ECORE
INTERNATIONAL ASSOCIATION FOR RESEARCH AND TEACHING
Economics, Finance, Operations Research, Econometrics and Statistics

++ research ++ teaching ++

ECORE DISCUSSION PAPER

2007/55

Optimal Tests for Homogeneity of Covariance, Scale,and Shape

Marc HALLIN
Davy PAINDAVEINE

CORE
Université catholique de Louvain

ECARES
Université Libre de Bruxelles

Optimal Tests for Homogeneity of Covariance, Scale, and Shape

Marc Hallin*† and Davy Paindaveine*‡

June 2007

Abstract

The assumption of homogeneity of covariance matrices is the fundamental prerequisite of a number of classical procedures in multivariate analysis. Despite its importance and long history, however, this problem so far has not been completely settled beyond the traditional and highly unrealistic context of multivariate Gaussian models. And the modified likelihood ratio tests (MLRT) that are used in everyday practice are known to be highly sensitive to violations of Gaussian assumptions. In this paper, we provide a complete and systematic study of the problem, and propose test statistics which, while preserving the optimality features of the MLRT under multinormal assumptions, remain valid under unspecified elliptical densities with finite fourth-order moments. As a first step, the Le Cam LAN approach is used for deriving locally and asymptotically optimal testing procedures $\phi_f^{(n)}$ for any specified m-tuple of radial densities $f = (f_1, \ldots, f_m)$. Combined with an estimation of the m densities $f_1, \ldots, f_m$, these procedures can be used to construct adaptive tests for the problem. Adaptive tests however typically require very large samples, and pseudo-Gaussian tests—namely, tests that are locally and asymptotically optimal at Gaussian densities while remaining valid under a much broader class of distributions—in general are preferable. We therefore construct two pseudo-Gaussian modifications of the Gaussian version $\phi_{\mathcal{N}}^{(n)}$ of the optimal test $\phi_f^{(n)}$. The first one, $\phi_{\mathcal{N}*}^{(n)}$, is valid under the class of homokurtic m-tuples f, while the validity of the second, $\phi_{\mathcal{N}\dagger}^{(n)}$, extends to the heterokurtic ones, that is, to arbitrary m-tuples of elliptical distributions with finite fourth-order moments. We moreover show that these tests are asymptotically equivalent to modified Wald tests recently proposed by Schott (2001). This settles the optimality properties of the latter. Our results however are much more informative than Schott's. They also allow for computing local powers, and for an ANOVA-type decomposition of the test statistics into two mutually independent parts providing tests against subalternatives of scale and shape heterogeneity, respectively, thus supplying additional insight into the reasons why rejection occurs. Reinforcing a result of Yanagihara et al. (2005), we further show why another approach, based on bootstrapped critical values of the Gaussian MLRT statistic, although producing asymptotically valid pseudo-Gaussian tests, is highly unsatisfactory in this context. We also develop optimal pseudo-Gaussian tests for scale homogeneity and for shape homogeneity, based on the same methodology. Finally, the small-sample properties the proposed procedures are investigated via a Monte-Carlo study.

*E.C.A.R.E.S., Institut de Recherche en Statistique, and Département de Mathématique, Université Libre de Bruxelles, Belgium. The authors are also members of ECORE, the recently created association between CORE and ECARES.

†Part of this work was completed while Marc Hallin was visiting ORFE and the Bendheim Center for Finance at Princeton University; the financial support of the Bendheim Center is gratefully acknowledged.

‡Research supported by a Crédit d'Impulsion of the Fonds National de la Recherche Scientifique, Communauté française de Belgique.

AMS 1980 subject classification : 62M15, 62G35.

Key words and phrases : Elliptical densities, Homogeneity of covariances, Local asymptotic normality, Locally asymptotically most stringent tests, Multivariate analysis of variance.

1 Introduction.

1.1 Homogeneity of 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 covariance matrices $\boldsymbol{\Sigma}_i$. The assumption $\mathcal{H}_0 : \boldsymbol{\Sigma}_1 = \ldots = \boldsymbol{\Sigma}_m$ of covariance homogeneity is central to the theory and practice of m-sample multivariate analysis, playing a major role in such models as multivariate m-sample location (MANOVA), m-sample multiple-output regression (MANOCOVA) or multivariate discriminant analysis. Testing for $\mathcal{H}_0$ therefore is a problem of fundamental importance, and for more than half a century has been a subject of continued interest in the statistical literature. 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; see Zhang and Boos (1992) for further reference.

The most classical test for this problem is the Gaussian likelihood ratio test $\phi^{(n)}_{\mathrm{LRT}}$ (Wilks 1932). 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 the Bartlett (1937) modification of the likelihood ratio test $\phi^{(n)}_{\mathrm{MLRT}}$, based on the asymptotically chi-square (under Gaussian assumptions) distribution of $Q^{(n)}_{\mathrm{MLRT}} := -2\log\dot{\Lambda}$, where

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

with $\dot{n}_i := n_i - 1$ and $\dot{n} := \sum_{i=1}^m \dot{n}_i = n - m$. This MLRT has been shown to be unbiased for $k = 1$ by Pitman (1939), in the multivariate two-sample case by Sugiura and Nagao (1968), and by Perlman (1980) in 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 $\dot{\Lambda}$ or $-2\log\dot{\Lambda}$ (Sugiura 1973, Khatri and Srivastava 1974, Srivastava et al. 1978), exact distribution of $\dot{\Lambda}$ (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 (asymptotically) valid even under elliptical densities with finite fourth-order moments; see, in particular, Tyler (1983), Yanagihara et al. (2005), and Gupta and Xu (2006).

This non-robustness to violations of normality assumptions places a severe limitation on the applicability of $\phi^{(n)}_{\mathrm{LRT}}$ and $\phi^{(n)}_{\mathrm{MLRT}}$, but is not uncommon in the context. 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 $\mathbf{\Sigma}$ takes some given value $\mathbf{\Sigma}_0$, 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) and Shapiro and Browne (1987).

In particular, for the problem under study, a recent result of Yanagihara et al. (2005) establishes that, under *homokurtic* elliptical densities (when referring to *homo-* or *heterokurticity*, we of course tacitly assume the existence of finite fourth-order moments), the asymptotic null distribution of $Q^{(n)}_{\rm MLRT}$ is that of

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

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, and κ_k stands for the common kurtosis of the m underlying elliptical distributions; see Section 5 for a definition. In the multinormal case, $\kappa_k=0$, and this yields the well-known Gaussian result that $Q^{(n)}_{\rm MLRT}$ is asymptotically chi-square with $(m-1)k(k+1)/2$ degrees of freedom under the null hypothesis; but for $\kappa_k\neq 0$, (1.3) is no longer chi-square (see also Gupta and Xu 2006).

The $(1+\kappa_k)$ factor sitting in front of (1.3) is not uncommon in the context of likelihood ratio testing for covariance matrices (see Theorem 1 of Shapiro and Browne (1987) for a general result about this), and very easily is dealt with by dividing $Q^{(n)}_{\rm MLRT}$ by some consistent estimator $(1+\hat{\kappa}_k)$. The presence of κ_k in the coefficient of Y_1, however, is more problematic. Several attempts therefore have been made (Zhang and Boos 1992, Goodnight and Schwartz 1997, recently followed by Zhu et al. 2002) to bootstrap the MLRT test statistic, but also other measures of covariance heterogeneity. The resampling method of Zhang and Boos (1992), in particular, reconstructs the exact critical values of $Q^{(n)}_{\rm MLRT}$, thus extending the asymptotic validity of the Gaussian MLRT to a broad class of non-Gaussian densities, including non-elliptical ones. In the homokurtic elliptical case, however, those bootstrapped critical values asymptotically coincide with those associated with (1.3). One of the findings of this paper (delayed, for technical reasons, until Section 6.2) is that this approach, while yielding perfectly valid pseudo-Gaussian tests, is nevertheless highly unsatisfactory.

Other Gaussian testing procedures also have been considered. Among them are the test $\phi^{(n)}_{\rm Nagao}$ proposed by Nagao (1973), and the Wald test $\phi^{(n)}_{\rm Schott}$ of Schott (2001). The Nagao test is based on a result by Sugiura (1969) stating that, under Gaussian assumptions, as $n\to\infty$,

$$n^{-1/2}\Big(-2\log\dot{\Lambda}+2\log\frac{\prod_{i=1}^m|\mathbf{\Sigma}_i|^{n_i/2}}{|\mathbf{\Sigma}|^{n/2}}\Big)\stackrel{\mathcal{L}}{\to}\mathcal{N}\Big(0,2\sum_{i=1}^m\lambda_i\,{\rm tr}\,[(\mathbf{\Sigma}_i\mathbf{\Sigma}^{-1}-\mathbf{I}_k)^2]\Big) \tag{1.4}$$

where $\mathbf{\Sigma}:=\sum_{i=1}^m\lambda_i\mathbf{\Sigma}_i$, with $\lambda_i^{(n)}:=n_i/n$ and $\lambda_i:=\lim_{n\to\infty}\lambda_i^{(n)}$. The Nagao test $\phi^{(n)}_{\rm Nagao}$ then rejects the null hypothesis for large values of

$$Q^{(n)}_{\rm Nagao}:=\frac{1}{2}\sum_{i=1}^m\dot{n}_i\,{\rm tr}\Big[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}-\mathbf{I}_k)^2\Big]. \tag{1.5}$$

Schott's Wald test $\phi^{(n)}_{\rm Schott}$ is based on the vector $((\text{vec}(\dot{\mathbf{S}}_1 - \dot{\mathbf{S}}_m))', \ldots, (\text{vec}(\dot{\mathbf{S}}_{m-1} - \dot{\mathbf{S}}_m))')$, and rejects the null hypothesis for large values of the statistic

$$Q^{(n)}_{\rm Schott} := \frac{\dot{n}}{2}\Big\{ \sum_{i=1}^m \dot{\lambda}^{(n)}_i \,\text{tr}\,[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2] - \sum_{i,i'=1}^m \dot{\lambda}^{(n)}_i \dot{\lambda}^{(n)}_{i'} \text{tr}\big[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1}\big]\Big\}. \qquad (1.6)$$

Both $Q^{(n)}_{\rm Nagao}$ and $Q^{(n)}_{\rm Schott}$ are to be compared with the quantiles of their asymptotically chi-square (with $(m-1)k(k+1)/2$ degrees of freedom) null distribution under Gaussian densities.

Whereas Nagao does not say much about the validity under non-Gaussian densities of $\phi^{(n)}_{\rm Nagao}$, Schott stresses the fact that his test is no longer valid in that case, and accordingly proposes (in his Sections 2.2 and 2.3) robustifying $\phi^{(n)}_{\rm Schott}$ into $\phi^{(n)}_{\rm Schott*}$ and $\phi^{(n)}_{\rm Schott\dagger}$ by using an adequate estimate of the underlying asymptotic covariance matrix involved in his Wald statistic. The robustified Schott test $\phi^{(n)}_{\rm Schott*}$ rejects $\mathcal{H}_0$ for large values of

$$\begin{aligned} Q^{(n)}_{\rm Schott*} &:= \frac{\dot{n}}{2(1+\hat{\kappa}_k)}\Big\{ \sum_{i=1}^m \dot{\lambda}^{(n)}_i \,\text{tr}\,[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2] - \sum_{i,i'=1}^m \dot{\lambda}^{(n)}_i \dot{\lambda}^{(n)}_{i'} \text{tr}\big[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1}\big]\Big\} \qquad (1.7) \\ &- \frac{\dot{n}\hat{\kappa}_k}{2(1+\hat{\kappa}_k)((k+2)\hat{\kappa}_k+2)}\Big\{ \sum_{i=1}^m \dot{\lambda}^{(n)}_i \,\text{tr}^2\,[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}] - \sum_{i,i'=1}^m \dot{\lambda}^{(n)}_i \dot{\lambda}^{(n)}_{i'} \text{tr}\big[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\big]\text{tr}\big[\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1}\big]\Big\}, \end{aligned}$$

the null distribution of which is still asymptotically chi-square with $(m-1)k(k+1)/2$ degrees of freedom, but now under any homokurtic m-tuple of elliptical distributions. As for $\phi^{(n)}_{\rm Schott\dagger}$, it allows for heterokurtic elliptical observations, and is based on

$$\begin{aligned} Q^{(n)}_{\rm Schott\dagger} &:= \dot{n}\Big\{ \sum_{i=1}^m \hat{\dot{\alpha}}_i \,\text{tr}\,[(\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1})^2] - \sum_{i,i'=1}^m (\hat{\dot{\alpha}}_i\hat{\dot{\alpha}}_{i'}/\hat{\dot{\alpha}})\,\text{tr}\big[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1}\big]\Big\} \qquad (1.8) \\ &+ \dot{n}\Big\{ \sum_{i=1}^m \hat{\dot{\beta}}_i \,\text{tr}^2\,[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}] - \sum_{i,i'=1}^m \hat{\dot{\tau}}_{i,i'}\,\text{tr}\big[\dot{\mathbf{S}}_i\dot{\mathbf{S}}^{-1}\big]\text{tr}\big[\dot{\mathbf{S}}_{i'}\dot{\mathbf{S}}^{-1}\big]\Big\}, \end{aligned}$$

where

$$\hat{\dot{\alpha}}_i := \frac{\dot{\lambda}^{(n)}_i}{2(1+\hat{\kappa}_{k,i})}, \quad \hat{\dot{\alpha}} := \sum_{i=1}^m \hat{\dot{\alpha}}_i, \quad \hat{\dot{\beta}}_i := \frac{-\dot{\lambda}^{(n)}_i\hat{\kappa}_{k,i}}{2(1+\hat{\kappa}_{k,i})((k+2)\hat{\kappa}_{k,i}+2)},$$

and $\hat{\dot{\tau}}_{i,i'} := \hat{\dot{\alpha}}^{-1}\hat{\dot{\alpha}}_i\hat{\dot{\beta}}_{i'} + (\hat{\dot{\alpha}}_i\hat{\dot{\rho}} + \hat{\dot{\alpha}}^{-1}\hat{\dot{\beta}}_i + k\hat{\dot{\beta}}_i\hat{\dot{\rho}})(\hat{\dot{\alpha}}_{i'} + k\hat{\dot{\beta}}_{i'})$, where $\hat{\dot{\rho}} := -\hat{\dot{\beta}}/(\hat{\dot{\alpha}}(\hat{\dot{\alpha}} + k\hat{\dot{\beta}}))$ and $\hat{\dot{\beta}} := \sum_{i=1}^m \hat{\dot{\beta}}_i$. The asymptotic null distribution of this heterokurtic test statistic coincides with that of $Q^{(n)}_{\rm Schott*}$, but still requires each population to be elliptically symmetric with finite fourth-order moments.

Apart from the bootstrapped versions of the MLRT, Schott's robustified tests $\phi^{(n)}_{\rm Schott*}$ and $\phi^{(n)}_{\rm Schott\dagger}$ are the first and, to the best of our knowledge, the only tests available in the literature that do not require multinormality. Schott, however, apparently is not aware of any asymptotic optimality of his tests (his methodology cannot provide any information about local powers; nor does it provide any rationale for choosing, e.g., between $\phi^{(n)}_{\rm Schott*}$ and the bootstrapped versions of the MLRT), while practitioners are not aware of the fact that $\phi^{(n)}_{\rm Schott*}$ and $\phi^{(n)}_{\rm Schott\dagger}$ are, except for the bootstrapped MLRT, the only available tests which resist non-Gaussian assumptions.

Up to this point, the theory, for this seventy year old fundamental problem, is rather confusing for applied statisticians, who are facing a choice of procedures ($\phi^{(n)}_{\rm LRT}$, (bootstrapped)

$\phi^{(n)}_{\text{MLRT}}$, $\phi^{(n)}_{\text{Nagao}}$, $\phi^{(n)}_{\text{Schott}}$, ...), the optimality features and respective performances of which are all but clear, along with somewhat helpless warnings about their validity that fail to point at any definite recommendation. In the absence of any clear picture, everyday practice keeps defaulting to the traditional $\phi^{(n)}_{\text{MLRT}}$, a procedure nobody would recommend ... It is high time, thus, for providing a complete and general picture of the situation, with clear directions allowing practitioners to select a method they safely can rely on. Providing such a picture, with clear practical recommendations, not only for testing homogeneity of covariances but also for the related problems of scale and shape homogeneity, under the general assumption of heterokurtic elliptical symmetry, is the objective of this paper.

1.2 Outline of the paper.

Sections 2 and 3 mainly introduce the notation and main assumptions, with a short discussion of parametrization and invariance issues. Applying Le Cam's local asymptotic normality (LAN) methodology, we then derive (Sections 4.1 and 4.2) a locally and asymptotically optimal (at any given, possibly heterokurtic m-tuple $f = (f_1, \ldots, f_m)$ of elliptical densities with finite Fisher information) testing procedure $\phi^{(n)}_f$ for the problem. LAN not only allows for characterizing parametric optimality at given f; it also serves as the main tool in studying the behavior of Gaussian test statistics under non-Gaussian densities. Particularizing f, Section 4.3 provides an explicit form of the optimal Gaussian procedure ($\phi^{(n)}_{\mathcal{N}}$, say). In Sections 5.1 (homokurtic case) and 5.2 (heterokurtic case), we solve the "hard" Muirhead and Waternaux (1980) problem of robustifying $\phi^{(n)}_{\mathcal{N}}$, then show (Section 5.3) that our robustifications $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\mathcal{N}\dagger}$ (with test statistic $Q^{(n)}_{\mathcal{N}*}$ and $Q^{(n)}_{\mathcal{N}\dagger}$) are asymptotically equivalent to $\phi^{(n)}_{\text{Schott}*}$ and $\phi^{(n)}_{\text{Schott}\dagger}$, respectively. Moreover, our homokurtic test statistic $Q^{(n)}_{\mathcal{N}*}$ decomposes into a sum $Q^{II(n)}_{\mathcal{N}*} + Q^{III(n)}_{\mathcal{N}*}$, where $Q^{II(n)}_{\mathcal{N}*}$ and $Q^{III(n)}_{\mathcal{N}*}$ are the asymptotically optimal test statistics against the Gaussian subalternatives of scale and shape heterogeneity, respectively; a similar decomposition also holds for the heterokurtic $Q^{(n)}_{\mathcal{N}\dagger}$. In Section 5.4, we first show that under the null and any distribution with finite fourth-order moments, $\phi^{(n)}_{\text{Nagao}}$, $\phi^{(n)}_{\text{Schott}}$, $\phi^{(n)}_{\text{LRT}}$, $\phi^{(n)}_{\text{MLRT}}$ all are asymptotically equivalent to $\phi^{(n)}_{\mathcal{N}}$ and thus share the same Gaussian optimality features, but also the same non-robustness against violations of Gaussian assumptions. Then we considerably reinforce the Yanagihara et al. (2005) result (under homokurtic ellipticity) by turning their convergence in distribution result (1.3) into a convergence in probability result, where Y_1 and Y_2 actually coincide with $Q^{II(n)}_{\mathcal{N}*}$ and $Q^{III(n)}_{\mathcal{N}*}$. Whereas our pseudo-Gaussian test statistic $Q^{(n)}_{\mathcal{N}*}$ under ellipticity reduces to an unweighted sum of the two asymptotically independent chi-square test statistics $Q^{II(n)}_{\mathcal{N}*}$ and $Q^{III(n)}_{\mathcal{N}*}$ respectively detecting scale and shape heterogeneity, the classical MLRT, in its bootstrapped version, is asymptotically equivalent to a weighted linear combination of the same, with weights that do not correspond to any sound decision-theoretic principle, and depend on the unknown underlying density.

This, in principle, settles the problem: under Gaussian assumptions, $\phi^{(n)}_{\text{LRT}}$, $\phi^{(n)}_{\text{MLRT}}$, $\phi^{(n)}_{\text{Nagao}}$, $\phi^{(n)}_{\text{Schott}}$ and $\phi^{(n)}_{\mathcal{N}}$ all are asymptotically equivalent. Under possibly non-Gaussian but homokurtic elliptical symmetry, the asymptotically equivalent $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\text{Schott}*}$ should be considered. Under heterokurtic elliptical symmetry, only the asymptotically equivalent $\phi^{(n)}_{\mathcal{N}\dagger}$ and $\phi^{(n)}_{\text{Schott}\dagger}$ are valid. Under elliptical symmetry but infinite moments of order four, the problem is still tractable via the rank-based tests described in Hallin and Paindaveine (2007), provided however that the m

standardized radial densities coincide. Finally, under possibly non-elliptical densities with finite fourth-order moments, bootstrapping $\phi^{(n)}_{\mathcal{N}*}$ or $\phi^{(n)}_{\mathcal{N}\dagger}$ (equivalently, $\phi^{(n)}_{\text{Schott}*}$ or $\phi^{(n)}_{\text{Schott}\dagger}$) is much preferable to bootstrapping $\phi^{(n)}_{\text{MLRT}}$.

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, for $q \geq 2$,

$$\mathcal{F}^q := \Big\{ h : \mathbb{R}^+_0 \to \mathbb{R}^+ : \mu_{k+q-1;h} < \infty \Big\} \quad \text{and} \quad \mathcal{F}^q_1 := \Big\{ h \in \mathcal{F}^q \ : \frac{\mu_{k+1;h}}{\mu_{k-1;h}} = k \Big\},$$

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

ASSUMPTION (A). The observations $\mathbf{X}_{ij}$, $j = 1, \ldots, n_i$ are mutually independent, with probability density function

$$\mathbf{x} \mapsto c_{k,f_i} |\mathbf{\Sigma}_i|^{-1/2} f_i \left(\left((\mathbf{x} - \boldsymbol{\theta}_i)' \mathbf{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\theta}_i) \right)^{1/2} \right), \quad i = 1, \ldots, m, \tag{2.1}$$

for some k-dimensional vector $\boldsymbol{\theta}_i$ (*location*), some positive definite $(k \times k)$ covariance matrix $\mathbf{\Sigma}_i$, and some f_i in the class $\mathcal{F}^2_1$ of *standardized radial densities*.

Define (throughout, $\mathbf{\Sigma}^{1/2}$ stands for the symmetric root of $\mathbf{\Sigma}$) the *elliptical coordinates*

$$\mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i) := \frac{\mathbf{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)}{\|\mathbf{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)\|} \quad \text{and} \quad d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i) := \|\mathbf{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)\|. \tag{2.2}$$

Under Assumption (A), 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 *standardized elliptical distances* d_{ij} are independent of the $\mathbf{U}_{ij}$, with density $\tilde{f}_{ik}(r) := (\mu_{k-1;f_i})^{-1} r^{k-1} f_i(r)$ (justifying the terminology *standardized radial density* for f_i) and distribution function $\tilde{F}_{ik}$. The condition that $f_i \in \mathcal{F}^2$ is therefore equivalent to the finiteness of d_{ij}'s second-order moments, while $f_i \in \mathcal{F}^2_1$ implies that f_i is standardized in such a way that $\mathrm{E}[d^2_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i)] = k$, hence that $\mathbf{\Sigma}_i = \mathrm{Var}[\mathbf{X}_{ij}]$ is the covariance matrix in population i. In the sequel, we write f for the m-tuples of radial densities $(f_1, \ldots, f_m) \in (\mathcal{F}^2_1)^m$.

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

The derivation of locally and asymptotically optimal tests at radial densities $f (= (f_1, \ldots, f_m))$ will be based on the *uniform local and asymptotic normality* (ULAN) of the model *at given* f. 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. More precisely, ULAN (see Proposition 4.1 below) requires f to belong to $(\mathcal{F}_a^2)^m$, where $\mathcal{F}_a^2$ stands for the collection of absolutely continuous and a.e. positive densities $f_i \in \mathcal{F}_1^2$ for which, letting $\varphi_{f_i} := -\dot{f}_i/f_i$ (with $\dot{f}_i$ the a.e.-derivative of f_i), the integrals

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

are finite. The quantities $\mathcal{I}_k(f_i)$ and $\mathcal{J}_k(f_i)$ play the roles of *radial Fisher information for location* and *radial Fisher information for shape/scale* in population i, 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 $(\mathbf{X}_{11}^{(n)}, \ldots, \mathbf{X}_{1n_1^{(n)}}^{(n)}, \mathbf{X}_{21}^{(n)}, \ldots, \mathbf{X}_{2n_2^{(n)}}^{(n)}, \ldots, \mathbf{X}_{m1}^{(n)}, \ldots, \mathbf{X}_{mn_m^{(n)}}^{(n)})$ indexed by the total sample size n, where the sequences $n_i^{(n)}$ of sample sizes 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 $n \to \infty$. 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 $i = 1, \ldots, m$, $\lambda_i^{(n)} := n_i^{(n)}/n \to \lambda_i \in (0, 1)$, as $n \to \infty$.

3 Parametrization of elliptical families.

3.1 Covariance, 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 $\boldsymbol{\theta}$, covariance $\mathbf{\Sigma}$, and radial density $f (\in \mathcal{F}_1^2)$. The model for this observation is generally parametrized by $(\boldsymbol{\theta}, \mathbf{\Sigma})$. The asymptotic study of this model is more easily conducted if the covariance $\mathbf{\Sigma}$ is decomposed into a product $\sigma^2 \mathbf{V}$, 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). In the testing problem under study, this decomposition moreover corresponds to a decomposition of the alternative into two "natural" subalternatives: heterogeneity of scale and heterogeneity of shape, respectively. When σ^2 is chosen as $|\mathbf{\Sigma}|^{1/k}$, this decomposition, as we shall see, plays an essential role in the interpretation and asymptotic behavior of all tests statistics considered, and induces (see Section 6.1) an ANOVA-type decomposition of the optimal ones.

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 unique 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, Salibián-Barrera et al. 2006, and Taskinen et al. 2006; under the terminology "generalized variance", this determinant-based measure of scale actually goes back to Wilks 1932).

In practice, all choices are essentially equivalent (see Hallin and Paindaveine 2006b). Paindaveine (2007) however shows that the information matrix for $\boldsymbol{\theta}$, σ_S, and $\mathbf{V}_S$ is block-diagonal iff the normalisation $S(\mathbf{\Sigma}) = |\mathbf{\Sigma}|^{1/k}$ is considered. This block-diagonality simplifies several arguments in statistical inference, and we therefore throughout adopt it, simply writing $\mathbf{V} \in \mathcal{V}_k$ and σ for the resulting shape and scale.

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

$$\boldsymbol{\vartheta} := (\boldsymbol{\vartheta}'_I, \boldsymbol{\vartheta}'_{II}, \boldsymbol{\vartheta}'_{III})' := (\boldsymbol{\theta}'_1, \ldots, \boldsymbol{\theta}'_m, \sigma^2_1, \ldots, \sigma^2_m, (\overset{\circ}{\mathrm{vech}}\,\mathbf{V}_1)', \ldots, (\overset{\circ}{\mathrm{vech}}\,\mathbf{V}_m)')',$$

where $L = mk(k+3)/2$ and $\overset{\circ}{\mathrm{vech}}\,(\mathbf{V})$ is defined by $\mathrm{vech}(\mathbf{V}) =: ((\mathbf{V})_{11}, (\overset{\circ}{\mathrm{vech}}\,\mathbf{V})')'$: indeed, $\mathbf{\Sigma}_i$ is entirely determined by σ_i^2 and $\overset{\circ}{\mathrm{vech}}\,(\mathbf{V}_i)$. Write $\mathbf{\Theta}$ for the set $\mathbb{R}^{mk} \times (\mathbb{R}_0^+)^m \times (\overset{\circ}{\mathrm{vech}}\,(\mathcal{V}_k))^m$ of admissible $\boldsymbol{\vartheta}$ values, and $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f}$ for the joint distribution of the n observations under parameter value $\boldsymbol{\vartheta}$ and standardized radial densities $f \in (\mathcal{F}_1^2)^m$. In the sequel, however, we write $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi}$ for the multinormal case ($f = (\phi, \ldots, \phi)$).

3.2 Invariance issues.

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

$$\mathbf{\Upsilon} := \begin{pmatrix} \mathbf{\Upsilon}_I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Upsilon}_{II} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Upsilon}_{III} \end{pmatrix} := \begin{pmatrix} \mathbf{I}_{mk} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_m \otimes \mathbf{I}_{k_0} \end{pmatrix}, \quad k_0 := \frac{k(k+1)}{2} - 1. \tag{3.1}$$

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

This hypothesis is invariant under the group of affine transformations of the observations, which generates the parametric families $\mathcal{P}^{(n)}_{\mathbf{\Upsilon},f} := \bigcup_{\boldsymbol{\vartheta} \in \mathcal{M}(\mathbf{\Upsilon})} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f}\}$. More precisely, this group is the group $\mathcal{G}^{m,k},\circ$ of 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 $\mathcal{G}^{m,k},\circ$ is the group $\tilde{\mathcal{G}}^{m,k},\circ$ of 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 $\mathcal{H}_0$ is invariant under $\mathcal{G}^{m,k},\circ$—meaning that $\mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}}(\mathcal{M}(\mathbf{\Upsilon})) = \mathcal{M}(\mathbf{\Upsilon})$ 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 $Q(\mathbf{A}\mathbf{X}_{11} + \mathbf{b}_1, \ldots \mathbf{A}\mathbf{X}_{mn_m} + \mathbf{b}_m) = Q(\mathbf{X}_{11}, \ldots, \mathbf{X}_{mn_m})$ for all $\mathbf{A}, \mathbf{b}_1, \ldots, \mathbf{b}_m$—yield tests that are *coordinate-free*.

4 Locally asymptotically optimal tests.

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 provide the uniform local asymptotic normality (ULAN) result (with respect to location, scale, and shape parameters, for fixed radial densities $f = (f_1, \ldots, f_m)$) on which optimality will be based.

Writing

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

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

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

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

$$\boldsymbol{\nu}^{(n)} := \begin{pmatrix} \boldsymbol{\nu}_I^{(n)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\nu}_{II}^{(n)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\nu}_{III}^{(n)} \end{pmatrix} := \begin{pmatrix} \boldsymbol{\Lambda}^{(n)} \otimes \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{(n)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}^{(n)} \otimes \mathbf{I}_{k_0} \end{pmatrix} \tag{4.1}$$

(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. Let $\mathrm{diag}(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_m)$ stand for the block-diagonal matrix with diagonal blocks $\mathbf{B}_1$, $\mathbf{B}_2, \ldots, \mathbf{B}_m$. 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 also $\mathbf{K}_k := \sum_{i,j=1}^k (\mathbf{e}_i\mathbf{e}_j') \otimes (\mathbf{e}_j\mathbf{e}_i')$ be the $k^2 \times k^2$ *commutation matrix*, and put $\mathbf{J}_k := (\mathrm{vec}\, \mathbf{I}_k)(\mathrm{vec}\, \mathbf{I}_k)'$. Define $\mathbf{M}_k(\mathbf{V})$ as the $(k_0 \times k^2)$ matrix such that $(\mathbf{M}_k(\mathbf{V}))'(\overset{\circ}{\mathrm{vech}}\, \mathbf{v}) = (\mathrm{vec}\, \mathbf{v})$ for any symmetric $k \times k$ matrix $\mathbf{v}$ such that $\mathrm{tr}(\mathbf{V}^{-1}\mathbf{v}) = 0$. As shown in Paindaveine (2007; Lemma 4.2(v)), $\mathbf{M}_k(\mathbf{V})(\mathrm{vec}\, \mathbf{V}^{-1}) = \mathbf{0}$ for all $\mathbf{V} \in \mathcal{V}_k$. Let further

$$\mathbf{H}_k(\mathbf{V}) := \frac{1}{4k(k+2)}\, \mathbf{M}_k(\mathbf{V})\, [\mathbf{I}_{k^2} + \mathbf{K}_k] \left(\mathbf{V}^{\otimes 2}\right)^{-1} (\mathbf{M}_k(\mathbf{V}))'.$$

Finally, for $f = (f_1, \ldots, f_m) \in (\mathcal{F}_a^2)^m$, write $\underline{\mathcal{I}_k}(f) := \mathrm{diag}(\mathcal{I}_k(f_1), \ldots, \mathcal{I}_k(f_m))$, $\underline{\mathcal{J}_k}(f) := \mathrm{diag}(\mathcal{J}_k(f_1), \ldots, \mathcal{J}_k(f_m))$, and $\underline{\mathcal{L}_k}(f) := \mathrm{diag}(\mathcal{L}_k(f_1), \ldots, \mathcal{L}_k(f_m))$, where $\mathcal{L}_k(f_i) := \mathcal{J}_k(f_i) - k^2$.

We then have the following ULAN result; the proof follows along the same lines as in Theorem 2.1 (which deals with the case $m = 1$ of one population) of Paindaveine (2007) and hence is omitted.

Proposition 4.1 *Assume that (A) and (B) hold, and that* $f(=(f_1,\ldots,f_m)) \in (\mathcal{F}_a^2)^m$. *Then the family* $\mathcal{P}_f^{(n)} := \{\mathrm{P}_{\boldsymbol{\vartheta};f}^{(n)} | \boldsymbol{\vartheta} \in \boldsymbol{\Theta}\}$ *is ULAN, with central sequence*

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

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

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

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

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

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f} := \operatorname{diag}(\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f}^{I}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f}^{II}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f}^{III}), \tag{4.2}$$

where, defining $\underline{\boldsymbol{\sigma}} := \operatorname{diag}(\sigma_1, \ldots, \sigma_m)$, $\underline{\mathbf{V}} := \operatorname{diag}(\mathbf{V}_1, \ldots, \mathbf{V}_m)$, *and* $\mathbf{H}_k(\underline{\mathbf{V}}) := \operatorname{diag}(\mathbf{H}_k(\mathbf{V}_1), \ldots, \mathbf{H}_k(\mathbf{V}_m))$, *we let*

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

and

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f}^{III} := (\underline{\mathcal{J}_k}(f) \otimes \mathbf{I}_{k_0})\, \mathbf{H}_k(\underline{\mathbf{V}}).$$

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

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

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

Via a redefinition of $\boldsymbol{\Sigma}_i$ and f_i (such as in Hallin and Paindaveine 2007), this ULAN result, which, since the problem of covariance homogeneity is void in the absence of second-order moments, we state for $f \in (\mathcal{F}_a^2)^m$, is actually valid under $f \in (\mathcal{F}_a)^m$, where $\mathcal{F}_a$ is defined in the same way as $\mathcal{F}_a^2$ except that finite second-order moments are not required. The null hypothesis of covariance homogeneity then is extended into a null hypothesis of *scatter homogeneity*.

4.2 Locally asymptotically optimal tests.

The classical theory of hypothesis testing in Gaussian shifts (see Section 11.9 of Le Cam 1986) 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

$$Q^{(n)}_{\boldsymbol{\vartheta};f} := \boldsymbol{\Delta}'_{\boldsymbol{\vartheta};f}\boldsymbol{\Gamma}^{-1/2}_{\boldsymbol{\vartheta};f}\Big[\mathbf{I} - \mathrm{proj}(\boldsymbol{\Gamma}^{1/2}_{\boldsymbol{\vartheta};f}(\boldsymbol{\nu}^{(n)})^{-1}\boldsymbol{\Upsilon})\Big]\boldsymbol{\Gamma}^{-1/2}_{\boldsymbol{\vartheta};f}\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f} \tag{4.3}$$

(with $\boldsymbol{\vartheta}$ replaced by an appropriate estimator $\hat{\boldsymbol{\vartheta}}$; see Assumption (C) below), where $\mathrm{proj}(\mathbf{A}) = \mathbf{A}[\mathbf{A}'\mathbf{A}]^{-1}\mathbf{A}'$, for any full rank $(L\times r)$ $(r \leq L)$ matrix $\mathbf{A}$, is the matrix projecting $\mathbb{R}^L$ onto $\mathcal{M}(\mathbf{A})$. Whenever $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f}$, $\boldsymbol{\nu}^{(n)}$ and $\boldsymbol{\Upsilon}$ all happen to be block-diagonal, which is the case in our problem, this projection matrix clearly is also block-diagonal, with diagonal blocks $\mathrm{proj}((\boldsymbol{\Gamma}^{I}_{\boldsymbol{\vartheta};f})^{1/2}(\boldsymbol{\nu}^{(n)}_{I})^{-1}\boldsymbol{\Upsilon}_{I})$, $\mathrm{proj}((\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};f})^{1/2}(\boldsymbol{\nu}^{(n)}_{II})^{-1}\boldsymbol{\Upsilon}_{II})$, and $\mathrm{proj}((\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};f})^{1/2}(\boldsymbol{\nu}^{(n)}_{III})^{-1}\boldsymbol{\Upsilon}_{III})$ denoting projections in $\mathbb{R}^{mk}$, $\mathbb{R}^{m}$, and $\mathbb{R}^{mk_0}$, respectively. Since moreover $\mathcal{M}((\boldsymbol{\Gamma}^{I}_{\boldsymbol{\vartheta};f})^{1/2}(\boldsymbol{\nu}^{(n)}_{I})^{-1}\boldsymbol{\Upsilon}_{I}) = \mathbb{R}^{mk}$, we have $\mathrm{proj}((\boldsymbol{\Gamma}^{I}_{\boldsymbol{\vartheta};f})^{1/2}$ $(\boldsymbol{\nu}^{(n)}_{I})^{-1}\boldsymbol{\Upsilon}_{I}) = \mathbf{I}_{mk}$, so that $Q^{(n)}_{\boldsymbol{\vartheta};f}$ reduces to

$$\begin{aligned} Q^{(n)}_{\boldsymbol{\vartheta};f} &= (\boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};f})'\,(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};f})^{-1/2}\Big[\mathbf{I} - \mathrm{proj}((\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};f})^{1/2}(\boldsymbol{\nu}^{(n)}_{II})^{-1}\boldsymbol{\Upsilon}_{II})\Big]\,(\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};f})^{-1/2}\boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};f} \\ &\qquad + (\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};f})'\,(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};f})^{-1/2}\Big[\mathbf{I} - \mathrm{proj}((\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};f})^{1/2}(\boldsymbol{\nu}^{(n)}_{III})^{-1}\boldsymbol{\Upsilon}_{III})\Big]\,(\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};f})^{-1/2}\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};f} \\ &=: Q^{II(n)}_{\boldsymbol{\vartheta};f} + Q^{III(n)}_{\boldsymbol{\vartheta};f}, \end{aligned} \tag{4.4}$$

where $\boldsymbol{\Delta}^{I}_{\boldsymbol{\vartheta};f}$ does not play any role; note that $Q^{II(n)}_{\boldsymbol{\vartheta};f}$ and $Q^{III(n)}_{\boldsymbol{\vartheta};f}$ are locally and asymptotically optimal for testing $\mathcal{H}^{(n)}_0 : \boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ against subalternatives of the form

$$\bigcup\Big\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f}|\boldsymbol{\vartheta}_{III;1} = \ldots = \boldsymbol{\vartheta}_{III;m} \text{ and } \boldsymbol{\vartheta} \notin \mathcal{M}(\boldsymbol{\Upsilon})\Big\} \text{ and } \bigcup\Big\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f}|\boldsymbol{\vartheta}_{II;1} = \ldots = \boldsymbol{\vartheta}_{II;m} \text{ and } \boldsymbol{\vartheta} \notin \mathcal{M}(\boldsymbol{\Upsilon})\Big\}, \tag{4.5}$$

that is, against scale and shape heterogeneity, respectively. Defining $\bar{\mathcal{J}}^{(n)}_k(f) := \sum_{i=1}^n \lambda^{(n)}_i \mathcal{J}_k(f_i)$ and $\bar{\mathcal{L}}^{(n)}_k(f) := \sum_{i=1}^n \lambda^{(n)}_i \mathcal{L}_k(f_i)$, and writing $\mathbf{C}^{(n)} = (C^{(n)}_{ii'})$ for the $(m \times m)$ matrix with entries $C^{(n)}_{ii'} := (\lambda^{(n)}_i \lambda^{(n)}_{i'})^{1/2}$, it is easy to check that

$$Q^{II(n)}_{\boldsymbol{\vartheta};f} = (\boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};f})'\left\{4\sigma^4\Big[(\underline{\mathcal{L}_k}(f))^{-1} - \frac{1}{\bar{\mathcal{L}}^{(n)}_k(f)}\mathbf{C}^{(n)}\Big]\right\}\boldsymbol{\Delta}^{II}_{\boldsymbol{\vartheta};f}$$

and

$$Q^{III(n)}_{\boldsymbol{\vartheta};f} = (\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};f})'\left\{\Big[(\underline{\mathcal{J}_k}(f))^{-1} - \frac{1}{\bar{\mathcal{J}}^{(n)}_k(f)}\mathbf{C}^{(n)}\Big] \otimes (\mathbf{H}_k(\mathbf{V}))^{-1}\right\}\boldsymbol{\Delta}^{III}_{\boldsymbol{\vartheta};f}.$$

Let $d_{ij} := d_{ij}(\boldsymbol{\theta}_i, \mathbf{V})$, $\mathbf{U}_{ij} := \mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{V})$ and σ be associated with the null parameter value $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$. Then, by using the inversion formula

$$(\mathbf{M}_k(\mathbf{V}))'(\mathbf{H}_k(\mathbf{V}))^{-1}\mathbf{M}_k(\mathbf{V}) = k(k+2)(\mathbf{V}^{\otimes 2})^{1/2}\Big[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\Big](\mathbf{V}^{\otimes 2})^{1/2} \tag{4.6}$$

(see Lemma 5.2 in Hallin and Paindaveine 2006b), one easily obtains

$$\begin{aligned} Q^{(n)}_{\boldsymbol{\vartheta};f} &= \sum_{i,i'=1}^{m}\Big[\frac{\delta_{i,i'}}{n_i\mathcal{L}_k(f_i)} - \frac{1}{n\bar{\mathcal{L}}^{(n)}_k(f)}\Big]\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\left(\varphi_{f_i}\Big(\frac{d_{ij}}{\sigma}\Big)\frac{d_{ij}}{\sigma} - k\right)\left(\varphi_{f_{i'}}\Big(\frac{d_{i'j'}}{\sigma}\Big)\frac{d_{i'j'}}{\sigma} - k\right) \\ &+ \frac{k(k+2)}{2}\sum_{i,i'=1}^{m}\Big[\frac{\delta_{i,i'}}{n_i\mathcal{J}_k(f_i)} - \frac{1}{n\bar{\mathcal{J}}^{(n)}_k(f)}\Big]\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\frac{d_{ij}d_{i'j'}}{\sigma^2}\varphi_{f_i}\Big(\frac{d_{ij}}{\sigma}\Big)\varphi_{f_{i'}}\Big(\frac{d_{i'j'}}{\sigma}\Big)\left((\mathbf{U}'_{ij}\mathbf{U}_{i'j'})^2 - \frac{1}{k}\right) \\ &= Q^{II(n)}_{\boldsymbol{\vartheta};f} + Q^{III(n)}_{\boldsymbol{\vartheta};f}, \end{aligned} \tag{4.7}$$

where $\delta_{i,i'} = 1$ if $i = i'$ and 0 otherwise. As $\boldsymbol{\vartheta}$ remains unspecified under the null, we will need replacing it with some estimate. For this purpose, the traditional LAN approach generally assumes the existence of $\hat{\boldsymbol{\vartheta}} := \hat{\boldsymbol{\vartheta}}^{(n)}$ satisfying

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

(C1) *constrained:* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}[\hat{\boldsymbol{\vartheta}}^{(n)} \in \mathcal{M}(\boldsymbol{\Upsilon})] = 1$ for all n, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and $g \in (\mathcal{F}^2_1)^m$;

(C2) $n^{1/2}(\boldsymbol{\nu}^{(n)})^{-1}$*-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 $\bigcup_{g\in(\mathcal{F}^2_1)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}\}$;

(C3) *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

(C4) *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 $\mathbf{g}^{m,k}_{\mathbf{A},\mathbf{B}} \in \tilde{\mathcal{G}}^{m,k}$.

There are many possible choices for $\hat{\boldsymbol{\vartheta}}$. Using the same notation as in (1.1), a possible choice is provided by

$$\hat{\boldsymbol{\vartheta}} := (\bar{\mathbf{X}}'_1, \ldots, \bar{\mathbf{X}}'_m, |\mathbf{S}|^{1/k}\mathbf{1}'_m, \mathbf{1}'_m \otimes (\overset{\circ}{\mathrm{vech}}\,(\mathbf{S}/|\mathbf{S}|^{1/k}))')'. \tag{4.8}$$

This estimator—which is the natural estimator in the Gaussian or pseudo-Gaussian context—clearly statisfies (C1), (C2), and (C4). After adequate discretization, it also would satisfy (C3). However, (C3) 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). Moreover, the highly regular form of (4.8) makes (C3) unnecessary when considering (pseudo-)Gaussian tests: see the comments after Lemma 5.2.

The locally asymptotically optimal test $\phi^{(n)}_f$ of covariance homogeneity, in the parametric ULAN family $\mathcal{P}^{(n)}_f := \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f} | \boldsymbol{\vartheta} \in \boldsymbol{\Theta}\}$ then consists in rejecting the null hypothesis whenever $Q^{(n)}_f := Q^{(n)}_{\hat{\boldsymbol{\vartheta}};f}$ exceeds the $(1-\alpha)$ quantile of the chi-square distribution with $(m-1)(k_0+1)$ degrees of freedom. Such tests can be combined with an adequate estimate $\hat{f}^{(n)} = (\hat{f}^{(n)}_1, \ldots, \hat{f}^{(n)}_m)$ of the radial density (see Liebscher 2005) in order to build, in the spirit of Section 6.2 of Hallin and Paindaveine (2004), adaptive tests $\phi^{(n)}_{\hat{f}^{(n)}}$, that are uniformly locally asymptotically optimal at any m-tuple f satisfying appropriate regularity conditions. Adaptive tests however typically require very large samples, and we will not pursue along that line. For moderate sample sizes, indeed, better results can be expected from pseudo-Gaussian tests—namely, tests that are asymptotically equivalent, under Gaussian conditions, to the Gaussian version $\phi^{(n)}_{\mathcal{N}}$ of the optimal tests $\phi^{(n)}_f$ but, contrary to $\phi^{(n)}_{\mathcal{N}}$, remain valid under a broad class of non-Gaussian densities.

The derivation of the optimal test $\phi^{(n)}_f$ only requires the family $\mathcal{P}^{(n)}_f$ to be ULAN, and therefore could have been made for any $f \in (\mathcal{F}_a)^m$, yielding optimal tests for scatter homogeneity (see the comment at the end of Section 4.1). We nevertheless restrict to $f \in (\mathcal{F}^2_a)^m$, since turning Gaussian tests into pseudo-Gaussian ones requires second-, actually fourth-order moment assumptions.

4.3 Locally asymptotically optimal Gaussian tests.

In this section, we describe in more details the Gaussian version $\phi_{\mathcal{N}}^{(n)}$ of the optimal tests $\phi_f^{(n)}$ derived in Section 4.2, that is, the one obtained under the assumption that each population is multinormal ($f = (\phi, \ldots, \phi)$). Writing $\hat{d}_{ij} := d_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{V}})$, $\hat{\mathbf{U}}_{ij} := \mathbf{U}_{ij}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{V}})$ for the quantities computed from $\hat{\boldsymbol{\vartheta}} := (\hat{\boldsymbol{\theta}}'_1, \ldots, \hat{\boldsymbol{\theta}}'_m, \hat{\sigma}^2\mathbf{1}'_m, \mathbf{1}'_m \otimes (\mathrm{v\mathring{e}ch}\,\hat{\mathbf{V}})')'$, the Gaussian test statistic

$$Q_{\mathcal{N}}^{(n)} := Q_{\phi}^{(n)} = \sum_{i,i'=1}^{m} \left[\frac{\delta_{i,i'}}{n_i} - \frac{1}{n}\right] \sum_{j=1}^{n_i} \sum_{j'=1}^{n_{i'}} \left\{ \frac{1}{2k}\left(\frac{\hat{d}_{ij}^2}{\hat{\sigma}^2} - k\right)\left(\frac{\hat{d}_{i'j'}^2}{\hat{\sigma}^2} - k\right) + \frac{\hat{d}_{ij}^2 \hat{d}_{i'j'}^2}{2\hat{\sigma}^4}\left((\hat{\mathbf{U}}'_{ij}\hat{\mathbf{U}}_{i'j'})^2 - \frac{1}{k}\right) \right\} =: Q_{\mathcal{N}}^{II(n)} + Q_{\mathcal{N}}^{III(n)} \tag{4.9}$$

can be rewritten as

$$Q_{\mathcal{N}}^{(n)} = \frac{1}{n} \sum_{1 \le i < i' \le m} (n_i + n_{i'})\, Q_{\mathcal{N};i,i'}^{(n)}, \tag{4.10}$$

where

$$Q_{\mathcal{N};i,i'}^{(n)} = \frac{n_i n_{i'}}{(n_i + n_{i'})\hat{\sigma}^4} \left\{ \frac{1}{2k}\left[\left(\frac{1}{n_i}\sum_{j=1}^{n_i} \hat{d}_{ij}^2\right) - \left(\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}} \hat{d}_{i'j'}^2\right)\right]^2 + \frac{1}{2}\left\| \left[\frac{1}{n_i}\sum_{j=1}^{n_i} \hat{d}_{ij}^2 \mathrm{vec}\left(\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij} - \frac{1}{k}\mathbf{I}_k\right)\right] - \left[\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}} \hat{d}_{i'j'}^2 \mathrm{vec}\left(\hat{\mathbf{U}}_{i'j'}\hat{\mathbf{U}}'_{i'j'} - \frac{1}{k}\mathbf{I}_k\right)\right] \right\|^2 \right\}$$
$$=: Q_{\mathcal{N};i,i'}^{II(n)} + Q_{\mathcal{N};i,i'}^{III(n)}$$

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. Most importantly, when the Gaussian estimator $\hat{\boldsymbol{\vartheta}}$ in (4.8) is used, one has

$$\mathrm{tr}\left[\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right] = \left(\frac{1}{n_i}\sum_{j=1}^{n_i} \frac{\hat{d}_{ij}^2}{\hat{\sigma}^2}\right) - \left(\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}} \frac{\hat{d}_{i'j'}^2}{\hat{\sigma}^2}\right) \tag{4.11}$$

and

$$\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\right] - \frac{1}{k}\mathrm{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right] = \left\| \left[\frac{1}{n_i}\sum_{j=1}^{n_i} \frac{\hat{d}_{ij}^2}{\hat{\sigma}^2}\mathrm{vec}\left(\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij} - \frac{1}{k}\mathbf{I}_k\right)\right] - \left[\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}} \frac{\hat{d}_{i'j'}^2}{\hat{\sigma}^2}\mathrm{vec}\left(\hat{\mathbf{U}}_{i'j'}\hat{\mathbf{U}}'_{i'j'} - \frac{1}{k}\mathbf{I}_k\right)\right] \right\|^2, \tag{4.12}$$

so that the Gaussian test statistic reduces to the very simple form

$$Q_{\mathcal{N}}^{(n)} = \frac{1}{n}\sum_{1 \le i < i' \le m} (n_i + n_{i'}) Q_{\mathcal{N};i,i'}^{(n)}, \quad \text{with} \quad Q_{\mathcal{N};i,i'}^{(n)} = \frac{n_i n_{i'}}{2(n_i + n_{i'})}\mathrm{tr}\Big[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\Big]. \tag{4.13}$$

We show in Theorem 5.4 that this $Q_{\mathcal{N}}^{(n)}$ asymptotically coincides with the MLRT test statistic $Q_{\mathrm{MLRT}}^{(n)}$, the Nagao statistic $Q_{\mathrm{Nagao}}^{(n)}$, and Schott's original test statistic $Q_{\mathrm{Schott}}^{(n)}$.

5 Optimal pseudo-Gaussian tests.

The Gaussian test $\phi_{\mathcal{N}}^{(n)}$ of the previous section is unfortunately valid at the multinormal only. In this section, we turn this test into a *pseudo-Gaussian* one, that is, we extend its validity to a broad class of distributions in such a way that its optimality properties at the multinormal are not affected. We actually define two pseudo-Gaussian procedures. The first one ($\phi_{\mathcal{N}*}^{(n)}$; homokurtic case, see Section 5.1) requires the m populations to share the same kurtosis, whereas the second one ($\phi_{\mathcal{N}\dagger}^{(n)}$; heterokurtic case, see Section 5.2) is valid even under kurtosis heterogeneity. In all cases, of course, finite fourth-order moments are needed.

In order to be more specific, we introduce the following notation. For any $g = (g_1, \ldots, g_m) \in (\mathcal{F}_1^4)^m$, write $E_k(g_i) := \mathrm{E}_{\boldsymbol{\vartheta};g_i}[d_{ij}^4(\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_i)] = \int_0^1 (\tilde{G}_{ik}^{-1}(u))^4 du$ and $C_k(g_i) := E_k(g_i) - k^2$. Then, following e.g. Anderson (2003; see page 54), we define the kurtosis of the ith elliptic population under $\mathrm{P}_{\boldsymbol{\vartheta};g}^{(n)}$ as $\kappa_k(g_i) := (k(k+2))^{-1} E_k(g_i) - 1$; note that no population-specific standardization of this kurtosis measure is required since $\mathrm{E}_{\boldsymbol{\vartheta};g_i}[d_{ij}^2(\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_i)] = k$ for all i. For Gaussian densities, $E_k(\phi) = k(k+2)$, $C_k(\phi) = 2k$, and $\kappa_k(\phi) = 0$. With this notation, $\phi_{\mathcal{N}\dagger}^{(n)}$ will be valid under any $g \in (\mathcal{F}_1^4)^m$, whereas $\phi_{\mathcal{N}*}^{(n)}$ will require the underlying g to belong to the homokurtic subset $(\mathcal{F}_1^4)^m_{\mathrm{homo}}$ of $(\mathcal{F}_1^4)^m$ containing m-tuples of densities for which $\kappa_k(g_i)$ (equivalently, $E_k(g_i)$ or $C_k(g_i)$) does not depend on $i = 1, \ldots, m$. In both cases, it is crucial to characterize the asymptotic behavior under non-Gaussian densities of (the scale and shape components of) the Gaussian central sequence, which we denote by $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\phi}$. This behavior is described in the following lemma (see the appendix for the proof).

Lemma 5.1 *Assume that (A) and (B) hold, and that $g \in (\mathcal{F}_1^4)^m$. Then, for any $\boldsymbol{\vartheta}$, $((\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\phi}^{II})', (\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\phi}^{III})')'$ is asymptotically normal with mean zero and mean*

$$\begin{pmatrix} \frac{k}{2} \underline{\boldsymbol{\sigma}}^{-4} \boldsymbol{\tau}_{II} \\ \\ k(k+2) \mathbf{H}_k(\underline{\mathbf{V}}) \boldsymbol{\tau}_{III} \end{pmatrix}$$

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

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\phi}^{g} := \mathrm{diag}(\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\phi}^{g,II}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\phi}^{g,III}) := \mathrm{diag}\left(\frac{1}{4} \underline{C_k}(g) \underline{\boldsymbol{\sigma}}^{-4}, (\underline{E_k}(g) \otimes \mathbf{I}_{k_0}) \mathbf{H}_k(\underline{\mathbf{V}}) \right),$$

where $\underline{E_k}(g) := \mathrm{diag}(E_k(g_1), \ldots, E_k(g_m))$ and $\underline{C_k}(g) := \mathrm{diag}(C_k(g_1), \ldots, C_k(g_m))$ (the claim under $\mathrm{P}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau};g}^{(n)}$ further requires that $g \in (\mathcal{F}_a^4)^m$, where $\mathcal{F}_a^4 := \mathcal{F}_a^2 \cap \mathcal{F}_1^4$).

In order to control for the non-specification of $\boldsymbol{\vartheta}$ in the Gaussian central sequence under arbitrary $g \in (\mathcal{F}_1^4)^m$, we will need the following asymptotic linearity result (see the appendix for the proof).

Lemma 5.2 *Assume that (A) and (B) hold, and that $g \in (\mathcal{F}_1^4)^m$. Let $\hat{\boldsymbol{\vartheta}}$ be the Gaussian estimator (4.8). Then, for any $\boldsymbol{\vartheta}$, (i) $\boldsymbol{\Delta}_{\hat{\boldsymbol{\vartheta}};\phi}^{II} - \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\phi}^{II} + (k/2) \underline{\boldsymbol{\sigma}}^{-4} (\boldsymbol{\nu}_{II}^{(n)})^{-1} n^{1/2} (\hat{\boldsymbol{\vartheta}}_{II}^{(n)} - \boldsymbol{\vartheta}_{II}) = o_{\mathrm{P}}(1)$ and (ii) $\boldsymbol{\Delta}_{\hat{\boldsymbol{\vartheta}};\phi}^{III} - \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\phi}^{III} + k(k+2) \mathbf{H}_k(\underline{\mathbf{V}}) (\boldsymbol{\nu}_{III}^{(n)})^{-1} n^{1/2} (\hat{\boldsymbol{\vartheta}}_{III}^{(n)} - \boldsymbol{\vartheta}_{III}) = o_{\mathrm{P}}(1)$, under $\mathrm{P}_{\boldsymbol{\vartheta};g}^{(n)}$, as $n \to \infty$.*

A glance at the proof of Lemma 5.2 shows that the result still holds if $\hat{\boldsymbol{\vartheta}} = (\bar{\mathbf{X}}_1', \ldots, \bar{\mathbf{X}}_m', \hat{\boldsymbol{\vartheta}}_{II}', \hat{\boldsymbol{\vartheta}}_{III}')'$ satisfies Assumption (C2), so that only the estimators of the location centres need be the Gaussian ones. Similarly, we point out that this lemma would also hold, at any $\mathrm{P}_{\boldsymbol{\vartheta};g}^{(n)}$, $g \in (\mathcal{F}_a^4)^m$, for any estimator $\hat{\boldsymbol{\vartheta}}$ satisfying Assumptions (C2) and (C3). Here, the highly regular form of the Gaussian central sequences for shape and scale and that of the Gaussian estimator of the location centres allow for (i) skipping this unpleasant discretization and for (ii) saving the regularity assumptions ensuring LAN.

5.1 From Gaussian tests to pseudo-Gaussian under homokurticity.

The most natural idea to obtain a g-valid (valid under $g \in (\mathcal{F}_1^4)^m_{\text{homo}}$) transformation of $\phi^{(n)}_{\mathcal{N}}$ is to replace $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\phi}$ with $\boldsymbol{\Gamma}^g_{\boldsymbol{\vartheta};\phi}$ in $Q^{(n)}_{\mathcal{N}}$, that is, to consider

$$Q^{g(n)}_{\mathcal{N}*} := Q^{g(n)}_{\hat{\boldsymbol{\vartheta}},\mathcal{N}*} := \boldsymbol{\Delta}'_{\hat{\boldsymbol{\vartheta}};\phi}\,(\boldsymbol{\Gamma}^g_{\hat{\boldsymbol{\vartheta}};\phi})^{-1/2}\Big[\mathbf{I} - \text{proj}((\boldsymbol{\Gamma}^g_{\hat{\boldsymbol{\vartheta}};\phi})^{1/2}(\boldsymbol{\nu}^{(n)})^{-1}\boldsymbol{\Upsilon})\Big](\boldsymbol{\Gamma}^g_{\hat{\boldsymbol{\vartheta}};\phi})^{-1/2}\boldsymbol{\Delta}_{\hat{\boldsymbol{\vartheta}};\phi}. \tag{5.1}$$

As we show in the proof of Theorem 5.1 below, this indeed provides a valid pseudo-Gaussian test under kurtosis homogeneity, that is, under $g \in (\mathcal{F}_1^4)^m_{\text{homo}}$. For any such g, standard algebra shows that

$$Q^{g(n)}_{\mathcal{N}*} = (\boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\phi})'\,\mathbf{P}^{g,II}_{\hat{\boldsymbol{\vartheta}};\phi*}\,\boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\phi} + (\boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi})'\,\mathbf{P}^{g,III}_{\hat{\boldsymbol{\vartheta}};\phi*}\,\boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi} =: Q^{g,II(n)}_{\mathcal{N}*} + Q^{g,III(n)}_{\mathcal{N}*}, \tag{5.2}$$

where (σ^2 and $\mathbf{V}$ still stand for the common null values of the scale and shape parameters under $\boldsymbol{\vartheta}$, $C_k(g_1)$, $E_k(g_1)$, etc. for the common values, under $g \in (\mathcal{F}_1^4)^m_{\text{homo}}$, of the $C_k(g_i)$'s, etc.),

$$\mathbf{P}^{g,II}_{\boldsymbol{\vartheta};\phi*} = \frac{4\sigma^4}{C_k(g_1)}\,[\mathbf{I}_m - \mathbf{C}^{(n)}] \quad \text{and} \quad \mathbf{P}^{g,III}_{\boldsymbol{\vartheta};\phi*} = \frac{1}{E_k(g_1)}\,[\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes (\mathbf{H}_k(\mathbf{V}))^{-1}. \tag{5.3}$$

With $\hat{d}_{ij}$, $\hat{\mathbf{U}}_{ij}$, and $\hat{\sigma}$ as in (4.9), this test statistic can be reformulated as

$$\begin{aligned} Q^{g(n)}_{\mathcal{N}*} \;=\; & \sum_{i,i'=1}^{m}\left[\frac{\delta_{i,i'}}{n_i} - \frac{1}{n}\right]\sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}}\Bigg\{\frac{1}{C_k(g_1)}\Big(\frac{\hat{d}^2_{ij}}{\hat{\sigma}^2} - k\Big)\Big(\frac{\hat{d}^2_{i'j'}}{\hat{\sigma}^2} - k\Big) \\ & \qquad + \frac{k(k+2)}{2E_k(g_1)}\,\frac{\hat{d}^2_{ij}\hat{d}^2_{i'j'}}{\hat{\sigma}^4}\left((\hat{\mathbf{U}}'_{ij}\hat{\mathbf{U}}_{i'j'})^2 - \frac{1}{k}\right)\Bigg\}, \end{aligned}$$

or

$$Q^{g(n)}_{\mathcal{N}*} = \frac{1}{n}\sum_{1\le i<i'\le m}(n_i + n_{i'})Q^{g(n)}_{\mathcal{N}*;i,i'}, \tag{5.4}$$

with

$$\begin{aligned} Q^{g(n)}_{\mathcal{N}*;i,i'} = & \frac{n_i n_{i'}}{n_i + n_{i'}}\,\frac{1}{\hat{\sigma}^4}\Bigg\{\frac{1}{C_k(g_1)}\left[\Big(\frac{1}{n_i}\sum_{j=1}^{n_i}\hat{d}^2_{ij}\Big) - \Big(\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}}\hat{d}^2_{i'j'}\Big)\right]^2 \\ & + \frac{k(k+2)}{2E_k(g_1)}\left\|\left[\frac{1}{n_i}\sum_{j=1}^{n_i}\hat{d}^2_{ij}\text{vec}\left(\hat{\mathbf{U}}_{ij}\hat{\mathbf{U}}'_{ij} - \frac{1}{k}\mathbf{I}_k\right)\right] - \left[\frac{1}{n_{i'}}\sum_{j'=1}^{n_{i'}}\hat{d}^2_{i'j'}\text{vec}\left(\hat{\mathbf{U}}_{i'j'}\hat{\mathbf{U}}'_{i'j'} - \frac{1}{k}\mathbf{I}_k\right)\right]\right\|^2\Bigg\} \\ & =: Q^{g,II(n)}_{\mathcal{N}*;i,i'} + Q^{g,III(n)}_{\mathcal{N}*;i,i'}. \end{aligned}$$

If the estimator $\hat{\boldsymbol{\vartheta}}$ in (4.8) is used, this can still be written as

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

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

Of course, at the multinormal case ($g = (\phi, \ldots, \phi)$), $Q^{g(n)}_{\mathcal{N}*}$ coincides with $Q^{(n)}_{\mathcal{N}}$ given in (4.13).

Clearly, in order to obtain a genuine test statistic $Q^{(n)}_{\mathcal{N}*}$ (that is, a random variable that does not depend on g anymore) which nevertheless, under any $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$ (with $g \in (\mathcal{F}^4_1)^m_{\text{homo}}$), is asymptotically equivalent to $Q^{g(n)}_{\mathcal{N}*}$, it is sufficient to replace $\kappa_k(g_1)$ with a consistent (still under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$, $g \in (\mathcal{F}^4_1)^m_{\text{homo}}$) estimator $\hat{\kappa}_k$. An obvious choice is $\hat{\kappa}^{(n)}_k := (k(k+2))^{-1}(n^{-1}\sum_{i=1}^m \sum_{j=1}^{n_i} \hat{d}^4_{ij}(\bar{\mathbf{X}}_i, \mathbf{S})) - 1$. The resulting pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ then rejects the null hypothesis (at asymptotic level α) as soon as

$$Q^{(n)}_{\mathcal{N}*} = \frac{1}{n} \sum_{1 \le i < i' \le m} (n_i + n_{i'})\, Q^{(n)}_{\mathcal{N}*;i,i'} > \chi^2_{(m-1)(k_0+1);1-\alpha}, \tag{5.5}$$

where

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

the notation $Q^{II(n)}_{\mathcal{N}*}$, $Q^{III(n)}_{\mathcal{N}*}$, $Q^{II(n)}_{\mathcal{N}*;i,i'}$, $Q^{III(n)}_{\mathcal{N}*;i,i'}$ will be used in similar fashion when $\hat{\kappa}^{(n)}_k$ is substituted for $\kappa_k(g_1)$ in $Q^{g,II(n)}_{\mathcal{N}*}$, etc.

This test statistic is clearly affine-invariant. The following theorem summarizes its asymptotic properties; see (ii) and (iii) for its pseudo-Gaussian nature. For the sake of simplicity, asymptotic powers are expressed under Assumption (B′) and perturbations $\boldsymbol{\tau}^{(n)}$ satisfying $\lim_{n\to\infty} \boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)} = \boldsymbol{\nu}\boldsymbol{\tau} (\notin \mathcal{M}(\boldsymbol{\Upsilon}))$, with

$$\boldsymbol{\nu}_{II}\boldsymbol{\tau}_{II} = (s_1^2/\sqrt{\lambda_1}, \ldots, s_m^2/\sqrt{\lambda_m})' \quad \text{and} \quad \boldsymbol{\nu}_{III}\boldsymbol{\tau}_{III} = ((\overset{\circ}{\mathrm{vech}}\, \mathbf{v}_1)'/\sqrt{\lambda_1}, \ldots, (\overset{\circ}{\mathrm{vech}}\, \mathbf{v}_m)'/\sqrt{\lambda_m})'.$$

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

$$r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} := \frac{1}{\sigma^4} \lim_{n\to\infty} \left\{ (\boldsymbol{\tau}^{(n)}_{II})' [\mathbf{I}_m - \mathbf{C}^{(n)}]\, \boldsymbol{\tau}^{(n)}_{II} \right\} = \sum_{1 \le i < i' \le m} \frac{\lambda_i \lambda_{i'}}{\sigma^4} \left(\frac{s_i^2}{\sqrt{\lambda_i}} - \frac{s_{i'}^2}{\sqrt{\lambda_{i'}}} \right)^2 \tag{5.6}$$

and

$$\begin{aligned} r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} &:= 2k(k+2) \lim_{n\to\infty} \left\{ (\boldsymbol{\tau}^{(n)}_{III})' \Big[[\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes \mathbf{H}_k(\mathbf{V}) \Big]\, \boldsymbol{\tau}^{(n)}_{III} \right\} \\ &= \sum_{1 \le i < i' \le m} \lambda_i \lambda_{i'} \operatorname{tr} \left[\Big(\mathbf{V}^{-1} \Big(\frac{\mathbf{v}_i}{\sqrt{\lambda_i}} - \frac{\mathbf{v}_{i'}}{\sqrt{\lambda_{i'}}} \Big) \Big)^2 \right]; \end{aligned} \tag{5.7}$$

recall that $\operatorname{tr}(\mathbf{V}^{-1}\mathbf{v}_i) = 0$ for all i (see the comments before Proposition 4.1).

Theorem 5.1 *Assume that (A) and (B) hold. Then,* *(i)* $Q^{(n)}_{\mathcal{N}*}$ *is asymptotically chi-square with* $(m-1)(k_0+1)$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in(\mathcal{F}_1^4)^m_{\rm homo}}\{{\rm P}^{(n)}_{\boldsymbol{\vartheta};g}\}$*, and (provided that (B) is reinforced into (B′)) asymptotically noncentral chi-square, still with* $(m-1)(k_0+1)$ *degrees of freedom but with noncentrality parameter*

$$\frac{k^2}{C_k(g_1)}\sum_{1\leq i<i'\leq m}\frac{\lambda_i\lambda_{i'}}{\sigma^4}\left(\frac{s_i^2}{\sqrt{\lambda_i}}-\frac{s_{i'}^2}{\sqrt{\lambda_{i'}}}\right)^2+\frac{k(k+2)}{2E_k(g_1)}\sum_{1\leq i<i'\leq m}\lambda_i\lambda_{i'}\,{\rm tr}\Big[\Big(\mathbf{V}^{-1}\Big(\frac{\mathbf{v}_i}{\sqrt{\lambda_i}}-\frac{\mathbf{v}_{i'}}{\sqrt{\lambda_{i'}}}\Big)\Big)^2\Big]$$
$$=\frac{k}{(k+2)\kappa_k(g_1)+2}\,r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}}+\frac{1}{2(1+\kappa_k(g_1))}\,r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}} \tag{5.8}$$

under ${\rm P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g}$*, with* $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$*,* $\boldsymbol{\nu\tau}:=\lim_{n\to\infty}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}\notin\mathcal{M}(\boldsymbol{\Upsilon})$*, and* $g\in(\mathcal{F}_a^4)^m_{\rm homo}$*, where* $(\mathcal{F}_a^4)^m_{\rm homo}:=(\mathcal{F}_1^4)^m_{\rm homo}\cap(\mathcal{F}_a^4)^m$*;*

(ii) the sequence of tests $\phi^{(n)}_{\mathcal{N}*}$ *has asymptotic size* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in(\mathcal{F}_1^4)^m_{\rm homo}}\{{\rm P}^{(n)}_{\boldsymbol{\vartheta};g}\}$*;*

(iii) the pseudo-Gaussian tests $\phi^{(n)}_{\mathcal{N}*}$ *are asymptotically equivalent, under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\{{\rm P}^{(n)}_{\boldsymbol{\vartheta};\phi}\}$ *and under contiguous alternatives, to the optimal parametric Gaussian tests* $\phi^{(n)}_{\mathcal{N}}$ *based on (4.13); hence, the sequence* $\phi^{(n)}_{\mathcal{N}*}$ *is locally and asymptotically most stringent, still at asymptotic level* α*, for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in(\mathcal{F}_1^4)^m_{\rm homo}}\{{\rm P}^{(n)}_{\boldsymbol{\vartheta};g}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{M}(\boldsymbol{\Upsilon})}\{{\rm P}^{(n)}_{\boldsymbol{\vartheta};\phi}\}$.

The proof is given in the appendix.

5.2 From Gaussian tests to pseudo-Gaussian under possible heterokurticity.

As announced, our main goal is to define a pseudo-Gaussian extension $\phi^{(n)}_{\mathcal{N}\dagger}$ of $\phi^{(n)}_{\mathcal{N}}$, which is valid at any $g=(g_1,\ldots,g_m)\in(\mathcal{F}_1^4)^m$, with possible kurtosis heterogeneity across populations, while maintaining the optimality properties of $\phi^{(n)}_{\mathcal{N}}$ at the multinormal. This goal is not achieved by the homokurtic pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ defined in the previous section, since it turns out that $Q^{(n)}_{\mathcal{N}*}$ under heterokurticity is *no longer (asymptotically) distribution-free* (this is established in the appendix, just before the proof of Theorem 5.2).

As an alternative test statistic, we propose (after due estimation of $\boldsymbol{\vartheta}$)

$$Q^{g(n)}_{\boldsymbol{\vartheta},\mathcal{N}\dagger}:=\boldsymbol{\Delta}'_{\boldsymbol{\vartheta};\phi}\,(\boldsymbol{\Gamma}^g_{\boldsymbol{\vartheta};\phi})^{-1/2}\Big[\mathbf{I}-{\rm proj}((\boldsymbol{\Gamma}^g_{\boldsymbol{\vartheta};\phi})^{-1/2}(\boldsymbol{\nu}^{(n)})^{-1}\boldsymbol{\Upsilon})\Big](\boldsymbol{\Gamma}^g_{\boldsymbol{\vartheta};\phi})^{-1/2}\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\phi}, \tag{5.9}$$

which is obtained from $Q^{g(n)}_{\boldsymbol{\vartheta},\mathcal{N}*}$ in (5.1) by replacing $(\boldsymbol{\Gamma}^g_{\boldsymbol{\vartheta};\phi})^{1/2}$ with $(\boldsymbol{\Gamma}^g_{\boldsymbol{\vartheta};\phi})^{-1/2}$ in the projection matrix. The motivation for this choice is twofold:

(a) unlike $Q^{g(n)}_{\mathcal{N}*}$, the statistic $Q^{g(n)}_{\mathcal{N}\dagger}:=Q^{g(n)}_{\hat{\boldsymbol{\vartheta}},\mathcal{N}\dagger}$ remains asymptotically distribution-free under heterokurticity (see the proof of Theorem 5.2);

(b) under homokurticity ($g\in(\mathcal{F}_1^4)^m_{\rm homo}$), the test based on $Q^{g(n)}_{\mathcal{N}\dagger}$ and the test based on $Q^{g(n)}_{\mathcal{N}*}$ coincide asymptotically (we skip the formal proof of this claim, which easily follows from the algebra of projection matrices, since the same claim readily follows from the explicit expressions for $Q^{g(n)}_{\mathcal{N}\dagger}$ below).

In order to derive such explicit expressions for $Q^{g(n)}_{\mathcal{N}\dagger}$, define the weighted harmonic means $\tilde{C}^{(n)}_k(g) := [\sum_{i=1}^n \lambda_i^{(n)}(C_k(g_i))^{-1}]^{-1}$ and $\tilde{E}^{(n)}_k(g) := [\sum_{i=1}^n \lambda_i^{(n)}(E_k(g_i))^{-1}]^{-1}$. It is then easy to check that

$$Q^{g(n)}_{\mathcal{N}\dagger} := Q^{g(n)}_{\hat{\boldsymbol{\vartheta}},\mathcal{N}\dagger} = \left(\boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\phi}\right)' \mathbf{P}^{g,II}_{\hat{\boldsymbol{\vartheta}};\phi\dagger}\, \boldsymbol{\Delta}^{II}_{\hat{\boldsymbol{\vartheta}};\phi} + \left(\boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi}\right)' \mathbf{P}^{g,III}_{\hat{\boldsymbol{\vartheta}};\phi\dagger}\, \boldsymbol{\Delta}^{III}_{\hat{\boldsymbol{\vartheta}};\phi} =: Q^{g,II(n)}_{\mathcal{N}\dagger} + Q^{g,III(n)}_{\mathcal{N}\dagger},$$

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

$$\mathbf{P}^{g,II}_{\boldsymbol{\vartheta};\phi\dagger} = 4\sigma^4\Big\{ \left(\underline{C_k}(g)\right)^{-1}\Big[\mathbf{I}_m - \tilde{C}^{(n)}_k(g)\,\mathbf{C}^{(n)}\left(\underline{C_k}(g)\right)^{-1}\Big]\Big\}$$

and

$$\mathbf{P}^{g,III}_{\hat{\boldsymbol{\vartheta}};\phi\dagger} = \Big\{ \left(\underline{E_k}(g)\right)^{-1}\Big[\mathbf{I}_m - \tilde{E}^{(n)}_k(g)\,\mathbf{C}^{(n)}\left(\underline{E_k}(g)\right)^{-1}\Big]\Big\} \otimes (\mathbf{H}_k(\mathbf{V}))^{-1};$$

compare with (5.3). With $\hat{d}_{ij}$, $\hat{\mathbf{U}}_{ij}$, and $\hat{\sigma}$ as in (4.9), this test statistic can be reformulated as

$$\begin{aligned} Q^{g(n)}_{\mathcal{N}\dagger} &= \sum_{i,i'=1}^{m} \Big[\frac{\delta_{i,i'}}{n_iC_k(g_i)} - \frac{\tilde{C}^{(n)}_k(g)}{nC_k(g_i)C_k(g_{i'})}\Big] \sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}} \left(\frac{\hat{d}^2_{ij}}{\hat{\sigma}^2} - k\right)\left(\frac{\hat{d}^2_{i'j'}}{\hat{\sigma}^2} - k\right) \\ &\quad + \frac{k(k+2)}{2} \sum_{i,i'=1}^{m} \Big[\frac{\delta_{i,i'}}{n_iE_k(g_i)} - \frac{\tilde{E}^{(n)}_k(g)}{nE_k(g_i)E_k(g_{i'})}\Big] \sum_{j=1}^{n_i}\sum_{j'=1}^{n_{i'}} \frac{\hat{d}^2_{ij}\hat{d}^2_{i'j'}}{\hat{\sigma}^4} \left((\hat{\mathbf{U}}'_{ij}\hat{\mathbf{U}}_{i'j'})^2 - \frac{1}{k}\right), \end{aligned}$$

which, along the same steps as in the derivation of (4.10), and using (4.11) and (4.12), can be written as (if the estimator $\hat{\boldsymbol{\vartheta}}$ in (4.8) is used)

$$\begin{aligned} Q^{g(n)}_{\mathcal{N}\dagger} &= \frac{1}{n} \sum_{1\le i<i'\le m} n_in_{i'} \left\{ \frac{\tilde{C}^{(n)}_k(g)}{C_k(g_i)C_k(g_{i'})} \operatorname{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right] \right. \\ &\quad \left. + \frac{k(k+2)\tilde{E}^{(n)}_k(g)}{2E_k(g_i)E_k(g_{i'})} \Big[\operatorname{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\right] - \frac{1}{k}\operatorname{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right]\Big] \right\}. \end{aligned} \tag{5.10}$$

Letting $\tilde{C}^{(n)}_{k;i,i'}(g) := [\frac{n_i}{n_i+n_{i'}}(C_k(g_i))^{-1} + \frac{n_{i'}}{n_i+n_{i'}}(C_k(g_{i'}))^{-1}]^{-1}$, $\tilde{E}^{(n)}_{k;i,i'}(g) := [\frac{n_i}{n_i+n_{i'}}(E_k(g_i))^{-1} + \frac{n_{i'}}{n_i+n_{i'}}(E_k(g_{i'}))^{-1}]^{-1}$, $\tilde{\kappa}^{(n)}_{k;i,i'}(g) := [\frac{n_i}{n_i+n_{i'}}(\kappa_k(g_i))^{-1} + \frac{n_{i'}}{n_i+n_{i'}}(\kappa_k(g_{i'}))^{-1}]^{-1}$, and $\kappa^{(n)}_{k;i,i'}(g) := \kappa_k(g_i)\kappa_k(g_{i'})/\tilde{\kappa}^{(n)}_{k;i,i'}(g)$, this may be written, in terms of pairwise "quadratic contrasts", as

$$Q^{g(n)}_{\mathcal{N}\dagger} = \frac{1}{n} \sum_{1\le i<i'\le m} (n_i + n_{i'}) \left[\frac{\tilde{C}^{(n)}_k(g)}{\tilde{C}^{(n)}_{k;i,i'}(g)}\, Q^{g,II(n)}_{\mathcal{N}\dagger;i,i'} + \frac{\tilde{E}^{(n)}_k(g)}{\tilde{E}^{(n)}_{k;i,i'}(g)}\, Q^{g,III(n)}_{\mathcal{N}\dagger;i,i'} \right], \tag{5.11}$$

where

$$\begin{aligned} Q^{g,II(n)}_{\mathcal{N}\dagger;i,i'} &= \frac{n_in_{i'}}{n_i + n_{i'}} \frac{\tilde{C}^{(n)}_{k;i,i'}(g)}{C_k(g_i)C_k(g_{i'})} \operatorname{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right] \\ &= \frac{n_in_{i'}}{n_i + n_{i'}} \frac{1}{k(k+2)\kappa^{(n)}_{k;i,i'}(g) + 2k} \operatorname{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'})\right] \end{aligned}$$

and

$$Q^{g,\textit{III}(n)}_{\mathcal{N}\dagger;i,i'} = \frac{n_i n_{i'}}{n_i+n_{i'}} \frac{k(k+2)\tilde{E}^{(n)}_{k;i,i'}(g)}{2E_k(g_i)E_k(g_{i'})} \Big[\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'}))^2\right] - \frac{1}{k}\mathrm{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'})\right]\Big]$$
$$= \frac{n_i n_{i'}}{n_i+n_{i'}} \frac{1}{2(1+\kappa^{(n)}_{k;i,i'}(g))} \Big[\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'}))^2\right] - \frac{1}{k}\mathrm{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'})\right]\Big].$$

It clearly follows from (5.11) that, as announced in the beginning of this subsection, $Q^{g(n)}_{\mathcal{N}\dagger} = Q^{g(n)}_{\mathcal{N}*}$ under homokurticity ($g \in (\mathcal{F}_1^4)^m_{\mathrm{homo}}$). In particular, at the multinormal case, $Q^{g(n)}_{\mathcal{N}\dagger} = Q^{g(n)}_{\mathcal{N}*} = Q^{(n)}_{\mathcal{N}}$. Also note that the two-sample test statistic for populations i and i' is precisely

$$Q^{g(n)}_{\mathcal{N}\dagger;i,i'} := Q^{g,\textit{II}(n)}_{\mathcal{N}\dagger;i,i'} + Q^{g,\textit{III}(n)}_{\mathcal{N}\dagger;i,i'} \tag{5.12}$$
$$= \frac{n_i n_{i'}}{n_i+n_{i'}} \frac{1}{2(1+\kappa^{(n)}_{k;i,i'}(g))}\Big\{\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'}))^2\right] - \frac{\kappa^{(n)}_{k;i,i'}(g)}{(k+2)\kappa^{(n)}_{k;i,i'}(g)+2}\mathrm{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'})\right]\Big\}.$$

Again, in order to obtain a genuine test statistic $Q^{(n)}_{\mathcal{N}\dagger}$, it is sufficient to replace, in (5.11), $E_k(g_i)$ and $C_k(g_i)$ with consistent (under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$, $g \in (\mathcal{F}_1^4)^m$) estimators $\hat{E}^{(n)}_{k,i}$ and $\hat{C}^{(n)}_{k,i}$, $i = 1,\ldots,m$. An obvious choice is $\hat{E}^{(n)}_{k,i} := n_i^{-1}\sum_{j=1}^{n_i}\hat{d}^4_{ij}(\bar{\mathbf{X}}_i,\mathbf{S})$ and $\hat{C}^{(n)}_{k,i} = \hat{E}^{(n)}_{k,i} - k^2$. The pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}\dagger}$ then rejects the null hypothesis (at asymptotic level α) as soon as $Q^{(n)}_{\mathcal{N}\dagger} > \chi^2_{(m-1)(k_0+1);1-\alpha}$, where $Q^{(n)}_{\mathcal{N}\dagger}$ stands for the statistic resulting from this substitution; the notation $Q^{\textit{II}(n)}_{\mathcal{N}\dagger}$, $Q^{\textit{III}(n)}_{\mathcal{N}\dagger}$, $Q^{\textit{II}(n)}_{\mathcal{N}\dagger;i,i'}$, $Q^{\textit{III}(n)}_{\mathcal{N}\dagger;i,i'}$ will be used in similar fashion when population fourth-order moments are replaced with consistent estimates in $Q^{g,\textit{II}(n)}_{\mathcal{N}\dagger}$, etc. The following theorem summarizes the asymptotic properties of the resulting test (see the appendix for the proof).

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

(i) $Q^{(n)}_{\mathcal{N}\dagger}$ *is asymptotically chi-square with* $(m-1)(k_0+1)$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in(\mathcal{F}_1^4)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}\}$*, and (provided that (B) is reinforced into (B′)) asymptotically noncentral chi-square, still with* $(m-1)(k_0+1)$ *degrees of freedom but with noncentrality parameter*

$$k^2\sum_{1\le i<i'\le m}\frac{\tilde{C}_k(g)}{C_k(g_i)C_k(g_{i'})}\frac{\lambda_i\lambda_{i'}}{\sigma^4}\left(\frac{s_i^2}{\sqrt{\lambda_i}}-\frac{s_{i'}^2}{\sqrt{\lambda_{i'}}}\right)^2 \tag{5.13}$$
$$+\frac{k(k+2)}{2}\sum_{1\le i<i'\le m}\frac{\tilde{E}_k(g)}{E_k(g_i)E_k(g_{i'})}\lambda_i\lambda_{i'}\,\mathrm{tr}\Big[\Big(\mathbf{V}^{-1}\Big(\frac{\mathbf{v}_i}{\sqrt{\lambda_i}}-\frac{\mathbf{v}_{i'}}{\sqrt{\lambda_{i'}}}\Big)\Big)^2\Big]$$
$$=\sum_{1\le i<i'\le m}\frac{\tilde{C}_k(g)}{\tilde{C}_{k;i,i'}(g)}\frac{k}{(k+2)\kappa_{k;i,i'}(g)+2}\frac{\lambda_i\lambda_{i'}}{\sigma^4}\left(\frac{s_i^2}{\sqrt{\lambda_i}}-\frac{s_{i'}^2}{\sqrt{\lambda_{i'}}}\right)^2$$
$$+\sum_{1\le i<i'\le m}\frac{\tilde{E}_k(g)}{\tilde{E}_{k;i,i'}(g)}\frac{1}{2(1+\kappa_{k;i,i'}(g))}\lambda_i\lambda_{i'}\,\mathrm{tr}\Big[\Big(\mathbf{V}^{-1}\Big(\frac{\mathbf{v}_i}{\sqrt{\lambda_i}}-\frac{\mathbf{v}_{i'}}{\sqrt{\lambda_{i'}}}\Big)\Big)^2\Big]$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g}$, *with* $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, $\boldsymbol{\nu\tau} := \lim_{n\to\infty} \boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)} \notin \mathcal{M}(\boldsymbol{\Upsilon})$, *and* $g \in (\mathcal{F}_a^4)^m$, *where,* $\tilde{C}_k(g)$, $\tilde{E}_k(g)$, $\tilde{C}_{k;i,i'}(g)$, $\tilde{E}_{k;i,i'}(g)$, *and* $\kappa_{k;i,i'}(g)$ *stand for the limiting values, as* $n \to \infty$, *under Assumption (B′), of* $\tilde{C}_k^{(n)}(g)$, $\tilde{E}_k^{(n)}(g)$, $\tilde{C}_{k;i,i'}^{(n)}(g)$, $\tilde{E}_{k;i,i'}^{(n)}(g)$, *and* $\kappa_{k;i,i'}^{(n)}(g)$, *respectively;*

(ii) the sequence of tests $\phi^{(n)}_{\mathcal{N}\dagger}$ *has asymptotic size* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in(\mathcal{F}_1^4)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}\}$;

(iii) the pseudo-Gaussian tests $\phi^{(n)}_{\mathcal{N}\dagger}$ *are asymptotically equivalent, under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi}\}$ *and under contiguous alternatives, to the optimal parametric Gaussian tests* $\phi^{(n)}_{\mathcal{N}}$ *based on (4.13); hence, the sequence* $\phi^{(n)}_{\mathcal{N}\dagger}$ *is locally and asymptotically most stringent, still at asymptotic level* α, *for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in(\mathcal{F}_1^4)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{M}(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi}\}$.

5.3 Relation between our optimal tests and Schott's tests.

The following result establishes the strong links between our optimal tests $\phi^{(n)}_{\mathcal{N}}$, $\phi^{(n)}_{\mathcal{N}*}$, and $\phi^{(n)}_{\mathcal{N}\dagger}$ and Schott (2001)'s tests $\phi^{(n)}_{\mathrm{Schott}}$, $\phi^{(n)}_{\mathrm{Schott}*}$, and $\phi^{(n)}_{\mathrm{Schott}\dagger}$, as defined in the introduction (see the appendix for the proof).

Theorem 5.3 *The Schott tests* $\phi^{(n)}_{\mathrm{Schott}}$, $\phi^{(n)}_{\mathrm{Schott}*}$, *and* $\phi^{(n)}_{\mathrm{Schott}\dagger}$ *are obtained from the optimal tests* $\phi^{(n)}_{\mathcal{N}}$, $\phi^{(n)}_{\mathcal{N}*}$, *and* $\phi^{(n)}_{\mathcal{N}\dagger}$, *respectively, in the same way the modified likelihood ratio test* $\phi^{(n)}_{\mathrm{MLRT}}$ *is obtained from the likelihood ratio test* $\phi^{(n)}_{\mathrm{LRT}}$*—that is, they follow from replacing* n_i, n, $\mathbf{S}_i$, *and* $\mathbf{S}$ *in* $\phi^{(n)}_{\mathcal{N}}$, $\phi^{(n)}_{\mathcal{N}*}$, *and* $\phi^{(n)}_{\mathcal{N}\dagger}$ *by* $\dot{n}_i$, $\dot{n}$, $\dot{\mathbf{S}}_i$, *and* $\dot{\mathbf{S}}$, *respectively.*

This replacement obviously has no impact on asymptotics, so that Theorems 5.1 and 5.2 also apply to $\phi^{(n)}_{\mathrm{Schott}*}$ and $\phi^{(n)}_{\mathrm{Schott}\dagger}$, respectively. This not only establishes the exact optimality properties of Schott's tests, but also provides their local powers, which do not follow from Schott's original derivation. Similarly, Theorem 5.4 below also applies to $\phi^{(n)}_{\mathrm{Schott}}$, which yields the optimality and (in)validity properties of the latter.

Now, despite the equivalence results of Theorem 5.3, our optimal tests dominate Schott's in the following respects. First of all, the structure of our optimal test statistics makes them readily interpretable as measures of covariance heterogeneity (compare, e.g., $Q^{(n)}_{\mathrm{Schott}\dagger}$ with $Q^{(n)}_{\mathcal{N}\dagger}$). More importantly, in view of the decompositions of $Q^{(n)}_{\mathcal{N}}$ into scale and shape contributions ($Q^{II(n)}_{\mathcal{N}}$ and $Q^{III(n)}_{\mathcal{N}}$), into pairwise comparisons ($Q^{(n)}_{\mathcal{N};i,i'}$), or into pairwise scale and shape comparisons ($Q^{II(n)}_{\mathcal{N};i,i'}$ and $Q^{III(n)}_{\mathcal{N};i,i'}$), $\phi^{(n)}_{\mathcal{N}}$, by providing further insight into the reasons for eventual rejection, allows for a substantially more refined analysis than $\phi^{(n)}_{\mathrm{Schott}}$ (see the discussion in Section 6.1). Similarly, $\phi^{(n)}_{\mathcal{N}*}$, and $\phi^{(n)}_{\mathcal{N}\dagger}$, in view of the corresponding decompositions, are much more informative than $\phi^{(n)}_{\mathrm{Schott}*}$ and $\phi^{(n)}_{\mathrm{Schott}\dagger}$, respectively.

5.4 Relation between the optimal Gaussian test $\phi^{(n)}_{\mathcal{N}}$ and other Gaussian tests.

The relation of our optimal tests 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 $Q^{(n)}_{\mathrm{LRT}} = -2\log\Lambda$ and $Q^{(n)}_{\mathrm{MLRT}} = -2\log\dot{\Lambda}$ (see (1.1) and (1.2), respectively), Schott's original test $\phi^{(n)}_{\rm Schott}$ based on $Q^{(n)}_{\rm Schott}$ (see (1.6)), the "original" Gaussian most stringent test $\phi^{(n)}_{\mathcal{N}}$ based on $Q^{(n)}_{\mathcal{N}}$ (see (4.13)), and the Nagao (1973) test $\phi^{(n)}_{\rm Nagao}$ based on $Q^{(n)}_{\rm Nagao}$ given in (1.5).

The following result establishes the asymptotic equivalence, under the null and any distribution with finite fourth-order moments (including the non-elliptical ones)—hence also under corresponding local alternatives—of all these statistics with the optimal Gaussian one $Q^{(n)}_{\mathcal{N}}$ (which entails their optimality in the Le Cam sense at the multinormal), and explains why none of them qualify as pseudo-Gaussian procedures. The proof is given in the appendix; Part (ii) considerably reinforces the result (1.3) (restricted to $Q^{(n)}_{\rm MLRT}$ and convergence in distribution) of Yanagihara et al. (2005).

Theorem 5.4 *(i) Under the null and any distribution with finite fourth-order moments, $Q^{(n)}_{\rm LRT}$, $Q^{(n)}_{\rm MLRT}$, $Q^{(n)}_{\rm Schott}$, $Q^{(n)}_{\rm Nagao}$ and the Gaussian most stringent test statistics $Q^{(n)}_{\mathcal{N}}$ all are asymptotically equivalent in probability (hence all share $Q^{(n)}_{\mathcal{N}}$'s (in)validity and Gaussian optimality properties).*

(ii) For any $g \in (\mathcal{F}^4_1)^m_{homo}$, the same five statistics, under $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})}\{{\rm P}^{(n)}_{\boldsymbol{\vartheta};g}\}$, all are asymptotically equivalent in probability to

$$(1+\kappa_k(g_1))\left\{\left[1+\frac{k\kappa_k(g_1)}{2(1+\kappa_k(g_1))}\right]Q^{II(n)}_{\mathcal{N}*}+Q^{III(n)}_{\mathcal{N}*}\right\} \tag{5.14}$$

where $Q^{II(n)}_{\mathcal{N}*} = Q^{g,II(n)}_{\mathcal{N}} + o_{\rm P}(1)$ and $Q^{III(n)}_{\mathcal{N}*} = Q^{g,III(n)}_{\mathcal{N}} + o_{\rm P}(1)$ (see (5.2)) asymptotically are independent chi-square random variables, with $(m-1)$ and $(m-1)k_0$ degrees of freedom, respectively.

Clearly, the null distribution of (5.14) is not asymptotically chi-square unless $\kappa_k(g_1) = 0$, that is, when g_1 and hence all g_i's have Gaussian kurtosis—in which case the five test statistics are asymptotically equivalent to the pseudo-Gaussian one $Q^{(n)}_{\mathcal{N}*} = Q^{II(n)}_{\mathcal{N}*} + Q^{III(n)}_{\mathcal{N}*}$.

6 Subalternatives, subhypotheses, and the bootstrapped MLRT.

6.1 Subalternatives and subhypotheses.

Subalternatives of scale and shape heterogeneity naturally enter into the picture via the block-diagonal structure of the information matrix. This block-diagonal structure indeed induces an ANOVA-type decomposition of the optimal test statistics $Q^{(n)}_f$ into $Q^{(n)}_f = Q^{II(n)}_f + Q^{III(n)}_f$ (whether all computed at $\boldsymbol{\vartheta}$ or $\hat{\boldsymbol{\vartheta}}$; see (4.4)), where $Q^{II(n)}_f$ and $Q^{III(n)}_f$ are locally and asymptotically optimal against subalternatives of scale and shape heterogeneity, respectively (see (4.5)). The chi-square p-values of these two asymptotically independent components thus provide an evaluation of the respective contributions of scale and shape heterogeneity in an eventual rejection of $\mathcal{H}^{(n)}_0$, hence an interesting insight into the reasons why rejection occurs. Further study of the statistics $Q^{II(n)}_f$ and $Q^{III(n)}_f$ (with pairwise decompositions on the model of (4.10)) moreover would allow for pairwise conclusions, in the traditional spirit of analysis of variance multiple-comparison methods. These conclusions readily extend to the pseudo-Gaussian procedures $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\mathcal{N}\dagger}$ described in Sections 5.1 and 5.2, respectively.

The same decomposition of covariances into scale and shape similarly leads to considering the subhypotheses $\mathcal{H}^{II}_0 : \sigma_1 = \ldots = \sigma_m$ and $\mathcal{H}^{III}_0 : \mathbf{V}_1 = \ldots = \mathbf{V}_m$ of scale homogeneity and shape homogeneity. Here again, $Q_f^{II(n)}$ and $Q_f^{III(n)}$ provide locally asymptotically optimal tests, provided however that the constraints on $\hat{\boldsymbol{\vartheta}}$ are relaxed appropriatedly. More precisely, using the natural Gaussian constrained estimator

$$\hat{\boldsymbol{\vartheta}}_{\text{scale}} := (\bar{\mathbf{X}}_1', \ldots, \bar{\mathbf{X}}_m', |\mathbf{S}|^{1/k}\mathbf{1}_m', (\overset{\circ}{\text{vech}}\,(\mathbf{S}_1/|\mathbf{S}_1|^{1/k}))', \ldots, , (\overset{\circ}{\text{vech}}\,(\mathbf{S}_m/|\mathbf{S}_m|^{1/k}))')',$$

the optimal pseudo-Gaussian test for scale homogeneity $\phi_{\mathcal{N}\dagger}^{\text{scale}(n)}$ rejects $\mathcal{H}_0^{II}$ (at asymptotic level α) whenever

$$Q_{\mathcal{N}\dagger}^{\text{scale}(n)} = \frac{1}{n} \sum_{1 \le i < i' \le m} (n_i + n_{i'}) \left(\hat{\tilde{C}}_k^{(n)} / \hat{\tilde{C}}_{k;i,i'}^{(n)}\right) Q_{\mathcal{N}\dagger;i,i'}^{\text{scale}(n)} > \chi^2_{m-1;1-\alpha}, \tag{6.1}$$

where

$$Q_{\mathcal{N}\dagger;i,i'}^{\text{scale}(n)} := Q_{\mathcal{N}\dagger;i,i'}^{II(n)}(\hat{\boldsymbol{\vartheta}}_{\text{scale}}) = \frac{n_i n_{i'}}{n_i + n_{i'}} \, \frac{k}{(k+2)\hat{\kappa}_{k;i,i'}^{(n)} + 2} \left\{ \frac{|\mathbf{S}_i|^{1/k} - |\mathbf{S}_{i'}|^{1/k}}{|\mathbf{S}|^{1/k}} \right\}^2 .$$

Similarly, when testing for shape homogeneity, the natural Gaussian constrained estimator is

$$\hat{\boldsymbol{\vartheta}}_{\text{shape}} := (\bar{\mathbf{X}}_1', \ldots, \bar{\mathbf{X}}_m', |\mathbf{S}_1|^{1/k}, \ldots, |\mathbf{S}_m|^{1/k}, \mathbf{1}_m' \otimes (\overset{\circ}{\text{vech}}\,(\mathbf{S}/|\mathbf{S}|^{1/k}))')',$$

and the optimal pseudo-Gaussian test $\phi_{\mathcal{N}\dagger}^{\text{shape}(n)}$ rejects $\mathcal{H}_0^{III}$ (still at asymptotic level α) whenever

$$Q_{\mathcal{N}\dagger}^{\text{shape}(n)} = \frac{1}{n} \sum_{1 \le i < i' \le m} (n_i + n_{i'}) \left(\hat{\tilde{E}}_k^{(n)} / \hat{\tilde{E}}_{k;i,i'}^{(n)}\right) Q_{\mathcal{N}\dagger;i,i'}^{\text{shape}(n)} > \chi^2_{(m-1)k_0;1-\alpha}, \tag{6.2}$$

where, letting $\hat{\mathbf{V}} := \mathbf{S}/|\mathbf{S}|^{1/k}$ and $\hat{\mathbf{V}}_i := \mathbf{S}_i/|\mathbf{S}_i|^{1/k}$, $i = 1, \ldots, m$,

$$Q_{\mathcal{N}\dagger;i,i'}^{\text{shape}(n)} := Q_{\mathcal{N}\dagger;i,i'}^{III(n)}(\hat{\boldsymbol{\vartheta}}_{\text{shape}}) = \frac{n_i n_{i'}}{n_i + n_{i'}} \, \frac{1}{2(1 + \hat{\kappa}_{k;i,i'}^{(n)})} \left\{ \text{tr}\left[(\hat{\mathbf{V}}^{-1}(\hat{\mathbf{V}}_i - \hat{\mathbf{V}}_{i'}))^2\right] - \frac{1}{k} \text{tr}^2 \left[\hat{\mathbf{V}}^{-1}(\hat{\mathbf{V}}_i - \hat{\mathbf{V}}_{i'})\right] \right\}.$$

Here, $\hat{\tilde{C}}_k^{(n)}$, $\hat{\tilde{C}}_{k;i,i'}^{(n)}$, $\hat{\tilde{E}}_k^{(n)}$, $\hat{\tilde{E}}_{k;i,i'}^{(n)}$, and $\hat{\kappa}_{k;i,i'}^{(n)}$ stand for consistent estimators of $\tilde{C}_k^{(n)}$, $\tilde{C}_{k;i,i'}^{(n)}$, $\tilde{E}_k^{(n)}$, $\tilde{E}_{k;i,i'}^{(n)}$, and $\kappa_{k;i,i'}^{(n)}$, respectively.

The pseudo-Gaussian tests $\phi_{\mathcal{N}\dagger}^{\text{scale}(n)}$ and $\phi_{\mathcal{N}\dagger}^{\text{shape}(n)}$ above allow for heterokurticity. If homokurticity can be assumed, one may rely on (i) the test for scale homogeneity $\phi_{\mathcal{N}*}^{\text{scale}(n)}$ rejecting $\mathcal{H}_0^{II}$ (at asymptotic level α) whenever

$$Q_{\mathcal{N}*}^{\text{scale}(n)} = \frac{1}{n} \sum_{1 \le i < i' \le m} (n_i + n_{i'})\, Q_{\mathcal{N}*;i,i'}^{\text{scale}(n)} > \chi^2_{m-1;1-\alpha},$$

where (denoting by $\hat{\kappa}_k^{(n)}$ a consistent estimator of the common value of the $\kappa_k(g_i)$'s)

$$Q_{\mathcal{N}*;i,i'}^{\text{scale}(n)} := \frac{n_i n_{i'}}{n_i + n_{i'}} \, \frac{k}{(k+2)\hat{\kappa}_k^{(n)} + 2} \left\{ \frac{|\mathbf{S}_i|^{1/k} - |\mathbf{S}_{i'}|^{1/k}}{|\mathbf{S}|^{1/k}} \right\}^2,$$

and (ii) the test $\phi_{\mathcal{N}*}^{\text{shape}(n)}$ rejecting $\mathcal{H}_0^{III}$ (still at asymptotic level α) whenever

$$Q_{\mathcal{N}*}^{\text{shape}(n)} = \frac{1}{n} \sum_{1 \le i < i' \le m} (n_i + n_{i'})\, Q_{\mathcal{N}*;i,i'}^{\text{shape}(n)} > \chi^2_{(m-1)k_0;1-\alpha},$$

with

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

Analogues of Theorems 5.1 and 5.2 are easily derived for $\phi^{\text{scale}(n)}_{\mathcal{N}\dagger}$, $\phi^{\text{shape}(n)}_{\mathcal{N}\dagger}$, $\phi^{\text{scale}(n)}_{\mathcal{N}*}$, and $\phi^{\text{shape}(n)}_{\mathcal{N}*}$. For instance, the noncentrality parameter of $\phi^{\text{scale}(n)}_{\mathcal{N}*}$ is $\frac{k}{(k+2)\kappa_k(g_1)+2}\, r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}}$ (with $(m-1)$ degrees of freedom and $r^{II}_{\boldsymbol{\vartheta},\boldsymbol{\tau}}$ given in (5.6)) and that of $\phi^{\text{shape}(n)}_{\mathcal{N}*}$ is $\frac{1}{2(1+\kappa_k(g_1))}\, r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}}$ (with $(m-1)k_0$ degrees of freedom and $r^{III}_{\boldsymbol{\vartheta},\boldsymbol{\tau}}$ given in (5.7)). Noncentrality parameters for $\phi^{\text{scale}(n)}_{\mathcal{N}\dagger}$ and $\phi^{\text{shape}(n)}_{\mathcal{N}\dagger}$ are similarly determined from Theorem 5.2. Details are left to the reader.

6.2 The bootstrapped MLRT.

As mentioned in the introduction, bootstrapping the MLRT test statistic is another way of obtaining valid pseudo-Gaussian critical values. Part (ii) of Theorem 5.4 and the discussion in Section 6.1 however tell us that the resulting test, under homokurticity, is asymptotically equivalent to a test based on a linear combination (5.14) of $Q^{II(n)}_{\mathcal{N}*}$ (optimal against local scale heterogeneity, locally insensitive to local shape heterogeneity) and $Q^{III(n)}_{\mathcal{N}*}$ (optimal against local shape heterogeneity, locally insensitive to local scale heterogeneity). These two subalternatives in (5.14) are weighted according to the common kurtosis $\kappa_k(g_1)$ of the m elliptical populations, in a way that does not correspond to any sound decision-theoretic principle. Although optimal under Gaussian densities (where the weights happen to be equal) and asymptotically valid under non-Gaussian densities, this test is thus highly unsatisfactory. The same conclusion would apply to a bootstrapped version of $\phi^{(n)}_{\mathcal{N}}$.

7 Simulations.

We conducted several simulations in the bivariate case ($k = 2$), for various types of alternatives and various radial densities. More precisely, we generated three couples of mutually independent absolutely continuous bivariate random vectors

$$\boldsymbol{\varepsilon}_{1j},\ j = 1,\ldots,n_1 = 200, \quad \text{and} \quad \boldsymbol{\varepsilon}_{2j},\ j = 1,\ldots,n_2 = 50,$$

with spherical densities centered at $\mathbf{0}$. In the first case, all $\boldsymbol{\varepsilon}_{ij}$'s have standard multinormal densities. In the second one (heterokurtic case), the $\boldsymbol{\varepsilon}_{1j}$'s have standardized t_5 densities, whereas the $\boldsymbol{\varepsilon}_{2j}$'s are multinormal. In the third one (non-Gaussian homokurtic case), all $\boldsymbol{\varepsilon}_{ij}$'s have standardized t_5 densities (standardized t_ν here refers to the bivariate t_ν distribution with unit covariance matrix, that is, the distribution of $\mathbf{Z}/(Y/(\nu-2))^{1/2}$, where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2)$ and $Y \sim \chi^2_\nu$ are independent). Starting from these three couples of populations with homogeneous covariances, we considered independent samples obtained from

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

where $\mathbf{A}_2(\ell)\mathbf{A}_2'(\ell) = (1+\ell s^2)(\mathbf{A}_1\mathbf{A}_1' + \ell\mathbf{v})$ ($\mathbf{v}$ a symmetric $(k\times k)$ matrix with $\text{tr}((\mathbf{A}_1\mathbf{A}_1')^{-1}\mathbf{v}) = 0$), $\ell = 0, 1, \ldots, 20$. The values of ℓ allow to produce distributions under the null ($\ell = 0$) and increasingly heterogeneous alternatives ($\ell = 1,\ldots,20$); 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 $N = 10,000$ independent samples, with $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (.05, (0,0))$, $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (.06, (0,0))$, and $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (.08, (0,0))$ under Gaussian-Gaussian, t_5-Gaussian, and t_5-t_5 densities, respectively; these values of s^2 were chosen in order to obtain rejection probabilities of the same order under each couple of densities. In the second simulation (pure shape alternatives), we similarly generated $N = 10,000$ independent samples, with $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (0, (0, .03))$—irrespective of the underlying densities. In the last simulation ("mixed" alternatives), $N = 10,000$ independent samples were considered, with $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (.04, (0, .02))$, $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (.05, (0, .02))$, and $(s^2, (\overset{\circ}{\text{vech}}\,\mathbf{v})') = (.06, (0, .02))$ under Gaussian-Gaussian, t_5-Gaussian, and t_5-t_5 densities, respectively—still in order to obtain rejection probabilities of the same order under each couple of densities.

Figure 1 reports rejection frequencies as functions of ℓ for the following four tests: (a) the Gaussian LRT $\phi^{(n)}_{\text{LRT}}$, based on $Q^{(n)}_{\text{LRT}} = -2\log\Lambda$ (solid line), (b) the Gaussian most stringent test $\phi^{(n)}_{\mathcal{N}}$, based on $Q^{(n)}_{\mathcal{N}}$ (dashed line), and (c-d) its pseudo-Gaussian versions $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\mathcal{N}\dagger}$, based on $Q^{(n)}_{\mathcal{N}*}$ and $Q^{(n)}_{\mathcal{N}\dagger}$, respectively (dot-dash line and dotted line, respectively). Rejection frequencies for the following five tests are reported in Figure 2: (a) the pseudo-Gaussian test for covariance homogeneity against scale alternatives $\phi^{II(n)}_{\mathcal{N}\dagger}$, based on $Q^{II(n)}_{\mathcal{N}\dagger}$ (dotted line), (b) the pseudo-Gaussian test for scale homogeneity $\phi^{\text{scale}(n)}_{\mathcal{N}\dagger}$, based on $Q^{\text{scale}(n)}_{\mathcal{N}*}$ (dot-dash line), (c) the pseudo-Gaussian test for covariance homogeneity against shape alternatives $\phi^{III(n)}_{\mathcal{N}\dagger}$, based on $Q^{III(n)}_{\mathcal{N}\dagger}$ (dashed line), (d) the pseudo-Gaussian test for shape homogeneity $\phi^{\text{shape}(n)}_{\mathcal{N}\dagger}$, based on $Q^{\text{shape}(n)}_{\mathcal{N}\dagger}$ (long-dash line), and—in order to compare with pseudo-Gaussian tests from Figure 1—(e) the pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}\dagger}$, based on $Q^{(n)}_{\mathcal{N}\dagger}$ (solid line). All tests were performed at asymptotic level $\alpha = 5\%$.

Inspection of Figure 1 reveals that

(i) $\phi^{(n)}_{\text{LRT}}$ and $\phi^{(n)}_{\mathcal{N}}$ are valid when both densities are multinormal only. They both strongly overreject under t_5-Gaussian and t_5-t_5 densities, with Type I risks close to 11% and 30%, respectively. These tests exhibit quite similar performances in all setups, although, for the moderate sample sizes considered, $\phi^{(n)}_{\text{LRT}}$ seems to be less (resp., more) powerful than $\phi^{(n)}_{\mathcal{N}}$ against scale (resp., shape) alternatives;

(ii) the pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}*}$ has a Type I risk close to 5% at Gaussian-Gaussian and t_5-t_5 densities, that is, under homokurticity. In the heterokurtic case, however, it is highly biased, with a Type I risk close to 2.5%;

(iii) unlike $\phi^{(n)}_{\mathcal{N}*}$, the pseudo-Gaussian test $\phi^{(n)}_{\mathcal{N}\dagger}$ has a Type I risk close to $\alpha = 5\%$ in all cases. Quite unexpectedly, $\phi^{(n)}_{\mathcal{N}\dagger}$ seems to be more powerful than $\phi^{(n)}_{\mathcal{N}*}$, not only under heterokurtic alternatives (t_5-Gaussian densities), but also under non-Gaussian homokurtic ones (t_5-t_5 densities).

These conclusions somewhat contradict those of Schott (2001). Schott indeed concludes from his simulations that $\phi^{(n)}_{\text{Schott}*}$ is to be preferred to $\phi^{(n)}_{\text{Schott}\dagger}$, since the former has a simpler structure and, in most cases, behaves at least as well as the latter. In view of Theorem 5.3, this statement should extend to our optimal tests $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\mathcal{N}\dagger}$. However, this is not the case, since our simulations clearly show that $\phi^{(n)}_{\mathcal{N}*}$ (hence also $\phi^{(n)}_{\text{Schott}*}$) may be severely biased. Schott's invalid conclusions can be explained by the fact that his simulations are restricted to equal sample sizes, irrespective of the underlying densities. In such cases, it is not very surprising that $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\mathcal{N}\dagger}$ behave quite similarly, *even under heterokurtic densities*: indeed, the overall estimate $\hat{\kappa}^{(n)}_k$ of the common kurtosis parameter then can be expected to be quite close to the natural estimate for $\kappa^{(n)}_{k;i,i'}(g)$, which implies that the two-sample test statistics $Q^{g(n)}_{\mathcal{N}*;i,i'}$ in (5.5) and $Q^{g(n)}_{\mathcal{N}\dagger;i,i'}$ in (5.12) tend to take very similar values.

Finally, Figure 2 shows that (i) the pseudo-Gaussian tests $\phi^{\mathit{II}(n)}_{\mathcal{N}\dagger}$, $\phi^{\text{scale}(n)}_{\mathcal{N}\dagger}$, $\phi^{\mathit{III}(n)}_{\mathcal{N}\dagger}$, and $\phi^{\text{shape}(n)}_{\mathcal{N}\dagger}$ remain valid under all densities considered; (ii) $\phi^{\mathit{II}(n)}_{\mathcal{N}\dagger}$ and $\phi^{\text{scale}(n)}_{\mathcal{N}\dagger}$ are best against scale alternatives, $\phi^{\mathit{III}(n)}_{\mathcal{N}\dagger}$ and $\phi^{\text{shape}(n)}_{\mathcal{N}\dagger}$ against shape alternatives; as expected, $\phi^{(n)}_{\mathcal{N}\dagger}$ is dominated by $\phi^{\mathit{II}(n)}_{\mathcal{N}\dagger}$ under scale alternatives and by $\phi^{\mathit{III}(n)}_{\mathcal{N}\dagger}$ under shape alternatives (same noncentrality parameters, but more degrees of freedom); (iii) whereas $\phi^{\text{shape}(n)}_{\mathcal{N}\dagger}$ is more powerful than $\phi^{\mathit{III}(n)}_{\mathcal{N}\dagger}$, $\phi^{\text{scale}(n)}_{\mathcal{N}\dagger}$ quite unexpectedly seems to be roughly equivalent to $\phi^{\mathit{II}(n)}_{\mathcal{N}\dagger}$, at the sample sizes considered.

A Appendix.

A.1 Proof of Lemma 5.2

In this section, we prove Lemma 5.2, that is, the asymptotic linearity with respect to estimated values $\hat{\boldsymbol{\vartheta}}_{I} := (\hat{\boldsymbol{\theta}}'_1, \ldots, \hat{\boldsymbol{\theta}}'_m)'$, $\hat{\boldsymbol{\vartheta}}_{II} := (\hat{\sigma}_1, \ldots, \hat{\sigma}_m)'$, and $\hat{\boldsymbol{\vartheta}}_{III} := ((\text{ve}\mathring{\text{c}}\text{h}\, \hat{\mathbf{V}}_1)', \ldots, (\text{ve}\mathring{\text{c}}\text{h}\, \hat{\mathbf{V}}_m)')'$, of $\boldsymbol{\vartheta}_{I}$, $\boldsymbol{\vartheta}_{II}$, and $\boldsymbol{\vartheta}_{III}$, of the Gaussian central sequences for scale and shape $\boldsymbol{\Delta}^{\mathit{II}}_{\hat{\boldsymbol{\vartheta}};\phi}$ and $\boldsymbol{\Delta}^{\mathit{III}}_{\hat{\boldsymbol{\vartheta}};\phi}$, when the Gaussian estimators $\hat{\boldsymbol{\theta}}_i = \bar{\mathbf{X}}_i$, $\hat{\sigma}_i = |\mathbf{S}_i|^{1/2k}$, and $\hat{\mathbf{V}}_i = \mathbf{S}_i/|\mathbf{S}_i|^{1/k}$ are used; the estimator $\hat{\boldsymbol{\Sigma}}_i$ of $\boldsymbol{\Sigma}_i$ is thus the empirical covariance matrix $\mathbf{S}_i$. This notation is used throughout the proof.

Proof of Lemma 5.2. *(i)* First note that, for any $1 \leq i \leq m$,

$$\begin{aligned}
\boldsymbol{\Delta}^{\mathit{II},i}_{\boldsymbol{\vartheta};\phi} &= \frac{n_i^{-1/2}}{2\sigma_i^2} \sum_{j=1}^{n_i} \Big(\frac{d^2_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)}{\sigma_i^2} - k \Big) = \frac{n_i^{-1/2}}{2\sigma_i^2} \sum_{j=1}^{n_i} \Big((\mathbf{X}_{ij} - \boldsymbol{\theta}_i)' \boldsymbol{\Sigma}_i^{-1} (\mathbf{X}_{ij} - \boldsymbol{\theta}_i) - k \Big) \\
&= \frac{n_i^{-1/2}}{2\sigma_i^2} \sum_{j=1}^{n_i} \text{tr}\Big[\boldsymbol{\Sigma}_i^{-1} \Big((\mathbf{X}_{ij} - \boldsymbol{\theta}_i)(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)' - \boldsymbol{\Sigma}_i \Big) \Big] \\
&= \frac{n_i^{-1/2}}{2\sigma_i^2} \Big(\text{vec}\, \boldsymbol{\Sigma}_i^{-1} \Big)' \sum_{j=1}^{n_i} \text{vec}\Big((\mathbf{X}_{ij} - \boldsymbol{\theta}_i)(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)' - \boldsymbol{\Sigma}_i \Big), \qquad \text{(A.1)}
\end{aligned}$$

where we used the fact that $\text{tr}(\mathbf{AB}) = (\text{vec}\, \mathbf{A})'(\text{vec}\, \mathbf{B})$. Note then that

$$\begin{aligned}
\boldsymbol{\Delta}^{\mathit{II},i}_{\hat{\boldsymbol{\vartheta}};\phi} &= \frac{n_i^{-1/2}}{2\hat{\sigma}_i^2} \Big(\text{vec}\, \hat{\boldsymbol{\Sigma}}_i^{-1} \Big)' \sum_{j=1}^{n_i} \text{vec}\Big((\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)' - \hat{\boldsymbol{\Sigma}}_i \Big) \\
&= \frac{1}{2\hat{\sigma}_i^2} \Big(\text{vec}\, \hat{\boldsymbol{\Sigma}}_i^{-1} \Big)' \Big\{ n_i^{-1/2} \sum_{j=1}^{n_i} \text{vec}\Big((\mathbf{X}_{ij} - \boldsymbol{\theta}_i)(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)' - \boldsymbol{\Sigma}_i \Big) \\
&\qquad + n_i^{1/2} \text{vec}((\bar{\mathbf{X}}_i - \boldsymbol{\theta}_i)(\bar{\mathbf{X}}_i - \boldsymbol{\theta}_i)') \Big\} - \frac{1}{2\hat{\sigma}_i^2} \Big(\text{vec}\, \hat{\boldsymbol{\Sigma}}_i^{-1} \Big)' n_i^{1/2} \text{vec}\Big(\hat{\boldsymbol{\Sigma}}_i - \boldsymbol{\Sigma}_i \Big).
\end{aligned}$$

Applying repeatedly Slutzky's Lemma and taking into account the fact that $(\text{vec}\, \mathbf{V}_i^{-1})'(\text{vec}\, \mathbf{V}_i) =$

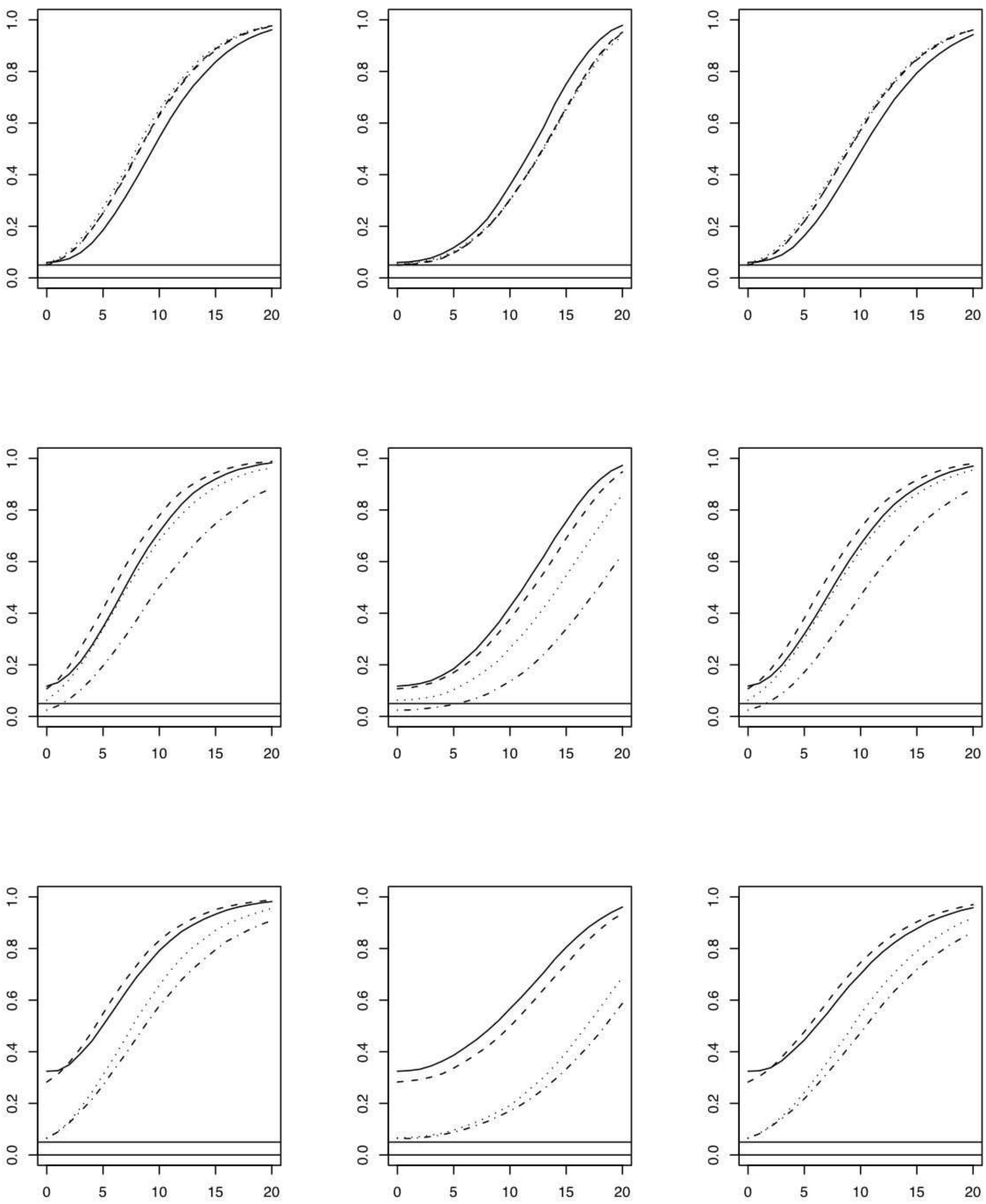

Figure 1: Rejection frequencies (out of $N = 10,000$ replications), under the null and under various scale (first column), shape (second column) and "mixed" (third column) alternatives (see Section 7 for details), of the Gaussian LRT $\phi^{(n)}_{\mathrm{LRT}}$ (solid line), the parametric Gaussian test $\phi^{(n)}_{\mathcal{N}}$ (dashed line), and its pseudo-Gaussian versions $\phi^{(n)}_{\mathcal{N}*}$ and $\phi^{(n)}_{\mathcal{N}\dagger}$ (dot-dash line and dotted line, respectively); in each case, the asymptotic nominal size α is 5%. There are $m = 2$ bivariate populations, with sample sizes $n_1 = 200$ and $n_2 = 50$; radial densities are Gaussian-Gaussian (first row), t_5-Gaussian (second row), and t_5-t_5 (third row), respectively.

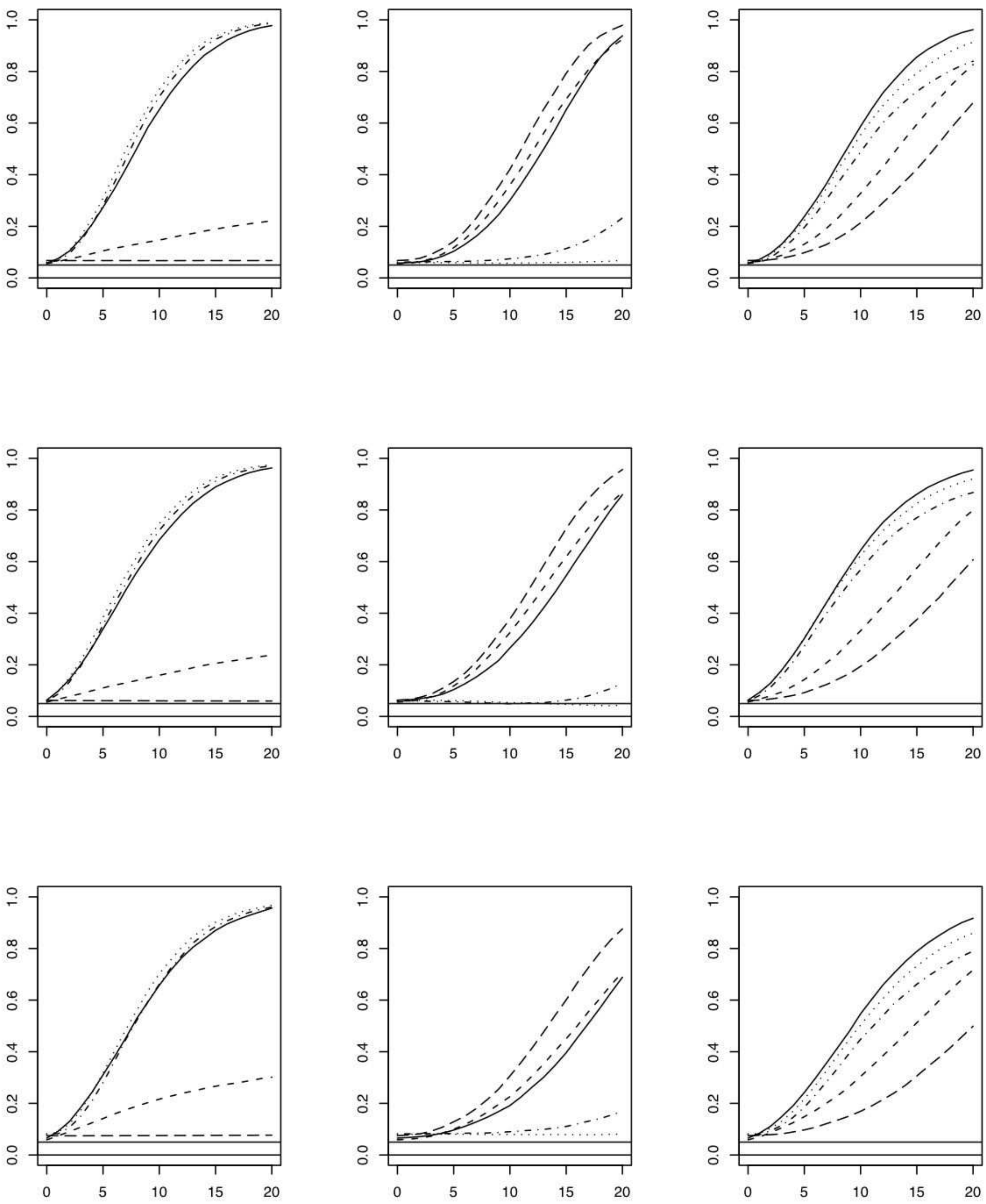

Figure 2: Rejection frequencies (out of $N = 10,000$ replications), under the null and under various scale (first column), shape (second column) and "mixed" (third column) alternatives (see Section 7 for details), of the pseudo-Gaussian tests for scale $\phi_{\mathcal{N}\dagger}^{II(n)}$ and $\phi_{\mathcal{N}\dagger}^{\mathrm{scale}(n)}$ (dotted line and dot-dash line, respectively), the pseudo-Gaussian tests for shape $\phi_{\mathcal{N}\dagger}^{III(n)}$ and $\phi_{\mathcal{N}\dagger}^{\mathrm{shape}(n)}$ (dashed line and long-dash line, respectively), and the pseudo-Gaussian test $\phi_{\mathcal{N}\dagger}^{(n)}$ (solid line); in each case, the asymptotic nominal size α is 5%. There are $m = 2$ bivariate populations, with sample sizes $n_1 = 200$ and $n_2 = 50$; radial densities are Gaussian-Gaussian (first row), t_5-Gaussian (second row), and t_5-t_5 (third row), respectively.

$\mathrm{tr}(\mathbf{V}_i^{-1}\mathbf{V}_i) = k$ yields

$$\begin{aligned}
\boldsymbol{\Delta}^{II,i}_{\hat{\boldsymbol{\vartheta}};\phi} &= \frac{1}{2\sigma_i^2}\Big(\mathrm{vec}\,\boldsymbol{\Sigma}_i^{-1}\Big)'\Big[n_i^{-1/2}\sum_{j=1}^{n_i}\mathrm{vec}\Big((\mathbf{X}_{ij}-\boldsymbol{\theta}_i)(\mathbf{X}_{ij}-\boldsymbol{\theta}_i)'-\boldsymbol{\Sigma}_i\Big)\Big] \\
&\qquad -\frac{1}{2\sigma_i^2}\Big(\mathrm{vec}\,\boldsymbol{\Sigma}_i^{-1}\Big)' n_i^{1/2}\mathrm{vec}\Big(\hat{\boldsymbol{\Sigma}}_i-\boldsymbol{\Sigma}_i\Big)+o_{\mathrm{P}}(1) \\
&= \boldsymbol{\Delta}^{II,i}_{\boldsymbol{\vartheta};\phi} - \frac{1}{2\sigma_i^4}\Big(\mathrm{vec}\,\mathbf{V}_i^{-1}\Big)' n_i^{1/2}\Big[(\hat{\sigma}_i^2-\sigma_i^2)\,\mathrm{vec}\,\hat{\mathbf{V}}_i+\sigma_i^2\,\mathrm{vec}(\hat{\mathbf{V}}_i-\mathbf{V}_i)\Big]+o_{\mathrm{P}}(1) \\
&= \boldsymbol{\Delta}^{II,i}_{\boldsymbol{\vartheta};\phi} - \frac{k}{2\sigma_i^4}\,n_i^{1/2}(\hat{\sigma}_i^2-\sigma_i^2) - \frac{1}{2\sigma_i^2}\Big(\mathrm{vec}\,\mathbf{V}_i^{-1}\Big)' n_i^{1/2}\mathrm{vec}(\hat{\mathbf{V}}_i-\mathbf{V}_i)+o_{\mathrm{P}}(1).
\end{aligned}$$

The result then follows by using successively $\mathrm{vec}(\hat{\mathbf{V}}_i-\mathbf{V}_i) = (\mathbf{M}_k(\mathbf{V}_i))'\,\overset{\circ}{\mathrm{vech}}\,(\hat{\mathbf{V}}_i-\mathbf{V}_i)+o_{\mathrm{P}}(n_i^{-1/2})$ (see Lemma 5.1 of Hallin and Paindaveine 2006b) and $\mathbf{M}_k(\mathbf{V}_i)\mathrm{vec}(\mathbf{V}_i^{-1}) = \mathbf{0}$.

(ii) For any $1 \leq i \leq m$, we have

$$\begin{aligned}
\boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\phi} &= \frac{n_i^{-1/2}}{2}\,\mathbf{M}_k(\mathbf{V}_i)\Big(\mathbf{V}_i^{\otimes 2}\Big)^{-1/2}\sum_{j=1}^{n_i}\frac{d_{ij}^2(\boldsymbol{\theta}_i,\mathbf{V}_i)}{\sigma_i^2}\mathrm{vec}(\mathbf{U}_{ij}(\boldsymbol{\theta}_i,\mathbf{V}_i)\mathbf{U}_{ij}'(\boldsymbol{\theta}_i,\mathbf{V}_i)) \\
&= \frac{n_i^{-1/2}}{2\sigma_i^2}\,\mathbf{M}_k(\mathbf{V}_i)\Big(\mathbf{V}_i^{\otimes 2}\Big)^{-1}\sum_{j=1}^{n_i}\mathrm{vec}\Big((\mathbf{X}_{ij}-\boldsymbol{\theta}_i)(\mathbf{X}_{ij}-\boldsymbol{\theta}_i)'-\boldsymbol{\Sigma}_i\Big), \qquad \text{(A.2)}
\end{aligned}$$

since $\mathbf{M}_k(\mathbf{V}_i)(\mathbf{V}_i^{\otimes 2})^{-1}\mathrm{vec}\,\mathbf{V}_i = \mathbf{M}_k(\mathbf{V}_i)\mathrm{vec}(\mathbf{V}_i^{-1}) = \mathbf{0}$. Hence, still from Slutzky's Lemma,

$$\begin{aligned}
\boldsymbol{\Delta}^{III,i}_{\hat{\boldsymbol{\vartheta}};\phi} &= \frac{n_i^{-1/2}}{2\hat{\sigma}_i^2}\,\mathbf{M}_k(\hat{\mathbf{V}}_i)\Big(\hat{\mathbf{V}}_i^{\otimes 2}\Big)^{-1}\sum_{j=1}^{n_i}\mathrm{vec}\Big((\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)'-\hat{\boldsymbol{\Sigma}}_i\Big) \\
&= \frac{n_i^{-1/2}}{2\hat{\sigma}_i^2}\,\mathbf{M}_k(\hat{\mathbf{V}}_i)(\hat{\mathbf{V}}_i^{\otimes 2})^{-1}\Big\{\sum_{j=1}^{n_i}\mathrm{vec}\Big((\mathbf{X}_{ij}-\boldsymbol{\theta}_i)(\mathbf{X}_{ij}-\boldsymbol{\theta}_i)'-\boldsymbol{\Sigma}_i\Big) \\
&\qquad + n_i\,\mathrm{vec}\Big((\bar{\mathbf{X}}_i-\boldsymbol{\theta}_i)(\bar{\mathbf{X}}_i-\boldsymbol{\theta}_i)'\Big) - n_i\,\mathrm{vec}(\hat{\boldsymbol{\Sigma}}_i-\boldsymbol{\Sigma}_i)\Big\} \\
&= \boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\phi} - \frac{1}{2}\sigma_i^2\,\mathbf{M}_k(\mathbf{V}_i)\Big(\mathbf{V}_i^{\otimes 2}\Big)^{-1} n_i^{1/2}\,\mathrm{vec}(\hat{\boldsymbol{\Sigma}}_i-\boldsymbol{\Sigma}_i)+o_{\mathrm{P}}(1),
\end{aligned}$$

Now, writing $\hat{\boldsymbol{\Sigma}}_i-\boldsymbol{\Sigma}_i$ under the form $(\hat{\sigma}_i^2-\sigma_i^2)\hat{\mathbf{V}}_i+\sigma_i^2(\hat{\mathbf{V}}_i-\mathbf{V}_i)$, applying Slutzky's Lemma again and using the fact that $\mathbf{M}_k(\mathbf{V}_i)(\mathbf{V}_i^{\otimes 2})^{-1}\mathrm{vec}\mathbf{V}_i = \mathbf{0}$ and $\mathbf{K}(\mathrm{vec}\,\mathbf{A}) = (\mathrm{vec}\,\mathbf{A}')$, we obtain

$$\begin{aligned}
\boldsymbol{\Delta}^{III,i}_{\hat{\boldsymbol{\vartheta}};\phi} &= \boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\phi} - \frac{1}{2}\mathbf{M}_k(\mathbf{V}_i)\Big(\mathbf{V}_i^{\otimes 2}\Big)^{-1} n_i^{1/2}\,\mathrm{vec}(\hat{\mathbf{V}}_i-\mathbf{V}_i)+o_{\mathrm{P}}(1) \\
&= \boldsymbol{\Delta}^{III,i}_{\boldsymbol{\vartheta};\phi} - \frac{1}{4}\mathbf{M}_k(\mathbf{V}_i)\Big(\mathbf{V}_i^{\otimes 2}\Big)^{-1}[\mathbf{I}_{k^2}+\mathbf{K}_k]\,n_i^{1/2}\,\mathrm{vec}(\hat{\mathbf{V}}_i-\mathbf{V}_i)+o_{\mathrm{P}}(1).
\end{aligned}$$

The desired result then follows from the fact that $\mathbf{K}_k(\mathbf{A}\otimes\mathbf{B}) = (\mathbf{B}\otimes\mathbf{A})\mathbf{K}_k$ and the definition of $\mathbf{M}_k(\mathbf{V})$ (see Section 4.1). □

A.2 Proofs of Lemma 5.1 and Theorems 5.1 and 5.2.

Proof of Lemma 5.1. The result straightforwardly follows, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$ with $\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})$, from the multivariate Central Limit Theorem. The result under local alternatives is obtained as usual, by establishing the joint normality under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$ of $((\boldsymbol{\Delta}^{\mathit{II}}_{\boldsymbol{\vartheta};\phi})', (\boldsymbol{\Delta}^{\mathit{III}}_{\boldsymbol{\vartheta};\phi})')'$ and $\Lambda^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}/\boldsymbol{\vartheta};g}$, 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)* The consistency of $\hat{\kappa}_k$, the continuity of the mapping $\boldsymbol{\vartheta} \mapsto (\mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi*}, \mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi*})$, Lemma 5.2 (jointly with Assumption (C1)), and the facts that $\mathbf{P}^{g,\mathit{II}}_{\hat{\boldsymbol{\vartheta}};\phi*}(\boldsymbol{\nu}^{(n)}_{\mathit{II}})^{-1}\boldsymbol{\Upsilon}_{\mathit{II}} = \mathbf{0}$ and $\mathbf{P}^{g,\mathit{III}}_{\hat{\boldsymbol{\vartheta}};\phi*}[\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})](\boldsymbol{\nu}^{(n)}_{\mathit{III}})^{-1}\boldsymbol{\Upsilon}_{\mathit{III}} = \mathbf{0}$, entail

$$Q^{(n)}_{\mathcal{N}*} = (\boldsymbol{\Delta}^{\mathit{II}}_{\boldsymbol{\vartheta};\phi})' \, \mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi*} \, \boldsymbol{\Delta}^{\mathit{II}}_{\boldsymbol{\vartheta};\phi} + (\boldsymbol{\Delta}^{\mathit{III}}_{\boldsymbol{\vartheta};\phi})' \, \mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi*} \, \boldsymbol{\Delta}^{\mathit{III}}_{\boldsymbol{\vartheta};\phi} + o_{\mathrm{P}}(1) \tag{A.3}$$

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

Now, since $(\boldsymbol{\Gamma}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi})^{1/2} \, \mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi*} (\boldsymbol{\Gamma}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi})^{1/2}$ is a symmetric idempotent matrix with rank $(m-1)$, Lemma 5.1 ensures that the first term in (A.3) is asymptotically chi-square with $(m-1)$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and asymptotically noncentral chi-square, still with $(m-1)$ degrees of freedom but with noncentrality parameter

$$\left(\frac{k}{2\sigma^4}\right)^2 \lim_{n\to\infty} \left\{ (\boldsymbol{\tau}^{(n)}_{\mathit{II}})' \mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi*} \, \boldsymbol{\tau}^{(n)}_{\mathit{II}} \right\} \tag{A.4}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g}$. Evaluation of (A.4) yields the first term in (5.8).

As for the shape part, using again Lemma 5.1 and the fact that $(\boldsymbol{\Gamma}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi})^{1/2} \, \mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi*} (\boldsymbol{\Gamma}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi})^{1/2}$ is symmetric and idempotent with rank $k_0(m-1)$, we obtain similarly that the second term in (A.3) is asymptotically chi-square with $k_0(m-1)$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$, $\boldsymbol{\vartheta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, and asymptotically noncentral chi-square, still with $k_0(m-1)$ degrees of freedom but with noncentrality parameter

$$k^2(k+2)^2 \lim_{n\to\infty} \left\{ (\boldsymbol{\tau}^{(n)}_{\mathit{III}})' [\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})] \mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi*} [\mathbf{I}_m \otimes \mathbf{H}_k(\mathbf{V})] \, \boldsymbol{\tau}^{(n)}_{\mathit{III}} \right\} \tag{A.5}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g}$. A straightforward evaluation of (A.5) yields the second term in (5.8). Since the two terms in (A.3) 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 size 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 Section 5.1, the consistency of $\hat{\kappa}^{(n)}_k$ entails the asymptotic equivalence, under Gaussian densities, of $Q^{(n)}_{\mathcal{N}*}$ with $Q^{(n)}_{\mathcal{N}}$, which has been derived from the general form of locally asymptotically optimal tests based on (4.3). □

Note that (A.3) also holds under heterokurticity, that is, at any $g \in (\mathcal{F}^4_1)^m$. However, at heterokurtic g, neither $(\boldsymbol{\Gamma}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi})^{1/2} \, \mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi*} (\boldsymbol{\Gamma}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi})^{1/2}$ nor $(\boldsymbol{\Gamma}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi})^{1/2} \, \mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi*} (\boldsymbol{\Gamma}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi})^{1/2}$ are idempotent, so that $Q^{(n)}_{\mathcal{N}*}$ is *not* asymptotically chi-square under the null (see, e.g., Rao and Mitra 1971, Theorem 9.2.1). Consequently, $\phi^{(n)}_{\mathcal{N}*}$ is not asymptotically distribution-free under $\cup_{\boldsymbol{\vartheta}\in\mathcal{M}(\boldsymbol{\Upsilon})} \cup_{g\in(\mathcal{F}^4_1)^m} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}\}$.

Proof of Theorem 5.2. The proof proceeds exactly along the same lines as in that of Theorem 5.1, by using the consistency (under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};g}$, $g \in (\mathcal{F}^4_1)^m$) of the estimators $\hat{E}_{k,i}$ and $\hat{C}_{k,i}$, $i =$

$1,\ldots,m$, the continuity of the mapping $\boldsymbol{\vartheta}\mapsto(\mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi\dagger},\mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi\dagger})$, the facts that $\mathbf{P}^{g,\mathit{II}}_{\hat{\boldsymbol{\vartheta}};\phi\dagger}(\boldsymbol{\nu}^{(n)}_{\mathit{II}})^{-1}\boldsymbol{\Upsilon}_{\mathit{II}}=\mathbf{0}$ and $\mathbf{P}^{g,\mathit{III}}_{\hat{\boldsymbol{\vartheta}};\phi\dagger}[\mathbf{I}_m\otimes\mathbf{H}_k(\mathbf{V})](\boldsymbol{\nu}^{(n)}_{\mathit{III}})^{-1}\boldsymbol{\Upsilon}_{\mathit{III}}=\mathbf{0}$, and that $(\boldsymbol{\Gamma}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi})^{1/2}\,\mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi\dagger}(\boldsymbol{\Gamma}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi})^{1/2}$ and $(\boldsymbol{\Gamma}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi})^{1/2}\,\mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi\dagger}(\boldsymbol{\Gamma}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi})^{1/2}$ are symmetric idempotent matrices with rank $m-1$ and $k_0(m-1)$, respectively. The resulting noncentrality parameter in the asymptotic chi-square distribution of $Q^{(n)}_{\mathcal{N}\dagger}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)};g}$ is given by

$$\left(\frac{k}{2\sigma^4}\right)^2\lim_{n\to\infty}\left\{(\boldsymbol{\tau}^{(n)}_{\mathit{II}})'\mathbf{P}^{g,\mathit{II}}_{\boldsymbol{\vartheta};\phi\dagger}\,\boldsymbol{\tau}^{(n)}_{\mathit{II}}\right\}+k^2(k+2)^2\lim_{n\to\infty}\left\{(\boldsymbol{\tau}^{(n)}_{\mathit{III}})'[\mathbf{I}_m\otimes\mathbf{H}_k(\mathbf{V})]\mathbf{P}^{g,\mathit{III}}_{\boldsymbol{\vartheta};\phi\dagger}[\mathbf{I}_m\otimes\mathbf{H}_k(\mathbf{V})]\,\boldsymbol{\tau}^{(n)}_{\mathit{III}}\right\}.$$

A direct evaluation of this expression yields the noncentrality parameter in (5.13). □

A.3 Proof of Theorems 5.3 and 5.4.

Proof of Theorem 5.3. From (5.10), we see that $Q^{g(n)}_{\mathcal{N}\dagger}=Q^{g(n),1}_{\mathcal{N}\dagger}+Q^{g(n),2}_{\mathcal{N}\dagger}$, where

$$Q^{g(n),1}_{\mathcal{N}\dagger}=n\sum_{1\le i<i'\le m}\lambda^{(n)}_i\lambda^{(n)}_{i'}\,\frac{k(k+2)\tilde{E}^{(n)}_k(g)}{2E_k(g_i)E_k(g_{i'})}\,\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'}))^2\right]$$

and

$$Q^{g(n),2}_{\mathcal{N}\dagger}=n\sum_{1\le i<i'\le m}\lambda^{(n)}_i\lambda^{(n)}_{i'}\left\{\frac{\tilde{C}^{(n)}_k(g)}{C_k(g_i)C_k(g_{i'})}-\frac{(k+2)\tilde{E}^{(n)}_k(g)}{2E_k(g_i)E_k(g_{i'})}\right\}\mathrm{tr}^2\left[\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'})\right].$$

Now, define

$$\alpha_i:=\frac{\lambda^{(n)}_i}{2(1+\kappa_k(g_i))},\ \beta_i:=\frac{-\lambda^{(n)}_i\kappa_k(g_i)}{2(1+\kappa_k(g_i))((k+2)\kappa_k(g_i)+2)},\ \alpha:=\sum_{i=1}^m\alpha_i,\ \beta:=\sum_{i=1}^m\beta_i,$$

$\rho:=-\beta/(\alpha(\alpha+k\beta))$, and $\tau_{i,i'}:=\alpha^{-1}\alpha_i\beta_j+(\alpha_i\rho+\alpha^{-1}\beta_i+k\beta_i\rho)(\alpha_{i'}+k\beta_{i'})$. We then have $\alpha=(k(k+2)/2)\sum_{i=1}^m\lambda^{(n)}_i(E_k(g_i))^{-1}=k(k+2)/(2\tilde{E}_k(g))$, so that

$$\begin{aligned}Q^{g(n),1}_{\mathcal{N}\dagger}&=\frac{n}{\alpha}\sum_{1\le i<i'\le m}\alpha_i\alpha_{i'}\,\mathrm{tr}\left[(\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'}))^2\right]\\&=\frac{n}{\alpha}\sum_{i,i'=1}^m\alpha_i\alpha_{i'}\left[\mathrm{tr}\big[(\mathbf{S}_i\mathbf{S}^{-1})^2\big]-\mathrm{tr}\big[\mathbf{S}_i\mathbf{S}^{-1}\mathbf{S}_{i'}\mathbf{S}^{-1}\big]\right]\\&=n\left\{\sum_{i=1}^m\alpha_i\,\mathrm{tr}\left[(\mathbf{S}_i\mathbf{S}^{-1})^2\right]-\sum_{i,i'=1}^m(\alpha_i\alpha_{i'}/\alpha)\,\mathrm{tr}\big[\mathbf{S}_i\mathbf{S}^{-1}\mathbf{S}_{i'}\mathbf{S}^{-1}\big]\right\}.\end{aligned}$$

Similarly, we obtain

$$
\begin{aligned}
Q^{g(n),2}_{\mathcal{N}\dagger} &= n \sum_{1\le i<i'\le m} \lambda_i^{(n)}\lambda_{i'}^{(n)} \left\{ \frac{\tilde{C}_k^{(n)}(g)}{C_k(g_i)C_k(g_{i'})} - \frac{(k+2)\tilde{E}_k^{(n)}(g)}{2E_k(g_i)E_k(g_{i'})} \right\} \mathrm{tr}^2\left(\mathbf{S}^{-1}(\mathbf{S}_i-\mathbf{S}_{i'})\right) \\
&= n \sum_{i,i'=1}^{m} \lambda_i^{(n)}\lambda_{i'}^{(n)} \left\{ \frac{\tilde{C}_k^{(n)}(g)}{C_k(g_i)C_k(g_{i'})} - \frac{(k+2)\tilde{E}_k^{(n)}(g)}{2E_k(g_i)E_k(g_{i'})} \right\} \Big[\mathrm{tr}^2[\mathbf{S}_i\mathbf{S}^{-1}] - \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}]\,\mathrm{tr}[\mathbf{S}_{i'}\mathbf{S}^{-1}]\Big], \\
&= n \sum_{i=1}^{m} \lambda_i^{(n)} \left\{ \frac{1}{C_k(g_i)} - \frac{k+2}{2E_k(g_i)} \right\} \mathrm{tr}^2[\mathbf{S}_i\mathbf{S}^{-1}] \\
&\quad - n \sum_{i,i'=1}^{m} \lambda_i^{(n)}\lambda_{i'}^{(n)} \left\{ \frac{\tilde{C}_k^{(n)}(g)}{C_k(g_i)C_k(g_{i'})} - \frac{(k+2)\tilde{E}_k^{(n)}(g)}{2E_k(g_i)E_k(g_{i'})} \right\} \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}]\,\mathrm{tr}[\mathbf{S}_{i'}\mathbf{S}^{-1}],
\end{aligned}
$$

which, after tedious yet straightforward computations, yields

$$
Q^{g(n),2}_{\mathcal{N}\dagger} = n \left\{ \sum_{i=1}^{m} \beta_i\, \mathrm{tr}^2\left[\mathbf{S}_i\mathbf{S}^{-1}\right] - \sum_{i,i'=1}^{m} \tau_{i,i'}\, \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}]\mathrm{tr}[\mathbf{S}_{i'}\mathbf{S}^{-1}] \right\}.
$$

Summing up, we have shown that

$$
\begin{aligned}
Q^{g(n)}_{\mathcal{N}\dagger} = n &\left\{ \sum_{i=1}^{m} \alpha_i\, \mathrm{tr}\left[(\mathbf{S}_i\mathbf{S}^{-1})^2\right] - \sum_{i,i'=1}^{m} (\alpha_i\alpha_{i'}/\alpha)\, \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}\mathbf{S}_{i'}\mathbf{S}^{-1}] \right\} \\
&+ n \left\{ \sum_{i=1}^{m} \beta_i\, \mathrm{tr}^2\left[\mathbf{S}_i\mathbf{S}^{-1}\right] - \sum_{i,i'=1}^{m} \tau_{i,i'}\, \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}]\mathrm{tr}[\mathbf{S}_{i'}\mathbf{S}^{-1}] \right\},
\end{aligned} \tag{A.6}
$$

which, after due substitution of consistent estimates for all g-related quantities, establishes the result for $\phi^{(n)}_{\mathrm{Schott}\dagger}$ (compare with (1.8)).

Under homokurticity (i.e., at any $g \in (\mathcal{F}_1^4)^m_{\mathrm{homo}}$), $Q^{g(n)}_{\mathcal{N}\dagger} = Q^{g(n)}_{\mathcal{N}*}$ and (A.6) becomes

$$
\begin{aligned}
Q^{g(n)}_{\mathcal{N}*} = &\frac{n}{2(1+\kappa_k(g_1))} \left\{ \sum_{i=1}^{m} \lambda_i^{(n)}\, \mathrm{tr}\left[(\mathbf{S}_i\mathbf{S}^{-1})^2\right] - \sum_{i,i'=1}^{m} \lambda_i^{(n)}\lambda_{i'}^{(n)}\, \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}\mathbf{S}_{i'}\mathbf{S}^{-1}] \right\} \\
&- \frac{n\kappa_k(g_1)}{2(1+\kappa_k(g_1))((k+2)\kappa_k(g_1)+2)} \left\{ \sum_{i=1}^{m} \lambda_i^{(n)}\, \mathrm{tr}^2\left[\mathbf{S}_i\mathbf{S}^{-1}\right] - \sum_{i,i'=1}^{m} \lambda_i^{(n)}\lambda_{i'}^{(n)}\, \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}]\mathrm{tr}[\mathbf{S}_{i'}\mathbf{S}^{-1}] \right\},
\end{aligned} \tag{A.7}
$$

which establishes the result for $\phi^{(n)}_{\mathrm{Schott}*}$ (compare with (1.7)).

Finally, the result for $\phi^{(n)}_{\mathrm{Schott}}$ follows by noting that, under multinormality (i.e., at $g = (\phi,\ldots,\phi)$), (A.7) reduces to

$$
Q^{(n)}_{\mathcal{N}} = \frac{n}{2} \left\{ \sum_{i=1}^{m} \lambda_i^{(n)}\, \mathrm{tr}\left[(\mathbf{S}_i\mathbf{S}^{-1})^2\right] - \sum_{i,j=1}^{m} \lambda_i^{(n)}\lambda_{i'}^{(n)}\, \mathrm{tr}[\mathbf{S}_i\mathbf{S}^{-1}\mathbf{S}_j\mathbf{S}^{-1}] \right\}, \tag{A.8}
$$

then comparing (A.8) with (1.6). □

Proof of Theorem 5.4. (i) We proceed by showing that the various test statistics are asymptotically equivalent to

$$Q_{\mathcal{N}}^{(n)} = \frac{1}{2n} \sum_{1\leq i<i'\leq m} n_i n_{i'} \, \mathrm{tr}\Big[(\mathbf{S}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\Big]; \qquad \text{(A.9)}$$

see (4.13). The asymptotic equivalence of $Q_{\text{Schott}}^{(n)}$ and $Q_{\mathcal{N}}^{(n)}$ directly follows from Theorem 5.3. Now, for the Nagao (1973) test, we have that

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

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

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

which establishes that $Q_{\text{Nagao}}^{(n)} = Q_{\mathcal{N}}^{(n)} + o_{\mathrm{P}}(1)$, as $n \to \infty$. As for the LRT (equivalently, the MLRT) test statistic, letting $\mathbf{\Sigma}^{1/2}\mathbf{Z}_i\mathbf{\Sigma}^{1/2} := n_i^{1/2}(\mathbf{S}_i - \mathbf{\Sigma}) =: n_i^{1/2}\mathbf{\Sigma}\mathbf{Y}_i$ (here, $\mathbf{\Sigma}$ stands for the common null value of the various covariance matrices), 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

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

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

$$\begin{aligned} &\frac{1}{2}\left\{\sum_{i=1}^m n_i \, \mathrm{tr}\,[\mathbf{Y}_i^2] - \frac{1}{n}\mathrm{tr}\left[\Big(\sum_{i=1}^m n_i\mathbf{Y}_i\Big)^2\right]\right\} = \frac{1}{2}\sum_{i=1}^m n_i \, \mathrm{tr}\left[\Big(\mathbf{Y}_i - \Big(\frac{1}{n}\sum_{r=1}^m n_r\mathbf{Y}_r\Big)\Big)^2\right] \\ &= \frac{1}{2}\sum_{i=1}^m n_i \, \mathrm{tr}\left[(\mathbf{\Sigma}^{-1}\,(\mathbf{S}_i - \mathbf{S}))^2\right] = \frac{1}{2}\sum_{i=1}^m n_i \, \mathrm{tr}\left[(\mathbf{S}_i\mathbf{\Sigma}^{-1} - \mathbf{I}_k)^2\right] + o_{\mathrm{P}}(1) = Q_{\text{Nagao}}^{(n)} + o_{\mathrm{P}}(1), \end{aligned}$$

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

(ii) Fix $g \in (\mathcal{F}_1^4)_{\text{homo}}^m$ and $\boldsymbol{\vartheta} = (\boldsymbol{\theta}_1', \ldots, \boldsymbol{\theta}_m', \sigma^2\mathbf{1}_m', \mathbf{1}_m' \otimes (\overset{\circ}{\mathrm{vech}}\,\mathbf{V})')' \in \mathcal{M}(\mathbf{\Upsilon})$, and write again $\mathbf{\Sigma} = \sigma^2\mathbf{V}$ for the common value of the various covariance matrices under $\mathrm{P}_{\boldsymbol{\vartheta};g}^{(n)}$. Then, by working as in the proof of Lemma 5.2, we obtain that

$$\begin{aligned} \mathrm{vec}\,\mathbf{Z}_i &= n_i^{1/2}\Big(\mathbf{\Sigma}^{\otimes 2}\Big)^{-1/2}\mathrm{vec}\,(\mathbf{S}_i - \mathbf{\Sigma}) \\ &= n_i^{-1/2}\Big(\mathbf{\Sigma}^{\otimes 2}\Big)^{-1/2}\sum_{j=1}^{n_i}\mathrm{vec}\Big((\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)' - \mathbf{\Sigma}\Big) \\ &= n_i^{-1/2}\Big(\mathbf{\Sigma}^{\otimes 2}\Big)^{-1/2}\sum_{j=1}^{n_i}\mathrm{vec}\Big((\mathbf{X}_{ij} - \boldsymbol{\theta}_i)(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)' - \mathbf{\Sigma}\Big) + o_{\mathrm{P}}(1), \end{aligned}$$

as $n \to \infty$. Letting $\mathbf{D}^{(n)} := ((\mathrm{vec}\,\mathbf{Z}_1)', \ldots, (\mathrm{vec}\,\mathbf{Z}_m)')'$, it then follows from (A.1) and (A.2) that

$$\mathbf{D}^{II(n)} := \frac{1}{2\sigma^2}\Big[\mathbf{I}_m \otimes (\mathrm{vec}\,\mathbf{I}_k)\Big]\mathbf{D}^{(n)} = \mathbf{\Delta}^{II}_{\boldsymbol{\vartheta};\phi} + o_{\mathrm{P}}(1) \tag{A.10}$$

and

$$\mathbf{D}^{III(n)} := \frac{1}{2}\Big[\mathbf{I}_m \otimes \mathbf{M}_k(\mathbf{V})(\mathbf{V}^{\otimes 2})^{-1/2}\Big]\mathbf{D}^{(n)} = \mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi} + o_{\mathrm{P}}(1), \tag{A.11}$$

as $n \to \infty$. Now,

$$\begin{aligned}
Q^{(n)}_{\mathcal{N}} &= \frac{1}{2n}\sum_{1\le i<i'\le m} n_i n_{i'}\ \mathrm{tr}\Big[(\mathbf{\Sigma}^{-1}(\mathbf{S}_i - \mathbf{S}_{i'}))^2\Big] + o_{\mathrm{P}}(1)\\
&= \frac{1}{2n}\sum_{1\le i<i'\le m} n_i n_{i'}\ \mathrm{tr}\Big[(n_i^{-1/2}\mathbf{Z}_i - n_{i'}^{-1/2}\mathbf{Z}_{i'})^2\Big] + o_{\mathrm{P}}(1)\\
&= \frac{1}{2}\sum_{1\le i<i'\le m} \mathrm{tr}\Big[((\lambda_{i'}^{(n)})^{1/2}\mathbf{Z}_i - (\lambda_i^{(n)})^{1/2}\mathbf{Z}_{i'})^2\Big] + o_{\mathrm{P}}(1)\\
&=: \frac{1}{2}\Big[(\mathbf{L}^{(n)}_m \otimes \mathbf{I}_{k^2})\mathbf{D}^{(n)}\Big]'\Big[(\mathbf{L}^{(n)}_m \otimes \mathbf{I}_{k^2})\mathbf{D}^{(n)}\Big] + o_{\mathrm{P}}(1).
\end{aligned} \tag{A.12}$$

Noting that $\mathbf{L}^{(n)\prime}_m\mathbf{L}^{(n)}_m = \mathbf{I}_m - \mathbf{C}^{(n)}$ and using the fact that $\mathbf{K}_k(\mathrm{vec}\,\mathbf{v}) = (\mathrm{vec}\,\mathbf{v})$ for any symmetric $k \times k$ matrix $\mathbf{v}$, we obtain

$$Q^{(n)}_{\mathcal{N}} = T^{(n)}_1 + T^{(n)}_2 + o_{\mathrm{P}}(1), \tag{A.13}$$

where

$$T^{(n)}_1 := \frac{1}{2}\Big(\mathbf{D}^{(n)}\Big)'\Big\{[\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes \Big(\frac{1}{k}\mathbf{J}_k\Big)\Big\}\,\mathbf{D}^{(n)}$$

and

$$T^{(n)}_2 := \frac{1}{4}\Big(\mathbf{D}^{(n)}\Big)'\Big\{[\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes \Big(\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\Big)\Big\}\,\mathbf{D}^{(n)}.$$

In view of (A.10), we straigthforwardly obtain

$$\begin{aligned}
T^{(n)}_1 &= \frac{2\sigma^4}{k}\Big(\mathbf{D}^{II(n)}\Big)'[\mathbf{I}_m - \mathbf{C}^{(n)}]\,\mathbf{D}^{II(n)} = \frac{C_k(g_1)}{2k}\Big(\mathbf{D}^{II(n)}\Big)'\mathbf{P}^{g,II}_{\boldsymbol{\vartheta};\phi*}\,\mathbf{D}^{II(n)}\\
&= \Big[1 + \kappa_k(g_1) + \frac{k\kappa_k(g_1)}{2}\Big]\Big(\mathbf{\Delta}^{II}_{\boldsymbol{\vartheta};\phi}\Big)'\mathbf{P}^{g,II}_{\boldsymbol{\vartheta};\phi*}\,\mathbf{\Delta}^{II}_{\boldsymbol{\vartheta};\phi} + o_{\mathrm{P}}(1).
\end{aligned} \tag{A.14}$$

Similarly, (4.6) and (A.11) entail

$$\begin{aligned}
T^{(n)}_2 &= \frac{1}{k(k+2)}\Big(\mathbf{D}^{III(n)}\Big)'\Big\{[\mathbf{I}_m - \mathbf{C}^{(n)}] \otimes (\mathbf{H}_k(\mathbf{V}))^{-1}\Big\}\Big(\mathbf{D}^{III(n)}\Big)'\\
&= \frac{E_k(g_1)}{k(k+2)}\Big(\mathbf{D}^{III(n)}\Big)'\mathbf{P}^{g,III}_{\boldsymbol{\vartheta};\phi*}\,\mathbf{D}^{III(n)} = (1 + \kappa_k(g_1))\Big(\mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi}\Big)'\mathbf{P}^{g,III}_{\boldsymbol{\vartheta};\phi*}\,\mathbf{\Delta}^{III}_{\boldsymbol{\vartheta};\phi} + o_{\mathrm{P}}(1),
\end{aligned}$$

which, together with (A.13) and (A.14) establishes the result. □

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

[5] 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.

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

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

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

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

[10] 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.

[11] 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.

[12] Hallin, M., H. Oja, and D. Paindaveine (2006). Semiparametrically efficient rank-based inference for shape. II. Optimal R-Estimation of Shape. *Annals of Statistics* **34**, 2757-2789.

[13] Hallin, M. and D. Paindaveine (2004). Rank-based optimal tests of the adequacy of an elliptic VARMA model. *Annals of Statistics* **32**, 2642-2678.

[14] Hallin, M. and D. Paindaveine (2006a). Semiparametrically efficient rank-based inference for shape I: Optimal rank-based tests for sphericity. *Annals of Statistics* **34**, 2707-2756.

[15] Hallin, M. and D. Paindaveine (2006b). Parametric and semiparametric inference for shape: the role of the scale functional. *Statistics and Decisions* **24**, 1001-1023.

[16] Hallin, M. and D. Paindaveine (2007). Optimal rank-based tests for homogeneity of scatter. *Annals of Statistics*, to appear.

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

[18] 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.

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

[20] Liebscher, E. (2005). A semiparametric density estimator based on elliptical distributions. *Journal of Multivariate Analysis* **92**, 205-225.

[21] 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.

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

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

[24] Paindaveine, D. (2007). A canonical definition of shape. Submitted.

[25] 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.

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

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

[28] Rao, C. R., and S. K. Mitra (1971). *Generalized Inverses of Matrices and its Applications.* J. Wiley, New York.

[29] Salibián-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.

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

[31] Shapiro, A., and M.W. Browne (1987). Analysis of covariance structures under elliptical distributions. *Journal of the American Statistical Association* **82**, 1092-1097.

[32] 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.

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

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

[35] Sugiura, N, and H. Nagao (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *The Annals of Mathematical Statistics* **39**, 1686-1692.

[36] 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.

[37] 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.

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

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

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

[41] Wilks, S.S. (1932). Certain generalizations in the analysis of variance. *Biometrika* **24**, 471-494.

[42] 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.

[43] 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.

[44] 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,
Institut de Recherche en Statistique,
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,
Institut de Recherche en Statistique,
and Département de Mathématique,
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
dpaindav@ulb.ac.be