.

The tests in (A.8) and (A.9) both require the estimation of the finite-dimensional nuisance parameter $\boldsymbol{\eta}$ to be feasible tests. It is therefore needed to control the replacement of the true unknown value of $\boldsymbol{\eta}$ by an estimator $\hat{\boldsymbol{\eta}}$, which may be challenging and is typically done by deriving a suitable *asymptotic linearity* result. In a triangular array framework, the situation is even more challenging since the value $\boldsymbol{\eta}_n$ of the finite-dimensional nuisance parameter to be estimated depends on the sample size n. To tackle this problem, we give here a second general result, that provides an asymptotic linearity result for triangular arrays. The result requires a reinforcement of the ULAN (Uniformly Locally Asymptotically Normal) structure. Like Proposition A.1, this new general result is of independent interest and may find applications outside the PCA problem we consider in this paper (see Section S.4 of the supplement for a proof).

PROPOSITION A.2. *Let $\mathbf{X}_{ni}$, $i=1,\ldots,n$, $n=1,2,\ldots$ be a triangular array of p-variate observations such that, for any n, $\mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}$ form a random sample from the distribution $\mathrm{P}_{\boldsymbol{\vartheta}_n,f}$. Denote as $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}$ the corresponding joint distribution of $\mathbf{X}_{n1},\ldots,\mathbf{X}_{nn}$. Assume that the sequence of models $\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta},f} : \boldsymbol{\vartheta}\in\boldsymbol{\Theta}\}$ is "super-ULAN" in the following sense: there exist a positive real sequence (ν_n) that is $o(1)$ and $k\times k$ matrices $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta},f}$ for any $\boldsymbol{\vartheta}\in\boldsymbol{\Theta}$ such that, for any $\boldsymbol{\vartheta}_0\in\boldsymbol{\Theta}$, any sequence $(\boldsymbol{\vartheta}_n)$ in $\boldsymbol{\Theta}$ converging to $\boldsymbol{\vartheta}_0$, and any bounded sequence $(\boldsymbol{\tau}_n)$ in $\mathbb{R}^k$, we have that, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}$,*

$$\log\frac{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n+\nu_n\boldsymbol{\tau}_n,f}}{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}} = \boldsymbol{\tau}_n'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}_n,f} - \frac{1}{2}\boldsymbol{\tau}_n'\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0,f}\boldsymbol{\tau}_n + o_{\mathrm{P}}(1) \tag{A.10}$$

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}_n,f} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0},\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0,f}), \tag{A.11}$$

where $\xrightarrow{\mathcal{L}}$ denotes weak convergence. Now, consider a sequence of random m-vectors

$$\mathbf{T}^{(n)}_{\boldsymbol{\vartheta}} := \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{T}_{\boldsymbol{\vartheta}}(\mathbf{X}_{ni}),$$

involving measurable functions $\mathbf{T}_{\boldsymbol{\vartheta}}:\mathbb{R}^p\to\mathbb{R}^m$ such that, for any $\boldsymbol{\vartheta}_0\in\boldsymbol{\Theta}$, any sequence $(\boldsymbol{\vartheta}_n)$ in $\boldsymbol{\Theta}$ converging to $\boldsymbol{\vartheta}_0$, and any bounded sequence $(\boldsymbol{\tau}_n)$ in $\mathbb{R}^k$, we have that, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}$,

(a) $\mathrm{E}[\mathbf{T}_{\boldsymbol{\vartheta}_n}(\mathbf{X}_{n1})]=\mathbf{0}$,
(b) $\mathrm{E}[\|\mathbf{T}_{\boldsymbol{\vartheta}_n+\nu_n\boldsymbol{\tau}_n}(\mathbf{X}_{n1})-\mathbf{T}_{\boldsymbol{\vartheta}_n}(\mathbf{X}_{n1})\|^2]=o(1)$, *and that*
(c) $((\mathbf{T}^{(n)}_{\boldsymbol{\vartheta}_n})',(\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}_n,f})')'$ *is asymptotically normal with mean zero and covariance matrix*

$$\begin{pmatrix}\boldsymbol{\Sigma}_{\boldsymbol{\vartheta}_0,f} & \mathbf{C}_{\boldsymbol{\vartheta}_0,f}\\ \mathbf{C}'_{\boldsymbol{\vartheta}_0,f} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0,f}\end{pmatrix}.$$

Then, for any $\boldsymbol{\vartheta}_0\in\boldsymbol{\Theta}$, any sequence $(\boldsymbol{\vartheta}_n)$ in $\boldsymbol{\Theta}$ converging to $\boldsymbol{\vartheta}_0$, and any bounded sequence $(\boldsymbol{\tau}_n)$ in $\mathbb{R}^k$,

$$\mathbf{T}^{(n)}_{\boldsymbol{\vartheta}_n+\nu_n\boldsymbol{\tau}_n} - \mathbf{T}^{(n)}_{\boldsymbol{\vartheta}_n} + \mathbf{C}_{\boldsymbol{\vartheta}_0,f}\boldsymbol{\tau}_n = o_{\mathrm{P}}(1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}$.

When testing $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ in a triangular array framework, this asymptotic linearity result will typically be applied to parameter sequences of the form $\boldsymbol{\vartheta}_n := (\boldsymbol{\theta}_0, \boldsymbol{\eta}_n)$, where $(\boldsymbol{\eta}_n)$ will converge to a fixed value $\boldsymbol{\eta}_0$. Note that although $\boldsymbol{\theta}_0$ will not depend on n below, our general results in Propositions A.1–A.2 above would also us to consider a sequence of null hypotheses associated with a sequence of null values $(\boldsymbol{\theta}_{0n})$ that would converge to $\boldsymbol{\theta}_0$. In this context, classical assumptions on estimators $\hat{\boldsymbol{\eta}}_n$ of $\boldsymbol{\eta}_n$ allow one to replace the deterministic local perturbation $\boldsymbol{\vartheta}_n + \nu_n \boldsymbol{\tau}_n = (\boldsymbol{\theta}_0, \boldsymbol{\eta}_n + \nu_n \mathbf{t}_n)$ by a ν_n^{-1}-consistent estimator $\hat{\boldsymbol{\vartheta}}_n = (\boldsymbol{\theta}_0, \hat{\boldsymbol{\eta}}_n)$; see, e.g., Lemma 4.4 in Kreiss (1987). It is important to note that, in Proposition A.2, $\boldsymbol{\vartheta}_n$ may converge to $\boldsymbol{\vartheta}$ at an arbitrary rate—when applying this result to the PCA context, this will allow us to consider arbitrarily weak identifiability. The price to pay to obtain the asymptotic linearity result in such a general asymptotic scheme is that we need to assume the "super-ULAN" structure in (A.10)–(A.11): having a fixed value $\boldsymbol{\vartheta}_0$ in (A.10)–(A.11) would correspond to LAN, whereas having (A.10)–(A.11) for any sequence $(\boldsymbol{\vartheta}_n)$ satisfying $\boldsymbol{\vartheta}_n = \boldsymbol{\vartheta}_0 + O(\nu_n)$ would correspond to ULAN (again, this super-ULAN allows for any sequence $(\boldsymbol{\vartheta}_n)$ such that $\boldsymbol{\vartheta}_n = \boldsymbol{\vartheta}_0 + o(1)$). In this regard, it may seem surprising at first that Proposition A.2 requires such a strong reinforcement of the LAN property given that, as already mentioned, even the usual LAN structure may fail under weak identifiability. In the PCA framework we considered in the paper, the way out is that while the LAN structure fails in the $(\boldsymbol{\theta}, \boldsymbol{\eta})$-parametrization that corresponds to the eigenvalues-eigenvectors parametrization in the PCA context, its super-ULAN reinforcement holds in the scatter/shape matrix parametrization[2], which will be enough for our purposes.

Acknowledgments. Davy Paindaveine is also affiliated at the Toulouse School of Economics, Université Toulouse 1 Capitole.

Funding. Davy Paindaveine and Thomas Verdebout are supported by two "Projet de Recherche" grants from the FNRS (Fonds National pour la Recherche Scientifique), Communauté Française de Belgique.

SUPPLEMENTARY MATERIAL

Supplement to "Rank-based inference for PCA under weak identifiability"
In the supplement, we prove all theoretical results of the present paper.

REFERENCES

ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34** 122–148.

BANERJEE, D. and MA, Z. (2022). Optimal signal detection in some spiked random matrix models: Likelihood ratio tests and linear spectral statistics. *Ann. Statist.* **50** 1910–1932.

BERNARD, G. and VERDEBOUT, T. (2024a). On testing the equality of latent roots of scatter matrices under ellipticity. *J. Multivariate Anal.* **199** 105232.

BERNARD, G. and VERDEBOUT, T. (2024b). Power enhancement for dimension detection of Gaussian signals. *Statist. Sinica* **34** 2161–2182.

CROUX, C. and HAESBROECK, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. *Biometrika* **87** 603–618.

DONOHO, D. L., GAVISH, M. and JOHNSTONE, I. M. (2018). Optimal shrinkage of eigenvalues in the spiked covariance model. *Ann. Statist.* **46** 1742–1778.

DÖRNEMANN, N. and DETTE, H. (2025). A CLT for the difference of eigenvalue statistics of sample covariance matrices. *Bernoulli,* to appear.

DÜMBGEN, L. (1995). Likelihood ratio tests for principal components. *J. Multivariate Anal.* **52** 245–258.

[2]This will actually be established in the supplement by applying again the general result in Proposition A.1.

FAN, J., FAN, Y., HAN, X. and LV, J. (2022). Asymptotic theory of eigenvectors for large random matrices. *J. Amer. Statist. Assoc.* **117** 996–1009.

FLURY, B. (1988). *Common Principal Components & Related Multivariate Models*. John Wiley & Sons, Inc., New York.

HALLIN, M. and PAINDAVEINE, D. (2004). Rank-based optimal tests of the adequacy of an elliptical VARMA model. *Ann. Statist.* **32** 2642–2678.

HALLIN, M. and PAINDAVEINE, D. (2006). Semiparametrically efficient rank-based inference for shape I: Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756.

HALLIN, M. and PAINDAVEINE, D. (2009). Optimal tests for homogeneity of covariance, scale, and shape. *J. Multivar. Anal.* **100** 422–444.

HALLIN, M., PAINDAVEINE, D. and VERDEBOUT, T. (2010a). Optimal rank-based testing for principal components. *Ann. Statist.* **38** 3245–3299.

HALLIN, M., PAINDAVEINE, D. and VERDEBOUT, T. (2010b). Testing for common principal components under heterokurticity. *J. Nonparametr. Stat.* **22** 879–895.

HALLIN, M., PAINDAVEINE, D. and VERDEBOUT, T. (2014). Efficient R-estimation of principal and common principal components. *J. Amer. Statist. Assoc.* **109** 1071–1083.

HALLIN, M. and WERKER, B. J. M. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9** 137–165.

HAN, F. and LIU, H. (2014). Scale-invariant sparse PCA on high-dimensional meta-elliptical data. *J. Amer. Statist. Assoc.* **109** 275–287.

HE, R., HU, B.-G., ZHENG, W.-S. and KONG, X.-W. (2011). Robust principal component analysis based on maximum correntropy criterion. *IEEE Trans. Signal Process.* **20** 1485–1494.

HETTMANSPERGER, T. P. and RANDLES, R. H. (2002). A practical affine equivariant multivariate median. *Biometrika* **89** 851–860.

HUBERT, M., ROUSSEEUW, P. J. and VANDEN BRANDEN, K. (2005). ROBPCA: a new approach to robust principal component analysis. *Technometrics* **47** 64–79.

ILMONEN, P. and PAINDAVEINE, D. (2011). Semiparametrically efficient inference based on signed ranks in symmetric independent component models. *Ann. Statist.* **39** 2448–2476.

JACKSON, J. E. (2005). *A User's Guide to Principal Components*. *Wiley Series in Probability and Statistics*. John Wiley & Sons, Hoboken, New Jersey.

JOHNSTONE, I. M. and LU, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *J. Amer. Statist. Assoc.* **104** 682–693.

JOLICOEUR, P. (1984). Principal components, factor analysis, and multivariate allometry: a small-sample direction test. *Biometrics* **40** 685–690.

KREISS, J. P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133.

MUIRHEAD, R. J. and WATERNAUX, C. M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika* **67** 31–43.

OJA, H. (2010). *Multivariate Nonparametric Methods with R. An Approach Based on Spatial Signs and Ranks*. Springer-Verlag, New York.

PAINDAVEINE, D. (2006). A Chernoff–Savage result for shape: On the non-admissibility of pseudo-Gaussian methods. *J. Multivariate Anal.* **97** 2206–2220.

PAINDAVEINE, D., PERALVO MAROTO, L. and VERDEBOUT, T. (2025). Supplement to "Rank tests for PCA under weak identifiability". Online supplement to the present submission.

PAINDAVEINE, D., REMY, J. and VERDEBOUT, T. (2020a). Testing for principal component directions under weak identifiability. *Ann. Statist.* **48** 324–345.

PAINDAVEINE, D., REMY, J. and VERDEBOUT, T. (2020b). Sign tests for weak principal directions. *Bernoulli* **26** 2987–3016.

SCHWARTZMAN, A., MASCARENHAS, W. F. and TAYLOR, J. E. (2008). Inference for eigenvalues and eigenvectors of Gaussian symmetric matrices. *Ann. Statist.* **36** 2886–2919.

SHAPIRO, A. and BROWNE, M. W. (1987). Analysis of covariance structures under elliptical distributions. *J. Amer. Statist. Assoc.* **82** 1092–1097.

SYLVESTER, A. D., KRAMER, P. A. and JUNGERS, W. L. (2008). Modern humans are not (quite) isometric. *Amer. J. Phys. Anthropol.* **137** 371–383.

TYLER, D. E. (1981). Asymptotic inference for eigenvectors. *Ann. Statist.* **9** 725–736.

TYLER, D. E. (1983). A class of asymptotic tests for principal component vectors. *Ann. Statist.* **11** 1243–1250.

TYLER, D. E. (1987). A distribution-free M-estimator of multivariate component vectors. *Ann. Statist.* **15** 1243–1250.

VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge Univ. Press, Cambridge.