T E C H N I C A L
R E P O R T

0629

OPTIMAL RANK-BASED TESTS FOR HOMOGENEITY OF SCATTER

HALLIN M., and D. PAINDAVEINE

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Optimal Rank-Based Tests for Homogeneity of Scatter

Marc Hallin*
and
Davy Paindaveine*

E.C.A.R.E.S., Institute for Research in Statistics, and Département de Mathématique
Université Libre de Bruxelles, Brussels, Belgium

Abstract

We propose a class of locally and asymptotically optimal tests, based on multivariate ranks and signs, for the homogeneity of scatter matrices in m elliptical populations. Contrary to the existing parametric procedures, these tests remain valid without any moment assumptions, and thus are perfectly robust against heavy-tailed distributions (*validity robustness*). Nevertheless, they reach semiparametric efficiency bounds at correctly specified densities (*efficiency robustness*). In most cases, their normal-score version outperforms Schott's pseudo-Gaussian test (*JSPI* **94**, 25-36, 2001), which, as we also show, actually is a robustified version of the traditional Gaussian likelihood ratio test.

AMS 1980 subject classification : 62M15, 62G35.

Key words and phrases : Elliptical densities, Scatter matrices, Multivariate ranks and signs, Multivariate analysis of variance, Local asymptotic normality, Locally asymptotically most stringent tests.

1 Introduction.

1.1 Homogeneity of variances and covariance matrices.

The assumption of variance homogeneity is central to the theory and practice of univariate m-sample inference, playing a major role in such models as m-sample location (ANOVA) or m-sample regression (ANOCOVA). The problem of testing the null hypothesis of variance homogeneity therefore is of fundamental importance, and for more than half a century has been a subject of continued interest in the statistical literature. The standard procedure, described in most textbooks, is Bartlett (1937)'s modified (Gaussian) likelihood ratio test (MLRT). This test however is well-known to be highly non-robust against violations of Gaussian assumptions, a fact that gave rise to a large number of "robustified" versions of the likelihood ratio procedure (Cochran 1941, Bartlett and Kendall 1946, Hartley 1950, Box 1953, to quote only a few). Soon,

*Research supported by a P.A.I. contract of the Belgian Federal Government and an Action de Recherche Concertée of the Communauté française de Belgique.

it was noticed that these "robustifications", if reasonably resistant to nonnormality, unfortunately were lacking power: in the convenient terminology of Heritier and Ronchetti (1994), they enjoy *validity robustness* but not *efficiency robustness*.

In an extensive simulation study, Conover et al. (1981) have investigated the validity-robustness (against nonnormal densities) and efficiency-robustness properties of 56 distinct tests, including several (signed) rank-based ones. Their conclusion is that only three out of 56 survive the examination, and that two of the three survivors are normal-score signed rank tests (adapted from Fligner and Killeen 1976).

In view of its applications in MANOVA, MANOCOVA, discriminant analysis, etc., the multivariate problem of testing for homogeneity of covariance matrices is certainly no less important than its univariate counterpart. The same problem moreover is of intrinsic interest in such fields as psychometrics or genetics where, for instance, the homogeneity of genetic covariance structure among species is a classical subject of investigation. Robustness and power issues however are even more delicate and complex in the multivariate context. So far, only validity robustness (resistance to violations of joint normality assumptions) has been investigated, leading to various robustifications of the Gaussian likelihood ratio test—see Section 1.2 for a brief account. But, although no multivariate equivalent of the Conover et al. (1981) study has been conducted so far, it is more than likely that those various robustifications also lose most of the power of the original tests, and thus suffer the same lack of efficiency robustness as in the univariate setting; our own simulations amply confirm this fact for the robustified version of Schott (2001)'s test—see below for details.

Since rank-based procedures are doing so well in the univariate case, it is tempting to extend them to the multivariate context. The problem however is that of choosing an adequate multivariate concept of (signed) ranks. The purpose of this paper is to develop such testing procedures for the null hypothesis of covariance homogeneity under elliptical distributions, based on the signs and ranks considered by Hallin and Paindaveine (2005 and 2006a) for one-sample inference for location, (auto)regression, and shape. Elliptical densities indeed constitute the most general class of densities under which classical multivariate techniques (after adequate modification) apply.

Contrary to all existing methods, our tests do not require any moment assumptions, so that the null hypothesis they address actually is the hypothesis of homogeneous *scatter matrices*, reducing to more classical homogeneity of covariance matrices under finite second-order moments. Being asymptotically distribution-free, our tests enjoy validity robustness (against nonnormal elliptic densities, including the heavy-tailed ones). They reach semiparametric efficiency at correctly specified densities, and therefore are efficiency-robust; when based on Gaussian scores, their asymptotic relative efficiency with respect to the various robustifications of the Gaussian likelihood ratio test is larger than one under almost all elliptical densities (see Section 6 for details).

1.2 Testing equality of scatter (covariance) matrices.

Denote by $(\mathbf{X}_{i1}, \ldots, \mathbf{X}_{in_i})$, $i = 1, \ldots, m$ a collection of m mutually independent samples of i.i.d. random k-dimensional vectors with location parameters $\boldsymbol{\theta}_i$ and scatter (under finite second-order moments, covariance) matrices $\mathbf{\Sigma}_i$, $i = 1, \ldots, m$. The purpose of this paper is to develop a signed rank-based solution to the problem of testing the null hypothesis $\mathcal{H}_0 : \mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_m$ of scatter (covariance) homogeneity against the alternative that the $\mathbf{\Sigma}_i$'s are not all equal.

The most classical test for this problem is the Gaussian likelihood ratio test (LRT). This test, which is based on the additional assumption that $\mathbf{X}_{ij} \sim \mathcal{N}_k(\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_i)$, rejects $\mathcal{H}_0$ for small values of

$$\Lambda := \frac{\prod_{i=1}^m |\mathbf{W}_i/n_i|^{n_i/2}}{|\mathbf{W}/n|^{n/2}} =: \frac{\prod_{i=1}^m |\mathbf{S}_i|^{n_i/2}}{|\mathbf{S}|^{n/2}}, \tag{1.1}$$

where $n = \sum_{i=1}^m n_i$ is the total sample size, $\bar{\mathbf{X}}_i := \frac{1}{n_i}\sum_{j=1}^{n_i} \mathbf{X}_{ij}$, $\mathbf{W}_i := \sum_{j=1}^{n_i}(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)'$, and $\mathbf{W} := \sum_{i=1}^m \mathbf{W}_i$. Even under Gaussian assumptions, this LRT is actually biased (see Brown 1939 and Das Gupta 1969), and one therefore usually relies on Bartlett's modified likelihood ratio test (MLRT), based on

$$\Lambda^* := \frac{\prod_{i=1}^m |\mathbf{W}_i/\dot{n}_i|^{\dot{n}_i/2}}{|\mathbf{W}/\dot{n}|^{\dot{n}/2}} =: \frac{\prod_{i=1}^m |\dot{\mathbf{S}}_i|^{n_i/2}}{|\dot{\mathbf{S}}|^{\dot{n}/2}}, \tag{1.2}$$

where $\dot{n}_i := n_i - 1$ and $\dot{n} := \sum_{i=1}^m \dot{n}_i = n - m$. The MLRT has been shown to be unbiased by Pitman (1939) for $k = 1$ and by Perlman (1980) for the general case. Much is known today about this test: monotonicity of the power function (Anderson and Das Gupta 1964, Das Gupta and Giri 1973), null and non-null expansions (both for fixed and local alternatives) of the distributions of Λ^* or $-2\log\Lambda^*$ (Sugiura 1973, Khatri and Srivastava 1974, Srivastava, Khatri, and Carter 1978), exact distribution of Λ^* (Gupta and Tang 1984), etc. All authors however insist on the extreme non-robustness to departures from normality of both the LRT and the MLRT, which are not even valid under non-Gaussian elliptical densities with finite fourth-order moments; see, in particular, Tyler (1983), Yanagihara et al. (2005), and Gupta and Xu (2006).

Such a poor resistance to non-normality is not uncommon in the context, and similar problems arise with most Gaussian likelihood ratio tests in multivariate analysis. In a classical reference, Muirhead and Waternaux (1980) provide an in-depth study of the problem of turning standard Gaussian tests about covariance matrices into pseudo-Gaussian ones remaining valid under elliptical densities (possibly with adequate moment assumptions). They clearly distinguish some "easy" cases—tests of sphericity, tests of equality of a subset of the characteristic roots of the covariance matrix (i.e., *subspace sphericity*), tests of block-diagonality—and some "harder" ones, among which the (apparently simpler) one-sample test of the hypothesis that the covariance matrix $\boldsymbol{\Sigma}$ takes some given value $\boldsymbol{\Sigma}_0$, the two-sample test of equality of covariance matrices, and the corresponding m-sample test (based on (1.1) or (1.2)). For these "hard" cases, they conclude that *"it is not possible in the more general elliptical case to adjust the (Gaussian likelihood ratio) test so that its limiting distribution agrees with that obtained under the normality assumption"*; see also Section 3 of Tyler (1983). In particular, for the problem under study, a recent result of Yanagihara et al. (2005) establishes that the asymptotic null distribution of $-2\log\Lambda^*/(1+\kappa_k)$ (where Λ^* is defined in (1.2) and κ_k stands for a measure of kurtosis of the underlying elliptical distribution; see Section 5.2 for a definition) is that of

$$\left[1 + \frac{k\kappa_k}{2(1+\kappa_k)}\right] Y_1 + Y_2, \tag{1.3}$$

where Y_1 and Y_2 are independent chi-square random variables with $m-1$ and $(m-1)(k-1)(k+2)/2$ degrees of freedom, respectively. In the multinormal case, $\kappa_k = 0$, and this yields the well-known Gaussian result that $-2\log\Lambda^*$ is asymptotically chi-square with $(m-1)k(k+1)/2$ degrees of freedom under the null hypothesis; but for $\kappa \neq 0$, (1.3) is no longer chi-square (see also Gupta and Xu 2006).

For the sake of comparison, in the problem of testing for sphericity, $-2\log\Lambda_{\text{spher}}/(1+\kappa_k)$ (where Λ_{spher} stands for the LRT statistic for sphericity) is asymptotically chi-square with $(m-$

$1)k(k+1)/2$ degrees of freedom under the null hypothesis, *irrespective of the underlying elliptic distribution* (with finite fourth-order moments and kurtosis κ_k). Consequently, robustifying the LRT for sphericity is easily achieved by adopting the modified test statistic $-2\log\Lambda_s/(1+\hat{\kappa}_k)$, where $\hat{\kappa}_k$ is a consistent estimate of κ_k. Clearly, in view of the null asymptotic distribution of $-2\log\Lambda^*/(1+\kappa_k)$ in (1.3), such an easy robustification is not possible when testing for covariance homogeneity.

Other Gaussian testing procedures also have been considered (see Section 5.3); to the best of our knowledge, they all suffer of the same lack of robustness against violations of Gaussian assumptions. Quite surprisingly thus, and except for some attempts to bootstrap the classical MLRT statistic (Goodnight and Schwartz 1997, Zhang and Boos 1992; Zhu et al. 2002), this important problem of testing for homogeneity of covariance matrices under possibly non-Gaussian elliptical densities, despite its considerable impact on applications, had remained an open problem until a recent paper by Schott (2001). In his Section 2.1, Schott first proposes a Gaussian Wald test based on the vector $((\mathrm{vec}(\dot{\mathbf{S}}_1-\dot{\mathbf{S}}_m))',\ldots,(\mathrm{vec}(\dot{\mathbf{S}}_{m-1}-\dot{\mathbf{S}}_m))')'$. This test rejects the null hypothesis for large values of the statistic

$$Q^{(n)}_{\mathrm{Schott}} := \frac{\dot{n}}{2}\left\{\sum_{i=1}^{m}\dot{\lambda}^{(n)}_i \,\mathrm{tr}\,[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2] - \sum_{i,i'=1}^{m}\dot{\lambda}^{(n)}_i\dot{\lambda}^{(n)}_{i'}\,\mathrm{tr}\,(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1})\right\}, \quad \text{with } \dot{\lambda}^{(n)}_i := \frac{\dot{n}_i}{\dot{n}}; \quad (1.4)$$

this statistic is asymptotically chi-square with $(m-1)k(k+1)/2$ degrees of freedom under homogeneity of covariance matrices and Gaussian densities. Schott himself stresses the poor resistance of his test to non-Gaussian densities, and proposes robustifying $Q^{(n)}_{\mathrm{Schott}}$ by using an adequate estimate of the underlying asymptotic covariance matrix involved in the Wald statistic. Letting $\hat{\delta}_1 := 1/(1+\hat{\kappa}_k)$ and $\hat{\delta}_2 := \hat{\kappa}_k/((k+2)\hat{\kappa}_k+2)$, the resulting test rejects $\mathcal{H}_0$ for large values of

$$Q^{(n)}_{\mathrm{Schott}*} := \hat{\delta}_1\left[Q^{(n)}_{\mathrm{Schott}} - \frac{\dot{n}\hat{\delta}_2}{2}\left\{\sum_{i=1}^{m}\dot{\lambda}^{(n)}_i\,\mathrm{tr}^2\,(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}) - \sum_{i,i'=1}^{m}\dot{\lambda}^{(n)}_i\dot{\lambda}^{(n)}_{i'}\,\mathrm{tr}\,(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})\,\mathrm{tr}\,(\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1})\right\}\right], \quad (1.5)$$

the null distribution of which is still asymptotically chi-square with $(m-1)k(k+1)/2$ degrees of freedom under the whole family of elliptical distributions with finite fourth-order moments and homogeneous (across samples) kurtosis value. In that respect, $Q^{(n)}_{\mathrm{Schott}*}$ is fairly robust. Being based on traditional covariance matrices, however, it remains invalid under heavy tails, and extremely sensitive to possible outliers.

Schott apparently is not aware of the asymptotic optimality of his test under Gaussian assumptions: in Section 5.3, we establish the asymptotic equivalence, under the null hypothesis of homogeneity and any density with finite fourth-order moments, of the MLRT, Nagao, and Schott (1.4) test statistics. This and the results of Section 5.2 imply that (i) all these tests share the Gaussian optimality properties of the MLRT, but (ii) only Schott's modified test based on (1.5), while asymptotically equivalent to the MLRT under Gaussian densities, remains valid under finite fourth-order moment non-Gaussian ones, and therefore can be considered a *pseudo-Gaussian* test.

Schott's modified test thus is an important step in the direction of a pseudo-Gaussian approach to multivariate analysis. Still, this is not entirely satisfactory. Being a multivariate extension of the univariate tests considered in the Conover et al. (1981) study, Schott's modified test is likely to be behave very poorly away from the multinormal case (this is confirmed by the simulation study of Section 7): validity robustness again is obtained to the detriment of efficiency robustness. The validity of Schott's modified test moreover is restricted to densities with finite fourth-order moments. In contrast with this, the rank-based approach we are developing here is robust on all counts. Due to the distribution-freeness of multivariate ranks and signs, our tests are valid under arbitrary elliptical densities while, when based on appropriate scores, they achieve semiparametric efficiency—the best one can hope for in the presence of unspecified densities—and almost always quite significantly outperform Schott's tests.

1.3 Our methodology.

Throughout, we assume that the m distributions are *elliptically symmetric*. More precisely, for all $i = 1, \ldots, m$, the n_i observations $\mathbf{X}_{ij}$, $j = 1, \ldots, n_i$ are assumed to have a probability density function of the form

$$\underline{f_i}(\mathbf{x}) := c_{k,f_1} |\boldsymbol{\Sigma}_i|^{-1/2} f_1 \left(\left((\mathbf{x} - \boldsymbol{\theta}_i)' \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\theta}_i) \right)^{1/2} \right), \quad \mathbf{x} \in \mathbb{R}^k, \tag{1.6}$$

for some k-dimensional vector $\boldsymbol{\theta}_i$ (*location*), some positive definite $(k \times k)$ matrix $\boldsymbol{\Sigma}_i$ (the *scatter* matrix), and some (duly standardized: see below) function $f_1 : \mathbb{R}_0^+ \to \mathbb{R}^+$ (the *radial density*).

The null hypothesis considered throughout is the hypothesis $\mathcal{H}_0 : \boldsymbol{\Sigma}_1 = \ldots = \boldsymbol{\Sigma}_m$ of *scatter homogeneity* (under finite variances, *covariance homogeneity*).

Let (throughout $\boldsymbol{\Sigma}^{1/2}$ denotes the symmetric root of $\boldsymbol{\Sigma}$)

$$\mathbf{U}_{ij}(\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_i) := \frac{\boldsymbol{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)}{\|\boldsymbol{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)\|} \quad \text{and} \quad d_{ij}(\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_i) := \|\boldsymbol{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)\|. \tag{1.7}$$

Writing $R_{ij}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_m)$ for the rank of $d_{ij}(\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_i)$ among $d_{11}(\boldsymbol{\theta}_1, \boldsymbol{\Sigma}_1), \ldots, d_{mn_m}(\boldsymbol{\theta}_m, \boldsymbol{\Sigma}_m)$, consider the signed rank scatter matrices

$$\underset{\sim}{\mathbf{S}}_{K;i} := \frac{1}{n_i} \sum_{j=1}^{n_i} K \left(\frac{R_{ij}(\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\Sigma}}, \ldots, \hat{\boldsymbol{\Sigma}})}{n+1} \right) \mathbf{U}_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\Sigma}})\, \mathbf{U}'_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\Sigma}}), \tag{1.8}$$

where $\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_m$ are consistent (under $\mathcal{H}_0$) estimates of the various location parameters, $\hat{\boldsymbol{\Sigma}}$ is a consistent (under $\mathcal{H}_0$) estimate of the common null value of the $\boldsymbol{\Sigma}_i$'s, and K is some appropriate score function. The proposed signed rank tests reject the null hypothesis of scatter homogeneity for large values of

$$\underset{\sim}{Q}_K := \frac{1}{n} \sum_{1<i<i'<m} (n_i + n_{i'})\, \underset{\sim}{Q}_{K;i,i'}, \tag{1.9}$$

where

$$\underset{\sim}{Q}_{K;i,i'} := \frac{n_i n_{i'}}{n_i + n_{i'}} \left\{ \alpha_{k,K} \operatorname{tr}\left[(\underset{\sim}{\mathbf{S}}_{K;i} - \underset{\sim}{\mathbf{S}}_{K;i'})^2 \right] + \beta_{k,K} \operatorname{tr}^2(\underset{\sim}{\mathbf{S}}_{K;i} - \underset{\sim}{\mathbf{S}}_{K;i'}) \right\};$$

$\alpha_{k,K}$ and $\beta_{k,K}$ are constants depending on the dimension k and the score function K. Although we derive (1.9) from Le Cam type optimality arguments, we show in Section 5.2 that this test statistic $\underset{\sim}{Q}_K^{(n)}$ also can be obtained by replacing, in Schott's robustified test statistic (1.5), the traditional sample covariance matrices with the signed rank scatter matrices (1.8). Being rank-based, our tests however remain valid under much broader conditions than Schott's (no finite moment requirements) and enjoy much better resistance to outliers.

The use of signed ranks is justified by the invariance principle: $\mathcal{H}_0$ indeed is invariant under groups of affine and (continuous monotone) radial transformations; see Section 3.2 for details. Beyond affine-invariance (all tests considered in this paper are affine-invariant), our rank tests—unlike their competitors—are also (asymptotically) invariant with respect to the groups of radial transformations; validity robustness follows from this latter invariance property.

As announced, our methodology combines validity and efficiency robustness. We will show that, for (essentially) any radial density f_1, it is possible to define a score function $K := K_{f_1}$ characterizing a signed rank test which is *locally and asymptotically optimal* (actually, *locally and asymptotically most stringent*, in the Le Cam sense) under radial density f_1. In particular, when based on Gaussian scores, our rank tests achieve the same asymptotic performances as Schott's (1.4) and the other Gaussian tests at the multinormal, while enjoying the validity robustness of (1.5) and even more, since no moment assumption is required. Moreover, the asymptotic relative efficiencies (AREs) of these normal-score tests are almost always larger than one with respect to their parametric competitors (see the AREs and simulations in Sections 6 and 7. The class of tests we are proposing thus in most cases dominates the existing parametric ones, both in terms of robustness and power.

1.4 Outline of the paper.

The paper is organized as follows. In Section 2, we collect the main assumptions needed in the sequel. Section 3.1 discusses semiparametric modelling issues and their relation to group invariance. Section 4.1 states the uniform local asymptotic normality result (ULAN) on which our construction of locally and asymptotically optimal tests is based. In Section 4.2, we construct rank-based versions of the *central sequences* appearing in this ULAN result. In Section 5.1, we derive and study the proposed nonparametric (signed-rank) tests (based on (1.9)) for scatter homogeneity. Section 5.2 presents the parametric Gaussian counterparts of these tests and shows how they can be turned into pseudo-Gaussian ones; their relation to Schott's modified tests based on (1.5) is also studied. In Section 5.3, we investigate the links between these tests and the main Gaussian procedures available in the literature. Asymptotic relative efficiencies with respect to the pseudo-Gaussian tests are derived in Section 6. Section 7 provides some simulation results confirming the theoretical ones. Finally, the appendix collects proofs of asymptotic linearity and other technical results.

2 Main assumptions.

For the sake of convenience, we are collecting here the main assumptions to be used in the sequel.

2.1 Elliptical symmetry.

As mentioned before, we throughout assume that all populations are elliptically symmetric. More precisely, defining the collections $\mathcal{F}$ of *radial densities* and $\mathcal{F}_1$ of *standardized radial densities* as

$$\mathcal{F} := \Big\{ f > 0 \ \text{ a.e. } : \mu_{k-1;f} < \infty \Big\} \quad \text{and} \quad \mathcal{F}_1 := \Big\{ f_1 \in \mathcal{F} \ : (\mu_{k-1;f_1})^{-1} \int_0^1 r^{k-1} f_1(r)\, dr = 1/2 \Big\},$$

respectively, where $\mu_{\ell;f} := \int_0^\infty r^\ell f(r)\, dr$, we require the following.

ASSUMPTION (A). The observations $\mathbf{X}_{ij}$, $j = 1, \ldots, n_i$, $i = 1, \ldots, m$ are mutually independent, with pdf $\underline{f_i}$, $i = 1, \ldots, m$, given in (1.6), for some $f_1 \in \mathcal{F}_1$.

Clearly, for the scatter matrices $\mathbf{\Sigma}_i$ in (1.6) to be well defined, identifiability restrictions are needed. This is why we impose that $f_1 \in \mathcal{F}_1$, which implies that $d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i)$ defined in (1.7) has median one and identifies $\mathbf{\Sigma}_i$ without requiring any moment assumptions (see Hallin and

Paindaveine 2006a for a discussion). Note however that, under finite second-order moments, $\mathbf{\Sigma}_i$ is proportional to the covariance matrix $\mathbf{\Sigma}_{0i}$ of $\mathbf{X}_{ij}$.

Special instances of elliptical densities 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 equidensity contours associated with (1.6) are hyper-ellipsoids centered at $\boldsymbol{\theta}_i$, whose shape and orientation are determined by the scatter matrix $\mathbf{\Sigma}_i$. The *multivariate signs* $\mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i)$ and *standardized radial distances* $d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i)$ defined in (1.7) are the (within-group) elliptical coordinates associated with those ellipsoids: the observation $\mathbf{X}_{ij}$ then decomposes into $\boldsymbol{\theta}_i + d_{ij}\mathbf{\Sigma}_i^{1/2}\mathbf{U}_{ij}$, where the $\mathbf{U}_{ij}$'s, $j = 1, \ldots, n_i$, $i = 1, \ldots, m$ are i.i.d. uniform over the unit sphere in $\mathbb{R}^k$, and the d_{ij}'s are i.i.d., independent of the $\mathbf{U}_{ij}$, with common density $\tilde{f}_{1k}(r) := (\mu_{k-1;f_1})^{-1} r^{k-1} f_1(r) I_{[r>0]}$ (justifying the terminology *standardized radial density* for f_1) and distribution function $\tilde{F}_{1k}$. In the sequel, the notation $\tilde{g}_{1k}$ and $\tilde{G}_{1k}$ will be used for the corresponding functions computed from a standardized radial density $g_1 (\in \mathcal{F}_1)$.

The derivation of locally and asymptotically optimal tests at 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 4.1—only holds under some further mild regularity conditions on f_1. More precisely, ULAN (see Proposition 4.1 below) requires f_1 to belong to the collection $\mathcal{F}_a$ of absolutely continuous densities in $\mathcal{F}_1$ such that, letting $\varphi_{f_1} := -\dot{f}_1/f_1$ (with $\dot{f}_1$ the a.e.-derivative of f_1), the integrals

$$\mathcal{I}_k(f_1) := \int_0^1 \varphi_{f_1}^2(\tilde{F}_{1k}^{-1}(u))\, du \quad \text{and} \quad \mathcal{J}_k(f_1) := \int_0^1 \varphi_{f_1}^2(\tilde{F}_{1k}^{-1}(u))(\tilde{F}_{1k}^{-1}(u))^2\, du$$

are finite. The quantities $\mathcal{I}_k(f_1)$ anf $\mathcal{J}_k(f_1)$ play the roles of *radial Fisher information for location* and *radial Fisher information for shape/scale*, respectively (see Hallin and Paindaveine 2006a).

2.2 Asymptotic behavior of sample sizes.

Although, for the sake of notational simplicity, we do not mention it explicitly, we actually consider sequences of statistical experiments, with triangular arrays of observations of the form $(\mathbf{X}_{1,1}^{(n)}, \ldots, \mathbf{X}_{1,n_1^{(n)}}^{(n)}, \mathbf{X}_{2,1}^{(n)}, \ldots, \mathbf{X}_{2,n_2^{(n)}}^{(n)}, \ldots, \mathbf{X}_{m,1}^{(n)}, \ldots, \mathbf{X}_{m,n_m^{(n)}}^{(n)})$ indexed by the total sample size n, where the sequences $n_i^{(n)}$ satisfy the following assumption.

ASSUMPTION (B). For all $i = 1, \ldots, m$, $n_i = n_i^{(n)} \to \infty$ as $n \to \infty$.

Note that this assumption is weaker than the corresponding classical assumption in (univariate or multivariate) multisample problems, which requires that n_i/n be bounded away from 0 and 1 for all i as $n \to \infty$. Letting $\lambda_i^{(n)} := n_i^{(n)}/n$, it is easy to check that Assumption (B) is actually equivalent to the Noether conditions

$$\max\left(\frac{1-\lambda_i^{(n)}}{\lambda_i^{(n)}}, \frac{\lambda_i^{(n)}}{1-\lambda_i^{(n)}}\right) = o(n) \quad \text{as } n \to \infty, \text{ for all } i,$$

that are needed for the representation result in Lemma 4.1(i) below. However, the following reinforcement of Assumption (B) is assumed to hold (mainly, for notational comfort) in the derivation of asymptotic distributions under local alternatives:

ASSUMPTION (B′). For all $i = 1, \ldots, m$, $\lambda_i^{(n)} \to \lambda_i \in (0, 1)$, as $n \to \infty$.

2.3 Score functions.

The score functions K appearing in the rank-based statistics (1.8) will be assumed to satisfy the following regularity assumptions.

ASSUMPTION (C). The score function $K : (0, 1) \to \mathbb{R}$ (C1) is a continuous, non-constant, and square-integrable mapping which (C2) can be expressed as the difference of two monotone increasing functions, and (C3) satisfies $\int_0^1 K(u)\, du = k$.

Assumption (C3) is a normalization constraint that is automatically satisfied by the score functions $K(u) = K_{f_1}(u) := \varphi_{f_1}(\tilde{F}_{1k}^{-1}(u))\tilde{F}_{1k}^{-1}(u)$ leading to local and asymptotic optimality at radial density f_1 (at which ULAN holds); see Section 4.1. For score functions K, K_1, K_2 satisfying Assumption (C), let

$$\mathcal{J}_k(K_1, K_2) := \mathrm{E}[K_1(U)K_2(U)] \quad \text{and} \quad \mathcal{L}_k(K_1, K_2) := \mathrm{Cov}[K_1(U), K_2(U)] = \mathcal{J}_k(K_1, K_2) - k^2$$

(throughout, U stands for a random variable uniformly distributed over $(0, 1)$), with $\mathcal{J}_k(K) := \mathcal{J}_k(K, K)$ and $\mathcal{L}_k(K) := \mathcal{L}_k(K, K)$. Also, for simplicity, we write $\mathcal{J}_k(K, f_1)$ for $\mathrm{E}[K(U)K_{f_1}(U)]$, $\mathcal{L}_k(f_1, g_1)$ for $\mathrm{E}[K_{f_1}(U)K_{g_1}(U)] - k^2$, etc.

The power score functions $K_a(u) := k(a+1)u^a$ $(a > 0)$ provide some traditional score functions satisfying Assumption (C), with $\mathcal{J}_k(K_a) = k^2(a+1)^2/(2a+1)$ and $\mathcal{L}_k(K_a) = k^2a^2/(2a+1)$: Wilcoxon and Spearman scores are obtained for $a = 1$ and $a = 2$, respectively. As for 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, denoting by Ψ_k the chi-square distribution function with k degrees of freedom,

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

Similarly, Student densities $f_1 = f_{1,\nu}^t$ yield

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

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

3 Semiparametric modeling of elliptical families.

3.1 Scatter, scale, and shape.

Consider an observed n-tuple $\mathbf{X}_1, \ldots, \mathbf{X}_n$ of i.i.d. k-dimensional elliptical random vectors, with location $\boldsymbol{\theta}$, scatter $\boldsymbol{\Sigma}$, and radial density $f_1 \in \mathcal{F}_1$ but otherwise unspecified. The family $\mathcal{P}^{(n)}$ of distributions for this observation is indexed by $(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f_1)$. As soon as a semiparametric point of view is adopted, or when rank-based methods are considered, the scatter matrix $\boldsymbol{\Sigma}$ naturally decomposes into $\boldsymbol{\Sigma} = \sigma^2\mathbf{V}$, where σ is a *scale parameter* (equivariant under multiplication by a positive constant) and $\mathbf{V}$ a *shape matrix* (invariant under multiplication by a positive constant).

A semiparametric model with specified σ and unspecified standardized radial density f_1 indeed would be highly artificial, and we therefore only consider the case under which σ and f_1 are jointly unspecified. This semiparametric setting is also the one that enjoys the group invariance structure in which the ranks and the signs to be used in our method spontaneously arise from invariance arguments; see Section 3.2 below.

The concepts of scale and shape however require a more careful definition. Denoting by $\mathcal{S}_k$ the collection of all $k \times k$ symmetric positive definite real matrices, consider a function $S : \mathcal{S}_k \to \mathbb{R}_0^+$ satisfying $S(\lambda\mathbf{\Sigma}) = \lambda S(\mathbf{\Sigma})$ for all $\lambda \in \mathbb{R}_0^+$, $\mathbf{\Sigma} \in \mathcal{S}_k$, and define scale and shape as $\sigma_S := (S(\mathbf{\Sigma}))^{1/2}$ and $\mathbf{V}_S := \mathbf{\Sigma}/S(\mathbf{\Sigma})$, respectively. Clearly, $\mathbf{V}_S$ is the only matrix in $\mathcal{S}_k$ which is proportional to $\mathbf{\Sigma}$ and satisfies $S(\mathbf{V}_S) = 1$: denote by $\mathcal{V}_k^S := \{\mathbf{V} \in \mathcal{S}_k : S(\mathbf{V}) = 1\}$ the set of all possible shape matrices associated with S.

Classical choices of S are

(i) $S(\mathbf{\Sigma}) = (\mathbf{\Sigma})_{11}$ (considered in Randles 2000, Hettmansperger and Randles 2002, Hallin and Paindaveine 2006a, and Hallin et al. 2006);

(ii) $S(\mathbf{\Sigma}) = k^{-1}\mathrm{tr}(\mathbf{\Sigma})$ (considered in Tyler 1987, Dümbgen 1998, and Ollila et al. 2004);

(iii) $S(\mathbf{\Sigma}) = |\mathbf{\Sigma}|^{1/k}$ (considered in Tatsuoka and Tyler 2000, Dümbgen and Tyler 2005, Salibian-Barrera et al. 2006, and Taskinen et al. 2006).

In practice, all choices of S are essentially equivalent. Although favoring a trace-based normalization of $\mathbf{\Sigma}^{-1}$, Bickel (1982, Example 4) actually shows that, irrespective of S, the asymptotic information matrix for $\mathbf{V}_S$ in the presence of unspecified $\boldsymbol{\theta}$ and σ_S is the same, at any $\boldsymbol{\theta} \in \mathbb{R}^k$, $\sigma_S \in \mathbb{R}_0^+$, $\mathbf{V}_S \in \mathcal{V}_k^S$ and f_1, whether f_1 is specified (parametric model) or not (semiparametric model): once $\boldsymbol{\theta}$ and σ_S are unspecified, an unspecified f_1 does not induce any additional loss for inference about $\mathbf{V}_S$. Paindaveine (2006b) establishes the stronger result that the information matrix for $\mathbf{V}_S$ in the presence of unspecified $\boldsymbol{\theta}$, σ_S and f_1 is strictly less, at any $\boldsymbol{\theta} \in \mathbb{R}^k$, $\sigma_S \in \mathbb{R}_0^+$, $\mathbf{V}_S \in \mathcal{V}_k^S$ and f_1, than in the corresponding parametric model with specified $\boldsymbol{\theta}$, σ_S and f_1—except for $S : \mathbf{\Sigma} \mapsto |\mathbf{\Sigma}|^{1/k}$, where those two information matrices coincide: under this determinant-based normalization, thus, the presence of nuisances ($\boldsymbol{\theta}$, σ_S, and f_1) (resp., $\boldsymbol{\theta}$, $\mathbf{V}_S$, and f_1) asymptotically has no effect on inference about shape (resp., inference about scale). In both cases, it can be said (adopting a point estimation terminology) that shape can be estimated *adaptively*. This *Paindaveine adaptivity*, where $\boldsymbol{\theta}$, σ_S and f_1 lie in the nuisance space of the semiparametric model, is much stronger than *Bickel adaptivity* where only f_1 does. This finding strongly pleads in favor of the determinant-based definition of shape which, with its block-diagonal information matrix for $\boldsymbol{\theta}, \sigma_S$, and $\mathbf{V}_S$, is also more convenient from the point of view of statistical inference. Therefore, we throughout adopt $S(\mathbf{\Sigma}) = |\mathbf{\Sigma}|^{1/k}$, and henceforth simply write $\mathbf{V} \in \mathcal{V}_k$ and σ for the resulting shape and scale.

The parameter in our problem then is the L-dimensional vector

$$\boldsymbol{\vartheta} := (\boldsymbol{\vartheta}'_I, \boldsymbol{\vartheta}'_{II}, \boldsymbol{\vartheta}'_{III})' := (\boldsymbol{\theta}'_1, \ldots, \boldsymbol{\theta}'_m, \sigma_1^2, \ldots, \sigma_m^2, (\mathrm{v\mathring{e}ch}\,\mathbf{V}_1)', \ldots, (\mathrm{v\mathring{e}ch}\,\mathbf{V}_m)')',$$

where $L = mk(k+3)/2$ and $\mathrm{v\mathring{e}ch}\,(\mathbf{V})$ is characterized by $\mathrm{vech}(\mathbf{V}) =: ((\mathbf{V})_{11}, (\mathrm{v\mathring{e}ch}\,\mathbf{V})')'$: indeed, $\mathbf{\Sigma}_i$ is entirely determined by σ_i^2 and $\mathrm{v\mathring{e}ch}\,(\mathbf{V}_i)$. Write $\mathbf{\Theta}$ for the set $\mathbb{R}^{mk} \times (\mathbb{R}_0^+)^m \times \mathrm{v\mathring{e}ch}\,(\mathcal{V}_k)$ of admissible $\boldsymbol{\vartheta}$ values, and $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$ or $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_I,\boldsymbol{\vartheta}_{II},\boldsymbol{\vartheta}_{III};f_1}$ for the joint distribution of the n observations under parameter value $\boldsymbol{\vartheta}$ and standardized radial density f_1 (always implicitly assumed to belong to $\mathcal{F}_1$, when notation f_1 is used).

Finally, note that for any $c(\mathbf{\Sigma}_i) > 0$, $\mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i) = \mathbf{U}_{ij}(\boldsymbol{\theta}_i, c(\mathbf{\Sigma}_i)\mathbf{\Sigma}_i) = \mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$ and $d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i) = c^{1/2}(\mathbf{\Sigma}_i) d_{ij}(\boldsymbol{\theta}_i, c(\mathbf{\Sigma}_i)\mathbf{\Sigma}_i) = \sigma_i^{-1} d_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$. It follows that the multivariate signs computed from the shape $\mathbf{V}_i$ and those computed from the scatter $\boldsymbol{\Sigma}_i$ coincide. Since, under null hypothesis $\mathcal{H}_0$ of scatter homogeneity, the (nonstandardized) radial distances computed from the common value $\mathbf{V}$ of the shape matrices are proportional to the standardized ones computed from the common value $\boldsymbol{\Sigma}$ of the scatter matrices, the corresponding ranks also coincide.

3.2 Invariance issues.

Denoting by $\mathcal{M}(\boldsymbol{\Upsilon})$ the vector space spanned by the columns of some $L \times r$ full-rank matrix $\boldsymbol{\Upsilon}$ ($r < L$), the null hypothesis of scatter homogeneity $\mathcal{H}_0 : \sigma_1^2\mathbf{V}_1 = \ldots = \sigma_m^2\mathbf{V}_m$ can be written as $\mathcal{H}_0 : \boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, with

$$\boldsymbol{\Upsilon} := \begin{pmatrix} \boldsymbol{\Upsilon}_{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Upsilon}_{II} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Upsilon}_{III} \end{pmatrix} := \begin{pmatrix} \mathbf{I}_{mk} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_m \otimes \mathbf{I}_K \end{pmatrix}, \quad K := \frac{k(k+1)}{2} - 1. \tag{3.1}$$

where $\mathbf{1}_m := (1, \ldots, 1)' \in \mathbb{R}^m$ and $\mathbf{I}_\ell$ denotes the ℓ-dimensional identity matrix.

Two distinct invariance structures play a role here. The first one is related with the group of affine transformations of the observations, which generates the parametric families $\mathcal{P}^{(n)}_{\boldsymbol{\Upsilon},f_1} := \bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}\}$. More precisely, this group is the group $\mathcal{G}^{m,k},_{\circ}$ of affine transformations of the form $\mathbf{X}_{ij} \mapsto \mathbf{A}\mathbf{X}_{ij} + \mathbf{b}_i$, where $\mathbf{A}$ is a full-rank $(k \times k)$ matrix and $\mathbf{B} := (\mathbf{b}_1, \ldots, \mathbf{b}_m)$ a $(k \times m)$ matrix. Associated with that group are the transformations $\boldsymbol{\vartheta} \mapsto \mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}}(\boldsymbol{\vartheta})$ of the parameter space, where

$$\mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}}(\boldsymbol{\vartheta}) := \Big((\mathbf{A}\boldsymbol{\theta}_1 + \mathbf{b}_1)', \ldots, (\mathbf{A}\boldsymbol{\theta}_m + \mathbf{b}_m)', |\mathbf{A}|^{2/k}\sigma_1^2, \ldots, |\mathbf{A}|^{2/k}\sigma_m^2,$$
$$(\overset{\circ}{\mathrm{vech}}\,(\mathbf{A}\mathbf{V}_1\mathbf{A}'))'/|\mathbf{A}|^{2/k}, \ldots, (\overset{\circ}{\mathrm{vech}}\,(\mathbf{A}\mathbf{V}_m\mathbf{A}'))'/|\mathbf{A}|^{2/k}\Big)'.$$

Clearly, the null hypothesis $\mathcal{H}_0$ of scatter homogeneity is invariant under $\mathcal{G}^{m,k},_{\circ}$—meaning that $\mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}}(\mathcal{M}(\boldsymbol{\Upsilon})) = \mathcal{M}(\boldsymbol{\Upsilon})$ for all $\mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}}$. Therefore, it is reasonable to restrict to affine-invariant tests of $\mathcal{H}_0$. Beyond their distribution-freeness with respect to the $\boldsymbol{\theta}_i$'s and the common null values σ and $\mathbf{V}$ of the scale and shape parameters, affine-invariant test statistics—that is, statistics Q such that $Q(\mathbf{A}\mathbf{X}_{11} + \mathbf{b}_1, \ldots \mathbf{A}\mathbf{X}_{mn_m} + \mathbf{b}_m) = Q(\mathbf{X}_{11}, \ldots, \mathbf{X}_{m,n_m})$ for all $\mathbf{A}, \mathbf{b}_1, \ldots, \mathbf{b}_m$—yield tests that are *coordinate-free*.

A second invariance structure is induced by the groups $\mathcal{G},_{\circ} := \mathcal{G}^{\boldsymbol{\vartheta}_I,\mathbf{V}},_{\circ}$ of *continuous monotone radial transformations*, of the form

$$\begin{aligned} \mathbf{X} \mapsto\ & \mathcal{G}_g(\mathbf{X}_{11}, \ldots, \mathbf{X}_{mn_m}) \\ =\ & \mathcal{G}_g(\boldsymbol{\theta}_1 + d_{11}(\boldsymbol{\theta}_1, \mathbf{V})\mathbf{V}^{1/2}\mathbf{U}_{11}(\boldsymbol{\theta}_1, \mathbf{V}), \ldots, \boldsymbol{\theta}_m + d_{mn_m}(\boldsymbol{\theta}_m, \mathbf{V})\mathbf{V}^{1/2}\mathbf{U}_{mn_m}(\boldsymbol{\theta}_m, \mathbf{V})) \\ :=\ & (\boldsymbol{\theta}_1 + g(d_{11}(\boldsymbol{\theta}_1, \mathbf{V}))\mathbf{V}^{1/2}\mathbf{U}_{11}(\boldsymbol{\theta}_1, \mathbf{V}), \ldots, \boldsymbol{\theta}_m + g(d_{mn_m}(\boldsymbol{\theta}_m, \mathbf{V}))\mathbf{V}^{1/2}\mathbf{U}_{mn_m}(\boldsymbol{\theta}_m, \mathbf{V})), \end{aligned}$$

where $g\colon \mathbb{R}^+ \!\to \mathbb{R}^+$ is continuous, monotone increasing, and such that $g(0) = 0$ and $\lim_{r\to\infty} g(r) = \infty$. For each $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, that is, for each $\boldsymbol{\vartheta}_I, \mathbf{V}$, this group $\mathcal{G}^{\boldsymbol{\vartheta}_I,\mathbf{V}},_{\circ}$ is a generating group for the nonparametric family $\mathcal{P}^{(n)}_{\boldsymbol{\vartheta}_I,\mathbf{V}} := \bigcup_\sigma \bigcup_{f_1} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_I,\sigma^2\mathbf{1}_m,\mathbf{1}_m\otimes(\overset{\circ}{\mathrm{vech}}\,\mathbf{V});f_1}\}$. In such families, the invariance principle suggests basing inference on statistics that are measurable with respect to the corresponding maximal invariant, namely the vectors $(\mathbf{U}_{11}, \ldots, \mathbf{U}_{mn_m})$ of signs and the vectors $(R_{11}, \ldots, R_{mn_m})$ of ranks, where $\mathbf{U}_{ij} = \mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{V})$, and $R_{ij} = R_{ij}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m, \mathbf{V}, \ldots, \mathbf{V})$. Such invariant statistics of course are distribution-free under $\mathcal{P}^{(n)}_{\boldsymbol{\vartheta}_I,\mathbf{V}}$.

4 Uniform local asymptotic normality, signs, and ranks.

4.1 Uniform local asymptotic normality (ULAN).

As mentioned in Section 1, we plan to develop tests that are optimal at correctly specified densities, in the sense of Le Cam's asymptotic theory of statistical experiments. In this section, we state the uniform local asymptotic normality (ULAN) result (with respect to location, scale, and shape parameters, for fixed radial density f_1) on which optimality will be based.

Writing

$$\boldsymbol{\vartheta}^{(n)} = (\boldsymbol{\vartheta}_{I}^{(n)\prime}, \boldsymbol{\vartheta}_{II}^{(n)\prime}, \boldsymbol{\vartheta}_{III}^{(n)\prime})' = (\boldsymbol{\theta}_1^{(n)\prime}, \ldots, \boldsymbol{\theta}_m^{(n)\prime}, \sigma_1^{2(n)}, \ldots, \sigma_m^{2(n)}, (\mathrm{v\overset{\circ}{e}ch}\, \mathbf{V}_1^{(n)})', \ldots, (\mathrm{v\overset{\circ}{e}ch}\, \mathbf{V}_m^{(n)})')'$$

for an arbitrary sequence of L-dimensional parameter values in $\boldsymbol{\Theta}$, consider sequences of "local alternatives" $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}$, where

$$\boldsymbol{\tau}^{(n)} = (\boldsymbol{\tau}_{I}^{(n)\prime}, \boldsymbol{\tau}_{II}^{(n)\prime}, \boldsymbol{\tau}_{III}^{(n)\prime})' = (\mathbf{t}_1^{(n)\prime}, \ldots, \mathbf{t}_m^{(n)\prime}, s_1^{2(n)}, \ldots, s_m^{2(n)}, (\mathrm{v\overset{\circ}{e}ch}\, \mathbf{v}_1^{(n)})', \ldots, (\mathrm{v\overset{\circ}{e}ch}\, \mathbf{v}_m^{(n)})')'$$

is such that $\sup_n \boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)} < \infty$ and where, denoting by $\boldsymbol{\Lambda}^{(n)} = (\Lambda_{rs}^{(n)})$ the $(m \times m)$ diagonal matrix with $\Lambda_{ii}^{(n)} := (\lambda_i^{(n)})^{-1/2}$ (see Section 2.2),

$$\boldsymbol{\nu}^{(n)} := \begin{pmatrix} \boldsymbol{\nu}_{I}^{(n)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\nu}_{II}^{(n)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\nu}_{III}^{(n)} \end{pmatrix} := \begin{pmatrix} \boldsymbol{\Lambda}^{(n)} \otimes \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{(n)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}^{(n)} \otimes \mathbf{I}_K \end{pmatrix} \tag{4.1}$$

(under Assumption (B′), we also write $\boldsymbol{\nu}$ for $\lim_{n\to\infty} \boldsymbol{\nu}^{(n)}$). Clearly, these local alternatives do not involve $(\mathbf{v}_i^{(n)})_{11}$, $i = 1, \ldots, m$. It is natural, though, to see that the perturbed shapes $\mathbf{V}_i^{(n)} + n_i^{-1/2}\mathbf{v}_i^{(n)}$ remain (up to $o(n_i^{-1/2})$'s) within the family $\mathcal{V}_k$ of shape matrices: this leads to defining $(\mathbf{v}_i^{(n)})_{11}$ in such a way that $\mathrm{tr}((\mathbf{V}_i^{(n)})^{-1}\mathbf{v}_i^{(n)}) = 0$, $i = 1, \ldots, m$, which entails $|\mathbf{V}_i^{(n)} + n_i^{-1/2}\mathbf{v}_i^{(n)}|^{1/k} = 1 + o(n_i^{-1/2})$ (see Hallin and Paindaveine 2006b, Section 4).

The following notation will be used throughout. Write $\mathbf{V}^{\otimes 2}$ for the Kronecker product $\mathbf{V} \otimes \mathbf{V}$. 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*, and put $\mathbf{J}_k := (\mathrm{vec}\, \mathbf{I}_k)(\mathrm{vec}\, \mathbf{I}_k)'$. Finally, let $\mathbf{M}_k(\mathbf{V})$ be the $(K \times k^2)$ matrix such that $(\mathbf{M}_k(\mathbf{V}))'(\mathrm{v\overset{\circ}{e}ch}\, \mathbf{v}) = (\mathrm{vec}\, \mathbf{v})$ for any symmetric $k \times k$ matrix $\mathbf{v}$ such that $\mathrm{tr}(\mathbf{V}^{-1}\mathbf{v}) = 0$. As shown in Paindaveine (2006b; Lemma 4.2(v)), $\mathbf{M}_k(\mathbf{V})(\mathrm{vec}\, \mathbf{V}^{-1}) = \mathbf{0}$ for all $\mathbf{V} \in \mathcal{V}_k$.

We then have the following ULAN result; the proof follows along the same lines as in Theorem 2.1 of Paindaveine (2006b) and hence is omitted.

Proposition 4.1 *Assume that (A) and (B) hold, and that $f_1 \in \mathcal{F}_a$. Then the family $\mathcal{P}_{f_1}^{(n)} := \{\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)} | \boldsymbol{\vartheta} \in \boldsymbol{\Theta}\}$ is ULAN, with central sequence*

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1} = \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{(n)} := \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I} \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{II} \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III} \end{pmatrix}, \ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I} = \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I,1} \\ \vdots \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I,m} \end{pmatrix}, \ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{II} = \begin{pmatrix} \Delta_{\boldsymbol{\vartheta};f_1}^{II,1} \\ \vdots \\ \Delta_{\boldsymbol{\vartheta};f_1}^{II,m} \end{pmatrix}, \ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III} = \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III,1} \\ \vdots \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III,m} \end{pmatrix},$$

where (with $d_{ij} = d_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$ and $\mathbf{U}_{ij} = \mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$)

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I,i} := \frac{n_i^{-1/2}}{\sigma_i}\sum_{j=1}^{n_i} \varphi_{f_1}\left(\frac{d_{ij}}{\sigma_i}\right) \mathbf{V}_i^{-1/2}\mathbf{U}_{ij}, \qquad \Delta_{\boldsymbol{\vartheta};f_1}^{II,i} := \frac{n_i^{-1/2}}{2\sigma_i^2}\sum_{j=1}^{n_i} \left(\varphi_{f_1}\left(\frac{d_{ij}}{\sigma_i}\right)\frac{d_{ij}}{\sigma_i} - k\right),$$

$$\boldsymbol{\Delta}^{\scriptscriptstyle III,i}_{\boldsymbol{\vartheta};f_1} := \frac{n_i^{-1/2}}{2}\mathbf{M}_k(\mathbf{V}_i)\left(\mathbf{V}_i^{\otimes 2}\right)^{-1/2}\sum_{j=1}^{n_i}\varphi_{f_1}\left(\frac{d_{ij}}{\sigma_i}\right)\frac{d_{ij}}{\sigma_i}\,\mathrm{vec}\left(\mathbf{U}_{ij}\mathbf{U}'_{ij}\right),$$

$i = 1,\ldots,m$, *and full-rank block-diagonal information matrix*

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1} := \begin{pmatrix} \boldsymbol{\Gamma}^{\scriptscriptstyle I}_{\boldsymbol{\vartheta};f_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Gamma}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1} \end{pmatrix}, \tag{4.2}$$

where, defining $\underline{\boldsymbol{\sigma}} := \mathrm{diag}(\sigma_1,\ldots,\sigma_m)$, $\underline{\mathbf{V}} := \mathrm{diag}(\mathbf{V}_1,\ldots,\mathbf{V}_m)$, $\mathbf{M}_k(\underline{\mathbf{V}}) := \mathrm{diag}(\mathbf{M}_k(\mathbf{V}_1),\ldots,\mathbf{M}_k(\mathbf{V}_m))$, *and* $\underline{\mathbf{V}^{\otimes 2}} := \mathrm{diag}(\mathbf{V}_1^{\otimes 2},\ldots,\mathbf{V}_m^{\otimes 2})$, *we let*

$$\boldsymbol{\Gamma}^{\scriptscriptstyle I}_{\boldsymbol{\vartheta};f_1} := \frac{1}{k}\mathcal{I}_k(f_1)(\underline{\boldsymbol{\sigma}}^{-2}\otimes\mathbf{I}_k)\underline{\mathbf{V}}^{-1}, \quad \boldsymbol{\Gamma}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1} := \frac{1}{4}\mathcal{L}_k(f_1)\,\underline{\boldsymbol{\sigma}}^{-4},$$

and

$$\boldsymbol{\Gamma}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1} := \frac{\mathcal{J}_k(f_1)}{4k(k+2)}\mathbf{M}_k(\underline{\mathbf{V}})\left[\mathbf{I}_m\otimes(\mathbf{I}_{k^2}+\mathbf{K}_k)\right]\left(\underline{\mathbf{V}^{\otimes 2}}\right)^{-1}(\mathbf{M}_k(\underline{\mathbf{V}}))'.$$

More precisely, for any $\boldsymbol{\vartheta}^{(n)} = \boldsymbol{\vartheta} + O(n^{-1/2})$ *and any bounded sequence* $\boldsymbol{\tau}^{(n)}$, *we have, under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}$,

$$\begin{aligned}\Lambda^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}/\boldsymbol{\vartheta}^{(n)};f_1} &:= \log\left(d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};f_1}/d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}\right) \\ &= (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1} - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1)\end{aligned}$$

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 classical theory of hypothesis testing in Gaussian shifts (see Section 11.9 of Le Cam 1986) then provides the general form for locally asymptotically optimal (namely, *most stringent*) tests of hypotheses in ULAN models. Such tests, for a null hypothesis of the form $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$, should be based on the asymptotically chi-square null distribution of

$$Q_{\boldsymbol{\Upsilon}} := (\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1})'\,\boldsymbol{\Gamma}^{-1/2}_{\boldsymbol{\vartheta};f_1}\Big[\mathbf{I} - \mathrm{proj}(\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta};f_1}(\boldsymbol{\nu}^{(n)})^{-1}\boldsymbol{\Upsilon})\Big]\boldsymbol{\Gamma}^{-1/2}_{\boldsymbol{\vartheta};f_1}\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1},$$

where $\mathrm{proj}(\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta};f_1}(\boldsymbol{\nu}^{(n)})^{-1}\boldsymbol{\Upsilon})$ is the matrix projecting $\mathbb{R}^L$ onto $\mathcal{M}(\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta};f_1}(\boldsymbol{\nu}^{(n)})^{-1}\boldsymbol{\Upsilon})$ (with $\boldsymbol{\vartheta}$ replaced by an appropriate estimator $\hat{\boldsymbol{\vartheta}}$; see Assumption (D) below). Whenever $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}$, $\boldsymbol{\nu}^{(n)}$ and $\boldsymbol{\Upsilon}$ all happen to be block-diagonal, which is the case in our problem, this projection matrix clearly is block-diagonal, with diagonal blocks

$$\mathrm{proj}((\boldsymbol{\Gamma}^{\scriptscriptstyle I}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle I})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle I}),\ \mathrm{proj}((\boldsymbol{\Gamma}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle II})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle II}),\ \text{ and }\ \mathrm{proj}((\boldsymbol{\Gamma}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle III})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle III})$$

denoting projections in $\mathbb{R}^{mk}$, $\mathbb{R}^m$, and $\mathbb{R}^{mK}$, respectively. Since moreover $\mathcal{M}((\boldsymbol{\Gamma}^{\scriptscriptstyle I}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle I})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle I}) = \mathbb{R}^{mk}$, $\mathrm{proj}((\boldsymbol{\Gamma}^{\scriptscriptstyle I}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle I})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle I}) = \mathbf{I}_{mk}$, so that $Q_{\boldsymbol{\Upsilon}}$ reduces to

$$\begin{aligned}Q_{\boldsymbol{\Upsilon}} = {} & (\boldsymbol{\Delta}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1})'\,(\boldsymbol{\Gamma}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1})^{1/2}\left[\mathbf{I} - \mathrm{proj}((\boldsymbol{\Gamma}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle II})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle II})\right](\boldsymbol{\Gamma}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1})^{1/2}\boldsymbol{\Delta}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1} \\ & + (\boldsymbol{\Delta}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1})'\,(\boldsymbol{\Gamma}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1})^{1/2}\left[\mathbf{I} - \mathrm{proj}((\boldsymbol{\Gamma}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1})^{1/2}(\boldsymbol{\nu}^{(n)}_{\scriptscriptstyle III})^{-1}\boldsymbol{\Upsilon}_{\scriptscriptstyle III})\right](\boldsymbol{\Gamma}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1})^{1/2}\boldsymbol{\Delta}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1}\end{aligned} \tag{4.3}$$

where $\boldsymbol{\Delta}^{\scriptscriptstyle I}_{\boldsymbol{\vartheta};f_1}$ does not play any role. Accordingly, in the next section, we proceed with rank-based analogues of $\boldsymbol{\Delta}^{\scriptscriptstyle II}_{\boldsymbol{\vartheta};f_1}$ and $\boldsymbol{\Delta}^{\scriptscriptstyle III}_{\boldsymbol{\vartheta};f_1}$ only.

4.2 A rank-based central sequence for scale and shape (scatter).

A general result by Hallin and Werker (2003) implies that, in adaptive models for which fixed-f submodels are ULAN and fixed-$\boldsymbol{\vartheta}$ submodels are generated by a group $\mathcal{G}_{\boldsymbol{\vartheta}}$, invariant versions of central sequences exist under very general assumptions. In the present context, this result would imply the existence, for the null values of $\boldsymbol{\vartheta}$ ($\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$), of central sequences based on the multivariate signs $\mathbf{U}_{ij}$ and the ranks R_{ij}. Although that result does not directly apply here, it is very likely that it still holds. This fact is confirmed by the asymptotic representation of Lemma 4.1(i) below.

Consider the signed rank statistic (associated with some score function K satisfying Assumption (C)) $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\vartheta};K} := ((\underset{\sim}{\boldsymbol{\Delta}}^{\mathit{II}}_{\boldsymbol{\vartheta};K})', (\underset{\sim}{\boldsymbol{\Delta}}^{\mathit{III}}_{\boldsymbol{\vartheta};K})')' := ((\underset{\sim}{\Delta}^{\mathit{II},1}_{\boldsymbol{\vartheta};K})', \ldots, (\underset{\sim}{\Delta}^{\mathit{II},m}_{\boldsymbol{\vartheta};K})', (\underset{\sim}{\boldsymbol{\Delta}}^{\mathit{III},1}_{\boldsymbol{\vartheta};K})', \ldots, (\underset{\sim}{\boldsymbol{\Delta}}^{\mathit{III},m}_{\boldsymbol{\vartheta};K})')'$, where

$$\underset{\sim}{\Delta}^{\mathit{II},i}_{\boldsymbol{\vartheta};K} := \frac{1}{2\sigma_i^2}\, n_i^{-1/2} \sum_{j=1}^{n_i} \left(K\!\left(\frac{R_{ij}}{n+1}\right) - k \right) \tag{4.4}$$

and

$$\underset{\sim}{\boldsymbol{\Delta}}^{\mathit{III},i}_{\boldsymbol{\vartheta};K} := \frac{1}{2}\, n_i^{-1/2} \mathbf{M}_k(\mathbf{V}_i) \left(\mathbf{V}_i^{\otimes 2}\right)^{-1/2} \sum_{j=1}^{n_i} K\!\left(\frac{R_{ij}}{n+1}\right) \operatorname{vec}\left(\mathbf{U}_{ij}\mathbf{U}'_{ij}\right). \tag{4.5}$$

The following lemma provides (i) an asymptotic representation and (ii) the asymptotic distribution of $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\vartheta};K}$ (see the appendix for the proof). An immediate corollary of (i) is that $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\vartheta};f_1} := \underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\vartheta};K_{f_1}}$, with $K = K_{f_1}$, actually constitutes a signed-rank version of the scatter part $((\boldsymbol{\Delta}^{\mathit{II}}_{\boldsymbol{\vartheta};f_1})', (\boldsymbol{\Delta}^{\mathit{III}}_{\boldsymbol{\vartheta};f_1})')'$ of the central sequence $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}$.

Lemma 4.1 *Assume that (A), (B), and (C) hold. Fix $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ (with common values σ and $\mathbf{V}$ of the scale and shape parameters). Let R_{ij} be the rank of $d_{ij} := d_{ij}(\boldsymbol{\theta}_i, \mathbf{V})$ among $d_{11}, \ldots, d_{mn_m}$, and let $\mathbf{U}_{ij} := \mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{V})$. Then, for all $g_1 \in \mathcal{F}_1$,*

(i) $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\vartheta};K} = \boldsymbol{\Delta}_{\boldsymbol{\vartheta};K;g_1} + o_{L^2}(1)$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, as $n \to \infty$, where

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};K;g_1} := ((\boldsymbol{\Delta}^{\mathit{II}}_{\boldsymbol{\vartheta};K;g_1})', (\boldsymbol{\Delta}^{\mathit{III}}_{\boldsymbol{\vartheta};K;g_1})')' := ((\Delta^{\mathit{II},1}_{\boldsymbol{\vartheta};K;g_1})', \ldots, (\Delta^{\mathit{II},m}_{\boldsymbol{\vartheta};K;g_1})', (\boldsymbol{\Delta}^{\mathit{III},1}_{\boldsymbol{\vartheta};K;g_1})', \ldots, (\boldsymbol{\Delta}^{\mathit{III},m}_{\boldsymbol{\vartheta};K;g_1})')',$$

with

$$\Delta^{\mathit{II},i}_{\boldsymbol{\vartheta};K;g_1} := \frac{1}{2\sigma^2}\, n_i^{-1/2} \sum_{j=1}^{n_i} \left(K\left(\tilde{G}_{1k}\left(\frac{d_{ij}}{\sigma}\right)\right) - k \right)$$

and

$$\boldsymbol{\Delta}^{\mathit{III},i}_{\boldsymbol{\vartheta};K;g_1} := \frac{1}{2}\, n_i^{-1/2} \mathbf{M}_k(\mathbf{V}) \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \sum_{j=1}^{n_i} K\left(\tilde{G}_{1k}\left(\frac{d_{ij}}{\sigma}\right)\right) \operatorname{vec}\left(\mathbf{U}_{ij}\mathbf{U}'_{ij}\right); \tag{4.6}$$

(ii) defining $\mathbf{H}_k(\mathbf{V}) := \frac{1}{4k(k+2)} \mathbf{M}_k(\mathbf{V}) \left[\mathbf{I}_{k^2} + \mathbf{K}_k\right] \left(\mathbf{V}^{\otimes 2}\right)^{-1} (\mathbf{M}_k(\mathbf{V}))'$, $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};K;g_1}$ is asymptotically normal with mean zero and mean

$$\begin{pmatrix} \frac{1}{4\sigma^4}\mathcal{L}_k(K, g_1)\boldsymbol{\tau}_{\mathit{II}} \\ \mathcal{J}_k(K, g_1)[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})]\boldsymbol{\tau}_{\mathit{III}} \end{pmatrix}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ and $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau};g_1}$, respectively, and covariance matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};K} := \begin{pmatrix} \boldsymbol{\Gamma}^{\mathit{II}}_{\boldsymbol{\vartheta};K} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}^{\mathit{III}}_{\boldsymbol{\vartheta};K} \end{pmatrix} := \begin{pmatrix} \frac{1}{4\sigma^4}\mathcal{L}_k(K)\mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathcal{J}_k(K)[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})] \end{pmatrix}$$

under both (the claim under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau};g_1}$ *further requires* $g_1 \in \mathcal{F}_a$*).*

As mentioned in the description of the most stringent tests (see the comments after Proposition 4.1), we will need replacing the parameter $\boldsymbol{\vartheta}$ with some estimate. For this purpose, we assume the existence of $\hat{\boldsymbol{\vartheta}} := \hat{\boldsymbol{\vartheta}}^{(n)}$ satisfying

Assumption (D). The sequence of estimators $(\hat{\boldsymbol{\vartheta}}^{(n)}, n \in \mathbb{N})$ is

(D1) *constrained:* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}[\hat{\boldsymbol{\vartheta}}^{(n)} \in \mathcal{M}(\boldsymbol{\Upsilon})] = 1$ for all n, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and $g_1 \in \mathcal{F}_1$;

(D2) $n^{1/2}(\boldsymbol{\nu}^{(n)})^{-1}$*-consistent:* for all $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, $n^{1/2}(\boldsymbol{\nu}^{(n)})^{-1}(\hat{\boldsymbol{\vartheta}}^{(n)} - \boldsymbol{\vartheta}) = O_{\mathrm{P}}(1)$, as $n \to \infty$, under $\bigcup_{g_1 \in \mathcal{F}_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$;

(D3) *locally asymptotically discrete:* for all $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ 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 $\{\mathbf{t} \in \mathbb{R}^L : n^{1/2}\|(\boldsymbol{\nu}^{(n)})^{-1}(\mathbf{t} - \boldsymbol{\vartheta})\| \leq c\}$ is bounded by M, uniformly as $n \to \infty$, and

(D4) *affine-equivariant:* denoting by $\hat{\boldsymbol{\vartheta}}^{(n)}(\mathbf{A}, \mathbf{B})$ the value of $\hat{\boldsymbol{\vartheta}}^{(n)}$ computed from the transformed sample $\mathbf{A}\mathbf{X}_{ij} + \mathbf{b}_i$, $j = 1, \ldots, n_i$, $i = 1, \ldots, m$, $\hat{\boldsymbol{\vartheta}}^{(n)}(\mathbf{A}, \mathbf{B}) = \mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}}(\hat{\boldsymbol{\vartheta}}^{(n)})$, for all $\mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}} \in \mathcal{G}^{m,k}$.

There are many possible choices for $\hat{\boldsymbol{\vartheta}}$. However, still in order to avoid moment assumptions, we propose the following estimators, related with the affine-equivariant median proposed by Hettmansperger and Randles (2002). For each $i = 1, \ldots, m$, let $\hat{\boldsymbol{\theta}}_i$ and $\hat{\mathbf{V}}_i$ be characterized by

$$\frac{1}{n_i}\sum_{j=1}^{n_i} \mathbf{U}_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{V}}_i) = \mathbf{0} \quad \text{and} \quad \frac{1}{n_i}\sum_{j=1}^{n_i} \mathbf{U}_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{V}}_i)(\mathbf{U}_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{V}}_i))' = \frac{1}{k}\mathbf{I}_k,$$

with $|\hat{\mathbf{V}}_i| = 1$. Then, under $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, the common value $\mathbf{V}$ of the shape matrices $\mathbf{V}_i$ is consistently estimated (as $n \to \infty$, under $\bigcup_{g_1}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$ and Assumptions (A1) and (B), and *without any moment assumption on* g_1), at the rate required by Assumption (D2), by the Tyler estimator $\hat{\mathbf{V}}$ computed from the n data points $\mathbf{X}_{ij} - \hat{\boldsymbol{\theta}}_i$ and normalized in such a way that $|\hat{\mathbf{V}}| = 1$. Under the same conditions, the common scale σ is the median of the i.i.d. radial distances $d_{ij}(\boldsymbol{\theta}_i, \mathbf{V})$, so that the empirical median $\hat{\sigma}$ of the estimated distances $d_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{V}})$ can be used as an estimator of σ. Consequently, the estimator

$$\hat{\boldsymbol{\vartheta}} := (\hat{\boldsymbol{\theta}}'_1, \ldots, \hat{\boldsymbol{\theta}}'_m, \hat{\sigma}^2\mathbf{1}'_m, \mathbf{1}'_m \otimes (\overset{\circ}{\mathrm{vech}}\,\hat{\mathbf{V}})')' \tag{4.7}$$

satisfies (D2) above—except perhaps for the $\hat{\boldsymbol{\vartheta}}_{II}$ part, which however is not involved in the test statistics below. This estimator also satisfies (D1) and (D4). As for (D3), it is a purely technical requirement, with little practical implications (for fixed sample size, any estimator indeed can be considered part of a locally asymptotically discrete sequence). Therefore, we henceforth assume that (4.7) satisfies Assumption (D).

The resulting ranks $\hat{R}_{ij} := R_{ij}(\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_m, \hat{\mathbf{V}}, \ldots, \hat{\mathbf{V}})$ are usually called *aligned ranks.* The following *asymptotic linearity* result describes the asymptotic behavior of the aligned versions $\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\vartheta}};K}$ of the rank statistics $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\vartheta};K}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$; see the appendix for the proof.

Proposition 4.2 *Assume that (A), (B), (C), and (D1)-(D3) hold, and that* $g_1 \in \mathcal{F}_a$. *Fix* $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ *(with common values* σ *and* $\mathbf{V}$ *for the scale and shape parameters). Then,*

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{II}_{\hat{\boldsymbol{\vartheta}};K} - \underset{\sim}{\boldsymbol{\Delta}}{}^{II}_{\boldsymbol{\vartheta};K} + \frac{1}{4\sigma^4}\,\mathcal{L}_k(K,g_1)\,(\boldsymbol{\nu}^{(n)}_{II})^{-1} n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)}_{II} - \boldsymbol{\vartheta}_{II})$$

and

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\hat{\boldsymbol{\vartheta}};K} - \underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\boldsymbol{\vartheta};K} + \mathcal{J}_k(K,g_1)\,[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})]\,(\boldsymbol{\nu}^{(n)}_{III})^{-1}\, n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)}_{III} - \boldsymbol{\vartheta}_{III})$$

are $o_{\mathrm{P}}(1)$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, *as* $n \to \infty$.

5 Optimal tests of scatter homogeneity.

5.1 Optimal rank-based tests.

For all $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ (with common values σ and $\mathbf{V}$ for the scale and shape parameters), define

$$\begin{aligned}(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{\perp} &:= (\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{-1} - (\boldsymbol{\nu}^{(n)}_{II})^{-1}\boldsymbol{\Upsilon}_{II}\left(\boldsymbol{\Upsilon}'_{II}(\boldsymbol{\nu}^{(n)}_{II})^{-1}\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K}(\boldsymbol{\nu}^{(n)}_{II})^{-1}\boldsymbol{\Upsilon}_{II}\right)^{-1}\boldsymbol{\Upsilon}'_{II}(\boldsymbol{\nu}^{(n)}_{II})^{-1}\\ &= \frac{4\sigma^4}{\mathcal{L}_k(K)}[\mathbf{I}_m - \mathbf{C}^{(n)}]\end{aligned}$$

and

$$\begin{aligned}(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{\perp} &:= (\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{-1} - (\boldsymbol{\nu}^{(n)}_{III})^{-1}\boldsymbol{\Upsilon}_{III}\left(\boldsymbol{\Upsilon}'_{III}(\boldsymbol{\nu}^{(n)}_{III})^{-1}\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K}(\boldsymbol{\nu}^{(n)}_{III})^{-1}\boldsymbol{\Upsilon}_{III}\right)^{-1}\boldsymbol{\Upsilon}'_{III}(\boldsymbol{\nu}^{(n)}_{III})^{-1}\\ &= (\mathcal{J}_k(K))^{-1}[\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes (\mathbf{H}_k(\mathbf{V}))^{-1},\end{aligned}$$

where $\mathbf{C}^{(n)} = (C^{(n)}_{ij})$ denotes the $m \times m$ matrix with entries $C^{(n)}_{ij} := (\lambda^{(n)}_i \lambda^{(n)}_j)^{1/2}$. The K-score version $\underset{\sim}{\phi}{}^{(n)}_K$ of the rank-based tests we are proposing rejects $\mathcal{H}_0 : \boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ as soon as

$$\begin{aligned}\underset{\sim}{Q}{}^{(n)}_K &:= \left(\underset{\sim}{\boldsymbol{\Delta}}{}^{II}_{\hat{\boldsymbol{\vartheta}};K}\right)'(\boldsymbol{\Gamma}^{II}_{\hat{\boldsymbol{\vartheta}};K})^{\perp}\underset{\sim}{\boldsymbol{\Delta}}{}^{II}_{\hat{\boldsymbol{\vartheta}};K} + \left(\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\hat{\boldsymbol{\vartheta}};K}\right)'(\boldsymbol{\Gamma}^{III}_{\hat{\boldsymbol{\vartheta}};K})^{\perp}\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\hat{\boldsymbol{\vartheta}};K}\\ &= \sum_{i,i'=1}^{m}\frac{\delta_{i,i'} - (\lambda^{(n)}_i\lambda^{(n)}_{i'})^{1/2}}{(n_i n_{i'})^{1/2}}\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\left\{\frac{1}{\mathcal{L}_k(K)}\left(K\Big(\frac{\hat{R}_{ij}}{n+1}\Big) - k\right)\left(K\Big(\frac{\hat{R}_{i'j'}}{n+1}\Big) - k\right)\right.\\ &\qquad\left. + \frac{k(k+2)}{2\mathcal{J}_k(K)}K\Big(\frac{\hat{R}_{ij}}{n+1}\Big)K\Big(\frac{\hat{R}_{i'j'}}{n+1}\Big)\left((\hat{\mathbf{U}}'_{ij}\hat{\mathbf{U}}_{i'j'})^2 - \frac{1}{k}\right)\right\} \qquad (5.1)\end{aligned}$$

exceeds the α-upper quantile $\chi^2_{(m-1)(K+1);1-\alpha}$ of the chi-square distribution with $(m-1)(K+1)$ degrees of freedom ($\delta_{i,i'}$ stands for the usual Kronecker symbol); the explicit form of $(\mathbf{H}_k(\mathbf{V}))^{-1}$ allowing for (5.1) can be found in Lemma 5.2 of Hallin and Paindaveine (2006b).

In the sequel, we write $\underset{\sim}{\phi}{}^{(n)}_{f_1}$ and $\underset{\sim}{Q}{}^{(n)}_{f_1}$ for $\underset{\sim}{\phi}{}^{(n)}_{K_{f_1}}$ and $\underset{\sim}{Q}{}^{(n)}_{K_{f_1}}$, respectively.

We are now ready to state the main result of this paper; for the sake of simplicity, asymptotic powers are expressed under Assumption (B$'$) and perturbations $\boldsymbol{\tau}^{(n)}$ such that $\lim_{n\to\infty}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)} = \boldsymbol{\nu\tau} \notin \mathcal{M}(\boldsymbol{\Upsilon})$, with $\boldsymbol{\nu}_{II}\boldsymbol{\tau}_{II} = (s_1^2/\sqrt{\lambda_1}, \ldots, s_m^2/\sqrt{\lambda_m})'$ and $\boldsymbol{\nu}_{III}\boldsymbol{\tau}_{III} = ((\overset{\circ}{\mathrm{vech}}\,\mathbf{v}_1)'/\sqrt{\lambda_1}, \ldots, (\overset{\circ}{\mathrm{vech}}\,\mathbf{v}_m)'/\sqrt{\lambda_m})'$. For any such $\boldsymbol{\tau}$ and any $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ (still with common values σ^2 and $\mathbf{V}$ of the scale and shape parameters), let

$$r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} := \frac{1}{\sigma^4}\lim_{n\to\infty}\left\{(\boldsymbol{\tau}^{(n)}_{II})'[\mathbf{I}_m - \mathbf{C}^{(n)}]\,\boldsymbol{\tau}^{(n)}_{II}\right\} = \sum_{1\le i<i'\le m}\frac{\lambda_i\lambda_{i'}}{\sigma^4}\left(\frac{s_i^2}{\sqrt{\lambda_i}} - \frac{s_{i'}^2}{\sqrt{\lambda_{i'}}}\right)^2$$

and

$$r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} := 2k(k+2)\lim_{n\to\infty}\Big\{(\boldsymbol{\tau}^{(n)}_{III})'\Big[[\mathbf{I}_m-\mathbf{C}^{(n)}]\otimes\mathbf{H}_k(\mathbf{V})\Big]\boldsymbol{\tau}^{(n)}_{III}\Big\} = \sum_{1\le i<i'\le m}\lambda_i\lambda_{i'}\,\mathrm{tr}\left[\Big(\mathbf{V}^{-1}\Big(\frac{\mathbf{v}_i}{\sqrt{\lambda_i}}-\frac{\mathbf{v}_{i'}}{\sqrt{\lambda_{i'}}}\Big)\Big)^2\right];$$

recall that $\mathrm{tr}(\mathbf{V}^{-1}\mathbf{v}_i^{(n)}) = 0$, $i = 1,\ldots,m$ (see the comments before Proposition 4.1).

Theorem 5.1 *Assume that (A), (B), (C), and (D1-D3) hold. Then,*

(i) $\underset{\sim}{Q}^{(n)}_K$ *is asymptotically chi-square with* $(m-1)(K+1)$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$*, and (provided that (B) is reinforced into (B′)) asymptotically noncentral chi-square, still with* $(m-1)(K+1)$ *degrees of freedom, but with noncentrality parameter*

$$\frac{\mathcal{L}^2_k(K,g_1)}{4\mathcal{L}_k(K)}\,r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} + \frac{\mathcal{J}^2_k(K,g_1)}{2k(k+2)\mathcal{J}_k(K)}\,r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} \tag{5.2}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$, $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$, $\boldsymbol{\nu\tau} := \lim_{n\to\infty}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)} \notin \mathcal{M}(\boldsymbol{\Upsilon})$*, and* $g_1\in\mathcal{F}_a$*;*

(ii) the sequence of tests $\underset{\sim}{\phi}^{(n)}_K$ *has asymptotic level* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$*;*

(iii) if $f_1\in\mathcal{F}_a$ *and* K_{f_1} *satisfies Assumption (C), the sequence of tests* $\underset{\sim}{\phi}^{(n)}_{f_1}$ *is locally and asymptotically most stringent, still at asymptotic level* α*, for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g_1\in\mathcal{F}_a}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{M}(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}\}$.

See the appendix for the proof.

Of course, provided that K_{f_1} satisfies (C), (i) holds for $\underset{\sim}{Q}^{(n)}_{f_1}$ with $\mathcal{L}_k(K,g_1)$ and $\mathcal{J}_k(K,g_1)$ replaced, in (5.2), by $\mathcal{L}_k(f_1,g_1)$ and $\mathcal{J}_k(f_1,g_1)$, respectively. After some algebra, one obtains

$$\underset{\sim}{Q}^{(n)}_K = \frac{1}{n}\sum_{1\le i<i'\le m}(n_i+n_{i'})\,\underset{\sim}{Q}^{(n)}_{K;i,i'},$$

where

$$\underset{\sim}{Q}^{(n)}_{K;i,i'} = \frac{n_in_{i'}}{n_i+n_{i'}}\Bigg\{\frac{1}{\mathcal{L}_k(K)}\left[\frac{1}{n_i}\sum_{j=1}^{n_i}K\left(\frac{\hat{R}_{ij}}{n+1}\right) - \frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}}K\left(\frac{\hat{R}_{i'j'}}{n+1}\right)\right]^2 \tag{5.3}$$
$$+\frac{k(k+2)}{2\mathcal{J}_k(K)}\left\|\left[\frac{1}{n_i}\sum_{j=1}^{n_i}K\Big(\frac{\hat{R}_{ij}}{n+1}\Big)\mathrm{vec}\Big(\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij}-\frac{1}{k}\mathbf{I}_k\Big)\right] - \left[\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}}K\Big(\frac{\hat{R}_{i'j'}}{n+1}\Big)\mathrm{vec}\Big(\hat{\mathbf{U}}_{i'j'}\hat{\mathbf{U}}'_{i'j'}-\frac{1}{k}\mathbf{I}_k\Big)\right]\right\|^2\Bigg\}$$

is the test statistic obtained in the two-sample case (for populations i and i'); see Um and Randles (1998) for a similar decomposition in MANOVA problems. Note that no estimate $\hat{\boldsymbol{\vartheta}}_{II}$ of the common scale appears in the test statistics. Also, letting

$$\underset{\sim}{\mathbf{S}}_{K;i} := \frac{1}{n_i}\sum_{j=1}^{n_i}K\left(\frac{\hat{R}_{ij}}{n+1}\right)\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij},$$

the statistics in (5.3) take the simple form

$$\begin{aligned}
\underset{\sim}{Q}^{(n)}_{K;i,i'} &= \frac{n_i n_{i'}}{n_i+n_{i'}}\left\{\frac{1}{\mathcal{L}_k(K)}\mathrm{tr}^2(\underset{\sim}{\mathbf{S}}_{K;i}-\underset{\sim}{\mathbf{S}}_{K;i'})+\frac{k(k+2)}{2\mathcal{J}_k(K)}\left[\mathrm{tr}\left[(\underset{\sim}{\mathbf{S}}_{K;i}-\underset{\sim}{\mathbf{S}}_{K;i'})^2\right]-\frac{1}{k}\mathrm{tr}^2(\underset{\sim}{\mathbf{S}}_{K;i}-\underset{\sim}{\mathbf{S}}_{K;i'})\right]\right\}\\
&= \frac{n_i n_{i'}}{n_i+n_{i'}}\left\{\frac{k(k+2)}{2\mathcal{J}_k(K)}\mathrm{tr}\left[(\underset{\sim}{\mathbf{S}}_{K;i}-\underset{\sim}{\mathbf{S}}_{K;i'})^2\right]-\frac{k(\mathcal{J}_k(K)-k(k+2))}{2\mathcal{J}_k(K)\mathcal{L}_k(K)}\mathrm{tr}^2(\underset{\sim}{\mathbf{S}}_{K;i}-\underset{\sim}{\mathbf{S}}_{K;i'})\right\}.
\end{aligned}$$

For Gaussian scores (i.e., for $K=K_{\phi_1}$; see Section 2.3), one obtains the *van der Waerden* test statistics

$$\underset{\sim}{Q}^{(n)}_{\mathrm{vdW}}=\frac{1}{n}\sum_{1\le i<i'\le m}(n_i+n_{i'})\underset{\sim}{Q}^{(n)}_{\mathrm{vdW};i,i'},\quad\text{where}\quad \underset{\sim}{Q}^{(n)}_{\mathrm{vdW};i,i'}=\frac{n_i n_{i'}}{2(n_i+n_{i'})}\,\mathrm{tr}\left[(\underset{\sim}{\mathbf{S}}_{\mathrm{vdW};i}-\underset{\sim}{\mathbf{S}}_{\mathrm{vdW};i'})^2\right] \tag{5.4}$$

with $\underset{\sim}{\mathbf{S}}_{\mathrm{vdW};i}:=n_i^{-1}\sum_{j=1}^{n_i}\Psi_k^{-1}(\hat{R}_{ij}/(n+1))\,\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij}$. The Student scores (i.e., $K=K_{f^t_{1,\nu}}$; see Section 2.3) yield

$$\underset{\sim}{Q}^{(n)}_{f^t_{1,\nu}}=\frac{1}{n}\sum_{1\le i<i'\le m}(n_i+n_{i'})\underset{\sim}{Q}^{(n)}_{f^t_{1,\nu};i,i'}, \tag{5.5}$$

where

$$\underset{\sim}{Q}^{(n)}_{f^t_{1,\nu};i,i'}=\frac{n_i n_{i'}}{n_i+n_{i'}}\,\frac{k+\nu+2}{2(k+\nu)}\left\{\mathrm{tr}\left[(\underset{\sim}{\mathbf{S}}_{f^t_{1,\nu};i}-\underset{\sim}{\mathbf{S}}_{f^t_{1,\nu};i'})^2\right]+\frac{1}{\nu}\mathrm{tr}^2(\underset{\sim}{\mathbf{S}}_{f^t_{1,\nu};i}-\underset{\sim}{\mathbf{S}}_{f^t_{1,\nu};i'})\right\}$$

with $\underset{\sim}{\mathbf{S}}_{f^t_{1,\nu};i}:=k(k+\nu)n_i^{-1}\sum_{j=1}^{n_i}G^{-1}_{k,\nu}(\hat{R}_{ij}/(n+1))/[\nu+kG^{-1}_{k,\nu}(\hat{R}_{ij}/(n+1))]\,\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij}$. As for the tests associated with the usual power score functions K_a ($a>0$), they are based on

$$\underset{\sim}{Q}^{(n)}_{K_a}=\frac{1}{n}\sum_{1\le i<i'\le m}(n_i+n_{i'})\,\underset{\sim}{Q}^{(n)}_{K_a;i,i'}, \tag{5.6}$$

where

$$\begin{aligned}
\underset{\sim}{Q}^{(n)}_{K_a;i,i'} &= \frac{n_i n_{i'}}{n_i+n_{i'}}\,\frac{2a+1}{2a^2(a+1)^2k^2}\\
&\quad\times\left\{a^2k(k+2)\mathrm{tr}\left[(\underset{\sim}{\mathbf{S}}_{K_a;i}-\underset{\sim}{\mathbf{S}}_{K_a;i'})^2\right]-(a^2k-4a-2)\,\mathrm{tr}^2(\underset{\sim}{\mathbf{S}}_{K_a;i}-\underset{\sim}{\mathbf{S}}_{K_a;i'})\right\}
\end{aligned}$$

with $\underset{\sim}{\mathbf{S}}_{K_a;i}:=k(a+1)(n+1)^{-a}n_i^{-1}\sum_{j=1}^{n_i}(\hat{R}_{ij})^a\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij}$.

Corollary 5.1 *Assume that the conditions of Theorem 5.1 hold. Then,*

(i) provided that $g_1\in\mathcal{F}_a$ is such that $\mathcal{L}_k(K,g_1)\neq 0\neq\mathcal{J}_k(K,g_1)$, $\underset{\sim}{\phi}^{(n)}_K$ is consistent under any local g_1-alternative (that is, under any $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$, $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$, $\lim_{n\to\infty}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}\notin\mathcal{M}(\boldsymbol{\Upsilon})$);

(ii) the same conclusion holds if $u\mapsto K(u)$ is differentiable with derivative K', and if $g_1\in\mathcal{F}_a$ is such that $\int_0^\infty K'(\tilde{G}_{1k}(r))\,r\,(\tilde{g}_{1k}(r))^2dr>0$ (in particular, if K is non-decreasing).

See the appendix for the proof. This corollary shows that the van der Waerden tests above, as well as those achieving local asymptotic stringency at prespecified Student or power-exponential densities, are universally (locally) consistent (since the corresponding score functions are strictly increasing). Of course, the same holds for the tests associated with the power functions K_a, $a>0$. Non-local consistency results can be obtained along the same lines as in Section 5.2 of Hallin and Paindaveine (2006a).

5.2 The optimal pseudo-Gaussian tests.

In this section, we introduce the pseudo-Gaussian counterpart of the rank-based tests defined in Section 5.1. The parametric Gaussian test $\phi_{\mathcal{N}}^{(n)}$ for $\mathcal{H}_0$ follows, in the same way as (4.3), from classical results on most stringent tests in Gaussian shift experiments. Denoting by $((\boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};\mathcal{N}})', (\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};\mathcal{N}})')' := ((\Delta^{II,1}_{\boldsymbol{\vartheta};\mathcal{N}})', \ldots, (\Delta^{II,m}_{\boldsymbol{\vartheta};\mathcal{N}})', (\boldsymbol{\Delta}^{III,1}_{\boldsymbol{\vartheta};\mathcal{N}})', \ldots, (\boldsymbol{\Delta}^{III,m}_{\boldsymbol{\vartheta};\mathcal{N}})')'$ the scale and shape components of the Gaussian central sequence, where (still with $d_{ij} := d_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$; see Section 2.3 for the definition of a_k)

$$\Delta^{II,i}_{\boldsymbol{\vartheta};\mathcal{N}} = \frac{n_i^{-1/2}}{2\sigma_i^2} \sum_{j=1}^{n_i} \left(a_k \frac{d_{ij}^2}{\sigma_i^2} - k \right) \quad \text{and} \quad \boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\mathcal{N}} = \frac{a_k n_i^{-1/2}}{2\sigma_i^2} \mathbf{M}_k(\mathbf{V}_i) \left(\mathbf{V}_i^{\otimes 2} \right)^{-1/2} \sum_{j=1}^{n_i} d_{ij}^2 \mathrm{vec} \left(\mathbf{U}_{ij} \mathbf{U}'_{ij} \right),$$

$i = 1, \ldots, m$, this parametric Gaussian test is based on a quadratic test statistic of the form

$$Q_{\mathcal{N}}^{(n)} := \left(\boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}} \right)' (\boldsymbol{\Gamma}^{II}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}})^{\perp} \boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}} + \left(\boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}} \right)' (\boldsymbol{\Gamma}^{III}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}})^{\perp} \boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}}, \tag{5.7}$$

computed at a sequence of estimators $\hat{\boldsymbol{\vartheta}}$ satisfying Assumptions (D1)-(D3), where, for any $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ (still with common values σ^2 and $\mathbf{V}$ of the scale and shape parameters), we let

$$(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};\mathcal{N}})^{\perp} := \frac{2\sigma^4}{k} [\mathbf{I}_m - \mathbf{C}^{(n)}] \quad \text{and} \quad (\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};\mathcal{N}})^{\perp} := (k(k+2))^{-1} [\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes (\mathbf{H}_k(\mathbf{V}))^{-1}.$$

Turning this Gaussian test $\phi_{\mathcal{N}}^{(n)}$ into a pseudo-Gaussian one $\phi_{\mathcal{N}*}^{(n)}$ will be possible under the existence of finite fourth-order moments only. This requirement, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, is satisfied iff $E_k(g_1) := \sigma_i^{-4} \mathrm{E}_{\boldsymbol{\vartheta};g_1}[d_{ij}^4(\boldsymbol{\theta}_i, \mathbf{V}_i)] = \int_0^1 (\tilde{G}_{1k}^{-1}(u))^4 du$ is finite, that is, iff

$$g_1 \in \mathcal{F}_a^{(4)} := \left\{ g_1 \in \mathcal{F}_a \, : \, \int_0^{\infty} r^{k+3} g_1(r) \, dr < \infty \right\};$$

for all $g_1 \in \mathcal{F}_a^{(4)}$, let $D_k(g_1) := \sigma_i^{-2} \mathrm{E}_{g_1}[d_{ij}^2(\boldsymbol{\theta}_i, \mathbf{V}_i)] = \int_0^1 (\tilde{G}_{1k}^{-1}(u))^2 du$. For Gaussian densities ($g_1 = \phi_1$), one easily obtains $E_k(\phi_1) = a_k^{-2} k(k+2)$ and $D_k(\phi_1) = a_k^{-1} k$.

The problem with $\phi_{\mathcal{N}}^{(n)}$ under non-Gaussian densities g_1 is that, whereas $\boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\mathcal{N}}$ remains correctly centered under any $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, the expectation of $\boldsymbol{\Delta}^{II,i}_{\boldsymbol{\vartheta};\mathcal{N}}$ under non-Gaussian g_1 is not zero anymore, which induces for $\boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}}$ the same type of shift as a perturbation of the σ_i's. To remedy this, define, for $g_1 \in \mathcal{F}_a^{(4)}$,

$$T^{II,i}_{\boldsymbol{\vartheta};g_1} := \boldsymbol{\Delta}^{II,i}_{\boldsymbol{\vartheta};\mathcal{N}} - \mathrm{E}_{\boldsymbol{\vartheta};g_1} \left[\boldsymbol{\Delta}^{II,i}_{\boldsymbol{\vartheta};\mathcal{N}} \right] = \frac{a_k n_i^{-1/2}}{2\sigma_i^2} \sum_{j=1}^{n_i} \left(\frac{d_{ij}^2}{\sigma_i^2} - D_k(g_1) \right) \quad \text{and} \quad \mathbf{T}^{III,i}_{\boldsymbol{\vartheta};g_1} := \boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\mathcal{N}}, \; i = 1, \ldots, m.$$

Since g_1 in practice remains unspecified, the $T^{II,i}_{\boldsymbol{\vartheta};g_1}$'s cannot be computed from the data; this however will be taken care of later on. The asymptotic distribution, under $\mathcal{H}_0$ and local alternatives, of $\mathbf{T}_{\boldsymbol{\vartheta};g_1} := ((\mathbf{T}^{II}_{\boldsymbol{\vartheta};g_1})', (\mathbf{T}^{III}_{\boldsymbol{\vartheta};g_1})')' := (T^{II,1}_{\boldsymbol{\vartheta};g_1}, \ldots, T^{II,m}_{\boldsymbol{\vartheta};g_1}, (\mathbf{T}^{III,1}_{\boldsymbol{\vartheta};g_1})', \ldots, (\mathbf{T}^{III,m}_{\boldsymbol{\vartheta};g_1})')'$, is provided in the following lemma (see the appendix for the proof).

Lemma 5.1 *Assume that (A) and (B) hold, and that $g_1 \in \mathcal{F}_a^{(4)}$. Fix $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ (with common values σ^2 and $\mathbf{V}$ for the scale and shape parameters, respectively). Then, letting $C_k(g_1) :=$*

$\sigma_i^{-4}\mathrm{Var}_{\boldsymbol{\vartheta};g_1}[d^2_{ij}(\boldsymbol{\theta}_i,\mathbf{V}_i)] = E_k(g_1) - D_k^2(g_1)$, $\mathbf{T}_{\boldsymbol{\vartheta};g_1}$ *is asymptotically normal with mean zero and mean*

$$\begin{pmatrix} \frac{a_k}{2\sigma^4}\, D_k(g_1)\,\boldsymbol{\tau}_{II} \\ \\ a_k(k+2)D_k(g_1)\,[\mathbf{I}_m\otimes \mathbf{H}_k(\mathbf{V})]\,\boldsymbol{\tau}_{III} \end{pmatrix}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ *and* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau};g_1}$, *respectively, and covariance matrix*

$$\boldsymbol{\Gamma}^{\mathcal{N}}_{\boldsymbol{\vartheta};g_1} := \begin{pmatrix} \boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta};g_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1}\end{pmatrix} := \begin{pmatrix} \frac{a_k^2}{4\sigma^4}C_k(g_1)\,\mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & a_k^2E_k(g_1)\,[\mathbf{I}_m\otimes\mathbf{H}_k(\mathbf{V})]\end{pmatrix}$$

under both.

Note that, under Gaussian g_1, $\mathbf{T}_{\boldsymbol{\vartheta};g_1}$ coincides with the Gaussian central sequence $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathcal{N}}$, and $\boldsymbol{\Gamma}^{\mathcal{N}}_{\boldsymbol{\vartheta};g_1}$ with the Gaussian information matrix $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\mathcal{N}}$.

The effect on $\mathbf{T}_{\boldsymbol{\vartheta};g_1}$ of a non-specification of $\boldsymbol{\vartheta}$ is dealt with in a similar way as for the rank-based statistics in Section 5.1. More precisely, the following parametric Gaussian analog of Proposition 4.2 can be established (the proof follows along the same lines and is omitted).

Proposition 5.1 *Assume that (A), (B), (D2), and (D3) hold, and that* $g_1 \in \mathcal{F}_a^{(4)}$. *Fix* $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$ *(with common values* σ^2 *and* $\mathbf{V}$ *of the scale and shape parameters, respectively). Then,*

$$\mathbf{T}^{II}_{\hat{\boldsymbol{\vartheta}};g_1} - \mathbf{T}^{II}_{\boldsymbol{\vartheta};g_1} + \frac{a_k}{2\sigma^4}\,D_k(g_1)\,(\boldsymbol{\nu}^{(n)}_{II})^{-1}n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)}_{II} - \boldsymbol{\vartheta}_{II})$$

and

$$\mathbf{T}^{III}_{\hat{\boldsymbol{\vartheta}};g_1} - \mathbf{T}^{III}_{\boldsymbol{\vartheta};g_1} + a_k(k+2)D_k(g_1)\,[\mathbf{I}_m\otimes\mathbf{H}_k(\mathbf{V})]\,(\boldsymbol{\nu}^{(n)}_{III})^{-1}\,n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)}_{III} - \boldsymbol{\vartheta}_{III})$$

are $o_{\mathrm{P}}(1)$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, *as* $n\to\infty$.

Under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, with $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$ and $g_1\in\mathcal{F}_a^{(4)}$, the pooled regular covariance matrix $\mathbf{S} := \frac{1}{n}\sum_{i=1}^m\sum_{j=1}^{n_i}(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)'$ is a root-n consistent estimator of $k^{-1}D_k(g_1)\sigma^2\mathbf{V}$, and therefore

$$\hat{\boldsymbol{\vartheta}}^{(n)}_{g_1} := (\bar{\mathbf{X}}_1',\ldots,\bar{\mathbf{X}}_m', k(D_k(g_1))^{-1}|\mathbf{S}|^{1/k}\mathbf{1}_m', \mathbf{1}_m'\otimes(\overset{\circ}{\mathrm{vech}}\,(\mathbf{S}/|\mathbf{S}|^{1/k}))')' \tag{5.8}$$

is an estimator satisfying (after due discretization) Assumptions (D1)-(D3), so that Proposition 5.1, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, applies.

Replacing $\boldsymbol{\Delta}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}}$ with $\mathbf{T}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1}$ and $\boldsymbol{\Gamma}_{\hat{\boldsymbol{\vartheta}};\mathcal{N}}$ with $\boldsymbol{\Gamma}^{\mathcal{N}}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1}$ in the Gaussian test statistic (5.7) yields

$$Q^{(n)}_{\mathcal{N},g_1} := \left(\mathbf{T}^{II}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1}\right)'(\boldsymbol{\Gamma}^{\mathcal{N},II}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1})^{\perp}\,\mathbf{T}^{II}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1} + \left(\mathbf{T}^{III}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1}\right)'(\boldsymbol{\Gamma}^{\mathcal{N},III}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1})^{\perp}\,\mathbf{T}^{III}_{\hat{\boldsymbol{\vartheta}}_{g_1};g_1},$$

where (σ^2 and $\mathbf{V}$ still stand for the common null values of the scale and shape parameters under $\boldsymbol{\vartheta}$)

$$(\boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta};g_1})^{\perp} = \frac{4\sigma^4}{a_k^2C_k(g_1)}[\mathbf{I}_m - \mathbf{C}^{(n)}] \quad\text{and}\quad (\boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1})^{\perp} = (a_k^2E_k(g_1))^{-1}[\mathbf{I}_m-\mathbf{C}^{(n)}]\otimes(\mathbf{H}_k(\mathbf{V}))^{-1}.$$

Writing $\hat{d}_{ij}$, $\hat{\mathbf{U}}_{ij}$, and $\hat{\sigma}^2_{g_1}$ for the quantities

$$\hat{d}_{ij} := d_{ij}(\bar{\mathbf{X}}_i, \mathbf{S}/|\mathbf{S}|^{1/k}),\ \hat{\mathbf{U}}_{ij} := \mathbf{U}_{ij}(\bar{\mathbf{X}}_i,\mathbf{S}/|\mathbf{S}|^{1/k}),\ \text{and}\ \hat{\sigma}^2_{g_1} := k(D_k(g_1))^{-1}|\mathbf{S}|^{1/k}$$

computed from $\hat{\boldsymbol{\vartheta}}_{g_1}$ given in (5.8), $Q^{(n)}_{\mathcal{N},g_1}$ can be reformulated as

$$Q^{(n)}_{\mathcal{N},g_1} = \sum_{i,i'=1}^{m} \frac{\delta_{i,i'} - (\lambda_i^{(n)}\lambda_j^{(n)})^{1/2}}{(n_i n_{i'})^{1/2}} \sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}} \left\{ \frac{1}{C_k(g_1)} \left(\frac{\hat{d}^2_{ij}}{\hat{\sigma}^2_{g_1}} - D_k(g_1)\right)\left(\frac{\hat{d}^2_{i'j'}}{\hat{\sigma}^2_{g_1}} - D_k(g_1)\right) + \frac{k(k+2)}{2E_k(g_1)} \frac{\hat{d}^2_{ij}\hat{d}^2_{i'j'}}{\hat{\sigma}^4_{g_1}} \left((\hat{\mathbf{U}}'_{ij}\hat{\mathbf{U}}_{i'j'})^2 - \frac{1}{k}\right)\right\},$$

or

$$Q^{(n)}_{\mathcal{N},g_1} = \frac{1}{n} \sum_{1\leq i<i'\leq m} (n_i + n_{i'}) Q^{(n)}_{\mathcal{N},g_1;i,i'}, \tag{5.9}$$

with

$$Q^{(n)}_{\mathcal{N},g_1;i,i'} = \frac{n_i n_{i'}}{(n_i+n_{i'})} \frac{1}{\hat{\sigma}^4_{g_1}} \left\{ \frac{1}{C_k(g_1)} \left[\left(\frac{1}{n_i} \sum_{j=1}^{n_i} \hat{d}^2_{ij}\right) - \left(\frac{1}{n_{i'}} \sum_{j'=1}^{n_{i'}} \hat{d}^2_{i'j'}\right)\right]^2 \right.$$

$$\left. + \frac{k(k+2)}{2E_k(g_1)} \left\| \left[\frac{1}{n_i} \sum_{j=1}^{n_i} \hat{d}^2_{ij} \mathrm{vec}\left(\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij} - \frac{1}{k}\mathbf{I}_k\right)\right] - \left[\frac{1}{n_{i'}} \sum_{j'=1}^{n_{i'}} \hat{d}^2_{i'j'} \mathrm{vec}\left(\hat{\mathbf{U}}_{i'j'}\hat{\mathbf{U}}'_{i'j'} - \frac{1}{k}\mathbf{I}_k\right)\right] \right\|^2 \right\}.$$

In terms of the intragroup covariance matrices $\mathbf{S}_i := \frac{1}{n_i}\sum_{j=1}^{n_i}(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)'$, this can still be written as

$$Q^{(n)}_{\mathcal{N},g_1;i,i'} = \frac{n_i n_{i'}}{(n_i+n_{i'})} \frac{D_k^2(g_1)}{k^2} \left\{ \frac{1}{C_k(g_1)} \mathrm{tr}^2\left(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right) + \frac{k(k+2)}{2E_k(g_1)} \left[\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\right] - \frac{1}{k} \mathrm{tr}^2\left(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right)\right]\right\}$$

$$= \frac{n_i n_{i'}}{n_i+n_{i'}} \frac{1}{2(1+\kappa_k(g_1))} \left\{ \mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\right] - \frac{\kappa_k(g_1)}{(k+2)\kappa_k(g_1)+2} \mathrm{tr}^2\left(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right)\right\},$$

where $\kappa_k(g_1) := kE_k(g_1)/((k+2)D_k^2(g_1)) - 1$ is the kurtosis common to the m elliptic populations under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ (see, e.g., page 54 of Anderson 2003). Note that, at the multinormal case ($g_1 = \phi_1$), this reduces to

$$Q^{(n)}_{\mathcal{N},\phi_1} = \frac{1}{n} \sum_{1\leq i<i'\leq m} (n_i+n_{i'}) Q^{(n)}_{\mathcal{N},\phi_1;i,i'}, \quad \text{with} \quad Q^{(n)}_{\mathcal{N},\phi_1:i,i'} = \frac{n_i n_{i'}}{2(n_i+n_{i'})} \mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\right], \tag{5.10}$$

which coincides with $Q^{(n)}_{\mathcal{N}}$ in (5.7) provided that $\hat{\boldsymbol{\vartheta}} := \hat{\boldsymbol{\vartheta}}_{g_1}$ (see (5.8)).

Clearly, in order to obtain a genuine test statistic $Q^{(n)}_{\mathcal{N}*}$ (that is, a statistic that does not depend on g_1 anymore) which nevertheless, under any $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ (with $g_1 \in \mathcal{F}_a^{(4)}$), is asymptotically equivalent to $Q^{(n)}_{\mathcal{N},g_1}$, it is sufficient to replace $\kappa_k(g_1)$ with a consistent (still under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $g_1 \in \mathcal{F}_a^{(4)}$) estimator. An obvious choice is

$$\hat{\kappa}_k := \frac{k}{(k+2)} \frac{\left(n^{-1}\sum_{i=1}^m \sum_{j=1}^{n_i} \hat{d}^4_{ij}\right)}{\left(n^{-1}\sum_{i=1}^m \sum_{j=1}^{n_i} \hat{d}^2_{ij}\right)^2} - 1.$$

The resulting pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ then rejects the null hypothesis (at asymptotic level α) as soon as

$$Q^{(n)}_{\mathcal{N}*} = \frac{1}{n} \sum_{1\leq i<i'\leq m} (n_i + n_{i'})\, Q^{(n)}_{\mathcal{N}*;i,i'} > \chi^2_{(m-1)(K+1);1-\alpha} \tag{5.11}$$

where, with $\mathbf{S}$ and $\mathbf{S}_i$, $i = 1, \ldots, m$ as above,

$$Q^{(n)}_{\mathcal{N}*;i,i'} := \frac{n_i n_{i'}}{n_i + n_{i'}} \frac{1}{2(1+\hat{\kappa}_k)} \Big\{ \mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2 \right] - \frac{\hat{\kappa}_k}{(k+2)\hat{\kappa}_k + 2} \mathrm{tr}^2\left(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right) \Big\}.$$

This test statistic is clearly affine-invariant; the following theorem summarizes its asymptotic properties (see the appendix for the proof).

Theorem 5.2 *Assume that (A) and (B) hold. Then,*

(i) $Q^{(n)}_{\mathcal{N}*}$ *is asymptotically chi-square with* $(m-1)(K+1)$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})} \bigcup_{g_1\in\mathcal{F}_a^{(4)}} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$*, and (provided that (B) is reinforced into (B′)) asymptotically noncentral chi-square, still with* $(m-1)(K+1)$ *degrees of freedom but with noncentrality parameter*

$$\frac{k}{(k+2)\kappa_k(g_1)+2}\, r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} + \frac{1}{2(1+\kappa_k(g_1))}\, r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} \tag{5.12}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$*, with* $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$*,* $\boldsymbol{\nu}\boldsymbol{\tau} := \lim_{n\to\infty} \boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)} \notin \mathcal{M}(\boldsymbol{\Upsilon})$*), and* $g_1 \in \mathcal{F}_a^{(4)}$*;*

(ii) the sequence of tests $\phi^{(n)}_{\mathcal{N}*}$ *has asymptotic level* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})} \bigcup_{g_1\in\mathcal{F}_a^{(4)}} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$*;*

(iii) the pseudo-Gaussian tests $\phi^{(n)}_{\mathcal{N}*}$ *are asymptotically equivalent, under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi_1}\}$ *and under contiguous alternatives, to the optimal parametric Gaussian tests* $\phi^{(n)}_{\mathcal{N}}$ *based on (5.7); hence, the sequence* $\phi^{(n)}_{\mathcal{N}*}$ *is locally and asymptotically most stringent, still at asymptotic level* α*, for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})} \bigcup_{g_1\in\mathcal{F}_a^{(4)}} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{M}(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi_1}\}$*.*

We conclude this section by showing that the pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ is essentially the robustified Schott test based on (1.5). Indeed, since $\sum_{i=1}^m \dot{\lambda}_i^{(n)}\, \mathrm{tr}\left[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2\right] = \sum_{i,i'=1}^m \dot{\lambda}_i^{(n)}\dot{\lambda}_{i'}^{(n)}\, \mathrm{tr}\left[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2\right]$, we obtain

$$Q^{(n)}_{\mathrm{Schott}} = \frac{\dot{n}}{2} \sum_{1\leq i\neq i'\leq m} \dot{\lambda}_i^{(n)}\dot{\lambda}_{i'}^{(n)} \Big[\mathrm{tr}\big[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2\big] - \mathrm{tr}(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1})\Big] = \frac{\dot{n}}{2} \sum_{1\leq i<i'\leq m} \dot{\lambda}_i^{(n)}\dot{\lambda}_{i'}^{(n)}\, \mathrm{tr}\left[(\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_{i'}))^2\right].$$

Working along exactly the same lines yields

$$\frac{\dot{n}}{2}\Big[\sum_{i=1}^m \dot{\lambda}_i^{(n)} \mathrm{tr}^2(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}) - \sum_{i,i'=1}^m \dot{\lambda}_i^{(n)}\dot{\lambda}_{i'}^{(n)} \mathrm{tr}(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})\mathrm{tr}(\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1})\Big] = \frac{\dot{n}}{2} \sum_{1\leq i<i'\leq m} \dot{\lambda}_i^{(n)}\dot{\lambda}_{i'}^{(n)} \mathrm{tr}^2(\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_{i'})).$$

Hence, $Q^{(n)}_{\mathrm{Schott}*} = \frac{1}{\dot{n}} \sum_{1\leq i<i'\leq m} (\dot{n}_i + \dot{n}_{i'})\, Q^{(n)}_{\mathrm{Schott};i,i'}$, where

$$Q^{(n)}_{\mathrm{Schott};i,i'} := \frac{\dot{n}_i\dot{n}_{i'}}{\dot{n}_i + \dot{n}_{i'}} \frac{1}{2(1+\hat{\kappa}_k)} \Big\{ \mathrm{tr}\left[(\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_{i'}))^2\right] - \frac{\hat{\kappa}_k}{(k+2)\hat{\kappa}_k + 2} \mathrm{tr}^2\left(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right) \Big\}.$$

Comparing with (5.11), this clearly relates our pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ to Schott's test $\phi^{(n)}_{\text{Schott}*}$ based on $Q^{(n)}_{\text{Schott}*}$ in the same way the LRT is related to the MLRT, that is, by replacing n_i, n, $\mathbf{S}_i$, and $\mathbf{S}$ by $\dot{n}_i$, $\dot{n}$, $\dot{\mathbf{S}}_i$, and $\dot{\mathbf{S}}$, respectively—a replacement which obviously has no impact on asymptotics. Therefore, Theorem 5.2 also holds for Schott's test; this not only establishes the exact optimality properties of the latter, but also provides its local powers (such results do not follow from Schott's original derivation).

5.3 Links with existing Gaussian tests.

In the previous section, we have derived optimal Gaussian and pseudo-Gaussian tests, and shown that our pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ essentially coincides with Schott's modified test $\phi^{(n)}_{\text{Schott}*}$. The relation to other Gaussian tests in the literature is less obvious; in this section, we investigate the behavior under non-Gaussian elliptical densities of the LRT/MLRT tests based on $-2\log\Lambda$ and $-2\log\Lambda^\star$ (see (1.1) and (1.2), respectively), the Schott unrobustified test based on $Q^{(n)}_{\text{Schott}}$ (see (1.4)), the Gaussian ("non-pseudo" or "unrobustified") most stringent test $\phi^{(n)}_{\mathcal{N}}$ based on $Q^{(n)}_{\mathcal{N}} = Q^{(n)}_{\mathcal{N},\phi_1}$ (see (5.10)), and the Nagao (1973) test—another popular Gaussian method. The latter is based on a result by Sugiura (1969) stating that, under Gaussian assumptions, as $n \to \infty$,

$$n^{-1/2}\Big(-2\log\Lambda^* + 2\log\frac{\prod_{i=1}^m |\mathbf{\Sigma}_{0i}|^{n_i/2}}{|\mathbf{\Sigma}_0|^{n/2}}\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, 2\sum_{i=1}^m \lambda_i \operatorname{tr}\left[(\mathbf{\Sigma}_{0i}\mathbf{\Sigma}_0^{-1} - \mathbf{I}_k)^2\right]\Big) \tag{5.13}$$

where $\mathbf{\Sigma}_{0i} := \text{Var}[\mathbf{X}_{ij}]$, $i = 1, \ldots, m$ and $\mathbf{\Sigma}_0 := \sum_{i=1}^m \lambda_i \mathbf{\Sigma}_{0i}$ (throughout this section, the quantities $\lambda_i^{(n)}$ and λ_i are as in Assumption (B′); the notation $\dot{\mathbf{S}}_i$ and $\dot{\mathbf{S}}$ is used as in Section 5.2). The Nagao (1973) test then rejects the null hypothesis for large values of

$$Q^{(n)}_{\text{Nagao}} := \frac{1}{2}\sum_{i=1}^m \dot{n}_i \operatorname{tr}\Big[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1} - \mathbf{I}_k)^2\Big].$$

The following result establishes the asymptotic equivalence, under Gaussian assumptions, of all these statistics with the optimal Gaussian statistic (which entails their optimality in the Le Cam sense at the multinormal), and explains why none of them qualifies as a pseudo-Gaussian procedure. The proof is given in the appendix; Part (ii) actually is a direct consequence of Part (i) and a more general result by Yanagihara et al. (2005).

Proposition 5.2 *(i) Under any null distribution with finite fourth-order moments, $-2\log\Lambda$, $-2\log\Lambda^\star$, $Q^{(n)}_{\text{Schott}}$, and $Q^{(n)}_{\text{Nagao}}$ all are asymptotically equivalent to the Gaussian most stringent test statistics $Q^{(n)}_{\mathcal{N},\phi_1}$ (hence inherit the optimality properties of the latter). (ii) For any $g_1 \in \mathcal{F}_a^{(4)}$, the asymptotic distribution, under $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\mathbf{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}\}$, of these test statistics is that of*

$$(1+\kappa_k)\left\{\left[1 + \frac{k\kappa_k(g_1)}{2(1+\kappa_k(g_1))}\right]Y_1 + Y_2\right\}, \tag{5.14}$$

where Y_1 and Y_2 are independent chi-square random variables, with $m-1$ and $(m-1)K$ degrees of freedom, respectively.

Clearly, (5.14) does not yield a chi-square distribution unless $\kappa_k(g_1) = 0$, that is, when g_1 has Gaussian kurtosis.

6 Asymptotic relative efficiencies.

The asymptotic relative efficiencies of the rank-based tests $\underset{\sim}{\phi}_K^{(n)}$ with respect to their Gaussian counterparts $\phi_{\mathcal{N}*}^{(n)}$ (equivalently, with respect to Schott's tests based on $Q_{\text{Schott}*}^{(n)}$) directly follow as the ratios of the noncentrality parameters in the asymptotic distributions of the various test statistics under local alternatives (see Theorems 5.1 and 5.2).

Proposition 6.1 *Assume that (A), (B′), (C), and (D) hold, and that $g_1 \in \mathcal{F}_a^{(4)}$. Then, the asymptotic relative efficiency of $\underset{\sim}{\phi}_K^{(n)}$ with respect to the pseudo-Gaussian test $\phi_{\mathcal{N}*}^{(n)}$, when testing $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$ against $\mathrm{P}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}^{(n)}$ (with $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ and $\boldsymbol{\nu\tau} := \lim_{n\to\infty} \boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)} \notin \mathcal{M}(\boldsymbol{\Upsilon})$), is*

$$\mathrm{ARE}_{\boldsymbol{\vartheta},\boldsymbol{\tau},k,g_1}(\underset{\sim}{\phi}_K^{(n)}/\phi_{\mathcal{N}*}^{(n)}) = (1-\xi)\,\mathrm{ARE}_{k,g_1}^{(\text{scale})}(\underset{\sim}{\phi}_K^{(n)}/\phi_{\mathcal{N}*}^{(n)}) + \xi\,\mathrm{ARE}_{k,g_1}^{(\text{shape})}(\underset{\sim}{\phi}_K^{(n)}/\phi_{\mathcal{N}*}^{(n)}), \tag{6.1}$$

where

$$\mathrm{ARE}_{k,g_1}^{(\text{scale})}(\underset{\sim}{\phi}_K^{(n)}/\phi_{\mathcal{N}*}^{(n)}) := \frac{((k+2)\kappa_k(g_1)+2)\mathcal{L}_k^2(K,g_1)}{4k\mathcal{L}_k(K)}, \tag{6.2}$$

$$\mathrm{ARE}_{k,g_1}^{(\text{shape})}(\underset{\sim}{\phi}_K^{(n)}/\phi_{\mathcal{N}*}^{(n)}) := \frac{(1+\kappa_k(g_1))\mathcal{J}_k^2(K,g_1)}{k(k+2)\mathcal{J}_k(K)}, \tag{6.3}$$

and $\xi := \xi_{\boldsymbol{\vartheta},\boldsymbol{\tau},k,g_1} := ((k+2)\kappa_k(g_1)+2)r_{\boldsymbol{\vartheta},\boldsymbol{\tau}}^{III}/[2k(1+\kappa_k(g_1))r_{\boldsymbol{\vartheta},\boldsymbol{\tau}}^{II} + ((k+2)\kappa_k(g_1)+2)r_{\boldsymbol{\vartheta},\boldsymbol{\tau}}^{III}] \in [0,1]$.

The "shape AREs" in (6.3) do coincide with those obtained in problems involving shape only—for instance, testing null hypotheses of the form $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ for some fixed $\mathbf{V}_0$ (see Hallin and Paindaveine 2006a). Proposition 6.1 shows that the AREs, with respect to the pseudo-Gaussian tests of Section 5.2, of the rank tests proposed in Section 5.1 are convex linear combinations of these "shape AREs" and the "scale AREs" in (6.2).

Numerical values of (6.2) and (6.3), for various values of the space dimension k and various radial densities (Student, Gaussian, and power-exponential), are given in Table 1 for the van der Waerden test $\underset{\sim}{\phi}_{\text{vdW}}^{(n)}$, the Wilcoxon test $\underset{\sim}{\phi}_{K_1}^{(n)}$, and the Spearman test $\underset{\sim}{\phi}_{K_2}^{(n)}$ (the score functions K_a, $a > 0$ were defined in Section 2.3). These ARE values are uniformly large (with the exception, possibly, of univariate scale Wilcoxon AREs), particularly so under heavy tails, as often in rank-based inference.

Also note that the AREs of the proposed van der Waerden tests with respect to the parametric Gaussian tests are larger than or equal to one for all distributions considered in Table 1. For pure shape alternatives, Paindaveine (2006a) has shown that a Chernoff-Savage property holds, that is, $\inf_{g_1} \mathrm{ARE}_{k,g_1}^{(\text{shape})}(\underset{\sim}{\phi}_{\text{vdW}}^{(n)}/\phi_{\mathcal{N}*}^{(n)}) = 1$.One may wonder whether this uniform dominance property of van der Waerden tests extends to the present situation. Although it does for all usual distributions, including all Student and power-exponential ones, the general answer unfortunately is negative; see Section 4 of Paindaveine 2006a for a (pathological) counterexample.

7 Simulations.

We conducted two simulations, one for pure scale alternatives and another one for pure shape alternatives, both in dimension $k = 2$. More precisely, starting from two sets of i.i.d. bivariate random vectors $\boldsymbol{\varepsilon}_{1j}$ $(j = 1, \ldots, n_1 = 100)$ and $\boldsymbol{\varepsilon}_{2j}$ $(j = 1, \ldots, n_2 = 100)$ with spherical densities (the standard bivariate normal and bivariate t-distributions with .5, 2, and 5 degrees of freedom, respectively) centered at $\mathbf{0}$, we considered independent samples obtained from

$$\mathbf{X}_{1j} = \mathbf{A}_1\boldsymbol{\varepsilon}_{1j} + \boldsymbol{\theta}_1, \quad j = 1, \ldots, n_1, \quad \text{and} \quad \mathbf{X}_{2j} = \mathbf{A}_{2,m}\boldsymbol{\varepsilon}_{2j} + \boldsymbol{\theta}_2, \quad j = 1, \ldots, n_2,$$

where $\mathbf{A}_{2,m}\mathbf{A}'_{2,m} = (1 + ms^2)(\mathbf{A}_1\mathbf{A}'_1 + m\mathbf{v})$ ($\mathbf{v}$ a symmetric $(k \times k)$ matrix with $\mathrm{tr}(\mathbf{v}) = 0$), $m = 0, 1, 2, 3$. The values of m allow to produce distributions under the null ($m = 0$) and increasingly heterogeneous alternatives ($m = 1, 2, 3$); all tests being affine-invariant, there is no loss of generality in letting $\mathbf{A}_1 = \mathbf{I}_2$ and $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 = \mathbf{0}$.

In the first simulation (pure scale alternatives), we generated $N = 2,500$ independent samples, with $\mathbf{v} = \mathbf{0}$ and $s^2 = .30$, .44, .56, and 1.50 under Gaussian, t_5, t_2, and $t_{0.5}$ alternatives, respectively; these values of s^2 have been chosen in order to obtain rejection probabilities of the same order under those various densities. In the second simulation (pure shape alternatives), we similarly generated $N = 2,500$ independent samples, with $s^2 = 0$ and $\overset{\circ}{\mathrm{vech}}\,\mathbf{v} = (0, .18)'$, $(0, .20)'$, $(0, .21)'$, and $(0, .22)'$ under Gaussian, t_5, t_2, and $t_{0.5}$ alternatives, respectively, still with the same objective of obtaining comparable powers under the various densities considered.

In each of these samples, we performed the following nine tests (all at asymptotic level $\alpha = 5\%$): (a) the Gaussian LRT ϕ_{LRT} based on (1.1); (b) its modified version ϕ_{MLRT} based on (1.2); (c) the parametric Gaussian test $\phi_{\mathcal{N}}$ based on (5.10) (equivalently, Schott's original test ϕ_{Schott}, based on (1.4)); (d) its pseudo-Gaussian version $\phi_{\mathcal{N}*}$, based on (5.11) (equivalently, the robustified Schott test $\phi_{\mathrm{Schott}*}$, based on (1.5)); (e) the van der Waerden test $\underset{\sim}{\phi}_{\mathrm{vdW}}$ (based on (5.4)); (f)-(h) t_ν-score tests $\underset{\sim}{\phi}_{f^t_{1,\nu}}$ with $\nu = 5$, 2, and .5 (based on (5.5)), as well as (i) the Spearman test (based on $\underset{\sim}{Q}_{K_2}$ in (5.6)). It can be checked that the Wilcoxon test $\underset{\sim}{Q}_{K_1}$, in this bivariate case, coincides with $\underset{\sim}{\phi}_{f^t_{1,2}}$.

Rejection frequencies are reported in Table 2 for pure scale alternatives, and in Table 3 for pure shape alternatives. The corresponding individual confidence intervals (for $N = 2,500$ replications), at confidence level .95, have half-widths .0044, .0080, and .0100, for frequencies of the order of .05 (.95), .20 (.80), and .50, respectively.

A glance at Tables 2 and 3 indicates that the rank tests, when based on their asymptotic chi-square critical values, are conservative and significantly biased at this moderate sample size (100 observations in each group). In order to remedy this, we also implemented bias-corrected versions of each of the rank procedures, by estimating the (distribution-free) quantile of the test statistic under known location θ and known common null value of the shape. These quantiles, just as the asymptotic chi-square quantile, are consistent approximations of the corresponding exact quantiles under the null. They were obtained, for each of the five rank tests under consideration in (e)-(i) above, as the empirical 0.05-upper quantiles $q_{.95}$ of the corresponding rank-based test statistics in a collection of 10^5 simulated multinormal samples, yielding $q_{.95} = 7.2117$, 7.6351, 7.7473, 7.7636, and 7.6773, respectively. These bias-corrected critical values all are smaller than the corresponding asymptotic chi-square one (three degrees of freedom) $\chi^2_{3;.95} = 7.8147$. The resulting tests therefore are uniformly less conservative than the original ones. The corresponding rejection frequencies are given in parentheses in Tables 2 and 3.

Inspection of Tables 2 and 3 confirms the fact that the parametric Gaussian tests $\phi_{\mathcal{N}}$, contrary to the pseudo-Gaussian ones $\phi_{\mathcal{N}*}$, are invalid under non-Gaussian densities (culminating, under $t_{0.5}$, with a size of .9992). However, even the pseudo-Gaussian tests $\phi_{\mathcal{N}*}$, though resisting non-Gaussian densities with finite fourth-order moments, are collapsing under the heavy-tailed $t_{0.5}$ and t_2 distributions (with power less than 10^{-4} under $t_{0.5}$). In sharp contrast with this, all rank-based tests appear to satisfy the 5% probability level constraint. They are conservative in their original versions (particularly so for van der Waerden scores), but seem to be reasonably unbiased (for sample sizes $n_1 = n_2 = 100$) after bias-correction: maximal bias-corrected size we obtain is 0.0616 for $\underset{\sim}{\phi}_{f^t_{1,.5}}$ under the very heavy tailed $t_{0.5}$). Empirical power rankings are essentially consistent with the corresponding ARE values; in order to allow for meaningful comparisons also under infinite fourth-order moments,we also provide AREs with respect to the van der Waerden rank test.

A Appendix.

A.1 Proofs of Lemma 4.1, Theorem 5.1, and Corollary 5.1.

Proof of Lemma 4.1. (i) Fix $r \in \{1, \ldots, m\}$. Clearly, $\underset{\sim}{\Delta}^{II,r}_{\boldsymbol{\vartheta};K} = \Delta^{II,r}_{\boldsymbol{\vartheta};K;g_1} + o_{L^2}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ iff

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} c^{(n)}_{ij;r}\, K\Big(\frac{R_{ij}}{n+1}\Big) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} c^{(n)}_{ij;r}\, K\left(\tilde{G}_{1k}\left(\frac{d_{ij}}{\sigma}\right)\right) + o_{L^2}(1), \tag{A.1}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, where $c^{(n)}_{ij;r} := n_i^{-1/2}\delta_{i,r}$. For $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, the R_{ij}'s are the ranks of the d_{ij}/σ's, which under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ are i.i.d. with distribution function $\tilde{G}_{1k}$. The asymptotic equivalence (A.1) thus follows from Hájek's classical projection result for linear rank statistics (see, e.g., Puri and Sen 1985, Chapter 2), since (a) the $c^{(n)}_{ij;r}$'s are not all equal and (b) the Noether condition

$$\frac{\max_{i,j}\,(c^{(n)}_{ij;r} - n^{-1}\sum_{i,j} c^{(n)}_{ij;r})^2}{\sum_{i,j}\,(c^{(n)}_{ij;r} - n^{-1}\sum_{i,j} c^{(n)}_{ij;r})^2} = n^{-1}\max\Big(\frac{1-\lambda^{(n)}_r}{\lambda^{(n)}_r}, \frac{\lambda^{(n)}_r}{1-\lambda^{(n)}_r}\Big) = o(1), \text{ as } n \to \infty$$

holds (see the comments after Assumption (B)).

Similarly, for the shape part, $\underset{\sim}{\boldsymbol{\Delta}}^{III,r}_{\boldsymbol{\vartheta};K} = \boldsymbol{\Delta}^{III,r}_{\boldsymbol{\vartheta};K;g_1} + o_{L^2}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ iff

$$n_r^{-1/2}\mathbf{M}_k(\mathbf{V})\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\sum_{j=1}^{n_r}\left[K\Big(\frac{R_{rj}}{n+1}\Big) - K\left(\tilde{G}_{1k}\left(\frac{d_{rj}}{\sigma}\right)\right)\right]\mathbf{J}_k^{\perp}\,\mathrm{vec}\left(\mathbf{U}_{rj}\mathbf{U}'_{rj}\right) = o_{L^2}(1)$$

(where $\mathbf{J}_k^{\perp} := \mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k$ satisfies $\mathbf{M}_k(\mathbf{V})(\mathbf{V}^{\otimes 2})^{-1/2}\mathbf{J}_k^{\perp} = \mathbf{M}_k(\mathbf{V})(\mathbf{V}^{\otimes 2})^{-1/2}$ and is such that $\mathbf{J}_k^{\perp}\,\mathrm{vec}(\mathbf{U}_{rj}\mathbf{U}'_{rj})$ is exactly centered), or equivalently iff, for all $\ell \in \{1, 2, \ldots, k^2\}$,

$$T^{(n)}_{r;l} := \sum_{i=1}^{m}\sum_{j=1}^{n_i} c^{(n)}_{ij;r}\left[K\Big(\frac{R_{ij}}{n+1}\Big) - K\left(\tilde{G}_{1k}\left(\frac{d_{ij}}{\sigma}\right)\right)\right]\left[\mathbf{J}_k^{\perp}\,\mathrm{vec}\left(\mathbf{U}_{ij}\mathbf{U}'_{ij}\right)\right]_{\ell} = o_{L^2}(1), \tag{A.2}$$

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

$$\mathrm{E}\left[\left(T^{(n)}_{r;\ell}\right)^2\right] = C_{\ell,k}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(c^{(n)}_{ij;r}\right)^2 \mathrm{E}\left[\left(K\left(\frac{R_i}{n+1}\right) - K\left(\tilde{G}_{1k}\left(\frac{d_i}{\sigma}\right)\right)\right)^2\right]$$

where, denoting by $U_{ij,s}$ the sth component of $\mathbf{U}_{ij}$, $C_{\ell,k} = \mathrm{Var}[U^2_{11,1}] = 2(k-1)/(k^2(k+2))$ for $\ell \in \mathfrak{L}_k := \{mk + m + 1,\ m = 0, 1, \ldots, k-1\}$ and $C_{\ell,k} = \mathrm{Var}\,[U_{11,1}U_{11,2}] = 1/k^2$ for $\ell \notin \mathfrak{L}_k$. Here, the Hájek projection result for linear *signed* rank statistics (see, e.g., Puri and Sen 1985, Chapter 3) yields (A.2), since $\max_{i,j}\,(c^{(n)}_{ij;r})^2/\sum_{i,j}\,(c^{(n)}_{ij;r})^2 = n_r^{-1} = o(1)$, as $n \to \infty$.

As for (ii), the result straightforwardly follows, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ with $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, from the multivariate CLT. The result under local alternatives is obtained as usual, by establishing the joint normality under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ of $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};K;g_1}$ and $\Lambda^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}/\boldsymbol{\vartheta};g_1}$, then applying Le Cam's third Lemma; the required joint normality follows from a routine application of the classical Cramér-Wold device. □

Proof of Theorem 5.1. *(i)* Using successively the continuity of the mapping $\boldsymbol{\vartheta} \mapsto \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};K}$, Proposition 4.2 (jointly with Assumption (D1) and the fact that $[\mathbf{I}_m - \mathbf{C}^{(n)}](\boldsymbol{\Lambda}^{(n)})^{-1}\mathbf{1}_m = \mathbf{0}$), and Lemma 4.1(i), we obtain that

$$
\begin{aligned}
\underset{\sim}{Q}{}^{(n)}_{K} &= \left(\underset{\sim}{\boldsymbol{\Delta}}{}^{II}_{\boldsymbol{\vartheta};K}\right)' (\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{\perp}\, \underset{\sim}{\boldsymbol{\Delta}}{}^{II}_{\boldsymbol{\vartheta};K} + \left(\underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\boldsymbol{\vartheta};K}\right)' (\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{\perp}\, \underset{\sim}{\boldsymbol{\Delta}}{}^{III}_{\boldsymbol{\vartheta};K} + o_{\mathrm{P}}(1) \\
&= \left(\boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};K;g_1}\right)' (\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{\perp}\, \boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};K;g_1} + \left(\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};K;g_1}\right)' (\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{\perp}\, \boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};K;g_1} + o_{\mathrm{P}}(1) \qquad \text{(A.3)}
\end{aligned}
$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ (and therefore, also under the contiguous sequence $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$).

Now, since $(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{1/2}\,(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{\perp}(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{1/2}$ is a symmetric idempotent matrix with rank $m-1$, it follows from Lemma 4.1(ii) that the first term in (A.3) is asymptotically chi-square with $m-1$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and asymptotically noncentral chi-square, still with $m-1$ degrees of freedom, but with noncentrality parameter

$$
\left(\frac{\mathcal{L}_k(K, g_1)}{4\sigma^4}\right)^2 \lim_{n\to\infty} \left\{(\boldsymbol{\tau}^{(n)}_{II})'(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};K})^{\perp}\,\boldsymbol{\tau}^{(n)}_{II}\right\} \qquad \text{(A.4)}
$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$. Evaluation of the limit in (A.4) yields the first term in (5.2).

As for the shape part, using again Lemma 4.1(ii) and the fact that $(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{1/2}\,(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{\perp}(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{1/2}$ is symmetric and idempotent with rank $K(m-1)$, we obtain similarly that the second term in (A.3) is asymptotically chi-square with $K(m-1)$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and asymptotically noncentral chi-square, still with $K(m-1)$ degrees of freedom but with noncentrality parameter

$$
(\mathcal{J}_k(K, g_1))^2 \lim_{n\to\infty} \left\{(\boldsymbol{\tau}^{(n)}_{III})'[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})](\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};K})^{\perp}[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})]\,\boldsymbol{\tau}^{(n)}_{III}\right\} \qquad \text{(A.5)}
$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$. Evaluation of the limit in (A.5) yields the second term in (5.2). As the two terms in (A.3) are asymptotically uncorrelated (see Lemma 4.1(ii) again), they can indeed be treated separately.

(ii) The fact that $\underset{\sim}{\phi}{}^{(n)}_{K}$ has asymptotic level α directly follows from the asymptotic null distribution in part *(i)* and the classical Helly-Bray theorem.

(iii) Optimality is a consequence of the asymptotic equivalence (A.3), under $g_1 = f_1$ satisfying Assumption (A2), of $\underset{\sim}{Q}{}^{(n)}_{f_1}$ and the locally asymptotically optimal test statistic $Q_{\boldsymbol{\Upsilon}}$, as described in (4.3). □

Proof of Corollary 5.1. (i) Fix $g_1 \in \mathcal{F}_a$, with $\mathcal{L}_k(K, g_1) \neq 0 \neq \mathcal{J}_k(K, g_1)$. Clearly, $\underset{\sim}{\phi}{}^{(n)}_{K}$ is consistent under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ iff the corresponding non-centrality parameter in (5.2) is non-zero. Assume the latter is zero. Then, the assumptions on g_1 imply that $s_i^2/\sqrt{\lambda_i} =$

$s_{i'}^2/\sqrt{\lambda_{i'}}$ and

$$\operatorname{tr}\Big[\Big(\mathbf{V}^{-1/2}\Big(\frac{\mathbf{v}_i}{\sqrt{\lambda_i}}-\frac{\mathbf{v}_{i'}}{\sqrt{\lambda_{i'}}}\Big)\mathbf{V}^{-1/2}\Big)^2\Big], \tag{A.6}$$

for all (i,i'). Now, since $\operatorname{tr}(\mathbf{A}^2)=0$ implies that $\mathbf{A}=\mathbf{0}$ for any symmetric $k\times k$ matrix $\mathbf{A}$, it follows from (A.6) that $\mathbf{v}_i/\sqrt{\lambda_i}=\mathbf{v}_{i'}/\sqrt{\lambda_{i'}}$ for all (i,i'). This is possible only for $\boldsymbol{\nu\tau}\in\mathcal{M}(\boldsymbol{\Upsilon})$, which establishes the result.

(ii) Going back to the definition of $g_1\mapsto\mathcal{J}_k(K,g_1)$, we have

$$\mathcal{J}_k(K,g_1)=\int_0^\infty K(\tilde{G}_{1k}(r))\,r\,\varphi_{g_1}(r)\,\tilde{g}_{1k}(r)\,dr=\frac{1}{\mu_{k-1;g_1}}\int_0^\infty K(\tilde{G}_{1k}(r))\,(-\dot{g}_1(r))\,r^k\,dr.$$

Integrating by parts yields

$$\mathcal{J}_k(K,g_1)=\int_0^\infty\Big[kK(\tilde{G}_{1k}(r))+K'(\tilde{G}_{1k}(r))r\tilde{g}_{1k}(r)\Big]\tilde{g}_{1k}(r)\,dr=k^2+\int_0^\infty K'(\tilde{G}_{1k}(r))r(\tilde{g}_{1k}(r))^2\,dr,$$

so that $\int_0^\infty K'(\tilde{G}_{1k}(r))\,r\,(\tilde{g}_{1k}(r))^2\,dr>0$ guarantees that $\mathcal{L}_k(K,g_1)=\mathcal{J}_k(K,g_1)-k^2>0$. Part (i) of the corollary therefore yields the result. □

A.2 Proof of Proposition 4.2.

Consider an arbitrary value $\boldsymbol{\vartheta}=(\boldsymbol{\vartheta}_{I}',\boldsymbol{\vartheta}_{II}',\boldsymbol{\vartheta}_{III}')'=(\boldsymbol{\theta}_1',\ldots,\boldsymbol{\theta}_m',\sigma^2\mathbf{1}_m',\mathbf{1}_m'\otimes(\overset{\circ}{\operatorname{vech}}\,\mathbf{V})')'\in\mathcal{M}(\boldsymbol{\Upsilon})$ of the parameter and a (bounded) sequence of corresponding local perturbations $\boldsymbol{\vartheta}^{(n)}:=\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}$, where

$$\boldsymbol{\tau}^{(n)}=(\boldsymbol{\tau}_{I}^{(n)\prime},\boldsymbol{\tau}_{II}^{(n)\prime},\boldsymbol{\tau}_{III}^{(n)\prime})'=(\mathbf{t}_1^{(n)\prime},\ldots,\mathbf{t}_m^{(n)\prime},s_1^{2(n)},\ldots,s_m^{2(n)},(\overset{\circ}{\operatorname{vech}}\,\mathbf{v}_1^{(n)})',\ldots,(\overset{\circ}{\operatorname{vech}}\,\mathbf{v}_m^{(n)})')'$$

is such that $\boldsymbol{\vartheta}^{(n)}\in\mathcal{M}(\boldsymbol{\Upsilon})$ for all n. To prove Proposition 4.2, it is sufficient to show that, under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$ (where g_1 is as in Proposition 4.2),

$$\underset{\sim}{\boldsymbol{\Delta}}{}_{\boldsymbol{\vartheta}^{(n)};K}^{II}-\underset{\sim}{\boldsymbol{\Delta}}{}_{\boldsymbol{\vartheta};K}^{II}+\frac{\mathcal{L}_k(K,g_1)}{4\sigma^4}\,\boldsymbol{\tau}_{II}^{(n)} \tag{A.7}$$

and

$$\underset{\sim}{\boldsymbol{\Delta}}{}_{\boldsymbol{\vartheta}^{(n)};K}^{III}-\underset{\sim}{\boldsymbol{\Delta}}{}_{\boldsymbol{\vartheta};K}^{III}+\mathcal{J}_k(K,g_1)\,[\mathbf{I}_m\otimes\mathbf{H}_k(\mathbf{V})]\,\boldsymbol{\tau}_{III}^{(n)} \tag{A.8}$$

are $o_{\mathrm{P}}(1)$ as $n\to\infty$, since the local discreteness of $\hat{\boldsymbol{\vartheta}}$ (see, e.g., Kreiss 1987, Lemma 4.4) allows to replace the nonrandom quantity $\boldsymbol{\vartheta}^{(n)}$ with the random one $\hat{\boldsymbol{\vartheta}}$ in (A.7) and (A.8) above. Note that the constraintness of $\hat{\boldsymbol{\vartheta}}$ indeed allows us to restrict to local perturbations $\boldsymbol{\vartheta}^{(n)}\in\mathcal{M}(\boldsymbol{\Upsilon})$. Looking at block i ($i\in\{1,\ldots,m\}$), this implies that Proposition 4.2 is a corollary of the following result.

Proposition A.1 Assume that (A), (B), and (C) hold, and that $g_1\in\mathcal{F}_a$. Fix $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$ and a sequence $\boldsymbol{\vartheta}^{(n)}\in\mathcal{M}(\boldsymbol{\Upsilon})$ as above. Then, for all $i=1,\ldots,m$,

$$\underset{\sim}{\Delta}{}_{\boldsymbol{\vartheta}^{(n)};K}^{II,i}-\underset{\sim}{\Delta}{}_{\boldsymbol{\vartheta};K}^{II,i}+\frac{\mathcal{L}_k(K,g_1)}{4\sigma^4}\,s_i^{2(n)} \tag{A.9}$$

and

$$\underset{\sim}{\boldsymbol{\Delta}}{}_{\boldsymbol{\vartheta}^{(n)};K}^{III,i}-\underset{\sim}{\boldsymbol{\Delta}}{}_{\boldsymbol{\vartheta};K}^{III,i}+\mathcal{J}_k(K,g_1)\,\mathbf{H}_k(\mathbf{V})\,(\overset{\circ}{\operatorname{vech}}\,\mathbf{v}_i^{(n)}) \tag{A.10}$$

are $o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, as $n\to\infty$.

Proof of Proposition A.1. In this proof, we let $\boldsymbol{\theta}_i^n := \boldsymbol{\theta}_i + n_i^{-1/2}\mathbf{t}_i^{(n)}$, $\mathbf{V}^n := \mathbf{V} + n_i^{-1/2}\mathbf{v}_i^{(n)}$, and $\sigma_n^2 := \sigma^2 + n_i^{-1/2}s_i^{2(n)}$ (since $\boldsymbol{\vartheta}, \boldsymbol{\vartheta}^{(n)} \in \mathcal{M}(\boldsymbol{\Upsilon})$, σ_n^2 and $\mathbf{V}^n$ do not depend on i, which explains the notation). Accordingly, let $\mathbf{Z}_{ij}^0 := \mathbf{V}^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)$, $d_{ij}^0 := \|\mathbf{Z}_{ij}^0\|$, $\mathbf{U}_{ij}^0 := \mathbf{Z}_{ij}^0/d_{ij}^0$, $\mathbf{Z}_{ij}^n := (\mathbf{V}^n)^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i^n)$, $d_{ij}^n := \|\mathbf{Z}_{ij}^n\|$, and $\mathbf{U}_{ij}^n := \mathbf{Z}_{ij}^n/d_{ij}^n$.

Following an argument that goes back to Jurečková (1969), consider the following truncation of the score function K: for all $\ell \in \mathbb{N}_0$, define

$$K^{(\ell)}(u) := K\left(\frac{2}{\ell}\right)\ell\left(u - \frac{1}{\ell}\right) I_{[\frac{1}{\ell} < u \le \frac{2}{\ell}]} + K(u)\, I_{[\frac{2}{\ell} < u \le 1 - \frac{2}{\ell}]}$$
$$+ K\left(1 - \frac{2}{\ell}\right)\ell\left(\left(1 - \frac{1}{\ell}\right) - u\right) I_{[1 - \frac{2}{\ell} < u \le 1 - \frac{1}{\ell}]},$$

where I_A denotes the indicator function of A. Since $u \mapsto K(u)$ is continuous, the functions $u \mapsto K^{(\ell)}(u)$ are also continuous on $(0,1)$. It follows that the truncated scores $K^{(\ell)}$ are bounded for all ℓ. Clearly, it can safely be assumed that K is a monotone increasing function (rather than the difference of two monotone increasing functions), so that there exists some L such that $|K^{(\ell)}(u)| \le |K(u)|$ for all $u \in (0,1)$ and all $\ell \ge L$.

We start with the proof that (A.9) is $o_{\mathrm{P}}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$. For the shape part $\underset{\sim}{\boldsymbol{\Delta}}^{III}_{\boldsymbol{\vartheta};K}$, the result is a straightforward m-sample extension of the corresponding result in Hallin et al. (2006); details are left to the reader. Turning to scale, Lemma 4.1(i) shows that $\underset{\sim}{\Delta}^{II,i}_{\boldsymbol{\vartheta};K} - \Delta^{II,i}_{\boldsymbol{\vartheta};K;g_1}$ is $o_{\mathrm{P}}(1)$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$. Similarly, the difference $\underset{\sim}{\Delta}^{II,i}_{\boldsymbol{\vartheta}^{(n)};K} - \Delta^{II,i}_{\boldsymbol{\vartheta}^{(n)};K;g_1}$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};g_1}$—hence, from contiguity, also under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$. Consequently, (A.9) is asymptotically equivalent, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, to

$$\Delta^{II,i}_{\boldsymbol{\vartheta}^{(n)};K;g_1} - \Delta^{II,i}_{\boldsymbol{\vartheta};K;g_1} + \frac{\mathcal{L}_k(K,g_1)}{4\sigma^4}\, s_i^{2(n)}. \tag{A.11}$$

Now, $\frac{1}{2}n_i^{-1/2}\sum_{j=1}^{n_i}\left(K(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - k\right)$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};g_1}$, is asymptotically normal as $n \to \infty$, with mean zero and variance $\frac{1}{4}\mathcal{L}_k(K)$, so that $\frac{1}{2}(\frac{1}{\sigma_n^2} - \frac{1}{\sigma^2})n_i^{-1/2}\sum_{j=1}^{n_i}\left(K(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - k\right)$ is $o_{\mathrm{P}}(1)$, as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};g_1}$, as well as under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ (by contiguity). Consequently, (A.11) is asymptotically equivalent, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, to

$$\mathbf{C}_i^{(n)} := \frac{1}{2\sigma^2}n_i^{-1/2}\sum_{j=1}^{n_i}\left(K(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - k\right) - \frac{1}{2\sigma^2}n_i^{-1/2}\sum_{j=1}^{n_i}\left(K(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - k\right) + \frac{\mathcal{L}_k(K,g_1)}{4\sigma^4}\, s_i^{2(n)}.$$

Decompose $\mathbf{C}_i^{(n)}$ into $\mathbf{C}_i^{(n)} = \mathbf{D}_{i1}^{(n;\ell)} + \mathbf{D}_{i2}^{(n;\ell)} - \mathbf{R}_{i1}^{(n;\ell)} + \mathbf{R}_{i2}^{(n;\ell)} + \mathbf{R}_{i3}^{(n;\ell)}$ where, denoting by E_0 expectation under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$,

$$
\begin{aligned}
\mathbf{D}_{i1}^{(n;\ell)} &:= \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - \mathrm{E}\big[K^{(\ell)}(U)\big] \Big) \\
&\quad - \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - \mathrm{E}\big[K^{(\ell)}(U)\big] \Big) \\
&\quad - \frac{1}{2\sigma^2} n_i^{-1/2} \mathrm{E}_0 \left[\sum_{j=1}^{n_i} \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - \mathrm{E}\big[K^{(\ell)}(U)\big] \Big) \right], \\
\mathbf{D}_{i2}^{(n;\ell)} &:= \frac{1}{2\sigma^2} n_i^{-1/2} \mathrm{E}_0 \left[\sum_{j=1}^{n_i} \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - \mathrm{E}\big[K^{(\ell)}(U)\big] \Big) \right] + \frac{\mathcal{L}_k(K^{(\ell)}, g_1)}{4\sigma^4}\, s_i^{2(n)}, \\
\mathbf{R}_{i1}^{(n;\ell)} &:= \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big[\Big(K(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - k \Big) - \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - \mathrm{E}\big[K^{(\ell)}(U)\big] \Big) \Big], \\
\mathbf{R}_{i2}^{(n;\ell)} &:= \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big[\Big(K(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - k \Big) - \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - \mathrm{E}\big[K^{(\ell)}(U)\big] \Big) \Big],
\end{aligned}
$$

and

$$
\mathbf{R}_{i3}^{(n;\ell)} := \frac{1}{4\sigma^4} \Big(\mathcal{L}_k(K, g_1) - \mathcal{L}_k(K^{(\ell)}, g_1) \Big)\, s_i^{2(n)}.
$$

We prove that $\mathbf{C}_i^{(n)} = o_{\mathrm{P}}(1)$ (thus completing the proof that (A.9) is $o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$) by establishing that $\mathbf{D}_{i1}^{(n;\ell)}$ and $\mathbf{D}_{i2}^{(n;\ell)}$ are $o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, as $n \to \infty$, for fixed ℓ, and that $\mathbf{R}_{i1}^{(n;\ell)}$, $\mathbf{R}_{i2}^{(n;\ell)}$, and $\mathbf{R}_{i3}^{(n;\ell)}$ are $o_{\mathrm{P}}(1)$ under the same sequence of hypotheses, as $\ell \to \infty$, uniformly in n. For the sake of convenience, these three results are treated separately (Lemmas A.1, A.2, and A.3).

Lemma A.1 *For any fixed ℓ, $\mathrm{E}_0\big[|\mathbf{D}_{i1}^{(n;\ell)}|^2\big] = o(1)$ as $n \to \infty$.*

Lemma A.2 *For any fixed ℓ, $\mathbf{D}_{i2}^{(n;\ell)} = o(1)$ as $n \to \infty$.*

Lemma A.3 *As $\ell \to \infty$, uniformly in n,*

(i) $\mathbf{R}_{i1}^{(n;\ell)}$ *is $o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$,*

(ii) $\mathbf{R}_{i2}^{(n;\ell)}$ *is $o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$ for n sufficiently large,*

(iii) $\mathbf{R}_{i3}^{(n;\ell)}$ *is $o(1)$.*

Proof of Lemma A.1. First note that

$$
\mathbf{D}_{i1}^{(n;\ell)} = \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big[\mathbf{T}_{ij}^{(n;\ell)} - \mathrm{E}_0\big[\mathbf{T}_{ij}^{(n;\ell)}\big] \Big],
$$

where $\mathbf{T}_{ij}^{(n;\ell)} := K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^0/\sigma))$, $j = 1, \ldots, n_i$ are i.i.d. Writing Var_0 for variances under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$,

$$
\mathrm{E}_0\big[|\mathbf{D}_{i1}^{(n;\ell)}|^2\big] = \mathrm{Var}_0\big[\mathbf{D}_{i1}^{(n;\ell)}\big] = \frac{1}{4\sigma^4} \mathrm{Var}_0\big[\mathbf{T}_{i1}^{(n;\ell)}\big] \leq \frac{1}{4\sigma^4}\, \mathrm{E}_0\big[|\mathbf{T}_{i1}^{(n;\ell)}|^2\big],
$$

and it only remains to show that

$$\mathrm{E}_0\big[|\mathbf{T}_{i1}^{(n;\ell)}|^2\big] = \mathrm{E}_0\Big[\Big(K^{(\ell)}(\tilde{G}_{1k}(d_{i1}^n/\sigma_n)) - K^{(\ell)}(\tilde{G}_{1k}(d_{i1}^0/\sigma))\Big)^2\Big] = o(1) \tag{A.12}$$

as $n \to \infty$. Now, $|d_{i1}^n/\sigma_n - d_{i1}^0/\sigma| \leq |d_{i1}^n - d_{i1}^0|/\sigma_n + |\sigma_n^{-1} - \sigma^{-1}|d_{i1}^0$ is $o_\mathrm{P}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$ since $|d_{i1}^n - d_{i1}^0|$ is $o_\mathrm{P}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$; see Lemma A.1 in Hallin et al. (2006). This and the continuity of $K^{(\ell)} \circ \tilde{G}_{1k}$ imply that $K^{(\ell)}(\tilde{G}_{1k}(d_{i1}^n/\sigma_n)) - K^{(\ell)}(\tilde{G}_{1k}(d_{i1}^0/\sigma)) = o_\mathrm{P}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, as $n \to \infty$. Since $K^{(\ell)}$ is bounded, this convergence to zero also holds in quadratic mean, which establishes the convergence in (A.12). □

Proof of Lemma A.2. Letting

$$\mathbf{B}_{i1}^{(n;\ell)} := \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - \mathrm{E}\big[K^{(\ell)}(U)\big]\Big)$$

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

$$\mathbf{B}_{i1}^{(n;\ell)} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(\mathbf{0}, \frac{1}{4\sigma^4}\mathrm{Var}[K^{(\ell)}(U)]\Big). \tag{A.13}$$

Under the sequence of local alternatives $\mathrm{P}_{\boldsymbol{\vartheta}^{(n)};g_1}^{(n)}$, as $n \to \infty$,

$$\mathbf{B}_{i1}^{(n;\ell)} - \frac{\mathcal{L}_k(K^{(\ell)}, g_1)}{4\sigma^4} s_i^{2(n)} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(\mathbf{0}, \frac{1}{4\sigma^4}\mathcal{L}_k(K^{(\ell)})\Big).$$

Defining $\mathbf{B}_{i2}^{(n;\ell)} := \frac{1}{2\sigma^2} n_i^{-1/2} \sum_{j=1}^{n_i} \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^n/\sigma_n)) - \mathrm{E}\big[K^{(\ell)}(U)\big]\Big)$, it follows from ULAN that, under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, as $n \to \infty$,

$$\mathbf{B}_{i2}^{(n;\ell)} + \frac{\mathcal{L}_k(K^{(\ell)}, g_1)}{4\sigma^4} s_i^{2(n)} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(\mathbf{0}, \frac{1}{4\sigma^4}\mathcal{L}_k(K^{(\ell)})\Big). \tag{A.14}$$

Now, from (A.13) and the fact that, under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$, $\mathbf{D}_{i1}^{(n;\ell)} = \mathbf{B}_{i2}^{(n;\ell)} - \mathbf{B}_{i1}^{(n;\ell)} - \mathrm{E}_0\big[\mathbf{B}_{i2}^{(n;\ell)}\big] = o_\mathrm{P}(1)$ as $n \to \infty$ (Lemma A.1), we obtain that

$$\mathbf{B}_{i2}^{(n;\ell)} - \mathrm{E}_0\big[\mathbf{B}_{i2}^{(n;\ell)}\big] \xrightarrow{\mathcal{L}} \mathcal{N}\Big(\mathbf{0}, \frac{1}{4\sigma^4}\mathcal{L}_k(K^{(\ell)})\Big), \tag{A.15}$$

as $n \to \infty$, under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$. The result then follows by comparing (A.14) and (A.15). □

We now complete the proof that (A.9) is $o_\mathrm{P}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$ by proving Lemma A.3.

Proof of Lemma A.3. (i) In view of the independence of the d_{ij}^0's (under $\mathrm{P}_{\boldsymbol{\vartheta};g_1}^{(n)}$), we obtain, for all n,

$$\begin{aligned}
\mathrm{E}_0\big[|\mathbf{R}_{i1}^{(n;\ell)}|^2\big] &= \frac{1}{4\sigma^4} n_i^{-1} \sum_{j=1}^{n_i} \mathrm{E}_0\Big[\Big[\Big(K(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - k\Big) - \Big(K^{(\ell)}(\tilde{G}_{1k}(d_{ij}^0/\sigma)) - \mathrm{E}\big[K^{(\ell)}(U)\big]\Big)\Big]^2\Big] \\
&= \frac{1}{4\sigma^4}\mathrm{Var}\Big[K(U) - K^{(\ell)}(U)\Big] \leq \frac{1}{4\sigma^4}\mathrm{E}\Big[(K(U) - K^{(\ell)}(U))^2\Big] \\
&= \frac{1}{4\sigma^4}\int_0^1 \Big[K(u) - K^{(\ell)}(u)\Big]^2 du.
\end{aligned} \tag{A.16}$$

Clearly, $K^{(\ell)}(u)$ converges to $K(u)$, for all $u \in (0,1)$. Also, since $|K^{(\ell)}(u)|$ is bounded by $|K(u)|$, for all $\ell \geq L$, the integrand in (A.16) is bounded (uniformly in ℓ) by $4\,|K(u)|^2$, which is integrable on $(0,1)$. The Lebesgue dominated convergence theorem thus implies that $\mathrm{E}_0\big[|\mathbf{R}_{i1}^{(n;\ell)}|^2\big] = o(1)$, as $\ell \to \infty$. This convergence is of course uniform in n.

(ii) The claim in (ii) is the same as in (i), except that d_{ij}^n/σ_n replaces d_{ij}^0/σ. Accordingly, (ii) holds under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};g_1}$. That it also holds under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ follows from Lemma 3.5 in Jurečková (1969).

(iii) Note that $|\mathcal{L}_k(K, g_1) - \mathcal{L}_k(K^{(\ell)}, g_1)|^2 = |\mathrm{Cov}[K(U) - K^{(\ell)}(U), K_{g_1}(U)]|^2 \leq \mathcal{L}_k(g_1) \times \mathrm{Var}\big[K(U) - K^{(\ell)}(U)\big]$, which is $o(1)$ as $\ell \to \infty$ (see (i) above). The result then follows from the boundedness of $(s_i^{2(n)})$. □

A.3 Proofs of Lemma 5.1, Theorem 5.2, and Proposition 5.2.

Proof of Lemma 5.1. As in the proof of Lemma 4.1, the result under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$ follows from the multivariate CLT, and under contiguous alternatives from Le Cam's third Lemma. □

Proof of Theorem 5.2. *(i)* The consistency of $\hat{\kappa}_k$, the continuity of the mapping $\boldsymbol{\vartheta} \mapsto \boldsymbol{\Gamma}^{\mathcal{N}}_{\boldsymbol{\vartheta};g_1}$, Lemma 5.1 (jointly with Assumption (D1)), and the fact that $[\mathbf{I}_m - \mathbf{C}^{(n)}](\boldsymbol{\Lambda}^{(n)})^{-1}\mathbf{1}_m = \mathbf{0})$ entail

$$Q^{(n)}_{\mathcal{N}*} = \Big(\mathbf{T}^{II}_{\boldsymbol{\vartheta};g_1}\Big)' (\boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta}_{g_1};g_1})^{\perp}\, \mathbf{T}^{II}_{\boldsymbol{\vartheta};g_1} + \Big(\mathbf{T}^{III}_{\boldsymbol{\vartheta};g_1}\Big)' (\boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1})^{\perp}\, \mathbf{T}^{III}_{\boldsymbol{\vartheta};g_1} + o_{\mathrm{P}}(1) \tag{A.17}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, hence also under the contiguous alternatives $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$.

The result then follows along the same lines as for Theorem 5.1, by noting that $(\boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta};g_1})^{1/2}\,(\boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta};g_1})^{\perp}(\boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta};g_1})^{1/2}$ is a symmetric idempotent matrix with rank $m-1$, which ensures (see Lemma 5.1) that the first term in (A.17) is asymptotically chi-square with $m-1$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and asymptotically noncentral chi-square, still with $m-1$ degrees of freedom but with noncentrality parameter

$$\left(\frac{a_k D_k(g_1)}{2\sigma^4}\right)^2 \lim_{n\to\infty} \Big\{(\boldsymbol{\tau}^{(n)}_{II})'(\boldsymbol{\Gamma}^{\mathcal{N},II}_{\boldsymbol{\vartheta};g_1})^{\perp}\,\boldsymbol{\tau}^{(n)}_{II}\Big\} \tag{A.18}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$. Evaluation of (A.18) yields the first term in (5.12).

As for the shape part, using again Lemma 5.1 and the fact that $(\boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1})^{1/2}\,(\boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1})^{\perp}(\boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1})^{1/2}$ is symmetric and idempotent with rank $K(m-1)$, we obtain similarly that the second term in (A.17) is asymptotically chi-square with $K(m-1)$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and asymptotically noncentral chi-square, still with $K(m-1)$ degrees of freedom but with noncentrality parameter

$$a_k^2(k+2)^2 D_k^2(g_1) \lim_{n\to\infty} \Big\{(\boldsymbol{\tau}^{(n)}_{III})'[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})](\boldsymbol{\Gamma}^{\mathcal{N},III}_{\boldsymbol{\vartheta};g_1})^{\perp}[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})]\,\boldsymbol{\tau}^{(n)}_{III}\Big\} \tag{A.19}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g_1}$. A straightforward evaluation of (A.19) yields the second term in (5.12). As the two terms in (A.17) are asymptotically uncorrelated (see Lemma 5.1 again), they can be treated separately; the result follows.

(ii) The fact that $\phi^{(n)}_{\mathcal{N}}$ has asymptotic level α directly follows from the asymptotic null distribution given in part *(i)* of the theorem and the classical Helly-Bray theorem.

(iii) As observed in the comments that follow (5.10), the consistency of $\hat{\kappa}$ entails the asymptotic equivalence, under Gaussian densities, of $Q^{(n)}_{\mathcal{N}*}$ with $Q^{(n)}_{\mathcal{N},\phi_1}$ (hence with $Q^{(n)}_{\mathcal{N}}$), which has been derived from the general form of locally asymptotically optimal tests based on (4.3). $\square$

Proof of Proposition 5.2. (i) We proceed by showing that the various test statistics are asymptotically equivalent to

$$Q^{(n)}_{\mathcal{N},\phi_1} = \frac{1}{2n} \sum_{1 \leq i < i' \leq m} n_i n_{i'} \; \mathrm{tr}\Big[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\Big]; \tag{A.20}$$

see (5.10). The same calculation as that used to show that the first term in $Q^{(n)}_{\mathrm{Schott}*}$ is asymptotically equivalent to the corresponding term in $Q^{(n)}_{\mathcal{N}*}$ (see Page 21) establishes that $Q^{(n)}_{\mathrm{Schott}}$ is asymptotically equivalent to $Q^{(n)}_{\mathcal{N},\phi_1}$. Now, for the Nagao (1973) test, we have that

$$\begin{aligned} Q^{(n)}_{\mathrm{Nagao}} &= \frac{1}{2}\sum_{i=1}^{m} \dot{n}_i \,\mathrm{tr}\left[(\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}))^2\right] = \frac{1}{2}\sum_{i=1}^{m} \dot{n}_i \,\mathrm{tr}\left[(\dot{\mathbf{S}}^{-1}\sum_{r=1}^{m}\frac{\dot{n}_r}{\dot{n}}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_r))^2\right] \\ &= \frac{1}{2\dot{n}^2}\sum_{i,r,s=1}^{m} \dot{n}_i\dot{n}_r\dot{n}_s \,\mathrm{tr}\left[\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_r)\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_s)\right]. \end{aligned}$$

Splitting $\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_s$ into $(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_r) + (\dot{\mathbf{S}}_r - \dot{\mathbf{S}}_s)$ then yields

$$Q^{(n)}_{\mathrm{Nagao}} = \frac{1}{2\dot{n}}\sum_{i,r=1}^{m} \dot{n}_i\dot{n}_r \,\mathrm{tr}\left[(\dot{\mathbf{S}}^{-1}(\dot{\mathbf{S}}_i - \dot{\mathbf{S}}_r))^2\right] - Q^{(n)}_{\mathrm{Nagao}},$$

which establishes that $Q^{(n)}_{\mathrm{Nagao}} = Q^{(n)}_{\mathcal{N},\phi_1} + o_{\mathrm{P}}(1)$, as $n \to \infty$. As for the LRT (equivalently, the MLRT) test statistics, letting $\mathbf{\Sigma}_0^{1/2}\mathbf{Z}_i\mathbf{\Sigma}_0^{1/2} = n_i^{1/2}(\mathbf{S}_i - \mathbf{\Sigma}_0) = n_i^{1/2}\mathbf{\Sigma}_0\mathbf{Y}_i$ (same notation as in (5.13)), in view of the fact that $\log|\mathbf{I}_k + \mathbf{A}| = \mathrm{tr}\,\mathbf{A} - \frac{1}{2}\mathrm{tr}\,(\mathbf{A}^2) + o(\|\mathbf{A}\|^2)$, as $\|\mathbf{A}\| \to 0$, we have that

$$\begin{aligned} -2\log\Lambda &= -\sum_{i=1}^{m} n_i \log|\mathbf{S}_i| + n\log|\mathbf{S}| = -\sum_{i=1}^{m} n_i \log|\mathbf{\Sigma}_0 + \mathbf{\Sigma}_0\mathbf{Y}_i| + n\,\log\Big|\mathbf{\Sigma}_0 + \mathbf{\Sigma}_0\Big(\frac{1}{n}\sum_{i=1}^{m} n_i\mathbf{Y}_i\Big)\Big| \\ &= \frac{1}{2}\Big\{\sum_{i=1}^{m} n_i\,\mathrm{tr}\,[\mathbf{Y}_i^2] - \frac{1}{n}\,\mathrm{tr}\Big[\Big(\sum_{i=1}^{m} n_i\mathbf{Y}_i\Big)^2\Big]\Big\} + o_{\mathrm{P}}(1), \end{aligned}$$

as $n \to \infty$, under any null distribution with finite fourth-order moments. This establishes the result, since

$$\frac{1}{2}\Big\{\sum_{i=1}^{m} n_i\,\mathrm{tr}\,[\mathbf{Y}_i^2] - \frac{1}{n}\,\mathrm{tr}\Big[\Big(\sum_{i=1}^{m} n_i\mathbf{Y}_i\Big)^2\Big]\Big\} = \frac{1}{2}\sum_{i=1}^{m} n_i\,\mathrm{tr}\Big[\Big(\mathbf{Y}_i - \Big(\frac{1}{n}\sum_{r=1}^{m} n_r\mathbf{Y}_r\Big)\Big)^2\Big]$$

$$= \frac{1}{2}\sum_{i=1}^{m} n_i\,\mathrm{tr}\left[(\mathbf{\Sigma}_0^{-1}\,(\mathbf{S}_i - \mathbf{S}))^2\right] = \frac{1}{2}\sum_{i=1}^{m} n_i\,\mathrm{tr}\left[(\mathbf{S}_i\mathbf{\Sigma}_0^{-1} - \mathbf{I}_k)^2\right] + o_{\mathrm{P}}(1) = Q^{(n)}_{\mathrm{Nagao}} + o_{\mathrm{P}}(1),$$

still as $n \to \infty$, under any null distribution with finite fourth-order moments.

(ii) For any $g_1 \in \mathcal{F}_a^{(4)}$, the result readily follows from Part (i) and Corollary 3.4.2 of Yanagihara et al. (2005) (β there, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g_1}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\mathbf{\Upsilon})$, coincides with $1 + \kappa_k(g_1)$ in our notation). $\square$

References

[1] Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd edition. Wiley-Interscience, New York.

[2] Anderson, T. W., and S. Das Gupta. (1964). A monotonicity property of the power functions of some tests of the equality of two covariance matrices. *Annals of Mathematical Statistics* **35**, 1059-1063.

[3] Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Statistical Society Series A* **160**, 268-282.

[4] Bartlett, M.S. and Kendall, D.G. (1946). The statistical analysis of variance-heterogeneity and the logarithmic transformation. *Supplement to the Journal of the Royal Statistical Society* **8**, 128-138.

[5] Bickel, P.J. (1982). On Adaptive Estimation. *Annals of Statistics* **10**, 647-671.

[6] Box, G.E.P. (1953). Non-normality and tests on variances. *Biometrika* **40**, 318-335.

[7] Brown, G.W. (1939). On the power of the L_1 test for equality of several variances. *Annals of Mathematical Statistics* **10**, 119-128.

[8] Cochran, W.G. (1941). The distribution of the largest of a set of estimated variances as a fraction of their total. *Annals of Eugenics* **11**, 47-52.

[9] Conover, W. J., M. E. Johnson, and M. M. Johnson (1981). Comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* **23**, 351-361.

[10] Das Gupta, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions. *Annals of Mathematical Statistics* **40**, 697-701.

[11] Das Gupta, S., and N. Giri (1973). Properties of tests concerning covariance matrices of normal distributions. *Annals of Statistics* **1**, 1222-1224.

[12] Dümbgen, L. (1998). On Tyler's M-functional of scatter in high dimension. *Annals of the Institute of Statistical Mathematics* **50**, 471–491.

[13] Dümbgen, L., and D.E. Tyler (2005). On the breakdown properties of some multivariate M-Functionals. *Scandinavian Journal of Statistics* **32**, 247–264.

[14] Fligner, M.A., and T.J. Killeen (1976). Distribution-free two-sample tests for scale. *Journal of the American Statistical Association* **71**, 210-213.

[15] Goodnight, C.J., and J.M. Schwartz (1997). A bootstrap comparison of genetic covariance matrices. *Biometrics* **53**, 1026-1035.

[16] Gupta, A.K., and J. Tang (1984). Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate Gaussian models. *Annals of Statistics* **71**, 555-559.

[17] Gupta, A.K., and J. Xu (2006). On some tests of the covariance matrix under general conditions. *Annals of the Institute of Statistical Mathematics* **58** 101-114.

[18] Hallin, M. and D. Paindaveine (2005). Affine invariant aligned rank tests for the multivariate general linear model with ARMA errors. *Journal of Multivariate Analysis* **93**, 122-163.

[19] Hallin, M. and D. Paindaveine (2006a). Semiparametrically efficient rank-based inference for shape I: Optimal rank-based tests for sphericity. *Annals of Statistics*, to appear.

[20] Hallin, M. and D. Paindaveine (2006b). Parametric and semiparametric inference for shape: the role of the scale functional. *Statistics and Decisions*, to appear.

[21] Hallin, M., H. Oja, and D. Paindaveine (2006). Semiparametrically efficient rank-based inference for shape. II. Optimal R-Estimation of Shape. *Annals of Statistics*, to appear.

[22] Hallin, M. and B.J.M. Werker (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9**, 137-165.

[23] Hartley, H.O. (1950). The maximum F-ratio as a shortcut test for heterogeneity of variance. *Biometrika* **37**, 187-194.

[24] Heritier, S. and E. Ronchetti (1994). Robust bounded-influence tests in general parametric models. *Journal of the American Statistical Association* **89**, 897-904.

[25] Hettmansperger, T.P. and R.H. Randles (2002). A practical affine equivariant multivariate median. *Biometrika* **89**, 851-860.

[26] Jurečková, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Annals of Mathematical Statististics* **40**, 1889-1900.

[27] Khatri, C.G., and M.S. Srivastava (1974). Asymptotic expansions of the non-null distributions of likelihood ratio criteria for covariance matrices. *Annals of Statistics* **2**, 109-117.

[28] Kreiss, J.P. (1987). On adaptive estimation in stationary ARMA processes. *Annals of Statistics* **15**, 112-133.

[29] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory.* Springer-Verlag, New York.

[30] Muirhead, R.J., and C.M. Waternaux (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika* **67**, 31-43.

[31] Nagao, H. (1973). On some test criteria for covariance matrix. *Annals of Statistics* **1**, 700-709.

[32] Ollila, E., T.P. Hettmansperger, and H. Oja (2004). Affine equivariant multivariate sign methods. Preprint, University of Jyväskylä.

[33] Paindaveine, D. (2006a). A Chernoff-Savage result for shape. On the non-admissibility of pseudo-Gaussian methods. *Journal of Multivariate Analysis*, to appear.

[34] Paindaveine, D. (2006b). A canonical definition of shape. Submitted.

[35] Perlman, M.D. (1980). Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations. *Annals of Statistics* **8**, 247-263.

[36] Pitman, E.J.G. (1939). Tests of hypothesis concerning location and scale parameters. *Biometrika* **31**, 200-215.

[37] Puri, M.L. and P.K. Sen (1985). *Nonparametric Methods in General Linear Models.* J. Wiley, New York.

[38] Randles, R.H. (2000). A simpler, affine-invariant, multivariate, distribution-free sign test. *Journal of the American Statistical Association* **95**, 1263-1268.

[39] Salibian-Barrera, M., S. Van Aelst, and G. Willems (2006). PCA based on multivariate MM-estimators with fast and robust bootstrap. *Journal of the American Statistical Association*, to appear.

[40] Schott, J. R. (2001). Some tests for the equality of covariance matrices, *Journal of Statistical Planning and Inference* **94**, 25-36.

[41] Srivastava, M.S., C.G. Khatri, and E.M. Carter (1978). On monotonicity of the modified likelihood ratio test for the equality of two covariances. *Journal of Multivariate Analysis* **8**, 262-267.

[42] Sugiura, N. (1969). Asymptotic expansions of the distributions of the likelihood ratio criteria for covariance matrix. *The Annals of Mathematical Statistics* **40**, 2051-2063.

[43] Sugiura, N. (1973). Asymptotic non-null distributions of the likelihood ratio criteria for covariance matrix under local alternatives. *Annals of Statistics* **1**, 718-728.

[44] Taskinen, S., C. Croux, A. Kankainen, E. Ollila, and H. Oja (2006). Influence functions and efficiencies of the canonical correlation and vector estimates based on scatter and shape matrices. *Journal of Multivariate Analysis* **97**, 359–384.

[45] Tatsuoka, K.S., and D.E. Tyler (2000). On the uniqueness of S-Functionals and M-functionals under nonelliptical distributions. *Annals of Statistics* **28**, 1219-1243.

[46] Tyler, D.E. (1983). Robustness and efficiency properties of scatter matrices. *Biometrika* **70**, 411-420.

[47] Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter. *Annals of Statistics* **15**, 234-251.

[48] Um, Y. and R. H. Randles (1998). Nonparametric tests for the multivariate multi-sample location problem. *Statistica Sinica* **8**, 801-812.

[49] Yanagihara, H., T. Tonda, and C. Matsumoto (2005). The effects of nonnormality on asymptotic distributions of some likelihood ratio criteria for testing covariance structures under normal assumption. *Journal of Multivariate Analysis* **96**, 237-264.

[50] Zhang, J., and D.D. Boos (1992). Bootstrap critical values for testing homogeneity of covariance matrices. *Journal of the American Statistical Association* **87**, 425-429.

[51] Zhu, L.X., K.W. Ng, and P. Jing (2002). Resampling methods for homogeneity tests of covariance matrices. *Statistica Sinica* **12**, 769-783.

Marc Hallin
E.C.A.R.E.S,
Institute for Research in Statistics,
and Département de Mathématique,
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
mhallin@ulb.ac.be

Davy Paindaveine
E.C.A.R.E.S,
Institute for Research in Statistics,
and Département de Mathématique,
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
dpaindav@ulb.ac.be

			underlying density						
	k	ξ	t_5	t_8	t_{12}	$\mathcal{N}$	e_2	e_3	e_5
vdW	1	0	2.321	1.230	1.082	1.000	1.151	1.376	1.822
		1	——	——	——	——	——	——	——
	2	0	2.551	1.280	1.102	1.000	1.115	1.296	1.669
		1	2.204	1.215	1.078	1.000	1.129	1.308	1.637
	3	0	2.732	1.322	1.120	1.000	1.092	1.241	1.558
		1	2.270	1.233	1.086	1.000	1.108	1.259	1.536
	4	0	2.881	1.358	1.136	1.000	1.077	1.202	1.475
		1	2.326	1.249	1.093	1.000	1.093	1.223	1.462
	6	0	3.108	1.416	1.163	1.000	1.057	1.151	1.361
		1	2.413	1.275	1.106	1.000	1.072	1.174	1.363
	10	0	3.403	1.498	1.204	1.000	1.037	1.099	1.239
		1	2.531	1.312	1.126	1.000	1.050	1.121	1.254
	∞	0	4.586	1.894	1.446	1.000	1.000	1.000	1.000
		1	3.000	1.500	1.250	1.000	1.000	1.000	1.000
W	1	0	1.993	0.939	0.769	0.608	0.519	0.509	0.517
		1	——	——	——	——	——	——	——
	2	0	2.604	1.185	0.959	0.750	0.694	0.703	0.743
		1	2.258	1.174	1.001	0.844	0.789	0.804	0.842
	3	0	2.929	1.304	1.045	0.811	0.775	0.795	0.854
		1	2.386	1.246	1.068	0.913	0.897	0.933	1.001
	4	0	3.140	1.377	1.096	0.844	0.820	0.844	0.911
		1	2.432	1.273	1.094	0.945	0.955	1.006	1.095
	6	0	3.407	1.467	1.156	0.879	0.866	0.892	0.961
		1	2.451	1.283	1.105	0.969	1.008	1.075	1.188
	10	0	3.685	1.560	1.216	0.908	0.903	0.925	0.984
		1	2.426	1.264	1.088	0.970	1.032	1.106	1.233
	∞	0	4.323	1.794	1.374	0.955	0.955	0.955	0.955
		1	2.250	1.125	0.938	0.750	0.750	0.750	0.750
SP	1	0	2.333	1.126	0.935	0.760	0.705	0.724	0.774
		1	——	——	——	——	——	——	——
	2	0	2.737	1.289	1.063	0.868	0.868	0.924	1.038
		1	2.301	1.230	1.067	0.934	0.965	1.042	1.168
	3	0	2.913	1.348	1.105	0.904	0.924	0.993	1.136
		1	2.277	1.225	1.070	0.957	1.033	1.141	1.317
	4	0	3.016	1.378	1.125	0.920	0.949	1.020	1.170
		1	2.225	1.200	1.051	0.956	1.057	1.179	1.383
	6	0	3.137	1.410	1.142	0.932	0.966	1.032	1.176
		1	2.128	1.146	1.007	0.936	1.057	1.189	1.414
	10	0	3.255	1.438	1.154	0.937	0.969	1.022	1.139
		1	2.001	1.068	0.936	0.891	1.017	1.144	1.365
	∞	0	3.507	1.503	1.176	0.895	0.895	0.895	0.895
		1	1.667	0.833	0.694	0.556	0.556	0.556	0.556

Table 1: AREs, for $\xi = 0$ (pure scale alternatives) and $\xi = 1$ (pure shape alternatives), of the van der Waerden (vdW), Wilcoxon (W), and Spearman (SP) rank-based tests with respect to the pseudo-Gaussian tests, under k-dimensional Student (with 5, 8, and 12 degrees of freedom), normal, and power-exponential densities (with parameter $\eta = 2, 3, 5$), for $k =$ 2, 3, 4, 6, 10, and $k \to \infty$.

		m					
test		0	1	2	3	$\mathrm{ARE}_{\mathcal{N}*}$	$\mathrm{ARE}_{\mathrm{vdW}}$
ϕ_{LRT}	$\mathcal{N}$	.0512	.3168	.7932	.9776	1.000	1.000
ϕ_{MLRT}		.0500	.3100	.7876	.9772	1.000	1.000
$\phi_{\mathcal{N}}$		.0464	.3008	.7760	.9756	1.000	1.000
$\phi_{\mathcal{N}*}$		.0472	.2944	.7568	.9736	1.000	1.000
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0348 (.0472)	.2388 (.2932)	.6912 (.7316)	.9520(.9676)	1.000	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$		.0444 (.0496)	.2604(.2724)	.7080 (.7200)	.9552 (.9600)	0.918	0.918
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0516 (.0524)	.2180(.2248)	.6360(.6404)	.9004(.9016)	0.750	0.750
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0476 (.0492)	.1224 (.1248)	.3252(.3260)	.5692(.5724)	0.360	0.360
$\underset{\sim}{\phi}_{K_2}$		.0432(.0480)	.2448(.2572)	.6956(.7060)	.9480 (.9508)	0.868	0.868
ϕ_{LRT}	t_5	.3288	.6308	.9168	.9872	ND	ND
ϕ_{MLRT}		.3244	.6260	.9144	.9868	ND	ND
$\phi_{\mathcal{N}}$		.3160	.6208	.9092	.9856	ND	ND
$\phi_{\mathcal{N}*}$		.0300	.1896	.5268	.7892	1.000	0.392
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0320 (.0444)	.2500 (.2956)	.7068(.7468)	.9396(.9560)	2.551	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$		.0428 (.0480)	.3004 (.3152)	.7740 (.7812)	.9636 (.9676)	2.778	1.089
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0488 (.0512)	.2916 (.2980)	.7456 (.7520)	.9528 (.9544)	2.604	1.021
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0512(.0516)	.1824 (.1848)	.4916 (.4972)	.7556 (.7612)	1.543	0.605
$\underset{\sim}{\phi}_{K_2}$		.0448 (.0484)	.3068 (.3184)	.7720(.7828)	.9644 (.9656)	2.737	1.073
ϕ_{LRT}	t_2	.8728	.9164	.9496	.9712	ND	ND
ϕ_{MLRT}		.8696	.9156	.9496	.9700	ND	ND
$\phi_{\mathcal{N}}$		.8648	.9120	.9480	.9684	ND	ND
$\phi_{\mathcal{N}*}$		.0120	.0300	.0672	.1276	ND	ND
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0428 (.0568)	.1880 (.2264)	.5368 (.5816)	.7988 (.8324)	ND	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$		.0536 (.0592)	.2532(.2644)	.6592 (.6704)	.9000 (.9072)	ND	1.250
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0508 (.0532)	.2732 (.2816)	.6912 (.6964)	.9212 (.9236)	ND	1.333
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0496 (.0500)	.2116(.2136)	.5404 (.5468)	.8128 (.8144)	ND	1.000
$\underset{\sim}{\phi}_{K_2}$		.0572 (.0588)	.2568 (.2652)	.6632 (.6708)	.9036 (.9080)	ND	1.250
ϕ_{LRT}	$t_{0.5}$	.9992	.9996	.9996	.9988	ND	ND
ϕ_{MLRT}		.9992	.9996	.9996	.9988	ND	ND
$\phi_{\mathcal{N}}$		.9992	.9996	.9988	.9988	ND	ND
$\phi_{\mathcal{N}*}$		0	0	0	0	ND	ND
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0388 (.0520)	.1464 (.1764)	.3096 (.3572)	.4608 (.5188)	ND	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$		.0496 (.0524)	.2328 (.2452)	.5000 (.5132)	.6920 (.7044)	ND	1.543
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0508 (.0528)	.3076 (.3136)	.6404 (.6448)	.8276 (.8316)	ND	2.083
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0604 (.0616)	.3928(.3972)	.7572 (.7600)	.9208 (.9212)	ND	2.778
$\underset{\sim}{\phi}_{K_2}$		.0488 (.0524)	.2136(.2228)	.4728 (0.4840)	.6672(.6792)	ND	1.435

Table 2: Rejection frequencies (out of $N = 2,500$ replications), under the null and various scale alternatives (see Section 7 for details), of the Gaussian LRT (ϕ_{LRT}), its modified version (ϕ_{MLRT}), the parametric Gaussian test ($\phi_{\mathcal{N}}$), its pseudo-Gaussian version ($\phi_{\mathcal{N}*}$), and the signed-rank van der Waerden ($\underset{\sim}{\phi}_{\mathrm{vdW}}$), t_ν-score ($\underset{\sim}{\phi}_{f^t_{1,\nu}}$, $\nu = .5,\ 2,\ 5$), Wilcoxon-type ($\underset{\sim}{\phi}_{K_1}$), and Spearman-type ($\underset{\sim}{\phi}_{K_2}$) tests, respectively. Sample sizes are $n_1 = n_2 = 100$. ARE values are provided with respect to the parametric pseudo-Gaussian ($\mathrm{ARE}_{\mathcal{N}*}$) and van der Waerden rank tests (($\mathrm{ARE}_{\mathcal{N}*}$)); "ND" means "not defined" (this occurs as soon as one the tests involved is not valid under the distribution under consideration).

		m					
test		0	1	2	3	ARE$_{\mathcal{N}*}$	ARE$_{\mathrm{vdW}}$
ϕ_{LRT}		.0512	.1564	.6032	.9668	1.000	1.000
ϕ_{MLRT}		.0500	.1532	.5984	.9656	1.000	1.000
$\phi_{\mathcal{N}}$		.0464	.1484	.5900	.9640	1.000	1.000
$\phi_{\mathcal{N}*}$		.0472	.1444	.5812	.9648	1.000	1.000
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0348(.0472)	.1212 (.1464)	.5248(.5828)	.9488(.9632)	1.000	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$	$\mathcal{N}$	.0444 (.0496)	.1452 (.1536)	.5456(.5596)	.9464 (.9496)	0.945	0.945
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0516 (.0524)	.1364(.1392)	.4928 (.5004)	.9272 (.9276)	0.844	0.844
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0476 (.0492)	.1120(.1140)	.3996 (.4036)	.8356(.8388)	0.648	0.648
$\underset{\sim}{\phi}_{K_2}$		.0432 (.0480)	.1440(.1508)	.5420(.5512)	.9460(.9488)	0.934	0.934
ϕ_{LRT}		.3288	.4632	.7840	.9808	ND	ND
ϕ_{MLRT}		.3244	.4600	.7816	.9800	ND	ND
$\phi_{\mathcal{N}}$		.3160	.4512	.7728	.9796	ND	ND
$\phi_{\mathcal{N}*}$		.0300	.1020	.4204	.8552	1.000	0.454
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0320 (.0444)	.1268 (.1592)	.5320(.5816)	.9576(.9692)	2.204	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$	t_5	.0428(.0480)	.1572(.1676)	.5928 (.6036)	.9720 (.9740)	2.333	1.059
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0488 (.0512)	.1608(.1632)	.5876 (.5916)	.9684(.9692)	2.258	1.024
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0512 (.0516)	.1376 (.1388)	.5088(.5132)	.9312(.9332)	1.896	0.860
$\underset{\sim}{\phi}_{K_2}$		.0448 (.0484)	.1612(.1704)	.5860(.5976)	.9700(.9716)	2.301	1.044
ϕ_{LRT}		.8728	.8912	.9376	.9768	ND	ND
ϕ_{MLRT}		.8696	.8892	.9364	.9768	ND	ND
$\phi_{\mathcal{N}}$		.8648	.8864	.9332	.9764	ND	ND
$\phi_{\mathcal{N}*}$		.0120	.0224	.0808	.2380	ND	ND
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0428(.0568)	.1180(.1488)	.4596(.5120)	.9216 (.9416)	ND	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$	t_2	.0536(.0592)	.1488 (.1560)	.5460(.5572)	.9576 (.9616)	ND	1.147
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0508(.0532)	.1584(.1612)	.5640(.5668)	.9668 (.9668)	ND	1.185
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0496 (.0500)	.1508(.1524)	.5212(.5256)	.9412(.9420)	ND	1.089
$\underset{\sim}{\phi}_{K_2}$		.0572(.0588)	.1440(.1500)	.5288(.5420)	.9516(.9564)	ND	1.111
ϕ_{LRT}		.9992	.9988	.9992	.9992	ND	ND
ϕ_{MLRT}		.9992	.9988	.9992	.9992	ND	ND
$\phi_{\mathcal{N}}$		.9992	.9988	.9992	.9992	ND	ND
$\phi_{\mathcal{N}*}$		0	0	.0004	.0008	ND	ND
$\underset{\sim}{\phi}_{\mathrm{vdW}}$		.0388(.0520)	.0964(.1208)	.3328(.3792)	.7960 (.8328)	ND	1.000
$\underset{\sim}{\phi}_{f^t_{1,5}}$	$t_{0.5}$	.0496 (.0524)	.1280 (.1356)	.4288(.4408)	.8928(.9004)	ND	1.254
$\underset{\sim}{\phi}_{f^t_{1,2}} = \underset{\sim}{\phi}_{K_1}$		.0508(.0528)	.1396(.1440)	.4840(.4880)	.9360(.9380)	ND	1.418
$\underset{\sim}{\phi}_{f^t_{1,.5}}$		.0604 (.0616)	.1644 (.1648)	.5356(.5388)	.9560(.9568)	ND	1.543
$\underset{\sim}{\phi}_{K_2}$		.0488(.0524)	.1208 (.1272)	.3968(.4064)	.8624(.8704)	ND	1.138

Table 3: Rejection frequencies (out of $N = 2,500$ replications), under the null and various shape alternatives (see Section 7 for details), of the Gaussian LRT (ϕ_{LRT}), its modified version (ϕ_{MLRT}), the parametric Gaussian test ($\phi_{\mathcal{N}}$), its pseudo-Gaussian version ($\phi_{\mathcal{N}*}$), and the signed-rank van der Waerden ($\underset{\sim}{\phi}_{\mathrm{vdW}}$), t_ν-score ($\underset{\sim}{\phi}_{f^t_{1,\nu}}$, $\nu = .5,\ 2,\ 5$), Wilcoxon-type ($\underset{\sim}{\phi}_{K_1}$), and Spearman-type ($\underset{\sim}{\phi}_{K_2}$) tests, respectively. Sample sizes are $n_1 = n_2 = 100$. ARE values are provided with respect to the parametric pseudo-Gaussian (ARE$_{\mathcal{N}*}$) and van der Waerden rank tests ((ARE$_{\mathcal{N}*}$)); "ND" means "not defined" (this occurs as soon as one the tests involved is not valid under the distribution under consideration).