Journal de la Société Française de Statistique
Vol. 153 No. 1 (2012)

Rank Tests for Elliptical Graphical Modeling

Titre : Tests de Rangs pour les Modèles Graphiques Elliptiques

Davy Paindaveine[1,2] and Thomas Verdebout[3]

Abstract: As a reaction to the restrictive Gaussian assumptions that are usually part of graphical models, Vogel and Fried [17] recently introduced *elliptical graphical models*, in which the vector of variables at hand is assumed to have an elliptical distribution. The present work introduces a class of rank tests in the context of elliptical graphical models. The proposed tests are valid under any elliptical density, and in particular do not require any moment assumption. They achieve local and asymptotic optimality under correctly specified densities. Their asymptotic properties are investigated both under the null and under sequences of local alternatives. Asymptotic relative efficiencies with respect to the corresponding pseudo-Gaussian competitors are derived, which allows to show that, when based on normal scores, the proposed rank tests uniformly dominate the pseudo-Gaussian tests in the Pitman sense. The asymptotic results are confirmed through a Monte-Carlo study.

Résumé : En réaction aux hypothèses gaussiennes restrictives qui accompagnent le plus souvent les modèles graphiques, Vogel et Fried [17] ont récemment introduit des modèles graphiques elliptiques, qui prévoient que les variables suivent conjointement une distribution elliptique. Le présent travail introduit une classe de tests de rangs dans le contexte de ces modèles graphiques elliptiques. Ces tests sont valides sous une densité elliptique quelconque, et en particulier ne requièrent aucune hypothèse de moment. Ils sont localement et asymptotiquement optimaux sous des densités correctement spécifiées. Leurs propriétés asymptotiques sont étudiées à la fois sous l'hypothèse nulle et sous des suites de contre-hypothèses locales. Leurs efficacités asymptotiques relatives par rapport à leurs compétiteurs pseudo-gaussiens sont calculées, ce qui permet de montrer que, lorsqu'ils sont basés sur des scores gaussiens, les tests de rangs proposés dominent uniformément les tests pseudo-gaussiens au sens de Pitman. Les résultats asymptotiques sont confirmés par une étude de Monte-Carlo.

Keywords: Conditional independence, Graphical models, Local asymptotic normality, Pseudo-Gaussian tests, Rank tests, Scatter matrix, Signed ranks
Mots-clés : Indépendance conditionnelle, Matrice de scatter, Modèles graphiques, Normalité locale asymptotique, Rangs signés, Tests de rangs, Tests pseudo-gaussiens
AMS 2000 subject classifications: 62H15, 62H20, 62G35

1. Introduction

Graphical modeling is one of the main tools that allow to understand the network of linear dependencies in a collection of random variables $X_1, \ldots, X_k$. It has many applications, especially in biometrics, where it is used to study gene association networks; see, e.g., [12, 13] and [16]. Classically, graphical modeling produces a graph G in which the k vertices are associated with the

[1] Université libre de Bruxelles
E-mail: `dpaindav@ulb.ac.be`
[2] Davy Paindaveine is also member of ECORE, the association between CORE and ECARES.
[3] Université Lille Nord de France
E-mail: `thomas.verdebout@univ-lille3.fr`

 ISSN: 2102-6238

random variables at hand and (undirected) edges indicate that the corresponding pair of variables are not conditionally independent (conditional on the $k-2$ remaining variables).

Standard graphical modeling is based on the assumption that $\mathbf{X} = (X_1, \ldots, X_k)'$ is multinormal. Conditional independence between X_j and X_ℓ then is equivalent to $(\mathbf{\Sigma}_{\rm cov}^{-1})_{j\ell} = 0$, where $\mathbf{\Sigma}_{\rm cov}$ stands for the covariance matrix of $\mathbf{X}$. This implies that determining the edges that should be part of G can be achieved by considering null hypotheses of the form

$$\mathscr{H}_0 : (\mathbf{\Sigma}_{\rm cov}^{-1})_{j\ell} = 0 \quad \text{or} \quad \mathscr{H}_0 : (\mathbf{\Sigma}_{\rm cov}^{-1})_{j_1\ell_1} = \ldots = (\mathbf{\Sigma}_{\rm cov}^{-1})_{j_p\ell_p} = 0, \tag{1}$$

which therefore are of primary importance in graphical modeling. As often in multivariate analysis, the daily-practice tests are the corresponding (Gaussian) likelihood ratio tests (LRTs).

Of course, multinormality is a very strong assumption, that is seldom compatible with real data. In the recent paper [17], Vogel and Fried introduced *elliptical graphical models*, that only impose that the distribution of $\mathbf{X}$ is elliptically symmetric. In that context, non-gaussianity implies that conditional independence is replaced with conditional uncorrelatedness, while the possible lack of finite second-order moments requires substituting a *scatter matrix* parameter $\mathbf{\Sigma}$ (see Section 2.1 below for a precise definition) for the covariance matrix $\mathbf{\Sigma}_{\rm cov}$. In that framework, Vogel and Fried defined robustified Gaussian LRTs that remain valid—in the sense that they asymptotically meet the level constraint—under a broad range of elliptical densities. More precisely, their tests, that are based on quadratic forms in an estimator $\hat{\mathbf{\Sigma}}$ of $\mathbf{\Sigma}$, remain valid under any elliptical distribution at which $\sqrt{n}\,\mathrm{vec}(\hat{\mathbf{\Sigma}} - \mathbf{\Sigma})$ is asymptotically normal with mean zero and a covariance matrix that can be consistently estimated from the sample. Choosing an adequate robust estimator $\hat{\mathbf{\Sigma}}$ then yields tests that do not require any moment assumption. However, to obtain a pseudo-Gaussian test (that is, a test that is robust to deviations from multinormality, yet is asymptotically equivalent to Gaussian LRTs in the multinormal case), one has to use the empirical covariance matrix for $\hat{\mathbf{\Sigma}}$, which leads to pseudo-Gaussian tests that require finite fourth-order moments.

The main objective of the present paper is to provide a class of tests that achieve local and asymptotic optimality under any fixed *target* density, yet remain valid under arbitrary elliptical densities, in the absence of any moment assumption. The proposed tests, that arise from invariance arguments, are based on the same multivariate signs and ranks as in [4, 5] and [7]. Denoting by $\mathbf{X}_1, \ldots, \mathbf{X}_n$ an observed n-tuple of k-dimensional elliptical vectors with location $\boldsymbol{\theta}$ and scatter matrix $\mathbf{\Sigma}$, these signs and ranks are (sample versions of)

1. the *standardized spatial signs*

$$\mathbf{U}_i := \frac{\mathbf{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})\|}, \qquad i = 1, \ldots, n,$$

   (throughout $\mathbf{A}^{1/2}$, for a symmetric and positive definite matrix $\mathbf{A}$, stands for the symmetric and positive definite root of $\mathbf{A}$) and

2. the *Mahalanobis ranks* $R_i^{(n)}$, $i = 1, \ldots, n$, where $R_i^{(n)}$ denotes the rank of $\|\mathbf{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})\|$ among $\|\mathbf{\Sigma}^{-1/2}(\mathbf{X}_1 - \boldsymbol{\theta})\|, \ldots, \|\mathbf{\Sigma}^{-1/2}(\mathbf{X}_n - \boldsymbol{\theta})\|$.

Of particular interest within the proposed class of rank tests is the *van der Waerden*—that is, *normal-score*—test, that (i) is asymptotically equivalent, under Gaussian densities, to the corresponding Gaussian LRTs, and (ii) uniformly dominates, in the Pitman sense, the locally and asymptotically optimal pseudo-Gaussian test.

The outline of the paper is as follows. In Section 2, we introduce the notation and the parametrization to be used in the sequel (Section 2.1) and exploit the *uniform local asymptotic normality* of the parametric submodels considered (Section 2.2) to build optimal parametric tests (Section 2.3). In Section 3, we turn the Gaussian version of these parametric tests into pseudo-Gaussian tests, that remain valid under any elliptical distribution with finite fourth-order moments. The rank tests we introduce in Section 4 are validity-robust in the sense that they remain valid under arbitrary elliptical densities, without any moment assumption. They are also efficiency-robust, as we show by deriving asymptotic relative efficiencies (Section 5.1) and by conducting a Monte-Carlo study (Section 5.2). Finally, the Appendix collects technical proofs.

2. Parametrization, ULAN, and optimal parametric tests

In this section, we describe the parametrization of elliptical families that will be relevant for graphical modeling and we state the corresponding uniform local asymptotic normality (ULAN) result, that is the key result in the construction of optimal tests in the context.

2.1. Parametrization

We throughout assume that the k-dimensional observations $\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}$ are elliptically symmetric. More precisely, defining

$$\mathscr{F} := \big\{ g : \mathbb{R}_0^+ \to \mathbb{R}^+ : \mu_{k-1,g} < \infty \big\},$$

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

$$\mathscr{F}_1 := \Big\{ g_1 \in \mathscr{F} \ : (\mu_{k-1,g_1})^{-1} \int_0^1 r^{k-1} g_1(r)\, dr = 1/2 \Big\},$$

we assume that the $\mathbf{X}_i^{(n)}$'s are mutually independent with a common probability density function of the form

$$\mathbf{x} \mapsto c_{k,g_1} (\det \mathbf{\Sigma})^{-1/2} g_1 \left(\sqrt{(\mathbf{x}-\boldsymbol{\theta})' \mathbf{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\theta})} \right), \qquad (2)$$

for some k-dimensional vector $\boldsymbol{\theta}$ (*location*), some symmetric and positive definite $(k \times k)$ matrix $\mathbf{\Sigma}$ (*scatter*), and some $g_1 \in \mathscr{F}_1$ (*standardized radial density*). This latter terminology is justified by the fact that the Mahalanobis distances $d_i(\boldsymbol{\theta}, \mathbf{\Sigma}) := ((\mathbf{X}_i - \boldsymbol{\theta})' \mathbf{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\theta}))^{1/2}$ are i.i.d. with probability density function $r \mapsto \tilde{g}_{1k}(r) := (\mu_{k-1,g_1})^{-1} r^{k-1} g_1(r) I_{[r>0]}$, where I_B stands for the indicator function of the set B. In the sequel, the corresponding cumulative distribution function will be denoted as $\tilde{G}_{1k}(r) = \int_0^r \tilde{g}_{1k}(s)\, ds$.

Special instances are the k-variate multinormal distribution, with radial density $g_1(r) = \phi_1(r) := \exp(-a_k r^2/2)$, the k-variate Student distributions, with radial densities (for $\nu \in \mathbb{R}_0^+$ degrees of freedom) $g_1(r) = f_{1,\nu}^t(r) := (1 + a_{k,\nu} r^2/\nu)^{-(k+\nu)/2}$, and the k-variate power-exponential distributions, with radial densities of the form $g_1(r) = f_{1,\eta}^e(r) := \exp(-b_{k,\eta} r^{2\eta})$, $\eta \in \mathbb{R}_0^+$; the positive constants a_k, $a_{k,\nu}$, and $b_{k,\eta}$ are such that $g_1 \in \mathscr{F}_1$.

Since graphical modeling is based on $\mathbf{\Sigma}^{-1}$, it is natural to replace the parametrization in $(\boldsymbol{\theta}, \mathbf{\Sigma}, g_1)$ with a parametrization in $(\boldsymbol{\theta}, \mathbf{\Sigma}^{-1}, g_1)$. Actually, we will further factorize $\mathbf{\Sigma}^{-1}$ into

$$\mathbf{\Sigma}^{-1} = \varsigma^2 \mathbf{W}, \qquad \text{with } \operatorname{tr}(\mathbf{W}) = k, \tag{3}$$

since the classical testing problems in graphical modeling, parallel to those associated with (1), are problems for which $\mathbf{W}$ is the parameter of interest, while ς^2 plays the role of a nuisance. The factorization in (3) mimics the decomposition of the scatter matrix $\mathbf{\Sigma} = \sigma^2 \mathbf{V}$ into a shape parameter $\mathbf{V}$ (with trace k, say) and a scale parameter σ^2, that has proved useful in many inference problems involving the shape of elliptical distributions; see, e.g., [4, 5, 6] and [7].

To sum up, the parametrization of elliptical families we will consider in the context of graphical modeling involves, beyond the infinite-dimensional radial density parameter $g_1 \in \mathscr{F}_1$, the finite-dimensional parameter

$$\boldsymbol{\xi} = (\boldsymbol{\theta}', \varsigma^2, (\overset{\circ}{\operatorname{vech}} \mathbf{W})')' \in \mathbf{\Theta} := \mathbb{R}^k \times \mathbb{R}_0^+ \times \overset{\circ}{\operatorname{vech}}(\mathscr{S}_k),$$

where $\mathscr{S}_k$ stands for the set of $k \times k$ symmetric and positive definite matrices with trace k and where $\operatorname{vech}(\mathbf{A}) := (\mathbf{A}_{11}, (\overset{\circ}{\operatorname{vech}} \mathbf{A})')' \in \mathbb{R}^{1+K}$ (with $K := k(k+1)/2 - 1$) denotes the vector stacking the upper-triangular elements of the matrix $\mathbf{A}$ (the upper left entry of $\mathbf{W}$ may be dropped from $\boldsymbol{\xi}$ since it can be obtained from the other entries of $\mathbf{W}$ by using the fact that $\mathbf{W}$ has trace k). In the sequel, we will write $\mathrm{P}^{(n)}_{\boldsymbol{\xi}, g_1}$ or $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \varsigma^2, \mathbf{W}, g_1}$ for the joint distribution of the mutually independent random vectors $\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}$ with common pdf

$$\mathbf{x} \mapsto c_{k, g_1} \varsigma^k (\det \mathbf{W})^{1/2} g_1 \left(\varsigma \sqrt{(\mathbf{x} - \boldsymbol{\theta})' \mathbf{W} (\mathbf{x} - \boldsymbol{\theta})} \right), \tag{4}$$

which is the pdf in (2) written in terms of the new parametrization. The semiparametric model considered for elliptical graphical modeling is then associated with the family of probability distributions

$$\mathscr{P}^{(n)} := \cup_{g_1 \in \mathscr{F}_1} \mathscr{P}^{(n)}_{g_1} := \cup_{g_1 \in \mathscr{F}_1} \cup_{\boldsymbol{\xi} \in \mathbf{\Theta}} \{\mathrm{P}^{(n)}_{\boldsymbol{\xi}, g_1}\}. \tag{5}$$

As mentioned in the Introduction, the derivation of optimal—at some fixed $g_1 = f_1$—tests for graphical modeling will be based on the ULAN property of the corresponding parametric submodel $\mathscr{P}^{(n)}_{f_1}$.

2.2. ULAN

As usual, ULAN requires some mild regularity conditions on f_1. More precisely, we need here that f_1 belongs to the collection $\mathscr{F}_a$ of all absolutely continuous densities in $\mathscr{F}_1$ 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

$$\mathscr{I}_k(f_1) := \int_0^\infty \varphi_{f_1}^2(r) \tilde{f}_{1k}(r)\, dr \quad \text{and} \quad \mathscr{J}_k(f_1) := \int_0^\infty r^2 \varphi_{f_1}^2(r) \tilde{f}_{1k}(r)\, dr \tag{6}$$

exist and are finite. The quantities $\mathscr{I}_k(f_1)$ and $\mathscr{J}_k(f_1)$ 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 [5], where we refer to for details.

Stating the ULAN result relevant for elliptical graphical models requires introducing some further notation. Denoting by $\mathbf{e}_\ell$ the ℓth vector of the canonical basis of $\mathbb{R}^k$, let $\mathbf{K}_k := \sum_{i,j=1}^k (\mathbf{e}_i\mathbf{e}_j') \otimes (\mathbf{e}_j\mathbf{e}_i')$ be the $k^2 \times k^2$ *commutation matrix*, put $\mathbf{J}_k := (\mathrm{vec}\,\mathbf{I}_k)(\mathrm{vec}\,\mathbf{I}_k)'$, and write $\mathbf{A}^{\otimes 2}$ for the Kronecker product $\mathbf{A} \otimes \mathbf{A}$. Finally, let $\mathbf{M}_k$ be the $K \times k^2$ matrix such that $\mathbf{M}_k'(\mathrm{v\mathring{e}ch}\,\mathbf{w}) = \mathrm{vec}\,\mathbf{w}$ for any symmetric $k \times k$ matrix $\mathbf{w}$ satisfying $\mathrm{tr}(\mathbf{w}) = 0$. Applying Lemma A.1 from [7] to Proposition 2.1 of [5] then yields the following result.

Theorem 2.1. *For any $f_1 \in \mathscr{F}_a$, the family $\mathscr{P}_{f_1}^{(n)} = \{\mathrm{P}_{\boldsymbol{\xi},f_1}^{(n)} : \boldsymbol{\xi} = (\boldsymbol{\theta}', \varsigma^2, (\mathrm{v\mathring{e}ch}\,\mathbf{W})')' \in \boldsymbol{\Theta}\}$ is ULAN, with* central sequence

$$\boldsymbol{\Delta}_{\boldsymbol{\xi},f_1}^{(n)} = \left(\boldsymbol{\Delta}_{\boldsymbol{\xi},f_1;1}^{(n)\prime}, \Delta_{\boldsymbol{\xi},f_1;2}^{(n)}, \boldsymbol{\Delta}_{\boldsymbol{\xi},f_1;3}^{(n)\prime}\right)',$$

where (letting $d_i = d_i(\boldsymbol{\theta},\mathbf{W}) := ((\mathbf{X}_i - \boldsymbol{\theta})'\mathbf{W}(\mathbf{X}_i - \boldsymbol{\theta}))^{1/2}$ and $\mathbf{U}_i = \mathbf{U}_i(\boldsymbol{\theta},\mathbf{W}) := \mathbf{W}^{1/2}(\mathbf{X}_i - \boldsymbol{\theta})/d_i$)

$$\boldsymbol{\Delta}_{\boldsymbol{\xi},f_1;1}^{(n)} := \frac{\varsigma}{\sqrt{n}} \sum_{i=1}^n \varphi_{f_1}(\varsigma d_i)\,\mathbf{W}^{1/2}\mathbf{U}_i, \quad \Delta_{\boldsymbol{\xi},f_1;2}^{(n)} := -\frac{\varsigma^2}{2\sqrt{n}} \sum_{i=1}^n (\varsigma d_i\, \varphi_{f_1}(\varsigma d_i) - k),$$

and

$$\boldsymbol{\Delta}_{\boldsymbol{\xi},f_1;3}^{(n)} := -\frac{1}{2\sqrt{n}} \mathbf{M}_k \left(\mathbf{W}^{\otimes 2}\right)^{-1/2} \sum_{i=1}^n \mathrm{vec}\left(\varsigma d_i\, \varphi_{f_1}(\varsigma d_i)\,\mathbf{U}_i\mathbf{U}_i' - \mathbf{I}_k\right),$$

and full-rank information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1} := \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Gamma_{\boldsymbol{\xi},f_1;22} & \boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;32}' \\ \mathbf{0} & \boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;32} & \boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;33} \end{pmatrix},$$

where

$$\boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;11} := \frac{\varsigma^2 \mathscr{I}_k(f_1)}{k}\mathbf{W},$$

$$\Gamma_{\boldsymbol{\xi},f_1;22} := \frac{\varsigma^4(\mathscr{I}_k(f_1) - k^2)}{4}, \quad \boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;32} := \frac{\varsigma^2(\mathscr{I}_k(f_1) - k^2)}{4k}\mathbf{M}_k(\mathrm{vec}\,\mathbf{W}^{-1}),$$

and

$$\boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1;33} := \frac{1}{4}\mathbf{M}_k\left(\mathbf{W}^{\otimes 2}\right)^{-1/2}\left[\frac{\mathscr{I}_k(f_1)}{k(k+2)}(\mathbf{I}_{k^2} + \mathbf{K}_k + \mathbf{J}_k) - \mathbf{J}_k\right]\left(\mathbf{W}^{\otimes 2}\right)^{-1/2}\mathbf{M}_k'.$$

More precisely, for any $\boldsymbol{\xi}^{(n)} = \boldsymbol{\xi} + O(n^{-1/2})$ and any bounded sequence $\boldsymbol{\tau}^{(n)} \in \mathbb{R}^{k+K+1}$, we have that, as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\xi}^{(n)},f_1}^{(n)}$,

$$\log\left(d\mathrm{P}_{\boldsymbol{\xi}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)},f_1}^{(n)} / d\mathrm{P}_{\boldsymbol{\xi}^{(n)},f_1}^{(n)}\right) = (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}_{\boldsymbol{\xi}^{(n)},f_1}^{(n)} - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1)$$

and $\boldsymbol{\Delta}_{\boldsymbol{\xi}^{(n)},f_1} \xrightarrow{\mathscr{L}} \mathscr{N}_{k+K+1}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\xi},f_1})$.

The block-diagonal structure of the information matrix $\mathbf{\Gamma}_{\boldsymbol{\xi},f_1}$ implies that the unspecification of $\boldsymbol{\theta}$ has no asymptotic cost when performing inference on ς^2 and/or $\mathbf{W}$. In contrast, the non-zero asymptotic covariance $\mathbf{\Gamma}_{\boldsymbol{\xi},f_1;32}$ between the ς^2- and $\mathbf{W}$-parts of the central sequence entails that the unspecification of ς^2 will have an asymptotic cost when performing inference on $\mathbf{W}$. In such a situation, asymptotic inference should be based on the f_1-*efficient central sequence*

$$\begin{aligned}\mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1} := \mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1;3} &:= \mathbf{\Delta}^{(n)}_{\boldsymbol{\xi},f_1;3} - \mathbf{\Gamma}_{\boldsymbol{\xi},f_1;32}\mathbf{\Gamma}^{-1}_{\boldsymbol{\xi},f_1;22}\Delta^{(n)}_{\boldsymbol{\xi},f_1;2} \\ &= -\frac{1}{2\sqrt{n}}\mathbf{M}_k\big(\mathbf{W}^{\otimes 2}\big)^{-1/2}\sum_{i=1}^{n}\varsigma d_i\,\varphi_{f_1}(\varsigma d_i)\,\mathrm{vec}\Big(\mathbf{U}_i\mathbf{U}_i' - \frac{1}{k}\mathbf{I}_k\Big),\end{aligned} \tag{7}$$

which, under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},f_1}$, is asymptotically normal with mean zero and covariance matrix

$$\begin{aligned}\mathbf{\Gamma}^{\star}_{\boldsymbol{\xi},f_1} := \mathbf{\Gamma}^{\star}_{\boldsymbol{\xi},f_1;33} &:= \mathbf{\Gamma}_{\boldsymbol{\xi},f_1;33} - \mathbf{\Gamma}_{\boldsymbol{\xi},f_1;32}\mathbf{\Gamma}^{-1}_{\boldsymbol{\xi},f_1;22}\mathbf{\Gamma}'_{\boldsymbol{\xi},f_1;32} \\ &= \frac{\mathscr{J}_k(f_1)}{4k(k+2)}\mathbf{M}_k\big(\mathbf{W}^{\otimes 2}\big)^{-1/2}\Big[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\Big]\big(\mathbf{W}^{\otimes 2}\big)^{-1/2}\mathbf{M}_k' \\ &=: \mathscr{J}_k(f_1)\mathbf{G}_k(\mathbf{W});\end{aligned} \tag{8}$$

throughout this paper, inference will be about $\mathbf{W}$, which allows to make the notation a bit lighter by dropping the indices 3 and 33 in efficient central sequences and information matrices, respectively.

2.3. Optimal parametric tests

We focus on the problem of testing *linear restrictions* on $\mathbf{W}$, which covers most testing problems of interest in graphical modeling. More precisely, we consider a generic problem of the form

$$\begin{cases}\mathscr{H}_0 : (\mathrm{v\mathring{e}ch}\,\mathbf{W}) \in \mathscr{M}(\mathbf{\Upsilon})\ (\cap\ \mathrm{v\mathring{e}ch}\,(\mathscr{S}_k)) \\ \mathscr{H}_1 : (\mathrm{v\mathring{e}ch}\,\mathbf{W}) \notin \mathscr{M}(\mathbf{\Upsilon})\ (\cap\ \mathrm{v\mathring{e}ch}\,(\mathscr{S}_k))\end{cases} \quad \left(\text{equivalently, } \begin{cases}\mathscr{H}_0 : \boldsymbol{\xi} \in \Theta_0(\mathbf{\Upsilon}) \\ \mathscr{H}_1 : \boldsymbol{\xi} \notin \Theta_0(\mathbf{\Upsilon})\end{cases}\right), \tag{9}$$

where $\mathbf{\Upsilon}$ is a given (arbitrary) $K \times (K-r)$ matrix with full rank $K-r$, and where $\mathscr{M}(\mathbf{\Upsilon})$ stands for the vector space that is spanned by the columns of $\mathbf{\Upsilon}$. Under the null, $\mathbf{W}$ satisfies a set of r independent linear constraints. Note that the null hypotheses in (1)—when formulated in terms of the scatter matrix $\mathbf{\Sigma}$—are of this form : for instance, the null hypotheses $\mathscr{H}_0 : (\mathbf{\Sigma}^{-1})_{12} = 0$ and $\mathscr{H}_0 : (\mathbf{\Sigma}^{-1})_{12} = (\mathbf{\Sigma}^{-1})_{13} = 0$—equivalently, $\mathscr{H}_0 : (\mathbf{W})_{12} = 0$ and $\mathscr{H}_0 : (\mathbf{W})_{12} = (\mathbf{W})_{13} = 0$—are respectively associated with

$$\mathbf{\Upsilon} = \begin{pmatrix}\mathbf{0}_{1\times(K-1)} \\ \mathbf{I}_{K-1}\end{pmatrix} \quad \text{and} \quad \mathbf{\Upsilon} = \mathbf{L}^{(2,3)}_K\mathbf{\Pi} := \mathbf{L}^{(2,3)}_K\begin{pmatrix}\mathbf{0}_{2\times(K-2)} \\ \mathbf{I}_{K-2}\end{pmatrix},$$

where $\mathbf{L}^{(2,3)}_K$ is the $K \times K$ permutation matrix that exchanges the second and third rows of $\mathbf{\Pi}$. More generally, this class of testing problems allows to test that any fixed collection of off-diagonal entries of $\mathbf{W}$ is only made of zeroes.

It follows from [11] that, when considering the testing problem in (9) at asymptotic level α, the locally and asymptotically f_1-optimal test $\phi_{\boldsymbol{\xi},f_1}$—more precisely, the locally and asymptotically most stringent test in $\mathscr{P}^{(n)}_{f_1}$—rejects the null whenever

$$Q_{\boldsymbol{\xi},f_1} := (\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1})' \left[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},f_1})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},f_1}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right]\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1} \tag{10}$$
$$= \frac{1}{\mathscr{J}_k(f_1)}(\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1})' \left[(\mathbf{G}_k(\mathbf{W}))^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right]\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1} > \chi^2_{r;1-\alpha},$$

where $\chi^2_{r;1-\alpha}$ denotes the α-upper quantile of the chi-square distribution with r degrees of freedom. Of course, $\boldsymbol{\xi}$ remains unspecified under the null, hence should be replaced with an appropriate estimator $\hat{\boldsymbol{\xi}}$, which leads to the statistic Q_{f_1} and to the test ϕ_{f_1}, say. While this test achieves (local and asymptotic) optimality at f_1, it is of rather limited practical value since it is usually not valid under $g_1 \neq f_1$—in the sense that, under $g_1 \neq f_1$, its asymptotic size is not α under the null.

3. Pseudo-Gaussian tests

In view of the central role played by multinormal densities in classical multivariate analysis, the Gaussian version of the tests defined above, namely ϕ_{f_1} with $f_1 = \phi_1$, is of particular interest. As already mentioned, however, this test is of limited practical value since it is valid under multinormal densities only—or more precisely, as we will show, under elliptical densities with Gaussian kurtosis only. Now, it turns out that it is possible to extend the validity of this Gaussian test to the whole class of elliptical densities with finite fourth-order moments, while maintaining its optimality properties at the multinormal. This section defines and studies the resulting so-called *pseudo-Gaussian* tests.

As, e.g., in [1, 2] or [3], we define the *kurtosis* of the elliptical density in (4) as

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

with

$$D_k(g_1) := \frac{\mu_{k+1,g_1}}{\mu_{k-1,g_1}} \left(= \mathrm{E}_{\boldsymbol{\xi},g_1}[\varsigma^2 d_i^2(\boldsymbol{\theta},\mathbf{W})] \right) \quad \text{and} \quad E_k(g_1) := \frac{\mu_{k+3,g_1}}{\mu_{k-1,g_1}} \left(= \mathrm{E}_{\boldsymbol{\xi},g_1}[\varsigma^4 d_i^4(\boldsymbol{\theta},\mathbf{W})] \right);$$

this clearly requires that $g_1 \in \mathscr{F}_1^{(4)} := \{g_1 \in \mathscr{F}_1 : \mu_{k+3,g_1} < \infty\}$, that is, the elliptical distributions considered need to have finite fourth-order moments. For Gaussian densities, $E_k(\phi_1) = k(k+2)/a_k^2$, $D_k(\phi_1) = k/a_k$, which leads to $\kappa_k(\phi_1) = 0$. Positive (resp., negative) kurtosis values $\kappa_k(g_1)$ would therefore indicate tails that are heavier (resp., lighter) than in the multinormal case.

Robustifying Gaussian tests into pseudo-Gaussian ones requires investigating the asymptotic behavior of the Gaussian efficient central sequence $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1}$. Letting $\mathbf{S}^{(n)}_{\boldsymbol{\theta}} := \frac{1}{n}\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})'$, first note that this central sequence rewrites

$$\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1} = -\frac{a_k\varsigma^2}{2\sqrt{n}}\mathbf{M}_k\left(\mathbf{W}^{\otimes 2}\right)^{-1/2}\sum_{i=1}^n d_i^2 \mathrm{vec}\left(\mathbf{U}_i\mathbf{U}_i' - \frac{1}{k}\mathbf{I}_k\right) \tag{11}$$
$$= -\frac{a_k\varsigma^2\sqrt{n}}{2}\mathbf{M}_k\left[\mathbf{I}_{k^2} - \frac{1}{k}(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}})(\mathrm{vec}\boldsymbol{\Sigma}^{-1}_{\mathrm{cov}})'\right]\mathrm{vec}\left(\mathbf{S}^{(n)}_{\boldsymbol{\theta}}\right) =: a_k\varsigma^2\mathbf{T}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathrm{cov}}},$$

where $\mathbf{\Sigma}_{\text{cov}}$ denotes the common covariance matrix of the $\mathbf{X}_i$'s. The asymptotic properties of $\mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1}$, under any elliptical density with finite fourth-order moments, are given in the following result.

Lemma 3.1. *Fix* $\boldsymbol{\xi} = (\boldsymbol{\theta}', \varsigma^2, (\overset{\circ}{\text{vech}}\,\mathbf{W})')' \in \mathbf{\Theta}$, $g_1 \in \mathscr{F}_a^{(4)} := \mathscr{F}_1^{(4)} \cap \mathscr{F}_a$, *and a bounded sequence* $\boldsymbol{\tau}^{(n)} = ((\mathbf{t}^{(n)})', s^{(n)}, (\overset{\circ}{\text{vech}}\,\mathbf{w}^{(n)})')'$ *in* $\mathbb{R}^{k+K+1}$ *such that* $\overset{\circ}{\text{vech}}\,\mathbf{w} := \lim_{n\to\infty} \overset{\circ}{\text{vech}}\,\mathbf{w}^{(n)}$ *exists. Then, under* $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$,

$$\mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1} \xrightarrow{\mathscr{L}} \mathscr{N}_K(\boldsymbol{\mu}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},\phi_1}, \mathbf{\Gamma}^{\star g_1}_{\boldsymbol{\xi},\phi_1}),$$

where $\boldsymbol{\mu}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},\phi_1} := a_k(k+2)D_k(g_1)\mathbf{G}_k(\mathbf{W})(\overset{\circ}{\text{vech}}\,\mathbf{w})$ *and* $\mathbf{\Gamma}^{\star g_1}_{\boldsymbol{\xi},\phi_1} := a_k^2 E_k(g_1)\mathbf{G}_k(\mathbf{W})$; *for* $\boldsymbol{\tau}^{(n)} = \mathbf{0}$, *the claim actually only requires that* $g_1 \in \mathscr{F}_1^{(4)}$.

In view of the Gaussian ($f_1 = \phi_1$) version of (10), Lemma 3.1 leads to considering the pseudo-Gaussian test "statistic"

$$\begin{aligned} Q^{g_1}_{\boldsymbol{\xi},\mathscr{N}} &:= (\mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1})' \left[(\mathbf{\Gamma}^{\star g_1}_{\boldsymbol{\xi},\phi_1})^{-1} - \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Gamma}^{\star g_1}_{\boldsymbol{\xi},\phi_1}\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}' \right] \mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1} \qquad (12)\\ &= \frac{1}{a_k^2 E_k(g_1)} (\mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1})' \left[(\mathbf{G}_k(\mathbf{W}))^{-1} - \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{G}_k(\mathbf{W})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}' \right] \mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1} \\ &= \frac{\varsigma^4}{E_k(g_1)} (\mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}})' \left[(\mathbf{G}_k(\mathbf{W}))^{-1} - \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{G}_k(\mathbf{W})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}' \right] \mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}}, \end{aligned}$$

which, by using the fact that $\mathbf{\Sigma}_{\text{cov}} = \frac{1}{k}D_k(g_1)\mathbf{\Sigma} = \frac{1}{k\varsigma^2}D_k(g_1)\mathbf{W}^{-1}$, rewrites as

$$\begin{aligned} Q^{g_1}_{\boldsymbol{\xi},\mathscr{N}} &= \frac{D_k^2(g_1)}{k^2 E_k(g_1)} (\mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}})' \left[(\mathbf{G}_k(\mathbf{\Sigma}^{-1}_{\text{cov}}))^{-1} - \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{G}_k(\mathbf{\Sigma}^{-1}_{\text{cov}})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}' \right] \mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}} \\ &= \frac{1}{k(k+2)(1+\kappa_k(g_1))} (\mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}})' \left[(\mathbf{G}_k(\mathbf{\Sigma}^{-1}_{\text{cov}}))^{-1} - \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{G}_k(\mathbf{\Sigma}^{-1}_{\text{cov}})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}' \right] \mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}}. \end{aligned}$$

In practice, adequate estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{\Sigma}}_{\text{cov}}$ need to be substituted for $\boldsymbol{\theta}$ and $\mathbf{\Sigma}_{\text{cov}}$, respectively. The following result allows to control the impact of this substitution.

Lemma 3.2. *Fix* $\boldsymbol{\xi} \in \Theta$ *and* $g_1 \in \mathscr{F}_1^{(4)}$, *and assume that* $\hat{\boldsymbol{\theta}}$ *and* $\hat{\mathbf{\Sigma}}_{\text{cov}}$ *are root-n consistent for* $\boldsymbol{\theta}$ *and* $\mathbf{\Sigma}_{\text{cov}}$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$. *Then, letting* $\hat{\mathbf{W}}_{\text{cov}} := k\hat{\mathbf{\Sigma}}^{-1}_{\text{cov}}/\text{tr}\,(\hat{\mathbf{\Sigma}}^{-1}_{\text{cov}})$,

$$\mathbf{T}^{(n)}_{\hat{\boldsymbol{\theta}},\hat{\mathbf{\Sigma}}_{\text{cov}}} - \mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma}_{\text{cov}}} = -(k+2)\,(\text{tr}\,\mathbf{\Sigma}^{-1}_{\text{cov}})\,\mathbf{G}_k(\mathbf{\Sigma}^{-1}_{\text{cov}})\sqrt{n}\,\overset{\circ}{\text{vech}}\,(\hat{\mathbf{W}}_{\text{cov}} - \mathbf{W}) + o_{\mathrm{P}}(1),$$

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

As shown in the proof of Theorem 3.1 below, it follows from Lemma 3.2 that the replacement of $\boldsymbol{\theta}$ and $\mathbf{\Sigma}_{\text{cov}}$ with root-n consistent estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{\Sigma}}_{\text{cov}}$ will have no effect on the null asymptotic behavior of $Q^{g_1}_{\boldsymbol{\xi},\mathscr{N}}$ in probability, provided that the estimator $\hat{\mathbf{\Sigma}}_{\text{cov}}$ is constrained—in the sense that the corresponding $\hat{\mathbf{W}}_{\text{cov}}$ satisfies the null constraint in (9). Such an estimate can of course be obtained by projecting a preliminary estimator $\hat{\mathbf{\Sigma}}_{\rm cov}$ on the null constraint, that is by replacing $\hat{\mathbf{\Sigma}}_{\rm cov}$ with

$$\hat{\mathbf{\Sigma}}_{\rm cov;0} := \frac{k}{\mathrm{tr}\,(\hat{\mathbf{\Sigma}}_{\rm cov}^{-1})}\hat{\mathbf{W}}_{\rm cov;0}^{-1},$$

where $\hat{\mathbf{W}}_{\rm cov;0}$ is the statistic with values in $\mathscr{S}_k$ defined through

$$(\overset{\circ}{\mathrm{vech}}\,\hat{\mathbf{W}}_{\rm cov;0}) := \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}'\,\overset{\circ}{\mathrm{vech}}\left(\frac{k}{\mathrm{tr}\,(\hat{\mathbf{\Sigma}}_{\rm cov}^{-1})}\hat{\mathbf{\Sigma}}_{\rm cov}^{-1}\right).$$

Clearly, root-n consistency of $\hat{\mathbf{\Sigma}}_{\rm cov}$ will entail root-n consistency of $\hat{\mathbf{\Sigma}}_{\rm cov;0}$ under the null. The resulting pseudo-Gaussian test—$\phi_{\mathscr{N}}$, say—rejects the null at asymptotic level α whenever

$$Q_{\mathscr{N}} := \frac{1}{k(k+2)(1+\hat{\kappa}_k)}(\mathbf{T}^{(n)}_{\hat{\boldsymbol{\theta}},\hat{\mathbf{\Sigma}}_{\rm cov;0}})'\left[(\mathbf{G}_k(\hat{\mathbf{\Sigma}}_{\rm cov;0}^{-1}))^{-1} - \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{G}_k(\hat{\mathbf{\Sigma}}_{\rm cov;0}^{-1})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}'\right]\mathbf{T}^{(n)}_{\hat{\boldsymbol{\theta}},\hat{\mathbf{\Sigma}}_{\rm cov;0}} > \chi^2_{r;1-\alpha},$$

where $\hat{\kappa}_k := \frac{k}{k+2}(\frac{1}{n}\sum_{i=1}^n((\mathbf{X}_i-\hat{\boldsymbol{\theta}})'\hat{\mathbf{\Sigma}}_{\rm cov;0}^{-1}(\mathbf{X}_i-\hat{\boldsymbol{\theta}}))^2)/(\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_i-\hat{\boldsymbol{\theta}})'\hat{\mathbf{\Sigma}}_{\rm cov;0}^{-1}(\mathbf{X}_i-\hat{\boldsymbol{\theta}}))^2-1$, for any $g_1 \in \mathscr{F}_a^{(4)}$, consistently estimates $\kappa_k(g_1)$ under $\cup_{\boldsymbol{\xi}\in\Theta_0(\mathbf{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$. The following result summarizes the asymptotic properties of $\phi_{\mathscr{N}}$.

Theorem 3.1. *(i) Under $\cup_{\boldsymbol{\xi}\in\Theta_0(\mathbf{\Upsilon})}\cup_{g_1\in\mathscr{F}_1^{(4)}}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$, $Q_{\mathscr{N}}$ is asymptotically chi-square with r degrees of freedom;*

(ii) under $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$, with $\boldsymbol{\tau}^{(n)}$ as in Lemma 3.1, $\boldsymbol{\xi}\in\Theta_0(\mathbf{\Upsilon})$, and $g_1\in\mathscr{F}_a^{(4)}$, $Q_{\mathscr{N}}$ is asymptotically non-central chi-square with r degrees of freedom and non-centrality parameter

$$\frac{k(k+2)}{1+\kappa_k(g_1)}(\overset{\circ}{\mathrm{vech}}\,\mathbf{w})'\left[\mathbf{G}_k(\mathbf{W}) - \mathbf{G}_k(\mathbf{W})\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{G}_k(\mathbf{W})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}'\mathbf{G}_k(\mathbf{W})\right](\overset{\circ}{\mathrm{vech}}\,\mathbf{w});$$

(iii) $\phi_{\mathscr{N}}$ has asymptotic level α under $\cup_{\boldsymbol{\xi}\in\Theta_0(\mathbf{\Upsilon})}\cup_{g_1\in\mathscr{F}_1^{(4)}}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$;

(iv) $\phi_{\mathscr{N}}$ is locally and asymptotically most stringent, still at asymptotic level α, for $\cup_{\boldsymbol{\xi}\in\Theta_0(\mathbf{\Upsilon})}\cup_{g_1\in\mathscr{F}_1^{(4)}}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$ against alternatives of the form $\cup_{\boldsymbol{\xi}\notin\Theta_0(\mathbf{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},\phi_1}\}$.

Wrapping up, we defined in this section a pseudo-Gaussian test $\phi_{\mathscr{N}}$ for $\mathscr{H}_0$. This test achieves local asymptotic optimality in the multinormal case and remains valid under any elliptical density with finite fourth-order moments.

4. Rank tests

In the previous section, pseudo-Gaussian test statistics were obtained by building quadratic forms in the Gaussian efficient central sequences $\mathbf{\Delta}^{\star(n)}_{\boldsymbol{\xi},\phi_1}$. Distribution-freeness there was achieved by estimating the asymptotic covariance matrix of this central sequence at the underlying radial density g_1, which required estimating the kurtosis $\kappa_k(g_1)$ of g_1.

Another, more elegant, way of eliminating the nuisance g_1 is to exploit the strong invariance structure of the model considered. For any fixed values of $\boldsymbol{\theta}$ and $\mathbf{W}$, the corresponding submodel

$$\mathscr{P}^{(n)}_{\boldsymbol{\theta},\mathbf{W}} := \cup_{\varsigma\in\mathbb{R}^+_0} \cup_{g_1\in\mathscr{F}_1} \{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\varsigma^2,\mathbf{W},g_1}\}$$

is indeed invariant under the group $\mathscr{G}^{(n)}_{\boldsymbol{\theta},\mathbf{W}} = \{g^{(n)h}_{\boldsymbol{\theta},\mathbf{W}} : h \in \mathscr{H}\},\circ$ of continuous monotone radial transformations of the form

$$\begin{aligned} g^{(n)h}_{\boldsymbol{\theta},\mathbf{W}}(\mathbf{X}_1,\ldots,\mathbf{X}_n) &= g_h(\boldsymbol{\theta}+d_1\mathbf{W}^{-1/2}\mathbf{U}_1,\ldots,\boldsymbol{\theta}+d_n\mathbf{W}^{-1/2}\mathbf{U}_n) \\ &:= (\boldsymbol{\theta}+h(d_1)\mathbf{W}^{-1/2}\mathbf{U}_1,\ldots,\boldsymbol{\theta}+h(d_n)\mathbf{W}^{-1/2}\mathbf{U}_n), \end{aligned}$$

where the $d_i = d_i(\boldsymbol{\theta},\mathbf{W})$'s and the $\mathbf{U}_i = \mathbf{U}_i(\boldsymbol{\theta},\mathbf{W})$'s are the quantities involved in Theorem 2.1, and where $\mathscr{H}$ denotes the collection of the mappings $h\colon \mathbb{R}^+ \to \mathbb{R}^+$ that are continuous, monotone increasing, and satisfy $h(0)=0$ and $\lim_{z\to\infty} h(z)=\infty$. This group actually is a generating group for $\mathscr{P}^{(n)}_{\boldsymbol{\theta},\mathbf{W}}$, which implies that invariant statistics are distribution-free in $\mathscr{P}^{(n)}_{\boldsymbol{\theta},\mathbf{W}}$. This is what leads to considering signed rank tests (below, we simply speak of *rank tests*), since invariant statistics are exactly those that are measurable with respect to

$$(\mathbf{U}_1,\ldots,\mathbf{U}_n,R^{(n)}_1,\ldots,R^{(n)}_n),$$

where $R^{(n)}_i = R^{(n)}_i(\boldsymbol{\theta},\mathbf{W})$ denotes the rank of d_i among $d_1,\ldots,d_n$; see [5] for details.

The rank tests we propose will be based on linear rank statistics of the form

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K} = -\frac{1}{2\sqrt{n}}\mathbf{M}_k\big(\mathbf{W}^{\otimes 2}\big)^{-1/2}\sum_{i=1}^{n} K\left(\frac{R^{(n)}_i}{n+1}\right)\mathrm{vec}\Big(\mathbf{U}_i\mathbf{U}_i' - \frac{1}{k}\mathbf{I}_k\Big),$$

where the *score function* $K : (0,1) \to \mathbb{R}$ is continuous and square-integrable, and can be written as the difference of two monotone increasing functions. As Lemma 4.1(i) below shows, $\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$, is asymptotically equivalent in probability to

$$\boldsymbol{\Delta}^{\star g_1(n)}_{\boldsymbol{\xi},K} := -\frac{1}{2\sqrt{n}}\mathbf{M}_k\big(\mathbf{W}^{\otimes 2}\big)^{-1/2}\sum_{i=1}^{n} K\big(\tilde{G}_{1k}(d_i(\boldsymbol{\theta},\mathbf{W}))\big)\mathrm{vec}\Big(\mathbf{U}_i\mathbf{U}_i' - \frac{1}{k}\mathbf{I}_k\Big).$$

In particular, $\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K_{f_1}}$, with $K_{f_1}(u) := \tilde{F}^{-1}_{1k}(u)\varphi_{f_1}(\tilde{F}^{-1}_{1k}(u))$ for all u, is a rank-based version of the efficient central sequence $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\xi},f_1}$ in (7), hence can serve as the basis for the construction of optimal (at f_1) rank tests.

The asymptotic behavior of $\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K}$ is described in the following result.

Lemma 4.1. Fix $\boldsymbol{\xi} = (\boldsymbol{\theta}',\varsigma^2,(\overset{\circ}{\mathrm{vech}}\,\mathbf{W})')' \in \boldsymbol{\Theta}$ and $\boldsymbol{\tau}^{(n)}$ as in Lemma 3.1. Then,

(i) *(asymptotic representation)* under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$, with $g_1 \in \mathscr{F}_1$, $\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K} = \boldsymbol{\Delta}^{\star g_1(n)}_{\boldsymbol{\xi},K} + o_{L^2}(1)$ as $n\to\infty$.

(ii) *(asymptotic normality)* under $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$, with $g_1 \in \mathscr{F}_a$, $\boldsymbol{\Delta}^{\star g_1(n)}_{\boldsymbol{\xi},K}$ is asymptotically normal with mean

$$\underset{\sim}{\boldsymbol{\mu}}{}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},K} := \underset{\sim}{\boldsymbol{\Gamma}}{}^{\star g_1}_{\boldsymbol{\xi},K}(\overset{\circ}{\mathrm{vech}}\,\mathbf{w}) := \mathscr{J}_k(K,g_1)\mathbf{G}_k(\mathbf{W})(\overset{\circ}{\mathrm{vech}}\,\mathbf{w})$$

and covariance matrix

$$\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K} := \mathscr{J}_k(K)\mathbf{G}_k(\mathbf{W}), \tag{13}$$

where $\mathscr{J}_k(K,g_1) := \int_0^1 K(u)K_{g_1}(u)du$ and $\mathscr{J}_k(K) := \int_0^1 K^2(u)du$;

(iii) *(asymptotic linearity)* under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$, with $g_1 \in \mathscr{F}_a$, $\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},K} - \underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\xi},K} = -\underset{\sim}{\boldsymbol{\Gamma}}^{\star g_1}_{\boldsymbol{\xi},K}(\overset{\circ}{\mathrm{vech}}\,\mathbf{w}^{(n)}) + o_{\mathrm{P}}(1)$ as $n \to \infty$.

Mimicking the form of the optimal parametric test in (10), we consider the rank test $\underset{\sim}{\phi}_K$ that rejects the null in (9) at asymptotic level α whenever

$$\underset{\approx}{Q}_K := \underset{\approx}{Q}_{\hat{\boldsymbol{\xi}}^{(n)},K} > \chi^2_{r;1-\alpha},$$

where

$$\begin{aligned}\underset{\approx}{Q}_{\boldsymbol{\xi},K} &:= (\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\xi},K})' \left[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right] \underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\xi},K} \\ &= \frac{1}{\mathscr{J}_k(K)}(\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\xi},K})' \left[(\mathbf{G}_k(\mathbf{W}))^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right] \underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\xi},K}\end{aligned} \tag{14}$$

and where $(\hat{\boldsymbol{\xi}}^{(n)})$ is an adequate sequence of estimators of $\boldsymbol{\xi}$. More precisely, we will assume the following.

Assumption (A). The sequence of estimators $(\hat{\boldsymbol{\xi}}^{(n)})$ is

(A1) *constrained:* for any $n \in \mathbb{N}_0$, $\boldsymbol{\xi} \in \Theta_0(\boldsymbol{\Upsilon})$ and $g_1 \in \mathscr{F}_a$, $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\left[\hat{\boldsymbol{\xi}}^{(n)} \in \Theta_0(\boldsymbol{\Upsilon})\right] = 1$;

(A2) $\sqrt{n}$*-consistent under the null:* for any $\boldsymbol{\xi} \in \Theta_0(\boldsymbol{\Upsilon})$ and $g_1 \in \mathscr{F}_a$, $\sqrt{n}(\hat{\boldsymbol{\xi}}^{(n)} - \boldsymbol{\xi}) = O_{\mathrm{P}}(1)$ as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$;

(A3) *locally asymptotically discrete:* for all $\boldsymbol{\xi} \in \Theta_0(\boldsymbol{\Upsilon})$ and all $c > 0$, there exists $M = M(c) > 0$ such that the number of possible values of $\hat{\boldsymbol{\xi}}^{(n)}$ in balls of the form $\{\mathbf{t} : n^{1/2}\|\mathbf{t} - \boldsymbol{\xi}\| \leq c\}$ is bounded by M, uniformly as $n \to \infty$.

Among the many possible estimators $\hat{\boldsymbol{\xi}}$, we propose the estimator

$$\begin{aligned}\hat{\boldsymbol{\xi}} &:= (\hat{\boldsymbol{\theta}}', \varsigma^2, (\overset{\circ}{\mathrm{vech}}\,\hat{\mathbf{W}}_0)')' \\ &:= (\hat{\boldsymbol{\theta}}', \varsigma^2, \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'(\overset{\circ}{\mathrm{vech}}\,\hat{\mathbf{W}}))'\end{aligned} \tag{15}$$

(note that ς^2 does not appear in $\underset{\approx}{Q}_{\boldsymbol{\xi},K}$, hence does not need to be estimated), where $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{W}}$, as in [8], are defined through

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{U}_i(\boldsymbol{\theta},\mathbf{W}) = \mathbf{0}, \quad \frac{1}{n}\sum_{i=1}^{n}\mathbf{U}_i(\boldsymbol{\theta},\mathbf{W})\mathbf{U}_i'(\boldsymbol{\theta},\mathbf{W}) = \frac{1}{k}\mathbf{I}_k, \quad \text{and} \quad \mathrm{tr}(\mathbf{W}) = k.$$

The estimator $\hat{\boldsymbol{\xi}}$ in (15)—or more precisely, the resulting sequence of estimators $(\hat{\boldsymbol{\xi}}^{(n)})$—satisfies Assumptions (A1)-(A2). After appropriate discretization, it would also satisfy Assumption (A3). In practical situations, however, where n is fixed, this discretization is superfluous, as one may always assume that $\hat{\boldsymbol{\xi}}$ is part of a locally and asymptotically discrete sequence of estimators; see, e.g., [5, 7] or [9] for a discussion.

The following theorem, which states the asymptotic properties of the rank test $\underset{\sim}{\phi}_K$, is the main result of this paper.

Theorem 4.1. *Let Assumption (A) hold. Then,*

(i) *under* $\cup_{\boldsymbol{\xi}\in\Theta_0(\boldsymbol{\Upsilon})}\cup_{g_1\in\mathscr{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$, $\underset{\sim}{Q}_K$ *is asymptotically chi-square with r degrees of freedom;*

(ii) *under* $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$, *with* $\boldsymbol{\tau}^{(n)}$ *as in Lemma 3.1,* $\boldsymbol{\xi}\in\Theta_0(\boldsymbol{\Upsilon})$, *and* $g_1\in\mathscr{F}_a$, $\underset{\sim}{Q}_K$ *is asymptotically non-central chi-square with r degrees of freedom and non-centrality parameter*

$$\frac{\mathscr{J}_k^2(K,g_1)}{\mathscr{J}_k(K)}(\overset{\circ}{\mathrm{vech}}\,\mathbf{w})'\left[\mathbf{G}_k(\mathbf{W})-\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\right](\overset{\circ}{\mathrm{vech}}\,\mathbf{w});$$

(iii) $\underset{\sim}{\phi}_K$ *has asymptotic level* α *under* $\cup_{\boldsymbol{\xi}\in\Theta_0(\boldsymbol{\Upsilon})}\cup_{g_1\in\mathscr{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$.

(iv) $\underset{\sim}{\phi}_{K_{f_1}}$ *is locally and asymptotically most stringent, still at asymptotic level* α*, for* $\cup_{\boldsymbol{\xi}\in\Theta_0(\boldsymbol{\Upsilon})}\cup_{g_1\in\mathscr{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}\}$ *against alternatives of the form* $\cup_{\boldsymbol{\xi}\notin\Theta_0(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\xi},f_1}\}$.

This result confirms that the proposed rank tests do not require any moment condition, as they are valid (and distribution-free) under any $g_1\in\mathscr{F}_a$. When based on the optimal score function K_{f_1}, they also achieve local and asymptotic optimality at f_1. In particular, the van der Waerden (normal-score) rank test $\underset{\sim}{\phi}_{\mathrm{vdW}}=\underset{\sim}{\phi}_{K_{\phi_1}}$ (see Section 5.2 for details) is locally and asymptotically optimal in the Gaussian case.

5. Asymptotic relatives efficiencies and simulations

In this section, we compare, through asymptotic relative efficiencies (AREs) and simulations, the pseudo-Gaussian tests from Section 3 and the rank-based tests from Section 4.

5.1. AREs

The results of both previous sections allow to compute in a straightforward way the asymptotic relative efficiencies of the proposed rank tests $\underset{\sim}{\phi}_K$ with respect to their pseudo-Gaussian competitors $\phi_{\mathscr{N}}$. These AREs indeed are simply obtained by dividing the non-centrality parameter in Theorem 4.1 by the one in Theorem 3.1.

Theorem 5.1. *The asymptotic relative efficiency of the rank tests* $\underset{\sim}{\phi}_K$ *with respect to the pseudo-Gaussian tests* $\phi_{\mathscr{N}}$*, under standardized radial density* $g_1(\in\mathscr{F}_a^{(4)})$*, is*

$$\mathrm{ARE}_{k,g_1}(\underset{\sim}{\phi}_K/\phi_{\mathscr{N}})=\frac{1+\kappa_k(g_1)}{k(k+2)}\,\frac{\mathscr{J}_k^2(K,g_1)}{\mathscr{J}_k(K)}. \tag{16}$$

Unlike their pseudo-Gaussian competitors, the proposed rank tests do not require finite fourth-order moments. If the underlying elliptical density has infinite fourth-order moments, the AREs of our rank tests with respect to pseudo-Gaussian tests may therefore be considered infinite. This implies that the assumption that $g_1 \in \mathscr{F}_a^{(4)}$ is not restrictive in Theorem 5.1.

The AREs in (16) do coincide with the ones obtained in [5], in the context of testing sphericity of the underlying elliptical distribution. As a direct corollary, the Chernoff-Savage result of [14] applies in the present context, and shows that the AREs of the proposed van der Waerden test $\underset{\sim}{\phi}_{\text{vdW}}$ with respect to the pseudo-Gaussian tests are uniformly larger than or equal to one, with equality under multinormal densities only.

Numerical values of these AREs are provided in Table 1, for various rank tests that are defined in Section 5.2 below. Further ARE values will be provided when presenting simulations results in Table 2.

5.2. Simulations

Simulations were conducted as follows. We generated $N = 1,500$ mutually independent samples of i.i.d. trivariate ($k = 3$) random vectors $\boldsymbol{\varepsilon}_{\ell;i}$, $\ell = 1, 2, 3, 4$, $i = 1,\ldots,n = 200$, with spherical Gaussian ($\boldsymbol{\varepsilon}_{1;i}$), t_8 ($\boldsymbol{\varepsilon}_{2;i}$), t_5 ($\boldsymbol{\varepsilon}_{3;i}$), and t_1 ($\boldsymbol{\varepsilon}_{4;i}$) densities, respectively. Each $\boldsymbol{\varepsilon}_{\ell;i}$ was successively transformed into

$$\mathbf{X}_{\ell;i;\eta} = (\varsigma^2(\mathbf{W}_0 + \mathbf{w}_\eta))^{-1/2}\boldsymbol{\varepsilon}_{\ell;i}, \quad \ell = 1,2,3,4, \quad i = 1,\ldots,n, \quad \eta = 0,1,2,3, \tag{17}$$

with

$$\varsigma^2 := 2, \ \ \mathbf{W}_0 := \begin{pmatrix} 1/2 & 1/4 & 0 \\ 1/4 & 1/2 & 1/2 \\ 0 & 1/2 & 2 \end{pmatrix}, \text{ and } \mathbf{w}_\eta := \begin{pmatrix} 0 & 0 & \eta/10 \\ 0 & 0 & 0 \\ \eta/10 & 0 & 0 \end{pmatrix}.$$

The value $\eta = 0$ corresponds to the null hypothesis $\mathscr{H}_0 : \mathbf{W}_{13} = 0$, while the values $\eta = 1,2,3$ provide increasingly severe alternatives.

We performed the pseudo-Gaussian test $\phi_{\mathscr{N}}$ and various rank-based tests $\underset{\sim}{\phi}_K$, all at asymptotic level $\alpha = 5\%$. The rank-based tests considered are

– the van der Waerden (normal-score) test $\underset{\sim}{\phi}_{\text{vdW}} (= \underset{\sim}{\phi}_{K_{\phi_1}})$, that uses

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

where Ψ_k stands for the χ_k^2 distribution function;

– the t_ν-score tests $\underset{\sim}{\phi}_{t_\nu}$, for $\nu = 1,5,8$, which are based on

$$K_{f_{1,\nu}^t}(u) = \frac{k(k+\nu)G_{k,\nu}^{-1}(u)}{\nu + kG_{k,\nu}^{-1}(u)} \quad \text{and} \quad \mathscr{J}_k(f_{1,\nu}^t) = \frac{k(k+2)(k+\nu)}{k+\nu+2},$$

where $G_{k,\nu}$ denotes the Fisher-Snedecor distribution function with k and ν degrees of freedom;

– the sign test $\underset{\sim}{\phi}{}^{(n)}_{\mathrm{S}}$, Wilcoxon test $\underset{\sim}{\phi}{}^{(n)}_{\mathrm{W}}$, and Spearman test $\underset{\sim}{\phi}{}^{(n)}_{\mathrm{SP}}$, that are obtained for

$$K_a(u) := u^a \quad \text{and} \quad \mathscr{J}_k(K_a) = 1/(2a+1),$$

with $a = 0$, $a = 1$ and $a = 2$, respectively;

Table 2, which reports the corresponding rejection frequencies, confirms our asymptotic results. The pseudo-Gaussian test $\phi_{\mathcal{N}}$ meets the level constraint under Gaussian, t_8 and t_5 distributions, but not under the t_1 (that has infinite fourth-order moments). In contrast, the rank tests always meet the nominal asymptotic level constraint. Despite the relatively small sample size ($n = 200$), the rankings based on empirical powers and AREs agree in most cases. In particular, under all Student distributions considered, the optimality of the rank tests based on correctly specified densities is confirmed (van der Waerden tests notoriously require larger sample sizes to show agreement with asymptotic results).

6. Appendix

PROOF OF LEMMA 3.2. In this proof, all stochastic convergences are as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$, for the values $\boldsymbol{\xi} \in \Theta$ and $g_1 \in \mathscr{F}_1^{(4)}$ fixed in the statement of the Lemma.

First note that, since $(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}}^{-1})'(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}}) = \mathrm{tr}(\mathbf{I}_k) = k$, $\mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathrm{cov}}} = \mathbf{T}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathrm{cov}}}$ rewrites

$$\mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathrm{cov}}} = -\frac{\sqrt{n}}{2}\mathbf{M}_k\left[\mathbf{I}_{k^2} - \frac{1}{k}(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}})(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}}^{-1})'\right]\mathrm{vec}\big(\mathbf{S}^{(n)}_{\boldsymbol{\theta}} - \boldsymbol{\Sigma}_{\mathrm{cov}}\big). \tag{18}$$

Decomposing $\mathbf{S}^{(n)}_{\hat{\boldsymbol{\theta}}} - \hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}$ into $t_1^{(n)} + t_2^{(n)} - t_3^{(n)}$, where $t_1^{(n)} := \mathbf{S}^{(n)}_{\hat{\boldsymbol{\theta}}} - \mathbf{S}^{(n)}_{\boldsymbol{\theta}} = o_{\mathrm{P}}(n^{-1/2})$, $t_2^{(n)} := \mathbf{S}^{(n)}_{\boldsymbol{\theta}} - \boldsymbol{\Sigma}_{\mathrm{cov}} = O_{\mathrm{P}}(n^{-1/2})$, and $t_3^{(n)} := \hat{\boldsymbol{\Sigma}}_{\mathrm{cov}} - \boldsymbol{\Sigma}_{\mathrm{cov}} = O_{\mathrm{P}}(n^{-1/2})$, (18) and the continuous mapping theorem yield

$$\begin{aligned}
\mathbf{T}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}} &= -\frac{\sqrt{n}}{2}\mathbf{M}_k\left[\mathbf{I}_{k^2} - \frac{1}{k}(\mathrm{vec}\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}})(\mathrm{vec}\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}^{-1})'\right]\mathrm{vec}\big(\mathbf{S}^{(n)}_{\hat{\boldsymbol{\theta}}} - \hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}\big)\\
&= \mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathrm{cov}}} + \frac{\sqrt{n}}{2}\mathbf{M}_k\left[\mathbf{I}_{k^2} - \frac{1}{k}(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}})(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}}^{-1})'\right]\mathrm{vec}\big(\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}} - \boldsymbol{\Sigma}_{\mathrm{cov}}\big) + o_{\mathrm{P}}(1).
\end{aligned}$$

By using first the delta method, then the identities $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})(\mathrm{vec}\mathbf{B})$ and $\mathbf{K}_k(\mathrm{vec}\mathbf{A}) = \mathrm{vec}(\mathbf{A}')$, this entails that

$$\begin{aligned}
\mathbf{T}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}} - \mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathrm{cov}}} &= -\frac{\sqrt{n}}{2}\mathbf{M}_k\left[\mathbf{I}_{k^2} - \frac{1}{k}(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}})(\mathrm{vec}\boldsymbol{\Sigma}_{\mathrm{cov}}^{-1})'\right](\boldsymbol{\Sigma}_{\mathrm{cov}}^{\otimes 2})\mathrm{vec}\big(\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}^{-1} - \boldsymbol{\Sigma}_{\mathrm{cov}}^{-1}\big) + o_{\mathrm{P}}(1)\\
&= -\frac{\sqrt{n}}{2}\mathbf{M}_k(\boldsymbol{\Sigma}_{\mathrm{cov}}^{\otimes 2})^{1/2}\left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right](\boldsymbol{\Sigma}_{\mathrm{cov}}^{\otimes 2})^{1/2}\mathrm{vec}\big(\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}^{-1} - \boldsymbol{\Sigma}_{\mathrm{cov}}^{-1}\big) + o_{\mathrm{P}}(1)\\
&= -\frac{\sqrt{n}}{4}\mathbf{M}_k(\boldsymbol{\Sigma}_{\mathrm{cov}}^{\otimes 2})^{1/2}\mathbf{H}_k(\boldsymbol{\Sigma}_{\mathrm{cov}}^{\otimes 2})^{1/2}\mathrm{vec}\big(\hat{\boldsymbol{\Sigma}}_{\mathrm{cov}}^{-1} - \boldsymbol{\Sigma}_{\mathrm{cov}}^{-1}\big) + o_{\mathrm{P}}(1),
\end{aligned}$$

where we let $\mathbf{H}_k := \mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k$. Clearly, $\hat{\mathbf{W}}_{\text{cov}}$ is a root-n consistent estimator for $\mathbf{W} = k\boldsymbol{\Sigma}_{\text{cov}}^{-1}/\text{tr}\,(\boldsymbol{\Sigma}_{\text{cov}}^{-1})$. Recalling that $\hat{\boldsymbol{\Sigma}}_{\text{cov}}^{-1} - \boldsymbol{\Sigma}_{\text{cov}}^{-1}$ is $O_{\text{P}}(n^{-1/2})$, the continuous mapping theorem then yields

$$\begin{aligned}
\mathbf{T}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}_{\text{cov}}} - \mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\text{cov}}} &= -\frac{\sqrt{n}}{4k}\mathbf{M}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}\mathbf{H}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2} \\
&\qquad \times \text{vec}\Big((\text{tr}\,(\hat{\boldsymbol{\Sigma}}_{\text{cov}}^{-1} - \boldsymbol{\Sigma}_{\text{cov}}^{-1}))\hat{\mathbf{W}}_{\text{cov}} + (\text{tr}\,\boldsymbol{\Sigma}_{\text{cov}}^{-1})(\hat{\mathbf{W}}_{\text{cov}} - \mathbf{W})\Big) + o_{\text{P}}(1) \\
&= -\frac{\sqrt{n}}{4k}(\text{tr}\,(\hat{\boldsymbol{\Sigma}}_{\text{cov}}^{-1} - \boldsymbol{\Sigma}_{\text{cov}}^{-1}))\,\mathbf{M}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}\mathbf{H}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}(\text{vec}\mathbf{W}) \\
&\qquad -\frac{\sqrt{n}}{4k}(\text{tr}\,\boldsymbol{\Sigma}_{\text{cov}}^{-1})\,\mathbf{M}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}\mathbf{H}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}\text{vec}(\hat{\mathbf{W}}_{\text{cov}} - \mathbf{W}) + o_{\text{P}}(1).
\end{aligned}$$

Note that $\mathbf{H}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}(\text{vec}\mathbf{W}) = \mathbf{H}_k\text{vec}(\boldsymbol{\Sigma}_{\text{cov}}^{1/2}\mathbf{W}\boldsymbol{\Sigma}_{\text{cov}}^{1/2}) = k\mathbf{H}_k(\text{vec}\,\mathbf{I}_k)/\text{tr}\,(\boldsymbol{\Sigma}_{\text{cov}}^{-1}) = \mathbf{0}$. This and the definition of $\mathbf{M}_k$ (note that $\hat{\mathbf{W}}_{\text{cov}} - \mathbf{W}$ has trace zero) finally provides

$$\begin{aligned}
\mathbf{T}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}_{\text{cov}}} - \mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\text{cov}}} &= -\frac{\sqrt{n}}{4k}(\text{tr}\,\boldsymbol{\Sigma}_{\text{cov}}^{-1})\,\mathbf{M}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}\mathbf{H}_k(\boldsymbol{\Sigma}_{\text{cov}}^{\otimes 2})^{1/2}\mathbf{M}_k'\,\overset{\circ}{\text{vech}}\,(\hat{\mathbf{W}}_{\text{cov}} - \mathbf{W}) + o_{\text{P}}(1) \\
&= -(k+2)\,(\text{tr}\,\boldsymbol{\Sigma}_{\text{cov}}^{-1})\,\mathbf{G}_k(\boldsymbol{\Sigma}_{\text{cov}}^{-1})\sqrt{n}\;\overset{\circ}{\text{vech}}\,(\hat{\mathbf{W}}_{\text{cov}} - \mathbf{W}) + o_{\text{P}}(1),
\end{aligned}$$

which establishes the result. □

Proof of Theorem 3.1. (i) Fix $\boldsymbol{\xi} = (\boldsymbol{\theta}', \varsigma^2, (\overset{\circ}{\text{vech}}\,\mathbf{W})')' \in \boldsymbol{\Theta}_0(\boldsymbol{\Upsilon})$ and $g_1 \in \mathscr{F}_1^{(4)}$. By construction, $(\overset{\circ}{\text{vech}}\,\hat{\mathbf{W}}_{\text{cov};0})$ takes its values in $\mathscr{M}(\boldsymbol{\Upsilon})$, so that Lemma 3.2 implies that

$$\begin{aligned}
\mathbf{T}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}_{\text{cov}}} - \mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\text{cov}}} &= -(k+2)\,(\text{tr}\,\boldsymbol{\Sigma}_{\text{cov}}^{-1})\,\mathbf{G}_k(\boldsymbol{\Sigma}_{\text{cov}}^{-1})\sqrt{n}\;\overset{\circ}{\text{vech}}\,(\hat{\mathbf{W}}_{\text{cov};0} - \mathbf{W}) + o_{\text{P}}(1) \\
&= -(k+2)\,(\text{tr}\,\boldsymbol{\Sigma}_{\text{cov}}^{-1})\,\mathbf{G}_k(\boldsymbol{\Sigma}_{\text{cov}}^{-1})\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\sqrt{n}\;\overset{\circ}{\text{vech}}\,(\hat{\mathbf{W}}_{\text{cov};0} - \mathbf{W}) + o_{\text{P}}(1),
\end{aligned}$$

as $n \to \infty$ under $\text{P}_{\boldsymbol{\xi},g_1}^{(n)}$. This entails that

$$\left[(\mathbf{G}_k(\boldsymbol{\Sigma}_{\text{cov}}^{-1}))^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\boldsymbol{\Sigma}_{\text{cov}}^{-1})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right]\left(\mathbf{T}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}_{\text{cov}}} - \mathbf{T}_{\boldsymbol{\theta},\boldsymbol{\Sigma}_{\text{cov}}}\right) = o_{\text{P}}(1), \tag{19}$$

still as $n \to \infty$ under $\text{P}_{\boldsymbol{\xi},g_1}^{(n)}$. Jointly with the consistency of $\hat{\kappa}_k$ and the continuity of the mapping $\mathbf{A} \mapsto \mathbf{G}_k(\mathbf{A})$, this yields that $Q_{\mathscr{N}} - Q_{\boldsymbol{\xi};\mathscr{N}}^{g_1} = o_{\text{P}}(1)$ as $n \to \infty$ under $\text{P}_{\boldsymbol{\xi},g_1}^{(n)}$. Equality (12), the $\boldsymbol{\tau}^{(n)} = \mathbf{0}$ version of Lemma 3.1, and the idempotence of

$$\mathbf{B} := \boldsymbol{\Gamma}_{\boldsymbol{\xi},\phi_1}^{\star g_1}\left[(\boldsymbol{\Gamma}_{\boldsymbol{\xi},\phi_1}^{\star g_1})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}_{\boldsymbol{\xi},\phi_1}^{\star g_1}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right]$$

then yields (see., e.g., Theorem 9.2.1 in [15]) that $Q_{\mathscr{N}}$, under $\text{P}_{\boldsymbol{\xi},g_1}^{(n)}$, is asymptotically chi-square with

$$\text{tr}(\mathbf{B}) = \text{tr}\Big(\mathbf{I}_K - (\boldsymbol{\Gamma}_{\boldsymbol{\xi},\phi_1}^{\star g_1})^{1/2}\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}_{\boldsymbol{\xi},\phi_1}^{\star g_1}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'(\boldsymbol{\Gamma}_{\boldsymbol{\xi},\phi_1}^{\star g_1})^{1/2}\Big) = K - (K - r) = r \tag{20}$$

degrees of freedom.

(ii) Assume now that $g_1 \in \mathscr{F}_a^{(4)}$ and fix a sequence $\boldsymbol{\tau}^{(n)}$ as in Lemma 3.1. From contiguity, $Q_{\mathscr{N}} - Q^{g_1}_{\boldsymbol{\xi};\mathscr{N}}$ is also $o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$. Applying again Theorem 9.2.1 in [15] and using Lemma 3.1 now shows that $Q_{\mathscr{N}}$, still under $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$, is indeed asymptotically non-central chi-square with r degrees of freedom and non-centrality parameter

$$(\boldsymbol{\mu}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},\phi_1})' \left[(\boldsymbol{\Gamma}^{\star g_1}_{\boldsymbol{\xi},\phi_1})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star g_1}_{\boldsymbol{\xi},\phi_1}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}' \right] \boldsymbol{\mu}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},\phi_1}$$

$$= (k+2)^2 \frac{D_k^2(g_1)}{E_k(g_1)} (\overset{\circ}{\mathrm{vech}}\,\mathbf{w})' \left[\mathbf{G}_k(\mathbf{W}) - \mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W}) \right] (\overset{\circ}{\mathrm{vech}}\,\mathbf{w})$$

$$= \frac{k(k+2)}{1+\kappa_k(g_1)} (\overset{\circ}{\mathrm{vech}}\,\mathbf{w})' \left[\mathbf{G}_k(\mathbf{W}) - \mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W}) \right] (\overset{\circ}{\mathrm{vech}}\,\mathbf{w}).$$

(iii)-(iv) Part (iii) of the result directly follows from the asymptotic null distribution given in Part (i) and the classical Helly-Bray theorem. As for Part (iv), the asymptotic equivalence $Q_{\mathscr{N}} - Q^{g_1}_{\boldsymbol{\xi};\mathscr{N}} = o_{\mathrm{P}}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$ shows that, in the multinormal case $g_1 = \phi_1$, $Q_{\mathscr{N}}$ is asymptotically equivalent in probability to $Q^{\phi_1}_{\boldsymbol{\xi},\mathscr{N}} = Q_{\boldsymbol{\xi},\phi_1}$; the local asymptotic optimality result in Part (iv) then is a consequence of the weak convergence of local experiments to Gaussian shifts (see, e.g., Section 11.9 of [11]). □

Proof of Theorem 4.1. (i) Fix $\boldsymbol{\xi} = (\boldsymbol{\theta}', \varsigma^2, (\overset{\circ}{\mathrm{vech}}\,\mathbf{W})')' \in \boldsymbol{\Theta}_0$ and $g_1 \in \mathscr{F}_a$. Lemma 4.1 in [10] allows to replace the deterministic perturbation $\boldsymbol{\tau}^{(n)}$ in Lemma 4.1(iii) with $\sqrt{n}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})$, which yields

$$\begin{aligned} \underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\hat{\boldsymbol{\xi}},K} - \underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K} &= -\underset{\sim}{\boldsymbol{\Gamma}}{}^{\star g_1}_{\boldsymbol{\xi},K} n^{1/2} \overset{\circ}{\mathrm{vech}}\,(\hat{\mathbf{W}}^{(n)}_0 - \mathbf{W}) + o_{\mathrm{P}}(1) \\ &= -\frac{\mathscr{J}_k(K,g_1)}{\mathscr{J}_k(K)} \boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K} \sqrt{n}\ \overset{\circ}{\mathrm{vech}}\,(\hat{\mathbf{W}}^{(n)}_0 - \mathbf{W}) + o_{\mathrm{P}}(1) \end{aligned} \tag{21}$$

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$. Applying the continuous mapping theorem, using (21) in conjunction with Assumption (A1), and then applying Lemma 4.1(i), we obtain that

$$\begin{aligned} \underset{\sim}{Q}_K &= (\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\hat{\boldsymbol{\xi}},K})' \left[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}' \right] \underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\hat{\boldsymbol{\xi}},K} + o_{\mathrm{P}}(1) \\ &= (\underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K})' \left[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}' \right] \underset{\sim}{\boldsymbol{\Delta}}{}^{\star(n)}_{\boldsymbol{\xi},K} + o_{\mathrm{P}}(1) \\ &= (\boldsymbol{\Delta}^{\star g_1(n)}_{\boldsymbol{\xi},K})' \left[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}' \right] \boldsymbol{\Delta}^{\star g_1(n)}_{\boldsymbol{\xi},K} + o_{\mathrm{P}}(1) \end{aligned} \tag{22}$$

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\xi},g_1}$. The result then follows again from Theorem 9.2.1 in [15], by using (the $\boldsymbol{\tau}^{(n)} = \mathbf{0}$ version of) Lemma 4.1(ii) and the fact that $\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K}[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}']$ is idempotent with trace r (the trace can be computed as in (20)).

(ii) From contiguity, (22) also holds under $\mathrm{P}^{(n)}_{\boldsymbol{\xi}+n^{-1/2}\boldsymbol{\tau}^{(n)},g_1}$, so that the same result from [15], via Lemma 4.1(ii), now yields that $\underset{\sim}{Q}_K$, under these local alternatives, indeed is asymptotically non-central chi-square with r degrees of freedom and non-centrality parameter

$$\big(\underset{\sim}{\boldsymbol{\mu}}{}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},K}\big)' \Big[(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K})^{-1} - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\xi},K}\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\Big] \underset{\sim}{\boldsymbol{\mu}}{}^{\star g_1}_{\boldsymbol{\xi},\boldsymbol{\tau},K}$$

$$= \frac{\mathscr{J}_k^2(K,g_1)}{\mathscr{J}_k(K)} (\overset{\circ}{\mathrm{vech}}\,\mathbf{w})' \left[\mathbf{G}_k(\mathbf{W}) - \mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\mathbf{G}_k(\mathbf{W})\right] (\overset{\circ}{\mathrm{vech}}\,\mathbf{w}).$$

(iii)-(iv) Part (iii) directly follows from the asymptotic null distribution given in Part (i) and the classical Helly-Bray theorem. As for Part (iv), note that the $K = K_{f_1}$ version of (22) shows that $\underset{\sim}{Q}_{f_1}$ and $Q_{\boldsymbol{\xi},f_1}$ in (10) are asymptotically equivalent in probability under standardized radial density f_1; as in the proof of Theorem 3.1, the local asymptotic optimality of $\underset{\sim}{\phi}_{K_{f_1}}$ is then a consequence of the weak convergence of local experiments to Gaussian shifts. □

References

[1] T W Anderson. *An Introduction to Multivariate Statistical Analysis (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 3 edition, 2003.

[2] PM Bentler. Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. *Psychometrika*, 48:493–517, 1983.

[3] PM Bentler and M Berkane. Greatest lower bound to the elliptical theory kurtosis parameter. *Biometrika*, 73:240, 1986.

[4] M Hallin, H Oja, and D Paindaveine. Semiparametrically efficient rank-based inference for shape II. Optimal R-estimation of shape. *The Annals of Statistics*, 34:2757–2789, 2006.

[5] M Hallin and D Paindaveine. Semiparametrically efficient rank-based inference for shape I. optimal rank-based tests for sphericity. *The Annals of Statistics*, 34:2707–2756, 2006.

[6] M Hallin and D Paindaveine. Optimal rank-based tests for homogeneity of scatter. *Annals of Statistics*, 36:1261–1298, 2008.

[7] M Hallin, D Paindaveine, and T Verdebout. Optimal rank-based testing for principal components. *The Annals of Statistics*, 38:3245–3299, 2010.

[8] TP Hettmansperger and RH Randles. A practical affine equivariant multivariate median. *Biometrika*, 89:851, 2002.

[9] P Ilmonen and D Paindaveine. Semiparametrically efficient inference based on signed ranks in symmetric independent component models. *Annals of Statistics*, 39:2448–2476, 2011.

[10] JP Kreiss. On adaptive estimation in stationary ARMA processes. *The Annals of Statistics*, 15:112–133, 1987.

[11] L Le Cam. *Asymptotic Methods in Statistical Decision Theory (Springer Series in Statistics)*. Springer, New York, 1986.

[12] H Li and J Gui. Gradient directed regularization for sparse Gaussian concentration graphs, with applications to inference of genetic networks. *Biostatistics*, 7:302–317, 2005.

[13] PM Magwene and J Kim. Estimating genomic coexpression networks using first-order conditional independence. *Genome Biology*, 5:R100, 2004.

[14] D Paindaveine. A Chernoff-Savage result for shape: on the non-admissibility of pseudo-Gaussian methods. *Journal of Multivariate Analysis*, 97:2206–2220, 2006.

[15] C R Rao and Sujit K Mitra. *Generalized Inverse of Matrices and Its Applications*. John Wiley & Sons Inc, New York, 1971.

[16] H Toh and K Horimoto. Inference of a genetic network by a combined approach of cluster analysis and graphical Gaussian modeling. *Bioinformatics*, 18:287–297, 2002.

[17] D Vogel and R Fried. Elliptical graphical modelling. *Biometrika*, 98:935–951, 2011.

		degrees of freedom of the underlying t density					
test	k	$\nu \leq 4$	$\nu = 5$	$\nu = 8$	$\nu = 15$	$\nu = 20$	$\nu \to \infty$
$\underset{\sim}{\phi}_{\mathrm{vdW}}$	2	$+\infty$	2.204	1.215	1.047	1.025	1.000
	3	$+\infty$	2.270	1.233	1.052	1.028	1.000
	4	$+\infty$	2.326	1.249	1.057	1.031	1.000
	6	$+\infty$	2.413	1.275	1.066	1.036	1.000
	10	$+\infty$	2.531	1.312	1.080	1.045	1.000
$\underset{\sim}{\phi}_{t_6}$	2	$+\infty$	2.331	1.248	1.045	1.013	0.957
	3	$+\infty$	2.398	1.267	1.052	1.018	0.957
	4	$+\infty$	2.453	1.284	1.058	1.023	0.958
	6	$+\infty$	2.537	1.311	1.070	1.031	0.959
	10	$+\infty$	2.646	1.349	1.087	1.044	0.963
$\underset{\sim}{\phi}_{\mathrm{S}}$	2	$+\infty$	1.500	0.750	0.591	0.563	0.500
	3	$+\infty$	1.800	0.900	0.709	0.675	0.600
	4	$+\infty$	2.000	(1.000	0.788	0.750	0.667
	6	$+\infty$	2.250	(1.125	0.886	0.844	0.750
	10	$+\infty$	2.500	1.250	0.985	0.938	0.833
$\underset{\sim}{\phi}_{\mathrm{W}}$	2	$+\infty$	2.258	1.174	0.956	0.919	0.844
	3	$+\infty$	2.386	1.246	1.022	0.985	0.913
	4	$+\infty$	2.432	1.273	1.048	1.012	0.945
	6	$+\infty$	2.451	1.283	1.060	1.026	0.969
	10	$+\infty$	2.426	1.264	1.045	1.013	0.970

TABLE 1. *AREs, with respect to the pseudo-Gaussian tests* $\phi_{\mathcal{N}}$*, of the van der Waerden (*$\underset{\sim}{\phi}_{\mathrm{vdW}}$*),* t_ν *(with* $\nu = 6$*), sign (*$\underset{\sim}{\phi}_{\mathrm{S}}$*), and Wilcoxon (*$\underset{\sim}{\phi}_{\mathrm{W}}$*) rank tests* $\underset{\sim}{\phi}_K$*, under k-dimensional Student (*$\nu \leq 4$*,* $\nu = 5$*, 8, 15, and 20 degrees of freedom) and normal densities, respectively, for* $k = 2, 3, 4, 6$*, and 10.*

test	g_1	$\eta=0$	$\eta=1$	$\eta=2$	$\eta=3$	ARE$(\cdot/\phi_{\mathcal{N}})$
$\phi_{\mathcal{N}}$		.0513	.2500	.7893	.9940	1.0000
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0500	.2487	.7713	.9920	1.0000
$\underset{\sim}{\phi}_{t_8}$		.0633	.2533	.7753	.9907	.9714
$\underset{\sim}{\phi}_{t_5}$	$\mathcal{N}$	.0647	.2487	.7660	.9867	.9460
$\underset{\sim}{\phi}_{t_1}$		.0600	.2253	.6867	.9667	.7824
$\underset{\sim}{\phi}_{\mathrm{S}}$		.0527	.1893	.5747	.9200	.6000
$\underset{\sim}{\phi}_{\mathrm{W}}$		.0653	.2413	.7533	.9833	.9130
$\underset{\sim}{\phi}_{\mathrm{SP}}$		.0567	.2500	.7640	.9887	.9568
$\phi_{\mathcal{N}}$		.0433	.1933	.6220	.9507	1.0000
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0453	.2207	.6880	.9713	1.2329
$\underset{\sim}{\phi}_{t_8}$		.0487	.2247	.7140	.9767	1.2692
$\underset{\sim}{\phi}_{t_5}$	t_8	.0487	.2227	.7180	.9747	1.2637
$\underset{\sim}{\phi}_{t_1}$		.0567	.1967	.6887	.9560	1.1319
$\underset{\sim}{\phi}_{\mathrm{S}}$		.0567	.1773	.5767	.9100	.9000
$\underset{\sim}{\phi}_{\mathrm{W}}$		.0500	.2180	.7147	.9740	1.2464
$\underset{\sim}{\phi}_{\mathrm{SP}}$		.0487	.2227	.6967	.9727	1.2249
$\phi_{\mathcal{N}}$		.0500	.1453	.4887	.8373	1.0000
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0507	.2200	.6747	.9660	2.2705
$\underset{\sim}{\phi}_{t_8}$		.0447	.2493	.7020	.9760	2.3895
$\underset{\sim}{\phi}_{t_5}$		.0453	.2540	.7120	.9753	2.4000
$\underset{\sim}{\phi}_{t_1}$	t_5	.0473	.2373	.6793	.9613	2.2244
$\underset{\sim}{\phi}_{\mathrm{S}}$		.0540	.1860	.5827	.9140	1.8000
$\underset{\sim}{\phi}_{\mathrm{W}}$		.0473	.2500	.7100	.9767	2.3858
$\underset{\sim}{\phi}_{\mathrm{SP}}$		.0507	.2300	.6780	.9727	2.2766
$\phi_{\mathcal{N}}$		.0233	.0307	.0293	.0627	–
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0467	.1340	.5247	.8787	$+\infty$
$\underset{\sim}{\phi}_{t_8}$		.0520	.1580	.5787	.9180	$+\infty$
$\underset{\sim}{\phi}_{t_5}$	t_1	.0527	.1640	.5913	.9227	$+\infty$
$\underset{\sim}{\phi}_{t_1}$		.0513	.1727	.6420	.9407	$+\infty$
$\underset{\sim}{\phi}_{\mathrm{S}}$		.0513	.1667	.5893	.9153	$+\infty$
$\underset{\sim}{\phi}_{\mathrm{W}}$		.0507	.1667	.5993	.9260	$+\infty$
$\underset{\sim}{\phi}_{\mathrm{SP}}$		.0553	.1413	.5207	.8740	$+\infty$

TABLE 2. *Rejection frequencies (out of $N = 1,500$ replications), under the null ($\eta = 0$) and increasingly severe alternatives ($\eta = 1,2,3$; see Subsection 5.2 for details), of the pseudo-Gaussian test $\phi_{\mathcal{N}}$, and the van der Waerden ($\underset{\sim}{\phi}_{\mathrm{vdW}}$), t_ν ($\nu = 8, 5, 1$), sign ($\underset{\sim}{\phi}_{\mathrm{S}}$), Wilcoxon ($\underset{\sim}{\phi}_{\mathrm{W}}$), and Spearman ($\underset{\sim}{\phi}_{\mathrm{SP}}$) rank tests. Sample size is $n = 200$. All tests were based on asymptotic 5% critical values. The last column provides numerical computations of the AREs in (5.1).*