ECARES

Optimal Tests for Elliptical Symmetry: Specified and Unspecified Location

Slađana Babić
Ghent University

Laetitia Gelbgras
Université libre de Bruxelles

Marc Hallin
SBS-EM, ECARES, Université libre de Bruxelles

Christophe Ley
Ghent University

November 2019

ECARES working paper 2019-26

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

arXiv: arXiv:0000.0000

Optimal tests for elliptical symmetry: specified and unspecified location

Slađana Babić[*], Laetitia Gelbgras, Marc Hallin, and Christophe Ley

Slađana Babić
Department of Applied Mathematics, Computer Science, and Statistics
Ghent University
Krijgslaan 281, S9
B-9000 Gent, Belgium
and
Vlerick Business School
e-mail: sladana.babic@ugent.be

Laetitia Gelbgras
Département de Mathématique
Université libre de Bruxelles
Campus de la Plaine CP210
B-1050 Bruxelles, Belgium
e-mail: laetitia.gelbgras@yahoo.fr

Marc Hallin
ECARES and
Département de Mathématique
Université libre de Bruxelles CP114/4
Ave. F.D. Roosevelt, 50
B-1050 Bruxelles, Belgium
e-mail: mhallin@ulb.ac.be

Christophe Ley
Department of Applied Mathematics, Computer Science, and Statistics
Ghent University
Krijgslaan 281, S9
B-9000 Gent, Belgium
e-mail: christophe.ley@ugent.be

Abstract: Although the assumption of elliptical symmetry is quite common in multivariate analysis and widespread in a number of applications, the problem of testing the null hypothesis of ellipticity so far has not been addressed in a fully satisfactory way. Most of the literature in the area indeed addresses the null hypothesis of elliptical symmetry with specified location and actually addresses location rather than non-elliptical alternatives. In this paper, we are proposing new classes of testing procedures, both for specified and unspecified location. The backbone of our construction is Le Cam's asymptotic theory of statistical experiments, and optimality is to be understood locally and asymptotically within the family of generalized skew-elliptical distributions. The tests we are proposing are meeting all the desired properties of a"good" test of elliptical symmetry: they have a

[*]Slađana Babić was supported by a grant (165880) as a PhD Fellow of the Research Foundation-Flanders (FWO).

simple asymptotic distribution under the entire null hypothesis of elliptical symmetry with unspecified radial density and shape parameter; they are affine-invariant, computationally fast, intuitively understandable, and not too demanding in terms of moments. While achieving optimality against generalized skew-elliptical alternatives, they remain quite powerful under a much broader class of non-elliptical distributions and significantly outperform the available competitors.

MSC 2010 subject classifications: Primary 62H15, 62H10.
Keywords and phrases: Elliptical symmetry, Local asymptotic normality, Maximin tests, Multivariate skewness, Semiparametric inference, Skew-elliptical densities..

1. Introduction

1.1. The ubiquitous assumption of elliptical symmetry

Elliptical symmetry is a fundamental structural assumption in multivariate analysis and econometrics. It has been popularized in the 1970's as a natural extension of the (overly restrictive) multinormal assumption. Since then, most multivariate analysis procedures have been extended under elliptical symmetry with unspecified and sometimes possibly heavy-tailed *radial density* (see below for a definition): one- and K-sample location and shape problems ([50, 19, 24, 18, 25, 26]), serial dependence and time series ([20, 21, 22]), linear models with VARMA errors ([23]), one- and K-sample principal component problems ([27, 28, 29, 30]), to cite but a few. Most tests proposed in those references are either pseudo-Gaussian or based on variations of *Mahalanobis* ranks and signs, *interdirections*, etc. Elliptical densities also are considered in capital asset pricing models [31], semiparametric density estimation [38], graphical models [51], multivariate tail estimation [11], and many other areas.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ denote a sample of n i.i.d. d-dimensional observations. A d-dimensional random vector $\mathbf{X}$ is said to be elliptically symmetric about some location parameter $\boldsymbol{\theta} \in \mathbb{R}^d$ if its density $\underline{f}$ is of the form

$$\mathbf{x} \mapsto \underline{f}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f) = c_{d,f} |\boldsymbol{\Sigma}|^{-1/2} f\left(\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})\|\right), \qquad \mathbf{x} \in \mathbb{R}^d, \tag{1.1}$$

where $\boldsymbol{\Sigma} \in \mathcal{S}_d$ (the class of symmetric positive definite real $d \times d$ matrices) is a *scatter* parameter, $f : \mathbb{R}_0^+ \to \mathbb{R}^+$ is an a.e. strictly positive function called *radial density*, and $c_{d,f}$ is a normalizing constant depending on f and the dimension d. Well-known instances are the multivariate normal, Student t and power-exponential distributions. The family of elliptical distributions has several appealing properties. For instance, it is closed under affine transformations, and its marginal and conditional distributions are also elliptically symmetric: see [42] for details. A salient feature is the stochastic representation of elliptical variables: an elliptically symmetric random vector $\mathbf{X}$ is conveniently represented as

$$\mathbf{X} =_d \boldsymbol{\theta} + \rho \boldsymbol{\Lambda} \mathbf{U}^{(r)}, \tag{1.2}$$

where $=_d$ stands for equality in distribution, $\mathbf{\Lambda} \in \mathbb{R}^{d\times r}$ has rank $r \leq d$ and is such that $\mathbf{\Lambda\Lambda}' = \mathbf{\Sigma}$, $\mathbf{U}^{(r)}$ is an r-dimensional random vector uniformly distributed over the unit hypersphere, and ρ is a nonnegative random variable independent of $\mathbf{U}^{(r)}$. Letting $\mu_{\ell,f} := \int_0^\infty r^\ell f(r)dr$, the density of ρ is

$$r \mapsto \tilde{f}_d(r) := \mu_{d-1,f}^{-1} r^{d-1} f(r), \qquad r > 0. \tag{1.3}$$

The existence of this density thus requires $\mu_{d-1,f}$ to be finite, and $\mathbf{X}$ admits finite moments of order $\alpha > 0$ if and only if $\mu_{d+\alpha-1,f} < \infty$. Inference in elliptically symmetric distributions has been abundantly studied: see [42] for a survey.

1.2. *Testing for elliptical symmetry*

Considering the omnipresence of the assumption of elliptical symmetry, it is of primary importance to be able to test whether that assumption actually holds true, and various tests have been proposed in the literature for this problem. We briefly mention the most popular of them, along with their respective pitfalls; later on, we will focus more closely on those used in our comparative Monte Carlo study (Section 5). We also mention tests for spherical symmetry, a special case of elliptical symmetry corresponding to $\boldsymbol{\mu} = \mathbf{0}$ and $\mathbf{\Sigma} = \mathbf{I}_d$, the $d\times d$ identity matrix. These tests in principle can be turned into elliptical symmetry tests by standardizing the data via $\hat{\mathbf{\Sigma}}^{-1/2}(\mathbf{X}_i - \hat{\boldsymbol{\theta}})$ where $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{\Sigma}}$ are location and scatter estimators.

(i) Beran [6] introduces a test based on marginal signs and ranks. That test is neither distribution-free nor affine-invariant; moreover, there are no practical guidelines to the choice of the basis functions involved in the test statistic.
(ii) Baringhaus [5] proposes a Cramér-von Mises type test for spherical symmetry based on the independence between norm and direction. It assumes the location parameter to be known and its asymptotic distribution is not simple to use. Dyckerhoff et al. [12] have shown by simulations that this test can be used as a test for elliptical symmetry in dimension 2.
(iii) Koltchinskii and Sakhanenko [34] consider bootstrap-type tests based on a class of functions closed under orthogonal transformations. Their tests have no known asymptotic distribution, which is why a bootstrap procedure is required to get the critical values.
(iv) Manzotti et al. [40] develop a test based on spherical harmonics to test whether the standardized vectors $\hat{\mathbf{\Sigma}}^{-1/2}(\mathbf{X}_i - \hat{\boldsymbol{\theta}})/||\hat{\mathbf{\Sigma}}^{-1/2}(\mathbf{X}_i - \hat{\boldsymbol{\theta}})||$ are uniformly distributed on the unit sphere. The test is computationally demanding and requires moments of order 4.
(v) Schott [46] builds a Wald-type test to compare the sample fourth-order moments with the expected theoretical ones under elliptical symmetry. Being based on fourth-order moments, the test is very simple to use but requires moments of order 8. Moreover, it has very low power against several alternatives.

(vi) Huffer and Park [32] propose a Pearson chi-square type test with multi-dimensional cells. Its asymptotic distribution exists only in case of normality, otherwise bootstrap techniques are required.
(vii) Cassart [8] and Cassart et al. [9] construct a pseudo-Gaussian test that is most efficient against a multivariate form of Fechner-type asymmetry. The test requires finite moments of order 4.

Tests based on Monte Carlo simulations can be found in Diks and Tong [10] and Zhu and Neuhaus [54]; Li, Fang and Zhu [37] recur to graphical methods and Zhu and Neuhaus [55] build conditional tests. We refer the reader to Serfling [47] and Sakhanenko [45] for extensive reviews and performance comparisons.

1.3. Goal and organization of the paper

Despite the practical importance of the problem and the many proposals made in the literature, all tests for elliptical symmetry are suffering from some serious drawbacks. None of them, except for Cassart [8], is based on efficiency arguments; and, to the best of our knowledge, none of them has been implemented in `R`.

This paper is filling this gap by building tests for elliptical symmetry that are optimal against the very popular class of *generalized skew-elliptical distributions* which we define more precisely in Section 2.1. It should be clear, however, that we never require the actual density of the observations to belong to that class, the choice of which is made because it encompasses many proposed skew distributions from the literature (see, e.g., Genton [14]). The `R`-code is available on request and an `R`-package under preparation.

The tests we are proposing are meeting all the desired properties of a "good" test of elliptical symmetry: they have simple asymptotic distributions under the entire null hypothesis of elliptical symmetry with unspecified radial density and shape parameter; they are affine-invariant, computationally fast, intuitively understandable, and not too demanding in terms of moments. The latter property is particularly important when dealing with possibly heavy-tailed data as is often the case in a financial context. All our tests are devised for specified and, most importantly, unspecified location parameter. The latter indeed is the "genuine" problem here, as specified-location tests for ellipticity typically run into major problems—see Section 5.5 and the empirical illustration in Section 6.

The approach we are adopting thus combines optimality and robustness concerns (distribution-freeness with respect to radial densities and minimal moment assumptions). The backbone of our construction is Le Cam's asymptotic theory of statistical experiments, and optimality is to be understood in the local asymptotic sense (against local generalized skew-elliptical deviations from ellipticity). Under each scenario (specified and unspecified location), we first build optimal parametric tests by assuming a given elliptical distribution. Then we make these tests valid under the entire semiparametric family of elliptically symmetric distributions, while preserving their (parametric) optimality. As we shall see, under specified location, the optimal parametric test statistics do not involve the radial density, hence have all the same expression which consequently is *uniformly* optimal across radial densities—a rather rare phenomenon, which does not hold in other problems involving elliptical densities. When the location is unspecified, this uniform optimality property gets lost, but we still obtain very simple and fast-to-compute test statistics that significantly outperform their competitors and do not require estimating the actual density, as is often the case. A detailed comparative study of the finite-sample performances of our tests is conducted in Section 5 and demonstrates the power of our procedures.

The rest of the paper is organized as follows. In Section 2, we describe the family of generalized skew-elliptical distributions and state some mild conditions on the radial density f which are required in order to establish uniform local asymptotic normality (ULAN) under given f. In Section 3, we derive, for given f, the locally and asymptotically optimal tests for symmetry about a specified location $\boldsymbol{\theta}$. These tests are parametric, and valid under the known radial density f only. We turn them into semiparametric tests that remain valid under a broad class of radial densities and, as already mentioned, also are uniformly optimal against alternatives involving the same class of densities. Section 4 deals with the unspecified location case, for which again we derive parametrically locally and asymptotically optimal tests, which we turn into semiparametric ones, the properties of which we provide under the null and contiguous alternatives. Asymptotic relative efficiencies with respect to the aforementioned pseudo-Gaussian test of Cassart [8] are calculated in Section 4.3. In Section 5, we conduct a Monte Carlo simulation study of the finite-sample performances of the proposed tests and their main competitors. Section 5.5 stresses the all too often overlooked pitfalls of specified-location methods. A real-data analysis is carried out in Section 6 and conclusions are provided in Section 7. Technical proofs are concentrated in the Appendix.

2. Generalized skew-elliptical families and Uniform Local Asymptotic Normality (ULAN)

2.1. Generalized skew-elliptical distributions

As mentioned in the Introduction, our goal is to propose efficient tests against a family of densities representative of a broad class of skewed densities. The family of *generalized skew-elliptical distributions* ([15]) is an ideal candidate for this role.

Let us assume that the radial density f in (1.1) belongs to

$$\mathcal{F} := \left\{ f : \mathbb{R}_0^+ \to \mathbb{R}^+ \,:\, f(r) > 0 \text{ a.e. and } \mu_{d-1;f} := \int_0^\infty r^{d-1} f(r)dr < \infty \right\}.$$

It is clear from (1.2) that ρ and $\boldsymbol{\Sigma}$ are not separately identifiable, and we therefore impose a further identification constraint:

$$f \in \mathcal{F}_1 := \left\{ f \in \mathcal{F} : \mu_{d-1;f}^{-1} \int_{\mathbb{R}^+} r^{d+1} f(r)dr = d \right\}; \tag{2.1}$$

Under this constraint, ρ has finite variance and $\text{Cov}[\mathbf{X}] = \mathbf{\Sigma}$, which fully identifies the scatter matrix $\mathbf{\Sigma}$. While imposing the existence of finite second-order moments, (2.1) does not imply any loss of generality, as second-order moments are needed anyway (see Section 2.3) to have finite Fisher information for skewness. It will be required in all statements involving ULAN (optimality, local powers, etc.), but is not necessary for statements made under the null hypothesis of ellipticity (mainly, the asymptotic size of a test and its validity). Gaussian densities clearly satisfy (2.1), but the Student ones do not, and need to be rescaled.

The *generalized skew-elliptical* alternatives we are interested in belong to the class of *Azzalini-type distributions*. That class contains all generalizations of the famous scalar skew-normal distribution introduced by Azzalini [1] with density function $x \mapsto 2\phi(x)\Phi(\lambda x), x \in \mathbb{R}$, where ϕ and Φ stand for the standard normal density and distribution functions, respectively, and $\lambda \in \mathbb{R}$ is a skewness parameter. The idea underpinning the definition of the skew-normal consists in perturbating or modulating a *symmetric kernel*, here the normal, by multiplying it with a *skewing function*, here $\Phi(\lambda x)$. Its multivariate generalization was introduced in Azzalini and Dalla Valle [4] by replacing the scalar normal density with the d-variate normal. Azzalini and Capitanio [2] and Branco and Dey [7] in turn extended the multivariate skew-normal into skew-elliptical distributions based on elliptically symmetric kernels. Azzalini and Capitanio [3] established a link between the distinct constructions of skew-elliptical distributions, extending them into a broader class of skewed distributions very similar to the generalized skew-elliptical distributions defined by Genton and Loperfido [15], with pdfs of the form

$$\begin{aligned} \mathbf{x} \mapsto & \, \underline{f}(\mathbf{x}; \boldsymbol{\theta}, \mathbf{\Sigma}, \boldsymbol{\lambda}, f) \\ & := 2\, c_{d,f} |\mathbf{\Sigma}|^{-1/2} f(\|\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})\|)\Pi(\boldsymbol{\lambda}'\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})), \ \mathbf{x} \in \mathbb{R}^d, \end{aligned} \tag{2.2}$$

where $\boldsymbol{\theta}$, $\mathbf{\Sigma}$, $c_{d,f}$, and f are defined as in (1.1); the *skewing function* Π has values in $[0, 1]$ and satisfies $\Pi(-r) = 1 - \Pi(r)$ for $r \in \mathbb{R}$; $\boldsymbol{\lambda} \in \mathbb{R}^d$ plays the role of a *skewness* parameter. The density (2.2) thus results from perturbing the elliptically symmetric *kernel* $\underline{f}(\mathbf{x}; \boldsymbol{\theta}, \mathbf{\Sigma}, f)$ into $2\underline{f}(\mathbf{x}; \boldsymbol{\theta}, \mathbf{\Sigma}, f)\Pi(\boldsymbol{\lambda}'\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta}))$ by multiplying it with a general skewing function $\Pi(\cdot)$; clearly, the original symmetric version is retrieved for $\boldsymbol{\lambda} = \mathbf{0}$. Typical choices for Π are univariate distribution functions with symmetric densities, such as the normal or Student ones; see the monograph by Genton [14]. We opted for this class of skew alternatives because of its popularity and its ability to closely approximate a large variety of skewed distributions.

2.2. Notation and some definitions

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. with density (2.2). Denote by $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \mathbf{\Sigma}, \boldsymbol{\lambda}; f, \Pi}$ the joint distribution of $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ which, in case $\boldsymbol{\lambda} = \mathbf{0}$, we simply write as $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \mathbf{\Sigma}, \mathbf{0}; f}$.

Any couple (f, Π) then induces a parametric location-scatter-skewness model

$$\mathrm{P}^{(n)}_{f,\Pi} := \left\{ \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi} : \boldsymbol{\theta} \in \mathbb{R}^d, \boldsymbol{\Sigma} \in \mathcal{S}_d, \boldsymbol{\lambda} \in \mathbb{R}^d \right\}.$$

We are interested in testing $\mathcal{H}_0 : \boldsymbol{\lambda} = \mathbf{0}$ against $\mathcal{H}_1 : \boldsymbol{\lambda} \neq \mathbf{0}$ in (2.2), in the presence of a variety of unspecified nuisances: Π and/or $\boldsymbol{\theta}$ and/or $\boldsymbol{\Sigma}$ and/or f ... Depending on the case, the problem is either parametric or semiparametric. The four types of testing problems we are considering are

(a) (specified f and specified $\boldsymbol{\theta}$)
$\mathcal{H}^{(n)}_{0;f,\boldsymbol{\theta}} := \bigcup_{\boldsymbol{\Sigma}\in\mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$ versus $\mathcal{H}^{(n)}_{1;f,\Pi,\boldsymbol{\theta}} := \bigcup_{\boldsymbol{\Sigma}\in\mathcal{S}_d,\boldsymbol{\lambda}\neq\mathbf{0}} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi}$,

(b) (specified f and unspecified $\boldsymbol{\theta}$)
$\mathcal{H}^{(n)}_{0;f} := \bigcup_{\boldsymbol{\theta}\in\mathbb{R}^d,\boldsymbol{\Sigma}\in\mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$ versus $\mathcal{H}^{(n)}_{1;f,\Pi} = \bigcup_{\boldsymbol{\theta}\in\mathbb{R}^d,\boldsymbol{\Sigma}\in\mathcal{S}_d,\boldsymbol{\lambda}\neq\mathbf{0}} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi}$,

(c) (unspecified f and specified $\boldsymbol{\theta}$)
$\mathcal{H}^{(n)}_{0;\boldsymbol{\theta}} := \bigcup_{f\in\mathcal{F}_1,\boldsymbol{\Sigma}\in\mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$ versus $\mathcal{H}^{(n)}_{1;\Pi,\boldsymbol{\theta}} := \bigcup_{f\in\mathcal{F}_1,\boldsymbol{\Sigma}\in\mathcal{S}_d,\boldsymbol{\lambda}\neq\mathbf{0}} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi}$, and

(d) (unspecified f and unspecified $\boldsymbol{\theta}$)
$\mathcal{H}^{(n)}_{0} := \bigcup_{f\in\mathcal{F}_1,\boldsymbol{\theta}\in\mathbb{R}^d,\boldsymbol{\Sigma}\in\mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$ versus $\mathcal{H}^{(n)}_{1;\Pi} := \bigcup_{f\in\mathcal{F}_1,\boldsymbol{\theta}\in\mathbb{R}^d,\boldsymbol{\Sigma}\in\mathcal{S}_d,\boldsymbol{\lambda}\neq\mathbf{0}} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi}$;

the skewing function Π and the scatter $\boldsymbol{\Sigma}$ throughout remain unspecified.

For all $i = 1, \ldots, n$, denote by $d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \|\boldsymbol{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})\|$ the Mahalanobis distance of $\mathbf{X}_i$ to $\boldsymbol{\theta}$ and by $\mathbf{U}_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \boldsymbol{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})/d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ its *multivariate sign* in the metric $\boldsymbol{\Sigma}$. Under elliptical symmetry, those signs are uniformly distributed on the unit hypersphere of $\mathbb{R}^d$ whereas the radial quantities $d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ have common density $\tilde{f}_d$, see (1.3). Any square root of $\boldsymbol{\Sigma}$ can be used in the previous definitions, but we throughout denote by $\boldsymbol{\Sigma}^{1/2}$ the unique symmetric positive definite one.

Let $\mathbf{S}$ be a $d \times d$ symmetric matrix. We throughout use the classical vec$\mathbf{S}$ notation for the d^2-vector obtained by stacking the columns of $\mathbf{S}$ on top of each other and write vech$\mathbf{S}$ for the $d(d+1)/2$-dimensional vector stacking its upper-triangular elements. We then denote by $\mathbf{P}_d$ the $(d(d+1)/2) \times d^2$ matrix such that $\mathbf{P}'_d(\text{vech}\mathbf{S}) = \text{vec}\mathbf{S}$. Write $\mathbf{S}^{\otimes 2}$ for the Kronecker product $\mathbf{S} \otimes \mathbf{S}$. Finally, denoting by $\mathbf{e}_i$ the i^{th} vector of the canonical basis of $\mathbb{R}^d$, define the $d^2 \times d^2$ *commutation matrix* $\mathbf{K}_d := \sum_{i,j=1}^d (\mathbf{e}_i\mathbf{e}'_j) \otimes (\mathbf{e}_j\mathbf{e}'_i)$ and the $d^2 \times d^2$ *projection matrix* $\mathbf{J}_d := \sum_{i,j=1}^d (\mathbf{e}_i\mathbf{e}'_j) \otimes (\mathbf{e}_i\mathbf{e}'_j) = (\text{vec}\mathbf{I}_d)(\text{vec}\mathbf{I}_d)'$.

2.3. Uniform Local Asymptotic Normality (ULAN)

The backbone of our construction of efficient tests in the subsequent sections is the ULAN property, at $\boldsymbol{\lambda} = \mathbf{0}$, of the parametric model $\mathrm{P}^{(n)}_{f,\Pi}$. This ULAN property requires some further regularity conditions on f. Let $(\Omega, \mathcal{B}^d_\Omega, \lambda)$ be a measure space, where λ is a measure on the open subset $\Omega \subseteq \mathbb{R}^d$ equipped with its Borel σ-field $\mathcal{B}^d_\Omega$. Denote by $L^2(\Omega, \lambda)$ the space of measurable functions $h : \Omega \to \mathbb{R}$ such that $\int_\Omega [h(\mathbf{x})]^2 d\lambda(\mathbf{x}) < \infty$, by $L^2(\mathbb{R}^+_0, \mu_j)$ the space of square-integrable functions with respect to the Lebesgue measure with weight r^j over $\mathbb{R}_0^+$, and by $L^2(\mathbb{R}, \nu_j)$ the space of square-integrable functions with respect to the Lebesgue measure with weight e^{rj} over $\mathbb{R}$. We say that $g \in L^2(\Omega, \lambda)$ admits a *weak partial derivative* T_i with respect to the i^{th} variable iff

$$\int_\Omega g(\mathbf{x}) \partial_i \varphi(\mathbf{x}) d\mathbf{x} = -\int_\Omega T_i(\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x}$$

for any function $\varphi \in C_0^\infty(\Omega)$, i.e. for any infinitely differentiable (in the classical sense) compactly supported function φ on Ω. If T_i exists for all i, the gradient $\mathbf{T} := (T_1, \ldots, T_d)$ is also called the *derivative of g in the sense of distributions* in $L^2(\Omega, \lambda)$. If, in addition, $\mathbf{T} \in L^2(\Omega, \lambda)$, then g belongs to $W^{1,2}(\Omega, \lambda)$, the *Sobolev space of order 1* on $L^2(\Omega, \lambda)$. This space is a Banach space when equipped with the norm

$$\|g\|_{W^{1,2}(\Omega,\lambda)} := (\|g\|^2_{L^2(\Omega,\lambda)} + \sum_{i=1}^d \|T_i\|^2_{L^2(\Omega,\lambda)})^{1/2}.$$

In particular, we will denote by $L^2(\Omega)$ and $W^{1,2}(\Omega)$ the case where λ is the Lebesgue measure on Ω.

With this in hand, let us state the regularity assumptions we need for ULAN.

Assumption (A1) The mapping $r \mapsto f^{1/2}(r)$ belongs to $W^{1,2}(\mathbb{R}_0, \mu_{d-1})$. Define $\varphi_f(r) := -2(f^{1/2})'(r)/f^{1/2}(r)$, where $(f^{1/2})'$ stands for the weak derivative of $f^{1/2}$ in $L^2(\mathbb{R}_0, \mu_{d-1})$. Assumption (A1) ensures finiteness of the *Fisher information for location*

$$\mathcal{I}_{d,f} := c_{d,f} \int_{\mathbb{R}^d} \varphi_f^2(\|\mathbf{x}\|) f(\|\mathbf{x}\|) d\mathbf{x}.$$

Assumption (A2) The mapping $r \mapsto f_{\exp}^{1/2}(r) := f^{1/2}(e^r)$ belongs to $W^{1,2}(\mathbb{R}_0, \nu_d)$.
Letting $\psi_f(r) := -2r^{-1}\left(f_{\exp}^{1/2}\right)'(\log r)/f^{1/2}(r)$, where $(f_{\exp}^{1/2})'$ stands for the weak derivative of $f_{\exp}^{1/2}$ in $L^2(\mathbb{R}_0, \nu_d)$, Assumption (A2) ensures finiteness of the *Fisher information for scatter*

$$\mathcal{J}_{d,f} := c_{d,f} \int_{\mathbb{R}^d} \|\mathbf{x}\|^2 \psi_f^2(\|\mathbf{x}\|) f(\|\mathbf{x}\|) d\mathbf{x}.$$

Now, if we assume the radial density f to be continuously differentiable, then φ_f and ψ_f both coincide with $-\dot{f}/f$ where $\dot{f}$ is the classical (strong) derivative of f.

Note that (2.1) is sufficient for the finiteness of the *Fisher information for skewness* (see Theorem 2.1 below), which only requires finite moments of order 2.

Finally, let $\boldsymbol{\vartheta} := (\boldsymbol{\theta}', (\mathrm{vech}\boldsymbol{\Sigma})', \boldsymbol{\lambda}')'$ and $\boldsymbol{\vartheta_0} := (\boldsymbol{\theta}', (\mathrm{vech}\boldsymbol{\Sigma})', \mathbf{0}')'$. We are now ready to state the ULAN property of the family $\mathrm{P}^{(n)}_{f,\Pi}$ *in the vicinity of symmetry.*

Theorem 2.1. *Let $f \in \mathcal{F}_1$. Suppose that Assumptions (A1) and (A2) hold, and that the skewing function Π is continuously differentiable at 0, with derivative $\dot{\Pi}(0) \neq 0$. Then, the family $\mathrm{P}_{f,\Pi}^{(n)}$ is ULAN at $\boldsymbol{\vartheta}_0$ with respect to $\boldsymbol{\theta}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\lambda}$, with central sequence*

$$\boldsymbol{\Delta}_f(\boldsymbol{\vartheta}_{\mathbf{0}}) = \begin{pmatrix} \boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_{\mathbf{0}}) \\ \boldsymbol{\Delta}_{f;2}(\boldsymbol{\vartheta}_{\mathbf{0}}) \\ \boldsymbol{\Delta}_{3}(\boldsymbol{\vartheta}_{\mathbf{0}}) \end{pmatrix}$$

$$:= \begin{pmatrix} n^{-1/2}\sum_{i=1}^{n}\varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\boldsymbol{\Sigma}^{-1/2}\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma}) \\ \frac{1}{2}n^{-1/2}\mathbf{P}_d(\boldsymbol{\Sigma}^{\otimes 2})^{-1/2}\sum_{i=1}^{n}\mathrm{vec}\left(\psi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))d_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\mathbf{U}_i^{'}(\boldsymbol{\theta},\boldsymbol{\Sigma}) - \mathbf{I}_d\right) \\ 2n^{-1/2}\dot{\Pi}(0)\sum_{i=1}^{n}d_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma}) \end{pmatrix}$$

and Fisher information matrix

$$\boldsymbol{\Gamma}_f(\boldsymbol{\vartheta}_{\mathbf{0}}) := \begin{pmatrix} \boldsymbol{\Gamma}_{f;11}(\boldsymbol{\vartheta}_{\mathbf{0}}) & \mathbf{0} & \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_{\mathbf{0}}) \\ \mathbf{0} & \boldsymbol{\Gamma}_{f;22}(\boldsymbol{\vartheta}_{\mathbf{0}}) & \mathbf{0} \\ \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_{\mathbf{0}}) & \mathbf{0} & \boldsymbol{\Gamma}_{f;33}(\boldsymbol{\vartheta}_{\mathbf{0}}) \end{pmatrix}, \tag{2.3}$$

where $\boldsymbol{\Gamma}_{f;11}(\boldsymbol{\vartheta}_{\mathbf{0}}) := \frac{1}{d}\mathcal{I}_{d,f}\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_{\mathbf{0}}) := 2\dot{\Pi}(0)\boldsymbol{\Sigma}^{-1/2}$, $\boldsymbol{\Gamma}_{f;33}(\boldsymbol{\vartheta}_{\mathbf{0}}) := 4(\dot{\Pi}(0))^2\mathbf{I}_d$, *and* $\boldsymbol{\Gamma}_{f;22}(\boldsymbol{\vartheta}_{\mathbf{0}}) := \frac{1}{4}\mathbf{P}_d(\boldsymbol{\Sigma}^{\otimes 2})^{-1/2}\left[\frac{\mathcal{J}_{d,f}}{d(d+2)}(\mathbf{I}_{d^2} + \mathbf{K}_d + \mathbf{J}_d) - \mathbf{J}_d\right](\boldsymbol{\Sigma}^{\otimes 2})^{-1/2}\mathbf{P}_d'$.

More precisely, for any sequence $\boldsymbol{\vartheta}_{\mathbf{0},n} = (\boldsymbol{\theta}_n^{'}, (\mathrm{vech}\boldsymbol{\Sigma}_n)', \mathbf{0}')'$, *where* $\boldsymbol{\theta}_n - \boldsymbol{\theta}$ *and* $\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}$ *are* $O(n^{-1/2})$, *and for any bounded sequence* $\boldsymbol{\tau}^{(n)}$ *of the form* $((\mathbf{t}^{(n)})', (\mathrm{vech}\mathbf{H}^{(n)})', (\boldsymbol{\ell}^{(n)})')' = ((\boldsymbol{\tau}_1^{(n)})', (\boldsymbol{\tau}_2^{(n)})', (\boldsymbol{\tau}_3^{(n)})')' \in \mathbb{R}^{2d+d(d+1)/2}$,

$$\begin{aligned} L^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0},n}+n^{-1/2}\boldsymbol{\tau}^{(n)}/\boldsymbol{\vartheta}_{\mathbf{0},n};f} &:= \log\left(\frac{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0},n}+n^{-1/2}\boldsymbol{\tau}^{(n)};f,\Pi}}{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0},n};f}}\right) \\ &= (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}_f(\boldsymbol{\vartheta}_{\mathbf{0},n}) - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_f(\boldsymbol{\vartheta}_{\mathbf{0}})\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1) \end{aligned}$$

and

$$\boldsymbol{\Delta}_f(\boldsymbol{\vartheta}_{\mathbf{0},n}) \xrightarrow{D} \mathcal{N}_{2d+d(d+1)/2}(\mathbf{0}, \boldsymbol{\Gamma}_f(\boldsymbol{\vartheta}_{\mathbf{0}}))$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0},n};f}$ *as* $n \to \infty$.

See Appendix A for the proof.

Note that the central sequence for skewness $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_{\mathbf{0}})$ does not depend on f; this, as we shall see, has strong implications on optimality properties.

An immediate consequence of the ULAN property is the *asymptotic linearity*, as $n \to \infty$, of the central sequence $\mathbf{\Delta}_f$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0}};f}$:

$$\mathbf{\Delta}_f(\boldsymbol{\vartheta}_{\mathbf{0}} + n^{-1/2}\boldsymbol{\tau}^{(n)}) - \mathbf{\Delta}_f(\boldsymbol{\vartheta}_{\mathbf{0}}) = -\mathbf{\Gamma}_f(\boldsymbol{\vartheta}_{\mathbf{0}})\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1). \tag{2.4}$$

This property classically plays a key role in the handling of nuisance parameters. Denote by $\hat{\boldsymbol{\theta}}^{(n)}$ and $\hat{\mathbf{\Sigma}}^{(n)}$ sequences of estimators of $\boldsymbol{\theta}$ and $\mathbf{\Sigma}$, respectively, satisfying the following conditions.

Assumption (B) For any $f \in \mathcal{F}_1$ and $\boldsymbol{\vartheta}_{\mathbf{0}}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0}};f}$, as $n \to \infty$, $\hat{\boldsymbol{\theta}}^{(n)}$ and $\hat{\mathbf{\Sigma}}^{(n)}$ *(i)* are root-n consistent: $n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta})$ and $n^{1/2}(\hat{\mathbf{\Sigma}}^{(n)} - \mathbf{\Sigma})$ are $O_{\mathrm{P}}(1)$, and *(ii)* are locally asymptotically discrete: the number of possible values of $\hat{\boldsymbol{\theta}}^{(n)}$ and $\mathrm{vech}\hat{\mathbf{\Sigma}}^{(n)}$ in any sequence of $O(n^{-1/2})$ balls centered around $\boldsymbol{\theta}$ and $\mathrm{vech}\mathbf{\Sigma}$, respectively, is uniformly bounded as $n \to \infty$.

This assumption, in combination with Lemma 4.4 of Kreiss [35], entails

$$\begin{aligned}\mathbf{\Delta}_f(\hat{\boldsymbol{\theta}}^{(n)}, \hat{\mathbf{\Sigma}}^{(n)}, \mathbf{0}) \quad &- \quad \mathbf{\Delta}_f(\boldsymbol{\theta}, \mathbf{\Sigma}, \mathbf{0}) \\ &= \quad -\mathbf{\Gamma}_f(\boldsymbol{\vartheta}_{\mathbf{0}}) n^{1/2}\big((\hat{\boldsymbol{\theta}}^{(n)\prime}, (\mathrm{vech}\hat{\mathbf{\Sigma}}^{(n)})', \mathbf{0}')' - \boldsymbol{\vartheta}_{\mathbf{0}}\big) + o_{\mathrm{P}}(1)\end{aligned} \tag{2.5}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_{\mathbf{0}};f}$ as $n \to \infty$. It should be noted that Assumption B*(ii)* 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: see Yang and Le Cam [53]).

In practice, it is desirable to restrict to affine-equivariant estimators: we will assume that $\hat{\boldsymbol{\theta}}^{(n)}$ and $\hat{\mathbf{\Sigma}}^{(n)}$ also satisfy

$$\hat{\boldsymbol{\theta}}^{(n)}(\mathbf{A}\mathbf{X}_1 + \mathbf{b}, \ldots, \mathbf{A}\mathbf{X}_n + \mathbf{b}) = \mathbf{A}\hat{\boldsymbol{\theta}}^{(n)}(\mathbf{X}_1, \ldots, \mathbf{X}_n) + \mathbf{b}$$

and

$$\hat{\mathbf{\Sigma}}^{(n)}(\mathbf{A}\mathbf{X}_1 + \mathbf{b}, \ldots, \mathbf{A}\mathbf{X}_n + \mathbf{b}) = \mathbf{A}\hat{\mathbf{\Sigma}}^{(n)}(\mathbf{X}_1, \ldots, \mathbf{X}_n)\mathbf{A}'$$

for any $d \times d$ matrix $\mathbf{A}$ and any d-vector $\mathbf{b}$. Under this natural requirement, our test statistics will enjoy affine-invariance. In the sequel, the lighter notation $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{\Sigma}}$ will be adopted.

We conclude this section on ULAN by noting the block-diagonal structure of the Fisher information matrix, implying that the $\mathbf{\Sigma}$- and $(\boldsymbol{\theta}, \boldsymbol{\lambda})$-parts of the central sequence are asymptotically independent.

3. Optimal parametric and semiparametric tests: specified $\boldsymbol{\theta}$

Fix $\boldsymbol{\theta} \in \mathbb{R}^d$. ULAN and the convergence of local sequences of experiments to a Gaussian shift experiment imply that a locally asymptotically optimal parametric test for $\mathcal{H}_{0;f,\boldsymbol{\theta}}$ against $\mathcal{H}_{1;f,\Pi,\boldsymbol{\theta}}$ can be based on a quadratic form involving the $\boldsymbol{\lambda}$-part $\mathbf{\Delta}_3(\boldsymbol{\vartheta}_{\mathbf{0}})$ of the central sequence. Of course, the nuisance scatter parameter $\mathbf{\Sigma}$ needs to be estimated. The block-diagonal structure of the Fisher information matrix, combined with (2.5) allows for substituting, without any loss of power, any $\hat{\mathbf{\Sigma}}$ satisfying Assumption (B) for the unknown $\mathbf{\Sigma}$. Thus, unlike Rao score/Lagrange multiplier tests or likelihood ratio tests, where $\hat{\mathbf{\Sigma}}$ has to be the MLE, we can accommodate various estimators and privilege computational convenience or robustness, or avoid higher-order moment assumptions. In the sequel, we are opting for Tyler [49]'s estimator of scatter (shape). Denote by $\mathbf{T}$ the unique (for $n > d(d-1)$) $d \times d$ upper-triangular matrix with positive diagonal elements and determinant equal to one satisfying

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mathbf{T}(\mathbf{X}_i-\boldsymbol{\theta})}{\|\mathbf{T}(\mathbf{X}_i-\boldsymbol{\theta})\|}\right)\left(\frac{\mathbf{T}(\mathbf{X}_i-\boldsymbol{\theta})}{\|\mathbf{T}(\mathbf{X}_i-\boldsymbol{\theta})\|}\right)' = \frac{1}{d}\mathbf{I}_d.$$

This matrix $\mathbf{T}$ is such that the covariance structure of

$$\left(\frac{\mathbf{T}(\mathbf{X}_1-\boldsymbol{\theta})}{\|\mathbf{T}(\mathbf{X}_1-\boldsymbol{\theta})\|},\ldots,\frac{\mathbf{T}(\mathbf{X}_n-\boldsymbol{\theta})}{\|\mathbf{T}(\mathbf{X}_n-\boldsymbol{\theta})\|}\right)$$

is that of an i.i.d. sample with uniform distribution over the unit sphere in $\mathbb{R}^d$. Tyler's estimator of shape is then $(\mathbf{T}\mathbf{T}')^{-1}$ which we turn into a scatter estimator in accordance with the integration condition in the definition of $\mathcal{F}_1$.

Another potential estimator of $\mathbf{\Sigma}$ is the *minimum covariance determinant* (MCD) estimator ([43],[44]). Both Tyler's and the MCD estimator are affine-invariant.

Letting $\hat{\boldsymbol{\vartheta}}_0 := (\boldsymbol{\theta}', (\text{vech}\hat{\mathbf{\Sigma}})', \mathbf{0}')'$ for some estimator $\hat{\mathbf{\Sigma}}$ satisfying Assumption (B), denote by $\phi^{(n)}_{\boldsymbol{\theta};f}$ the test rejecting the null hypothesis $\mathcal{H}_{0;f,\boldsymbol{\theta}}$ whenever

$$Q^{(n)}_{\boldsymbol{\theta};f} := (\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0))'(\mathbf{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0)$$

exceeds the α-upper quantile $\chi^2_{d;1-\alpha}$ of the chi-squared distribution with d degrees of freedom. This asymptotic null distribution easily follows from the asymptotic normality of $\mathbf{\Delta}_3(\boldsymbol{\vartheta}_0)$ and the fact that $\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0) - \mathbf{\Delta}_3(\boldsymbol{\vartheta}_0)$ is $o_{\mathrm{P}}(1)$ under $\mathcal{H}_{0;f,\boldsymbol{\theta}}$ as $n \to \infty$. The test $\phi^{(n)}_{\boldsymbol{\theta};f}$ is locally and asymptotically optimal for $\mathcal{H}_{0;f,\boldsymbol{\theta}}$ against $\mathcal{H}_{1;f,\Pi,\boldsymbol{\theta}}$ (see Theorem 3.1 for its precise optimality properties).

Elementary algebra yields

$$Q^{(n)}_{\boldsymbol{\theta};f} = n(\bar{\mathbf{X}}-\boldsymbol{\theta}_0)'\hat{\mathbf{\Sigma}}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\theta}_0) =: Q^{(n)}_{\boldsymbol{\theta}}.$$

This expression is particularly striking, as it does *not* depend on the underlying radial density f. In other words, every parametric specified-f experiment leads to the same optimal test statistic $Q^{(n)}_{\boldsymbol{\theta}}$, so that $\phi^{(n)}_{\boldsymbol{\theta}} := \phi^{(n)}_{\boldsymbol{\theta};f}$ is *uniformly* (in f) optimal in the semiparametric unspecified-f experiment. This is an extremely rare feature. Another remarkable fact is that the skewing function Π plays no role in $Q^{(n)}_{\boldsymbol{\theta}}$, which means that optimality holds uniformly against *all* skew-elliptical alternatives. Finally, the alert reader has noticed the familiar form of $Q^{(n)}_{\boldsymbol{\theta}}$, which is nothing else but the classical Hotelling test statistic for location. Optimal testing for ellipticity with specified location thus, somewhat disappointingly, mostly boils down to testing for location.

The following theorem summarizes the properties of $\phi^{(n)}_{\boldsymbol{\theta}}$.

Theorem 3.1. *Let $f \in \mathcal{F}_1$ and suppose that Assumptions (A1), (A2), and (B) hold, and that the skewing function Π is continuously differentiable at 0, with $\dot{\Pi}(0) \neq 0$. Then,*
(i) under $\mathcal{H}_{0;\boldsymbol{\theta}}$, $Q_{\boldsymbol{\theta}}^{(n)} \xrightarrow{\mathcal{D}} \chi_d^2$ as $n \to \infty$, so that $\phi_{\boldsymbol{\theta}}^{(n)}$ has asymptotic level α;
(ii) under $\bigcup_{\boldsymbol{\Sigma} \in \mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, n^{-1/2}\boldsymbol{\tau}_3^{(n)}; g, \Pi}$ with $g \in \mathcal{F}_1$, $Q_{\boldsymbol{\theta}}^{(n)}$ is asymptotically non-central chi-square with d degrees of freedom and non-centrality parameter $4(\dot{\Pi}(0))^2 \boldsymbol{\tau}_3' \boldsymbol{\tau}_3$, where $\boldsymbol{\tau}_3 = \lim_{n\to\infty} \boldsymbol{\tau}_3^{(n)}$[1]*;*
(iii) the test $\phi_{\boldsymbol{\theta}}^{(n)}$ is locally and asymptotically maximin, at asymptotic level α, for testing $\mathcal{H}_{0;\boldsymbol{\theta}}$ against $\mathcal{H}_{1;\Pi,\boldsymbol{\theta}} = \bigcup_{f \in \mathcal{F}_1, \boldsymbol{\Sigma} \in \mathcal{S}_d, \boldsymbol{\lambda} \in \mathbb{R}^d \setminus \{\mathbf{0}\}} \mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}; f, \Pi}$. The test is thus uniformly (in f) optimal against any type of generalized skew-elliptical alternative as defined in (2.2)*.*

The proof is provided in Appendix B. The explicit expression

$$1 - F_{\chi_d'^2}(\chi^2_{d;1-\alpha}, 4(\dot{\Pi}(0))^2 \boldsymbol{\tau}_3' \boldsymbol{\tau}_3) = Q_{d/2}\left(2|\dot{\Pi}(0)|(\boldsymbol{\tau}_3'\boldsymbol{\tau}_3)^{1/2}, (\chi^2_{d;1-\alpha})^{1/2}\right)$$

of the asymptotic power of $\phi_{\boldsymbol{\theta}}^{(n)}$ against local alternatives of the form $\bigcup_{\boldsymbol{\Sigma} \in \mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, n^{-1/2}\boldsymbol{\tau}_3^{(n)}; g, \Pi}$ readily follows from part *(ii)* of the theorem ($F_{\chi_d'^2}$ stands for the distribution function of the non-central chi-square distribution with d degrees of freedom, $Q_M(\cdot, \cdot)$ for the *Marcum Q-function*).

4. Optimal parametric and semiparametric tests: unspecified $\boldsymbol{\theta}$

In some applications, maintaining a specified value of $\boldsymbol{\theta}$ (often, $\boldsymbol{\theta} = \mathbf{0}$) under the alternative does make sense. The test described in Theorem 3.1 then is a genuine test of ellipticity. In most cases, however, that assumption of a specified center is impossible or unrealistic—or just unclear: what is the "center" of an asymmetric distribution? The same test then no longer qualifies as a test of ellipticity. Moreover, as shown in Section 5, the impacts of location shift and non-ellipticity may cancel each other, with the consequence that obviously non-elliptical shifted distributions remain completely undetected (see Sections 5.5 and 6 for numerical evidence). Therefore, let us consider the case of an unspecified $\boldsymbol{\theta}$.

Instances of estimators of $\boldsymbol{\theta}$ that satisfy Assumption (B) and turn out to be useful in this section are the *spatial median* of Möttönen and Oja [41] or the (fast) MCD-based location estimator (Rousseeuw and Driessen [44]). Again, we shall first construct Le Cam efficient parametric tests (Section 4.1) and then turn them into semiparametrically efficient tests (Section 4.2).

Inspection of the Fisher information matrix (2.3) reveals that the scores for location and skewness are not asymptotically independent. Estimating the unknown location thus has a cost in terms of power against ellipticity. The family

[1]Here and in the sequel, several asymptotic results are established for sequences of perturbations of the form $n^{-1/2}\boldsymbol{\tau}_3^{(n)}$ such that $\boldsymbol{\tau}_3^{(n)}$ converges to $\boldsymbol{\tau}_3$. Clearly, since $\boldsymbol{\tau}_3^{(n)}$ is bounded, converging subsequences always exist; the asymptotic statement then holds along any such subsequence. This is tacitly assumed below whenever defining $\boldsymbol{\tau}_3$ as the limit of a sequence $\boldsymbol{\tau}_3^{(n)}$.

of generalized skew-elliptical distributions, moreover, is infamous for yielding singular Fisher information matrices in the vicinity of symmetry, which is precisely the situation we are interested in. In presence of such a singularity, the scores for skewness and location are perfectly colinear, with the consequence that the corresponding α-level optimal test for symmetry is the trivial test $\phi = \alpha$. Fortunately, this extreme situation only occurs at the multinormal distribution ([36], [16], [17]). Testing for multinormality against generalized skew-normality thus requires a special treatment (reparametrization and ULAN with slower contiguity rates), which is beyond the scope of this paper.

4.1. Optimal parametric tests: unspecified $\boldsymbol{\theta}$

Fix a radial density $f \in \mathcal{F}_1$ that is not Gaussian. The impact on the central sequence for skewness $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0)$ of a root-n perturbation of $\boldsymbol{\theta}$ is classically neutralized by projecting $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0)$ onto the subspace orthogonal to $\boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_0)$ in the metric of the information matrix, yielding the *f-efficient central sequence for skewness*

$$\boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\vartheta}_0) := \boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0) - \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\boldsymbol{\vartheta}_0)\boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_0).$$

Clearly, this new central sequence remains orthogonal to $\boldsymbol{\Delta}_{f;2}(\boldsymbol{\vartheta}_0)$. This orthogonality to $\boldsymbol{\Delta}_{f;1}$ and $\boldsymbol{\Delta}_{f;2}$, combined with (2.5), allows us to replace the unknown parameters $\boldsymbol{\Sigma}$ and $\boldsymbol{\theta}$ with any consistent estimators $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\theta}}$ satisfying Assumption (B) without altering the asymptotic behavior of $\boldsymbol{\Delta}^{\dagger}_{f;3}$ under the null and under local alternatives. Under $\mathcal{H}_{0;f}$, $\boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\vartheta}_0)$, hence also $\boldsymbol{\Delta}^{\dagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0)$, is asymptotically normal with mean zero and covariance (the *f-efficient Fisher information for skewness*)

$$\boldsymbol{\Gamma}^{\dagger}_{f;33}(\boldsymbol{\vartheta}_0) := \boldsymbol{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0) - \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0).$$

Note that this matrix would be the zero matrix if f were Gaussian. The resulting optimal f-parametric test statistic then is of the form

$$\begin{aligned} Q^{(n)}_f \quad &:= \quad (\boldsymbol{\Delta}^{\dagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0))'\Big(\boldsymbol{\Gamma}^{\dagger}_{f;33}(\hat{\boldsymbol{\vartheta}}_0)\Big)^{-1}\boldsymbol{\Delta}^{\dagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0) \\ &= \quad \frac{\mathcal{J}_{d,f}}{(\mathcal{J}_{d,f}-d)}\frac{1}{n}\sum_{i,j=1}^{n}\Big[d_i(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}) - \frac{d}{\mathcal{J}_{d,f}}\varphi_f(d_i(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}))\Big] \\ &\qquad \times\Big[d_j(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}) - \frac{d}{\mathcal{J}_{d,f}}\varphi_f(d_j(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}))\Big]\mathbf{U}_i(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}))'\mathbf{U}_j(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}), \end{aligned}$$

and the corresponding test $\phi^{(n)}_f$ rejects $\mathcal{H}_{0;f}$ at asymptotic level α whenever $Q^{(n)}_f$ exceeds the chi-square quantile $\chi^2_{d;1-\alpha}$. The next theorem, the proof of which we give in Appendix C, summarizes the asymptotic properties of this test.

Theorem 4.1. *Let $f \in \mathcal{F}_1$ and suppose that Assumptions (A1), (A2), and (B) hold, and that the skewing function Π is continuously differentiable at 0, with derivative $\dot{\Pi}(0) \neq 0$. Then,* *(i) under* $\mathcal{H}_{0;f}$, $Q_f^{(n)} \xrightarrow{\mathcal{D}} \chi^2_d$ *as* $n \to \infty$, *so that* $\phi_f^{(n)}$ *has asymptotic level* α*;*

(ii) under $\bigcup_{\boldsymbol{\theta}\in\mathbb{R}^d}\bigcup_{\boldsymbol{\Sigma}\in\mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},n^{-1/2}\boldsymbol{\tau}_3^{(n)};f,\Pi}$, $Q_f^{(n)}$ *is asymptotically non-central chi-square with* d *degrees of freedom and non-centrality parameter* $4(\dot{\Pi}(0))^2\big((\mathcal{I}_{d,f}-d)/\mathcal{I}_{d,f}\big)\boldsymbol{\tau}_3'\boldsymbol{\tau}_3$, *where* $\boldsymbol{\tau}_3 = \lim_{n\to\infty}\boldsymbol{\tau}_3^{(n)}$

(iii) the test $\phi_f^{(n)}$ *is locally and asymptotically maximin, at asymptotic level* α, *for* $\mathcal{H}_{0;f}$ *against* $\mathcal{H}_{1;f,\Pi} = \bigcup_{\boldsymbol{\theta}\in\mathbb{R}^d,\boldsymbol{\Sigma}\in\mathcal{S}_d,\boldsymbol{\lambda}\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi}$. *The test is thus optimal against any type of generalized skew-*f *alternative.*

Summing up, the test $\phi_f^{(n)}$ is (parametrically) optimal against any type of generalized skew-f alternative (f specified).

4.2. Optimal semiparametric tests: unspecified $\boldsymbol{\theta}$

Consider now the general null hypothesis $\mathcal{H}_0$ of elliptical symmetry with unspecified center $\boldsymbol{\theta}$. Since the central sequence for skewness $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0)$ does not depend on the actual radial density, the ideal test for the case of unspecified f and $\boldsymbol{\theta}$ should be based on $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0)$. But $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0)$ also depends on $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$, which therefore have to be replaced with estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ and, unfortunately, the impact of that substitution does depend on the actual radial density (denote it as g).

Let $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ satisfy Assumption (B). The asymptotic linearity property (note that (2.5) applies under any $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$ thanks to the fact that $\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0)$ does not depend on g) yields, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$ as $n\to\infty$,

$$\boldsymbol{\Delta}_3(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}},\mathbf{0}) - \boldsymbol{\Delta}_3(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) = -\mathrm{Cov}_g\left[\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0),\boldsymbol{\Delta}_{g;1}(\boldsymbol{\vartheta}_0)\right] n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) + o_{\mathrm{P}}(1)$$

where $\mathrm{Cov}_g\left[\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0),\boldsymbol{\Delta}_{g;1}(\boldsymbol{\vartheta}_0)\right] = 2\dot{\Pi}(0)\boldsymbol{\Sigma}^{-1/2}$. This is a non-zero quantity the projection of the previous section cannot cancel out for all g. Therefore, a “deeper projection” is required to obtain an f-efficient central sequence that is orthogonal, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$, to the g-based central sequence $\boldsymbol{\Delta}_{g;1}(\boldsymbol{\vartheta}_0)$, for any g. This deeper projection is taken care of by

$$\boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0) := \boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0) - 2\dot{\Pi}(0)\boldsymbol{\Sigma}^{-1/2}\left[\mathrm{Cov}_g\left[\boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_0),\boldsymbol{\Delta}_{g;1}(\boldsymbol{\vartheta}_0)\right]\right]^{-1}\boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_0)$$

which, unfortunately, depends on the unspecified g again. Simple algebra yields $\mathrm{Cov}_g\left[\boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_0),\boldsymbol{\Delta}_{g;1}(\boldsymbol{\vartheta}_0)\right] = \frac{1}{d}\mathcal{K}_{d,f,g}\boldsymbol{\Sigma}^{-1}$, with

$$\mathcal{K}_{d,f,g} := \int_0^\infty \left(\varphi_f'(r) + \frac{d-1}{r}\varphi_f(r)\right)\frac{1}{\mu_{d-1;g}} r^{d-1} g(r)\,dr$$

where we denote by φ_f' the weak derivative of $r\mapsto\varphi_f(r)$ coinciding, in case φ_f is differentiable, with the usual derivative of $r\mapsto\varphi_f(r)$.

The existence of this latter quantity, however, requires a slight reinforcement of the assumptions on the reference radial densities f and the actual radial density g.

Assumption (A3) The mapping $r \mapsto f^{1/2}(r)$ belongs to $W^{2,2}(\mathbb{R}_0, \mu_{d-1})$, $0 \neq \left| \int_0^\infty \left(\varphi_f'(r) + \frac{d-1}{r}\varphi_f(r) \right) r^{d-1} f(r) dr \right| < \infty$, and $\int_0^\infty (\varphi(r))^{2+\epsilon} r^{d-1} f(r) dr < \infty$ for some $\epsilon > 0$.

It follows from the definition that, for any $f \in \mathcal{F}_1$ satisfying Assumptions (A1-A3), there exists a class of densities

$$\begin{aligned}\mathcal{F}_{1;f} \quad := \quad & \left\{ g \in \mathcal{F}_1 : 0 \neq \left| \int_0^\infty \left(\varphi_f'(r) + \frac{d-1}{r}\varphi_f(r) \right) r^{d-1} g(r) dr \right| < \infty \right. \\ & \qquad \left. \text{and} \int_0^\infty (\varphi_f(r))^{2+\epsilon_g} r^{d-1} g(r) dr < \infty \, \text{for some } \epsilon_g > 0 \right\}\end{aligned}$$

such that, for $g \in \mathcal{F}_{1;f}$, $\mathcal{K}_{d,f,g}$, hence $\boldsymbol{\Delta}^{\ddagger}_{fg;3}$, are well defined. Clearly, under Assumptions (A1) and (A3), f itself belongs to $\mathcal{F}_{1;f}$. The resulting projected central sequence is

$$\boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0) = 2n^{-1/2}\dot{\Pi}(0) \sum_{i=1}^n \left[d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}) - \frac{d}{\mathcal{K}_{d,f,g}} \varphi_f(d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})) \right] \mathbf{U}_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}).$$

This, through $\mathcal{K}_{d,f,g}$, still depends on the unknown g. But $\mathcal{K}_{d,f,g}$ can be estimated via

$$\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \frac{1}{n} \sum_{i=1}^n \left[\varphi_f'(d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})) + \frac{d-1}{d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})} \varphi_f(d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})) \right],$$

hence, *in fine*, just as for the entire test statistic, by $\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})$ with $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ satisfying Assumption (B). The following lemma establishes the consistency of $\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})$ as an estimator of $\mathcal{K}_{d,f,g}$.

Lemma 4.1. *Let $f \in \mathcal{F}_1$ and suppose that Assumptions (A1-A3) and (B) hold. Then, for any $g \in \mathcal{F}_{1;f}$, $\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}) - \mathcal{K}_{d,f,g} = o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}; g}$.*

The proof is provided in Appendix C.

With this estimator of $\mathcal{K}_{d,f,g}$, the efficient central sequence for skewness takes the final form

$$\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0) = 2n^{-1/2}\dot{\Pi}(0) \sum_{i=1}^n \left[d_i(\hat{\boldsymbol{\theta}}, \hat{\Sigma}) - \frac{d}{\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})} \varphi_f(d_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})) \right] \mathbf{U}_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}).$$

The corresponding test $\phi_f^{\ddagger(n)}$ rejects $\mathcal{H}_0$ at asymptotic level α whenever the test statistic $Q_f^{\ddagger(n)} := (\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0))'(\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\hat{\boldsymbol{\vartheta}}_0))^{-1}\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0)$, with

$$\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\hat{\boldsymbol{\vartheta}}_0) := \frac{4(\dot{\Pi}(0))^2}{nd} \sum_{i=1}^n \left[d_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}) - \frac{d}{\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})} \varphi_f(d_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})) \right]^2 \mathbf{I}_d$$

exceeds the chi-square quantile $\chi^2_{d;1-\alpha}$. The asymptotic distribution of $Q_f^{\ddagger(n)}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$ for any $g \in \mathcal{F}_{1;f}$ and its optimality properties are formally established in Theorem 4.2. For the sake of exposition, we first establish the following lemma (see Appendix C for a proof).

Lemma 4.2. *Let $f \in \mathcal{F}_1$ and suppose that Assumptions (A1-A3) and (B) hold. Then,*

(i) $\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0) - \boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$ *and*

(ii) $\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\hat{\boldsymbol{\vartheta}}_0) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$ for any $g \in \mathcal{F}_{1;f}$, where

$$\boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_0) := \frac{4(\dot{\Pi}(0))^2}{nd}\sum_{i=1}^{n}\left[d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}) - \frac{d}{\mathcal{K}_{d,f,g}}\varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\right]^2 \mathbf{I}_d.$$

With this result in hand, we finally can state the announced asymptotic results about $\phi_f^{\ddagger(n)}$ and $Q_f^{\ddagger(n)}$.

Theorem 4.2. *Let $f \in \mathcal{F}_1$ and suppose that Assumptions (A1-A3) and (B) hold, and that the skewing function Π is continuously differentiable at 0, with $\dot{\Pi}(0) \neq 0$. Then,*

(i) under $\bigcup_{g\in\mathcal{F}_{1;f},\,\boldsymbol{\theta}\in\mathbb{R}^d,\,\boldsymbol{\Sigma}\in\mathcal{S}_d}\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$, the test statistic $Q_f^{\ddagger(n)}$ is asymptotically χ^2_d as $n \to \infty$, so that the test $\phi_f^{\ddagger(n)}$ has asymptotic level α;

(ii) under $\bigcup_{\boldsymbol{\theta}\in\mathbb{R}^d}\bigcup_{\boldsymbol{\Sigma}\in\mathcal{S}_d}\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},n^{-1/2}\boldsymbol{\tau}_3^{(n)};g,\Pi}$ with $g \in \mathcal{F}_{1;f}$, $Q_f^{\ddagger(n)}$ is asymptotically non-central chi-square with d degrees of freedom and non-centrality parameter $4(\dot{\Pi}(0))^2 d\gamma^{-1}_{d,f,g}(1-\alpha_{d,f,g}/\mathcal{K}_{d,f,g})^2\boldsymbol{\tau}_3'\boldsymbol{\tau}_3$, where $\boldsymbol{\tau}_3 = \lim_{n\to\infty}\boldsymbol{\tau}_3^{(n)}$, $\alpha_{d,f,g} := \frac{1}{\mu_{d-1;g}}\int_0^\infty r\varphi_f(r)r^{d-1}g(r)dr$, and

$$\gamma_{d,f,g} := \frac{1}{\mu_{d-1;g}}\int_0^\infty \left[r - \frac{d}{\mathcal{K}_{d,f,g}}\varphi_f(r)\right]^2 r^{d-1}g(r)dr.$$

(iii) the test $\phi_f^{\ddagger(n)}$ is locally and asymptotically maximin, at asymptotic level α, when testing $\bigcup_{g\in\mathcal{F}_{1;f},\boldsymbol{\theta}\in\mathbb{R}^d,\boldsymbol{\Sigma}\in\mathcal{S}_d}\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$ against alternatives of the form $\mathcal{H}_{1;f,\Pi} = \bigcup_{\boldsymbol{\theta}\in\mathbb{R}^d,\,\boldsymbol{\Sigma}\in\mathcal{S}_d,\,\boldsymbol{\lambda}\in\mathbb{R}^d\setminus\{\mathbf{0}\}}\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\boldsymbol{\lambda};f,\Pi}$, irrespective of Π.

Part *(i)* of this Theorem easily follows from Lemma 4.2. The rest of the proof follows along the same lines as the proofs of Theorems 3.1 and 4.1; details are left to the reader. Note that the finiteness of $\gamma_{d,f,g}$ follows from our assumptions on g.

The test $\phi_f^{\ddagger(n)}$ thus is valid under any $g \in \mathcal{F}_{1;f}$—the entire nonparametric hypothesis of elliptical symmetry with unspecified center—and uniformly optimal against *any* type of generalized skew-f alternative. For each radial density f satisfying Assumptions (A1-A3), we thus get such a test $\phi_f^{\ddagger(n)}$. These tests are the main contribution of this paper, and achieve all our objectives: they have a simple asymptotic chi-squared distribution under the null hypothesis of ellipticity, they are affine-invariant (this follows directly from the affine-invariance of $d_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})$ and $\mathbf{U}_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})$), computationally fast, have a simple and intuitive form, only require finite moments of order 2, and offer much flexibility in the choice of the radial density f at which optimality is achieved (recall that a Gaussian f is excluded, though).

The choice of f can be guided by asymptotic relative efficiency profiles, which we now provide for various choices of f.

4.3. Asymptotic Relative Efficiencies

In this section, we compute Asymptotic Relative Efficiencies (AREs) for $\phi_f^{\ddagger(n)}$ with respect to the pseudo-Gaussian test of [8] as a common benchmark.

Define $m_k^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := n^{-1} \sum_{i=1}^n (d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}))^k$ and

$$\begin{aligned} \mathbf{S}_i^{\mathbf{U}}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := ((\mathbf{U}_{i1}(\boldsymbol{\theta}, \boldsymbol{\Sigma}))^2 \text{sign}(\mathbf{U}_{i1}(\boldsymbol{\theta}, \boldsymbol{\Sigma})), \ldots \\ \ldots, (\mathbf{U}_{id}(\boldsymbol{\theta}, \boldsymbol{\Sigma}))^2 \text{sign}(\mathbf{U}_{id}(\boldsymbol{\theta}, \boldsymbol{\Sigma})))'. \end{aligned} \tag{4.1}$$

When the location $\boldsymbol{\theta}$ is unspecified, the Gaussian efficient central sequence for Cassart's Fechner-asymmetry model is

$$\boldsymbol{\Delta}_{\mathcal{G}}(\boldsymbol{\vartheta}_0) = n^{-1/2} \sum_{i=1}^n d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma})(c_d(d+1) m_1^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \mathbf{U}_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}) - d_i(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \mathbf{S}_i^{\mathbf{U}}(\boldsymbol{\theta}, \boldsymbol{\Sigma}))$$

where $c_d = 4\Gamma(d/2)/((d^2-1)\sqrt{\pi}\Gamma(\frac{d-1}{2}))$, with Fisher information matrix under radial density g (note that $m_k^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ converges to $\frac{\mu_{d+k-1;g}}{\mu_{d-1;g}}$ under g)

$$\boldsymbol{\Gamma}_{\mathcal{G}}(\boldsymbol{\vartheta}_0) := \left(\frac{3}{d(d+2)} \frac{\mu_{d+3;g}}{\mu_{d-1;g}} - 2c_d^2(d+1) \frac{\mu_{d;g}\mu_{d+2;g}}{(\mu_{d-1;g})^2} + c_d^2 \frac{(d+1)^2}{d} \frac{\mu_{d;g}^2 \mu_{d+1;g}}{(\mu_{d-1;g})^3} \right) \mathbf{I}_d.$$

The expectation of $\boldsymbol{\Delta}_{\mathcal{G}}(\boldsymbol{\vartheta}_0)$ remains $\mathbf{0}$, and the asymptotic normality with covariance $\boldsymbol{\Gamma}_{\mathcal{G}}(\boldsymbol{\vartheta}_0)$ holds, under any g with finite fourth-order moment, that is, under $g \in \mathcal{F}_{p\mathcal{G}} := \left\{ f \in \mathcal{F}_1 : \mu_{d+3;f} = \int_{\mathbb{R}^+} r^{d+3} f(r) dr < \infty \right\}$.

The Gaussian test based on $\boldsymbol{\Delta}_{\mathcal{G}}(\boldsymbol{\vartheta}_0)$ thus can be used as a pseudo-Gaussian test: denote it as $\phi_{p\mathcal{G}}^{(n)}$. That test rejects the null hypothesis $\bigcup_{g \in \mathcal{F}_{p\mathcal{G}}, \boldsymbol{\theta} \in \mathbb{R}^d, \boldsymbol{\Sigma} \in \mathcal{S}_d} \mathrm{P}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}; g}^{(n)}$ of elliptical symmetry with unspecified g and $\boldsymbol{\theta}$ at asymptotic level α whenever the test statistic (with $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ satisfying Assumption (B)) $Q_{p\mathcal{G}}^{(n)} := (\boldsymbol{\Delta}_{\mathcal{G}}(\hat{\boldsymbol{\vartheta}}_0))'(\boldsymbol{\Gamma}_{\mathcal{G}}(\hat{\boldsymbol{\vartheta}}_0))^{-1} \boldsymbol{\Delta}_{\mathcal{G}}(\hat{\boldsymbol{\vartheta}}_0)$ exceeds $\chi_{d;1-\alpha}^2$. We refer to Chapter 3 of [8] for formal details.

In order to compute AREs with respect to $\phi_{p\mathcal{G}}^{(n)}$, we need its asymptotic distribution under the local skew-elliptical alternatives considered in this paper. This is the purpose of the following result, the proof of which is similar to those of Theorems 3.1 and 4.1 and is left to the reader.

Theorem 4.3. *Suppose that Assumptions (A1), (A2), and (B) hold, and that the skewing function* Π *is continuously differentiable at* 0 *with* $\dot{\Pi}(0) \neq 0$*. Then,*
(i) under $\bigcup_{g \in \mathcal{F}_{p\mathcal{G}}, \boldsymbol{\theta} \in \mathbb{R}^d, \boldsymbol{\Sigma} \in \mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}; g}$, $Q^{(n)}_{p\mathcal{G}}$ *is asymptotically* χ^2_d *as* $n \to \infty$*, so that* $\phi^{(n)}_{p\mathcal{G}}$ *has asymptotic level* α*;*
(ii) under $\bigcup_{\boldsymbol{\theta} \in \mathbb{R}^d} \cup_{\boldsymbol{\Sigma} \in \mathcal{S}_d} \mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, n^{-1/2}\boldsymbol{\tau}^{(n)}_3; g, \Pi}$ *with* $g \in \mathcal{F}_{p\mathcal{G}}$, $Q^{(n)}_{p\mathcal{G}}$ *is asymptotically non-central chi-square with non-centrality parameter*

$$\frac{64(\dot{\Pi}(0))^2(\Gamma(d/2)((d+1)\frac{\mu_{d;g}\mu_{d+1;g}}{(\mu_{d-1;g})^2} - d\frac{\mu_{d+2;g}}{\mu_{d-1;g}}))^2}{\pi((d^2-1)\Gamma((d-1)/2))^2 d^2 \gamma_{\mathcal{G}}} \boldsymbol{\tau}_3' \boldsymbol{\tau}_3,$$

where $\boldsymbol{\tau}_3 = \lim_{n\to\infty} \boldsymbol{\tau}^{(n)}_3$ *and*

$$\gamma_{\mathcal{G}} := \frac{3}{d(d+2)} \frac{\mu_{d+3;g}}{\mu_{d-1;g}} - 2c_d^2(d+1)\frac{\mu_{d;g}\mu_{d+2;g}}{(\mu_{d-1;g})^2} + c_d^2 \frac{(d+1)^2}{d} \frac{\mu^2_{d;g}\mu_{d+1;g}}{(\mu_{d-1;g})^3}.$$

Theorems 4.2 and 4.3 allow for computing the desired ARE values as squared ratios of local shifts.

Theorem 4.4. *Let* $f \in \mathcal{F}_1$*; suppose that Assumptions (A1-A3) and (B) hold, and that the skewing function* Π *is continuously differentiable at* 0 *with* $\dot{\Pi}(0) \neq 0$*. Then, the ARE of* $\phi^{\ddagger(n)}_f$ *with respect to* $\phi^{(n)}_{p\mathcal{G}}$ *under local alternatives of the form* $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, n^{-1/2}\boldsymbol{\tau}^{(n)}_3; g, \Pi}$ *with* $g \in \mathcal{F}_{1;f} \cap \mathcal{F}_{p\mathcal{G}}$ *is*

$$ARE_g(\phi^{\ddagger(n)}_f / \phi^{(n)}_{p\mathcal{G}}) = \frac{d^3\pi\left(1 - \alpha_{d,f,g}/\mathcal{K}_{d,f,g}\right)^2 \left((d^2-1)\Gamma((d-1)/2)\right)^2 \gamma_{\mathcal{G}}}{16\left(\Gamma(d/2)((d+1)\frac{\mu_{d;g}\mu_{d+1;g}}{(\mu_{d-1;g})^2} - d\frac{\mu_{d+2;g}}{\mu_{d-1;g}})\right)^2 \gamma_{d,f,g}}.$$

TABLE 4.1
AREs, with respect to $\phi^{(n)}_{p\mathcal{G}}$ and under several skew-t alternatives, of our tests $\phi^{\ddagger(n)}_{t_\nu}$ for various values of ν and the dimension d.

		Degrees of freedom of the underlying t density				
d	test	4.1	5	7	10	20
2	$\phi^{\ddagger(n)}_{t_4}$	10.968	1.964	1.305	1.156	1.085
	$\phi^{\ddagger(n)}_{t_5}$	10.912	1.978	1.342	1.208	1.155
	$\phi^{\ddagger(n)}_{t_7}$	10.630	1.955	1.358	1.249	1.223
	$\phi^{\ddagger(n)}_{t_{10}}$	10.172	1.892	1.345	1.261	1.264
	$\phi^{\ddagger(n)}_{t_{20}}$	8.997	1.705	1.262	1.231	1.287
3	$\phi^{\ddagger(n)}_{t_4}$	11.780	2.149	1.473	1.341	1.300
	$\phi^{\ddagger(n)}_{t_5}$	11.725	2.164	1.511	1.397	1.383
	$\phi^{\ddagger(n)}_{t_7}$	11.449	2.140	1.528	1.442	1.462
	$\phi^{\ddagger(n)}_{t_{10}}$	10.993	2.076	1.513	1.455	1.510
	$\phi^{\ddagger(n)}_{t_{20}}$	9.804	1.882	1.424	1.420	1.539
5	$\phi^{\ddagger(n)}_{t_4}$	12.867	2.410	1.729	1.646	1.706
	$\phi^{\ddagger(n)}_{t_5}$	12.818	2.423	1.765	1.703	1.794
	$\phi^{\ddagger(n)}_{t_7}$	12.564	2.401	1.783	1.751	1.886
	$\phi^{\ddagger(n)}_{t_{10}}$	12.132	2.338	1.767	1.766	1.945
	$\phi^{\ddagger(n)}_{t_{20}}$	10.964	2.141	1.670	1.724	1.983
10	$\phi^{\ddagger(n)}_{t_4}$	7.486	2.759	2.117	2.170	2.548
	$\phi^{\ddagger(n)}_{t_5}$	14.202	2.770	2.143	2.215	2.626
	$\phi^{\ddagger(n)}_{t_7}$	14.008	2.752	2.158	2.256	2.719
	$\phi^{\ddagger(n)}_{t_{10}}$	13.654	2.699	2.143	2.270	2.786
	$\phi^{\ddagger(n)}_{t_{20}}$	12.618	2.519	2.047	2.224	2.832

Table 4.1 provides numerical values of the AREs for various skew-t alternatives. All ARE values are larger than one, sometimes quite significantly; as a rule, they decrease with the degrees of freedom of the underlying Student, and increase with the dimension. The test for which the reference f coincides with the actual g yields the maximal value of ARE_g, as it should. Note that we deliberately opted for the test $\phi^{\ddagger(n)}_{t_4}$ instead of $\phi^{\ddagger(n)}_{t_{4.1}}$: hence, the highest values of $ARE_{t_{4.1}}$ are not shown here.

5. Comparative finite-sample study

In this section we investigate, via Monte Carlo simulations, the finite-sample properties of the tests we are proposing and some of their competitors—first for specified location (Section 5.3) and then for unspecified location (Section 5.4). We start with a brief description of the competing methods to be considered in this study.

5.1. Competing methods: specified location

Most tests proposed in the literature are dealing with the specified-location problem We selected the following two, proposed by Baringhaus [5] and Cassart [8], respectively.

(a) Baringhaus [5] proposes a class of tests $\phi^{(n)}_{\text{Bar},\boldsymbol{\theta}}$ based on

$$B^{(n)} := \frac{1}{n^2}\sum_{i,j=1}^{n} h(\mathbf{U}_i(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}})'\mathbf{U}_j(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}}))(n-\max{(R_i,R_j)}+1), \tag{5.1}$$

where h is defined over $[-1,1]$ and satisfies some regularity conditions, $\hat{\boldsymbol{\Sigma}}$ is Tyler's estimator of scatter, and R_i is the rank of $\|\hat{\boldsymbol{\Sigma}}^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta})\|$ among $\|\hat{\boldsymbol{\Sigma}}^{-1/2}(\mathbf{X}_1 - \boldsymbol{\theta})\|,\ldots,\|\hat{\boldsymbol{\Sigma}}^{-1/2}(\mathbf{X}_n - \boldsymbol{\theta})\|$. In our simulations we chose $h(t) = \left(\frac{2}{17/8-t}\right)^{1/2} - 1$, $t\in[-1,1]$ because the asymptotic null distribution of $B^{(n)}$ then coincides (up to a multiplicative constant) with that of the squared Kolmogorov-Smirnov statistic for the problem under study (other choices of h would require simulation-based approximations of limiting null distributions). No moment assumptions are required. Baringhaus [5] actually introduced $\phi^{(n)}_{\text{Bar},\boldsymbol{\theta}}$ as a test for spherical symmetry (with $\mathbf{I}_d$ instead of $\hat{\boldsymbol{\Sigma}}$ in (5.1)). Empirical sphericization via the Tyler estimator $\hat{\boldsymbol{\Sigma}}$ turns it into a test for elliptical symmetry; this has been proposed by [12] who establishes (via simulations) the validity of the procedure in dimension $d = 2$.

(b) The pseudo-Gaussian tests $\phi^{(n)}_{p\mathcal{G},\boldsymbol{\theta}}$ described by [8] achieve Le Cam optimality against the Fechner-type multinormal alternatives defined there (Chapter 3). When the location $\boldsymbol{\theta}$ is known, the test $\phi^{(n)}_{p\mathcal{G},\boldsymbol{\theta}}$ rejects the hypothesis of elliptical symmetry with location $\boldsymbol{\theta}$ at asymptotic level α whenever

$$\frac{8}{3nm_4^{(n)}}\sum_{i,j=1}^{n}(d_i(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}}))^2(d_j(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}}))^2\mathbf{S}'_{\mathbf{U}_i(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}})}\mathbf{S}_{\mathbf{U}_j(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}})}$$

($\mathbf{S}_{\mathbf{U}_i(\boldsymbol{\theta},\hat{\boldsymbol{\Sigma}})}$ defined in (4.1)) exceeds the $(1-\alpha)$ chi-quare quantile $\chi^2_{d;1-\alpha}$. Finite moments of order four are required. For $\hat{\boldsymbol{\Sigma}}$, we still use Tyler's estimator.

5.2. Competing methods: unspecified location

The list of competitors is shorter in the unspecified-location case—despite the importance of the problem. Below, we are considering the unspecified-location pseudo-Gaussian tests $\phi^{(n)}_{p\mathcal{G}}$ proposed by Cassart [8], the Schott test $\phi^{(n)}_{\text{Schott}}$ [46], and the Koltchinskii–Sakhanenko test $\phi^{(n)}_{\text{K-S}}$ [34].

(c) Cassart's location-unspecified test $\phi^{(n)}_{p\mathcal{G}}$ is described in Section 4.3, where we refer to for details; its validity requires finite moments of order four.

(d) Schott's test $\phi^{(n)}_{\text{Schott}}$ [46] involves a test statistic based on fourth-order moments; its validity requires finite eighth-order moments. The underlying idea is that the fourth-order moment structure of an elliptical distribution is a scalar multiple of that of a normal distribution. Therefore, to test whether a given population has an elliptical distribution, it is sufficient to test whether its fourth-order moment structure matches that of a Gaussian population. A closed-form of the test statistic involves a long list of notations which we are skipping here—see [46] for details; its asymptotic distribution is chi-square with $d^2 + d(d-1)(d^2+7d-6)/24 - 1$ degrees of freedom.

(e) The Koltchinskii–Sakhanenko [34] test statistics $\phi^{(n)}_{\text{K-S}}$ are obtained as functionals of empirical processes indexed by special classes of functions. Let $\mathcal{F}_B$ be a class of Borel functions from $\mathbb{R}^d$ to $\mathbb{R}$. Their test statistics are functionals (for example, sup-norms) of the stochastic process

$$n^{-1/2}\sum_{i=1}^{n}\left(f(\hat{\mathbf{\Sigma}}^{-1/2}(\mathbf{X}_i-\hat{\boldsymbol{\theta}}))-m_f(d_i(\hat{\boldsymbol{\theta}},\hat{\mathbf{\Sigma}}))\right),$$

where $f \in \mathcal{F}_B$, $m_f(\rho)$ is the average value of f on the sphere with radius $\rho > 0$, and $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{\Sigma}}$ denote the sample average and covariance matrix, respectively. Several examples of classes $\mathcal{F}_B$ and test statistics based on the sup-norm of the above process are considered in [34]. Here we restrict to $\mathcal{F}_B := \left\{I_{0<||\mathbf{x}||\leq t}\psi\left(\frac{\mathbf{x}}{||\mathbf{x}||}\right) : \psi \in G_l, ||\psi||_2 \leq 1, t > 0\right\}$ where I_A stands for the indicator function of A, G_l for the linear space of spherical harmonics of degree less than or equal to l in $\mathbb{R}^d$, and $||\cdot||_2$ is the L^2-norm on the unit sphere $\mathcal{S}^{d-1}$ in $\mathbb{R}^d$. Critical values are obtained via a bootstrap procedure.

5.3. Finite-sample performance: specified location (Table 5.1)

Without loss of generality, fix $\boldsymbol{\theta} = \mathbf{0}$. In order to compare the null and non-null finite-sample behavior of our optimal semiparametric test $\phi^{(n)}_{\mathbf{0}}$ with that of the Baringhaus and pseudo-Gaussian tests $\phi^{(n)}_{\text{Bar},\boldsymbol{\theta}}$ and $\phi^{(n)}_{p\mathcal{G},\boldsymbol{\theta}}$, we consider samples of size $n = 100$ from various distributions in dimension $d = 3$, and calculate their rejection frequencies on the basis of $N = 3000$ replications. Under the null hypothesis, we consider the three-dimensional normal and Student t elliptical distributions with ν =2.1, 4.1, and 8 degrees of freedom, all with scatter $\mathbf{\Sigma} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$; the degrees of freedom 2.1 and 4.1 were selected as having finite moments of orders 2 and 4, respectively.

Alternatives are of four different types: normal and Student skew-elliptical (increasing $\boldsymbol{\lambda}$ values) in Table 5.1(a), sinh-arcsinh- (SAS-) transformed normal and $t_{4.1}$ (same $\mathbf{\Sigma}$ matrix as above; skewness parameters as indicated; kurtosis parameters all fixed to 1), location-scale Gaussian mixtures (LSGM), and mixtures of Gaussian distributions in Table 5.1(b).

The skew-elliptical alternatives are those against which $\phi^{(n)}_{\mathbf{0}}$ is optimal. The sinh-arcsinh-transformed families are families of skewed distributions in dimension d (see [33]) indexed by a d-dimensional parameter $\boldsymbol{\lambda}$ with the same interpretation as in skew-elliptical families.

As proposed by [52], we are considering a particular case of multivariate location-scale Gaussian mixtures (LSGM) yielding the so-called *multiple scaled generalized hyperbolic* (MSGH) distributions. Those distributions are indexed by parameters $\boldsymbol{\mu}$, $\mathbf{D}$, $\mathbf{A}$, $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\delta}$. More specifically, in Table 5.1(b), we chose the three-dimensional MSGH with $\boldsymbol{\gamma} = (2, 2, 2)'$, $\boldsymbol{\delta} = (1, 1, 1)'$, $\boldsymbol{\lambda} = (-1/2, 2, 1)'$, $\mathbf{A} = \mathbf{I}_3$, and $\mathbf{D} = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Finally, the mixtures of Gaussian distributions in Table 5.1(b) are of the form $\frac{1}{2}\mathcal{N}_3\big(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\big) + \frac{1}{2}\mathcal{N}_3\big(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2\big)$, with various locations $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and scatter matrices $\boldsymbol{\Sigma}_1 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$ and $\boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$, respectively. For each case, we considered increasingly skewed alternatives.

Inspection of Table 5.1 indicates that $\phi^{(n)}_{\boldsymbol{\theta}}$ (here $\phi^{(n)}_{\mathbf{0}}$) uniformly satisfies[2] the 5% level constraint and yields excellent powers for almost all settings. It is outperformed in two cases only:

(i) by Baringhaus' $\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$ test under skew-elliptical $t_{2.1}$; the same $\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$, however, is much weaker under all other skew distributions; this might be due to slow convergence, under heavy tails, to limit distributions;

(ii) by Cassart's pseudo-Gaussian test $\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$ under SAS-normal and LSGM distributions—the latter case, however, is explained by severe over-rejection (rejection frequency 21% at 5% nominal level!) under the null.

The results under Gaussian mixtures (bottom of Table 5.1) deserve some further comments. Note that the corresponding first column does not address a null hypothesis situation: although $\boldsymbol{\mu}_1 = \mathbf{0} = \boldsymbol{\mu}_2$, the resulting mixture is not elliptical. A comparison between columns 3 ($\boldsymbol{\mu}_1 = (0, 0, 0)'$, $\boldsymbol{\mu}_2 = (-1, 0, 0)'$) and 4 ($\boldsymbol{\mu}_1 = (1, 0, 0)'$, $\boldsymbol{\mu}_2 = (-1, 0, 0)'$) is particularly intriguing. The distribution in column 4 indeed is strictly "less elliptical" than in column 3; nevertheless, the power of $\phi^{(n)}_{\mathbf{0}}$, which is almost one in column 3, reduces to the nominal level in column 4. This is an illustration of the fact that specified-location tests cannot be considered as genuine ellipticity tests (see Section 5.5). Baringhaus apparently is less sensitive to that phenomenon—at the price, however, of very low powers under most values of $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$.

5.4. Finite-sample performance: unspecified location (Tables 5.2 and A.1–A.2)

The tests considered here are our optimal tests $\phi^{\ddagger(n)}_{f}$ (f elliptical Student with $\nu = 2.1$, 4, and 8 degrees of freedom), Schott's test $\phi^{(n)}_{\mathrm{Schott}}$, Cassart's pseudo-

[2]Within the confidence limits of the Monte Carlo experiment: with 3000 replications, a 5% confidence interval centered at the rejection frequencies shown in all tables in this section has approximate length 0.015.

Table 5.1

Rejection frequencies (out of $N = 3,000$ replications), under (a) various three-dimensional elliptical ($\boldsymbol{\lambda} = (0,0,0)$) and related skewed densities (increasing $\boldsymbol{\lambda}$ values) and (b) skewed SAS-normal, SAS-$t_{4.1}$, location-scale Gaussian mixtures (LSGM) and location Gaussian mixtures (increasing $\boldsymbol{\lambda}$) values), of our optimal specified-location($\boldsymbol{\theta}_0 = \mathbf{0}$) test $\phi^{(n)}_{\mathbf{0}}$, the Baringhaus test $\phi^{(n)}_{Bar,\mathbf{0}}$, and Cassart's pseudo-Gaussian test $\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$. The sample size is $n = 100$, the nominal probability level 5%.

(a) $\boldsymbol{\lambda}'$	$(0,0,0)$	$(0.1,-0.2,0)$	$(0.3,-0.6,0)$	$(0.1,0.1,0.1)$	$(0.2,0.2,0.2)$	$(0.3,0.3,0.3)$
test			Skew-normal			
$\phi^{(n)}_{\mathbf{0}}$	0.055	0.193	0.934	0.293	0.847	0.992
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.038	0.088	0.625	0.132	0.467	0.831
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.055	0.165	0.873	0.243	0.756	0.975
			Skew-$t_{2.1}$			
$\phi^{(n)}_{\mathbf{0}}$	0.039	0.106	0.651	0.145	0.522	0.815
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.035	0.092	0.675	0.147	0.521	0.874
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.012	0.036	0.157	0.040	0.126	0.235
			Skew-$t_{4.1}$			
$\phi^{(n)}_{\mathbf{0}}$	0.040	0.142	0.864	0.239	0.753	0.964
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.037	0.090	0.662	0.131	0.501	0.857
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.034	0.078	0.460	0.121	0.377	0.650
			Skew-t_8			
$\phi^{(n)}_{\mathbf{0}}$	0.050	0.170	0.902	0.277	0.813	0.990
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.034	0.081	0.638	0.131	0.475	0.862
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.042	0.119	0.688	0.186	0.587	0.879

(b) $\boldsymbol{\lambda}'$	$(0,0,0)$	$(0.05,-0.1,0.05)$	$(0.15,-0.3,0.15)$	$(0.1,0.1,0.1)$	$(0.2,0.2,0.2)$	$(0.3,0.3,0.3)$
test			SAS-normal			
$\phi^{(n)}_{\mathbf{0}}$	0.057	0.276	0.997	0.170	0.587	0.935
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.034	0.065	0.553	0.038	0.0583	0.113
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.052	0.363	0.999	0.273	0.842	0.998
			SAS-$t_{4.1}$			
$\phi^{(n)}_{\mathbf{0}}$	0.044	0.257	0.992	0.166	0.548	0.893
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.029	0.133	0.874	0.058	0.119	0.290
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.038	0.159	0.834	0.117	0.441	0.752
			LSGM			
$\phi^{(n)}_{\mathbf{0}}$	0.048	0.098	0.539	0.176	0.601	0.926
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.034	0.063	0.409	0.117	0.460	0.851
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.121	0.166	0.368	0.210	0.418	0.677
			Gaussian Mixture			
$\boldsymbol{\mu}'_1$	$(0,0,0)$	$(0,0,0)$	$(0,0,0)$	$(1,0,0)$	$(1,0,0)$	$(1,0,0)$
$\boldsymbol{\mu}'_2$	$(0,0,0)$	$(-0.5,0,0)$	$(-1,0,0)$	$(-1,0,0)$	$(-2,0,0)$	$(-3,0,0)$
$\phi^{(n)}_{\mathbf{0}}$	0.048	0.379	0.936	0.050	0.474	0.939
$\phi^{(n)}_{\mathrm{Bar},\mathbf{0}}$	0.034	0.080	0.264	0.710	0.395	0.340
$\phi^{(n)}_{p\mathcal{G},\mathbf{0}}$	0.052	0.272	0.836	0.063	0.417	0.893

Gaussian test $\phi^{(n)}_{p\mathcal{G}}$, and Koltchinskii and Sakhanenko's $\phi^{(n)}_{\text{K-S}}$ test. Table 5.2 is dealing with dimension $d = 2$, Tables A.1 and A.2 (in Appendix D) with $d = 3$. Because of its computational complexity, the Koltchinskii-Sakhanenko test is considered for $d = 2$ only.

We still consider samples of size $n = 100$, from the same distributions[3] as in 5.3, and calculate the rejection frequencies on the basis of $N = 3000$ replications. Again, our tests $\phi^{\ddagger(n)}_{f}$ outperform the other tests for almost all settings. The pseudo-Gaussian test performs very well for the SAS-normal distribution. In all other settings, the $\phi^{\ddagger(n)}_{t_\nu}$ tests yield the best results. Quite remarkably, $\phi^{\ddagger(n)}_{f}$ under Gaussian mixtures does not suffer at all the problems its specified-location counterpart was exhibiting in Table 5.1, and uniformly dominates all its competitors.

5.5. The pitfalls of specified-location tests

We already stressed the fact that most tests available in the literature are dealing with the null hypothesis of specified-location ellipticity. Those tests, as a rule, are reasonably powerful at detecting either elliptical location alternatives (a simple shift in the null distribution) or fixed-location violations of ellipticity. Problems occur when both violations are present, with opposite impacts on the test statistic: powers then completely collapse.

To showcase this, we ran our tests $\phi^{(n)}_{\mathbf{0}}$ and $\phi^{\ddagger;(n)}_{t_4}$ against simulated (3000 replications) 10-dimensional[4] Gaussian mixtures of the form

$$0.8\,\mathcal{N}_{10}(10\,\mathbf{e}_{1;10}, \mathbf{I}_{10}) + 0.2\,\mathcal{N}_{10}(-10\,\mathbf{e}_{1;10}, \mathbf{I}_{10}) - \boldsymbol{\delta}_i, \qquad i = 1, 2,$$

[3]For $d = 2$, we considered the MSGH distribution with parameter values $\boldsymbol{\gamma} = (2, 2)'$, $\boldsymbol{\delta} = (1, 1)'$, $\boldsymbol{\lambda} = (-1/2, 2)'$, $\mathbf{A} = \mathbf{I}_2$, and $\mathbf{D} = \begin{bmatrix} \sqrt{2+\sqrt{2}}/2 & -\sqrt{2-\sqrt{2}}/2 \\ \sqrt{2-\sqrt{2}}/2 & \sqrt{2+\sqrt{2}}/2 \end{bmatrix}$.

[4]The higher the dimension, the more serious the problem.

TABLE 5.2

Rejection frequencies (out of $N = 3{,}000$ replications), under various bivariate elliptical ($\boldsymbol{\lambda} = (0,0)$) and related skewed densities (increasing $\boldsymbol{\lambda}$ values), of our unspecified-location optimal tests $\phi_f^{\ddagger;(n)}$ (f the bivariate t distributions with 2.1, 4, and 8 degrees of freedom), Schott 's test $\phi_{Schott}^{(n)}$, Cassart's pseudo-Gaussian test $\phi_{p\mathcal{G}}^{(n)}$, and the Koltchinskii–Sakhanenko test $\phi_{K\text{-}S}^{(n)}$ for the null hypothesis of ellipticity with unspecified location. The sample size is $n = 100$, the nominal probability level 5%.

(b) $\boldsymbol{\lambda}'$	$(0,0)$	$(1,-1)$	$(1,1)$	$(2,2)$	$(3,3)$	$(0,0)$	$(0.15,-0.2)$	$(0.15,0.15)$	$(0.3,0.3)$	$(0.45,0.45)$
test	Skew-normal					SAS-normal				
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.044	0.052	0.129	0.501	0.684	0.049	0.254	0.229	0.732	0.973
$\phi_{t_4}^{\ddagger(n)}$	0.043	0.053	0.131	0.510	0.691	0.046	0.261	0.233	0.731	0.974
$\phi_{t_8}^{\ddagger(n)}$	0.044	0.055	0.129	0.502	0.679	0.049	0.255	0.230	0.710	0.964
$\phi_{\text{Schott}}^{(n)}$	0.034	0.040	0.035	0.036	0.042	0.040	0.035	0.034	0.028	0.034
$\phi_{p\mathcal{G}}^{(n)}$	0.045	0.040	0.064	0.139	0.192	0.051	0.420	0.284	0.808	0.985
$\phi_{\text{K-S}}^{(n)}$	0.048	0.047	0.065	0.095	0.116	0.056	0.096	0.069	0.121	0.213
	Skew-$t_{2.1}$					SAS-$t_{4.1}$				
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.033	0.294	0.445	0.621	0.675	0.040	0.168	0.120	0.368	0.698
$\phi_{t_4}^{\ddagger(n)}$	0.030	0.230	0.349	0.509	0.561	0.038	0.139	0.109	0.325	0.618
$\phi_{t_8}^{\ddagger(n)}$	0.022	0.165	0.260	0.376	0.431	0.037	0.119	0.088	0.267	0.511
$\phi_{\text{Schott}}^{(n)}$	0.265	0.285	0.309	0.340	0.324	0.060	0.061	0.059	0.061	0.067
$\phi_{p\mathcal{G}}^{(n)}$	0.020	0.101	0.152	0.221	0.265	0.037	0.171	0.096	0.308	0.578
$\phi_{\text{K-S}}^{(n)}$	0.057	0.211	0.341	0.473	0.539	0.059	0.086	0.075	0.131	0.229
	Skew-$t_{4.1}$					LSGM				
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.043	0.291	0.535	0.846	0.912	0.038	0.085	0.094	0.188	0.377
$\phi_{t_4}^{\ddagger(n)}$	0.040	0.266	0.482	0.775	0.844	0.038	0.076	0.079	0.162	0.328
$\phi_{t_8}^{\ddagger(n)}$	0.036	0.222	0.409	0.675	0.734	0.036	0.052	0.072	0.135	0.268
$\phi_{\text{Schott}}^{(n)}$	0.058	0.064	0.067	0.085	0.093	0.285	0.288	0.289	0.278	0.290
$\phi_{p\mathcal{G}}^{(n)}$	0.035	0.153	0.244	0.369	0.386	0.033	0.058	0.053	0.108	0.209
$\phi_{\text{K-S}}^{(n)}$	0.056	0.082	0.110	0.179	0.213	0.241	0.251	0.245	0.293	0.334
						Gaussian Mixture				
	Skew-t_8					$\boldsymbol{\mu}_1'$ $(0,0)$	$(0,0)$	$(1,0)$	$(1,0)$	$(1,0)$
						$\boldsymbol{\mu}_2'$ $(0,0)$	$(-1,0)$	$(-1,0)$	$(-2,0)$	$(-3,0)$
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.046	0.170	0.374	0.767	0.871	0.044	0.199	0.475	0.575	0.552
$\phi_{t_4}^{\ddagger(n)}$	0.046	0.167	0.357	0.734	0.845	0.043	0.199	0.452	0.520	0.482
$\phi_{t_8}^{\ddagger(n)}$	0.045	0.158	0.326	0.674	0.790	0.046	0.192	0.407	0.444	0.391
$\phi_{\text{Schott}}^{(n)}$	0.043	0.030	0.033	0.036	0.054	0.082	0.085	0.126	0.276	0.570
$\phi_{p\mathcal{G}}^{(n)}$	0.038	0.101	0.175	0.312	0.365	0.036	0.102	0.151	0.123	0.092
$\phi_{\text{K-S}}^{(n)}$	0.052	0.053	0.080	0.117	0.146	0.049	0.077	0.110	0.162	0.306

with $\mathbf{e}_{1;10} = (1,0,0,0,0,0,0,0,0,0)'$, $\boldsymbol{\delta}_1 = (-6,0,0,0,0,0,0,0,0,0)'$, and $\boldsymbol{\delta}_2 = (6,0,0,0,0,0,0,0,0,0)'$, that is, we consider two distinct scenarios, *(a)* and *(b)*, say. Note that non-ellipticity is strictly the same under both scenarios: only locations differ, with *(b)* remaining centered at $\mathbf{0}$. The rejection frequencies are as follows: under *(a)*, $\phi_{\mathbf{0}}^{(n)}$ reaches 0.042 and $\phi_{t_4}^{\ddagger;(n)}$ 0.681, while under *(b)*, $\phi_{\mathbf{0}}^{(n)}$ reaches 1.000 and $\phi_{t_4}^{\ddagger;(n)}$ 0.685. It appears very clearly that the unspecified-location test makes no distinction between *(a)* and *(b)*, detecting asymmetry under both, while the location-specified test fails to detect non-ellipticity under *(a)*. The reason is that non-ellipticity and location shift under *(a)* have opposite effects on the test statistic, which cancel each other. On the contrary, under *(b)*, the specified-location test is stronger as it does not suffer from the loss of power due to the estimation of $\boldsymbol{\theta}$.

The conclusion is that one should be extremely cautious before concluding that ellipticity can or cannot be rejected on the basis of a specified-location test, and rather check whether the unspecified-location procedure does not lead to the opposite conclusion. This warning is all the more important in higher dimensions, where a plot of the observations does not help much: Section 6 provides a real-life example of this in dimension $d = 17$.

6. An empirical analysis of financial returns data

Elliptical symmetry with respect to the origin is a common assumption in the multivariate analysis of financial data. In this section, we are testing whether such assumption is acceptable on a dataset consisting of 18 years of daily returns from 17 major financial indexes from America (S&P500, NASDAQ, TSX, Merval, Bovespa and IPC), Europe/Middle East (AEX, ATX, BEL, DAX and CAC40), and East Asia/Oceania (HgSg, Nikkei, BSE, KOSPI , TSEC and AlOrd). The sample consists of 4619 observations, from January 7, 2000 through September 20, 2017. Those observations, of course, are serially dependent. In order to neutralize conditional heteroskedasticity, following the suggestion of [39] for elliptical and possibly heavy-tailed data, they were adjusted via AR(2)-GARCH(1,1) filtering.

We shall test for elliptical symmetry both about the fixed location $\boldsymbol{\theta}_0 = \mathbf{0}$ (a natural choice) and without specifying the center of symmetry. We thus compare our test $\phi_{\mathbf{0}}^{(n)}$ with our test $\phi_{t_4}^{\ddagger(n)}$ based on the elliptical t distribution with 4 degrees of freedom. For the entire 17-dimensional data set, we obtain for $\phi_{\mathbf{0}}^{(n)}$ a p-value of 0.18, hence do not reject elliptical symmetry with respect to $\mathbf{0}$. If the location is not specified, $\phi_{t_4}^{\ddagger(n)}$, with p-value virtually zero, very significantly rejects ellipticity. Now, we investigate this in more details, using a rolling window of three years. Table A.3 in Appendix D contains the p-values corresponding to the resulting 16 three-year periods. We still observe quite opposite conclusions of the two tests: the specified-location test essentially never rejects, while the unspecified-location test consistently does. The only explanation for this, which illustrates our warnings from Section 5.5, is that the actual location is not

0. The unspecified-location test, in case $\phi^{\ddagger(n)}_{t_4}$ and $\phi^{(n)}_{\mathbf{0}}$ yield strongly opposite conclusions, is thus far more reliable than the specified-location one, from which we can conclude that the assumption of ellipticity in this dataset is unlikely to be satisfied.

7. Conclusion

Based on a family of generalized skew-elliptical distributions, we are proposing tests for the null hypothesis of elliptical symmetry under specified and unspecified location, respectively. Theoretical ARE values and finite-sample simulations demonstrate their excellent performance, well beyond the context of skew-elliptical alternatives. The inherent unreliability of specified-location methods is stressed.

References

[1] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12** 171–178.
[2] Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61** 579–602.
[3] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65** 367–389.
[4] Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83** 715–726.
[5] Baringhaus, L. (1991). Testing for spherical symmetry of a multivariate distribution. *The Annals of Statistics* **19** 899–917.
[6] Beran, R. (1979). Testing for ellipsoidal symmetry of a multivariate density. *The Annals of Statistics* **7** 150–162.
[7] Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis* **79** 99–113.
[8] Cassart, D. (2007). Optimal tests for symmetry, PhD thesis, Univ. libre de Bruxelles, Brussels.
[9] Cassart, D., Hallin, M. and Paindaveine, D. (2008). Optimal detection of Fechner-asymmetry. *Journal of Statistical Planning and Inference* **138** 2499–2525.
[10] Diks, C. and Tong, H. (1999). A test for symmetries of multivariate probability distributions. *Biometrika* **86** 605–614.
[11] Dominicy, Y., Ilmonen, P. and Veredas, D. (2017). Multivariate Hill estimators. *International Statistical Review* **85** 108–142.
[12] Dyckerhoff, R., Ley, C. and Paindaveine, D. (2015). Depth-based runs tests for bivariate central symmetry. *Annals of the Institute of Statistical Mathematics* **67** 917–941.

[13] Garel, B. and Hallin, M. (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend. *Annals of the Institute of Statistical Mathematics* **47** 551–579.

[14] Genton, M. G. (2004). *Skew-elliptical distributions and their applications: a journey beyond normality*. CRC Press, Boca Raton, FL.

[15] Genton, M. G. and Loperfido, N. M. (2005). Generalized skew-elliptical distributions and their quadratic forms. *Annals of the Institute of Statistical Mathematics* **57** 389–401.

[16] Hallin, M. and Ley, C. (2012). Skew-symmetric distributions and Fisher information–a tale of two densities. *Bernoulli* **18** 747–763.

[17] Hallin, M. and Ley, C. (2014). Skew-symmetric distributions and Fisher information: the double sin of the skew-normal. *Bernoulli* **20** 1432–1453.

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

[19] Hallin, M. and Paindaveine, D. (2002a). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks. *The Annals of Statistics* **30** 1103–1133.

[20] Hallin, M. and Paindaveine, D. (2002b). Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence. *Bernoulli* **8** 87–815.

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

[22] Hallin, M. and Paindaveine, D. (2004b). Multivariate signed rank tests in vector autoregressive order identification. *Statistical Science* **19** 697–711.

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

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

[25] Hallin, M. and Paindaveine, D. (2008). Optimal rank-based tests for homogeneity of scatter. *Annals of Statistics* **36** 261–1298.

[26] Hallin, M. and Paindaveine, D. (2009). Optimal tests for the homogeneity of covariance, scale, and shape. *Journal of Multivariate Analysis* **100** 422–444.

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

[28] Hallin, M., Paindaveine, D. and Verdebout, T. (2010b). Testing for common principal components under heterokurticity. *Journal of Nonparametric Statistics* **22** 4879–895.

[29] Hallin, M., Paindaveine, D. and Verdebout, T. (2013). Optimal rank-based tests for common principal components. *Bernoulli* **22** 2524–2556.

[30] Hallin, M., Paindaveine, D. and Verdebout, T. (2014). Efficient R-

estimation of principal and common principal components. *Journal of the American Statistical Association* **109** 1071–1083.

[31] Hodgson, D. J., Linton, O. and Vorkink, K. (2002). Testing the capital asset pricing model efficiently under elliptical symmetry: A semiparametric approach. *Journal of Applied Econometrics* **17** 617–639.

[32] Huffer, F. W. and Park, C. (2007). A test for elliptical symmetry. *Journal of Multivariate Analysis* **98** 256–281.

[33] Jones, M. and Pewsey, A. (2009). Sinh-arcsinh distributions. *Biometrika* **96** 761–780.

[34] Koltchinskii, V. and Sakhanenko, L. (2000). Testing for ellipsoidal symmetry of a multivariate distribution. In *High Dimensional Probability II* 493–510. Springer.

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

[36] Ley, C. and Paindaveine, D. (2010). On the singularity of multivariate skew-symmetric models. *Journal of Multivariate Analysis* **101** 1434–1444.

[37] Li, R.-Z., Fang, K.-T. and Zhu, L.-X. (1997). Some QQ probability plots to test spherical and elliptical symmetry. *Journal of Computational and Graphical Statistics* **6** 435–450.

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

[39] Lombardi, M. J. and Veredas, D. (2009). Indirect estimation of elliptical stable distributions. *Computational Statistics & Data Analysis* **53** 2309–2324.

[40] Manzotti, A., Pérez, F. J. and Quiroz, A. J. (2002). A statistic for testing the null hypothesis of elliptical symmetry. *Journal of Multivariate Analysis* **81** 274–285.

[41] Möttönen, R. J. and Oja, H. (1995). Multivariate spatial sign and rank methods. *Journal of Nonparametric Statistics* **5** 201–213.

[42] Paindaveine, D. (2014). Elliptical symmetry. In *Encyclopedia of Environmetrics, 2nd edition* (A. H. El-Shaarawi and W. Piegorsch, eds.) 802–807. John Wiley & Sons, Chichester, UK.

[43] Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association* **79** 871–880.

[44] Rousseeuw, P. J. and Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics* **41** 212–223.

[45] Sakhanenko, L. (2008). Testing for ellipsoidal symmetry: A comparison study. *Computational Statistics & Data Analysis* **53** 565–581.

[46] Schott, J. R. (2002). Testing for elliptical symmetry in covariance-matrix-based analyses. *Statistics & Probability Letters* **60** 395–404.

[47] Serfling, R. J. (2006). Multivariate symmetry and asymmetry. In *Encyclopedia of Statistical Sciences, Second Edition*, (S. Kotz, N. Balakrishnan, C. B. Read and B. Vidakovic, eds.) **8** 5338–5345. Wiley Online Library.

[48] Swensen, A. R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *Journal of Multivariate Analysis* **16** 54–70.

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

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

[51] Vogel, D. and Fried, R. (2011). Elliptical graphical modelling. *Biometrika* **98** 935–951.

[52] Wraith, D. and Forbes, F. (2015). Location and scale mixtures of Gaussians with flexible tail behaviour: Properties, inference and application to multivariate clustering. *Computational Statistics & Data Analysis* **90** 61–73.

[53] Yang, G. L. and Le Cam, L. (2000). *Asymptotics in Statistics: some basic concepts*. Springer, Berlin, Germany.

[54] Zhu, L.-X. and Neuhaus, G. (2000). Nonparametric Monte Carlo tests for multivariate distributions. *Biometrika* **87** 919–928.

[55] Zhu, L.-X. and Neuhaus, G. (2004). Conditional tests for elliptical symmetry. *Journal of Multivariate Analysis* **84** 284–298.

Appendix A: Proof of Theorem 2.1

Our proof of Theorem 2.1 relies on Swensen [48], Lemma 1—more precisely, on its extension by Garel and Hallin [13]. Checking most of the conditions from Garel and Hallin [13] is a routine task, which we leave to the reader, and the only difficulty consists in establishing the quadratic mean differentiability of $(\boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}) \mapsto \underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)$, which follows from the fillowing lemma.

Lemma A.1. *Letting $f \in \mathcal{F}_1$, suppose that Assumptions (A1) and (A2) hold and that the skewing function Π is continuously differentiable at 0, with $\dot{\Pi}(0) \neq 0$. Let*

$$D_{\boldsymbol{\theta}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}, f, \Pi) := \frac{1}{2}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f)\, \varphi_f(\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\|)\boldsymbol{\Sigma}^{-1/2}\mathbf{U}(\boldsymbol{\theta}, \boldsymbol{\Sigma}),$$
$$D_{\boldsymbol{\Sigma}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}, f, \Pi) := \frac{1}{4}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f)\mathbf{P}_d(\boldsymbol{\Sigma}^{\otimes 2})^{-1/2} \times \operatorname{vec}\left(\psi_f(\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\|)\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\|\mathbf{U}(\boldsymbol{\theta}, \boldsymbol{\Sigma})\mathbf{U}'(\boldsymbol{\theta}, \boldsymbol{\Sigma}) - \mathbf{I}_d\right),$$

and

$$D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}} := \underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f)\dot{\Pi}(0)\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\|\mathbf{U}(\boldsymbol{\theta}, \boldsymbol{\Sigma}),$$

where $\mathbf{U}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})/\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\|$. Then,

$$(i) \int_{\mathbb{R}^d} \left\{\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\ell}, f, \Pi) - \underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f) - \boldsymbol{\ell}' D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\right\}^2 d\mathbf{x} = o(\|\boldsymbol{\ell}\|^2)$$

and

$$(ii) \int_{\mathbb{R}^d} \left\{ \underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}+\mathbf{t}, \boldsymbol{\Sigma}+\mathbf{H}, \boldsymbol{\ell}, f, \Pi) - \underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f) - \begin{pmatrix}\mathbf{t}\\ \operatorname{vech}\mathbf{H}\\ \boldsymbol{\ell}\end{pmatrix}' \begin{pmatrix} D_{\boldsymbol{\theta}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}, f, \Pi)\\ D_{\boldsymbol{\Sigma}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}, f, \Pi)\\ D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\end{pmatrix}\right\}^2 d\mathbf{x} = o\left(\left\|\begin{pmatrix}\mathbf{t}\\ \operatorname{vech}\mathbf{H}\\ \boldsymbol{\ell}\end{pmatrix}\right\|^2\right),$$

where $\mathbf{t} \in \mathbb{R}^d$, $\mathbf{H} \in \mathcal{S}_d$, $\boldsymbol{\ell} \in \mathbb{R}^d$, and $o(\|\cdot\|)$'s are taken for $\|\cdot\| \to 0$.

Proof. All $o(\|\cdot\|)$'s below are to be understood as $\|\cdot\| \to 0$. Starting with (i) and letting $\mathbf{y} := \boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})$, the integral takes the form

$$\int_{\mathbb{R}^d} \left\{\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\ell}, f, \Pi) - \underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, f) - \boldsymbol{\ell}' D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\right\}^2 d\mathbf{x}$$
$$= \int_{\mathbb{R}^d} \left\{\Pi^{1/2}(\boldsymbol{\ell}'\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})) - \Pi^{1/2}(0) - \Pi^{1/2}(0)\dot{\Pi}(0)\boldsymbol{\ell}'\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\right\}^2 \times 2c_{d,f}|\boldsymbol{\Sigma}|^{-1/2} f(\|\boldsymbol{\Sigma}^{-1/2}\mathbf{x}-\boldsymbol{\theta}\|)d\mathbf{x}$$
$$= \int_{\mathbb{R}^d} \left\{\Pi^{1/2}(\boldsymbol{\ell}'\mathbf{y}) - \Pi^{1/2}(0) - \Pi^{1/2}(0)\dot{\Pi}(0)\boldsymbol{\ell}'\mathbf{y}\right\}^2 2c_{d,f} f(\|\mathbf{y}\|)d\mathbf{y}.$$

Since $\Pi(0) = 1/2$, $\Pi^{1/2}(0)\dot{\Pi}(0) = (\dot{\Pi^{1/2}})(0)$. Using the fact that Π is bounded, we obtain, for some real constant C,

$$\left\{\Pi^{1/2}(\boldsymbol{\ell}'\mathbf{y}) - \Pi^{1/2}(0) - \Pi^{1/2}(0)\dot{\Pi}(0)\boldsymbol{\ell}'\mathbf{y}\right\}^2 \leq (2C + C\Pi(0)(\dot{\Pi}(0))^2\|\mathbf{y}\|^2) =: C^+(\|\mathbf{y}\|^2), \text{ say,}$$

where $\int C^+(\|\mathbf{y}\|^2)2c_{d,f}f(\|\mathbf{y}\|)d\mathbf{y} < \infty$ since $f \in \mathcal{F}_1$. The result follows from Lebesgue's dominated convergence theorem combined with the fact that

$$\left\{\Pi^{1/2}(\boldsymbol{\ell}'\mathbf{y}) - \Pi^{1/2}(0) - \Pi^{1/2}(0)\dot{\Pi}(0)\boldsymbol{\ell}'\mathbf{y}\right\}^2 = o(\|\boldsymbol{\ell}\|^2).$$

Turning to (ii), the integral there is bounded by $C_3(S_1 + S_2 + \|\boldsymbol{\ell}\|^2 S_3)$, where

$$S_1 := \int_{\mathbb{R}^d} \Bigg\{\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta} + \mathbf{t}, \boldsymbol{\Sigma} + \mathbf{H}, f) - \underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta}, \boldsymbol{\Sigma}, f) - \begin{pmatrix}\mathbf{t}\\ \mathrm{vech}\mathbf{H}\end{pmatrix}' \begin{pmatrix} D_{\boldsymbol{\theta}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}, f, \Pi)\\ D_{\boldsymbol{\Sigma}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0}, f, \Pi)\end{pmatrix}\Bigg\}^2 d\mathbf{x},$$

$$S_2 := \int_{\mathbb{R}^d} \Big\{\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta} + \mathbf{t}, \boldsymbol{\Sigma} + \mathbf{H}, \boldsymbol{\ell}, f, \Pi) - \underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta} + \mathbf{t}, \boldsymbol{\Sigma} + \mathbf{H}, f) - \boldsymbol{\ell}' D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta} + \mathbf{t}, \boldsymbol{\Sigma} + \mathbf{H}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\Big\}^2 d\mathbf{x},$$

$$S_3 := \int_{\mathbb{R}^d} \Big\| D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta} + \mathbf{t}, \boldsymbol{\Sigma} + \mathbf{H}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}} - D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\Big\|^2 d\mathbf{x}$$

and C_3 is a strictly positive real constant.

From Lemma A.1 of Hallin and Paindaveine [24], we know that, under Assumptions (A1) and (A2), S_1 is $o\left(\left\|\mathbf{t}', \mathrm{vech}'\mathbf{H}\right\|^2\right)$, hence also $o\left(\left\|\mathbf{t}', \mathrm{vech}'\mathbf{H}, \boldsymbol{\ell}'\right\|^2\right)$. It follows from (i) above that the same holds true for S_2. It thus remains to show that S_3 is $o(1)$ to complete the proof. This, however, follows from the quadratic mean continuity of $(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \mapsto D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)$, since $||D_{\boldsymbol{\lambda}}\underline{f}^{1/2}(\mathbf{x};\boldsymbol{\theta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, f, \Pi)||$ belongs to $L^2(\mathbb{R}^d, d\mathbf{x})$ in view of the fact that f admits finite moments of order 2. □

Appendix B: Proof of Theorem 3.1

The following notation will be convenient here and in Appendix C. Letting $\boldsymbol{\theta}_n := \boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}_1^{(n)}$ for some bounded sequence of d-dimensional vectors $\boldsymbol{\tau}_1^{(n)}$ and $\boldsymbol{\Sigma}_n := \boldsymbol{\Sigma} + n^{-1/2}\boldsymbol{\tau}_2^{(n)}$ for some bounded sequence of $d \times d$ matrices $\boldsymbol{\tau}_2^{(n)}$, define $\boldsymbol{\vartheta}_{0n} := (\boldsymbol{\theta}_n', \mathrm{vech}'\boldsymbol{\Sigma}_n, \mathbf{0}')'$.

Lemma A.2. *Let $h : \mathbb{R}^+ \to \mathbb{R}^+$ be such that h^p for $p = 1$ (resp., $p = 2$) is integrable with respect to the measure ν, where ν is absolutely continuous with respect to the Lebesgue measure. Then,*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} |h(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) - h(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||)|^p d\nu(\mathbf{x}) = 0$$

for $p = 1$ (resp., $p = 2$).

Proof. For any $\epsilon > 0$, we can choose h_ϵ from $C_c^\infty(\mathbb{R}^+)$ such that $||h - h_\epsilon||_{L^p(d\nu)} < \epsilon$. Then,

$$\begin{aligned}
&\lim_{n\to\infty} \int_{\mathbb{R}^d} \left| h(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) - h(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||) \right|^p d\nu(\mathbf{x}) \\
&\leq \lim_{n\to\infty} \int_{\mathbb{R}^d} \left| h(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) - h_\epsilon(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) \right|^p d\nu(\mathbf{x}) \\
&\quad + \lim_{n\to\infty} \int_{\mathbb{R}^d} \left| h(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||) - h_\epsilon(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||) \right|^p d\nu(\mathbf{x}) \\
&\quad + \lim_{n\to\infty} \int_{\mathbb{R}^d} \left| h_\epsilon(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) - h_\epsilon(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||) \right|^p d\nu(\mathbf{x}) \\
&\leq 2\epsilon^p + \lim_{n\to\infty} \int_{\mathbb{R}^d} \left| h_\epsilon(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) - h_\epsilon(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||) \right|^p d\nu(\mathbf{x}).
\end{aligned}$$

Given that $h_\epsilon \in C_c^\infty(\mathbb{R}^+)$, Lebesgue's dominated convergence theorem implies that the latter limit is zero. Now, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \left| h(||\mathbf{\Sigma}_n^{-1/2}(\mathbf{x} - \boldsymbol{\theta}_n)||) - h(||\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})||) \right|^p d\nu(\mathbf{x}) < 2\epsilon^p.$$

The claim follows. □

We now turn to the proof of Theorem 3.1.

Proof of Theorem 3.1

(i) To start with, let us show that

$$(\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0))'(\mathbf{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0) - (\mathbf{\Delta}_3(\boldsymbol{\vartheta}_0))'(\mathbf{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\mathbf{\Delta}_3(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma},\mathbf{0};f}$. The asymptotic linearity property combined with Lemma 4.4 of Kreiss [35] entails that $\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0) - \mathbf{\Delta}_3(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$ as $n \to \infty$. Hence,

$$(\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0))'(\mathbf{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\mathbf{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0) = (\mathbf{\Delta}_3(\boldsymbol{\vartheta}_0))'(\mathbf{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\mathbf{\Delta}_3(\boldsymbol{\vartheta}_0) + o_{\mathrm{P}}(1)$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\mathbf{\Sigma},\mathbf{0};f}$ and the asymptotic normality of $\mathbf{\Delta}_3(\boldsymbol{\vartheta}_0)$ yields the desired result for given f. This, however, holds for any $f \in \mathcal{F}_1$ and $\mathbf{\Sigma} \in \mathcal{S}_d$, so that *(i)* follows under the entire $\mathcal{H}_{0;\boldsymbol{\theta}}$.

(ii) By contiguity,

$$(\boldsymbol{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0))'(\boldsymbol{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\boldsymbol{\Delta}_3(\hat{\boldsymbol{\vartheta}}_0) - (\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0))'(\boldsymbol{\Gamma}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\boldsymbol{\Delta}_3(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},n^{-1/2}\boldsymbol{\tau}_3^{(n)};g,\Pi}$ for every $\boldsymbol{\Sigma}$. By the Central Limit Theorem, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$ and as $n \to \infty$

$$\begin{pmatrix} \boldsymbol{\Delta}_3(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) \\ (\boldsymbol{\tau}_2^{(n)})'\boldsymbol{\Delta}_{g;2}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) + (\boldsymbol{\tau}_3^{(n)})'\boldsymbol{\Delta}_3(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) \\ -\frac{1}{2}(\boldsymbol{\tau}_2^{(n)})'\boldsymbol{\Gamma}_{g;22}(\boldsymbol{\vartheta}_0)\boldsymbol{\tau}_2^{(n)} - \frac{1}{2}(\boldsymbol{\tau}_3^{(n)})'\boldsymbol{\Gamma}_{g;33}(\boldsymbol{\vartheta}_\mathbf{0})\boldsymbol{\tau}_3^{(n)} + o_{\mathrm{P}}(1) \end{pmatrix}$$
$$\xrightarrow{\mathcal{D}} \mathcal{N}_{d+1}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} 4(\dot{\Pi}(0))^2\mathbf{I}_d & 4(\dot{\Pi}(0))^2\boldsymbol{\tau}_3 \\ 4(\dot{\Pi}(0))^2\boldsymbol{\tau}_3' & \boldsymbol{\tau}_2'\boldsymbol{\Gamma}_{g;22}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0})\boldsymbol{\tau}_2 + \boldsymbol{\tau}_3'\boldsymbol{\tau}_3 4(\dot{\Pi}(0))^2 \end{pmatrix}\right)$$

for $\boldsymbol{\tau}_3 = \lim_{n\to\infty}\boldsymbol{\tau}_3^{(n)}$ and $\boldsymbol{\tau}_2 = \lim_{n\to\infty}\boldsymbol{\tau}_2^{(n)}$. The asymptotic distribution of$(\boldsymbol{\Gamma}_{f;33}(\boldsymbol{\vartheta}_\mathbf{0}))^{-1/2}\boldsymbol{\Delta}_3(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0})$ under the alternative then follows from Le Cam's Third Lemma.

(iii) The asymptotic level α of $\phi^{(n)}_{\boldsymbol{\theta}}$ under $\mathcal{H}_{0;\boldsymbol{\theta}}$ follows from the asymptotic normality provided under *(i)*. Local asymptotic maximinity is a consequence of the weak convergence to Gaussian shifts of the local skewness experiments. □

Appendix C: Proof of Theorem 4.1, Lemma 4.1, and Lemma 4.2

Proof of Theorem 4.1

(i) Let us show that

$$(\boldsymbol{\Delta}^\dagger_{f;3}(\hat{\boldsymbol{\vartheta}}_0))'\left(\boldsymbol{\Gamma}^\dagger_{f;33}(\hat{\boldsymbol{\vartheta}}_0)\right)^{-1}\boldsymbol{\Delta}^\dagger_{f;3}(\hat{\boldsymbol{\vartheta}}_0) - (\boldsymbol{\Delta}^\dagger_{f;3}(\boldsymbol{\vartheta}_0))'\left(\boldsymbol{\Gamma}^\dagger_{f;33}(\boldsymbol{\vartheta}_0)\right)^{-1}\boldsymbol{\Delta}^\dagger_{f;3}(\boldsymbol{\vartheta}_0) \quad \text{(A.1)}$$

is $o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$. Continuity of the Fisher information matrices and asymptotic linearity yield

$$\begin{aligned}
\boldsymbol{\Delta}^\dagger_{f;3}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}},\mathbf{0}) &= \boldsymbol{\Delta}_3(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}},\mathbf{0}) - \boldsymbol{\Gamma}_{f;13}(\hat{\boldsymbol{\vartheta}}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\hat{\boldsymbol{\vartheta}}_0)\boldsymbol{\Delta}_{f;1}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}},\mathbf{0}) \\
&= \boldsymbol{\Delta}_3(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}},\mathbf{0}) - \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\boldsymbol{\vartheta}_0)\boldsymbol{\Delta}_{f;1}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}},\mathbf{0}) + o_{\mathrm{P}}(1) \\
&= \boldsymbol{\Delta}_3(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) - \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) - \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\boldsymbol{\vartheta}_0)\boldsymbol{\Delta}_{f;1}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) \\
&+ \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}_{f;11}(\boldsymbol{\vartheta}_0)n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) + o_{\mathrm{P}}(1) \\
&= \boldsymbol{\Delta}_3(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) - \boldsymbol{\Gamma}_{f;13}(\boldsymbol{\vartheta}_0)\boldsymbol{\Gamma}^{-1}_{f;11}(\boldsymbol{\vartheta}_0)\boldsymbol{\Delta}_{f;1}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) + o_{\mathrm{P}}(1) \\
&= \boldsymbol{\Delta}^\dagger_{f;3}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) + o_{\mathrm{P}}(1)
\end{aligned}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$. The continuous mapping theorem implies that $\boldsymbol{\Gamma}^\dagger_{f;33}(\hat{\boldsymbol{\vartheta}}_0) - \boldsymbol{\Gamma}^\dagger_{f;33}(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$, so that $\left(\boldsymbol{\Gamma}^\dagger_{f;33}(\hat{\boldsymbol{\vartheta}}_0)\right)^{-1} - \left(\boldsymbol{\Gamma}^\dagger_{f;33}(\boldsymbol{\vartheta}_0)\right)^{-1}$ is $o_{\mathrm{P}}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$. A simple application of Slutsky's Lemma then yields the desired result that (A.1) is $o_{\rm P}(1)$; the asymptotic normality of $\boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\vartheta}_0)$ completes the proof of this part of the theorem.

(ii) By contiguity,

$$(\boldsymbol{\Delta}^{\dagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0))'(\boldsymbol{\Gamma}^{\dagger}_{f;33}(\hat{\boldsymbol{\vartheta}}_0))^{-1}\boldsymbol{\Delta}^{\dagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0) - (\boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\vartheta}_0))'(\boldsymbol{\Gamma}^{\dagger}_{f;33}(\boldsymbol{\vartheta}_0))^{-1}\boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\vartheta}_0) = o_{\rm P}(1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},n^{-1/2}\boldsymbol{\tau}_3^{(n)};f,\Pi}$ for every $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$. The Central Limit Theorem entails

$$\begin{pmatrix} \boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0}) \\ (\boldsymbol{\tau}^{(n)})'((\boldsymbol{\Delta}_{f;1}(\boldsymbol{\vartheta}_0))',(\boldsymbol{\Delta}_{f;2}(\boldsymbol{\vartheta}_0))',(\boldsymbol{\Delta}_{3}(\boldsymbol{\vartheta}_0))')' - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_f(\boldsymbol{\vartheta}_0)\boldsymbol{\tau}^{(n)} + o_{\rm P}(1) \end{pmatrix}$$
$$\xrightarrow{\mathcal{D}} \mathcal{N}_{d+1}\left(\begin{pmatrix}\mathbf{0}\\ 0\end{pmatrix}, \begin{pmatrix} 4(\dot{\Pi}(0))^2\dfrac{\mathcal{I}_{d,f}-d}{\mathcal{I}_{d,f}}\mathbf{I}_d & 4(\dot{\Pi}(0))^2\dfrac{\mathcal{I}_{d,f}-d}{\mathcal{I}_{d,f}}\boldsymbol{\tau}_3 \\ 4(\dot{\Pi}(0))^2\dfrac{\mathcal{I}_{d,f}-d}{\mathcal{I}_{d,f}}\boldsymbol{\tau}_3' & \boldsymbol{\tau}'\boldsymbol{\Gamma}_f(\boldsymbol{\vartheta}_0)\boldsymbol{\tau}\end{pmatrix}\right)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};f}$ for $\boldsymbol{\tau} = (\boldsymbol{\tau}_1',\boldsymbol{\tau}_2',\boldsymbol{\tau}_3')'$ with $\boldsymbol{\tau}_j = \lim_{n\to\infty}\boldsymbol{\tau}_j^{(n)}$ for $j = 1,2,3$. The asymptotic distribution of $(\boldsymbol{\Gamma}^{\dagger}_{f;33}(\boldsymbol{\vartheta}_0))^{-1/2}\boldsymbol{\Delta}^{\dagger}_{f;3}(\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0})$ under the alternative follows from Le Cam's Third Lemma.

(iii) The asymptotic level α of $\phi_f^{(n)}$ under $\mathcal{H}_{0;f}$ follows from the asymptotic normality provided under (i). Local asymptotic maximinity is a consequence of the weak convergence to Gaussian shifts of the local skewness experiments. □

Proof of Lemma 4.1

Rewrite the difference $\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}) - \mathcal{K}_{d,f,g}$ as

$$\widehat{\mathcal{K}}_{d,f}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\Sigma}}) - \widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma}) + \widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma}) - \mathcal{K}_{d,f,g}.$$

The Law of Large Numbers implies that $\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma}) - \mathcal{K}_{d,f,g} = o_{\rm P}(1)$ as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$. Letting $h(r) = \varphi_f'(r) + \frac{d-1}{r}\varphi_f(r)$ in Lemma A.2 with $p = 1$ (integrability w.r.t. $r^{d-1}g(r)dr$ holds since $g \in \mathcal{F}_{1;f}$), we get the L^1-convergence to zero of $\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n) - \widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})$, hence also

$$\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n) - \widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma}) = o_{\rm P}(1) \quad \text{as } n\to\infty \text{ under } \mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}. \tag{A.2}$$

This, combined with Lemma 4.4 of Kreiss [35], concludes the proof. □

Proof of Lemma 4.2

(i) Rewrite $\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0) - \boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0)$ as

$$\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0) - \boldsymbol{\Delta}^{\ddagger}_{f;3}(\boldsymbol{\vartheta}_0) + \boldsymbol{\Delta}^{\ddagger}_{f;3}(\boldsymbol{\vartheta}_0) - \boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0).$$

We have

$$\begin{aligned}\boldsymbol{\Delta}^{\ddagger}_{f;3}(\boldsymbol{\vartheta}_0)-&\boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0)\\ &= -2\dot{\Pi}(0)dn^{-1/2}\sum_{i=1}^{n}\varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\left(\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})}-\frac{1}{\mathcal{K}_{d,f,g}}\right).\end{aligned}$$

The continuous mapping theorem combined with the Law of Large Numbers, the fact that $\mathcal{K}_{d,f,g}\neq 0$, and the integrability of φ_f w.r.t. $r^{d-1}g(r)dr$ yield

$$\boldsymbol{\Delta}^{\ddagger}_{f;3}(\boldsymbol{\vartheta}_0)-\boldsymbol{\Delta}^{\ddagger}_{fg;3}(\boldsymbol{\vartheta}_0)=o_{\mathrm{P}}(1)$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$.

Next, let us show that $\boldsymbol{\Delta}^{\ddagger}_{f;3}(\hat{\boldsymbol{\vartheta}}_0)-\boldsymbol{\Delta}^{\ddagger}_{f;3}(\boldsymbol{\vartheta}_0)=o_{\mathrm{P}}(1)$ as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$. Therefore, note that (in view of the existence of finite second-order moments)

$$n^{-1/2}\sum_{i=1}^{n}\left[d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)-d_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\right]=o_{L^2}(1)$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$, which directly implies the convergence in probability. Let us show that, similarly,

$$\begin{aligned}n^{-1/2}\sum_{i=1}^{n}\Bigg[&\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)}\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)\\ &-\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})}\varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\Bigg]=o_{\mathrm{P}}(1).\end{aligned}$$

The latter expression can be rewritten as

$$\begin{aligned}&\left[\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)}-\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})}\right]n^{-1/2}\sum_{i=1}^{n}\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)\\ &\quad+\frac{n^{-1/2}}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})}\sum_{i=1}^{n}\left[\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)-\varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\right].\end{aligned} \tag{A.3}$$

Combined with the continuous mapping theorem, (A.2) implies that $\dfrac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)}-\dfrac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})}$ is $o_{\mathrm{P}}(1)$ as $n\to\infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$, which takes care of the first term in (A.3) provided that

$$n^{-1/2}\sum_{i=1}^{n}\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)=O_{\mathrm{P}}(1)$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$. This fact, however, follows from the Central Limit Theorem applied to $n^{-1/2}\sum_{i=1}^n \varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})$ and the L^2 convergence to zero of

$$n^{-1/2}\sum_{i=1}^n \left[\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n) - \varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\right],$$

which we shall establish now (that proof is also required for showing that the second term above is $o_{\mathrm{P}}(1)$). It is sufficient to show that

$$\begin{aligned}\mathrm{E}\big[\|\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\left[\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n) - \mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\right]\\ + \left[\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)) - \varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma}))\right]\mathbf{U}_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\|^2\big] =: E_1 + E_2 = o(1)\end{aligned}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$. Applying Hölder's inequality for $p = (2+\epsilon)/2$ and $q = (2+\epsilon)/\epsilon$, then using the fact that $g \in \mathcal{F}_{1;f}$ and $\|\mathbf{U}_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)\| \leq 1$, together with Lebesgue's dominated convergence theorem, one easily obtains that $E_1 = o(1)$. The convergence to zero of E_2 follows from Lemma A.2 with $h(r) = \varphi_f(r)$ and $p = 2$ (integrability w.r.t. $r^{d-1}g(r)dr$ holds for $g \in \mathcal{F}_{1;f}$). Since L^2 convergence implies convergence in probability, the Law of Large Numbers and the continuous mapping theorem applied to $\widehat{\mathcal{K}}^{-1}_{d,f}(\boldsymbol{\theta},\boldsymbol{\Sigma})$ complete the proof of part *(i)*.

(ii) We still have to show that $\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\hat{\boldsymbol{\vartheta}}_0) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$. In view of Lemma 4.4 of Kreiss [35], this reduces to proving that $\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_0) = o_{\mathrm{P}}(1)$. The latter rewrites as

$$\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) + \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_0) =: \frac{4(\dot{\Pi}(0))^2}{d}(\mathbf{A}_1 + \mathbf{A}_2).$$

Term $\mathbf{A}_1$ takes the form

$$\begin{aligned}\mathbf{A}_1 = &-\frac{2d}{n}\sum_{i=1}^n d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))\left(\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)} - \frac{1}{\mathcal{K}_{d,f,g}}\right)\\ &+\frac{d^2}{n}\sum_{i=1}^n(\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)))^2\left(\frac{1}{(\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))^2} - \frac{1}{(\mathcal{K}_{d,f,g})^2}\right).\end{aligned}$$

The proof of Lemma 4.1, combined with the continuous mapping theorem, implies that both

$$\frac{1}{\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)} - \frac{1}{\mathcal{K}_{d,f,g}} \quad\text{and}\quad \frac{1}{(\widehat{\mathcal{K}}_{d,f}(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n))^2} - \frac{1}{(\mathcal{K}_{d,f,g})^2}$$

are $o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\Sigma},\mathbf{0};g}$. Using similar arguments as above, one can show that

$$\frac{1}{n}\sum_{i=1}^n d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)\varphi_f(d_i(\boldsymbol{\theta}_n,\boldsymbol{\Sigma}_n)) - \frac{1}{n}\sum_{i=1}^n d_i(\boldsymbol{\theta},\boldsymbol{\Sigma})\varphi_f(d_i(\boldsymbol{\theta},\boldsymbol{\Sigma})) = o_{L^1}(1),$$

hence that $\frac{1}{n}\sum_{i=1}^{n} d_i(\boldsymbol{\theta}_n, \boldsymbol{\Sigma}_n)\varphi_f(d_i(\boldsymbol{\theta}_n, \boldsymbol{\Sigma}_n))$ is $O_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0};g}$. A similar conclusion holds for $\frac{1}{n}\sum_{i=1}^{n}(\varphi_f(d_i(\boldsymbol{\theta}_n, \boldsymbol{\Sigma}_n)))^2$, which is also $O_{\mathrm{P}}(1)$. It follows that $\widehat{\boldsymbol{\Gamma}}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) = o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0};g}$.

By Lemma A.2 with $h(r) = (r - \frac{d}{\mathcal{K}_{d,f,g}}\varphi_f(r))^2$ and $p = 1$ (integrability with respect to $r^{d-1}g(r)dr$ follows from the square integrability of r and $\varphi_f(r)$), we get the L^1 convergence, hence the convergence to zero in probability of $\boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_{0n}) - \boldsymbol{\Gamma}^{\ddagger}_f(\boldsymbol{\vartheta}_0)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}, \boldsymbol{\Sigma}, \mathbf{0};g}$. □

Appendix D: Additional numerical results

Tables A.1 and A.2 below are providing the finite-sample rejection frequencies, as described in Section 5.4, of the unspecified-location tests: our optimal tests $\phi^{\ddagger(n)}_f$ (f elliptical Student with $\nu = 2.1$, 4, and 8 degrees of freedom), Schott's test $\phi^{(n)}_{\mathrm{Schott}}$, and Cassart's pseudo-Gaussian test $\phi^{(n)}_{p\mathcal{G}}$, in dimension $d = 3$.

Table A.3 shows the p-values for the optimal semiparametric test $\phi^{(n)}_{\mathbf{0}}$ (specified location $\boldsymbol{\theta}_0 = \mathbf{0}$) and the optimal semiparametric test for unspecified location $\phi^{\ddagger(n)}_{t_4}$, both applied to three-year subseries of the 17-dimensional financial return data described in Section 6.

Acknowledgements

We thank Yves Dominicy for sharing the dataset of daily returns with us.

Table A.1

Rejection frequencies (out of $N = 3,000$ replications), under various three-dimensional elliptical ($\boldsymbol{\lambda} = (0,0,0)$) and related skewed densities (increasing $\boldsymbol{\lambda}$ values), of our unspecified-location optimal tests $\phi_f^{\ddagger;(n)}$ (f the trivariate elliptical t distributions with 2.1, 4, and 8 degrees of freedom), Schott's test $\phi^{(n)}_{Schott}$ and Cassart's pseudo-Gaussian test $\phi^{(n)}_{p\mathcal{G}}$ for the null hypothesis of ellipticity with unspecified location. The sample size is $n = 100$, the nominal probability level 5%.

Method\\$\boldsymbol{\lambda}$	$(0,0,0)$	$(1,-2,0)$	$(1,1,1)$	$(2,2,2)$	$(3,3,3)$
			Skew-normal		
$\phi^{\ddagger(n)}_{t_{2.1}}$	0.044	0.110	0.199	0.463	0.558
$\phi^{\ddagger(n)}_{t_4}$	0.045	0.118	0.207	0.468	0.572
$\phi^{\ddagger(n)}_{t_8}$	0.046	0.119	0.211	0.466	0.564
$\phi^{(n)}_{\text{Schott}}$	0.038	0.038	0.038	0.038	0.049
$\phi^{(n)}_{p\mathcal{G}}$	0.043	0.122	0.062	0.083	0.088
			Skew-$t_{2.1}$		
$\phi^{\ddagger(n)}_{t_{2.1}}$	0.023	0.369	0.436	0.571	0.578
$\phi^{\ddagger(n)}_{t_4}$	0.018	0.276	0.328	0.446	0.455
$\phi^{\ddagger(n)}_{t_8}$	0.015	0.202	0.227	0.320	0.328
$\phi^{(n)}_{\text{Schott}}$	0.270	0.293	0.315	0.317	0.324
$\phi^{(n)}_{p\mathcal{G}}$	0.014	0.106	0.107	0.158	0.153
			Skew-$t_{4.1}$		
$\phi^{\ddagger(n)}_{t_{2.1}}$	0.040	0.507	0.637	0.845	0.867
$\phi^{\ddagger(n)}_{t_4}$	0.037	0.446	0.557	0.773	0.802
$\phi^{\ddagger(n)}_{t_8}$	0.032	0.364	0.459	0.659	0.692
$\phi^{(n)}_{\text{Schott}}$	0.047	0.042	0.051	0.055	0.057
$\phi^{(n)}_{p\mathcal{G}}$	0.033	0.224	0.195	0.267	0.274
			Skew-t_8		
$\phi^{\ddagger(n)}_{t_{2.1}}$	0.038	0.367	0.507	0.781	0.839
$\phi^{\ddagger(n)}_{t_4}$	0.040	0.349	0.483	0.753	0.807
$\phi^{\ddagger(n)}_{t_8}$	0.036	0.316	0.436	0.692	0.746
$\phi^{(n)}_{\text{Schott}}$	0.039	0.041	0.034	0.049	0.053
$\phi^{(n)}_{p\mathcal{G}}$	0.043	0.223	0.158	0.209	0.228
			Skew-t_{10}		
$\phi^{\ddagger(n)}_{t_{2.1}}$	0.045	0.310	0.439	0.724	0.799
$\phi^{\ddagger(n)}_{t_4}$	0.047	0.304	0.420	0.706	0.778
$\phi^{\ddagger(n)}_{t_8}$	0.049	0.281	0.393	0.658	0.730
$\phi^{(n)}_{\text{Schott}}$	0.034	0.039	0.036	0.040	0.049
$\phi^{(n)}_{p\mathcal{G}}$	0.054	0.213	0.134	0.184	0.208

TABLE A.2

Rejection frequencies (out of $N = 3,000$ replications), under various three-dimensional elliptical ($\boldsymbol{\lambda} = (0,0,0)$) and related skewed densities (increasing $\boldsymbol{\lambda}$ values), of our unspecified-location optimal tests $\phi_f^{\ddagger;(n)}$ (f the trivariate elliptical t distributions with 2.1, 4, and 8 degrees of freedom), Schott's test $\phi_{Schott}^{(n)}$ and Cassart's pseudo-Gaussian test $\phi_{p\mathcal{G}}^{(n)}$ for the null hypothesis of ellipticity with unspecified location. The sample size is $n = 100$, the nominal probability level 5%.

Method\ $\boldsymbol{\lambda}$	$(0,0,0)$	$(0.15,-2,0)$	$(0.15,0.15,0.15)$	$(0.3,0.3,0.3)$	$(0.45,0.45,0.45)$
			SAS-normal		
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.049	0.175	0.266	0.840	0.993
$\phi_{t_4}^{\ddagger(n)}$	0.050	0.174	0.271	0.844	0.991
$\phi_{t_8}^{\ddagger(n)}$	0.046	0.173	0.263	0.827	0.986
$\phi_{\text{Schott}}^{(n)}$	0.037	0.043	0.032	0.032	0.041
$\phi_{p\mathcal{G}}^{(n)}$	0.054	0.354	0.386	0.936	0.998
			SAS-$t_{4.1}$		
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.039	0.108	0.123	0.400	0.701
$\phi_{t_4}^{\ddagger(n)}$	0.038	0.095	0.104	0.345	0.606
$\phi_{t_8}^{\ddagger(n)}$	0.035	0.083	0.087	0.269	0.489
$\phi_{\text{Schott}}^{(n)}$	0.045	0.040	0.045	0.049	0.057
$\phi_{p\mathcal{G}}^{(n)}$	0.032	0.124	0.113	0.365	0.633
			LSGM		
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.050	0.054	0.073	0.182	0.398
$\phi_{t_4}^{\ddagger(n)}$	0.046	0.051	0.069	0.149	0.335
$\phi_{t_8}^{\ddagger(n)}$	0.042	0.047	0.058	0.123	0.269
$\phi_{\text{Schott}}^{(n)}$	0.461	0.452	0.450	0.450	0.429
$\phi_{p\mathcal{G}}^{(n)}$	0.058	0.057	0.073	0.132	0.254
			Gaussian Mixture		
$\boldsymbol{\mu}_1$	$(0,0,0)$	$(0,0,0)$	$(1,0,0)$	$(1,0,0)$	$(1,0,0)$
$\boldsymbol{\mu}_2$	$(0,0,0)$	$(-1,0,0)$	$(-1,0,0)$	$(-2,0,0)$	$(-3,0,0)$
$\phi_{t_{2.1}}^{\ddagger(n)}$	0.044	0.436	0.905	0.955	0.957
$\phi_{t_4}^{\ddagger(n)}$	0.044	0.404	0.860	0.915	0.911
$\phi_{t_8}^{\ddagger(n)}$	0.044	0.365	0.777	0.832	0.815
$\phi_{\text{Schott}}^{(n)}$	0.081	0.083	0.120	0.236	0.383
$\phi_{p\mathcal{G}}^{(n)}$	0.057	0.118	0.142	0.114	0.093

Table A.3

p-values for testing for elliptical symmetry in 17-dimensional financial return data for rolling windows over three years. We compare the optimal semiparametric test $\phi_{\mathbf{0}}^{(n)}$ for fixed $\boldsymbol{\theta}_0 = \mathbf{0}$ with the optimal semiparametric test for unspecified location $\phi_{t_4}^{\ddagger(n)}$ based on the multivariate t distribution with 4 degrees of freedom.

Start	End	p-value $\phi_{t_4}^{\ddagger(n)}$	p-value $\phi_{\mathbf{0}}^{(n)}$	Number of observations
2000-01-07	2002-12-31	0.107659	0.041650	778
2001-01-01	2003-12-31	0.571561	0.480889	783
2002-01-01	2004-12-31	0.251470	0.527236	784
2003-01-01	2005-12-30	0.028470	0.276275	783
2004-01-01	2006-12-29	0.007923	0.243174	782
2005-01-03	2007-12-31	0.000157	0.286752	781
2006-01-02	2008-12-31	0.000152	0.100240	783
2007-01-01	2009-12-31	0.000146	0.183233	784
2008-01-01	2010-12-31	0.005695	0.872441	784
2009-01-01	2011-12-30	0.000010	0.904906	782
2010-01-01	2012-12-31	0.000103	0.626142	782
2011-01-03	2013-12-31	0.011507	0.109618	782
2012-01-02	2014-12-31	0.035069	0.204622	783
2013-01-01	2015-12-31	0.000004	0.027661	783
2014-01-01	2016-12-30	0.000011	0.380901	783
2015-01-01	2017-09-20	0.006327	0.766111	710