The Annals of Statistics
2010, Vol. 38, No. 6, 3245–3299
DOI: 10.1214/10-AOS810

OPTIMAL RANK-BASED TESTING FOR PRINCIPAL COMPONENTS

BY MARC HALLIN[1,3,4], DAVY PAINDAVEINE[2,4] AND THOMAS VERDEBOUT

Université Libre de Bruxelles and Princeton University, Université Libre de Bruxelles and Université Lille Nord de France

This paper provides parametric and rank-based optimal tests for eigenvectors and eigenvalues of covariance or scatter matrices in elliptical families. The parametric tests extend the Gaussian likelihood ratio tests of Anderson (1963) and their pseudo-Gaussian robustifications by Davis (1977) and Tyler (1981, 1983). The rank-based tests address a much broader class of problems, where covariance matrices need not exist and principal components are associated with more general scatter matrices. The proposed tests are shown to outperform daily practice both from the point of view of validity as from the point of view of efficiency. This is achieved by utilizing the Le Cam theory of locally asymptotically normal experiments, in the nonstandard context, however, of a *curved* parametrization. The results we derive for curved experiments are of independent interest, and likely to apply in other contexts.

1. Introduction. This fairly detailed introduction aims at providing a comprehensive and nontechnical overview of the paper, including its asymptotic theory aspects, and a rough description of some of the rank-based test statistics to be derived. It is expected to be accessible to a broad readership. It should be sufficiently informative for the reader not interested in the technical aspects of asymptotic theory, to proceed to Sections 5 (Gaussian and pseudo-Gaussian tests) and 6 (rank-based tests), where the proposed testing procedures are described, and for the reader mainly interested in asymptotics, to decide whether he/she is interested in the treatment of a LAN family with curved parametrization developed in Sections 3 and 4.

Received April 2009; revised January 2010.

[1]Supported by the Sonderforschungsbereich "Statistical modelling of nonlinear dynamic processes" (SFB 823) of the Deutsche Forschungsgemeinschaft, and by a Discovery grant of the Australian Research Council.

[2]Supported by a contract of the National Bank of Belgium and a Mandat d'Impulsion Scientifique of the Fonds National de la Recherche Scientifique, Communauté française de Belgique.

[3]Member of Académie Royale de Belgique and extra-muros Fellow of CenTER, Tilburg University.

[4]Member of ECORE, the association between CORE and ECARES.

AMS 2000 subject classifications. Primary 62H25; secondary 62G35.

Key words and phrases. Principal components, tests for eigenvectors, tests for eigenvalues, elliptical densities, scatter matrix, shape matrix, multivariate ranks and signs, local asymptotic normality, curved experiments.

1.1. *Hypothesis testing for principal components.* Principal components are probably the most popular and widely used device in the traditional multivariate analysis toolkit. Introduced by Pearson (1901), principal component analysis (PCA) was rediscovered by Hotelling (1933), and ever since has been an essential part of daily statistical practice, basically in all domains of application.

The general objective of PCA is to reduce the dimension of some observed k-dimensional random vector $\mathbf{X}$ while preserving most of its total variability. This is achieved by considering an adequate number q of linear combinations of the form $\boldsymbol{\beta}_1'\mathbf{X}, \ldots, \boldsymbol{\beta}_q'\mathbf{X}$, where $\boldsymbol{\beta}_j$, $j = 1, \ldots, k$, are the eigenvectors associated with the eigenvalues $\lambda_1, \ldots, \lambda_k$ of $\mathbf{X}$'s covariance matrix $\boldsymbol{\Sigma}_{\text{cov}}$, ranked in decreasing order of magnitude. Writing $\boldsymbol{\beta}$ for the orthogonal $k \times k$ matrix with columns $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k$ and $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}_{\text{cov}}}$ for the diagonal matrix of eigenvalues $\lambda_1, \ldots, \lambda_k$, the matrix $\boldsymbol{\Sigma}_{\text{cov}}$ thus factorizes into $\boldsymbol{\Sigma}_{\text{cov}} = \boldsymbol{\beta}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}_{\text{cov}}}\boldsymbol{\beta}'$. The random variable $\boldsymbol{\beta}_j'\mathbf{X}$, with variance λ_j, is known as $\mathbf{X}$'s *jth principal component*.

Chapters on inference for eigenvectors and eigenvalues can be found in most textbooks on multivariate analysis, and mainly cover Gaussian maximum likelihood estimation (MLE) and the corresponding Wald and Gaussian likelihood ratio tests (LRT). The MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}_{\text{cov}}}$ are the eigenvectors and eigenvalues of the empirical covariance matrix

$$\mathbf{S}^{(n)} := \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}}^{(n)})(\mathbf{X}_i - \bar{\mathbf{X}}^{(n)})' \qquad \text{with } \bar{\mathbf{X}}^{(n)} := \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i,$$

while testing problems classically include testing for sphericity (equality of eigenvalues), testing for *subsphericity* (equality among some given subset of eigenvalues—typically, the last $k-q$ ones), testing that the ℓth eigenvector has some specified direction, or that the proportion of variance accounted for by the last $k-q$ principal components is larger than some fixed proportion of the total variance: see, for instance, Anderson (2003) or Jolliffe (1986).

Gaussian MLEs and the corresponding tests (Wald or likelihood ratio tests—since they are asymptotically equivalent, in the sequel we indistinctly refer to LRTs) for covariance matrices and functions thereof are notoriously sensitive to violations of Gaussian assumptions; see Muirhead and Waternaux (1980) for a classical discussion of this fact, or Yanagihara, Tonda and Matsumoto (2005) for a more recent overview. The problems just mentioned about the eigenvectors and eigenvalues of $\boldsymbol{\Sigma}_{\text{cov}}$ are no exception to that rule, although belonging, in Muirhead and Waternaux's terminology, to the class of "easily robustifiable" ones. For such problems, *adjusted* LRTs remaining valid under the whole class of elliptical distributions with finite fourth-order moments can be obtained via a correction factor involving estimated kurtosis coefficients [see Shapiro and Browne (1987) for a general result on the "easy" cases, and Hallin and Paindaveine (2008b) for the "harder" ones]. Such adjusted LRTs were obtained by Tyler (1981, 1983) for eigenvector problems and by Davis (1977) for eigenvalues.

Tyler actually constructs tests for the *scatter matrix* $\mathbf{\Sigma}$ characterizing the density contours [of the form $(\mathbf{x}-\boldsymbol{\theta})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\theta}) = \text{constant}$] of an elliptical family. His tests are the Wald tests associated with any available estimator $\hat{\mathbf{\Sigma}}$ of $\mathbf{\Sigma}$ such that $n^{1/2}\,\text{vec}(\hat{\mathbf{\Sigma}}-\mathbf{\Sigma})$ is asymptotically normal, with mean zero and covariance matrix $\mathbf{\Psi}_f$, say, under $f \in \mathcal{F}$, where $\mathcal{F}$ denotes some class of elliptical densities and $\mathbf{\Psi}_f$ either is known or (still, under $f \in \mathcal{F}$) can be estimated consistently. The resulting tests then are valid under the class $\mathcal{F}$. When the estimator $\hat{\mathbf{\Sigma}}$ is the empirical covariance matrix $\mathbf{S}^{(n)}$, these tests under Gaussian densities are asymptotically equivalent to Gaussian LRTs. Unlike the latter, however, they remain (asymptotically) valid under the class $\mathcal{F}^4$ of all elliptical distributions with finite moments of order four, and hence qualify as *pseudo-Gaussian* versions of the Gaussian LRTs.

Due to their importance for applications, throughout this paper, we concentrate on the following two problems:

(a) testing the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ that the first principal direction $\boldsymbol{\beta}_1$ coincides (up to the sign) with some specified unit vector $\boldsymbol{\beta}^0$ (the choice of the *first* principal direction here is completely arbitrary, and made for the simplicity of exposition only), and

(b) testing the null hypothesis $\mathcal{H}_0^{\mathbf{\Lambda}}$ that $\sum_{j=q+1}^k \lambda_j / \sum_{j=1}^k \lambda_j = p$ against the one-sided alternative under which $\sum_{j=q+1}^k \lambda_j / \sum_{j=1}^k \lambda_j < p$, $p \in (0,1)$ given.

The Gaussian LRT for (a) was introduced in a seminal paper by Anderson (1963). Denoting by $\lambda_{j;\mathbf{S}}$ and $\boldsymbol{\beta}_{j;\mathbf{S}}$, $j = 1, \ldots, k$, respectively, the eigenvalues and eigenvectors of $\mathbf{S}^{(n)}$, this test—denote it by $\phi^{(n)}_{\boldsymbol{\beta};\text{Anderson}}$—rejects $\mathcal{H}_0^{\boldsymbol{\beta}}$ (at asymptotic level α) as soon as

$$Q^{(n)}_{\text{Anderson}} := n[\lambda_{1;\mathbf{S}}\boldsymbol{\beta}^{0\prime}(\mathbf{S}^{(n)})^{-1}\boldsymbol{\beta}^0 + \lambda_{1;\mathbf{S}}^{-1}\boldsymbol{\beta}^{0\prime}\mathbf{S}^{(n)}\boldsymbol{\beta}^0 - 2]$$
$$= \frac{n}{\lambda_{1;\mathbf{S}}}\sum_{j=2}^k \frac{(\lambda_{j;\mathbf{S}}-\lambda_{1;\mathbf{S}})^2}{\lambda_{j;\mathbf{S}}^3}(\boldsymbol{\beta}'_{j;\mathbf{S}}\mathbf{S}^{(n)}\boldsymbol{\beta}^0)^2 \tag{1.1}$$

exceeds the α upper-quantile of the chi-square distribution with $(k-1)$ degrees of freedom. The behavior of this test being particularly poor under non-Gaussian densities, Tyler (1981, 1983) proposed a pseudo-Gaussian version $\phi^{(n)}_{\boldsymbol{\beta};\text{Tyler}}$, which he obtains via an empirical kurtosis correction

$$Q^{(n)}_{\text{Tyler}} := (1+\hat{\kappa}^{(n)})^{-1} Q^{(n)}_{\text{Anderson}} \tag{1.2}$$

of (1.1) (same asymptotic distribution), where $\hat{\kappa}^{(n)}$ is some consistent estimator of the underlying kurtosis parameter κ_k; see Section 5.3 for a definition.

A related test of Schott (1991) addresses the same problem where however $\boldsymbol{\beta}_1$ is the first eigenvector of the *correlation* matrix.

The traditional Gaussian test for problem (b) was introduced in the same paper by Anderson (1963). For any $k \times k$ diagonal matrix $\mathbf{\Lambda}$ with diagonal entries

$\lambda_1, \ldots, \lambda_k$, let $a_{p,q}(\mathbf{\Lambda}) := 2(p^2 \sum_{j=1}^q \lambda_j^2 + (1-p)^2 \sum_{j=q+1}^k \lambda_j^2)$. Defining $\boldsymbol{\beta}_{\mathbf{S}} := (\boldsymbol{\beta}_{1;\mathbf{S}}, \ldots, \boldsymbol{\beta}_{k;\mathbf{S}})$ and $\mathbf{c}_{p,q} := (-p\mathbf{1}_q' \vdots (1-p)\mathbf{1}_{k-q}')'$, with $\mathbf{1}_\ell := (1, \ldots, 1)' \in \mathbb{R}^\ell$, and denoting by $\mathrm{dvec}(\mathbf{A})$ the vector obtained by stacking the diagonal elements of a square matrix $\mathbf{A}$, Anderson's test, $\phi_{\mathbf{\Lambda};\mathrm{Anderson}}^{(n)}$, say, rejects the null hypothesis at asymptotic level α whenever

$$
\begin{aligned}
T_{\mathrm{Anderson}}^{(n)} &:= n^{1/2}(a_{p,q}(\mathbf{\Lambda}_{\mathbf{S}}))^{-1/2}\mathbf{c}_{p,q}' \,\mathrm{dvec}(\boldsymbol{\beta}_{\mathbf{S}}' \mathbf{S}^{(n)} \boldsymbol{\beta}_{\mathbf{S}}) \\
&= n^{1/2}(a_{p,q}(\mathbf{\Lambda}_{\mathbf{S}}))^{-1/2}\left((1-p)\sum_{j=q+1}^{k} \lambda_{j;\mathbf{S}} - p\sum_{j=1}^{q} \lambda_{j;\mathbf{S}}\right)
\end{aligned}
\tag{1.3}
$$

is less than the standard normal α-quantile. Although he does not provide any explicit form, Davis (1977) briefly explains how to derive the pseudo-Gaussian version

$$
T_{\mathrm{Davis}}^{(n)} := \left(1 + \hat{\kappa}^{(n)}\right)^{-1/2} T_{\mathrm{Anderson}}^{(n)}
\tag{1.4}
$$

of (1.3), where $\hat{\kappa}^{(n)}$ again is any consistent estimator of the underlying kurtosis parameter κ_k. The resulting test (same asymptotic standard normal distribution) will be denoted as $\phi_{\mathbf{\Lambda};\mathrm{Davis}}^{(n)}$.

Being based on empirical covariances, though, the pseudo-Gaussian tests based on (1.2) and (1.4) unfortunately remain poorly robust. They still are very sensitive to the presence of outliers—an issue which we do not touch here; see, for example, Croux and Haesbroeck (2000), Salibián-Barrera, Van Aelst and Willems (2006), and the references therein. Moreover, they do require finite moments of order four—hence lose their validity under heavy tails, and only address the traditional covariance-based concept of principal components.

This limitation is quite regrettable, as principal components, irrespective of any moment conditions, clearly depend on the elliptical geometry of underlying distributions only. Recall that an elliptical density over $\mathbb{R}^k$ is determined by a *location vector* $\boldsymbol{\theta} \in \mathbb{R}^k$, a *scale* parameter $\sigma \in \mathbb{R}_0^+$ (where σ^2 is not necessarily a variance), a real-valued $k \times k$ symmetric and positive definite matrix $\mathbf{V}$ called the *shape* matrix, and a *standardized radial density* f_1 (whenever the elliptical density has finite second-order moments, the shape and covariance matrices $\mathbf{V}$ and $\mathbf{\Sigma}_{\mathrm{cov}}$ are proportional, hence share the same collection of eigenvectors and, up to a positive factor, the same collection of eigenvalues). Although traditionally described in terms of the covariance matrix $\mathbf{\Sigma}_{\mathrm{cov}}$, most inference problems in multivariate analysis naturally extend to arbitrary elliptical models, with the shape matrix $\mathbf{V}$ or the *scatter matrix* $\mathbf{\Sigma} := \sigma^2\mathbf{V}$ playing the role of $\mathbf{\Sigma}_{\mathrm{cov}}$. Principal components are no exception; in particular, problems (a) and (b) indifferently can be formulated in terms of shape or covariance eigenvectors and eigenvalues. Below, $\mathbf{\Lambda}_{\mathbf{V}} := \mathrm{diag}(\lambda_{1;\mathbf{V}}, \ldots, \lambda_{k;\mathbf{V}})$ and $\boldsymbol{\beta} := (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k)$ collect the eigenvalues and eigenvectors of the shape matrix $\mathbf{V}$.

Our objective in this paper is to provide a class of signed-rank tests which remain valid under arbitrary elliptical densities, in the absence of *any* moment assumption, and hence are not limited to the traditional covariance-based concept of principal components. Of particular interest within that class are the *van der Waerden*—that is, *normal-score*—tests, which are asymptotically equivalent, under Gaussian densities, to the corresponding Gaussian LRTs (the asymptotic optimality of which we moreover establish in Section 5, along with local powers). Under non-Gaussian conditions, however, these van der Waerden tests uniformly dominate, in the Pitman sense, the pseudo-Gaussian tests based on (1.2) and (1.4) above, which, as a result, turn out to be nonadmissible (see Section 7).

Our tests are based on the multivariate signs and ranks previously considered by Hallin and Paindaveine (2006a, 2008a) and Hallin, Oja and Paindaveine (2006). Denote by $\mathbf{X}_1, \dots, \mathbf{X}_n$ an observed n-tuple of k-dimensional elliptical vectors with location $\boldsymbol{\theta}$ and shape $\mathbf{V}$. Let $\mathbf{Z}_i := \mathbf{V}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})$ denote the *sphericized* version of $\mathbf{X}_i$ (throughout $\mathbf{A}^{1/2}$, for a symmetric and positive definite matrix $\mathbf{A}$, stands for the symmetric and positive definite root of $\mathbf{A}$): the corresponding multivariate signs are defined as the unit vectors $\mathbf{U}_i = \mathbf{U}_i(\boldsymbol{\theta}, \mathbf{V}) := \mathbf{Z}_i/\|\mathbf{Z}_i\|$, while the ranks $R_i^{(n)} = R_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})$ are those of the norms $\|\mathbf{Z}_i\|$, $i = 1, \dots, n$. Our rank tests are based on signed-rank covariance matrices of the form

$$\underset{\sim}{\mathbf{S}}_K^{(n)} := \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{R_i^{(n)}}{n+1}\right) \mathbf{U}_i \mathbf{U}_i',$$

where $K : (0, 1) \to \mathbb{R}$ stands for some *score function*, and $\mathbf{U}_i = \mathbf{U}_i(\hat{\boldsymbol{\theta}}, \hat{\mathbf{V}})$ and $R_i^{(n)} = R_i^{(n)}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{V}})$ are computed from appropriate estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{V}}$ of $\boldsymbol{\theta}$ and $\mathbf{V}$. More precisely, for the testing problem (a), the rank-based test $\underset{\sim}{\phi}_{\boldsymbol{\beta};K}^{(n)}$ rejects the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ (at asymptotic level α) whenever

$$\underset{\sim}{Q}_K^{(n)} := \frac{nk(k+2)}{\mathcal{J}_k(K)} \sum_{j=2}^{k} (\tilde{\boldsymbol{\beta}}_j' \underset{\sim}{\mathbf{S}}_K^{(n)} \boldsymbol{\beta}^0)^2$$

exceeds the α upper-quantile of the chi-square distribution with $(k - 1)$ degrees of freedom; here, $\mathcal{J}_k(K)$ is a standardizing constant and $\tilde{\boldsymbol{\beta}}_j$ stands for a constrained estimator of $\mathbf{V}$'s jth eigenvector; see (5.2) for details. As for problem (b), our rank tests $\underset{\sim}{\phi}_{\boldsymbol{\Lambda};K}^{(n)}$ are based on statistics of the form

$$\underset{\sim}{T}_K^{(n)} := \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2} (a_{p,q}(\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}))^{-1/2} \mathbf{c}_{p,q}' \operatorname{dvec}(\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}^{1/2} \hat{\boldsymbol{\beta}}' \underset{\sim}{\mathbf{S}}_K^{(n)} \hat{\boldsymbol{\beta}} \tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}^{1/2}),$$

where $\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}$ and $\hat{\boldsymbol{\beta}}$ are adequate estimators of $\boldsymbol{\Lambda}_{\mathbf{V}}$ and $\boldsymbol{\beta}$, respectively. The null hypothesis $\mathcal{H}_0^{\boldsymbol{\Lambda}}$ is to be rejected at asymptotic level α whenever $\underset{\sim}{T}_K^{(n)}$ is smaller than the standard normal α-quantile.

These tests are not just validity-robust, they also are efficient. For any smooth radial density f_1, indeed, the score function $K = K_{f_1}$ (see Section 2.2) provides a signed-rank test which is *locally and asymptotically optimal* (*locally and asymptotically most stringent*, in the Le Cam sense) under radial density f_1. In particular, when based on *normal or van der Waerden* scores $K = K_{\phi_1} := \Psi_k^{-1}$, where Ψ_k denotes the chi-square distribution function with k degrees of freedom, our rank tests achieve the same asymptotic performances as the optimal Gaussian ones at the multinormal, while enjoying maximal validity robustness, since no assumption is required on the underlying density beyond ellipticity. Moreover, the asymptotic relative efficiencies (AREs) under non-Gaussian densities of these van der Waerden tests are uniformly larger than one with respect to their pseudo-Gaussian parametric competitors; see Section 7. On all counts, validity, robustness, and efficiency, our van der Waerden tests thus perform uniformly better than the daily practice Anderson tests and their pseudo-Gaussian extensions.

1.2. *Local asymptotic normality for principal components*. The methodological tool we are using throughout is Le Cam's theory of *locally asymptotically normal* (LAN) *experiments* [for background reading on LAN, we refer to Le Cam (1986), Le Cam and Yang (2000) or van der Vaart (1998); see also Strasser (1985) or Rieder (1994)]. Although this powerful method has been used quite successfully in inference problems for elliptical families [Hallin and Paindaveine (2002, 2004, 2005, 2006a), Hallin, Oja and Paindaveine (2006) and Hallin and Paindaveine (2008a) for location, VARMA dependence, linear models, shape and scatter, resp.], it has not been considered so far in problems involving eigenvectors and eigenvalues, and, as a result, little is known about optimality issues in that context. The main reason, probably, is that the eigenvectors $\boldsymbol{\beta}$ and eigenvalues $\boldsymbol{\Lambda}$ are complicated functions of the covariance or scatter matrix $\boldsymbol{\Sigma}$, with unpleasant identification problems at possibly multiple eigenvalues. These special features of eigenvectors and eigenvalues, as we shall see, make the LAN approach more involved than in standard cases.

LAN (actually, ULAN) has been established, under appropriate regularity assumptions on radial densities, in Hallin and Paindaveine (2006a), for elliptical families when parametrized by a location vector $\boldsymbol{\theta}$ and a scatter matrix $\boldsymbol{\Sigma}$ [more precisely, the vector $\operatorname{vech}(\boldsymbol{\Sigma})$ resulting from stacking the upper diagonal elements of $\boldsymbol{\Sigma}$]. Recall, however, that LAN or ULAN are properties of the parametrization of a family of distributions, not of the family itself. Now, due to the complicated relation between $(\boldsymbol{\theta}, \operatorname{vech}\boldsymbol{\Sigma})$ and the quantities of interest $\boldsymbol{\Lambda}$ and $\boldsymbol{\beta}$, the $(\boldsymbol{\theta}, \operatorname{vech}(\boldsymbol{\Sigma}))$-parametrization is not convenient in the present context. Another parametrization, involving location, scale, and shape eigenvalues and eigenvectors is much preferable, as the hypotheses to be tested then take simple forms. Therefore, we show (Lemma A.1) how the ULAN result of Hallin and Paindaveine (2006a) carries over to this new parametrization where, moreover, the information matrix, very conveniently, happens to be block-diagonal—a structure that greatly simplifies inference in the presence of nuisance parameters. Unfortunately, this new parametrization, where $\boldsymbol{\beta}$ ranges over the set $\mathcal{S}O_k$ of $k \times k$ real orthogonal matrices with determinant one, raises problems of another nature. The subparameter $\mathrm{vec}(\boldsymbol{\beta})$ indeed ranges over $\mathrm{vec}(\mathcal{S}O_k)$, a nonlinear manifold of $\mathbb{R}^{k^2}$, yielding a *curved* ULAN experiment. By a *curved experiment*, we mean a parametric model indexed by a ℓ-dimensional parameter ranging over some nonlinear manifold of $\mathbb{R}^\ell$, such as in *curved* exponential families, for instance. Under a $\mathrm{vec}(\boldsymbol{\beta})$-parametrization, the local experiments are not the traditional *Gaussian shifts* anymore, but *curved* Gaussian location ones, that is, Gaussian location models under which the mean of a multinormal observation with specified covariance structure ranges over a nonlinear manifold of $\mathbb{R}^\ell$, so that the simple local asymptotic optimality results associated with local Gaussian shifts no longer hold. To the best of our knowledge, such experiments never have been considered in the LAN literature.

A third parametrization, however, can be constructed from the fact that $\boldsymbol{\beta}$ is in $\mathcal{S}O_k$ if it can be expressed as the exponential of a $k \times k$ skew-symmetric matrix $\boldsymbol{\iota}$. Denoting by $\mathrm{vech}^+(\boldsymbol{\iota})$ the vector resulting from stacking the upper off-diagonal elements of $\boldsymbol{\iota}$, this yields a parametrization involving location, scale, shape eigenvalues and $\mathrm{vech}^+(\boldsymbol{\iota})$; the latter subparameter ranges freely over $\mathbb{R}^{k(k-1)/2}$, yielding a well-behaved ULAN parametrization where local experiments converge to the classical Gaussian shifts, thereby allowing for the classical construction [Le Cam (1986), Section 11.9] of locally asymptotically optimal tests. The trouble is that translating null hypotheses (a) and (b) into the $\boldsymbol{\iota}$-space in practice seems unfeasible.

Three distinct ULAN structures are thus coexisting on the same families of distributions:

(ULAN1) proved in Hallin and Paindaveine (2006a) for the $(\boldsymbol{\theta}, \mathrm{vech}(\boldsymbol{\Sigma}))$-parametrization, serving as the mother of all subsequent ones;

(ULAN2) for the location-scale-eigenvalues–eigenvectors parametrization, where the null hypotheses of interest take simple forms, but the local experiments happen to be *curved* ones;

(ULAN3) for the location-scale-eigenvalues–skew symmetric matrix parametrization, where everything is fine from a decision-theoretical point of view, with, however, the major inconvenience that explicit solutions cannot be obtained in terms of original parameters.

The main challenge of this paper was the delicate interplay between these three structures. Basically, we are showing (Lemma A.1) how ULAN can be imported from the first parametrization, and (Section 3.3) optimality results from the third parametrization, both to the second one. These results then are used in order to derive locally asymptotically optimal Gaussian, pseudo-Gaussian and rank-based tests for eigenvectors and eigenvalues of shape. This treatment we are giving of curved ULAN experiments, to the best of our knowledge, is original, and likely to apply in a variety of other contexts.

1.3. *Outline of the paper.* Section 2 contains, for easy reference, some basic notation and fundamental assumptions to be used later on. The main ULAN result, of a nonstandard curved nature, is established in Section 3, and its consequences for testing developed in Section 4. As explained in the Introduction, optimality is imported from an untractable parametrization involving skew-symmetric matrices. This is elaborated, in some detail, in Section 3.3, where a general result is derived, and in (Section 4.1), where that result is applied to the particular case of eigenvectors and (Section 4.2) eigenvalues of shape, under arbitrary radial density f_1. Special attention is given, in Sections 5.1 and 5.2, to the Gaussian case ($f_1 = \phi_1$); in Sections 5.3 and 5.4, those Gaussian tests are extended to a pseudo-Gaussian context with finite fourth-order moments. Then, in Section 6, rank-based procedures, which do not require any moment assumptions, are constructed: Section 6.1 provides a general asymptotic representation result [Proposition 6.1(i)] in the Hájek style; asymptotic normality, under the null as well as under local alternatives, follows as a corollary [Proposition 6.1(ii)]. Based on these results, Sections 6.2 and 6.3 provide optimal rank-based tests for the eigenvector and eigenvalue problems considered throughout; Sections 7 and 8 conclude with asymptotic relative efficiencies and simulations. Technical proofs are concentrated in the Appendix.

The reader interested in inferential results and principal components only (the form of the tests, their optimality properties and local powers) may skip Sections 3 and 4, which are devoted to curved LAN experiments, and concentrate on Section 5 for the "parametric" procedures, on Section 6 for the rank-based ones, on Sections 7 and 8 for their asymptotic and finite-sample performances.

1.4. *Notation.* The following notation will be used throughout. For any $k \times k$ matrix $\mathbf{A} = (A_{ij})$, write $\mathrm{vec}(\mathbf{A})$ for the k^2-dimensional vector obtained by stacking the columns of $\mathbf{A}$, $\mathrm{vech}(\mathbf{A})$ for the $[k(k+1)/2]$-dimensional vector obtained by stacking the upper diagonal elements of those columns, $\mathrm{vech}^+(\mathbf{A})$ for the $[k(k-1)/2]$-dimensional vector obtained by stacking the upper off-diagonal elements of the same, and $\mathrm{dvec}(\mathbf{A}) =: (A_{11}, (\mathrm{d\mathring{v}ec}(\mathbf{A}))')'$ for the k-dimensional vector obtained by stacking the diagonal elements of $\mathbf{A}$; $\mathrm{d\mathring{v}ec}(\mathbf{A})$ thus is $\mathrm{dvec}(\mathbf{A})$ deprived of its first component. Let $\mathbf{H}_k$ be the $k \times k^2$ matrix such that $\mathbf{H}_k \,\mathrm{vec}(\mathbf{A}) = \mathrm{dvec}(\mathbf{A})$. Note that we then have that $\mathbf{H}_k' \,\mathrm{dvec}(\mathbf{A}) = \mathrm{vec}(\mathbf{A})$ for any $k \times k$ diagonal matrix $\mathbf{A}$, which implies that $\mathbf{H}_k \mathbf{H}_k' = \mathbf{I}_k$. Write $\mathrm{diag}(\mathbf{B}_1, \ldots, \mathbf{B}_m)$ for the block-diagonal matrix with blocks $\mathbf{B}_1, \ldots, \mathbf{B}_m$ and $\mathbf{A}^{\otimes 2}$ for the Kronecker product $\mathbf{A} \otimes \mathbf{A}$. Finally, denoting by $\mathbf{e}_\ell$ the ℓth vector in the canonical basis of $\mathbb{R}^k$, write $\mathbf{K}_k := \sum_{i,j=1}^k (\mathbf{e}_i \mathbf{e}_j') \otimes (\mathbf{e}_j \mathbf{e}_i')$ for the $k^2 \times k^2$ *commutation matrix*.

2. Main assumptions.

2.1. *Elliptical densities.* We throughout assume that the observations are elliptically symmetric. More precisely, defining

$$\mathcal{F} := \{h : \mathbb{R}_0^+ \to \mathbb{R}^+ : \mu_{k-1;h} < \infty\},$$

where $\mu_{\ell;h} := \int_0^\infty r^\ell h(r)\,dr$, and

$$\mathcal{F}_1 := \left\{ h_1 \in \mathcal{F} : (\mu_{k-1;h_1})^{-1} \int_0^1 r^{k-1} h_1(r)\,dr = 1/2 \right\},$$

we denote by $\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}$ an observed n-tuple of mutually independent k-dimensional random vectors with probability density function of the form

$$f(\mathbf{x}) := c_{k,f_1} |\boldsymbol{\Sigma}|^{-1/2} f_1\big(((\mathbf{x}-\boldsymbol{\theta})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\theta}))^{1/2}\big), \qquad \mathbf{x} \in \mathbb{R}^k, \tag{2.1}$$

for some k-dimensional vector $\boldsymbol{\theta}$ (*location*), some symmetric and positive definite $(k \times k)$ *scatter* matrix $\boldsymbol{\Sigma}$, and some f_1 in the class $\mathcal{F}_1$ of *standardized radial densities*; throughout, $|\mathbf{A}|$ stands for the determinant of the square matrix $\mathbf{A}$.

Define the *elliptical coordinates* of $\mathbf{X}_i^{(n)}$ as

$$\begin{aligned} \mathbf{U}_i^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) &:= \frac{\boldsymbol{\Sigma}^{-1/2}(\mathbf{X}_i^{(n)} - \boldsymbol{\theta})}{\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{X}_i^{(n)} - \boldsymbol{\theta})\|} \quad \text{and} \\ d_i^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) &:= \|\boldsymbol{\Sigma}^{-1/2}(\mathbf{X}_i^{(n)} - \boldsymbol{\theta})\|. \end{aligned} \tag{2.2}$$

Under the assumption of ellipticity, the *multivariate signs* $\mathbf{U}_i^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, $i = 1, \ldots, n$, are i.i.d. uniform over the unit sphere in $\mathbb{R}^k$, and independent of the *standardized elliptical distances* $d_i^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$. Imposing that $f_1 \in \mathcal{F}_1$ implies that the $d_i^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$'s, which have common density $\tilde{f}_{1k}(r) := (\mu_{k-1;f_1})^{-1} r^{k-1} f_1(r) I_{[r>0]}$, with distribution function $\tilde{F}_{1k}$, have median one [$\tilde{F}_{1k}(1) = 1/2$]—a constraint which identifies $\boldsymbol{\Sigma}$ without requiring any moment assumptions [see Hallin and Paindaveine (2006a) for a discussion]. Under finite second-order moments, the scatter matrix $\boldsymbol{\Sigma}$ is proportional to the traditional covariance matrix $\boldsymbol{\Sigma}_{\text{cov}}$.

Special instances are the k-variate multinormal distribution, with radial density $f_1(r) = \phi_1(r) := \exp(-a_k r^2/2)$, the k-variate Student distributions, with radial densities (for $\nu \in \mathbb{R}_0^+$ degrees of freedom) $f_1(r) = f_{1,\nu}^t(r) := (1 + a_{k,\nu} r^2/\nu)^{-(k+\nu)/2}$, and the k-variate power-exponential distributions, with radial densities of the form $f_1(r) = f_{1,\eta}^e(r) := \exp(-b_{k,\eta} r^{2\eta})$, $\eta \in \mathbb{R}_0^+$; the positive constants a_k, $a_{k,\nu}$, and $b_{k,\eta}$ are such that $f_1 \in \mathcal{F}_1$.

The derivation of locally and asymptotically optimal tests at standardized radial density f_1 will be based on the *uniform local and asymptotic normality* (ULAN) of the model *at given* f_1. This ULAN property—the statement of which requires some further preparation and is delayed to Section 3—only holds under some further mild regularity conditions on f_1. More precisely, we require f_1 to belong to the collection $\mathcal{F}_a$ of all absolutely continuous densities in $\mathcal{F}_1$ for which, denoting by $\dot{f}_1$ the a.e. derivative of f_1 and letting $\varphi_{f_1} := -\dot{f}_1/f_1$, the integrals

$$\mathcal{I}_k(f_1) := \int_0^1 \varphi_{f_1}^2(r) \tilde{f}_{1k}(r)\,dr \quad \text{and} \quad \mathcal{J}_k(f_1) := \int_0^1 r^2 \varphi_{f_1}^2(r) \tilde{f}_{1k}(r)\,dr \tag{2.3}$$

are finite. The quantities $\mathcal{I}_k(f_1)$ and $\mathcal{J}_k(f_1)$ play the roles of *radial Fisher information for location* and *radial Fisher information for shape/scale*, respectively. Slightly less stringent assumptions, involving derivatives in the sense of distributions, can be found in Hallin and Paindaveine (2006a), where we refer to for details. The intersection of $\mathcal{F}_a$ and $\mathcal{F}_1^4 := \{f_1 \in \mathcal{F}_1 : \int_0^\infty r^4 \tilde{f}_{1k}(r)\, dr < \infty\}$ will be denoted as $\mathcal{F}_a^4$.

2.2. *Score functions*. The various *score functions* K appearing in the rank-based statistics to be introduced in Section 6 will be assumed to satisfy a few regularity assumptions which we are listing here for convenience.

ASSUMPTION (S). The score function $K : (0, 1) \to \mathbb{R}$ (S1) is continuous and square-integrable, (S2) can be expressed as the difference of two monotone increasing functions, and (S3) satisfies $\int_0^1 K(u)\, du = k$.

Assumption (S3) is a normalization constraint that is automatically satisfied by the score functions $K(u) = K_{f_1}(u) := \varphi_{f_1}(\tilde{F}_{1k}^{-1}(u))\tilde{F}_{1k}^{-1}(u)$ associated with any radial density $f_1 \in \mathcal{F}_a$ (at which ULAN holds); see Section 3. For score functions K, K_1, K_2 satisfying Assumption (S), let [throughout, U stands for a random variable uniformly distributed over (0, 1)]

$$\mathcal{J}_k(K_1, K_2) := \mathrm{E}[K_1(U)K_2(U)], \qquad \mathcal{J}_k(K) := \mathcal{J}_k(K, K) \tag{2.4}$$

and

$$\mathcal{J}_k(K, f_1) := \mathcal{J}_k(K, K_{f_1}); \tag{2.5}$$

with this notation, $\mathcal{J}_k(f_1) = \mathcal{J}_k(K_{f_1}, K_{f_1})$.

The *power* score functions $K_a(u) := k(a+1)u^a$ ($a \geq 0$), with $\mathcal{J}_k(K_a) = k^2(a+1)^2/(2a+1)$, provide some traditional score functions satisfying Assumption (S): the sign, Wilcoxon, and Spearman scores are obtained for $a = 0$, $a = 1$ and $a = 2$, respectively. As for the score functions of the form K_{f_1}, an important particular case is that of van der Waerden or *normal* scores, obtained for $f_1 = \phi_1$. Then

$$K_{\phi_1}(u) = \Psi_k^{-1}(u) \quad \text{and} \quad \mathcal{J}_k(\phi_1) = k(k+2), \tag{2.6}$$

where Ψ_k was defined in page 3250. Similarly, Student densities $f_1 = f_{1,\nu}^t$ (with ν degrees of freedom) yield the scores

$$K_{f_{1,\nu}^t}(u) = \frac{k(k+\nu)G_{k,\nu}^{-1}(u)}{\nu + kG_{k,\nu}^{-1}(u)}$$

and

$$\mathcal{J}_k(f_{1,\nu}^t) = \frac{k(k+2)(k+\nu)}{k+\nu+2},$$

where $G_{k,\nu}$ stands for the Fisher–Snedecor distribution function with k and ν degrees of freedom.

3. Uniform local asymptotic normality (ULAN) and curved Gaussian location local experiments.

3.1. *Semiparametric modeling of elliptical families.* Consider an i.i.d. n-tuple $\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}$ with elliptical density (2.1) characterized by $\boldsymbol{\theta}$, $\boldsymbol{\Sigma}$, and $f_1 : (\boldsymbol{\theta}, \boldsymbol{\Sigma})$ or, if a vector is to be preferred, $(\boldsymbol{\theta}', (\mathrm{vech}\,\boldsymbol{\Sigma})')'$, provides a perfectly valid parametrization of the elliptical family with standardized radial density f_1. However, in the problems we are considering in this paper, it will be convenient to have eigenvalues and eigenvectors appearing explicitly in the vector of parameters. Decompose therefore the scatter matrix $\boldsymbol{\Sigma}$ into $\boldsymbol{\Sigma} = \sigma^2\mathbf{V} = \boldsymbol{\beta}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}\boldsymbol{\beta}' = \boldsymbol{\beta}\sigma^2\boldsymbol{\Lambda}_{\mathbf{V}}\boldsymbol{\beta}'$, where $\sigma \in \mathbb{R}_0^+$ is a *scale parameter* (equivariant under multiplication by a positive constant), and $\mathbf{V}$ a *shape matrix* (invariant under multiplication by a positive constant) with eigenvalues $\boldsymbol{\Lambda}_{\mathbf{V}} = \mathrm{diag}(\lambda_{1;\mathbf{V}}, \ldots, \lambda_{k;\mathbf{V}}) = \sigma^{-2}\,\mathrm{diag}(\lambda_{1;\boldsymbol{\Sigma}}, \ldots, \lambda_{k;\boldsymbol{\Sigma}}) = \sigma^{-2}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}$; $\boldsymbol{\beta}$ is an element of the so-called *special orthogonal group* $\mathcal{SO}_k := \{\mathbf{O} | \mathbf{O}'\mathbf{O} = \mathbf{I}_k, |\mathbf{O}| = 1\}$ diagonalizing both $\boldsymbol{\Sigma}$ and $\mathbf{V}$, the columns $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k$ of which are the eigenvectors (common to $\boldsymbol{\Sigma}$ and $\mathbf{V}$) we are interested in.

Such decomposition of scatter into scale and shape can be achieved in various ways. Here, we adopt the determinant-based definition of scale

$$\sigma := |\boldsymbol{\Sigma}|^{1/2k} = \prod_{j=1}^{k} \lambda_{j;\boldsymbol{\Sigma}}^{1/2k} \qquad \text{hence } \mathbf{V} := \boldsymbol{\Sigma}/\sigma^2 = \boldsymbol{\Sigma}/|\boldsymbol{\Sigma}|^{1/k},$$

which implies that $|\mathbf{V}| = \prod_{j=1}^{k} \lambda_{j;\mathbf{V}} = 1$. As shown by Paindaveine (2008), this choice indeed is the only one for which the information matrix for scale and shape is block-diagonal, which greatly simplifies inference. The parametric families of elliptical distributions with specified standardized radial density f_1 then are indexed by the $L = k(k+2)$-dimensional parameter

$$\boldsymbol{\vartheta} := (\boldsymbol{\theta}', \sigma^2, (\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_{\mathbf{V}})', (\mathrm{vec}\,\boldsymbol{\beta})')' =: (\boldsymbol{\vartheta}_I', \vartheta_{II}, \boldsymbol{\vartheta}_{III}', \boldsymbol{\vartheta}_{IV}')',$$

where $\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}(\boldsymbol{\Lambda}_{\mathbf{V}}) = (\lambda_{2;\mathbf{V}}, \ldots, \lambda_{k;\mathbf{V}})'$ since $\lambda_{1;\mathbf{V}} = \prod_{j=2}^{k} \lambda_{j;\mathbf{V}}^{-1}$.

This $\boldsymbol{\vartheta}$-parametrization however requires a fully identified k-tuple of eigenvectors, which places the following restriction on the eigenvalues $\boldsymbol{\Lambda}_{\mathbf{V}}$.

ASSUMPTION (A). The eigenvalues $\lambda_{j;\mathbf{V}}$ of the shape matrix $\mathbf{V}$ are all distinct, that is, since $\boldsymbol{\Sigma}$ (hence also $\mathbf{V}$) is positive definite, $\lambda_{1;\mathbf{V}} > \lambda_{2;\mathbf{V}} > \cdots > \lambda_{k;\mathbf{V}} > 0$.

Denote by $\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)}$ the joint distribution of $\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}$ under parameter value $\boldsymbol{\vartheta}$ and standardized radial density $f_1 \in \mathcal{F}_1$; the parameter space [the definition of which includes Assumption (A)] then is

$$\boldsymbol{\Theta} := \mathbb{R}^k \times \mathbb{R}_0^+ \times \mathcal{C}^{k-1} \times \mathrm{vec}(\mathcal{SO}_k),$$

where $\mathcal{C}^{k-1}$ is the open cone of $(\mathbb{R}_0^+)^{k-1}$ with strictly ordered (from largest to smallest) coordinates.

Since $\mathrm{vec}(\mathcal{S}O_k)$ is a nonlinear manifold of $\mathbb{R}^{k^2}$: the $\mathrm{vec}(\boldsymbol{\beta})$-parametrized experiments are *curved experiments*, in which the standard methods [see Section 11.9 of Le Cam (1986)] for constructing locally asymptotically optimal tests do not apply. It is well known, however [see, e.g., Khuri and Good (1989)], that any element $\boldsymbol{\beta}$ of $\mathcal{S}O_k$ can be expressed as the exponential $\exp(\boldsymbol{\iota})$ of a $k \times k$ *skew-symmetric matrix* $\boldsymbol{\iota}$, itself characterized by the $k(k-1)/2$-vector $\mathrm{vech}^+(\boldsymbol{\iota})$ of its upper off-diagonal elements. The differentiable mapping $\hbar : \mathrm{vech}^+(\boldsymbol{\iota}) \mapsto \hbar(\mathrm{vech}^+(\boldsymbol{\iota})) := \mathrm{vec}(\exp(\boldsymbol{\iota}))$ from $\mathbb{R}^{k(k-1)/2}$ to $\mathcal{S}O_k$ is one-to-one, so that $\mathrm{vech}^+(\boldsymbol{\iota}) \in \mathbb{R}^{k(k-1)/2}$ also can be used as a parametrization instead of $\mathrm{vec}(\boldsymbol{\beta}) \in \mathrm{vec}(\mathcal{S}O_k)$. Both parametrizations yield uniform local asymptotic normality (ULAN). Unlike the $\mathrm{vec}(\boldsymbol{\beta})$-parametrized one, the $\mathrm{vech}^+(\boldsymbol{\iota})$-parametrized experiment is not curved, as $\mathrm{vech}^+(\boldsymbol{\iota})$ freely ranges over $\mathbb{R}^{k(k-1)/2}$, so that the standard methods for constructing locally asymptotically optimal tests apply—which is not the case with curved experiments. On the other hand, neither the $\mathrm{vech}^+(\boldsymbol{\iota})$-part of the central sequence, nor the image in the $\mathrm{vech}^+(\boldsymbol{\iota})$-space of the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ yield tractable forms. Therefore, we rather state ULAN for the curved $\mathrm{vec}(\boldsymbol{\beta})$-parametrization. Then (Section 3.3), we develop a general theory of locally asymptotically optimal tests for differentiable hypotheses in curved ULAN experiments.

Without Assumption (A), the $\boldsymbol{\vartheta}$-parametrization is not valid, and cannot enjoy LAN nor ULAN; optimality properties (of a local and asymptotic nature) then cannot be obtained. As far as validity issues (irrespective of optimality properties) are considered, however, this assumption can be weakened. If the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ is to make any sense, the first eigenvector $\boldsymbol{\beta}_1$ clearly should be identifiable, but not necessarily the remaining ones. The following assumption on the $\lambda_{j;\mathbf{V}}$'s, under which $\boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_k$ need not be identified, is thus minimal in that case.

ASSUMPTION (A_1'). The eigenvalues of the shape matrix $\mathbf{V}$ are such that $\lambda_{1;\mathbf{V}} > \lambda_{2;\mathbf{V}} \geq \cdots \geq \lambda_{k;\mathbf{V}} > 0$.

Under Assumption (A_1'), $\boldsymbol{\Theta}$ is broadened into a larger parameter space $\boldsymbol{\Theta}_1'$, which does not provide a valid parametrization anymore, and for which the ULAN property of Proposition 3.1 below no longer holds. As we shall see, all the tests we are proposing for $\mathcal{H}_0^{\boldsymbol{\beta}}$ nevertheless remains valid under the extended null hypothesis $\mathcal{H}_{0;1}^{\boldsymbol{\beta}\prime}$ resulting from weakening (A) into (A_1'). Note that, in case the null hypothesis is dealing with $\boldsymbol{\beta}_q$ instead of $\boldsymbol{\beta}_1$, the appropriate weakening of Assumption (A) is the following.

ASSUMPTION (A_q'). The eigenvalues of the shape matrix $\mathbf{V}$ are such that $\lambda_{1;\mathbf{V}} \geq \cdots \geq \lambda_{q-1;\mathbf{V}} > \lambda_{q;\mathbf{V}} > \lambda_{q+1;\mathbf{V}} \geq \cdots \geq \lambda_{k;\mathbf{V}} > 0$.

This yields enlarged parameter space $\mathbf{\Theta}'_q$ and null hypothesis $\mathcal{H}_{0;q}^{\boldsymbol{\beta}\prime}$.

Similarly, the null hypothesis $\mathcal{H}_0^{\mathbf{\Lambda}}$ requires the identifiability of the groups of q largest (hence $k-q$ smallest) eigenvalues; within each group, however, eigenvalues may coincide, yielding the following assumption.

ASSUMPTION (A''_q). The eigenvalues of the shape matrix $\mathbf{V}$ are such that $\lambda_{1;\mathbf{V}} \geq \cdots \geq \lambda_{q-1;\mathbf{V}} \geq \lambda_{q;\mathbf{V}} > \lambda_{q+1;\mathbf{V}} \geq \cdots \geq \lambda_{k;\mathbf{V}} > 0$.

This yields enlarged parameter space $\mathbf{\Theta}''_q$ and null hypothesis $\mathcal{H}_{0;q}^{\mathbf{\Lambda}\prime\prime}$, say. As we shall see, the tests we are proposing for $\mathcal{H}_0^{\mathbf{\Lambda}}$ remain valid under $\mathcal{H}_{0;q}^{\mathbf{\Lambda}\prime\prime}$.

3.2. *Curved ULAN experiments*. Uniform local asymptotic normality (ULAN) for the parametric families or *experiments* $\mathcal{P}_{f_1}^{(n)} := \{\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)} : \boldsymbol{\vartheta} \in \mathbf{\Theta}\}$, with classical root-$n$ rate, is the main technical tool of this paper. For any $\boldsymbol{\vartheta} := (\boldsymbol{\theta}', \sigma^2, (\mathrm{d}\overset{\circ}{\mathrm{ve}}\mathrm{c}\, \mathbf{\Lambda}_{\mathbf{V}})', (\mathrm{vec}\, \boldsymbol{\beta})')' \in \mathbf{\Theta}$, a *local alternative* is a sequence $\boldsymbol{\vartheta}^{(n)} \in \mathbf{\Theta}$ such that $(\boldsymbol{\vartheta}^{(n)} - \boldsymbol{\vartheta})$ is $O(n^{-1/2})$. For any such $\boldsymbol{\vartheta}^{(n)}$, consider a further sequence $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}^{(n)}$, with $\boldsymbol{\tau}^{(n)} = ((\boldsymbol{\tau}^{I(n)})', \tau^{II(n)}, (\boldsymbol{\tau}^{III(n)})', (\boldsymbol{\tau}^{IV(n)})')'$ such that $\sup_n \boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)} < \infty$ and $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}^{(n)} \in \mathbf{\Theta}$ for all n. Note that such $\boldsymbol{\tau}^{(n)}$ exist: $\boldsymbol{\tau}^{I(n)}$ can be any bounded sequence of $\mathbb{R}^k$, $\tau^{II(n)}$ any bounded sequence with $\tau^{II(n)} > -n^{1/2}\sigma^{2(n)}$, $\boldsymbol{\tau}^{III(n)}$ any bounded sequence of real $(k-1)$-tuples $(\tau_1^{III(n)}, \ldots, \tau_{k-1}^{III(n)})$ such that

$$
\begin{aligned}
0 &< \lambda_{k;\mathbf{V}}^{(n)} + n^{-1/2}\tau_{k-1}^{III(n)} < \cdots \\
&< \lambda_{3;\mathbf{V}}^{(n)} + n^{-1/2}\tau_2^{III(n)} < \lambda_{2;\mathbf{V}}^{(n)} + n^{-1/2}\tau_1^{III(n)} \\
&< \prod_{j=2}^{k} (\lambda_{j;\mathbf{V}}^{(n)} + n^{-1/2}\tau_{j-1}^{III(n)})^{-1},
\end{aligned}
$$

which ensures that the perturbed eigenvalues $\lambda_{j;\mathbf{V}}^{(n)} + n^{-1/2}\ell_j^{(n)}$, with

$$
\begin{aligned}
\ell_1^{(n)} &:= n^{1/2}\left(\prod_{j=2}^{k} (\lambda_{j;\mathbf{V}}^{(n)} + n^{-1/2}\tau_{j-1}^{III(n)})^{-1} - \lambda_{1;\mathbf{V}}^{(n)}\right) \\
&= -\lambda_{1;\mathbf{V}}^{(n)} \sum_{j=2}^{k} (\lambda_{j;\mathbf{V}}^{(n)})^{-1}\tau_{j-1}^{III(n)} + O(n^{-1/2})
\end{aligned}
\tag{3.1}
$$

and $(\ell_2^{(n)}, \ldots, \ell_k^{(n)}) := \boldsymbol{\tau}^{III(n)}$, still satisfy Assumption (A) and yield determinant value one. Writing $\boldsymbol{\ell}^{(n)}$ for the diagonal $k \times k$ matrix with diagonal elements

$\ell_1^{(n)}, \dots, \ell_k^{(n)}$, we then have

$$\begin{aligned}\operatorname{tr}\big((\mathbf{\Lambda}_{\mathbf{V}}^{(n)})^{-1}\boldsymbol{\ell}^{(n)}\big) &= \big(\lambda_{1;\mathbf{V}}^{(n)}\big)^{-1}\Bigg[-\lambda_{1;\mathbf{V}}^{(n)}\sum_{j=2}^{k}\big(\lambda_{j;\mathbf{V}}^{(n)}\big)^{-1}\tau_{j-1}^{III(n)} + O(n^{-1/2})\Bigg] \\ &\qquad + \sum_{j=2}^{k}\big(\lambda_{j;\mathbf{V}}^{(n)}\big)^{-1}\tau_{j-1}^{III(n)} \\ &= O(n^{-1/2}).\end{aligned}$$

Finally, denote by $\mathbf{M}_k'(\lambda_2, \dots, \lambda_k) = (-\lambda_1(\lambda_2^{-1}, \dots, \lambda_k^{-1})' \vdots \mathbf{I}_{k-1})'$ the value at $(\lambda_2, \dots, \lambda_k)$ of the Jacobian matrix of

$$(\lambda_2, \dots, \lambda_k) \mapsto \left(\lambda_1 := \prod_{j=2}^{k}\lambda_j^{-1}, \lambda_2, \dots, \lambda_k\right).$$

Letting $\mathbf{\Lambda} := \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$, we have $\mathbf{M}_k'(\lambda_2, \dots, \lambda_k)\, \mathrm{d}\mathring{\mathrm{vec}}(\mathbf{l}) = \mathrm{dvec}(\mathbf{l})$ for any $k \times k$ real matrix $\mathbf{l}$ such that $\operatorname{tr}(\mathbf{\Lambda}^{-1}\mathbf{l}) = 0$. Indeed,

$$\begin{aligned}\mathbf{M}_k'(\lambda_2, \dots, \lambda_k)\, \mathrm{d}\mathring{\mathrm{vec}}(\mathbf{l}) &= \big(-\lambda_1(\lambda_2^{-1}, \dots, \lambda_k^{-1})' \vdots \mathbf{I}_{k-1}\big)'(\mathrm{d}\mathring{\mathrm{vec}}\,\mathbf{l}) \\ &= \big(-\lambda_1\big(\operatorname{tr}(\mathbf{\Lambda}^{-1}\mathbf{l}) - (\lambda_1)^{-1}\mathbf{l}_{11}\big) \vdots (\mathrm{d}\mathring{\mathrm{vec}}\,\mathbf{l})'\big)' \\ &= \mathrm{dvec}(\mathbf{l}),\end{aligned}$$

an identity that will be used later on for $\mathbf{M}_k^{\mathbf{\Lambda}_{\mathbf{V}}} := \mathbf{M}_k(\mathrm{d}\mathring{\mathrm{vec}}(\mathbf{\Lambda}_{\mathbf{V}}))$.

The problem is slightly more delicate for $\boldsymbol{\tau}^{IV(n)}$, which must be such that $\mathrm{vec}(\boldsymbol{\beta}^{(n)}) + n^{-1/2}\boldsymbol{\tau}^{IV(n)}$ remains in $\mathrm{vec}(\mathcal{S}O_k)$. That is, $\boldsymbol{\tau}^{IV(n)}$ must be of the form $\boldsymbol{\tau}^{IV(n)} = \mathrm{vec}(\mathbf{b}^{(n)})$, with

$$\begin{aligned}\mathbf{0} &= \big(\boldsymbol{\beta}^{(n)} + n^{-1/2}\mathbf{b}^{(n)}\big)'\big(\boldsymbol{\beta}^{(n)} + n^{-1/2}\mathbf{b}^{(n)}\big) - \mathbf{I}_k \\ &= n^{-1/2}\big(\boldsymbol{\beta}^{(n)\prime}\mathbf{b}^{(n)} + \mathbf{b}^{(n)\prime}\boldsymbol{\beta}^{(n)}\big) + n^{-1}\mathbf{b}^{(n)\prime}\mathbf{b}^{(n)}.\end{aligned} \tag{3.2}$$

That is, $\boldsymbol{\beta}^{(n)\prime}\mathbf{b}^{(n)} + n^{-1/2}\mathbf{b}^{(n)\prime}\mathbf{b}^{(n)}/2$ should be skew-symmetric. Such local perturbations admit an intuitive interpretation: we have indeed

$$\boldsymbol{\beta}^{(n)} + n^{-1/2}\mathbf{b}^{(n)} = \boldsymbol{\beta}^{(n)}\boldsymbol{\beta}^{(n)\prime}\big(\boldsymbol{\beta}^{(n)} + n^{-1/2}\mathbf{b}^{(n)}\big) = \boldsymbol{\beta}^{(n)}\big(\mathbf{I}_k + n^{-1/2}\boldsymbol{\beta}^{(n)\prime}\mathbf{b}^{(n)}\big)$$

an expression in which $\mathbf{I}_k + n^{-1/2}\boldsymbol{\beta}^{(n)\prime}\mathbf{b}^{(n)}$, up to a $O(n^{-1})$ quantity, coincides with the first-order approximation of the exponential of a skew-symmetric matrix, and therefore can be interpreted as an infinitesimal rotation. Identity (3.2) provides a characterization of $\mathcal{S}O_k$ in the vicinity of $\boldsymbol{\beta}^{(n)}$. The tangent space [in $\mathbb{R}^{k^2}$, at

$\mathrm{vec}(\boldsymbol{\beta})$] to $\mathrm{vec}(\mathcal{S}O_k)$ is obtained by linearizing (3.2). More precisely, this tangent space is of the form

$$\begin{aligned}(3.3)\qquad &\{\mathrm{vec}(\boldsymbol{\beta}+\mathbf{b})\,|\,\mathrm{vec}(\mathbf{b})\in\mathbb{R}^{k^2}\text{ and }\boldsymbol{\beta}'\mathbf{b}+\mathbf{b}'\boldsymbol{\beta}=\mathbf{0}\}\\ &\quad=\{\mathrm{vec}(\boldsymbol{\beta}+\mathbf{b})\,|\,\mathrm{vec}(\mathbf{b})\in\mathbb{R}^{k^2}\text{ and }\boldsymbol{\beta}'\mathbf{b}\text{ skew-symmetric}\}.\end{aligned}$$

We then have the following result (see the Appendix for the proof).

PROPOSITION 3.1. *The experiment* $\mathcal{P}_{f_1}^{(n)} := \{\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)}\,|\,\boldsymbol{\vartheta}\in\boldsymbol{\Theta}\}$ *is ULAN, with central sequence* $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{(n)} := (\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I\prime}, \Delta_{\boldsymbol{\vartheta};f_1}^{II}, \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III\prime}, \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{IV\prime})'$, *where* [*with* $d_i := d_i^{(n)}(\boldsymbol{\theta},\mathbf{V})$ *and* $\mathbf{U}_i := \mathbf{U}_i^{(n)}(\boldsymbol{\theta},\mathbf{V})$ *as defined in* (2.2), *and letting* $\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}} := \mathbf{M}_k(\mathrm{d}\overset{\circ}{\mathrm{vec}}\,\boldsymbol{\Lambda}_\mathbf{V})$],

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I} := \frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{n}\varphi_{f_1}\left(\frac{d_i}{\sigma}\right)\mathbf{V}^{-1/2}\mathbf{U}_i,$$

$$\Delta_{\boldsymbol{\vartheta};f_1}^{II} := \frac{1}{2\sqrt{n}\sigma^2}\sum_{i=1}^{n}\left(\varphi_{f_1}\left(\frac{d_i}{\sigma}\right)\frac{d_i}{\sigma}-k\right),$$

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III} := \frac{1}{2\sqrt{n}}\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}}\mathbf{H}_k(\boldsymbol{\Lambda}_\mathbf{V}^{-1/2}\boldsymbol{\beta}')^{\otimes 2}\sum_{i=1}^{n}\mathrm{vec}\left(\varphi_{f_1}\left(\frac{d_i}{\sigma}\right)\frac{d_i}{\sigma}\mathbf{U}_i\mathbf{U}_i'\right)$$

and

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{IV} := \frac{1}{2\sqrt{n}}\mathbf{G}_k^{\boldsymbol{\beta}}\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_\mathbf{V}}(\mathbf{V}^{\otimes 2})^{-1/2}\sum_{i=1}^{n}\mathrm{vec}\left(\varphi_{f_1}\left(\frac{d_i}{\sigma}\right)\frac{d_i}{\sigma}\mathbf{U}_i\mathbf{U}_i'\right),$$

with $\mathbf{G}_k^{\boldsymbol{\beta}} := (\mathbf{G}_{k;12}^{\boldsymbol{\beta}}\mathbf{G}_{k;13}^{\boldsymbol{\beta}}\cdots\mathbf{G}_{k;(k-1)k}^{\boldsymbol{\beta}})$, $\mathbf{G}_{k;jh}^{\boldsymbol{\beta}} := \mathbf{e}_j\otimes\boldsymbol{\beta}_h - \mathbf{e}_h\otimes\boldsymbol{\beta}_j$ *and*

$$\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_\mathbf{V}} := \left(\mathbf{L}_{k;12}^{\boldsymbol{\beta},\boldsymbol{\Lambda}_\mathbf{V}}\mathbf{L}_{k;13}^{\boldsymbol{\beta},\boldsymbol{\Lambda}_\mathbf{V}}\cdots\mathbf{L}_{k;(k-1)k}^{\boldsymbol{\beta},\boldsymbol{\Lambda}_\mathbf{V}}\right)', \qquad \mathbf{L}_{k;jh}^{\boldsymbol{\beta},\boldsymbol{\Lambda}_\mathbf{V}} := (\lambda_{h;\mathbf{V}}-\lambda_{j;\mathbf{V}})(\boldsymbol{\beta}_h\otimes\boldsymbol{\beta}_j),$$

and with block-diagonal information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1} = \mathrm{diag}(\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}^{I}, \Gamma_{\boldsymbol{\vartheta};f_1}^{II}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}^{III}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}^{IV}), \tag{3.4}$$

where, defining $\mathbf{D}_k(\boldsymbol{\Lambda}_\mathbf{V}) := \frac{1}{4}\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}}\mathbf{H}_k[\mathbf{I}_{k^2}+\mathbf{K}_k](\boldsymbol{\Lambda}_\mathbf{V}^{-1})^{\otimes 2}\mathbf{H}_k'(\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}})'$,

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}^{I} = \frac{\mathcal{I}_k(f_1)}{k\sigma^2}\mathbf{V}^{-1}, \qquad \Gamma_{\boldsymbol{\vartheta};f_1}^{II} = \frac{\mathcal{J}_k(f_1)-k^2}{4\sigma^4}, \qquad \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}^{III} = \frac{\mathcal{J}_k(f_1)}{k(k+2)}\mathbf{D}_k(\boldsymbol{\Lambda}_\mathbf{V})$$

and

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}^{IV} := \frac{1}{4}\frac{\mathcal{J}_k(f_1)}{k(k+2)}\mathbf{G}_k^{\boldsymbol{\beta}}\,\mathrm{diag}\big(\nu_{12}^{-1},\nu_{13}^{-1},\ldots,\nu_{(k-1)k}^{-1}\big)(\mathbf{G}_k^{\boldsymbol{\beta}})',$$

where $\nu_{jh} := \lambda_{j;\mathbf{V}}\lambda_{h;\mathbf{V}}/(\lambda_{j;\mathbf{V}} - \lambda_{h;\mathbf{V}})^2$. *More precisely, for any local alternative* $\boldsymbol{\vartheta}^{(n)}$ *and any bounded sequence* $\boldsymbol{\tau}^{(n)}$ *such that* $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}^{(n)} \in \boldsymbol{\Theta}$, *we have, under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}$,

$$\Lambda^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)}/\boldsymbol{\vartheta}^{(n)};f_1} := \log\big(d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)};f_1}/d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}\big)$$
$$= \big(\boldsymbol{\tau}^{(n)}\big)'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1} - \tfrac{1}{2}\big(\boldsymbol{\tau}^{(n)}\big)'\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1)$$

and $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}^{(n)};f_1} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1})$, *as* $n \to \infty$.

The block-diagonal structure of the information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}$ implies that inference on $\boldsymbol{\beta}$ (resp., $\boldsymbol{\Lambda}_{\mathbf{V}}$) can be conducted under unspecified $\boldsymbol{\theta}$, σ and $\boldsymbol{\Lambda}_{\mathbf{V}}$ (resp., $\boldsymbol{\beta}$) as if the latter were known, at no asymptotic cost. The orthogonality between the eigenvalue and eigenvector parts of the central sequence is structural, while that between the eigenvalue and eigenvector parts on one hand and the scale parameter part on the other hand is entirely due to the determinant-based parametrization of scale [see Hallin and Paindaveine (2006b) or Paindaveine (2008)]. Note that $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};f_1}$, with rank $k(k-1)/2 < k^2$, is not invertible.

3.3. *Locally asymptotically optimal tests for differentiable hypotheses in curved ULAN experiments.* Before addressing testing problems involving eigenvalues and eigenvectors, we need a general theory for locally asymptotically optimal tests in curved ULAN experiments, which we are developing in this section.

Consider a ULAN sequence of experiments $\{\mathrm{P}^{(n)}_{\boldsymbol{\xi}} : \boldsymbol{\xi} \in \Xi\}$, where Ξ is an open subset of $\mathbb{R}^m$, with central sequence $\boldsymbol{\Delta}_{\boldsymbol{\xi}}$ and information $\boldsymbol{\Gamma}_{\boldsymbol{\xi}}$. For the simplicity of exposition, assume that $\boldsymbol{\Gamma}_{\boldsymbol{\xi}}$ for any $\boldsymbol{\xi}$ has full rank m. Let $\hbar : \Xi \to \mathbb{R}^p$, $p \geq m$, be a continuously differentiable mapping such that the Jacobian matrix $D\hbar(\boldsymbol{\xi})$ has full rank m for all $\boldsymbol{\xi}$, and consider the experiments $\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}} : \boldsymbol{\vartheta} \in \boldsymbol{\Theta} := \hbar(\Xi)\}$, where, with a slight abuse of notation, $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}} := \mathrm{P}^{(n)}_{\boldsymbol{\xi}}$ for $\boldsymbol{\vartheta} = \hbar(\boldsymbol{\xi})$. This sequence also is ULAN, with central sequence $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}}$ and information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}}$ such that [see (A.5) and the proof of Lemma A.1], at $\boldsymbol{\vartheta} = \hbar(\boldsymbol{\xi})$, and [up to $o_{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}}}(1)$'s which, for simplicity, we omit here] $\boldsymbol{\Delta}_{\boldsymbol{\xi}} = D\hbar'(\boldsymbol{\xi})\boldsymbol{\Delta}_{\boldsymbol{\vartheta}}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\xi}} = D\hbar'(\boldsymbol{\xi})\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}} D\hbar(\boldsymbol{\xi})$—throughout, we write $D\hbar'(\cdot)$, $D\bar{b}'(\cdot)$, etc., instead of $(D\hbar(\cdot))'$, $(D\bar{b}(\cdot))'$, etc. In general, $\boldsymbol{\Theta}$ is a nonlinear manifold of $\mathbb{R}^p$; the experiment parametrized by $\boldsymbol{\Theta}$ then is a *curved* experiment.

Next, denoting by C an r-dimensional manifold in $\mathbb{R}^p$, $r < p$, consider the null hypothesis $\mathcal{H}_0 : \boldsymbol{\vartheta} \in C \cap \boldsymbol{\Theta}$—in general, a nonlinear restriction of the parameter space $\boldsymbol{\Theta}$. The same hypothesis can be expressed in the $\boldsymbol{\xi}$-parametrization as $\mathcal{H}_0 : \boldsymbol{\xi} \in \Xi_0$, where $\Xi_0 := \hbar^{-1}(C \cap \boldsymbol{\Theta})$ is a (ℓ-dimensional, say) submanifold of Ξ. Fix $\boldsymbol{\xi}_0 = \hbar^{-1}(\boldsymbol{\vartheta}_0) \in \Xi_0$, and let $\bar{l} : B \subset \mathbb{R}^\ell \to \Xi$ be a local (at $\boldsymbol{\xi}_0$) chart for this manifold.

Define $\boldsymbol{\alpha}_0 := \bar{l}^{-1}(\boldsymbol{\xi}_0)$. At $\boldsymbol{\xi}_0$, $\mathcal{H}_0$ is linearized into $\mathcal{H}_{\boldsymbol{\xi}_0} : \boldsymbol{\xi} \in \boldsymbol{\xi}_0 + \mathcal{M}(D\bar{l}(\boldsymbol{\alpha}_0))$, where $D\bar{l}(\boldsymbol{\alpha}_0)$ is the Jacobian matrix of $\bar{l}$ (with rank ℓ) computed at $\boldsymbol{\alpha}_0$ and $\mathcal{M}(\mathbf{A})$

denotes the vector space spanned by the columns of a matrix $\mathbf{A}$. At $\boldsymbol{\alpha}_0$, a locally asymptotically most stringent test statistic (at $\boldsymbol{\xi}_0$) for $\mathcal{H}_{\boldsymbol{\xi}_0}$ is

$$Q_{\boldsymbol{\xi}_0} := \boldsymbol{\Delta}'_{\boldsymbol{\xi}_0}\big(\boldsymbol{\Gamma}^{-1}_{\boldsymbol{\xi}_0} - D\bar{l}(\boldsymbol{\alpha}_0)(D\bar{l}'(\boldsymbol{\alpha}_0)\boldsymbol{\Gamma}_{\boldsymbol{\xi}_0}D\bar{l}(\boldsymbol{\alpha}_0))^{-1}D\bar{l}'(\boldsymbol{\alpha}_0)\big)\boldsymbol{\Delta}_{\boldsymbol{\xi}_0} \tag{3.5}$$

[see Section 11.9 of Le Cam (1986)]. This test statistic is nothing else but the squared Euclidean norm of the orthogonal projection, onto the linear space orthogonal to $\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\xi}_0}D\bar{l}(\boldsymbol{\alpha}_0)$, of the standardized central sequence $\boldsymbol{\Gamma}^{-1/2}_{\boldsymbol{\xi}_0}\boldsymbol{\Delta}_{\boldsymbol{\xi}_0}$. In view of ULAN, the asymptotic behavior of $\boldsymbol{\Delta}_{\boldsymbol{\xi}_0}$ is the same under local alternatives in Ξ_0 as under local alternatives in $\boldsymbol{\xi}_0 + \mathcal{M}(D\bar{l}(\boldsymbol{\alpha}_0))$, so that the same test statistic $Q_{\boldsymbol{\xi}_0}$, which (at $\boldsymbol{\xi}_0$) is locally asymptotically most stringent for $\mathcal{H}_{\boldsymbol{\xi}_0}$, is also locally asymptotically most stringent for $\mathcal{H}_0$.

In many cases, however, it is highly desirable to express the most stringent statistic in the curved $\boldsymbol{\Theta}$-parametrization, which, as is the case for the eigenvalues/eigenvectors problems considered in this work, is the natural parametrization. This is the objective of the following result (see the Appendix for the proof).

PROPOSITION 3.2. *With the same notation as above, a locally asymptotically most stringent statistic (at $\boldsymbol{\vartheta}_0$) for testing $\mathcal{H}_0 : \boldsymbol{\vartheta} \in C \cap \boldsymbol{\Theta}$ is*

$$Q_{\boldsymbol{\xi}_0} = Q_{\boldsymbol{\vartheta}_0} := \boldsymbol{\Delta}'_{\boldsymbol{\vartheta}_0}\big(\boldsymbol{\Gamma}^{-}_{\boldsymbol{\vartheta}_0} - D\bar{b}(\boldsymbol{\eta}_0)(D\bar{b}'(\boldsymbol{\eta}_0)\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0}D\bar{b}(\boldsymbol{\eta}_0))^{-}D\bar{b}'(\boldsymbol{\eta}_0)\big)\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0}, \tag{3.6}$$

where $\bar{b} : A \subset \mathbb{R}^{\ell} \to \mathbb{R}^p$ is a local (at $\boldsymbol{\vartheta}_0$) chart for the tangent (still at $\boldsymbol{\vartheta}_0$) to the manifold $C \cap \boldsymbol{\Theta}$, $\boldsymbol{\eta}_0 := \bar{b}^{-1}(\boldsymbol{\vartheta}_0)$, and $\mathbf{A}^{-}$ denotes the Moore–Penrose inverse of $\mathbf{A}$.

Hence, a locally asymptotically most stringent (at $\boldsymbol{\xi}_0$ or $\boldsymbol{\vartheta}_0$, depending on the parametrization) test for $\mathcal{H}_0$ can be based on either of the two quadratic forms $Q_{\boldsymbol{\xi}_0}$ or $Q_{\boldsymbol{\vartheta}_0}$, which coincide, and are asymptotically chi-square [$(m-\ell)$ degrees of freedom] under $\mathrm{P}^{(n)}_{\boldsymbol{\xi}_0} = \mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0}$, for $\boldsymbol{\vartheta}_0 = \hbar(\boldsymbol{\xi}_0)$. For practical implementation, of course, an adequately discretized root-n consistent estimator has to be substituted for the unknown $\boldsymbol{\vartheta}_0$ or $\boldsymbol{\xi}_0$—which asymptotically does not affect the test statistic.

Provided that Ξ remains an open subset of $\mathbb{R}^m$, the assumption of a full-rank information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\xi}}$ is not required. Hallin and Puri [(1994), Lemma 5.12] indeed have shown, in the case of ARMA experiments, that (3.5) remains locally asymptotically most stringent provided that generalized inverses (not necessarily Moore–Penrose ones) are substituted for the inverses of noninvertible matrices, yielding

$$Q_{\boldsymbol{\xi}_0} := \boldsymbol{\Delta}'_{\boldsymbol{\xi}_0}\big(\boldsymbol{\Gamma}^{-}_{\boldsymbol{\xi}_0} - D\bar{l}(\boldsymbol{\alpha}_0)(D\bar{l}'(\boldsymbol{\alpha}_0)\boldsymbol{\Gamma}_{\boldsymbol{\xi}_0}D\bar{l}(\boldsymbol{\alpha}_0))^{-}D\bar{l}'(\boldsymbol{\alpha}_0)\big)\boldsymbol{\Delta}_{\boldsymbol{\xi}_0}.$$

The same reasoning as in the proof of Proposition 3.2 then applies, mutatis mutandis, when "translating" $Q_{\boldsymbol{\xi}_0}$ into $Q_{\boldsymbol{\vartheta}_0}$ (with appropriate degrees of freedom).

4. Parametrically optimal tests for principal components.

4.1. *Optimal parametric tests for eigenvectors.* Testing the hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ on eigenvectors is a particular case of the problem considered in the previous section. The $\mathrm{vech}^+(\boldsymbol{\iota})$ parametrization [$\boldsymbol{\iota}$ an arbitrary skew-symmetric $(k \times k)$ matrix] yields a standard ULAN experiment, with parameter

$$\boldsymbol{\xi} := (\boldsymbol{\theta}', \sigma^2, (\mathrm{d}\overset{\circ}{\mathrm{v}}\mathrm{ec}(\boldsymbol{\Lambda}_{\mathbf{V}}))', (\mathrm{vech}^+(\boldsymbol{\iota}))')' \in \mathbb{R}^k \times \mathbb{R}^+ \times \mathcal{C}^{k-1} \times \mathbb{R}^{k(k-1)/2} =: \boldsymbol{\Xi},$$

hence $m = k(k+3)/2$, while Proposition 3.1 provides the curved ULAN experiment, with parameter $\boldsymbol{\vartheta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ and $p = k(k+2)$. ULAN for the $\boldsymbol{\xi}$-experiment readily follows from the fact that the mapping $\mathrm{vech}^+(\boldsymbol{\iota}) \mapsto \mathrm{vec}(\boldsymbol{\beta}) = \mathrm{vec}(\exp(\boldsymbol{\iota}))$ is continuously differentiable.

As explained before, the block-diagonal structure of the information matrix (3.4) implies that locally asymptotically optimal inference about $\boldsymbol{\beta}$ can be based on $\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta};f_1}$ only, as if $\boldsymbol{\theta}$, σ^2 and $\mathrm{d}\overset{\circ}{\mathrm{v}}\mathrm{ec}(\boldsymbol{\Lambda}_{\mathbf{V}})$ were specified. Since this also allows for simpler exposition and lighter notation, let us assume that these parameters take on specified values $\boldsymbol{\theta}$, σ^2 and $(\lambda_{2;\mathbf{V}}, \dots, \lambda_{k;\mathbf{V}})$, respectively. The resulting experiment then is parametrized either by $\mathrm{vec}\,\boldsymbol{\beta} \in \mathrm{vec}(\mathcal{SO}_k) \subset \mathbb{R}^{k^2}$ (playing the role of $\boldsymbol{\vartheta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ in the notation of Proposition 3.2) or by $\mathrm{vech}^+(\boldsymbol{\iota}) \in \mathbb{R}^{k(k-1)/2}$ (playing the role of $\boldsymbol{\xi}$).

In this experiment, the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ consists in the intersection of the linear manifold $C := (\boldsymbol{\beta}^{0\prime}, \mathbf{0}_{1\times(k-1)k})' + \mathcal{M}(\boldsymbol{\Upsilon})$, where $\boldsymbol{\Upsilon} := (\mathbf{0}_{k(k-1)\times k}, \mathbf{I}_{k(k-1)})'$, with the nonlinear manifold $\mathrm{vec}(\mathcal{SO}_k)$. Let $\boldsymbol{\beta}_0 := (\boldsymbol{\beta}^0, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_k)$ be such that $\mathrm{vec}(\boldsymbol{\beta}_0)$ belongs to that intersection. In view of Proposition 3.2, a most stringent test statistic [at $\mathrm{vec}(\boldsymbol{\beta}_0)$] for $\mathcal{H}_0^{\boldsymbol{\beta}}$ requires a chart for the tangent to $C \cap \mathrm{vec}(\mathcal{SO}_k)$ at $\mathrm{vec}(\boldsymbol{\beta}_0)$. It follows from (3.3) that this tangent space reduces to

$$\{\mathrm{vec}(\boldsymbol{\beta}_0 + \mathbf{b}) | \mathbf{b} := (\mathbf{0}, \mathbf{b}_2, \dots, \mathbf{b}_k) \text{ such that } \boldsymbol{\beta}_0'\mathbf{b} + \mathbf{b}'\boldsymbol{\beta}_0 = \mathbf{0}\}.$$

Solving for $\mathrm{vec}(\mathbf{b}) = (\mathbf{0}', \mathbf{b}_2', \dots, \mathbf{b}_k')'$ the system of constraints $\boldsymbol{\beta}_0'\mathbf{b} + \mathbf{b}'\boldsymbol{\beta}_0 = \mathbf{0}$ yields $\mathrm{vec}(\mathbf{b}) \in \mathcal{M}(\mathbf{P}_k^{\boldsymbol{\beta}_0})$, where

$$\mathbf{P}_k^{\boldsymbol{\beta}_0} := \begin{pmatrix} \mathbf{0}_{k\times k(k-1)} \\ \mathbf{I}_{k-1} \otimes [\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime}] - \displaystyle\sum_{i,j=1}^{k-1} [\mathbf{e}_{i;k-1}\mathbf{e}_{j;k-1}' \otimes \boldsymbol{\beta}_{j+1}\boldsymbol{\beta}_{i+1}'] \end{pmatrix} \tag{4.1}$$

(with $\mathbf{e}_{i;k-1}$ denoting the ith vector of the canonical basis of $\mathbb{R}^{k-1}$). A local chart for the tangent space of interest is then simply $\hbar : \boldsymbol{\eta} \in \mathbb{R}^{(k-1)k} \mapsto \hbar(\boldsymbol{\eta}) := \mathrm{vec}(\boldsymbol{\beta}_0) + \mathbf{P}_k^{\boldsymbol{\beta}_0}\boldsymbol{\eta}$, with $\boldsymbol{\eta}_0 = \hbar^{-1}(\mathrm{vec}(\boldsymbol{\beta}_0)) = \mathbf{0}_{(k-1)k}$ and $D\hbar(\boldsymbol{\eta}_0) = \mathbf{P}_k^{\boldsymbol{\beta}_0}$. Letting

$\boldsymbol{\vartheta}_0 := (\boldsymbol{\theta}', \sigma^2, (\mathrm{d\overset{\circ}{vec}}\,\boldsymbol{\Lambda}_{\mathbf{V}})', (\mathrm{vec}\,\boldsymbol{\beta}_0)')'$, the test statistic (3.6) takes the form

$$\begin{aligned} Q^{(n)}_{\boldsymbol{\vartheta}_0;f_1} &= \boldsymbol{\Delta}^{IV\prime}_{\boldsymbol{\vartheta}_0;f_1}[(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;f_1})^- - \mathbf{P}^{\boldsymbol{\beta}_0}_k((\mathbf{P}^{\boldsymbol{\beta}_0}_k)'\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;f_1}\mathbf{P}^{\boldsymbol{\beta}_0}_k)^-(\mathbf{P}^{\boldsymbol{\beta}_0}_k)']\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;f_1} \\ &= \frac{nk(k+2)}{\mathcal{J}_k(f_1)}\sum_{j=2}^k(\boldsymbol{\beta}'_j\mathbf{S}^{(n)}_{\boldsymbol{\vartheta}_0;f_1}\boldsymbol{\beta}^0)^2, \end{aligned} \tag{4.2}$$

with

$$\mathbf{S}^{(n)}_{\boldsymbol{\vartheta};f_1} := \frac{1}{n}\sum_{i=1}^n \varphi_{f_1}\left(\frac{d_i(\boldsymbol{\theta},\mathbf{V})}{\sigma}\right)\frac{d_i(\boldsymbol{\theta},\mathbf{V})}{\sigma}\mathbf{U}_i(\boldsymbol{\theta},\mathbf{V})\mathbf{U}'_i(\boldsymbol{\theta},\mathbf{V}), \tag{4.3}$$

where $\mathbf{V}$ denotes the unique shape value associated with the parameter $\boldsymbol{\vartheta}$.

After simple algebra, we obtain

$$\begin{aligned} &\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;f_1}[(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;f_1})^- - \mathbf{P}^{\boldsymbol{\beta}_0}_k((\mathbf{P}^{\boldsymbol{\beta}_0}_k)'\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;f_1}\mathbf{P}^{\boldsymbol{\beta}_0}_k)^-(\mathbf{P}^{\boldsymbol{\beta}_0}_k)'] \\ &\qquad = \tfrac{1}{2}\mathbf{G}^{\boldsymbol{\beta}_0}_k\,\mathrm{diag}\big(\mathbf{I}_{k-1}, \mathbf{0}_{(k-2)(k-1)/2\times(k-2)(k-1)/2}\big)(\mathbf{G}^{\boldsymbol{\beta}_0}_k)', \end{aligned} \tag{4.4}$$

which is idempotent with rank $(k-1)$. Since, moreover, $\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;f_1}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$, is asymptotically $\mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;f_1})$, Theorem 9.2.1 in Rao and Mitra (1971) then shows that $Q^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$, still under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$, is asymptotically chi-square with $(k-1)$ degrees of freedom.

The resulting test, which rejects $\mathcal{H}^{\boldsymbol{\beta}}_0$ at asymptotic level α whenever $Q^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$ exceeds the α-upper quantile $\chi^2_{k-1,1-\alpha}$ of the χ^2_{k-1} distribution, will be denoted as $\phi^{(n)}_{\boldsymbol{\beta};f_1}$. It is locally asymptotically most stringent, at $\boldsymbol{\vartheta}_0$ and under correctly specified standardized radial density f_1 (an unrealistic assumption). Of course, even if f_1 were supposed to be known, $Q^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$ still depends on the unspecified $\boldsymbol{\theta}, \sigma^2, \boldsymbol{\Lambda}_{\mathbf{V}}$ and $\boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_k$. In order to obtain a genuine test statistic, providing a locally asymptotically most stringent test at *any* $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}}_0$ (with an obvious abuse of notation), we would need replacing those nuisance parameters with adequate estimates. We will not pursue any further with this problem here, as it is of little practical interest for arbitrary density f_1. The same problem will be considered in Section 5 for the Gaussian and pseudo-Gaussian versions of (4.2), then in Section 6 for the rank-based ones.

4.2. *Optimal parametric tests for eigenvalues.* We now turn to the problem of testing the null hypothesis $\mathcal{H}^{\boldsymbol{\Lambda}}_0 : \sum_{j=q+1}^k \lambda_{j;\mathbf{V}} - p\sum_{j=1}^k \lambda_{j;\mathbf{V}} = 0$ against alternatives of the form $\mathcal{H}^{\boldsymbol{\Lambda}}_1 : \sum_{j=q+1}^k \lambda_{j;\mathbf{V}} - p\sum_{j=1}^k \lambda_{j;\mathbf{V}} < 0$, for given $p \in (0,1)$. Letting

$$h : (\lambda_2, \lambda_3, \dots, \lambda_k)' \in \mathcal{C}^{k-1} \mapsto \sum_{j=q+1}^k \lambda_j - p\left(\prod_{j=2}^k \lambda_j^{-1} + \sum_{j=2}^k \lambda_j\right)$$

and recalling that $\prod_{j=1}^{k} \lambda_{j;\mathbf{V}} = 1$, $\mathcal{H}_0^{\mathbf{\Lambda}}$ rewrites, in terms of $\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V})$, as $\mathcal{H}_0^{\mathbf{\Lambda}} : h(\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V})) = 0$, a highly nonlinear but smooth constraint on $\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V})$. It is easy to check that, when computed at $\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V})$, the gradient of h is

$$\begin{aligned}
&\mathrm{grad}\, h(\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V}))\\
&\quad = \big(p(\lambda_{1;\mathbf{V}}\lambda_{2;\mathbf{V}}^{-1} - 1), \ldots, p(\lambda_{1;\mathbf{V}}\lambda_{q;\mathbf{V}}^{-1} - 1),\\
&\qquad 1 + p(\lambda_{1;\mathbf{V}}\lambda_{q+1;\mathbf{V}}^{-1} - 1), \ldots, 1 + p(\lambda_{1;\mathbf{V}}\lambda_{k;\mathbf{V}}^{-1} - 1)\big)'.
\end{aligned}$$

Here again, in view of the block-diagonal form of the information matrix, we may restrict our attention to the $\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V})$-part $\mathbf{\Delta}_{\boldsymbol{\vartheta};f_1}^{III}$ of the central sequence as if $\boldsymbol{\theta}$, σ^2 and $\boldsymbol{\beta}$ were known; the parameter space then reduces to the $(k-1)$-dimensional open cone $\mathcal{C}^{k-1}$. Testing a nonlinear constraint on a parameter ranging over an open subset of $\mathbb{R}^{k-1}$ is much easier however than the corresponding problem involving a curved experiment, irrespective of the possible noninvertibility of the information matrix. In the noncurved experiment, indeed, a linearized version $\mathcal{H}_{0,\mathrm{lin}}^{\mathbf{\Lambda}} : \mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda_V}) \in \mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda}_0) + \mathcal{M}^{\perp}(\mathrm{grad}\, h(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0))$ of $\mathcal{H}_0^{\mathbf{\Lambda}}$ in the vicinity of $\mathrm{d\overset{\circ}{v}ec}(\mathbf{\Lambda}_0)$ satisfying $h(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0) = 0$ makes sense [$\mathcal{M}^{\perp}(\mathbf{A})$ denotes the orthogonal complement of $\mathcal{M}(\mathbf{A})$]. And, as mentioned in Section 3.3, under ULAN, the asymptotic behavior of $\mathbf{\Delta}_{\boldsymbol{\vartheta}_0;f_1}^{III}$, with $\boldsymbol{\vartheta}_0 = (\boldsymbol{\theta}', \sigma^2, (\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0)', (\mathrm{vec}\, \boldsymbol{\beta})')'$, is locally the same under $\mathcal{H}_{0,\mathrm{lin}}^{\mathbf{\Lambda}}$ as under $\mathcal{H}_0^{\mathbf{\Lambda}}$. As for the "linearized alternative" $\mathcal{H}_{1,\mathrm{lin}}^{\mathbf{\Lambda}}$ consisting of all $\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}$ values such that $(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda} - \mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0)'\, \mathrm{grad}\, h(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0) < 0$, it locally and asymptotically coincides with $\mathcal{H}_1^{\mathbf{\Lambda}}$: indeed, although the symmetric difference $\mathcal{H}_1^{\mathbf{\Lambda}} \Delta \mathcal{H}_{1,\mathrm{lin}}^{\mathbf{\Lambda}}$, for fixed n, is not empty, any $\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0 + n^{-1/2}\boldsymbol{\tau}^{III} \in \mathcal{H}_{1,\mathrm{lin}}^{\mathbf{\Lambda}}$ eventually belongs to $\mathcal{H}_1^{\mathbf{\Lambda}}$, and conversely. Therefore, a locally (at $\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0$) asymptotically optimal test for $\mathcal{H}_{0,\mathrm{lin}}^{\mathbf{\Lambda}}$ against $\mathcal{H}_{1,\mathrm{lin}}^{\mathbf{\Lambda}}$ is also locally asymptotically optimal for $\mathcal{H}_0^{\mathbf{\Lambda}}$ against $\mathcal{H}_1^{\mathbf{\Lambda}}$, and conversely, whatever the local asymptotic optimality concept adopted. Now, in the problem of testing $\mathcal{H}_{0,\mathrm{lin}}^{\mathbf{\Lambda}}$ against $\mathcal{H}_{1,\mathrm{lin}}^{\mathbf{\Lambda}}$ the null hypothesis is (locally) a hyperplane of $\mathbb{R}^{k-1}$, with an alternative consisting of the halfspace lying "below" that hyperplane. For such one-sided problems (locally and asymptotically, still at $\boldsymbol{\vartheta}_0$) uniformly most powerful tests exist; a *most powerful* test statistic is [Le Cam (1986), Section 11.9]

$$\begin{aligned}
T_{\boldsymbol{\vartheta}_0;f_1}^{(n)} &:= (\mathrm{grad}'\, h(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0)(\mathbf{\Gamma}_{\boldsymbol{\vartheta}_0;f_1}^{III})^{-1}\, \mathrm{grad}\, h(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0))^{-1/2}\\
&\qquad \times \mathrm{grad}'\, h(\mathrm{d\overset{\circ}{v}ec}\, \mathbf{\Lambda}_0)(\mathbf{\Gamma}_{\boldsymbol{\vartheta}_0;f_1}^{III})^{-1}\mathbf{\Delta}_{\boldsymbol{\vartheta}_0;f_1}^{III},
\end{aligned} \tag{4.5}$$

which, under $\mathrm{P}_{\boldsymbol{\vartheta}_0;f_1}^{(n)}$, is asymptotically standard normal. An explicit form of $T_{\boldsymbol{\vartheta}_0;f_1}^{(n)}$ requires a closed form expression of the inverse of $\mathbf{\Gamma}_{\boldsymbol{\vartheta};f_1}^{III} = (\mathcal{J}_k(f_1)/k(k+2)) \times \mathbf{D}_k(\mathbf{\Lambda_V})$. The following lemma provides such an expression for the inverse of $\mathbf{D}_k(\mathbf{\Lambda_V})$ (see the Appendix for the proof).

LEMMA 4.1. *Let* $\mathbf{P}_k^{\mathbf{\Lambda_V}} := \mathbf{I}_{k^2} - \frac{1}{k}\mathbf{\Lambda}_{\mathbf{V}}^{\otimes 2}\operatorname{vec}(\mathbf{\Lambda}_{\mathbf{V}}^{-1})(\operatorname{vec}(\mathbf{\Lambda}_{\mathbf{V}}^{-1}))'$ *and* $\mathbf{N}_k := (\mathbf{0}_{(k-1)\times 1}, \mathbf{I}_{k-1})$. *Then*, $(\mathbf{D}_k(\mathbf{\Lambda_V}))^{-1} = \mathbf{N}_k\mathbf{H}_k\mathbf{P}_k^{\mathbf{\Lambda_V}}(\mathbf{I}_{k^2} + \mathbf{K}_k)\mathbf{\Lambda}_{\mathbf{V}}^{\otimes 2}(\mathbf{P}_k^{\mathbf{\Lambda_V}})'\mathbf{H}_k'\mathbf{N}_k'$.

Using this lemma, it follows after some algebra that, for any $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0^{\mathbf{\Lambda}}$,

$$\begin{aligned}&\operatorname{grad}' h(\operatorname{d\mathring{v}ec}\mathbf{\Lambda}_0)(\mathbf{D}_k(\mathbf{\Lambda}_0))^{-1}\operatorname{grad} h(\operatorname{d\mathring{v}ec}\mathbf{\Lambda}_0)\\ &\quad = 2\left\{p^2\sum_{j=1}^{q}\lambda_{j;0}^2 + (1-p)^2\sum_{j=q+1}^{k}\lambda_{j;0}^2\right\} = a_{p,q}(\mathbf{\Lambda}_0)\end{aligned}$$

[where $\mathbf{\Lambda} \mapsto a_{p,q}(\mathbf{\Lambda})$ is the mapping defined in (1.3)], and

$$\operatorname{grad}' h(\operatorname{d\mathring{v}ec}\mathbf{\Lambda}_0)(\mathbf{D}_k(\mathbf{\Lambda}_0))^{-1}\mathbf{M}_k^{\mathbf{\Lambda}_0}\mathbf{H}_k(\mathbf{\Lambda}_0^{-1/2})^{\otimes 2} = \mathbf{c}_{p,q}'\mathbf{H}_k(\mathbf{\Lambda}_0^{1/2})^{\otimes 2}.$$

This and the definition of $\mathbf{H}_k$ yields

$$T_{\boldsymbol{\vartheta}_0;f_1}^{(n)} = \left(\frac{nk(k+2)}{\mathcal{J}_k(f_1)}\right)^{1/2}(a_{p,q}(\mathbf{\Lambda}_0))^{-1/2}\mathbf{c}_{p,q}'\operatorname{dvec}\big(\mathbf{\Lambda}_0^{1/2}\boldsymbol{\beta}'\mathbf{S}_{\boldsymbol{\vartheta}_0;f_1}^{(n)}\boldsymbol{\beta}\mathbf{\Lambda}_0^{1/2}\big) \tag{4.6}$$

with $\mathbf{S}_{\boldsymbol{\vartheta};f_1}^{(n)}$ defined in (4.3). The corresponding test, which rejects $\mathcal{H}_0^{\mathbf{\Lambda}}$ for small values of $T_{\boldsymbol{\vartheta}_0;f_1}^{(n)}$, will be denoted as $\phi_{\mathbf{\Lambda};f_1}^{(n)}$.

4.3. *Estimation of nuisance parameters*. The tests $\phi_{\boldsymbol{\beta};f_1}^{(n)}$ and $\phi_{\mathbf{\Lambda};f_1}^{(n)}$ derived in Sections 4.1 and 4.2 typically are valid under standardized radial density f_1 only; they mainly settle the optimality bounds at given density f_1, and are of little practical value. Due to its central role in multivariate analysis, the Gaussian case ($f_1 = \phi_1$) is an exception. In this subsection devoted to the treatment of nuisance parameters, we therefore concentrate on the Gaussian tests $\phi_{\boldsymbol{\beta};\phi_1}^{(n)}$ and $\phi_{\mathbf{\Lambda};\phi_1}^{(n)}$, to be considered in more detail in Section 5.

The test statistics derived in Sections 4.1 and 4.2 indeed still involve nuisance parameters which in practice have to be replaced with estimators. The traditional way of handling this substitution in ULAN families consists in assuming, for a null hypothesis of the form $\boldsymbol{\vartheta} \in \mathcal{H}_0$, the existence of a sequence $\hat{\boldsymbol{\vartheta}}^{(n)}$ of estimators of $\boldsymbol{\vartheta}$ satisfying all or part of the following assumptions (in the notation of this paper).

ASSUMPTION (B). We say that a sequence of estimators $(\hat{\boldsymbol{\vartheta}}^{(n)}, n \in \mathbb{N})$ satisfies Assumption (B) for the null $\mathcal{H}_0$ and the density f_1 if $\hat{\boldsymbol{\vartheta}}^{(n)}$ is:

(B1) *constrained*: $\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)}[\hat{\boldsymbol{\vartheta}}^{(n)} \in \mathcal{H}_0] = 1$ for all n and all $\boldsymbol{\vartheta} \in \mathcal{H}_0$;

(B2) $n^{1/2}$*-consistent*: for all $\boldsymbol{\vartheta} \in \mathcal{H}_0$, $n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)} - \boldsymbol{\vartheta}) = O_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)}$, as $n \to \infty$;

(B3) *locally asymptotically discrete*: for all $\boldsymbol{\vartheta} \in \mathcal{H}_0$ and all $c > 0$, there exists $M = M(c) > 0$ such that the number of possible values of $\hat{\boldsymbol{\vartheta}}^{(n)}$ in balls of the form $\{\mathbf{t} : n^{1/2}\|(\mathbf{t} - \boldsymbol{\vartheta})\| \le c\}$ is bounded by M, uniformly as $n \to \infty$.

These assumptions will be used later on. In the Gaussian or pseudo-Gaussian context we are considering here, however, Assumption (B3) can be dispensed with under arbitrary densities with finite fourth-order moments. The following asymptotic linearity result characterizes the asymptotic impact, on $\mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi_1}$ and $\mathbf{\Delta}^{IV}_{\boldsymbol{\vartheta};\phi_1}$, under any elliptical density g_1 with finite fourth-order moments, of estimating $\boldsymbol{\vartheta}$ (see the Appendix for the proof).

LEMMA 4.2. *Let Assumption (A) hold, fix $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and $g_1 \in \mathcal{F}_1^4$, and write $D_k(g_1) := \mu_{k+1;g_1}/\mu_{k-1;g_1}$. Then, for any root-n consistent estimator $\hat{\boldsymbol{\vartheta}} := (\hat{\boldsymbol{\vartheta}}^{I\prime}, \hat{\vartheta}^{II}, \hat{\boldsymbol{\vartheta}}^{III\prime}, \hat{\boldsymbol{\vartheta}}^{IV\prime})'$ of $\boldsymbol{\vartheta}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, both $\mathbf{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi_1} - \mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi_1} + a_k(D_k(g_1)/k) \times \mathbf{\Gamma}^{III}_{\boldsymbol{\vartheta};\phi_1} n^{1/2}(\hat{\boldsymbol{\vartheta}}^{III} - \boldsymbol{\vartheta}^{III})$ and $\mathbf{\Delta}^{IV}_{\hat{\boldsymbol{\vartheta}};\phi_1} - \mathbf{\Delta}^{IV}_{\boldsymbol{\vartheta};\phi_1} + a_k(D_k(g_1)/k)\mathbf{\Gamma}^{IV}_{\boldsymbol{\vartheta};\phi_1} n^{1/2}(\hat{\boldsymbol{\vartheta}}^{IV} - \boldsymbol{\vartheta}^{IV})$ are $o_{\mathrm{P}}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, as $n \to \infty$, where a_k was defined in Section 2.1.*

5. Optimal Gaussian and pseudo-Gaussian tests for principal components.

5.1. *Optimal Gaussian tests for eigenvectors.* For $f_1 = \phi_1$, the test statistic (4.2) takes the form

$$Q^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1} = n \sum_{j=2}^{k} (\boldsymbol{\beta}'_j \mathbf{S}^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1} \boldsymbol{\beta}^0)^2 = n \boldsymbol{\beta}^{0\prime} \mathbf{S}^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1} (\mathbf{I}_k - \boldsymbol{\beta}^0 \boldsymbol{\beta}^{0\prime}) \mathbf{S}^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1} \boldsymbol{\beta}^0, \tag{5.1}$$

with $\mathbf{S}^{(n)}_{\boldsymbol{\vartheta};\phi_1} := \frac{a_k}{n\sigma^2} \sum_{i=1}^n \mathbf{V}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})'\mathbf{V}^{-1/2}$. This statistic still depends on nuisance parameters, to be replaced with estimators. Letting $\mathbf{S}^{(n)} = \frac{1}{n}\sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$, a natural choice for such estimators would be $\hat{\boldsymbol{\theta}} = \bar{\mathbf{X}} := \frac{1}{n}\sum_{i=1}^n \mathbf{X}_i$ and

$$\mathbf{S}^{(n)} =: |\mathbf{S}^{(n)}|^{1/k} \hat{\mathbf{V}} = \left(\frac{|\mathbf{S}^{(n)}|^{1/k}}{\hat{\sigma}^2}\right) \hat{\sigma}^2 \hat{\mathbf{V}} =: \left(\frac{|\mathbf{S}^{(n)}|^{1/k}}{\hat{\sigma}^2}\right) \hat{\sigma}^2 \hat{\boldsymbol{\beta}}_{\mathbf{V}} \hat{\mathbf{\Lambda}}_{\mathbf{V}} \hat{\boldsymbol{\beta}}'_{\mathbf{V}},$$

where $\hat{\mathbf{\Lambda}}_{\mathbf{V}}$ is the diagonal matrix collecting the eigenvalues of $\hat{\mathbf{V}}$ (ranked in decreasing order), $\hat{\boldsymbol{\beta}}_{\mathbf{V}} := (\hat{\boldsymbol{\beta}}_{1;\mathbf{V}}, \ldots, \hat{\boldsymbol{\beta}}_{k;\mathbf{V}})$ is the corresponding matrix of eigenvectors, and $\hat{\sigma}^2$ stands for the empirical median of $d_i^2(\bar{\mathbf{X}}, \hat{\mathbf{V}})$, $i = 1, \ldots, n$. For $\boldsymbol{\beta}$, however, we need a constrained estimator $\tilde{\boldsymbol{\beta}}$ satisfying Assumption (B) for $\mathcal{H}_0^{\boldsymbol{\beta}}$ ($\hat{\boldsymbol{\beta}}_{\mathbf{V}}$ in general does not). Thus, we rather propose estimating $\boldsymbol{\vartheta}$ by

$$\hat{\boldsymbol{\vartheta}} := (\bar{\mathbf{X}}', \hat{\sigma}^2, (\mathrm{d}\overset{\circ}{\mathrm{vec}}\, \hat{\mathbf{\Lambda}}_{\mathbf{V}})', (\mathrm{vec}\, \tilde{\boldsymbol{\beta}}_0)')', \tag{5.2}$$

where $\tilde{\boldsymbol{\beta}}_0 := (\boldsymbol{\beta}^0, \tilde{\boldsymbol{\beta}}_2, \ldots, \tilde{\boldsymbol{\beta}}_k)$ can be obtained from $(\hat{\boldsymbol{\beta}}_{1;\mathbf{V}}, \ldots, \hat{\boldsymbol{\beta}}_{k;\mathbf{V}})$ via the following Gram–Schmidt technique. Let $\tilde{\boldsymbol{\beta}}_2 := (\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime})\hat{\boldsymbol{\beta}}_{2;\mathbf{V}} / \|(\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime})\hat{\boldsymbol{\beta}}_{2;\mathbf{V}}\|$. By construction, $\tilde{\boldsymbol{\beta}}_2$ is the unit-length vector proportional to the projection of the second eigenvector of $\mathbf{S}^{(n)}$ onto the space which is orthogonal to $\boldsymbol{\beta}^0$. Iterating this procedure, define

$$\tilde{\boldsymbol{\beta}}_j = \frac{(\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime} - \sum_{h=2}^{j-1}\tilde{\boldsymbol{\beta}}_h\tilde{\boldsymbol{\beta}}_h')\hat{\boldsymbol{\beta}}_{j;\mathbf{V}}}{\|(\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime} - \sum_{h=2}^{j-1}\tilde{\boldsymbol{\beta}}_h\tilde{\boldsymbol{\beta}}_h')\hat{\boldsymbol{\beta}}_{j;\mathbf{V}}\|}, \qquad j = 3, \ldots, k.$$

The corresponding (constrained) estimator of the scatter $\boldsymbol{\Sigma}$ is $\tilde{\boldsymbol{\Sigma}} := \hat{\sigma}^2\tilde{\mathbf{V}} := \hat{\sigma}^2\tilde{\boldsymbol{\beta}}_0\hat{\boldsymbol{\Lambda}}_{\mathbf{V}}\tilde{\boldsymbol{\beta}}_0'$.

It is easy to see that $\tilde{\boldsymbol{\beta}}_0$, under $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0^{\boldsymbol{\beta}}$, inherits $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$'s root-$n$ consistency, which holds under any elliptical density g_1 with finite fourth-order moments. Lemma 4.2 thus applies. Combining Lemma 4.2 with (4.4) and the fact that

$$\mathbf{G}_k^{\boldsymbol{\beta}_0}\operatorname{diag}(\mathbf{I}_{k-1}, \mathbf{0}_{(k-2)(k-1)/2\times(k-2)(k-1)/2})(\mathbf{G}_k^{\boldsymbol{\beta}_0})'\operatorname{vec}(\tilde{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0) = \mathbf{0}$$

(where $\boldsymbol{\beta}_0$ is the matrix of eigenvectors associated with $\boldsymbol{\vartheta}_0$), one easily obtains that substituting $\hat{\boldsymbol{\vartheta}}$ for $\boldsymbol{\vartheta}_0$ in (5.1) has no asymptotic impact on $Q^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1}$—more precisely, $Q^{(n)}_{\hat{\boldsymbol{\vartheta}};\phi_1} - Q^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1}$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, with $g_1 \in \mathcal{F}_1^4$. It follows that $Q^{(n)}_{\hat{\boldsymbol{\vartheta}};\phi_1}$ shares the same asymptotic optimality properties as $Q^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1}$, irrespective of the value of $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0^{\boldsymbol{\beta}}$. Thus, a locally and asymptotically *most stringent* Gaussian test of $\mathcal{H}_0^{\boldsymbol{\beta}}$—denote it by $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$—can be based on the asymptotic chi-square distribution [with $(k-1)$ degrees of freedom] of

$$\begin{aligned}
Q^{(n)}_{\hat{\boldsymbol{\vartheta}};\phi_1} &= \frac{na_k^2}{\hat{\sigma}^4}\sum_{j=2}^{k}(\tilde{\boldsymbol{\beta}}_j'\tilde{\mathbf{V}}^{-1/2}\mathbf{S}^{(n)}\tilde{\mathbf{V}}^{-1/2}\boldsymbol{\beta}^0)^2 \\
&= \frac{na_k^2}{\hat{\sigma}^4\hat{\lambda}_{1;\mathbf{V}}}\sum_{j=2}^{k}\hat{\lambda}_{j;\mathbf{V}}^{-1}(\tilde{\boldsymbol{\beta}}_j'\mathbf{S}^{(n)}\boldsymbol{\beta}^0)^2 \qquad (5.3)\\
&= \frac{na_k^2|\mathbf{S}^{(n)}|^{2/k}}{\hat{\sigma}^4\lambda_{1;\mathbf{S}}}\sum_{j=2}^{k}\lambda_{j;\mathbf{S}}^{-1}(\tilde{\boldsymbol{\beta}}_j'\mathbf{S}^{(n)}\boldsymbol{\beta}^0)^2 =: Q^{(n)}_{\mathcal{N}}.
\end{aligned}$$

Since $\hat{\sigma}^2/|\mathbf{S}^{(n)}|^{1/k}$ converges to a_k as $n \to \infty$ under the null $\mathcal{H}_0^{\boldsymbol{\beta}}$ and Gaussian densities, one can equivalently use the statistic

$$\bar{Q}^{(n)}_{\mathcal{N}} := \frac{n}{\lambda_{1;\mathbf{S}}}\sum_{j=2}^{k}\lambda_{j;\mathbf{S}}^{-1}(\tilde{\boldsymbol{\beta}}_j'\mathbf{S}^{(n)}\boldsymbol{\beta}^0)^2,$$

which, of course, is still a locally and asymptotically *most stringent* Gaussian test statistic. For results on local powers, we refer to Proposition 5.1.

This test is valid under Gaussian densities only (more precisely, under radial densities with Gaussian kurtosis). On the other hand, it remains valid in case Assumption (A) is weakened [as in Anderson (1963) and Tyler (1981, 1983)] into Assumption (A_1'). Indeed, the consistency of $\tilde{\mathbf{\Sigma}}$ remains unaffected under the null, and $\boldsymbol{\beta}^0$ still is an eigenvector for both $\tilde{\mathbf{\Sigma}}^{-1/2}$ and $\mathbf{\Sigma}$, so that $[\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime}]\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{\Sigma}\tilde{\mathbf{\Sigma}}^{-1/2}\boldsymbol{\beta}^0 = \mathbf{0}$. Hence,

$$
\begin{aligned}
Q_{\mathcal{N}}^{(n)} &= na_k^2 \sum_{j=2}^{k} (\tilde{\boldsymbol{\beta}}_j' \tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{S}^{(n)}\tilde{\mathbf{\Sigma}}^{-1/2}\boldsymbol{\beta}^0)^2 \\
&= na_k^2 \boldsymbol{\beta}^{0\prime}\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{S}^{(n)}\tilde{\mathbf{\Sigma}}^{-1/2}\left(\sum_{j=2}^{k}\tilde{\boldsymbol{\beta}}_j\tilde{\boldsymbol{\beta}}_j'\right)\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{S}^{(n)}\tilde{\mathbf{\Sigma}}^{-1/2}\boldsymbol{\beta}^0 \\
&= na_k^2 \boldsymbol{\beta}^{0\prime}\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{S}^{(n)}\tilde{\mathbf{\Sigma}}^{-1/2}[\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime}]\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{S}^{(n)}\tilde{\mathbf{\Sigma}}^{-1/2}\boldsymbol{\beta}^0 \\
&= na_k^2 \boldsymbol{\beta}^{0\prime}\tilde{\mathbf{\Sigma}}^{-1/2}(\mathbf{S}^{(n)} - a_k^{-1}\mathbf{\Sigma})\tilde{\mathbf{\Sigma}}^{-1/2} \\
&\quad \times [\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime}]\tilde{\mathbf{\Sigma}}^{-1/2}(\mathbf{S}^{(n)} - a_k^{-1}\mathbf{\Sigma})\tilde{\mathbf{\Sigma}}^{-1/2}\boldsymbol{\beta}^0 \\
&= na_k^2 \boldsymbol{\beta}^{0\prime}\mathbf{\Sigma}^{-1/2}(\mathbf{S}^{(n)} - a_k^{-1}\mathbf{\Sigma})\mathbf{\Sigma}^{-1/2} \\
&\quad \times [\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime}]\mathbf{\Sigma}^{-1/2}(\mathbf{S}^{(n)} - a_k^{-1}\mathbf{\Sigma})\mathbf{\Sigma}^{-1/2}\boldsymbol{\beta}^0 + o_{\mathrm{P}}(1),
\end{aligned}
$$

as $n \to \infty$ under $\mathcal{H}_{0;1}^{\boldsymbol{\beta}\prime}$. Since $n^{1/2}a_k\mathbf{\Sigma}^{-1/2}(\mathbf{S}^{(n)} - a_k^{-1}\mathbf{\Sigma})\mathbf{\Sigma}^{-1/2}\boldsymbol{\beta}^0$ is asymptotically $\mathcal{N}(\mathbf{0}, \mathbf{I}_k + \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime})$ as $n \to \infty$ under $\mathcal{H}_{0;1}^{\boldsymbol{\beta}\prime}$ and Gaussian densities, this idempotent quadratic form remains asymptotically chi-square, with $(k-1)$ degrees of freedom, even when (A) is weakened into (A_1'), as was to be shown.

This test is also invariant under the group of transformations $\mathcal{G}_{\mathrm{rot},\circ}$ mapping $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ onto $(\mathbf{O}\mathbf{X}_1 + \mathbf{t}, \ldots, \mathbf{O}\mathbf{X}_n + \mathbf{t})$, where $\mathbf{t}$ is an arbitrary k-vector and $\mathbf{O} \in \mathcal{SO}_k^{\boldsymbol{\beta}^0} := \{\mathbf{O} \in \mathcal{SO}_k | \mathbf{O}\boldsymbol{\beta}^0 = \boldsymbol{\beta}^0\}$, provided that the estimator of $\boldsymbol{\vartheta}_0$ used is equivariant under the same group—which the estimator $\hat{\boldsymbol{\vartheta}}$ proposed in (5.2) is. Indeed, denoting by $Q_{\mathcal{N}}^{(n)}(\mathbf{O}, \mathbf{t})$, $\hat{\boldsymbol{\vartheta}}(\mathbf{O}, \mathbf{t})$, $\mathbf{\Lambda}_{\mathbf{S}}(\mathbf{O}, \mathbf{t})$, $\tilde{\boldsymbol{\beta}}(\mathbf{O}, \mathbf{t})$, $\tilde{\mathbf{\Sigma}}(\mathbf{O}, \mathbf{t})$ and $\mathbf{S}^{(n)}(\mathbf{O}, \mathbf{t})$ the statistics $Q_{\mathcal{N}}^{(n)}$, $\hat{\boldsymbol{\vartheta}}$, $\mathbf{\Lambda}_{\mathbf{S}}$, $\tilde{\boldsymbol{\beta}}$, $\tilde{\mathbf{\Sigma}}$ and $\mathbf{S}^{(n)}$ computed from the transformed sample $(\mathbf{O}\mathbf{X}_1^{(n)} + \mathbf{t}, \ldots, \mathbf{O}\mathbf{X}_n^{(n)} + \mathbf{t})$, one easily checks that, for any $\mathbf{O} \in \mathcal{SO}_k^{\boldsymbol{\beta}^0}$, $\mathbf{\Lambda}_{\mathbf{S}}(\mathbf{O}, \mathbf{t}) = \mathbf{\Lambda}_{\mathbf{S}}$, $\tilde{\boldsymbol{\beta}}(\mathbf{O}, \mathbf{t}) = \mathbf{O}\tilde{\boldsymbol{\beta}}$, $\tilde{\mathbf{\Sigma}}(\mathbf{O}, \mathbf{t}) = \mathbf{O}\tilde{\mathbf{\Sigma}}\mathbf{O}'$ and $\mathbf{S}^{(n)}(\mathbf{O}, \mathbf{t}) = \mathbf{O}\mathbf{S}^{(n)}\mathbf{O}'$, so that (noting that $\mathbf{O}'\boldsymbol{\beta}^0 = \boldsymbol{\beta}^0$)

$$
Q_{\mathcal{N}}^{(n)}(\mathbf{O}, \mathbf{t}) = na_k^2 \sum_{j=2}^{k} (\tilde{\boldsymbol{\beta}}_j'\mathbf{O}'\mathbf{O}\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{O}'\mathbf{O}\mathbf{S}^{(n)}\mathbf{O}'\mathbf{O}\tilde{\mathbf{\Sigma}}^{-1/2}\mathbf{O}'\boldsymbol{\beta}^0)^2 = Q_{\mathcal{N}}^{(n)}.
$$

Finally, let us show that $Q_{\mathrm{Anderson}}^{(n)}$ and $Q_{\mathcal{N}}^{(n)}$ asymptotically coincide, under $\mathcal{H}_{0;1}^{\boldsymbol{\beta}\prime}$ and Gaussian densities, hence also under contiguous alternatives. This asymptotic equivalence indeed is not a straightforward consequence of the definitions (1.1) and (5.3). Since $\sum_{j=2}^{k}\lambda_{j;\mathbf{S}}^{-1}(\boldsymbol{\beta}_{j;\mathbf{S}}\boldsymbol{\beta}_{j;\mathbf{S}}' - \tilde{\boldsymbol{\beta}}_j\tilde{\boldsymbol{\beta}}_j')$ is $o_{\mathrm{P}}(1)$ and $n^{1/2}(\mathbf{S}^{(n)} -$

$\tilde{\boldsymbol\beta}_0\boldsymbol\Lambda_{\mathbf S}\tilde{\boldsymbol\beta}_0')$ is $O_{\mathrm P}(1)$ as $n\to\infty$, under $\mathcal H_{0;1}^{\boldsymbol\beta\prime}$ and Gaussian densities [with $\boldsymbol\Lambda_{\mathbf S}:=\mathrm{diag}(\lambda_{1;\mathbf S},\dots,\lambda_{k;\mathbf S})$], it follows from Slutsky's lemma that

$$\begin{aligned}
Q_{\mathrm{Anderson}}^{(n)} &= \frac{n}{\lambda_{1;\mathbf S}}\sum_{j=2}^{k}\lambda_{j;\mathbf S}^{-1}[(\lambda_{j;\mathbf S}-\lambda_{1;\mathbf S})\boldsymbol\beta_{j;\mathbf S}'\boldsymbol\beta^0]^2\\
&= \frac{n}{\lambda_{1;\mathbf S}}\sum_{j=2}^{k}\lambda_{j;\mathbf S}^{-1}[\boldsymbol\beta_{j;\mathbf S}'(\mathbf S^{(n)}-\tilde{\boldsymbol\beta}_0\boldsymbol\Lambda_{\mathbf S}\tilde{\boldsymbol\beta}_0')\boldsymbol\beta^0]^2\\
&= \frac{n}{\lambda_{1;\mathbf S}}\sum_{j=2}^{k}\lambda_{j;\mathbf S}^{-1}[\tilde{\boldsymbol\beta}_j'(\mathbf S^{(n)}-\tilde{\boldsymbol\beta}_0\boldsymbol\Lambda_{\mathbf S}\tilde{\boldsymbol\beta}_0')\boldsymbol\beta^0]^2+o_{\mathrm P}(1)\\
&= \frac{n}{\lambda_{1;\mathbf S}}\sum_{j=2}^{k}\lambda_{j;\mathbf S}^{-1}(\tilde{\boldsymbol\beta}_j'\mathbf S^{(n)}\boldsymbol\beta^0)^2+o_{\mathrm P}(1)\\
&= \bar Q_{\mathcal N}^{(n)}+o_{\mathrm P}(1)
\end{aligned}$$

as $n\to\infty$, still under $\mathcal H_{0;1}^{\boldsymbol\beta\prime}$ and Gaussian densities. The equivalence between $Q_{\mathrm{Anderson}}^{(n)}$ and $Q_{\mathcal N}^{(n)}$ in the Gaussian case then follows since $\bar Q_{\mathcal N}^{(n)}=Q_{\mathcal N}^{(n)}+o_{\mathrm P}(1)$ as $n\to\infty$, under $\mathcal H_{0;1}^{\boldsymbol\beta\prime}$ and Gaussian densities.

5.2. *Optimal Gaussian tests for eigenvalues.* Turning to $\mathcal H_0^{\boldsymbol\Lambda}$, we now consider the Gaussian version of the test statistic $T_{\boldsymbol\vartheta_0;f_1}^{(n)}$ obtained in Section 4.2. In view of (4.6), we have

$$T_{\boldsymbol\vartheta_0;\phi_1}^{(n)}=n^{1/2}(a_{p,q}(\boldsymbol\Lambda_0))^{-1/2}\mathbf c_{p,q}'\,\mathrm{dvec}\big(\boldsymbol\Lambda_0^{1/2}\boldsymbol\beta'\mathbf S_{\boldsymbol\vartheta_0;\phi_1}^{(n)}\boldsymbol\beta\boldsymbol\Lambda_0^{1/2}\big)\tag{5.4}$$

[recall that $\mathcal J_k(\phi_1)=k(k+2)$; see (2.6)]. Here also we have to estimate $\boldsymbol\vartheta_0$ in order to obtain a genuine test statistic. By using the fact that $\boldsymbol\beta\boldsymbol\Lambda_0\boldsymbol\beta'=\mathbf V_0$ (where all parameter values refer to those in $\boldsymbol\vartheta_0$), we obtain that, in (5.4),

$$\begin{aligned}
&n^{1/2}\mathbf c_{p,q}'\,\mathrm{dvec}\big(\boldsymbol\Lambda_0^{1/2}\boldsymbol\beta'\mathbf S_{\boldsymbol\vartheta_0;\phi_1}^{(n)}\boldsymbol\beta\boldsymbol\Lambda_0^{1/2}\big)\\
&\qquad=\frac{n^{1/2}a_k}{\sigma^2}\mathbf c_{p,q}'\,\mathrm{dvec}\left(\boldsymbol\beta'\frac{1}{n}\sum_{i=1}^{n}(\mathbf X_i-\boldsymbol\theta)(\mathbf X_i-\boldsymbol\theta)'\boldsymbol\beta\right),
\end{aligned}\tag{5.5}$$

a $O_{\mathrm P}(1)$ expression which does not depend on $\boldsymbol\Lambda_0$. In view of Lemma 4.2 and the block-diagonal form of the information matrix, estimation of $\boldsymbol\theta$, σ^2 and $\boldsymbol\beta$ has no asymptotic impact on the eigenvalue part $\boldsymbol\Delta_{\boldsymbol\vartheta;\phi_1}^{III}$ of the central sequence, hence on $T_{\boldsymbol\vartheta_0;\phi_1}^{(n)}$. As for $a_{p,q}(\boldsymbol\Lambda_0)$, it is a continuous function of $\boldsymbol\Lambda_0$, so that, in view of Slutsky's lemma, plain consistency of the estimator of $\boldsymbol\Lambda_0$ is sufficient. Consequently, we safely can use here the unconstrained estimator

$$\hat{\boldsymbol\vartheta}:=(\bar{\mathbf X}',\hat\sigma^2,(\mathrm{d\mathring{v}ec}\,\hat{\boldsymbol\Lambda}_{\mathbf V})',(\mathrm{vec}\,\hat{\boldsymbol\beta}_{\mathbf V})')';\tag{5.6}$$

see the beginning of Section 5.1. Using again the fact that, under Gaussian densities, $\hat{\sigma}^2/|\mathbf{S}^{(n)}|^{1/k}$ converges to a_k as $n \to \infty$, a locally and asymptotically *most powerful* Gaussian test statistic therefore is given by

$$
\begin{aligned}
T_{\mathcal{N}}^{(n)} &:= \frac{n^{1/2} a_k}{\hat{\sigma}^2} (a_{p,q}(\hat{\mathbf{\Lambda}}_{\mathbf{V}}))^{-1/2} \mathbf{c}'_{p,q} \operatorname{dvec}(\hat{\boldsymbol{\beta}}'_{\mathbf{V}} \mathbf{S}^{(n)} \hat{\boldsymbol{\beta}}_{\mathbf{V}}) \\
&= \frac{n^{1/2} a_k |\mathbf{S}^{(n)}|^{1/k}}{\hat{\sigma}^2} (a_{p,q}(\mathbf{\Lambda}_{\mathbf{S}}))^{-1/2} \mathbf{c}'_{p,q} \operatorname{dvec}(\hat{\boldsymbol{\beta}}'_{\mathbf{V}} \mathbf{S}^{(n)} \hat{\boldsymbol{\beta}}_{\mathbf{V}}) \\
&= n^{1/2} (a_{p,q}(\mathbf{\Lambda}_{\mathbf{S}}))^{-1/2} \left((1-p) \sum_{j=q+1}^{k} \lambda_{j;\mathbf{S}} - p \sum_{j=1}^{q} \lambda_{j;\mathbf{S}} \right) \\
&\quad + o_{\mathrm{P}}(1),
\end{aligned}
\tag{5.7}
$$

under Gaussian densities as $n \to \infty$. The corresponding test, $\phi_{\mathbf{\Lambda};\mathcal{N}}^{(n)}$ say, rejects $\mathcal{H}_0^{\mathbf{\Lambda}}$ whenever $T_{\mathcal{N}}^{(n)}$ is smaller than the standard normal α-quantile; (5.7) shows that $T_{\mathcal{N}}^{(n)}$ coincides [up to $o_{\mathrm{P}}(1)$] with $T_{\text{Anderson}}^{(n)}$ given in (1.3), which entails that (i) $\phi_{\mathbf{\Lambda};\text{Anderson}}^{(n)}$ is also locally and asymptotically most powerful under Gaussian densities, and that (ii) the validity of $\phi_{\mathbf{\Lambda};\mathcal{N}}^{(n)}$ extends to $\mathcal{H}_{0;q}^{\mathbf{\Lambda}\prime\prime}$ (since the validity of $\phi_{\mathbf{\Lambda};\text{Anderson}}^{(n)}$ does).

5.3. *Optimal pseudo-Gaussian tests for eigenvectors*. The Gaussian tests $\phi_{\boldsymbol{\beta};\mathcal{N}}^{(n)}$ and $\phi_{\mathbf{\Lambda};\mathcal{N}}^{(n)}$ of Sections 5.1 and 5.2 unfortunately are valid under multinormal densities only (more precisely, as we shall see, under densities with Gaussian kurtosis). It is not difficult, however, to extend their validity to the whole class of elliptical populations with finite fourth-order moments, while maintaining their optimality properties at the multinormal.

Let us first introduce the following notation. For any $g_1 \in \mathcal{F}_1^4$, let (as in Lemma 4.2) $D_k(g_1) := \mu_{k+1;g_1}/\mu_{k-1;g_1} = \sigma^{-2} \mathrm{E}_{\boldsymbol{\vartheta};g_1}[d_i^2(\boldsymbol{\theta}, \mathbf{V})] = \int_0^1 (\tilde{G}_{1k}^{-1}(u))^2 \, du$ and $E_k(g_1) := \sigma^{-4} \mathrm{E}_{\boldsymbol{\vartheta};g_1}[d_i^4(\boldsymbol{\theta}, \mathbf{V})] = \int_0^1 (\tilde{G}_{1k}^{-1}(u))^4 \, du$, where $\tilde{G}_{1k}(r) := (\mu_{k-1;g_1})^{-1} \int_0^r s^{k-1} g_1(s) \, ds$; see Section 2.1. Then

$$
\kappa_k(g_1) := \frac{k}{k+2} \frac{E_k(g_1)}{D_k^2(g_1)} - 1
$$

is the *kurtosis* of the elliptic population with radial density g_1 [see, e.g., page 54 of Anderson (2003)]. For Gaussian densities, $E_k(\phi_1) = k(k+2)/a_k^2$, $D_k(\phi_1) = k/a_k$ and $\kappa_k(\phi_1) = 0$.

Since the asymptotic covariance matrix of $\mathbf{\Delta}_{\boldsymbol{\vartheta};\phi_1}^{IV}$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$ (with $\boldsymbol{\vartheta} \in \mathcal{H}_0^{\boldsymbol{\beta}}$ and $g_1 \in \mathcal{F}_1^4$) is $(a_k^2 E_k(g_1)/k(k+2)) \mathbf{\Gamma}_{\boldsymbol{\vartheta};\phi_1}^{IV}$, it is natural to base our pseudo-Gaussian tests on statistics of the form [compare with the $f_1 = \phi_1$ version of (5.1)]

$$Q^{(n)}_{\boldsymbol{\vartheta}_0,\mathcal{N}*} := \frac{k(k+2)}{a_k^2 E_k(g_1)} \boldsymbol{\Delta}^{IV\prime}_{\boldsymbol{\vartheta}_0;\phi_1}[(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;\phi_1})^- - \mathbf{P}^{\boldsymbol{\beta}_0}_k((\mathbf{P}^{\boldsymbol{\beta}_0}_k)'\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;\phi_1}\mathbf{P}^{\boldsymbol{\beta}_0}_k)^-(\mathbf{P}^{\boldsymbol{\beta}_0}_k)']\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;\phi_1}$$

$$= \bigl(1+\kappa_k(g_1)\bigr)^{-1} \frac{k^2}{D_k^2(g_1)a_k^2} Q^{(n)}_{\boldsymbol{\vartheta}_0,\phi_1} =: \bigl(1+\kappa_k(g_1)\bigr)^{-1} Q_{\boldsymbol{\vartheta}_0,\mathcal{N}}(g_1).$$

As in the Gaussian case, and with the same $\hat{\boldsymbol{\vartheta}}$ as in (5.2), Lemma 4.2 entails that $Q^{(n)}_{\hat{\boldsymbol{\vartheta}},\phi_1} = Q^{(n)}_{\boldsymbol{\vartheta}_0,\phi_1} + o_{\mathrm{P}}(1)$, as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, with $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}}_0$ and $g_1 \in \mathcal{F}^4_1$. Since $\hat{\sigma}^2/|\mathbf{S}^{(n)}|^{1/k}$ consistently estimates $k/D_k(g_1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, with $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}}_0$ and $g_1 \in \mathcal{F}^4_1$, it follows from Slutsky's lemma that

$$\hat{Q}_{\hat{\boldsymbol{\vartheta}},\mathcal{N}} := \frac{\hat{\sigma}^2}{|\mathbf{S}^{(n)}|^{1/k} a_k^2} Q^{(n)}_{\hat{\boldsymbol{\vartheta}},\phi_1}$$

satisfies $\bar{Q}^{(n)}_{\mathcal{N}} = \hat{Q}_{\hat{\boldsymbol{\vartheta}},\mathcal{N}} = Q_{\boldsymbol{\vartheta}_0,\mathcal{N}}(g_1) + o_{\mathrm{P}}(1)$ as $n \to \infty$, still under $\mathcal{H}^{\boldsymbol{\beta}}_0$, $g_1 \in \mathcal{F}^4_1$. The pseudo-Gaussian test $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$ we propose is based on

$$Q^{(n)}_{\mathcal{N}*} := (1+\hat{\kappa}_k)^{-1}\bar{Q}^{(n)}_{\mathcal{N}}, \tag{5.8}$$

where $\hat{\kappa}_k := (kn^{-1}\sum_{i=1}^n \hat{d}_i^4)/((k+2)(n^{-1}\sum_{i=1}^n \hat{d}_i^2)^2) - 1$, with $\hat{d}_i := d_i(\bar{\mathbf{X}}, \mathbf{S}^{(n)})$. The statistic $Q^{(n)}_{\mathcal{N}*}$ indeed remains asymptotically chi-square [$(k-1)$ degrees of freedom] under $\mathcal{H}^{\boldsymbol{\beta}\prime}_{0;1}$ for any $g_1 \in \mathcal{F}^4_1$. Note that $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$ is obtained from $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$ by means of the standard kurtosis correction of Shapiro and Browne (1987), and asymptotically coincides with $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Tyler}}$; see (1.2).

Local powers for $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$ classically follow from applying Le Cam's third lemma. Let $\boldsymbol{\tau}^{(n)} := ((\boldsymbol{\tau}^{I(n)})', \tau^{II(n)}, (\boldsymbol{\tau}^{III(n)})', (\boldsymbol{\tau}^{IV(n)})')'$, with $\boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)}$ uniformly bounded, where $\boldsymbol{\tau}^{IV(n)} = \mathrm{vec}(\mathbf{b}^{(n)})$ is a perturbation of $\mathrm{vec}(\boldsymbol{\beta}_0) = \mathrm{vec}(\boldsymbol{\beta}^0, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_k)$ such that $\boldsymbol{\beta}_0'\mathbf{b}$, with $\mathbf{b} = (\mathbf{b}_1', \ldots, \mathbf{b}_k')' := \lim_{n\to\infty}\mathbf{b}^{(n)}$, is skew-symmetric; see (3.2) and (3.3). Assume furthermore that the corresponding perturbed value of $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}}_0$ does not belong to $\mathcal{H}^{\boldsymbol{\beta}}_0$, that is, $\mathbf{b}_1 \neq \mathbf{0}$, and define

$$\begin{aligned} r^{\boldsymbol{\beta}}_{\boldsymbol{\vartheta}_0;\boldsymbol{\tau}} &:= \lim_{n\to\infty}\bigl(\mathrm{vec}\,\mathbf{b}^{(n)}\bigr)'\mathbf{G}^{\boldsymbol{\beta}_0}_k \,\mathrm{diag}\bigl(\nu_{12}^{-1}, \ldots, \nu_{1k}^{-1}, \mathbf{0}_{1\times(k-2)(k-1)/2}\bigr) \\ &\qquad \times (\mathbf{G}^{\boldsymbol{\beta}_0}_k)'\bigl(\mathrm{vec}\,\mathbf{b}^{(n)}\bigr) \\ &= 4\sum_{j=2}^k \nu_{1j}^{-1}(\boldsymbol{\beta}_j'\mathbf{b}_1)^2. \end{aligned} \tag{5.9}$$

The following result summarizes the asymptotic properties of the pseudo-Gaussian tests $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$. Note that optimality issues involve $\mathcal{H}^{\boldsymbol{\beta}}_0$ [hence require Assumption (A)], while validity extends to $\mathcal{H}^{\boldsymbol{\beta}\prime}_{0;1}$ [which only requires Assumption (A_1')].

PROPOSITION 5.1. (i) $Q^{(n)}_{\mathcal{N}*}$ *is asymptotically chi-square with* $(k-1)$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}^{\boldsymbol{\beta}'}_{0;1}}\bigcup_{g_1\in\mathcal{F}^4_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$, *and asymptotically noncentral chi-square*, *still with* $(k-1)$ *degrees of freedom*, *but with noncentrality parameter* $r^{\boldsymbol{\beta}}_{\boldsymbol{\vartheta};\boldsymbol{\tau}}/4(1+\kappa_k(g_1))$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$, *with* $\boldsymbol{\vartheta}\in\mathcal{H}^{\boldsymbol{\beta}}_0$, $g_1\in\mathcal{F}^4_a$, *and* $\boldsymbol{\tau}^{(n)}$ *as described above*;

(ii) *the sequence of tests* $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$ *rejecting the null whenever* $Q^{(n)}_{\mathcal{N}*}$ *exceeds the* α *upper-quantile* $\chi^2_{k-1;1-\alpha}$ *of the chi-square distribution with* $(k-1)$ *degrees of freedom has asymptotic size* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}^{\boldsymbol{\beta}'}_{0;1}}\bigcup_{g_1\in\mathcal{F}^4_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$;

(iii) *the pseudo-Gaussian tests* $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$ *are asymptotically equivalent*, *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}^{\boldsymbol{\beta}'}_{0;1}}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi_1}\}$ *and contiguous alternatives*, *to the optimal parametric Gaussian tests* $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$; *hence*, *the sequence* $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$ *is locally and asymptotically most stringent*, *still at asymptotic level* α, *for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}^{\boldsymbol{\beta}'}_{0;1}}\bigcup_{g_1\in\mathcal{F}^4_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{H}^{\boldsymbol{\beta}}_0}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi_1}\}$.

Of course, since $\hat{\kappa}_k$ is invariant under $\mathcal{G}_{\mathrm{rot},\circ}$, the pseudo-Gaussian test inherits the $\mathcal{G}_{\mathrm{rot},\circ}$-invariance features of the Gaussian one.

5.4. *Optimal pseudo-Gaussian tests for eigenvalues.* As in the previous section, the asymptotic null distribution of the Gaussian test statistic $T^{(n)}_{\mathcal{N}}$ is not standard normal anymore under radial density g_1 as soon as $\kappa_k(g_1)\neq\kappa_k(\phi_1)$. The Gaussian test $\phi^{(n)}_{\boldsymbol{\Lambda};\mathcal{N}}$ thus is not valid (does not have asymptotic level α) under such densities. The same reasoning as before leads to a similar kurtosis correction, yielding a pseudo-Gaussian test statistic

$$T^{(n)}_{\mathcal{N}*}:=(1+\hat{\kappa}_k)^{-1/2}\tilde{T}^{(n)}_{\mathcal{N}},$$

where $\tilde{T}^{(n)}_{\mathcal{N}}:=n^{1/2}(a_{p,q}(\boldsymbol{\Lambda}_{\mathbf{S}}))^{-1/2}((1-p)\sum_{j=q+1}^{k}\lambda_{j;\mathbf{S}}-p\sum_{j=1}^{q}\lambda_{j;\mathbf{S}})$ and $\hat{\kappa}_k$ is as in Section 5.3. This statistic coincides with $T^{(n)}_{\mathrm{Davis}}$ given in (1.4).

Here also, local powers are readily obtained via Le Cam's third lemma. Let $\boldsymbol{\tau}^{(n)}:=((\boldsymbol{\tau}^{I(n)})',\tau^{II(n)},(\boldsymbol{\tau}^{III(n)})',(\boldsymbol{\tau}^{IV(n)})')'$, with $\boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)}$ uniformly bounded, where $\boldsymbol{\tau}^{III(n)}:=\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}(\mathbf{l}^{(n)})$ is such that $\mathbf{l}:=\lim_{n\to\infty}\mathbf{l}^{(n)}:=\lim_{n\to\infty}\mathrm{diag}(\ell^{(n)}_1,\dots,\ell^{(n)}_k)$ satisfies $\mathrm{tr}(\boldsymbol{\Lambda}^{-1}_{\mathbf{V}}\mathbf{l})=0$ [see (3.1) and the comments thereafter], and define

$$r^{\boldsymbol{\Lambda}_{\mathbf{V}}}_{\boldsymbol{\vartheta};\boldsymbol{\tau}}:=\lim_{n\to\infty}\mathrm{grad}\,h(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}(\boldsymbol{\Lambda}_{\mathbf{V}}))'\boldsymbol{\tau}^{III(n)}=(1-p)\sum_{j=q+1}^{k}\mathbf{l}_j-p\sum_{j=1}^{q}\mathbf{l}_j. \tag{5.10}$$

The following proposition summarizes the asymptotic properties of the resulting pseudo-Gaussian tests $\phi^{(n)}_{\boldsymbol{\Lambda}_{\mathbf{V}};\mathcal{N}*}$.

PROPOSITION 5.2. (i) $T_{\mathcal{N}*}^{(n)}$ *is asymptotically normal, with mean zero under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_{0;q}^{\boldsymbol{\Lambda}''}}\bigcup_{g_1\in\mathcal{F}_1^4}\{\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}\}$, *mean* $(4a_{p,q}(\boldsymbol{\Lambda}_{\mathbf{V}})(1+\kappa_k(g_1))^{-1/2}r_{\boldsymbol{\vartheta};\boldsymbol{\tau}}^{\boldsymbol{\Lambda}_{\mathbf{V}}}$ *under* $\mathrm{P}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}^{(n)}$, $\boldsymbol{\vartheta}\in\mathcal{H}_0^{\boldsymbol{\Lambda}}$, $g_1\in\mathcal{F}_a^4$ *and* $\boldsymbol{\tau}^{(n)}$ *as described above, and variance one under both*;

(ii) *the sequence of tests* $\phi_{\boldsymbol{\Lambda};\mathcal{N}*}^{(n)}$ *rejecting the null whenever* $T_{\mathcal{N}*}^{(n)}$ *is less than the standard normal* α-*quantile* z_α *has asymptotic size* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_{0;q}^{\boldsymbol{\Lambda}''}}\bigcup_{g_1\in\mathcal{F}_1^4}\{\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}\}$;

(iii) *the pseudo-Gaussian tests* $\phi_{\boldsymbol{\Lambda};\mathcal{N}*}^{(n)}$ *are asymptotically equivalent, under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_{0;q}^{\boldsymbol{\Lambda}''}}\{\mathrm{P}_{\boldsymbol{\vartheta};\phi_1}^{(n)}\}$ *and contiguous alternatives, to the optimal parametric Gaussian tests* $\phi_{\boldsymbol{\Lambda};\mathcal{N}}^{(n)}$; *hence, the sequence* $\phi_{\boldsymbol{\Lambda};\mathcal{N}*}^{(n)}$ *is locally and asymptotically most powerful, still at asymptotic level* α, *for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_{0;q}^{\boldsymbol{\Lambda}''}}\bigcup_{g_1\in\mathcal{F}_1^4}\{\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{H}_0^{\boldsymbol{\Lambda}}}\{\mathrm{P}_{\boldsymbol{\vartheta};\phi_1}^{(n)}\}$.

6. Rank-based tests for principal components.

6.1. *Rank-based statistics*: *Asymptotic representation and asymptotic normality.* The parametric tests proposed in Section 4 are valid under specified radial densities f_1 only, and therefore are of limited practical value. The importance of the Gaussian tests of Sections 5.1 and 5.2 essentially follows from the fact that they belong to usual practice, but Gaussian assumptions are quite unrealistic in most applications. The pseudo-Gaussian procedures of Sections 5.3 and 5.4 are more appealing, as they only require finite fourth-order moments. Still, moments of order four may be infinite and, being based on empirical covariances, pseudo-Gaussian procedures remain poorly robust. A straightforward idea would consist in robustifying them by substituting some robust estimate of scatter for empirical covariance matrices. This may take care of validity-robustness issues, but has a negative impact on powers, and would not achieve efficiency-robustness. The picture is quite different with the rank-based procedures we are proposing in this section. While remaining valid under completely arbitrary radial densities, these methods indeed also are efficiency-robust; when based on Gaussian scores, they even uniformly outperform, in the Pitman sense, their pseudo-Gaussian counterparts (see Section 7). Rank-based inference, thus, in this problem as in many others, has much to offer, and enjoys an extremely attractive combination of robustness and efficiency properties.

The natural framework for principal component analysis actually is the semiparametric context of elliptical families in which $\boldsymbol{\theta}$, $\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}(\boldsymbol{\Lambda}_{\mathbf{V}})$, and $\boldsymbol{\beta}$ (not σ^2) are the parameters of interest, while the radial density f [equivalently, the couple (σ^2, f_1)] plays the role of an infinite-dimensional nuisance. This semiparametric model enjoys the double structure considered in Hallin and Werker (2003), which allows for efficient rank-based inference: the fixed-f_1 subexperiments, as shown in Proposition 3.1 are ULAN, while the fixed-($\boldsymbol{\theta}$, $\mathrm{d}\overset{\circ}{\mathrm{vec}}(\boldsymbol{\Lambda}_{\mathbf{V}})$, $\boldsymbol{\beta}$) subexperiments [equivalently, the fixed-($\boldsymbol{\theta}, \mathbf{V}$) subexperiments] are generated by groups of transformations acting on the observation space. Those groups here are of the form $\mathcal{G}^{(n)}_{\boldsymbol{\theta},\mathbf{V}}, \circ$ and consist of the *continuous monotone radial transformations* $\mathcal{G}^{(n)}_h$

$$\begin{aligned}\mathcal{G}^{(n)}_h(\mathbf{X}_1,\dots,\mathbf{X}_n) &= \mathcal{G}^{(n)}_h\big(\boldsymbol{\theta}+d_1(\boldsymbol{\theta},\mathbf{V})\mathbf{V}^{1/2}\mathbf{U}_1(\boldsymbol{\theta},\mathbf{V}),\dots,\\ &\qquad\qquad \boldsymbol{\theta}+d_n(\boldsymbol{\theta},\mathbf{V})\mathbf{V}^{1/2}\mathbf{U}_n(\boldsymbol{\theta},\mathbf{V})\big)\\ &:= \big(\boldsymbol{\theta}+h(d_1(\boldsymbol{\theta},\mathbf{V}))\mathbf{V}^{1/2}\mathbf{U}_1(\boldsymbol{\theta},\mathbf{V}),\dots,\\ &\qquad\qquad \boldsymbol{\theta}+h(d_n(\boldsymbol{\theta},\mathbf{V}))\mathbf{V}^{1/2}\mathbf{U}_n(\boldsymbol{\theta},\mathbf{V})\big),\end{aligned}$$

where $h:\mathbb{R}^+\to\mathbb{R}^+$ is continuous, monotone increasing, and satisfies $\lim_{r\to\infty}h(r)=\infty$ and $h(0)=0$. The group $\mathcal{G}^{(n)}_{\boldsymbol{\theta},\mathbf{V}}, \circ$ generates the fixed-($\boldsymbol{\theta},\mathbf{V}$) family of distributions $\bigcup_{\sigma^2}\bigcup_{f_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathrm{d}\overset{\circ}{\mathrm{vec}}(\boldsymbol{\Lambda}_{\mathbf{V}}),\mathrm{vec}(\boldsymbol{\beta});f_1}\}$. The general results of Hallin and Werker (2003) thus indicate that efficient inference can be based on the corresponding maximal invariants, namely the vectors

$$\big(R^{(n)}_1(\boldsymbol{\theta},\mathbf{V}),\dots,R^{(n)}_n(\boldsymbol{\theta},\mathbf{V}),\mathbf{U}_1(\boldsymbol{\theta},\mathbf{V}),\dots,\mathbf{U}_n(\boldsymbol{\theta},\mathbf{V})\big)$$

of ranks and multivariate signs, where $R^{(n)}_i(\boldsymbol{\theta},\mathbf{V})$ denotes the rank of $d_i(\boldsymbol{\theta},\mathbf{V})$ among $d_1(\boldsymbol{\theta},\mathbf{V}),\dots,d_n(\boldsymbol{\theta},\mathbf{V})$. Test statistics based on such invariants automatically are distribution-free under $\bigcup_{\sigma^2}\bigcup_{f_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathrm{d}\overset{\circ}{\mathrm{vec}}(\boldsymbol{\Lambda}_{\mathbf{V}}),\mathrm{vec}(\boldsymbol{\beta});f_1}\}$.

Letting $R_i := R_i(\boldsymbol{\theta},\mathbf{V})$ and $\mathbf{U}_i := \mathbf{U}_i(\boldsymbol{\theta},\mathbf{V})$, define

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\boldsymbol{\vartheta};K} := \frac{1}{2\sqrt{n}}\mathbf{M}^{\boldsymbol{\Lambda}_{\mathbf{V}}}_k\mathbf{H}_k(\boldsymbol{\Lambda}_{\mathbf{V}}^{-1/2}\boldsymbol{\beta}')^{\otimes 2}\sum_{i=1}^n K\left(\frac{R^{(n)}_i}{n+1}\right)\mathrm{vec}(\mathbf{U}_i\mathbf{U}_i')$$

and

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K} := \frac{1}{2\sqrt{n}}\mathbf{G}^{\boldsymbol{\beta}}_k\mathbf{L}^{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}}_k(\mathbf{V}^{\otimes 2})^{-1/2}\sum_{i=1}^n K\left(\frac{R^{(n)}_i}{n+1}\right)\mathrm{vec}(\mathbf{U}_i\mathbf{U}_i').$$

Associated with $\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\boldsymbol{\vartheta};K}$ and $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K}$, let

$$\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};K,g_1} := \frac{1}{2\sqrt{n}}\mathbf{M}^{\boldsymbol{\Lambda}_{\mathbf{V}}}_k\mathbf{H}_k(\boldsymbol{\Lambda}_{\mathbf{V}}^{-1/2}\boldsymbol{\beta}')^{\otimes 2}\sum_{i=1}^n K\left(\tilde{G}_{1k}\left(\frac{d_i(\boldsymbol{\theta},\mathbf{V})}{\sigma}\right)\right)\mathrm{vec}(\mathbf{U}_i\mathbf{U}_i')$$

and

$$\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta};K,g_1} := \frac{1}{2\sqrt{n}}\mathbf{G}^{\boldsymbol{\beta}}_k\mathbf{L}^{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}}_k(\mathbf{V}^{\otimes 2})^{-1/2}\sum_{i=1}^n K\left(\tilde{G}_{1k}\left(\frac{d_i(\boldsymbol{\theta},\mathbf{V})}{\sigma}\right)\right)\mathrm{vec}(\mathbf{U}_i\mathbf{U}_i'),$$

where $\tilde{G}_{1k}$ is as in Section 5.3. The following proposition provides an asymptotic representation and asymptotic normality result for $\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\boldsymbol{\vartheta};K}$ and $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K}$.

PROPOSITION 6.1. *Let Assumption (S) hold for the score function K. Then*:

(i) (*asymptotic representation*) $(\underset{\sim}{\mathbf{\Delta}}^{III\prime}_{\boldsymbol{\vartheta};K}, \underset{\sim}{\mathbf{\Delta}}^{IV\prime}_{\boldsymbol{\vartheta};K})' = (\mathbf{\Delta}^{III\prime}_{\boldsymbol{\vartheta};K,g_1}, \mathbf{\Delta}^{IV\prime}_{\boldsymbol{\vartheta};K,g_1})' + o_{L^2}(1)$ *as* $n \to \infty$, *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, *for any* $\boldsymbol{\vartheta} \in \mathbf{\Theta}$ *and* $g_1 \in \mathcal{F}_1$;

(ii) (*asymptotic normality*) *let Assumption* (A) *hold and consider a bounded sequence* $\boldsymbol{\tau}^{(n)} := ((\boldsymbol{\tau}^{I(n)})', \tau^{II(n)}, (\boldsymbol{\tau}^{III(n)})', (\boldsymbol{\tau}^{IV(n)})')'$ *such that both* $\boldsymbol{\tau}^{III} := \lim_{n\to\infty} \boldsymbol{\tau}^{III(n)}$ *and* $\boldsymbol{\tau}^{IV} := \lim_{n\to\infty} \boldsymbol{\tau}^{IV(n)}$ *exist. Then* $(\mathbf{\Delta}^{III\prime}_{\boldsymbol{\vartheta};K,g_1}, \mathbf{\Delta}^{IV\prime}_{\boldsymbol{\vartheta};K,g_1})'$ *is asymptotically normal*, *with mean zero and mean*

$$\frac{\mathcal{J}_k(K,g_1)}{k(k+2)} \begin{pmatrix} \mathbf{D}_k(\mathbf{\Lambda}_{\mathbf{V}})\boldsymbol{\tau}^{III} \\ \frac{1}{4}\mathbf{G}_k^{\boldsymbol{\beta}} \operatorname{diag}(\nu_{12}^{-1}, \dots, \nu_{(k-1)k}^{-1})(\mathbf{G}_k^{\boldsymbol{\beta}})'\boldsymbol{\tau}^{IV} \end{pmatrix}$$

[*where* $\mathcal{J}_k(K,g_1)$ *was defined in* (2.5)], *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ (*any* $\boldsymbol{\vartheta} \in \mathbf{\Theta}$ *and* $g_1 \in \mathcal{F}_1$) *and* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$ (*any* $\boldsymbol{\vartheta} \in \mathbf{\Theta}$ *and* $g_1 \in \mathcal{F}_a$), *respectively*, *and block-diagonal covariance matrix* $\operatorname{diag}(\mathbf{\Gamma}^{III}_{\boldsymbol{\vartheta};K}, \mathbf{\Gamma}^{IV}_{\boldsymbol{\vartheta};K})$ *under both*, *with*

$$\mathbf{\Gamma}^{III}_{\boldsymbol{\vartheta};K} := \frac{\mathcal{J}_k(K)}{k(k+2)}\mathbf{D}_k(\mathbf{\Lambda}_{\mathbf{V}})$$

and

$$\mathbf{\Gamma}^{IV}_{\boldsymbol{\vartheta};K} := \frac{\mathcal{J}_k(K)}{4k(k+2)}\mathbf{G}_k^{\boldsymbol{\beta}} \operatorname{diag}(\nu_{12}^{-1}, \dots, \nu_{(k-1)k}^{-1})(\mathbf{G}_k^{\boldsymbol{\beta}})'. \tag{6.1}$$

The proofs of parts (i) and (ii) of this proposition are entirely similar to those of Lemma 4.1 and Proposition 4.1, respectively, in Hallin and Paindaveine (2006a), and therefore are omitted.

In case $K = K_{f_1}$ is the score function associated with $f_1 \in \mathcal{F}_a$, and provided that Assumption (A) holds (in order for the central sequence $\mathbf{\Delta}_{\boldsymbol{\vartheta};f_1}$ of Proposition 3.1 to make sense), $\mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};K_{f_1},f_1}$ and $\mathbf{\Delta}^{IV}_{\boldsymbol{\vartheta};K_{f_1},f_1}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$ clearly coincide with $\mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};f_1}$ and $\mathbf{\Delta}^{IV}_{\boldsymbol{\vartheta};f_1}$. Therefore, $\underset{\sim}{\mathbf{\Delta}}^{III}_{\boldsymbol{\vartheta};K_{f_1}}$ and $\underset{\sim}{\mathbf{\Delta}}^{IV}_{\boldsymbol{\vartheta};K_{f_1}}$ constitute rank-based, hence distribution-free, versions of those central sequence components. Exploiting this, we now construct signed-rank tests for the two problems we are interested in.

6.2. *Optimal rank-based tests for eigenvectors*. Proposition 6.1 provides the theoretical tools for constructing rank-based tests for $\mathcal{H}_0^{\boldsymbol{\beta}}$ and computing their local powers. Letting again $\boldsymbol{\vartheta}_0 := (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\operatorname{dvec}}\, \mathbf{\Lambda}_{\mathbf{V}})', (\operatorname{vec} \boldsymbol{\beta}_0)')'$, with $\boldsymbol{\beta}_0 = (\boldsymbol{\beta}^0, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_k)$, define the rank-based analog of (4.2)

$$\begin{aligned} \underset{\sim}{Q}^{(n)}_{\boldsymbol{\vartheta}_0;K} &:= \underset{\sim}{\mathbf{\Delta}}^{IV\prime}_{\boldsymbol{\vartheta}_0;K}[(\mathbf{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K})^- - \mathbf{P}_k^{\boldsymbol{\beta}_0}((\mathbf{P}_k^{\boldsymbol{\beta}_0})'\mathbf{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K}\mathbf{P}_k^{\boldsymbol{\beta}_0})^-(\mathbf{P}_k^{\boldsymbol{\beta}_0})']\underset{\sim}{\mathbf{\Delta}}^{IV}_{\boldsymbol{\vartheta}_0;K} \\ &= \frac{nk(k+2)}{\mathcal{J}_k(f_1)} \sum_{j=2}^{k} (\boldsymbol{\beta}_j' \underset{\sim}{\mathbf{S}}^{(n)}_{\boldsymbol{\vartheta}_0;K}\boldsymbol{\beta}^0)^2, \end{aligned} \tag{6.2}$$

where $\underset{\sim}{\mathbf{S}}^{(n)}_{\boldsymbol{\vartheta};K} := \frac{1}{n}\sum_{i=1}^{n} K(\frac{R_i^{(n)}(\boldsymbol{\theta},\mathbf{V})}{n+1})\mathbf{U}_i(\boldsymbol{\theta},\mathbf{V})\mathbf{U}_i'(\boldsymbol{\theta},\mathbf{V})$.

In order to turn $\underset{\sim}{Q}^{(n)}_{\boldsymbol{\vartheta}_0;K}$ into a genuine test statistic, as in the parametric case, we still have to replace $\boldsymbol{\vartheta}_0$ with some adequate estimator $\hat{\boldsymbol{\vartheta}}$ satisfying, under as large as possible a class of densities, Assumption (B) for $\mathcal{H}_0^{\boldsymbol{\beta}}$. In particular, root-$n$ consistency should hold without any moment assumptions. Denote by $\hat{\boldsymbol{\theta}}_{\mathrm{HR}}$ the Hettmansperger and Randles (2002) affine-equivariant median, and by $\hat{\mathbf{V}}_{\mathrm{Tyler}}$ the shape estimator of Tyler (1987), normalized so that it has determinant one: both are root-n consistent under any radial density g_1. Factorize $\hat{\mathbf{V}}_{\mathrm{Tyler}}$ into $\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}}\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}\hat{\boldsymbol{\beta}}'_{\mathrm{Tyler}}$. The estimator we are proposing (among many possible ones) is $\hat{\boldsymbol{\vartheta}} = (\hat{\boldsymbol{\theta}}'_{\mathrm{HR}}, \sigma^2, (\mathrm{d}\overset{\circ}{\mathrm{vec}}\,\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}})', (\mathrm{vec}\,\tilde{\boldsymbol{\beta}}_0)')'$, where the constrained estimator $\tilde{\boldsymbol{\beta}}_0 := (\boldsymbol{\beta}^0, \tilde{\boldsymbol{\beta}}_2, \dots, \tilde{\boldsymbol{\beta}}_k)$ is constructed from $\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}}$ via the same Gram–Schmidt procedure as was applied in Section 5.1 to the eigenvectors $\hat{\boldsymbol{\beta}}_{\mathbf{V}}$ of $\hat{\mathbf{V}} := \mathbf{S}^{(n)}/|\mathbf{S}^{(n)}|^{1/k}$; note that σ^2 does not even appear in $\underset{\sim}{Q}^{(n)}_{\boldsymbol{\vartheta}_0;K}$, hence needs not be estimated.

In view of (6.2),

$$
\begin{aligned}
\underset{\sim}{Q}^{(n)}_{K} &= \underset{\sim}{Q}^{(n)}_{\hat{\boldsymbol{\vartheta}};K} \\
&= \frac{nk(k+2)}{\mathcal{J}_k(K)}\boldsymbol{\beta}^{0\prime}\underset{\sim}{\mathbf{S}}^{(n)}_{\hat{\boldsymbol{\vartheta}};K}(\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime})\underset{\sim}{\mathbf{S}}^{(n)}_{\hat{\boldsymbol{\vartheta}};K}\boldsymbol{\beta}^0 \\
&= \frac{nk(k+2)}{\mathcal{J}_k(K)}\big\|[\boldsymbol{\beta}^{0\prime}\otimes(\mathbf{I}_k - \boldsymbol{\beta}^0\boldsymbol{\beta}^{0\prime})](\mathrm{vec}\,\underset{\sim}{\mathbf{S}}^{(n)}_{\hat{\boldsymbol{\vartheta}};K})\big\|^2,
\end{aligned}
\tag{6.3}
$$

where the ranks and signs in $\underset{\sim}{\mathbf{S}}_{\hat{\boldsymbol{\vartheta}};K}$ are computed at $\hat{\boldsymbol{\vartheta}}$, that is, $R_i^{(n)} := R_i^{(n)}(\hat{\boldsymbol{\theta}}_{\mathrm{HR}}, \tilde{\boldsymbol{\beta}}_0\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}\tilde{\boldsymbol{\beta}}_0')$ and $\mathbf{U}_i = \mathbf{U}_i(\hat{\boldsymbol{\theta}}_{\mathrm{HR}}, \tilde{\boldsymbol{\beta}}_0\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}\tilde{\boldsymbol{\beta}}_0')$.

Let us show that substituting $\hat{\boldsymbol{\vartheta}}$ for $\boldsymbol{\vartheta}_0$ in (6.2) has no asymptotic impact on $\underset{\sim}{Q}^{(n)}_{\boldsymbol{\vartheta}_0;K}$—that is, $\underset{\sim}{Q}^{(n)}_{\boldsymbol{\vartheta}_0;K} - \underset{\sim}{Q}^{(n)}_{K} = o_{\mathrm{P}}(1)$ as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, with $g_1 \in \mathcal{F}_a$, $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}\prime}_{0;1}$. The proof, as usual, relies on an asymptotic linearity property which, in turn, requires ULAN. The ULAN property of Proposition 3.1, which was motivated by optimality issues in tests involving $\boldsymbol{\beta}$ and $\boldsymbol{\Lambda}_{\mathbf{V}}$, here cannot help us, as it does not hold under Assumption (A$_1'$). Another ULAN property, however, where Assumption (A) is not required, has been obtained by Hallin and Paindaveine (2006a) for another parametrization—based on $(\boldsymbol{\theta}, \sigma^2, \mathbf{V})$—of the same families of distributions, and perfectly fits our needs here.

Defining $\mathbf{J}_k := (\mathrm{vec}\,\mathbf{I}_k)(\mathrm{vec}\,\mathbf{I}_k)'$ and $\mathbf{J}_k^{\perp} := \mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k$, it follows from Proposition A.1 in Hallin, Oja and Paindaveine (2006) and Lemma 4.4 in Kreiss (1987) that, for any locally asymptotically discrete [Assumption (B3)] and root-n consistent [Assumption (B2)] sequence $(\hat{\boldsymbol{\theta}}^{(n)}, \hat{\mathbf{V}}^{(n)})$ of estimators of location and shape, one has

$$\begin{aligned}
&\mathbf{J}_k^{\perp}\sqrt{n}\,\mathrm{vec}(\underset{\sim}{\mathbf{S}}_{\hat{\boldsymbol\vartheta};K} - \underset{\sim}{\mathbf{S}}_{\boldsymbol\vartheta;K})\\
&\qquad + \frac{\mathcal{J}_k(K,g_1)}{4k(k+2)}\left[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\right](\mathbf{V}^{-1/2})^{\otimes 2} n^{1/2}\,\mathrm{vec}(\hat{\mathbf{V}}^{(n)} - \mathbf{V})\\
&\qquad = o_{\mathrm{P}}(1)
\end{aligned} \tag{6.4}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol\vartheta_0;g_1}$, with $\boldsymbol\vartheta_0 \in \boldsymbol\Theta$ and $g_1 \in \mathcal{F}_a$. This result readily applies to any adequately discretized version of $(\hat{\boldsymbol\theta}_{\mathrm{HR}}, \tilde{\boldsymbol\beta}_0\hat{\boldsymbol\Lambda}_{\mathrm{Tyler}}\tilde{\boldsymbol\beta}_0')$ at $\boldsymbol\vartheta_0 \in \mathcal{H}^{\boldsymbol\beta\prime}_{0;1}$. It is well known, however, that discretization, although necessary for asymptotic statements, is not required in practice [see pages 125 or 188 of Le Cam and Yang (2000) for a discussion on this point]; we therefore do not emphasize discretization any further in the notation, and henceforth assume that $\hat{\boldsymbol\vartheta}$, whenever needed, has been adequately discretized.

Using (6.4) and the fact that $[\boldsymbol\beta^{0\prime} \otimes (\mathbf{I}_k - \boldsymbol\beta^0\boldsymbol\beta^{0\prime})]\mathbf{J}_k = \mathbf{0}$, we obtain, under $\mathrm{P}^{(n)}_{\boldsymbol\vartheta_0;g_1}$ with $\boldsymbol\vartheta_0 \in \mathcal{H}^{\boldsymbol\beta\prime}_{0;1}$ and $g_1 \in \mathcal{F}_a$, since $\mathbf{K}_k(\mathrm{vec}\,\mathbf{A}) = \mathrm{vec}(\mathbf{A}')$ for any $k \times k$ matrix $\mathbf{A}$ and since $\boldsymbol\beta^0$ under $\boldsymbol\vartheta_0 \in \mathcal{H}^{\boldsymbol\beta\prime}_{0;1}$ is an eigenvector of $\mathbf{V}^{-1/2}\tilde{\boldsymbol\beta}_0\hat{\boldsymbol\Lambda}_{\mathrm{Tyler}}\tilde{\boldsymbol\beta}_0'\mathbf{V}^{-1/2}$,

$$\begin{aligned}
&\sqrt{n}[\boldsymbol\beta^{0\prime} \otimes (\mathbf{I}_k - \boldsymbol\beta^0\boldsymbol\beta^{0\prime})]\,\mathrm{vec}(\underset{\sim}{\mathbf{S}}{}^{(n)}_{\hat{\boldsymbol\vartheta};K} - \underset{\sim}{\mathbf{S}}{}^{(n)}_{\boldsymbol\vartheta;K})\\
&\quad = -\frac{\mathcal{J}_k(K,g_1)}{2k(k+2)} n^{1/2}[\boldsymbol\beta^{0\prime} \otimes (\mathbf{I}_k - \boldsymbol\beta^0\boldsymbol\beta^{0\prime})](\mathbf{V}^{-1/2})^{\otimes 2}\,\mathrm{vec}(\tilde{\boldsymbol\beta}_0\hat{\boldsymbol\Lambda}_{\mathrm{Tyler}}\tilde{\boldsymbol\beta}_0' - \mathbf{V})\\
&\qquad + o_{\mathrm{P}}(1)\\
&\quad = -\frac{\mathcal{J}_k(K,g_1)}{2k(k+2)} n^{1/2}\,\mathrm{vec}((\mathbf{I}_k - \boldsymbol\beta^0\boldsymbol\beta^{0\prime})[\mathbf{V}^{-1/2}\tilde{\boldsymbol\beta}_0\hat{\boldsymbol\Lambda}_{\mathrm{Tyler}}\tilde{\boldsymbol\beta}_0'\mathbf{V}^{-1/2} - \mathbf{I}_k]\boldsymbol\beta^0)\\
&\qquad + o_{\mathrm{P}}(1)\\
&\quad = o_{\mathrm{P}}(1),
\end{aligned}$$

as $n \to \infty$. In view of (6.3), we therefore conclude that $\underset{\sim}{Q}{}^{(n)}_K - \underset{\sim}{Q}{}^{(n)}_{\boldsymbol\vartheta_0;K} = o_{\mathrm{P}}(1)$ as $n \to \infty$, still under under $\boldsymbol\vartheta_0 \in \mathcal{H}^{\boldsymbol\beta\prime}_{0;1}$, as was to be shown.

The following result summarizes the results of this section.

PROPOSITION 6.2. *Let Assumption (S) hold for the score function K. Then*:

(i) $\underset{\sim}{Q}{}^{(n)}_K$ *is asymptotically chi-square with $(k-1)$ degrees of freedom under $\bigcup_{\boldsymbol\vartheta \in \mathcal{H}^{\boldsymbol\beta\prime}_{0;1}} \bigcup_{g_1 \in \mathcal{F}_a} \{\mathrm{P}^{(n)}_{\boldsymbol\vartheta;g_1}\}$, and asymptotically noncentral chi-square, still with $(k-1)$ degrees of freedom, and noncentrality parameter*

$$\frac{\mathcal{J}_k^2(K,g_1)}{4k(k+2)\mathcal{J}_k(K)} r^{\boldsymbol\beta}_{\boldsymbol\vartheta;\boldsymbol\tau}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$, *for* $\boldsymbol{\vartheta}\in\mathcal{H}_0^{\boldsymbol{\beta}}$ *and* $g_1\in\mathcal{F}_a$, *with* $\boldsymbol{\tau}^{(n)}$ *as in Proposition* 5.1 *and* $r^{\boldsymbol{\beta}}_{\boldsymbol{\vartheta};\boldsymbol{\tau}}$ *defined in* (5.9);

(ii) *the sequence of tests* $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K}$ *rejecting the null when* $\underset{\sim}{Q}^{(n)}_K$ *exceeds the* α *upper-quantile of the chi-square distribution with* $(k-1)$ *degrees of freedom has asymptotic size* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0^{\boldsymbol{\beta}'}}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$;

(iii) *for scores* $K=K_{f_1}$, *with* $f_1\in\mathcal{F}_a$, $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K}$ *is locally asymptotically most stringent, at asymptotic level* α, *for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0^{\boldsymbol{\beta}'}}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{H}_0^{\boldsymbol{\beta}}}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}\}$.

Being measurable with respect to signed-ranks, $\underset{\sim}{Q}^{(n)}_K$ is asymptotically invariant under continuous monotone radial transformations, in the sense that it is asymptotically equivalent (in probability) to a random variable that is strictly invariant under such transformations. Furthermore, it is easy to show that it enjoys the same $\mathcal{G}_{\mathrm{rot},\circ}$-invariance features as the parametric, Gaussian, or pseudo-Gaussian test statistics.

6.3. *Optimal rank-based tests for eigenvalues*. Finally, still from the results of Proposition 6.1, we construct signed-rank tests for the null hypothesis $\mathcal{H}_0^{\boldsymbol{\Lambda}}$. A rank-based counterpart of (4.5) and (4.6) [at $\boldsymbol{\vartheta}_0=(\boldsymbol{\theta}',\sigma^2,(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_0)',(\mathrm{vec}\,\boldsymbol{\beta})')'\in\mathcal{H}_0^{\boldsymbol{\Lambda}}$] is, writing $\mathbf{V}_0$ for $\boldsymbol{\beta}\boldsymbol{\Lambda}_0\boldsymbol{\beta}'$,

$$
\begin{aligned}
\underset{\sim}{T}^{(n)}_{\boldsymbol{\vartheta}_0;K} &= (\mathrm{grad}'\,h(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_0)(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta}_0;K})^{-1}\,\mathrm{grad}\,h(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_0))^{-1/2}\\
&\qquad\times \mathrm{grad}'\,h(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_0)(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta}_0;K})^{-1}\underset{\sim}{\boldsymbol{\Delta}}^{III}_{\boldsymbol{\vartheta}_0;K} \qquad (6.5)\\
&= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\boldsymbol{\Lambda}_0))^{-1/2}\mathbf{c}'_{p,q}\,\mathrm{dvec}\big(\boldsymbol{\Lambda}_0^{1/2}\boldsymbol{\beta}'\underset{\sim}{\mathbf{S}}^{(n)}_{\boldsymbol{\vartheta}_0,K}\boldsymbol{\beta}\boldsymbol{\Lambda}_0^{1/2}\big).
\end{aligned}
$$

Here again, we have to estimate $\boldsymbol{\vartheta}_0$. Note that, unlike the quantity $\mathrm{grad}'\,h(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_0)(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta}_0;\phi_1})^{-1}\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta}_0;\phi_1}$ appearing in the Gaussian or pseudo-Gaussian cases, $\mathrm{grad}'\,h(\mathrm{d}\mathring{\mathrm{v}}\mathrm{ec}\,\boldsymbol{\Lambda}_0)(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta}_0;K})^{-1}\underset{\sim}{\boldsymbol{\Delta}}^{III}_{\boldsymbol{\vartheta}_0;K}$ does depend on $\boldsymbol{\Lambda}_0$ [see the comments below (5.4)]. Consequently, we have to carefully select an estimator $\hat{\boldsymbol{\vartheta}}$ that has no influence on the asymptotic behavior of $\underset{\sim}{T}^{(n)}_{\boldsymbol{\vartheta}_0;K}$ under $\mathcal{H}^{\boldsymbol{\Lambda}\prime\prime}_{0;q}$.

To this end, consider Tyler's estimator of shape $\hat{\mathbf{V}}_{\mathrm{Tyler}}(=:\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}}\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}\hat{\boldsymbol{\beta}}'_{\mathrm{Tyler}}$, with obvious notation) and define

$$
\mathrm{dvec}(\tilde{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}) := \big(\mathbf{I}_k-\mathbf{c}_{p,q}(\mathbf{c}'_{p,q}\mathbf{c}_{p,q})^{-1}\mathbf{c}'_{p,q}\big)(\mathrm{dvec}\,\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}}).
$$

Then the estimator of shape $\tilde{\mathbf{\Lambda}}_{\mathbf{V}} := \tilde{\mathbf{\Lambda}}_{\text{Tyler}}/|\tilde{\mathbf{\Lambda}}_{\text{Tyler}}|^{1/k}$ is clearly constrained: $\mathbf{c}'_{p,q}(\text{dvec}\,\tilde{\mathbf{\Lambda}}_{\mathbf{V}}) = 0$ and $|\tilde{\mathbf{\Lambda}}_{\mathbf{V}}| = 1$. The resulting preliminary estimator $\hat{\boldsymbol{\vartheta}}$ is

$$\hat{\boldsymbol{\vartheta}} := (\hat{\boldsymbol{\theta}}'_{\text{HR}}, \sigma^2, (\overset{\circ}{\text{dvec}}\,\tilde{\mathbf{\Lambda}}_{\mathbf{V}})', (\text{vec}\,\hat{\boldsymbol{\beta}}_{\text{Tyler}})')', \tag{6.6}$$

where $\hat{\boldsymbol{\theta}}_{\text{HR}}$ still denotes the Hettmansperger and Randles (2002) affine-equivariant median. The test statistic we propose is then

$$\begin{aligned}
\underset{\sim}{T}_K^{(n)} &:= \underset{\sim}{T}_{\hat{\boldsymbol{\vartheta}};K}^{(n)} \\
&= (\text{grad}'\, h(\overset{\circ}{\text{dvec}}\,\tilde{\mathbf{\Lambda}}_{\mathbf{V}})(\mathbf{\Gamma}^{III}_{\hat{\boldsymbol{\vartheta}};K})^{-1}\,\text{grad}\, h(\overset{\circ}{\text{dvec}}\,\tilde{\mathbf{\Lambda}}_{\mathbf{V}}))^{-1/2} \\
&\qquad \times \text{grad}'\, h(\overset{\circ}{\text{dvec}}\,\tilde{\mathbf{\Lambda}}_{\mathbf{V}})(\mathbf{\Gamma}^{III}_{\hat{\boldsymbol{\vartheta}};K})^{-1}\underset{\sim}{\mathbf{\Delta}}^{III}_{\hat{\boldsymbol{\vartheta}};K} \\
&= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\tilde{\mathbf{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q} \\
&\qquad \times \text{dvec}\big(\tilde{\mathbf{\Lambda}}_{\mathbf{V}}^{1/2}\hat{\boldsymbol{\beta}}'_{\text{Tyler}}\underset{\sim}{\mathbf{S}}^{(n)}_{\hat{\boldsymbol{\vartheta}};K}\hat{\boldsymbol{\beta}}_{\text{Tyler}}\tilde{\mathbf{\Lambda}}_{\mathbf{V}}^{1/2}\big),
\end{aligned} \tag{6.7}$$

where $\underset{\sim}{\mathbf{S}}^{(n)}_{\hat{\boldsymbol{\vartheta}};K} := \frac{1}{n}\sum_{i=1}^n K(\frac{\hat{R}_i^{(n)}}{n+1})\hat{\mathbf{U}}_i\hat{\mathbf{U}}_i'$, with $\hat{R}_i^{(n)} := R_i^{(n)}(\hat{\boldsymbol{\theta}}_{\text{HR}}, \hat{\boldsymbol{\beta}}_{\text{Tyler}}\tilde{\mathbf{\Lambda}}_{\mathbf{V}}\hat{\boldsymbol{\beta}}'_{\text{Tyler}})$ and $\hat{\mathbf{U}}_i := \mathbf{U}_i(\hat{\boldsymbol{\theta}}_{\text{HR}}, \hat{\boldsymbol{\beta}}_{\text{Tyler}}\tilde{\mathbf{\Lambda}}_{\mathbf{V}}\hat{\boldsymbol{\beta}}'_{\text{Tyler}})$. The following lemma shows that the substitution of $\hat{\boldsymbol{\vartheta}}$ for $\boldsymbol{\vartheta}$ in (6.6) has no asymptotic effect on $\underset{\sim}{T}_{\boldsymbol{\vartheta};K}^{(n)}$ (see the Appendix for a proof).

LEMMA 6.1. *Fix $\boldsymbol{\vartheta} \in \mathcal{H}_{0;q}^{\mathbf{\Lambda}''}$ and $g_1 \in \mathcal{F}_a$, and let $\hat{\boldsymbol{\vartheta}}$ be the estimator in* (6.6). *Then $\underset{\sim}{T}_K^{(n)} - \underset{\sim}{T}_{\boldsymbol{\vartheta};K}^{(n)}$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$.*

The following result summarizes the results of this section.

PROPOSITION 6.3. *Let Assumption* (S) *hold for the score function K. Then*:

(i) *$\underset{\sim}{T}_K^{(n)}$ is asymptotically standard normal under $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_{0;q}^{\mathbf{\Lambda}''}}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$, and asymptotically normal with mean*

$$\frac{\mathcal{J}_k(K, g_1)}{\sqrt{4k(k+2)a_{p,q}(\mathbf{\Lambda}_{\mathbf{V}})\mathcal{J}_k(K)}}r^{\mathbf{\Lambda}_{\mathbf{V}}}_{\boldsymbol{\vartheta};\boldsymbol{\tau}}$$

and variance 1 *under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$, with $\boldsymbol{\vartheta} \in \mathcal{H}_0^{\mathbf{\Lambda}_{\mathbf{V}}}$, $g_1 \in \mathcal{F}_a$, $\boldsymbol{\tau}^{(n)}$ as in Proposition* 5.2, *and $r^{\mathbf{\Lambda}_{\mathbf{V}}}_{\boldsymbol{\vartheta};\boldsymbol{\tau}}$ defined in* (5.10);

(ii) *the sequence of tests* $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};K}$ *rejecting the null whenever* $\underset{\sim}{T}^{(n)}_{K}$ *is less than the standard normal* α*-quantile* z_α *has asymptotic size* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}^{\mathbf{\Lambda}''}_{0;q}}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$;

(iii) *for scores* $K = K_{f_1}$ *with* $f_1 \in \mathcal{F}_a$, *the sequence of tests* $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};K}$ *is locally and asymptotically most powerful*, *still at asymptotic level* α, *for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}^{\mathbf{\Lambda}''}_{0;q}}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{H}^{\mathbf{\Lambda}}_{0}}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}\}$.

7. Asymptotic relative efficiencies. The asymptotic relative efficiencies (AREs) of the rank-based tests of Section 6 with respect to their Gaussian and pseudo-Gaussian competitors of Sections 5 are readily obtained as ratios of noncentrality parameters under local alternatives (squared ratios of standardized asymptotic shifts for the one-sided problems on eigenvalues). Denoting by $\mathrm{ARE}^{\boldsymbol{\vartheta},\boldsymbol{\tau}}_{k,g_1}(\phi^{(n)}_1/\phi^{(n)}_2)$ the ARE, under local alternatives of the form $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau};g_1}$, of a sequence of tests $\phi^{(n)}_1$ with respect to the sequence $\phi^{(n)}_2$, we thus have the following result.

PROPOSITION 7.1. *Let Assumptions* (S) *and* (B) *hold for the score function* K *and* (*with the appropriate null hypotheses and densities*) *for the estimators* $\hat{\boldsymbol{\vartheta}}$ *described in the previous sections*. *Then*, *for any* $g_1 \in \mathcal{F}^4_a$,

$$\mathrm{ARE}^{\boldsymbol{\vartheta},\boldsymbol{\tau}}_{k,g_1}(\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K}/\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}) = \mathrm{ARE}^{\boldsymbol{\vartheta},\boldsymbol{\tau}}_{k,g_1}(\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};K}/\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*}) := \frac{(1+\kappa_k(g_1))\mathcal{J}^2_k(K,g_1)}{k(k+2)\mathcal{J}_k(K)}.$$

Table 1 provides numerical values of these AREs for various values of the space dimension k and selected radial densities g_1 (Student, Gaussian and power-exponential), and for the van der Waerden tests $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{vdW}}$ and $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{vdW}}$, the Wilcoxon tests $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K_1}$ and $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};K_1}$, and the Spearman tests $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K_2}$ and $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};K_2}$ (the score functions K_a, $a > 0$ were defined in Section 2.2). These values coincide with the "AREs for shape" obtained in Hallin and Paindaveine (2006a), which implies [Paindaveine (2006)] that the AREs of van der Waerden tests with respect to their pseudo-Gaussian counterparts are uniformly larger than or equal to one (an extension of the classical Chernoff–Savage property):

$$\inf_{g_1}\mathrm{ARE}^{\boldsymbol{\vartheta},\boldsymbol{\tau}}_{k,g_1}(\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{vdW}}/\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}) = \inf_{g_1}\mathrm{ARE}^{\boldsymbol{\vartheta},\boldsymbol{\tau}}_{k,g_1}(\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{vdW}}/\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*}) = 1.$$

8. Simulations. In this section, we investigate via simulations the finite-sample performances of the following tests:

(i) the Anderson test $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Anderson}}$, the optimal Gaussian test $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$, the pseudo-Gaussian test $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$, the robust test $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Tyler}}$ based on $Q^{(n)}_{\mathrm{Tyler}}$, and various rank-based tests $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K}$ (with van der Waerden, Wilcoxon, Spearman and sign scores,

TABLE 1

AREs of the van der Waerden (vdW), Wilcoxon (W) and Spearman (SP) rank-based tests $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K}$ and $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\Lambda};K}$ with respect to their pseudo-Gaussian counterparts, under k-dimensional Student (with 5, 8 and 12 degrees of freedom), Gaussian, and power-exponential densities (with parameter $\eta = 2, 3, 5$), for $k = 2, 3, 4, 6, 10$, and $k \to \infty$

		Underlying density						
K	**k**	**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
W	2	2.258	1.174	1.001	0.844	0.789	0.804	0.842
	3	2.386	1.246	1.068	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

but also with scores achieving optimality at t_1, t_3 and t_5 densities), all for the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}}$ on eigenvectors;

(ii) the optimal Anderson test $\phi^{(n)}_{\boldsymbol{\Lambda};\text{Anderson}} = \phi^{(n)}_{\boldsymbol{\Lambda};\mathcal{N}}$, the pseudo-Gaussian test $\phi^{(n)}_{\boldsymbol{\Lambda};\mathcal{N}*} = \phi^{(n)}_{\boldsymbol{\Lambda};\text{Davis}}$ based on $T^{(n)}_{\text{Davis}}$, and various rank-based tests $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\Lambda};K}$ (still with van der Waerden, Wilcoxon, Spearman, sign, t_1, t_3 and t_5 scores), for the null hypothesis $\mathcal{H}_0^{\boldsymbol{\Lambda}}$ on eigenvalues.

Simulations were conducted as follows. We generated $N = 2500$ mutually independent samples of i.i.d. trivariate ($k = 3$) random vectors $\boldsymbol{\varepsilon}_{\ell;j}$, $\ell = 1, 2, 3, 4$, $j = 1, \ldots, n = 100$, with spherical Gaussian ($\boldsymbol{\varepsilon}_{1;j}$), t_5 ($\boldsymbol{\varepsilon}_{2;j}$), t_3 ($\boldsymbol{\varepsilon}_{3;j}$) and t_1 ($\boldsymbol{\varepsilon}_{4;j}$) densities, respectively. Letting

$$\boldsymbol{\Lambda} := \begin{pmatrix} 10 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{B}_\xi := \begin{pmatrix} \cos(\pi\xi/12) & -\sin(\pi\xi/12) & 0 \\ \sin(\pi\xi/12) & \cos(\pi\xi/12) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\mathbf{L}_\xi := \begin{pmatrix} 3\xi & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

each $\boldsymbol{\varepsilon}_{\ell;j}$ was successively transformed into

$$\mathbf{X}_{\ell;j;\xi} = \mathbf{B}_\xi \boldsymbol{\Lambda}^{1/2} \boldsymbol{\varepsilon}_{\ell;j}, \qquad \ell = 1,2,3,4, j = 1,\dots,n, \xi = 0,\dots,3, \tag{8.1}$$

and

$$\mathbf{Y}_{\ell;j;\xi} = (\boldsymbol{\Lambda} + \mathbf{L}_\xi)^{1/2} \boldsymbol{\varepsilon}_{\ell;j}, \qquad \ell = 1,2,3,4, j = 1,\dots,n, \xi = 0,\dots,3. \tag{8.2}$$

The value $\xi = 0$ corresponds to the null hypothesis $\mathcal{H}_0^{\boldsymbol{\beta}} : \boldsymbol{\beta}_1 = (1,0,0)'$ for the $\mathbf{X}_{\ell;j;\xi}$'s and the null hypothesis $\mathcal{H}_0^{\boldsymbol{\Lambda}} : \sum_{j=q+1}^k \lambda_{j;\mathbf{V}} / \sum_{j=1}^k \lambda_{j;\mathbf{V}} = 1/3$ (with $q = 1$ and $k = 3$) for the $\mathbf{Y}_{\ell;j;\xi}$'s; $\xi = 1,2,3$ characterizes increasingly distant alternatives. We then performed the tests listed under (i) and (ii) above in $N = 2500$ independent replications of such samples. Rejection frequencies are reported in Table 2 for $\mathcal{H}_0^{\boldsymbol{\beta}}$ and in Table 3 for $\mathcal{H}_0^{\boldsymbol{\Lambda}}$.

Inspection of Table 2 confirms our theoretical results. Anderson's $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Anderson}}$ meets the level constraint at Gaussian densities only; $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Tyler}}$ (equivalently, $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$) further survives the t_5 but not the t_3 or t_1 densities which have infinite fourth-order moments. In contrast, the rank-based tests for eigenvectors throughout satisfy the nominal asymptotic level condition (a 95% confidence interval here has half-width 0.0085). Despite the relatively small sample size $n = 100$, empirical power and ARE rankings almost perfectly agree.

The results for eigenvalues, shown in Table 3, are slightly less auspicious. While the Gaussian and pseudo-Gaussian tests remain hopelessly sensitive to the violations of Gaussian and fourth-order moments, respectively, the rank tests, when based on asymptotic critical values, all significantly overreject, indicating that asymptotic conditions are not met for $n = 100$. We therefore propose an alternative construction for critical values. Lemma 6.1 indeed implies that the asymptotic distribution of the test statistic $\underset{\sim}{T}{}^{(n)}_K$ (based on the ranks and signs of estimated residuals) is the same, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\Lambda}\prime\prime}_{0;q}$, as that of $\underset{\sim}{T}{}^{(n)}_{\boldsymbol{\vartheta}_0,K}$ (based on the ranks and signs of *exact* residuals, which are distribution-free). The latter distribution can be simulated, and its simulated quantiles provide valid approximations of the exact ones. The following critical values were obtained from $M = 100{,}000$ replications: -1.7782 for van der Waerden, -1.8799 for t_5-scores, -1.8976 for t_3-scores, -1.9439 for t_1-scores, -1.9320 for sign scores, -1.8960 for Wilcoxon and -1.8229 for Spearman. Note that they all are smaller than -1.645, which is consistent with the overrejection phenomenon. The corresponding rejection frequencies are reported in parentheses in Table 3. They all are quite close to the

TABLE 2

Rejection frequencies (out of $N = 2500$ replications), under the null $\mathcal{H}_0^{\boldsymbol{\beta}}$ and increasingly distant alternatives (see Section 8 for details), of the Anderson test $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Anderson}}$, the Tyler test $\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Tyler}}$, the parametric Gaussian test $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$, its pseudo-Gaussian version $\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$, and the signed-rank tests with van der Waerden, t_ν ($\nu = 1, 3, 5$), sign, Wilcoxon, and Spearman scores, $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{vdW}}$, $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,\nu}}$, $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{S}}$, $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{W}}$, and $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{SP}}$, respectively. Sample size is $n = 100$. All tests were based on asymptotic 5% critical values

	ξ							
Test	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
	$\mathcal{N}$				t_5			
$\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Anderson}}$	0.0572	0.3964	0.8804	0.9852	0.2408	0.4940	0.8388	0.9604
$\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$	0.0528	0.3724	0.8568	0.9752	0.2284	0.4716	0.8168	0.9380
$\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Tyler}}$	0.0572	0.3908	0.8740	0.9856	0.0612	0.2520	0.6748	0.8876
$\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$	0.0524	0.3648	0.8512	0.9740	0.0544	0.2188	0.6056	0.8156
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{vdW}}$	0.0368	0.2960	0.8032	0.9608	0.0420	0.2328	0.6908	0.9056
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,5}}$	0.0452	0.3204	0.8096	0.9596	0.0476	0.2728	0.7440	0.9284
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,3}}$	0.0476	0.3104	0.7988	0.9532	0.0496	0.2760	0.7476	0.9280
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,1}}$	0.0488	0.2764	0.7460	0.9220	0.0552	0.2652	0.7184	0.9024
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{S}}$	0.0448	0.2268	0.6204	0.8392	0.0496	0.2164	0.6236	0.8324
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{W}}$	0.0456	0.3144	0.8012	0.9556	0.0484	0.2808	0.7464	0.9320
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{SP}}$	0.0444	0.3096	0.8160	0.9576	0.0464	0.2548	0.7068	0.9152
	t_3				t_1			
$\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Anderson}}$	0.4772	0.6300	0.8532	0.9452	0.9540	0.9580	0.9700	0.9740
$\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}}$	0.4628	0.6040	0.8304	0.9168	0.9320	0.9384	0.9472	0.9480
$\phi^{(n)}_{\boldsymbol{\beta};\mathrm{Tyler}}$	0.0892	0.2248	0.5364	0.7508	0.5704	0.5980	0.6584	0.7444
$\phi^{(n)}_{\boldsymbol{\beta};\mathcal{N}*}$	0.0616	0.1788	0.4392	0.6092	0.4516	0.4740	0.5160	0.5624
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{vdW}}$	0.0444	0.2172	0.6464	0.8676	0.0472	0.1656	0.5104	0.7720
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,5}}$	0.0488	0.2628	0.7120	0.9076	0.0560	0.2100	0.6068	0.8508
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,3}}$	0.0500	0.2728	0.7156	0.9116	0.0576	0.2156	0.6292	0.8672
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};t_{1,1}}$	0.0476	0.2688	0.7100	0.9084	0.0548	0.2256	0.6600	0.8856
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{S}}$	0.0492	0.2202	0.6188	0.8352	0.0512	0.2116	0.6172	0.8448
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{W}}$	0.0520	0.2708	0.7136	0.9120	0.0552	0.2148	0.6148	0.8604
$\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};\mathrm{SP}}$	0.0544	0.2436	0.6648	0.8776	0.0580	0.1824	0.5200	0.7740

TABLE 3
Rejection frequencies (out of $N = 2500$ replications), under the null $\mathcal{H}_0^{\mathbf{\Lambda}}$ and increasingly distant alternatives (see Section 8), of the optimal Gaussian test $\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}} = \phi^{(n)}_{\mathbf{\Lambda};\mathrm{Anderson}}$, its pseudo-Gaussian version $\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*} = \phi^{(n)}_{\mathbf{\Lambda};\mathrm{Davis}}$, and the signed-rank tests with van der Waerden, t_ν ($\nu = 1, 3, 5$), sign, Wilcoxon, and Spearman scores $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{vdW}}$, $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,\nu}}$, $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{S}}$, $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{W}}$, $\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{SP}}$. Sample size is $n = 100$. All tests were based on asymptotic 5% critical values and (in parentheses) simulated ones

Test	ξ			
	0	**1**	**2**	**3**
	$\mathcal{N}$			
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}} = \phi^{(n)}_{\mathbf{\Lambda};\mathrm{Anderson}}$	0.0460	0.4076	0.8308	0.9604
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*} = \phi^{(n)}_{\mathbf{\Lambda};\mathrm{Davis}}$	0.0432	0.3976	0.8220	0.9572
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{vdW}}$	0.0608 (0.0480)	0.4604 (0.4116)	0.8576 (0.8280)	0.9668 (0.9596)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,5}}$	0.0728 (0.0480)	0.4804 (0.3972)	0.8572 (0.8116)	0.9644 (0.9504)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,3}}$	0.0748 (0.0496)	0.4804 (0.3884)	0.8524 (0.7964)	0.9612 (0.9432)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,1}}$	0.0780 (0.0504)	0.4532 (0.3572)	0.8160 (0.7320)	0.9448 (0.9112)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{S}}$	0.0864 (0.0508)	0.3980 (0.3088)	0.7384 (0.6408)	0.9028 (0.8552)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{W}}$	0.0744 (0.0480)	0.4816 (0.3908)	0.8544 (0.8012)	0.9640 (0.9464)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{SP}}$	0.0636 (0.0460)	0.4664 (0.4096)	0.8564 (0.8200)	0.9668 (0.9584)
	t_5			
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}} = \phi^{(n)}_{\mathbf{\Lambda};\mathrm{Anderson}}$	0.1432	0.4624	0.7604	0.9180
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*} = \phi^{(n)}_{\mathbf{\Lambda};\mathrm{Davis}}$	0.0504	0.2768	0.5732	0.7988
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{vdW}}$	0.0692 (0.0548)	0.4256 (0.3772)	0.7720 (0.7404)	0.9444 (0.9324)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,5}}$	0.0736 (0.0492)	0.4544 (0.3772)	0.7980 (0.7372)	0.9524 (0.9332)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,3}}$	0.0732 (0.0452)	0.4576 (0.3748)	0.7968 (0.7320)	0.9524 (0.9288)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,1}}$	0.0776 (0.0416)	0.4448 (0.3484)	0.7832 (0.6952)	0.9436 (0.9116)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{S}}$	0.0768 (0.0436)	0.4060 (0.3172)	0.7180 (0.6360)	0.9100 (0.8592)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{W}}$	0.0732 (0.0456)	0.4512 (0.3756)	0.7972 (0.7364)	0.9524 (0.9308)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\mathrm{SP}}$	0.0764 (0.0544)	0.4360 (0.3736)	0.7776 (0.7304)	0.9480 (0.9300)

nominal probability level $\alpha = 5\%$, while empirical powers are in line with theoretical ARE values.

TABLE 3
(*Continued.*)

Test	ξ 0	1	2	3
	t_3			
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}} = \phi^{(n)}_{\mathbf{\Lambda};\text{Anderson}}$	0.2572	0.5308	0.7200	0.8596
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*} = \phi^{(n)}_{\mathbf{\Lambda};\text{Davis}}$	0.0368	0.1788	0.3704	0.5436
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{vdW}}$	0.0708 (0.0560)	0.4088 (0.3260)	0.7540 (0.7040)	0.9304 (0.9140)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,5}}$	0.0812 (0.0544)	0.4472 (0.3524)	0.7936 (0.7240)	0.9416 (0.9208)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,3}}$	0.0832 (0.0560)	0.4556 (0.3568)	0.7944 (0.7256)	0.9452 (0.9192)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,1}}$	0.0924 (0.0548)	0.4464 (0.3400)	0.7812 (0.7024)	0.9364 (0.8996)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{S}}$	0.0936 (0.0604)	0.4104 (0.2928)	0.7320 (0.6404)	0.9012 (0.8528)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{W}}$	0.0832 (0.0572)	0.4488 (0.3580)	0.7956 (0.7272)	0.9448 (0.9180)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{SP}}$	0.0796 (0.0576)	0.4212 (0.3412)	0.7572 (0.7020)	0.9276 (0.9044)
	t_1			
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}} = \phi^{(n)}_{\mathbf{\Lambda};\text{Anderson}}$	0.7488	0.8000	0.8288	0.8528
$\phi^{(n)}_{\mathbf{\Lambda};\mathcal{N}*} = \phi^{(n)}_{\mathbf{\Lambda};\text{Davis}}$	0.0072	0.0080	0.0172	0.0296
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{vdW}}$	0.0724 (0.0596)	0.3500 (0.3032)	0.6604 (0.6176)	0.8600 (0.8332)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,5}}$	0.0824 (0.0512)	0.3836 (0.3120)	0.7312 (0.6492)	0.9036 (0.8664)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,3}}$	0.0828 (0.0532)	0.3936 (0.3108)	0.7488 (0.6644)	0.9168 (0.8776)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};t_{1,1}}$	0.0864 (0.0532)	0.4088 (0.3104)	0.7612 (0.6720)	0.9264 (0.8832)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{S}}$	0.0920 (0.0556)	0.3896 (0.3028)	0.7336 (0.6488)	0.9092 (0.8564)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{W}}$	0.0824 (0.0524)	0.3872 (0.3072)	0.7376 (0.6552)	0.9108 (0.8728)
$\underset{\sim}{\phi}^{(n)}_{\mathbf{\Lambda};\text{SP}}$	0.0752 (0.0588)	0.3536 (0.2992)	0.6648 (0.6064)	0.8604 (0.8220)

APPENDIX

We start with the proof of Proposition 3.1. To this end, note that although generally stated as a property of a parametric sequence of families of the form $\mathcal{P}^{(n)} = \{\mathrm{P}^{(n)}_{\boldsymbol{\omega}} | \boldsymbol{\omega} \in \boldsymbol{\Omega}\}$ ($n \in \mathbb{N}$), LAN (ULAN) actually is a property of the parametrization $\boldsymbol{\omega} \mapsto \mathrm{P}^{(n)}_{\boldsymbol{\omega}}$, $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ of $\mathcal{P}^{(n)}$ (i.e., of a bijective map from $\boldsymbol{\Omega}$ to $\mathcal{P}^{(n)}$). When parametrized with $\boldsymbol{\omega} := (\boldsymbol{\theta}', (\text{vech}\,\boldsymbol{\Sigma})')'$, $\boldsymbol{\omega} \in \boldsymbol{\Omega} := \mathbb{R}^k \times \text{vech}(\mathcal{S}_k)$, where $\mathcal{S}_k$ stands for the class of positive definite symmetric real $k \times k$ matrices, the elliptical families we are dealing with here have been shown to be ULAN in Hallin and Paindaveine (2006a), with central sequence

$$\boldsymbol{\Delta}_{\boldsymbol{\omega}}^{(n)} := \begin{pmatrix} n^{-1/2}\sum_{i=1}^{n}\varphi_{f_1}(d_i)\boldsymbol{\Sigma}^{-1/2}\mathbf{U}_i \\ \frac{1}{2\sqrt{n}}\mathbf{P}_k(\boldsymbol{\Sigma}^{\otimes 2})^{-1/2}\sum_{i=1}^{n}\operatorname{vec}\big(\varphi_{f_1}(d_i)d_i\mathbf{U}_i\mathbf{U}_i' - \mathbf{I}_k\big) \end{pmatrix}, \tag{A.1}$$

with $d_i = d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ and $\mathbf{U}_i = \mathbf{U}_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, where $\mathbf{P}_k'$ denotes the *duplication matrix* [such that $\mathbf{P}_k' \operatorname{vech}(\mathbf{A}) = \operatorname{vec}(\mathbf{A})$ for any $k \times k$ symmetric matrix $\mathbf{A}$].

The families we are considering in this proposition are slightly different, because the $\boldsymbol{\vartheta}$-parametrization requires k identifiable eigenvectors. However, denoting by $\boldsymbol{\Omega}^B := \mathbb{R}^k \times \operatorname{vech}(\mathcal{S}_k^B)$, where $\mathcal{S}_k^B$ is the set of all matrices in $\mathcal{S}_k$ compatible with Assumption (A), the mapping

$$\begin{aligned} \bar{d} : \boldsymbol{\omega} = (\boldsymbol{\theta}', (\operatorname{vech}\boldsymbol{\Sigma})')' \in \boldsymbol{\Omega}^B \\ \mapsto \quad \bar{d}(\boldsymbol{\omega}) := \big(\boldsymbol{\theta}', (\det\boldsymbol{\Sigma})^{1/k}, (\operatorname{d\overset{\circ}{v}ec}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}})'/(\det\boldsymbol{\Sigma})^{1/k}, (\operatorname{vec}\boldsymbol{\beta})'\big)' \in \boldsymbol{\Theta} \end{aligned}$$

from the open subset $\boldsymbol{\Omega}^B$ of $\mathbb{R}^{k+k(k+1)/2}$ to $\boldsymbol{\Theta}$ is a differentiable mapping such that, with a small abuse of notation, $\mathrm{P}_{\boldsymbol{\omega};f_1}^{(n)} = \mathrm{P}_{\boldsymbol{\vartheta}=\bar{d}(\boldsymbol{\omega});f_1}^{(n)}$ and $\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)} = \mathrm{P}_{\boldsymbol{\omega}=\bar{d}^{-1}(\boldsymbol{\vartheta});f_1}^{(n)}$ (with $\bar{d}^{-1}$ defined on $\boldsymbol{\Theta}$ only). The proof of Proposition 3.1 consists in showing how ULAN in the $\boldsymbol{\omega}$-parametrization implies ULAN in the $\boldsymbol{\vartheta}$-parametrization, and how the central sequences and information matrices are related to each other. Let us start with a lemma.

LEMMA A.1. *Let the parametrization $\boldsymbol{\omega} \mapsto \mathrm{P}_{\boldsymbol{\omega}}^{(n)}$, $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is an open subset of $\mathbb{R}^{k_1}$ be ULAN for $\mathcal{P}^{(n)} = \{\mathrm{P}_{\boldsymbol{\omega}}^{(n)} | \boldsymbol{\omega} \in \boldsymbol{\Omega}\}$, with central sequence $\boldsymbol{\Delta}_{\boldsymbol{\omega}}^{(n)}$ and information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\omega}}$. Let $\bar{d} : \boldsymbol{\omega} \mapsto \boldsymbol{\vartheta} := \bar{d}(\boldsymbol{\omega})$ be a continuously differentiable mapping from $\mathbb{R}^{k_1}$ to $\mathbb{R}^{k_2}$ ($k_2 \geq k_1$) with full column rank Jacobian matrix $D\bar{d}(\boldsymbol{\omega})$ at every $\boldsymbol{\omega}$. Write $\boldsymbol{\Theta} := \bar{d}(\boldsymbol{\Omega})$, and assume that $\boldsymbol{\vartheta} \mapsto \mathrm{P}_{\boldsymbol{\vartheta}}^{(n)}$, $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ provides another parametrization of $\mathcal{P}^{(n)}$. Then, $\boldsymbol{\vartheta} \mapsto \mathrm{P}_{\boldsymbol{\vartheta}}^{(n)}$, $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ is also ULAN, with [at $\boldsymbol{\vartheta} = \bar{d}(\boldsymbol{\omega})$] central sequence $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}}^{(n)} = (D^-\bar{d}(\boldsymbol{\omega}))'\boldsymbol{\Delta}_{\boldsymbol{\omega}}^{(n)}$ and information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}} = (D^-\bar{d}(\boldsymbol{\omega}))'\boldsymbol{\Gamma}_{\boldsymbol{\omega}}D^-\bar{d}(\boldsymbol{\omega})$, where $D^-\bar{d}(\boldsymbol{\omega}) := ((D\bar{d}(\boldsymbol{\omega}))'D\bar{d}(\boldsymbol{\omega}))^{-1}(D\bar{d}(\boldsymbol{\omega}))'$ is the Moore–Penrose inverse of $D\bar{d}(\boldsymbol{\omega})$.*

PROOF. Throughout, let $\boldsymbol{\vartheta}$ and $\boldsymbol{\omega}$ be such that $\boldsymbol{\vartheta} = \bar{d}(\boldsymbol{\omega})$. Consider $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and an arbitrary sequence $\boldsymbol{\vartheta}^{(n)} = \boldsymbol{\vartheta} + O(n^{-1/2}) \in \boldsymbol{\Theta}$. The characterization of ULAN for the $\boldsymbol{\vartheta}$-parametrization involves bounded sequence $\boldsymbol{\tau}_{**}^{(n)} \in \mathbb{R}^{k_2}$ such that the perturbation $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}_{**}^{(n)}$ still belongs to $\boldsymbol{\Theta}$. In order for $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}_{**}^{(n)}$ to belong to $\boldsymbol{\Theta}$, it is necessary that $\boldsymbol{\tau}_{**}^{(n)}$ be of the form $\boldsymbol{\tau}_{*}^{(n)} + o(1)$, with $\boldsymbol{\tau}_{*}^{(n)}$ in the tangent space to $\boldsymbol{\Theta}$ at $\boldsymbol{\vartheta}^{(n)}$, hence of the form $\boldsymbol{\tau}^{(n)} + o(1)$ with $\boldsymbol{\tau}^{(n)}$ in the tangent space to $\boldsymbol{\Theta}$ at $\boldsymbol{\vartheta}$, that is, $\boldsymbol{\tau}^{(n)} = D\vec{d}(\boldsymbol{\omega})\mathbf{w}^{(n)}$ for some bounded sequence $\mathbf{w}^{(n)} \in \mathbb{R}^{k_1}$. It follows from differentiability that, letting $\boldsymbol{\omega}^{(n)} = \vec{d}^{-1}(\boldsymbol{\vartheta}^{(n)})$,

$$
\begin{aligned}
\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}_{**}^{(n)} &= \boldsymbol{\vartheta}^{(n)} + n^{-1/2}D\vec{d}(\boldsymbol{\omega})\mathbf{w}^{(n)} + o(n^{-1/2}) \\
&= \vec{d}\big(\boldsymbol{\omega}^{(n)}\big) + n^{-1/2}D\vec{d}(\boldsymbol{\omega})\mathbf{w}^{(n)} + o(n^{-1/2}) \\
&= \vec{d}\big(\boldsymbol{\omega}^{(n)}\big) + n^{-1/2}D\vec{d}\big(\boldsymbol{\omega}^{(n)}\big)\mathbf{w}^{(n)} + o(n^{-1/2}) \\
&= \vec{d}\big(\boldsymbol{\omega}^{(n)} + n^{-1/2}\mathbf{w}^{(n)} + o(n^{-1/2})\big).
\end{aligned} \tag{A.2}
$$

Hence, turning to local log-likelihood ratios, in view of ULAN for the $\boldsymbol{\omega}$-parametrization,

$$
\begin{aligned}
&\log\big(d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}_{**}^{(n)}}/d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}}\big) \\
&\qquad = \log\big(d\mathrm{P}^{(n)}_{\boldsymbol{\omega}^{(n)}+n^{-1/2}\mathbf{w}^{(n)}+o(n^{-1/2})}/d\mathrm{P}^{(n)}_{\boldsymbol{\omega}^{(n)}}\big) \\
&\qquad = \mathbf{w}^{(n)\prime}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\omega}^{(n)}} - \tfrac{1}{2}\mathbf{w}^{(n)\prime}\boldsymbol{\Gamma}_{\boldsymbol{\omega}}\mathbf{w}^{(n)} + o_{\mathrm{P}}(1)
\end{aligned} \tag{A.3}
$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\omega}^{(n)}} = \mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}}$-probability, as $n \to \infty$. Now, the LAQ part of ULAN for the $\boldsymbol{\vartheta}$-parametrization requires, for some random vector $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)}}$ and constant matrix $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}}$,

$$
\log\big(d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}_{**}^{(n)}}/d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}}\big) = \boldsymbol{\tau}_{**}^{(n)\prime}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)}} - \tfrac{1}{2}\boldsymbol{\tau}_{**}^{(n)\prime}\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}}\boldsymbol{\tau}_{**}^{(n)} + o_{\mathrm{P}}(1) \tag{A.4}
$$

under the same $\mathrm{P}^{(n)}_{\boldsymbol{\omega}^{(n)}} = \mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}}$ probability distributions with, in view of (A.2), $\boldsymbol{\tau}_{**}^{(n)} = D\vec{d}(\boldsymbol{\omega})\mathbf{w}^{(n)} + o(1)$. Identifying (A.3) and (A.4), we obtain that LAQ is satisfied for the $\boldsymbol{\vartheta}$-parametrization, with any $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}}$ satisfying

$$
(D\vec{d}(\boldsymbol{\omega}))'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}} = \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\omega}}. \tag{A.5}
$$

Now, let $\mathbf{t}_i$ be the ith column of $D\vec{d}(\boldsymbol{\omega})$, $i = 1, \dots, k_1$, and choose $\mathbf{t}_{k_1+1}, \dots, \mathbf{t}_{k_2} \in \mathbb{R}^{k_2}$ in such a way that they span the orthogonal complement of $\mathcal{M}(D\vec{d}(\boldsymbol{\omega}))$. Then $\{\mathbf{t}_i,\ i = 1, \dots, k_2\}$ is a basis of $\mathbb{R}^{k_2}$, so that there exists a unique k_2-tuple $(\delta^{(n)}_{\boldsymbol{\vartheta};1}, \dots, \delta^{(n)}_{\boldsymbol{\vartheta};k_2})'$ such that $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}} = \sum_{i=1}^{k_2}\delta^{(n)}_{\boldsymbol{\vartheta};i}\mathbf{t}_i$. With this notation, (A.5) yields

$$
\begin{aligned}
\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\omega}} &= (D\vec{d}(\boldsymbol{\omega}))'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}} \\
&= \sum_{i=1}^{k_2}\delta^{(n)}_{\boldsymbol{\vartheta};i}(D\vec{d}(\boldsymbol{\omega}))'\mathbf{t}_i = \sum_{i=1}^{k_1}\delta^{(n)}_{\boldsymbol{\vartheta};i}(D\vec{d}(\boldsymbol{\omega}))'\mathbf{t}_i \\
&= (D\vec{d}(\boldsymbol{\omega}))'D\vec{d}(\boldsymbol{\omega})\underline{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\vartheta}},
\end{aligned}
$$

where we let $\underline{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\vartheta}} := (\delta^{(n)}_{\boldsymbol{\vartheta};1}, \dots, \delta^{(n)}_{\boldsymbol{\vartheta};k_1})'$. Since $D\vec{d}(\boldsymbol{\omega})$ has full column rank, this entails (i) $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}} = D\vec{d}(\boldsymbol{\omega})\underline{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\vartheta}}$ and (ii) $\underline{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\vartheta}} = ((D\vec{d}(\boldsymbol{\omega}))'D\vec{d}(\boldsymbol{\omega}))^{-1}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\omega}}$, hence

$\mathbf{\Delta}_{\boldsymbol{\vartheta}}^{(n)} = (D^{-}\bar{d}(\boldsymbol{\omega}))'\mathbf{\Delta}_{\boldsymbol{\omega}}^{(n)}$. As a linear transformation of $\mathbf{\Delta}_{\boldsymbol{\omega}}^{(n)}$, $\mathbf{\Delta}_{\boldsymbol{\vartheta}}^{(n)}$ clearly also satisfies the asymptotic normality part of ULAN, with the desired $\mathbf{\Gamma}_{\boldsymbol{\vartheta}}$. $\square$

The following slight extension of Lemma A.1 plays a role in the proof of Proposition 3.1 below. Consider a parametrization $\boldsymbol{\omega} = (\boldsymbol{\omega}_a', \boldsymbol{\omega}_b')' \mapsto \mathrm{P}_{\boldsymbol{\omega}}^{(n)}$, $\boldsymbol{\omega} \in \mathbf{\Omega} \times \mathcal{V}$, where $\mathbf{\Omega}$ is an open subset of $\mathbb{R}^{k_1}$ and $\mathcal{V} \subset \mathbb{R}^m$ is a ℓ-dimensional manifold in $\mathbb{R}^m$, and assume that it is ULAN for $\mathcal{P}^{(n)} = \{\mathrm{P}_{\boldsymbol{\omega}}^{(n)} | \boldsymbol{\omega} \in \mathbf{\Omega} \times \mathcal{V}\}$, with central sequence $\mathbf{\Delta}_{\boldsymbol{\omega}}^{(n)}$ and information matrix $\mathbf{\Gamma}_{\boldsymbol{\omega}}$. Let $\bar{d}_a$ be a continuously differentiable mapping from $\mathbb{R}^{k_1}$ to $\mathbb{R}^{k_2}$ ($k_2 \geq k_1$) with full column rank Jacobian matrix $D\bar{d}_a(\boldsymbol{\omega}_a)$ at every $\boldsymbol{\omega}_a$, and assume that $\boldsymbol{\vartheta} := \bar{d}(\boldsymbol{\omega}) \mapsto \mathrm{P}_{\boldsymbol{\vartheta}}^{(n)}$, $\boldsymbol{\vartheta} \in \mathbf{\Theta} \times \mathcal{V}$ [with $\mathbf{\Theta} := \bar{d}_a(\mathbf{\Omega})$], where

$$\bar{d} : \mathbf{\Omega} \times \mathcal{V} \to \mathbf{\Theta} \times \mathcal{V} \qquad \boldsymbol{\omega} = (\boldsymbol{\omega}_a, \boldsymbol{\omega}_b)' \quad \mapsto \quad \bar{d}(\boldsymbol{\omega}) = (\bar{d}_a(\boldsymbol{\omega}_a), \boldsymbol{\omega}_b)'$$

provides another parametrization of $\mathcal{P}^{(n)}$. Then the proof of Lemma A.1 straightforwardly extends to show that $\boldsymbol{\vartheta} \mapsto \mathrm{P}_{\boldsymbol{\vartheta}}^{(n)}$, $\boldsymbol{\vartheta} \in \mathbf{\Theta} \times \mathcal{V}$ is also ULAN, still with [at $\boldsymbol{\vartheta} = \bar{d}(\boldsymbol{\omega})$] central sequence $\mathbf{\Delta}_{\boldsymbol{\vartheta}}^{(n)} = (D^{-}\bar{d}(\boldsymbol{\omega}))'\mathbf{\Delta}_{\boldsymbol{\omega}}^{(n)}$ and information matrix $\mathbf{\Gamma}_{\boldsymbol{\vartheta}} = (D^{-}\bar{d}(\boldsymbol{\omega}))'\mathbf{\Gamma}_{\boldsymbol{\omega}} D^{-}\bar{d}(\boldsymbol{\omega})$.

PROOF OF PROPOSITION 3.1. Consider the differentiable mappings $\bar{d}_1 : \boldsymbol{\omega} := (\boldsymbol{\theta}', (\operatorname{vech} \mathbf{\Sigma})')' \mapsto \bar{d}_1(\omega) = (\boldsymbol{\theta}', (\operatorname{dvec} \mathbf{\Lambda}_{\mathbf{\Sigma}})', (\operatorname{vec} \boldsymbol{\beta})')'$ and $\bar{d}_2 : \bar{d}_1(\boldsymbol{\omega}) = (\boldsymbol{\theta}', (\operatorname{dvec} \mathbf{\Lambda}_{\mathbf{\Sigma}})', (\operatorname{vec} \boldsymbol{\beta})')' \mapsto \bar{d}_2(\bar{d}_1(\boldsymbol{\omega})) = (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\operatorname{dvec}} \mathbf{\Lambda}_{\mathbf{V}})', (\operatorname{vec} \boldsymbol{\beta})')' \in \mathbf{\Theta}$, the latter being invertible. Applying Lemma A.1 twice (the second time in its "extended form," since the $\boldsymbol{\beta}$-part of the parameter is invariant under $\bar{d}_2$) then yields

$$\begin{aligned}
\mathbf{\Delta}_{\boldsymbol{\vartheta}}^{(n)} &= (D\bar{d}_2(\bar{d}_1(\boldsymbol{\omega})))'^{-1} D\bar{d}_1(\boldsymbol{\omega})((D\bar{d}_1(\boldsymbol{\omega}))' D\bar{d}_1(\boldsymbol{\omega}))^{-1} \mathbf{\Delta}_{\boldsymbol{\omega}}^{(n)} \\
&= (D\bar{d}_2^{-1}(\bar{d}(\boldsymbol{\omega})))' D\bar{d}_1(\boldsymbol{\omega})((D\bar{d}_1(\boldsymbol{\omega}))' D\bar{d}_1(\boldsymbol{\omega}))^{-1} \mathbf{\Delta}_{\boldsymbol{\omega}}^{(n)}.
\end{aligned}$$

In view of the definition of $\mathbf{M}_k^{\mathbf{\Lambda}_{\mathbf{V}}}$ (Section 3.2), the Jacobian matrix, computed at $\boldsymbol{\vartheta}$, of the inverse mapping $\bar{d}_2^{-1}$ is

$$D\bar{d}_2^{-1}(\boldsymbol{\vartheta}) = \begin{pmatrix} \mathbf{I}_k & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \operatorname{dvec}(\mathbf{\Lambda}_{\mathbf{V}}) & \sigma^2 (\mathbf{M}_k^{\mathbf{\Lambda}_{\mathbf{V}}})' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{k^2} \end{pmatrix}.$$

An explicit expression for $D\bar{d}_1(\boldsymbol{\omega})$ was obtained by Kollo and Neudecker [(1993), page 288]:

$$D\bar{d}_1(\boldsymbol{\omega}) = \begin{pmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_{\boldsymbol{\beta}, \mathbf{\Lambda}_{\mathbf{\Sigma}}} \mathbf{P}_k' \end{pmatrix},$$
$$\text{with } \mathbf{\Xi}_{\boldsymbol{\beta}, \mathbf{\Lambda}_{\mathbf{\Sigma}}} := \begin{pmatrix} \mathbf{H}_k (\boldsymbol{\beta}')^{\otimes 2} \\ \boldsymbol{\beta}_1' \otimes [\boldsymbol{\beta}(\lambda_{1;\mathbf{\Sigma}} \mathbf{I}_k - \mathbf{\Lambda}_{\mathbf{\Sigma}})^{-} \boldsymbol{\beta}'] \\ \vdots \\ \boldsymbol{\beta}_k' \otimes [\boldsymbol{\beta}(\lambda_{k;\mathbf{\Sigma}} \mathbf{I}_k - \mathbf{\Lambda}_{\mathbf{\Sigma}})^{-} \boldsymbol{\beta}'] \end{pmatrix}. \tag{A.6}$$

The result then follows from a direct, though painful, computation, using the fact that

$$(\mathbf{P}_k \boldsymbol{\Xi}'_{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}} \boldsymbol{\Xi}_{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}} \mathbf{P}'_k)^{-1} = (\mathbf{P}'_k)^{-}(\boldsymbol{\beta} \otimes \boldsymbol{\beta}) \operatorname{diag}(l_{11;\boldsymbol{\Sigma}}, l_{12;\boldsymbol{\Sigma}}, \dots, l_{kk;\boldsymbol{\Sigma}})(\boldsymbol{\beta}' \otimes \boldsymbol{\beta}')\mathbf{P}_k^{-},$$

with $l_{ij;\boldsymbol{\Sigma}} = 1$ if $i = j$ and $l_{ij;\boldsymbol{\Sigma}} = (\lambda_{i;\boldsymbol{\Sigma}} - \lambda_{j;\boldsymbol{\Sigma}})^{-2}$ if $i \neq j$; $(\mathbf{P}'_k)^{-}$ here stands for the Moore–Penrose inverse of $\mathbf{P}_k$ [note that $(\mathbf{P}'_k)^{-}$ is such that $\mathbf{P}'_k(\mathbf{P}'_k)^{-} \operatorname{vec}(\mathbf{A}) = \operatorname{vec}(\mathbf{A})$ for any symmetric matrix $\mathbf{A}$]. $\square$

PROOF OF PROPOSITION 3.2. Proceeding as in the proof of Lemma A.1, let $\mathbf{v}_i$ be the ith column of $D\hbar(\boldsymbol{\xi}_0)$, $i = 1, \dots, m$, and choose $\mathbf{v}_{m+1}, \dots, \mathbf{v}_p \in \mathbb{R}^p$ spanning the orthogonal complement of $\mathcal{M}(D\hbar(\boldsymbol{\xi}_0))$. Then there exists a unique p-tuple $(\delta_{\boldsymbol{\vartheta}_0;1}, \dots, \delta_{\boldsymbol{\vartheta}_0;p})'$ such that $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0} = \sum_{i=1}^p \delta_{\boldsymbol{\vartheta}_0;i}\mathbf{v}_i$ (since $\mathbf{v}_i, i = 1, \dots, p$ spans $\mathbb{R}^p$) and

$$\begin{aligned}\boldsymbol{\Delta}_{\boldsymbol{\xi}_0} = D\hbar'(\boldsymbol{\xi}_0)\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0} &= \sum_{i=1}^p \delta_{\boldsymbol{\vartheta}_0;i} D\hbar'(\boldsymbol{\xi}_0)\mathbf{v}_i = \sum_{i=1}^m \delta_{\boldsymbol{\vartheta}_0;i} D\hbar'(\boldsymbol{\xi}_0)\mathbf{v}_i \\ &= \mathbf{C}_\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0},\end{aligned} \tag{A.7}$$

where $\mathbf{C}_\hbar(\boldsymbol{\xi}_0) := D\hbar'(\boldsymbol{\xi}_0)D\hbar(\boldsymbol{\xi}_0)$ and $\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} := (\delta_{\boldsymbol{\vartheta}_0;1}, \dots, \delta_{\boldsymbol{\vartheta}_0;m})'$. Hence, we also have $\boldsymbol{\Gamma}_{\boldsymbol{\xi}_0} = \mathbf{C}_\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0}\mathbf{C}_\hbar(\boldsymbol{\xi}_0)$, where $\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0}$ is the asymptotic covariance matrix of $\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0}$. Using the fact that $\mathbf{C}_\hbar(\boldsymbol{\xi}_0)$ is invertible, this yields

$$\begin{aligned}Q_{\boldsymbol{\xi}_0} &:= (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})'\mathbf{C}_\hbar(\boldsymbol{\xi}_0)(\mathbf{C}_\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0}\mathbf{C}_\hbar(\boldsymbol{\xi}_0))^{-1}\mathbf{C}_\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\ &\quad - (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})'\mathbf{C}_\hbar(\boldsymbol{\xi}_0)D\bar{l}(\boldsymbol{\alpha}_0)(D\bar{l}'(\boldsymbol{\alpha}_0)\mathbf{C}_\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0}\mathbf{C}_\hbar(\boldsymbol{\xi}_0)D\bar{l}(\boldsymbol{\alpha}_0))^{-1} \\ &\qquad \times D\bar{l}'(\boldsymbol{\alpha}_0)\mathbf{C}_\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\ &= (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})'(\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{-1}\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\ &\quad - (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})'(\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{-1/2}\boldsymbol{\Pi}((\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{1/2}\mathbf{C}_\hbar(\boldsymbol{\xi}_0)D\bar{l}(\boldsymbol{\alpha}_0))(\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{-1/2}\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\ &=: Q_{\boldsymbol{\xi}_0,1} - Q_{\boldsymbol{\xi}_0,2},\end{aligned}$$

where $\boldsymbol{\Pi}(\mathbf{P}) := \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'$ denotes the projection matrix on $\mathcal{M}(\mathbf{P})$.

Let $\bar{b} : A \subset \mathbb{R}^\ell \to \mathbb{R}^p$ be a local (at $\boldsymbol{\vartheta}_0$) chart for the manifold $C \cap \boldsymbol{\Theta}$, and assume, without loss of generality, that $\boldsymbol{\eta}_0 = \bar{b}^{-1}(\boldsymbol{\vartheta}_0)$. Since $D\hbar(\boldsymbol{\xi}_0)$ has maximal rank, it follows from (A.7) that $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0} = D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0}$. Hence, the statistic

$$\bar{Q}_{\boldsymbol{\vartheta}_0} := \boldsymbol{\Delta}'_{\boldsymbol{\vartheta}_0}\big(\boldsymbol{\Gamma}^{-}_{\boldsymbol{\vartheta}_0} - D\bar{b}(\boldsymbol{\eta}_0)(D\bar{b}'(\boldsymbol{\eta}_0)\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0}D\bar{b}(\boldsymbol{\eta}_0))^{-}D\bar{b}'(\boldsymbol{\eta}_0)\big)\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0} \tag{A.8}$$

[the squared Euclidean norm of the orthogonal projection, onto the linear space orthogonal to $\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta}_0}D\bar{b}(\boldsymbol{\eta}_0)$, of the standardized central sequence $(\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta}_0})^{-}\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0}$] can be written as

$$\begin{aligned}
\bar{Q}_{\boldsymbol{\vartheta}_0} &= (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})' D\hbar'(\boldsymbol{\xi}_0)(D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0} D\hbar'(\boldsymbol{\xi}_0))^- D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\
&\quad - (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})' D\hbar'(\boldsymbol{\xi}_0) D\bar{b}(\boldsymbol{\eta}_0)(D\bar{b}'(\boldsymbol{\eta}_0) D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0} D\hbar'(\boldsymbol{\xi}_0) D\bar{b}(\boldsymbol{\eta}_0))^- \\
&\qquad \times D\bar{b}'(\boldsymbol{\eta}_0) D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\
&= (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})' D\hbar'(\boldsymbol{\xi}_0)(D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0} D\hbar'(\boldsymbol{\xi}_0))^- D\hbar(\boldsymbol{\xi}_0)\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\
&\quad - (\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0})' (\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{-1/2} \boldsymbol{\Pi}((\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{1/2} D\hbar'(\boldsymbol{\xi}_0) D\bar{b}(\boldsymbol{\eta}_0))(\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{-1/2}\boldsymbol{\Delta}^m_{\boldsymbol{\vartheta}_0} \\
&=: \bar{Q}_{\boldsymbol{\vartheta}_0,1} - \bar{Q}_{\boldsymbol{\vartheta}_0,2}.
\end{aligned}$$

Since $D\hbar(\boldsymbol{\xi}_0)$ has full rank, the standard properties of Moore–Penrose inverses entail $Q_{\boldsymbol{\xi}_0,1} = \bar{Q}_{\boldsymbol{\vartheta}_0,1}$. As for $Q_{\boldsymbol{\xi}_0,2}$ and $\bar{Q}_{\boldsymbol{\vartheta}_0,2}$, they are equal if

$$\mathcal{M}((\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{1/2}\mathbf{C}_\hbar(\boldsymbol{\xi}_0) D\bar{l}(\boldsymbol{\alpha}_0)) = \mathcal{M}((\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0})^{1/2} D\hbar'(\boldsymbol{\xi}_0) D\bar{b}(\boldsymbol{\eta}_0)).$$

Since $\boldsymbol{\Gamma}^m_{\boldsymbol{\vartheta}_0}$ and $\mathbf{C}_\hbar(\boldsymbol{\xi}_0)$ are invertible, the latter equality holds if $\mathcal{M}(D\bar{l}(\boldsymbol{\alpha}_0)) = \mathcal{M}((\mathbf{C}_\hbar(\boldsymbol{\xi}_0))^{-1} D\hbar'(\boldsymbol{\xi}_0) D\bar{b}(\boldsymbol{\eta}_0))$, or, since $D\hbar(\boldsymbol{\xi}_0)$ has full rank, if

$$\begin{aligned}
\mathcal{M}(D\hbar(\boldsymbol{\xi}_0) D\bar{l}(\boldsymbol{\alpha}_0)) &= \mathcal{M}(D\hbar(\boldsymbol{\xi}_0)(\mathbf{C}_\hbar(\boldsymbol{\xi}_0))^{-1} D\hbar'(\boldsymbol{\xi}_0) D\bar{b}(\boldsymbol{\eta}_0)) \\
&\qquad \big(= \mathcal{M}(\boldsymbol{\Pi}(D\hbar(\boldsymbol{\xi}_0)) D\bar{b}(\boldsymbol{\eta}_0))\big),
\end{aligned}$$

which trivially holds true. Hence, $Q_{\boldsymbol{\xi}_0,2} = \bar{Q}_{\boldsymbol{\vartheta}_0,2}$, so that $Q_{\boldsymbol{\xi}_0} = \bar{Q}_{\boldsymbol{\vartheta}_0}$.

Eventually, the linear spaces orthogonal to $\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta}_0} D\bar{b}(\boldsymbol{\eta}_0)$ and to $\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta}_0} D\bar{b}(\boldsymbol{\eta}_0)$ do coincide, so that the statistic $Q_{\boldsymbol{\vartheta}_0}$, which is obtained by substituting $\bar{b}$ for $\bar{b}$ in (A.8), is equal to $Q_{\boldsymbol{\vartheta}_0}(=Q_{\boldsymbol{\xi}_0})$. This establishes the result. $\square$

We now turn to the proofs of Lemmas 4.1 and 4.2.

PROOF OF LEMMA 4.1. The proof consists in checking that postmultiplying $\mathbf{D}_k(\boldsymbol{\Lambda}_\mathbf{V})$ with $\mathbf{N}_k\mathbf{H}_k\mathbf{P}_k^{\boldsymbol{\Lambda}_\mathbf{V}}(\mathbf{I}_{k^2}+\mathbf{K}_k)\boldsymbol{\Lambda}_\mathbf{V}^{\otimes 2}(\mathbf{P}_k^{\boldsymbol{\Lambda}_\mathbf{V}})'\mathbf{H}_k'\mathbf{N}_k'$ yields the $(k-1)$-dimensional identity matrix ($\mathbf{P}_k^{\boldsymbol{\Lambda}_\mathbf{V}}$ and $\mathbf{N}_k$ are defined in the statement of the lemma). That is, we show that

$$\begin{aligned}
&\tfrac{1}{4}\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}}\mathbf{H}_k(\mathbf{I}_{k^2}+\mathbf{K}_k)(\boldsymbol{\Lambda}_\mathbf{V}^{-1})^{\otimes 2}\mathbf{H}_k'(\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}})'\mathbf{N}_k\mathbf{H}_k \\
&\quad \times \mathbf{P}_k^{\boldsymbol{\Lambda}_\mathbf{V}}(\mathbf{I}_{k^2}+\mathbf{K}_k)\boldsymbol{\Lambda}_\mathbf{V}^{\otimes 2}(\mathbf{P}_k^{\boldsymbol{\Lambda}_\mathbf{V}})'\mathbf{H}_k'\mathbf{N}_k' \\
&\quad = \mathbf{I}_{k-1}.
\end{aligned} \tag{A.9}$$

First of all, note that the definition of $\mathbf{M}_k^{\mathbf{V}}$ (see Section 3.2) entails that, for any $k\times k$ real matrix $\mathbf{l}$ such that $\mathrm{tr}(\boldsymbol{\Lambda}_\mathbf{V}^{-1}\mathbf{l}) = 0$, $(\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}})'\mathbf{N}_k\mathbf{H}_k(\mathrm{vec}\,\mathbf{l}) = (\mathbf{M}_k^{\boldsymbol{\Lambda}_\mathbf{V}})'(\mathrm{d}\overset{\circ}{\mathrm{v}}\mathrm{ec}\,\mathbf{l}) =$

$\mathrm{dvec}(\mathbf{l}) = \mathbf{H}_k(\mathrm{vec}\,\mathbf{l})$. Hence, since (letting $\mathbf{E}_{ij} := \mathbf{e}_i\mathbf{e}_j' + \mathbf{e}_j\mathbf{e}_i'$)

$$\begin{aligned}\mathbf{P}_k^{\mathbf{\Lambda_V}}(\mathbf{I}_{k^2} + \mathbf{K}_k) &= \mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{\Lambda}_{\mathbf{V}}^{\otimes 2}\,\mathrm{vec}(\mathbf{\Lambda}_{\mathbf{V}}^{-1})(\mathrm{vec}(\mathbf{\Lambda}_{\mathbf{V}}^{-1}))' \\ &= \sum_{i,j=1}^{k}\mathrm{vec}\Big(\frac{1}{2}\mathbf{E}_{ij} - \frac{1}{k}(\mathbf{\Lambda}_{\mathbf{V}}^{-1})_{ij}\mathbf{\Lambda_V}\Big)(\mathrm{vec}\,\mathbf{E}_{ij})' \\ &=: \sum_{i,j=1}^{k}(\mathrm{vec}\,\mathbf{F}_{ij}^{\mathbf{\Lambda_V}})(\mathrm{vec}\,\mathbf{E}_{ij})',\end{aligned}$$

with $\mathrm{tr}(\mathbf{\Lambda}_{\mathbf{V}}^{-1}\mathbf{F}_{ij}^{\mathbf{\Lambda_V}}) = 0$, for all $i, j = 1, \ldots, k$, we obtain that $(\mathbf{M}_k^{\mathbf{\Lambda_V}})'\mathbf{N}_k\mathbf{H}_k \times \mathbf{P}_k^{\mathbf{\Lambda_V}}(\mathbf{I}_{k^2} + \mathbf{K}_k) = \mathbf{H}_k\mathbf{P}_k^{\mathbf{\Lambda_V}}(\mathbf{I}_{k^2} + \mathbf{K}_k)$. Now, using the fact that $\mathbf{H}_k'\mathbf{H}_k(\mathbf{\Lambda}_{\mathbf{V}}^{-1})^{\otimes 2}(\mathbf{I}_{k^2} + \mathbf{K}_k)\mathbf{H}_k' = (\mathbf{\Lambda}_{\mathbf{V}}^{-1})^{\otimes 2}(\mathbf{I}_{k^2} + \mathbf{K}_k)\mathbf{H}_k'$, the left-hand side of (A.9) reduces to

$$\frac{1}{4}\mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{H}_k(\mathbf{I}_{k^2} + \mathbf{K}_k)(\mathbf{\Lambda}_{\mathbf{V}}^{-1})^{\otimes 2}\mathbf{P}_k^{\mathbf{\Lambda_V}}(\mathbf{I}_{k^2} + \mathbf{K}_k)\mathbf{\Lambda}_{\mathbf{V}}^{\otimes 2}(\mathbf{P}_k^{\mathbf{\Lambda_V}})'\mathbf{H}_k'\mathbf{N}_k'. \tag{A.10}$$

After straightforward computation, using essentially the well-known property of the Kronecker product $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{B})$ and the fact that $\mathbf{M}_k^{\mathbf{\Lambda_V}} \times \mathbf{H}_k(\mathrm{vec}\,\mathbf{\Lambda}_{\mathbf{V}}^{-1}) = \mathbf{0}$ and $\mathbf{H}_k\mathbf{K}_k = \mathbf{H}_k$, (A.10) reduces to $\mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{H}_k\mathbf{H}_k'\mathbf{N}_k'$. The result follows, since $\mathbf{H}_k\mathbf{H}_k' = \mathbf{I}_k$ and $\mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{N}_k' = \mathbf{I}_{k-1}$. □

PROOF OF LEMMA 4.2. All stochastic convergences in this proof are as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, for some fixed $\boldsymbol{\vartheta} \in \mathbf{\Theta}$ and $g_1 \in \mathcal{F}_1^4$. It follows from

$$\mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{H}_k(\boldsymbol{\beta}')^{\otimes 2}(\mathbf{V}^{-1})^{\otimes 2}\,\mathrm{vec}\,\mathbf{V} = \mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{H}_k(\mathrm{vec}\,\mathbf{\Lambda}_{\mathbf{V}}^{-1}) = \mathbf{0} \tag{A.11}$$

and

$$\mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}}(\mathbf{V}^{-1})^{\otimes 2}\,\mathrm{vec}\,\mathbf{V} = \mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}}(\mathrm{vec}\,\mathbf{V}^{-1}) = \mathbf{0}, \tag{A.12}$$

that

$$\begin{aligned}\mathbf{\Delta}_{\boldsymbol{\vartheta};\phi_1}^{III} &= \frac{a_k}{2\sqrt{n}}\mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{H}_k(\mathbf{\Lambda}_{\mathbf{V}}^{-1/2}\boldsymbol{\beta}')^{\otimes 2} \\ &\quad\times\sum_{i=1}^{n}\frac{d_i^2(\boldsymbol{\theta},\mathbf{V})}{\sigma^2}\,\mathrm{vec}(\mathbf{U}_i(\boldsymbol{\theta},\mathbf{V})\mathbf{U}_i'(\boldsymbol{\theta},\mathbf{V})) \\ &= \frac{a_k}{2\sqrt{n}\sigma^2}\mathbf{M}_k^{\mathbf{\Lambda_V}}\mathbf{H}_k(\boldsymbol{\beta}')^{\otimes 2}(\mathbf{V}^{-1})^{\otimes 2} \\ &\quad\times\sum_{i=1}^{n}\mathrm{vec}\big((\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})' - \big(D_k(g_1)/k\big)\mathbf{\Sigma}\big)\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta};\phi_1} &= \frac{a_k}{2\sqrt{n}}\mathbf{G}_k^{\boldsymbol{\beta}}\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}}(\mathbf{V}^{-1/2})^{\otimes 2}\sum_{i=1}^n \frac{d_i^2(\boldsymbol{\theta},\mathbf{V})}{\sigma^2}\operatorname{vec}(\mathbf{U}_i(\boldsymbol{\theta},\mathbf{V})\mathbf{U}_i'(\boldsymbol{\theta},\mathbf{V}))\\
&= \frac{a_k}{2\sqrt{n}\sigma^2}\mathbf{G}_k^{\boldsymbol{\beta}}\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}}(\mathbf{V}^{-1})^{\otimes 2}\\
&\quad\times\sum_{i=1}^n \operatorname{vec}\bigl((\mathbf{X}_i-\boldsymbol{\theta})(\mathbf{X}_i-\boldsymbol{\theta})' - \bigl(D_k(g_1)/k\bigr)\boldsymbol{\Sigma}\bigr).
\end{aligned}$$

Hence, using a root-n consistent estimator $\hat{\boldsymbol{\vartheta}} := (\hat{\boldsymbol{\theta}}', \hat{\sigma}^2, (\mathrm{d}\overset{\circ}{\mathrm{ve}}\mathrm{c}\,\hat{\boldsymbol{\Lambda}}_{\mathbf{V}})', (\operatorname{vec}\hat{\boldsymbol{\beta}})')'$ and letting $\hat{\boldsymbol{\Sigma}} := \hat{\sigma}^2\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\Lambda}}_{\mathbf{V}}\hat{\boldsymbol{\beta}}'$, Slutsky's lemma yields

$$\begin{aligned}
\boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi_1} &= \frac{a_k}{2\sqrt{n}\hat{\sigma}^2}\mathbf{M}_k^{\hat{\boldsymbol{\Lambda}}_{\mathbf{V}}}\mathbf{H}_k(\hat{\boldsymbol{\beta}}')^{\otimes 2}(\hat{\mathbf{V}}^{-1})^{\otimes 2}\\
&\quad\times\sum_{i=1}^n \operatorname{vec}\bigl((\mathbf{X}_i-\hat{\boldsymbol{\theta}})(\mathbf{X}_i-\hat{\boldsymbol{\theta}})' - \bigl(D_k(g_1)/k\bigr)\hat{\boldsymbol{\Sigma}}\bigr)\\
&= \frac{a_k}{2\sqrt{n}\hat{\sigma}^2}\mathbf{M}_k^{\hat{\boldsymbol{\Lambda}}_{\mathbf{V}}}\mathbf{H}_k(\hat{\boldsymbol{\beta}}')^{\otimes 2}(\hat{\mathbf{V}}^{-1})^{\otimes 2}\\
&\quad\times\Biggl\{\sum_{i=1}^n \operatorname{vec}\bigl((\mathbf{X}_i-\boldsymbol{\theta})(\mathbf{X}_i-\boldsymbol{\theta})' - \bigl(D_k(g_1)/k\bigr)\boldsymbol{\Sigma}\bigr)\\
&\qquad - n\operatorname{vec}\bigl((\bar{\mathbf{X}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})'\bigr) - n\operatorname{vec}\bigl((\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\bar{\mathbf{X}}-\boldsymbol{\theta})'\bigr)\\
&\qquad + n\operatorname{vec}\bigl((\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})'\bigr) - n\bigl(D_k(g_1)/k\bigr)\operatorname{vec}(\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma})\Biggr\}\\
&= \boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};\phi_1} - \frac{a_k D_k(g_1)}{2k\sigma^2}\mathbf{M}_k^{\boldsymbol{\Lambda}_{\mathbf{V}}}\mathbf{H}_k(\boldsymbol{\beta}')^{\otimes 2}(\mathbf{V}^{-1})^{\otimes 2}n^{1/2}\operatorname{vec}(\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma})\\
&\quad + o_{\mathrm{P}}(1),
\end{aligned}$$

and, similarly,

$$\begin{aligned}
\boldsymbol{\Delta}^{IV}_{\hat{\boldsymbol{\vartheta}};\phi_1} &= \frac{a_k}{2\sqrt{n}\hat{\sigma}^2}\mathbf{G}_k^{\hat{\boldsymbol{\beta}}}\mathbf{L}_k^{\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\Lambda}}_{\mathbf{V}}}(\hat{\mathbf{V}}^{-1})^{\otimes 2}\\
&\quad\times\sum_{i=1}^n \operatorname{vec}\bigl((\mathbf{X}_i-\hat{\boldsymbol{\theta}})(\mathbf{X}_i-\hat{\boldsymbol{\theta}})' - \bigl(D_k(g_1)/k\bigr)\hat{\boldsymbol{\Sigma}}\bigr)\\
&= \boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta};\phi_1} - \frac{a_k D_k(g_1)}{2k\sigma^2}\mathbf{G}_k^{\boldsymbol{\beta}}\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}}(\mathbf{V}^{-1})^{\otimes 2}n^{1/2}\operatorname{vec}(\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma})\\
&\quad + o_{\mathrm{P}}(1).
\end{aligned}$$

Writing $\hat{\mathbf{\Sigma}} - \mathbf{\Sigma} = (\hat{\sigma}^2 - \sigma^2)\hat{\mathbf{V}} + \sigma^2(\hat{\mathbf{V}} - \mathbf{V})$, applying Slutsky's lemma again, and using (A.11), (A.12) and the fact that $\mathbf{K}_k \operatorname{vec}(\mathbf{A}) = \operatorname{vec}(\mathbf{A}')$, we obtain

$$\begin{aligned}\mathbf{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi_1} &= \mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi_1} - \frac{a_k D_k(g_1)}{2k} \mathbf{M}_k^{\mathbf{\Lambda_V}} \mathbf{H}_k (\boldsymbol{\beta}')^{\otimes 2} (\mathbf{V}^{-1})^{\otimes 2} n^{1/2} \operatorname{vec}(\hat{\mathbf{V}} - \mathbf{V}) \\ &\quad + o_{\mathrm{P}}(1) \\ &= \mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi_1} - \frac{a_k D_k(g_1)}{4k} \mathbf{M}_k^{\mathbf{\Lambda_V}} \mathbf{H}_k (\boldsymbol{\beta}')^{\otimes 2} (\mathbf{V}^{-1})^{\otimes 2} [\mathbf{I}_{k^2} + \mathbf{K}_k] \\ &\qquad \times n^{1/2} \operatorname{vec}(\hat{\mathbf{V}} - \mathbf{V}) + o_{\mathrm{P}}(1)\end{aligned} \tag{A.13}$$

and

$$\begin{aligned}\mathbf{\Delta}^{IV}_{\hat{\boldsymbol{\vartheta}};\phi_1} &= \mathbf{\Delta}^{IV}_{\boldsymbol{\vartheta};\phi_1} - \frac{a_k D_k(g_1)}{2k} \mathbf{G}_k^{\boldsymbol{\beta}} \mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}} (\mathbf{V}^{-1})^{\otimes 2} n^{1/2} \operatorname{vec}(\hat{\mathbf{V}} - \mathbf{V}) + o_{\mathrm{P}}(1) \\ &= \mathbf{\Delta}^{IV}_{\boldsymbol{\vartheta};\phi_1} - \frac{a_k D_k(g_1)}{4k} \mathbf{G}_k^{\boldsymbol{\beta}} \mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}} (\mathbf{V}^{-1})^{\otimes 2} [\mathbf{I}_{k^2} + \mathbf{K}_k] n^{1/2} \operatorname{vec}(\hat{\mathbf{V}} - \mathbf{V}) \\ &\quad + o_{\mathrm{P}}(1).\end{aligned} \tag{A.14}$$

Now, Kollo and Neudecker (1993) showed that

$$n^{1/2} \begin{pmatrix} \operatorname{dvec}(\hat{\mathbf{\Lambda}}_{\mathbf{V}} - \mathbf{\Lambda}_{\mathbf{V}}) \\ \operatorname{vec}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \end{pmatrix} = n^{1/2} \mathbf{\Xi}_{\boldsymbol{\beta},\mathbf{\Lambda_V}} \operatorname{vec}(\hat{\mathbf{V}} - \mathbf{V}) + o_{\mathrm{P}}(1),$$

where $\mathbf{\Xi}_{\boldsymbol{\beta},\mathbf{\Lambda_V}}$ was defined in (A.6). Similar computations as in the proof of Proposition 3.1 then yield

$$\begin{aligned} & n^{1/2} \operatorname{vec}(\hat{\mathbf{V}} - \mathbf{V}) \\ &\quad = n^{1/2} (\mathbf{\Xi}'_{\boldsymbol{\beta},\mathbf{\Lambda_V}} \mathbf{\Xi}_{\boldsymbol{\beta},\mathbf{\Lambda_V}})^{-1} \mathbf{\Xi}'_{\boldsymbol{\beta},\mathbf{\Lambda_V}} \begin{pmatrix} \operatorname{dvec}(\hat{\mathbf{\Lambda}}_{\mathbf{V}} - \mathbf{\Lambda}_{\mathbf{V}}) \\ \operatorname{vec}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \end{pmatrix} + o_{\mathrm{P}}(1) \\ &\quad = (\mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}})' (\mathbf{G}_k^{\boldsymbol{\beta}})' n^{1/2} \operatorname{vec}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &\qquad + \boldsymbol{\beta}^{\otimes 2} \mathbf{H}'_k n^{1/2} \operatorname{dvec}(\hat{\mathbf{\Lambda}}_{\mathbf{V}} - \mathbf{\Lambda}_{\mathbf{V}}) + o_{\mathrm{P}}(1).\end{aligned} \tag{A.15}$$

The result for $\mathbf{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi_1}$ then follows by plugging (A.15) into (A.13) and using the facts that $\mathbf{H}_k(\boldsymbol{\beta}')^{\otimes 2}(\mathbf{V}^{-1})^{\otimes 2}(\mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}})' = \mathbf{0}$ and $n^{1/2} \operatorname{dvec}(\hat{\mathbf{\Lambda}}_{\mathbf{V}} - \mathbf{\Lambda}_{\mathbf{V}}) = n^{1/2} (\mathbf{M}_k^{\mathbf{\Lambda_V}})' \operatorname{d\mathring{v}ec}(\hat{\mathbf{\Lambda}}_{\mathbf{V}} - \mathbf{\Lambda}_{\mathbf{V}}) + o_{\mathrm{P}}(1)$ as $n \to \infty$ (the latter is a direct consequence of the definition of $\mathbf{M}_k^{\mathbf{\Lambda_V}}$ and the delta method). As for the result for $\mathbf{\Delta}^{IV}_{\hat{\boldsymbol{\vartheta}};\phi_1}$, it follows similarly by plugging (A.15) into (A.14) by noting that $\mathbf{G}_k^{\boldsymbol{\beta}} \mathbf{L}_k^{\boldsymbol{\beta},\mathbf{\Lambda_V}} (\mathbf{V}^{-1})^{\otimes 2} [\mathbf{I}_{k^2} + \mathbf{K}_k] \boldsymbol{\beta}^{\otimes 2} \mathbf{H}'_k = \mathbf{0}$. $\square$

PROOF OF LEMMA 6.1. Throughout fix $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\operatorname{d\mathring{v}ec} \mathbf{\Lambda})', (\operatorname{vec} \boldsymbol{\beta})')' \in \mathcal{H}_{0;q}^{\mathbf{\Lambda}\prime\prime}$ and $g_1 \in \mathcal{F}_a$, and define $\tilde{\mathbf{V}} := \hat{\boldsymbol{\beta}}_{\mathrm{Tyler}} \tilde{\mathbf{\Lambda}}_{\mathbf{V}} \hat{\boldsymbol{\beta}}'_{\mathrm{Tyler}}$. Since $\mathbf{K}_k \operatorname{vec}(\mathbf{A}) = \operatorname{vec}(\mathbf{A}')$

and $\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}(\operatorname{vec}\mathbf{I}_k)=\mathbf{0}$, we obtain, from (6.4),

$$\begin{aligned}
\underset{\sim}{T}^{(n)}_{K} &= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}\mathbf{J}_k^{\perp}\operatorname{vec}\big(\underset{\sim}{\mathbf{S}}^{(n)}_{\hat{\boldsymbol{\vartheta}};K}\big)\\
&= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}\mathbf{J}_k^{\perp}\operatorname{vec}\big(\underset{\sim}{\mathbf{S}}^{(n)}_{\boldsymbol{\vartheta};K}\big)\\
&\quad - \left(\frac{\mathcal{J}_k^2(K,g_1)}{4k(k+2)\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}\\
&\quad \times(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\operatorname{vec}(\tilde{\mathbf{V}}-\mathbf{V})+o_{\mathrm{P}}(1)
\end{aligned}\tag{A.16}$$

as $n\to\infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$.

We now show that the second term in (A.16) is $o_{\mathrm{P}}(1)$ as $n\to\infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$. Since $n^{1/2}\operatorname{vec}(\tilde{\mathbf{V}}-\mathbf{V})$ is $O_{\mathrm{P}}(1)$, Slutsky's lemma yields

$$\begin{aligned}
&(a_{p,q}(\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\operatorname{vec}(\tilde{\mathbf{V}}-\mathbf{V})\\
&\quad=(a_{p,q}(\boldsymbol{\Lambda}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}n^{1/2}\operatorname{vec}(\tilde{\mathbf{V}}-\mathbf{V})+o_{\mathrm{P}}(1).
\end{aligned}$$

By construction of the estimator $\tilde{\boldsymbol{\Lambda}}_{\mathbf{V}}$, $\mathbf{c}'_{p,q}\mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\operatorname{vec}\tilde{\mathbf{V}})=0$, so that we have to show that $n^{1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\operatorname{vec}(\hat{\boldsymbol{\beta}}'_{\mathrm{Tyler}}\mathbf{V}\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}})$ is $o_{\mathrm{P}}(1)$. We only do so for $\boldsymbol{\vartheta}$ values such that $\lambda_{1;\mathbf{V}}=\cdots=\lambda_{q;\mathbf{V}}=:\lambda_1^*>\lambda_2^*:=\lambda_{q+1;\mathbf{V}}=\cdots=\lambda_{k;\mathbf{V}}$, which is the most difficult case (extension to the general case is straightforward, although notationally more tricky). Note that the fact that $\boldsymbol{\vartheta}\in\mathcal{H}^{\boldsymbol{\Lambda}''}_{0;q}$ then implies that

$$-pq\lambda_1^*+(1-p)(k-q)\lambda_2^*=0.\tag{A.17}$$

Partition $\mathbf{E}:=\boldsymbol{\beta}'\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}}$ into

$$\mathbf{E}=\begin{pmatrix}\mathbf{E}_{11} & \mathbf{E}_{12}\\ \mathbf{E}_{21} & \mathbf{E}_{22}\end{pmatrix},\tag{A.18}$$

where $\mathbf{E}_{11}$ is $q\times q$ and $\mathbf{E}_{22}$ is $(k-q)\times(k-q)$. As shown in Anderson [(1963), page 129] $n^{1/2}(\mathbf{E}_{11}\mathbf{E}'_{11}-\mathbf{I}_q)=o_{\mathrm{P}}(1)=n^{1/2}(\mathbf{E}_{22}\mathbf{E}'_{22}-\mathbf{I}_{k-q})$ and $n^{1/2}\mathbf{E}_{12}=O_{\mathrm{P}}(1)=n^{1/2}\mathbf{E}'_{21}$ as $n\to\infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ [actually, Anderson (1963) proves this only for $\mathbf{E}=\boldsymbol{\beta}'\boldsymbol{\beta}_{\mathbf{S}}$ and under Gaussian densities, but his proof readily extends to the present situation]. Hence, still as $n\to\infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$,

$$\begin{aligned}
&n^{1/2}\mathbf{c}'_{p,q}\mathbf{H}_k\operatorname{vec}(\hat{\boldsymbol{\beta}}'_{\mathrm{Tyler}}\mathbf{V}\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}})\\
&\quad=-p\{n^{1/2}\lambda_1^*\operatorname{tr}(\mathbf{E}'_{11}\mathbf{E}_{11})+n^{1/2}\lambda_2^*\operatorname{tr}(\mathbf{E}'_{21}\mathbf{E}_{21})\}\\
&\qquad+(1-p)\{n^{1/2}\lambda_1^*\operatorname{tr}(\mathbf{E}'_{12}\mathbf{E}_{12})+n^{1/2}\lambda_2^*\operatorname{tr}(\mathbf{E}'_{22}\mathbf{E}_{22})\}\\
&\quad=-p\{n^{1/2}\lambda_1^*\operatorname{tr}(\mathbf{I}_q)\}+(1-p)\{n^{1/2}\lambda_2^*\operatorname{tr}(\mathbf{I}_{k-q})\}+o_{\mathrm{P}}(1)\\
&\quad=o_{\mathrm{P}}(1);
\end{aligned}\tag{A.19}$$

see (A.17). We conclude that the second term in (A.16) is $o_{\mathrm{P}}(1)$, so that

$$\begin{aligned}\underset{\sim}{T}_K^{(n)} &= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\tilde{\mathbf{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\\ &\quad\times \mathbf{H}_k\hat{\boldsymbol{\beta}}'^{\otimes 2}_{\mathrm{Tyler}}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}\mathbf{J}_k^{\perp}\operatorname{vec}\big(\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)}\big)\\ &\quad+o_{\mathrm{P}}(1)\\ &= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\tilde{\mathbf{\Lambda}}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\\ &\quad\times \mathbf{H}_k\mathbf{E}'^{\otimes 2}(\boldsymbol{\beta}')^{\otimes 2}(\tilde{\mathbf{V}}^{1/2})^{\otimes 2}\mathbf{J}_k^{\perp}\operatorname{vec}\big(\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)}\big)\\ &\quad+o_{\mathrm{P}}(1).\end{aligned}$$

Since $n^{1/2}\mathbf{J}_k^{\perp}\operatorname{vec}(\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)})$ is $O_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, Slutsky's lemma entails

$$\begin{aligned}\underset{\sim}{T}_K^{(n)} &= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\mathbf{\Lambda}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\\ &\quad\times \mathbf{H}_k(\operatorname{diag}(\mathbf{E}'_{11},\mathbf{E}'_{22}))^{\otimes 2}(\boldsymbol{\beta}')^{\otimes 2}\\ &\quad\times(\mathbf{V}^{1/2})^{\otimes 2}\mathbf{J}_k^{\perp}\operatorname{vec}\big(\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)}\big)+o_{\mathrm{P}}(1)\\ &= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2}(a_{p,q}(\mathbf{\Lambda}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\\ &\quad\times \mathbf{H}_k(\operatorname{diag}(\mathbf{E}'_{11},\mathbf{E}'_{22}))^{\otimes 2}(\boldsymbol{\beta}')^{\otimes 2}\\ &\quad\times(\mathbf{V}^{1/2})^{\otimes 2}\operatorname{vec}\big(\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)}\big)+o_{\mathrm{P}}(1),\end{aligned}\tag{A.20}$$

where we used the facts that $\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)}$ is $O_{\mathrm{P}}(1)$ and that

$$\begin{aligned}&n^{1/2}\mathbf{c}'_{p,q}\mathbf{H}_k(\operatorname{diag}(\mathbf{E}'_{11},\mathbf{E}'_{22}))^{\otimes 2}(\boldsymbol{\beta}')^{\otimes 2}(\mathbf{V}^{1/2})^{\otimes 2}(\operatorname{vec}\mathbf{I}_k)\\ &\quad= n^{1/2}\{-p\lambda_1^*\operatorname{tr}(\mathbf{E}_{11}\mathbf{E}'_{11})+(1-p)\lambda_2^*\operatorname{tr}(\mathbf{E}_{22}\mathbf{E}'_{22})\}\\ &\quad= n^{1/2}\{-p\lambda_1^*\operatorname{tr}(\mathbf{I}_q)+(1-p)\lambda_2^*\operatorname{tr}(\mathbf{I}_{k-q})\}+o_{\mathrm{P}}(1)\\ &\quad= o_{\mathrm{P}}(1).\end{aligned}$$

Then, putting [with the same partitioning as in (A.18)]

$$\boldsymbol{\beta}'\mathbf{V}^{1/2}\underset{\sim}{\mathbf{S}}_{\boldsymbol{\vartheta};K}^{(n)}\mathbf{V}^{1/2}\boldsymbol{\beta} =: \mathbf{D}_{\boldsymbol{\vartheta};K}^{(n)} =: \begin{pmatrix}(\mathbf{D}_{\boldsymbol{\vartheta};K}^{(n)})_{11} & (\mathbf{D}_{\boldsymbol{\vartheta};K}^{(n)})_{12}\\ (\mathbf{D}_{\boldsymbol{\vartheta};K}^{(n)})_{21} & (\mathbf{D}_{\boldsymbol{\vartheta};K}^{(n)})_{22}\end{pmatrix},$$

the asymptotic properties of $\underset{\sim}{\mathbf{S}}^{(n)}_{\boldsymbol{\vartheta};K}$ and $\mathbf{E}_{jj}$, $j = 1, 2$ imply that

$$\begin{aligned}
\underset{\sim}{T}^{(n)}_{K} &= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2} (a_{p,q}(\boldsymbol{\Lambda}_{\mathbf{V}}))^{-1/2}\{-p\,\mathrm{tr}(\mathbf{E}'_{11}(\mathbf{D}^{(n)}_{\boldsymbol{\vartheta};K})_{11}\mathbf{E}_{11}) \\
&\qquad + (1-p)\,\mathrm{tr}(\mathbf{E}'_{22}(\mathbf{D}^{(n)}_{\boldsymbol{\vartheta};K})_{22}\mathbf{E}_{22})\} + o_{\mathrm{P}}(1) \\
&= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2} (a_{p,q}(\boldsymbol{\Lambda}_{\mathbf{V}}))^{-1/2}\{-p\,\mathrm{tr}((\mathbf{D}^{(n)}_{\boldsymbol{\vartheta};K})_{11}) \\
&\qquad + (1-p)\,\mathrm{tr}((\mathbf{D}^{(n)}_{\boldsymbol{\vartheta};K})_{22})\} + o_{\mathrm{P}}(1) \\
&= \left(\frac{nk(k+2)}{\mathcal{J}_k(K)}\right)^{1/2} (a_{p,q}(\boldsymbol{\Lambda}_{\mathbf{V}}))^{-1/2}\mathbf{c}'_{p,q}\mathbf{H}_k(\boldsymbol{\beta}')^{\otimes 2}(\mathbf{V}^{1/2})^{\otimes 2}\,\mathrm{vec}(\underset{\sim}{\mathbf{S}}^{(n)}_{\boldsymbol{\vartheta};K}) + o_{\mathrm{P}}(1) \\
&= \underset{\sim}{T}^{(n)}_{\boldsymbol{\vartheta};K} + o_{\mathrm{P}}(1)
\end{aligned}$$

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, which establishes the result. □

PROOF OF PROPOSITION 6.2. Fix $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}'}_{0;1}$ and $g_1 \in \mathcal{F}_a$. We have already shown in Section 6.2 that $\underset{\sim}{Q}^{(n)}_{K} - \underset{\sim}{Q}^{(n)}_{\boldsymbol{\vartheta}_0,K} = o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$. Proposition 6.1(i) then yields

$$\begin{aligned}
\underset{\sim}{Q}^{(n)}_{K} &= \boldsymbol{\Delta}^{IV\prime}_{\boldsymbol{\vartheta}_0;K,g_1}[(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K})^{-} - \mathbf{P}^{\boldsymbol{\beta}^0}_{k}((\mathbf{P}^{\boldsymbol{\beta}^0}_{k})'\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K}\mathbf{P}^{\boldsymbol{\beta}^0}_{k})^{-}(\mathbf{P}^{\boldsymbol{\beta}^0}_{k})']\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;K,g_1} \\
&\qquad + o_{\mathrm{P}}(1),
\end{aligned} \tag{A.21}$$

still as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$. Now, since

$$\begin{aligned}
&\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K}[(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K})^{-} - \mathbf{P}^{\boldsymbol{\beta}^0}_{k}((\mathbf{P}^{\boldsymbol{\beta}^0}_{k})'\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K}\mathbf{P}^{\boldsymbol{\beta}^0}_{k})^{-}(\mathbf{P}^{\boldsymbol{\beta}^0}_{k})'] \\
&\qquad = \tfrac{1}{2}\mathbf{G}^{\boldsymbol{\beta}_0}_{k}\,\mathrm{diag}(\mathbf{I}_{k-1}, \mathbf{0}_{(k-2)(k-1)/2\times(k-2)(k-1)/2})(\mathbf{G}^{\boldsymbol{\beta}_0}_{k})'
\end{aligned}$$

is idempotent with rank $(k-1)$ [compare with (4.4)], it follows that $\underset{\sim}{Q}^{(n)}_{K}$ is asymptotically chi-square with $(k-1)$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, which establishes the null-hypothesis part of (i). For local alternatives, we restrict to those parameter values $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\beta}}_{0}$ for which we have ULAN. From contiguity, (A.21), also holds under alternatives of the form $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$. Le Cam's third lemma then implies that $\underset{\sim}{Q}^{(n)}_{K}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$, is asymptotically noncentral chi-square, still with $(k-1)$ degrees of freedom, but with noncentrality parameter

$$\begin{aligned}
&\lim_{n\to\infty}\{(\boldsymbol{\tau}^{IV(n)})' \\
&\quad\times [\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K,g_1}[(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K})^{-} - \mathbf{P}^{\boldsymbol{\beta}^0}_{k}((\mathbf{P}^{\boldsymbol{\beta}^0}_{k})'\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K}\mathbf{P}^{\boldsymbol{\beta}^0}_{k})^{-}(\mathbf{P}^{\boldsymbol{\beta}^0}_{k})']\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K,g_1}]\boldsymbol{\tau}^{IV(n)}\}.
\end{aligned}$$

Evaluation of this limit completes part (i) of the proof.

As for parts (ii) and (iii), the fact that $\underset{\sim}{\phi}^{(n)}_{\boldsymbol{\beta};K}$ has asymptotic level α directly follows from the asymptotic null distribution just established and the classical Helly–Bray theorem, while asymptotic optimality under K_{f_1} scores is a consequence of the asymptotic equivalence, under density f_1, of $\underset{\approx}{Q}^{(n)}_{K_{f_1}}$ and the optimal parametric test statistic for density f_1. $\square$

PROOF OF PROPOSITION 6.3. Fix $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\Lambda}\prime\prime}_{0;q}$ and $g_1 \in \mathcal{F}_a$. It directly follows from Lemma 6.1 and Proposition 6.1 that

$$\underset{\sim}{T}^{(n)}_{K} = (\operatorname{grad}' h(\operatorname{d\overset{\circ}{v}ec}\boldsymbol{\Lambda}^0_{\mathbf{V}})(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta}_0;K})^{-1}\operatorname{grad} h(\operatorname{d\overset{\circ}{v}ec}\boldsymbol{\Lambda}^0_{\mathbf{V}}))^{-1/2}$$

$$\times \operatorname{grad}' h(\operatorname{d\overset{\circ}{v}ec}\boldsymbol{\Lambda}^0_{\mathbf{V}})(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta}_0;K})^{-1}\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta}_0;K,g_1} + o_{\mathrm{P}}(1)$$

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$, hence also—provided that $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\Lambda}}_0$—under the contiguous sequences $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$. Parts (i) and (ii) result from the fact that $\underset{\sim}{\boldsymbol{\Delta}}^{III}_{\boldsymbol{\vartheta}_0;K,g_1}$ is asymptotically normal with mean zero under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;g_1}$ and mean

$$\lim_{n\to\infty}\{\mathcal{J}_k(K,g_1)/\big(k(k+2)\big)\mathbf{D}_k(\boldsymbol{\Lambda}_{\mathbf{V}})\boldsymbol{\tau}^{III(n)}\}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)};g_1}$ (Le Cam's third lemma; again, for $\boldsymbol{\vartheta}_0 \in \mathcal{H}^{\boldsymbol{\Lambda}}_0$), and with covariance matrix $\mathcal{J}_k(K)/(k(k+2))\mathbf{D}_k(\boldsymbol{\Lambda}_{\mathbf{V}})$ under both. Parts (iii) and (iv) follow as in the previous proof. $\square$

Acknowledgments. The authors very gratefully acknowledge the extremely careful and insightful editorial handling of this unusually long and technical paper. The original version received very detailed and constructive comments from two anonymous referees and a (no less anonymous) Associate Editor. Their remarks greatly helped improving the exposition.

REFERENCES

ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34** 122–148. MR0145620

ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley, New York. MR1990662

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. MR1789812

DAVIS, A. W. (1977). Asymptotic theory for principal component analysis: The nonnormal case. *Austral. J. Statist.* **19** 206–212. MR0501590

HALLIN, M., OJA, H. and PAINDAVEINE, D. (2006). Semiparametrically efficient rank-based inference for shape II: Optimal R-estimation of shape. *Ann. Statist.* **34** 2757–2789. MR2329466

HALLIN, M. and PAINDAVEINE, D. (2002). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks. *Ann. Statist.* **30** 1103–1133. MR1926170

HALLIN, M. and PAINDAVEINE, D. (2004). Rank-based optimal tests of the adequacy of an elliptic VARMA model. *Ann. Statist.* **32** 2642–2678. MR2153998

HALLIN, M. and PAINDAVEINE, D. (2005). Affine invariant aligned rank tests for the multivariate general linear model with ARMA errors. *J. Multivariate Anal.* **93** 122–163. MR2119768

HALLIN, M. and PAINDAVEINE, D. (2006a). Semiparametrically efficient rank-based inference for shape I: Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756. MR2329465

HALLIN, M. and PAINDAVEINE, D. (2006b). Parametric and semiparametric inference for shape: The role of the scale functional. *Statist. Decisions* **24** 1001–1023. MR2305111

HALLIN, M. and PAINDAVEINE, D. (2008a). Optimal rank-based tests for homogeneity of scatter. *Ann. Statist.* **36** 1261–1298. MR2418657

HALLIN, M. and PAINDAVEINE, D. (2008b). A general method for constructing pseudo-Gaussian tests. *J. Japan Statist. Soc.* **38** 27–40. MR2458315

HALLIN, M. and PURI, M. L. (1994). Aligned rank tests for linear models with autocorrelated errors. *J. Multivariate Anal.* **50** 175–237. MR1293044

HALLIN, M. and WERKER, B. J. M. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9** 137–165. MR1963675

HETTMANSPERGER, T. P. and RANDLES, R. H. (2002). A practical affine equivariant multivariate median. *Biometrika* **89** 851–860. MR1946515

HOTELLING, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24** 417–441.

JOLLIFFE, I. T. (1986). *Principal Component Analysis*. Springer, New York. MR0841268

KHURI, A. I. and GOOD, I. J. (1989). The parametrization of orthogonal matrices: A review mainly for statisticians. *South African Statist. J.* **23** 231–250. MR1054642

KOLLO, T. and NEUDECKER, H. (1993). Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices. *J. Multivariate Anal.* **47** 283–300. MR1247379

KREISS, J. P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133. MR0885727

LE CAM, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer, New York. MR0856411

LE CAM, L. and YANG, G. L. (2000). *Asymptotics in Statistics*, 2nd ed. Springer, New York. MR1784901

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. MR0570502

PAINDAVEINE, D. (2006). A Chernoff–Savage result for shape. On the nonadmissibility of pseudo-Gaussian methods. *J. Multivariate Anal.* **97** 2206–2220. MR2301635

PAINDAVEINE, D. (2008). A canonical definition of shape. *Statist. Probab. Lett.* **78** 2240–2247. MR2458033

PEARSON, K. (1901). On lines and planes of closest fit to system of points in space. *Philos. Magazine* **2** 559–572.

RAO, C. R. and MITRA, S. K. (1971). *Generalized Inverse of Matrices and Applications*. Wiley, New York. MR0338013

RIEDER, H. (1994). *Robust Asymptotic Statistics*. Springer, New York. MR1284041

SALIBIÁN-BARRERA, M., VAN AELST, S. and WILLEMS, G. (2006). PCA based on multivariate MM-estimators with fast and robust bootstrap. *J. Amer. Statist. Assoc.* **101** 1198–1211. MR2328307

SHAPIRO, A. and BROWNE, M. W. (1987). Analysis of covariance structures under elliptical distributions. *J. Amer. Statist. Assoc.* **82** 1092–1097. MR0922173

SCHOTT, J. R. (1991). A test for a specific principal component of a correlation matrix. *J. Amer. Statist. Assoc.* **86** 747–751. MR1147100

STRASSER, H. (1985). *Mathematical Theory of Statistics*. de Gruyter, Berlin. MR0812467

TYLER, D. E. (1981). Asymptotic inference for eigenvectors. *Ann. Statist.* **9** 725–736. MR0619278

TYLER, D. E. (1983). A class of asymptotic tests for principal component vectors. *Ann. Statist.* **11** 1243–1250. MR0720269

TYLER, D. E. (1987). A distribution-free M-estimator of multivariate scatter. *Ann. Statist.* **15** 234–251. MR0885734

VAN DER VAART, A. (1998). *Asymptotic Statistics*. Cambridge Univ. Press, Cambridge. MR1652247

YANAGIHARA, H., TONDA, T. and MATSUMOTO, C. (2005). The effects of nonnormality on asymptotic distributions of some likelihood ratio criteria for testing covariance structures under normal assumption. *J. Multivariate Anal.* **96** 237–264. MR2204977

M. HALLIN
INSTITUT DE RECHERCHE EN STATISTIQUE
E.C.A.R.E.S., CP114
UNIVERSITÉ LIBRE DE BRUXELLES
50, AVENUE F. D. ROOSEVELT
B-1050 BRUXELLES
BELGIUM
E-MAIL: mhallin@ulb.ac.be

D. PAINDAVEINE
INSTITUT DE RECHERCHE EN STATISTIQUE
E.C.A.R.E.S., CP114
AND DÉPARTEMENT DE MATHÉMATIQUE
UNIVERSITÉ LIBRE DE BRUXELLES
50, AVENUE F. D. ROOSEVELT
B-1050 BRUXELLES
BELGIUM
E-MAIL: dpaindav@ulb.ac.be
URL: http://homepages.ulb.ac.be/~dpaindav

T. VERDEBOUT
EQUIPPE-GREMARS
UNIVERSITÉ LILLE III
DOMAINE UNIVERSITAIRE DU PONT DE BOIS, BP 60149
F-59653 VILLENEUVE D'ASCQ CEDEX
FRANCE
E-MAIL: thomas.verdebout@univ-lille3.fr