



# Goodness-of-fit tests for multivariate skewed distributions based on the characteristic function

Maicon J. Karling<sup>1</sup> · Marc G. Genton<sup>1</sup> · Simos G. Meintanis<sup>2,3</sup>

Received: 23 February 2023 / Accepted: 24 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

We employ a general Monte Carlo method to test composite hypotheses of goodness-of-fit for several popular multivariate models that can accommodate both asymmetry and heavy tails. Specifically, we consider weighted L2-type tests based on a discrepancy measure involving the distance between empirical characteristic functions and thus avoid the need for employing corresponding population quantities which may be unknown or complicated to work with. The only requirements of our tests are that we should be able to draw samples from the distribution under test and possess a reasonable method of estimation of the unknown distributional parameters. Monte Carlo studies are conducted to investigate the performance of the test criteria in finite samples for several families of skewed distributions. Real-data examples are also included to illustrate our method.

**Keywords** Empirical characteristic function · Goodness-of-fit tests · Heavy tails · Skewed distributions · Skew-normal distribution · Tukey g-and-h distribution

**Mathematics Subject Classification** 62F03 · 62H12 · 62H15

## 1 Introduction

Since the late 1980s, L2-type tests for goodness-of-fit based on the characteristic function (CF) have witnessed increasing popularity. The main reason is that the CF uniquely determines the underlying distribution and that it may be consistently estimated by the empirical CF. For multivariate distributions, there is the extra advantage that multivariate CFs and empirical CFs are well-defined and smooth, unlike the cumulative distribution function and its empirical counterpart, and thus it is easier to work with, even when the population distribution function is known.

Not surprisingly, testing for multivariate normality occupies a prominent place in this setting (see, e.g., Chen and Genton 2022; Ebner et al. 2021; Henze 2002; Henze et al. 2019; Henze and Wagner 1997; Pudeřko 2005) as a wide range of procedures is available, including CF-based tests. Outside the multivariate normal context, however, the range of CF-based goodness-of-fit procedures is limited to only a handful of distributions, most of them belonging to the elliptical class (see, e.g., Fragiadakis and Meintanis 2011; Meintanis et al. 2015; Székely and Rizzo 2013). One of the main reasons for this lack of available procedures is that CFs and empirical CFs, despite being smooth, are often required to be numerically integrated in the L2 setting, a task that may be problematic in higher dimensions, let alone the fact that the analytic form of the population CFs may be altogether unknown for most multivariate distributions under test.

Recently Chen et al. (2022) proposed a procedure that is based on a Monte Carlo approximation of the CF under test, thereby avoiding the use of corresponding population quantities. However, the elliptical families considered by Chen et al. (2022), as important as they may be, render a range of shapes that limit their potential application, given the fact that asymmetry, in addition to excess kurtosis, is typically

---

Simos G. Meintanis: On sabbatical leave from the University of Athens.

---

✉ Maicon J. Karling  
maicon.karling@kaust.edu.sa

<sup>1</sup> Statistics Program, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

<sup>2</sup> Department of Economics, National and Kapodistrian University of Athens, Athens, Greece

<sup>3</sup> Pure and Applied Analytics, North-West University, Potchefstroom, South Africa

expected in real data analysis from Economics, Finance, and most other disciplines.

In this paper, we follow the approach suggested by Chen et al. (2022), but, at the same time, abandon the context of ellipticity adopted therein towards more general shapes. Specifically, we consider goodness-of-fit tests for certain popular families of multivariate skewed distributions. In this connection, an extra element that needs to be addressed in implementing the tests compared to Chen et al. (2022) is that, unlike the parameters-free tests proposed in that paper, in the current setting the presence of shape parameters necessitates an additional re-sampling cycle to replicate the empirical distribution of the test statistic for a given parameter configuration. In doing so we take advantage of the canonical form of the distributional family under test, whenever available.

The remainder of this work unfolds as follows. In Sect. 2, we revisit some of the main ideas of the work by Chen et al. (2022) and the background for our tests shall be provided. Section 3 specifies the actual implementation of the new test procedure using bootstrap re-sampling. In Sect. 4 we introduce and provide a short review of the collection of families that shall be used for the simulations and goodness-of-fit tests and study their respective canonical forms. An extensive Monte Carlo study is presented to illustrate the finite-sample properties of the tests in Sect. 5. The paper concludes with several real-data applications in Sect. 6 and a discussion of the overall results in Sect. 7. An online Supplement contains some extra Monte Carlo results.

## 2 Characteristic function-based tests

Let  $X \in \mathbb{R}^p$  ( $p \geq 1$ ) be a random vector with an absolutely continuous distribution function  $F_X$ . We are interested in the composite goodness-of-fit testing problem represented by the null hypothesis

$$\mathcal{H}_0 : \text{the law of } X \in \mathcal{F}_\vartheta, \quad \text{for some } \vartheta \in \Theta, \tag{2.1}$$

where  $\mathcal{F}_\vartheta = \{F_\vartheta, \vartheta \in \Theta\}$  denotes a specific parametric family of distributions admitting a parameterization in terms of the parameter vector  $\vartheta$ . The corresponding parameter space  $\Theta$  will be taken as an open subset of  $\mathbb{R}^q$  ( $q \geq 1$ ). Given the uniqueness of CFs, we may equivalently state the null hypothesis in (2.1) as

$$\varphi_X(\mathbf{t}) = \varphi_\vartheta(\mathbf{t}), \quad \forall \mathbf{t} \in \mathbb{R}^p, \text{ for some } \vartheta \in \Theta, \tag{2.2}$$

where  $\varphi_X(\mathbf{t}) = \mathbb{E}(e^{i\mathbf{t}^\top X})$  denotes the CF of  $X$  and  $\varphi_\vartheta(\mathbf{t})$  corresponds to the CF of some random vector in the family  $\mathcal{F}_\vartheta$ . Here  $i = \sqrt{-1}$  and  $\top$  means transposition of vectors and matrices.

A CF-based statistic for goodness-of-fit is typically formulated in terms of  $\|\varphi_n - \varphi_{\widehat{\vartheta}_n}\|_w^2$ , where

$$\|f - g\|_w^2 := \int_{\mathbb{R}^p} |f(\mathbf{t}) - g(\mathbf{t})|^2 w(\mathbf{t}) d\mathbf{t} \tag{2.3}$$

is an L2-type weighted distance between the pair of complex-valued functions  $(f, g)$ ,

$$\varphi_n(\mathbf{t}) = \frac{1}{n} \sum_{j=1}^n e^{i\mathbf{t}^\top X_j} \tag{2.4}$$

is the empirical CF computed from a collection  $(X_1, \dots, X_n)$  of independent and identically distributed (i.i.d.) copies of  $X$ , and  $\varphi_{\widehat{\vartheta}_n}$  is the CF corresponding to the null hypothesis  $\mathcal{H}_0$  with the parameter  $\vartheta$  replaced by an estimator  $\widehat{\vartheta}_n := \widehat{\vartheta}_n(X_1, \dots, X_n)$ . The weight function  $w > 0$  will be further specified below.

There exist cases though of distributions, some of which will be considered herein, for which the null CF  $\varphi_\vartheta(\cdot)$  is either completely unknown or too complicated to work with. In such cases we suggest formulating a test statistic analogously but without direct reference to the CF of the distribution under test. Specifically, and in line with Chen et al. (2022), we suggest the test statistic

$$T_{n,m}^{(w)} = \|\varphi_n - \widehat{\varphi}_{0,m}\|_w^2, \tag{2.5}$$

where  $\varphi_n(\cdot)$  is as in (2.4), while

$$\widehat{\varphi}_{0,m}(\mathbf{t}) = \frac{1}{m} \sum_{j=1}^m e^{i\mathbf{t}^\top X_{0,j}} \tag{2.6}$$

is an empirical CF computed from a sample  $(X_{0,1}, \dots, X_{0,m})$  which is drawn from  $F_{\widehat{\vartheta}_n}$ , i.e., from a sample of size  $m$  ( $m \geq n$ ) taken from the distribution under test with parameter estimated by a consistent estimator  $\widehat{\vartheta}_n := \widehat{\vartheta}_n(X_1, \dots, X_n)$ . In other words,  $\widehat{\varphi}_{0,m}$  is a Monte Carlo approximation of the null CF  $\varphi_\vartheta(\cdot)$ . Rejection is thus obtained for large values of  $T_{n,m}^{(w)}$ .

A clear advantage of using the test statistic  $T_{n,m}^{(w)}$  is its computational simplicity. To see this, write  $|\cdot|$  for the modulus of a complex number and, thereafter, by using standard algebra, we obtain

$$\begin{aligned} |\varphi_n(\mathbf{t}) - \widehat{\varphi}_{0,m}(\mathbf{t})|^2 &= \frac{1}{n^2} \sum_{j,k=1}^n \cos \mathbf{t}^\top (X_j - X_k) \\ &\quad + \frac{1}{m^2} \sum_{j,k=1}^m \cos \mathbf{t}^\top (X_{0,j} - X_{0,k}) \\ &\quad - \frac{2}{nm} \sum_{j=1}^n \sum_{k=1}^m \cos \mathbf{t}^\top (X_j - X_{0,k}). \end{aligned}$$

Now suppose that the weight function  $w(\cdot)$ , figuring in (2.3) and (2.5), is chosen as the density of a random vector  $\mathbf{W} \in \mathbb{R}^p$  following a certain spherical distribution. Then it is well known that the CF of  $\mathbf{W}$  simplifies to  $\varphi_{\mathbf{W}}(\mathbf{t}) = \mathbb{E}(\cos \mathbf{t}^\top \mathbf{W})$  and it is eventually given by  $\Psi(\|\mathbf{t}\|^2)$ , where  $\Psi(\cdot)$  is called the “kernel” associated with  $\mathbf{W}$  and  $\|\cdot\|$  stands for the standard Euclidean norm in  $\mathbb{R}^p$  (see Fang et al. 1990). By using the last equation and such a weight function  $w(\cdot)$  in (2.5), we end up with the test statistic

$$T^\Psi(\mathcal{X}_n; \mathcal{X}_{0,m}) = \frac{1}{n^2} \sum_{j,k=1}^n \Psi(\|\mathbf{X}_j - \mathbf{X}_k\|^2) + \frac{1}{m^2} \sum_{j,k=1}^m \Psi(\|\mathbf{X}_{0,j} - \mathbf{X}_{0,k}\|^2) - \frac{2}{nm} \sum_{j=1}^n \sum_{k=1}^m \Psi(\|\mathbf{X}_j - \mathbf{X}_{0,k}\|^2), \quad (2.7)$$

for fixed  $(n, m, \Psi)$ , where  $\mathcal{X}_n = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  denotes the observed data and  $\mathcal{X}_{0,m} = (\mathbf{X}_{0,1}, \dots, \mathbf{X}_{0,m})$  denotes a set of generated data from  $\mathcal{F}_{\hat{\boldsymbol{\theta}}_n}$ , i.e., from the null distribution with estimated parameters. Note that we have made the dependence of the test statistic on the kernel  $\Psi(\cdot)$  explicit, so that, provided that this kernel is simple enough, (2.7) can be readily computed. Some prominent examples of simple kernels at our disposal are:

- the standard normal kernel  $\Psi(\xi) = e^{-\xi/2}$ ;
- the kernel  $\Psi(\xi) = e^{-\xi^{b/2}}$ ,  $b \in (0, 2)$ , derived from the stable distributions (see Nolan 2013);
- and the generalized Laplace kernel  $\Psi(\xi) = (1 + \xi)^{-b}$ ,  $b > 0$  (see Kozubowski et al. 2013).

In the present work, we shall restrict our tests by making use only of the standard normal kernel. For a more in-depth discussion on kernels, we refer to Micchelli et al. (2006).

### 3 Test implementation by re-sampling

When some of the component parameters occurring in  $\boldsymbol{\vartheta}$  can be standardized out, the asymptotic null distribution of the proposed test statistic  $T^\Psi(\mathcal{X}_n; \mathcal{X}_{0,m})$  in (2.7) does not depend on them. Such parameters are typically location and scatter parameters, while others, labeled as shape parameters, such as skewness and kurtosis, cannot usually be standardized out and, therefore, will ultimately affect the asymptotic null distribution of the test statistic (see Meintanis and Swanepoel 2007). In such cases, we can decompose the parameter vector as  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}, \boldsymbol{\lambda})$ , where  $\boldsymbol{\theta}$  denotes the non-shape parameters and  $\boldsymbol{\lambda}$  denotes the part of  $\boldsymbol{\vartheta}$  that contains only shape parameters. In the presence of a canonical form of the distribution

under test (see Sect. 4), the asymptotic null distribution of the test statistic may be simulated by setting  $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ , where  $\boldsymbol{\theta}_0$  is some standard value of  $\boldsymbol{\theta}$ , and  $\boldsymbol{\lambda}$  is set equal to its value  $\tilde{\boldsymbol{\lambda}} := \tilde{\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{\theta})$  in the canonical form.

In the following, we outline the re-sampling procedure used within this work to approximate the test statistic’s asymptotic distribution under the null hypothesis and indicate how the test can be carried out in practice. We consider two cases of null hypotheses, one “composite” with all parameters being estimated, while the other will be labeled “simple”, although in this second case too, some, but not all, parameters are estimated.

#### 3.1 Simple null hypothesis

Here we are interested in the goodness-of-fit testing problem associated with the simple null hypothesis

$$\mathcal{H}_0^s : \text{the law of } \mathbf{X} \in \mathcal{F}_{\boldsymbol{\theta}, \boldsymbol{\lambda}}, \text{ for a fixed } \boldsymbol{\lambda} = \boldsymbol{\lambda}_0, \text{ and for some } \boldsymbol{\theta}; \quad (3.1)$$

and alternative hypothesis

$$\mathcal{H}_1^s : \text{the law of } \mathbf{X} \notin \mathcal{F}_{\boldsymbol{\theta}, \boldsymbol{\lambda}}, \text{ for a fixed } \boldsymbol{\lambda} = \boldsymbol{\lambda}_0, \text{ and any } \boldsymbol{\theta}. \quad (3.2)$$

Although we labeled (3.1) as a simple null hypothesis, it should be pointed out that the parameter  $\boldsymbol{\theta}$  is left unspecified in  $\mathcal{H}_0^s$ , and that our test procedure incorporates an estimation step for this parameter. For this case, the computation of critical points is based on simple Monte Carlo sampling from the distribution figuring in the null hypothesis. The steps of this Monte Carlo run are as follows:

- Step 1: Generate a random sample  $\mathcal{X}_n = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  from  $\mathcal{F}_{\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0}$ .
- Step 2: Compute the estimate  $\hat{\boldsymbol{\theta}}_n$  from  $\mathcal{X}_n$  and, whenever possible, obtain the standardized sample  $\hat{\mathcal{X}}_n = (\hat{\mathbf{X}}_1, \dots, \hat{\mathbf{X}}_n)$ , where  $\hat{\mathbf{X}}_j = \hat{\mathbf{X}}_j(\mathbf{X}_j, \hat{\boldsymbol{\theta}}_n)$ , for  $j \in \{1, \dots, n\}$ , denotes a standardization applied to  $\mathbf{X}_j$  which is induced by  $\hat{\boldsymbol{\theta}}_n$ .
- Step 3: Generate a random sample  $\mathcal{X}_{0,m} = (\mathbf{X}_{0,1}, \dots, \mathbf{X}_{0,m})$  from  $\mathcal{F}_{\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0}$ .
- Step 4: Compute the test statistic  $T := T^\Psi(\hat{\mathcal{X}}_n; \mathcal{X}_{0,m})$ , according to (2.7).
- Step 5: Repeat Steps 1–4 several times, say  $M$ , and obtain the set of test statistics  $\{T_1, \dots, T_M\}$ . Then the critical point, say  $\hat{c}_\delta$ , is defined as the  $(1 - \delta)\%$  quantile of  $(T_m, m = 1, \dots, M)$ .

Having obtained the empirical critical point,  $\hat{c}_\delta$  is used to compute the test’s empirical powers. In this connection,

we generate a random sample  $\mathcal{X}_n = (X_1, \dots, X_n)$  from any distribution belonging to the set of alternatives in the alternative hypothesis  $\mathcal{H}_1^s$ , and perform Steps 1–4 above, thereby computing the test statistic  $T$ . We reject the null hypothesis  $\mathcal{H}_0^s$  if  $T > \hat{c}_\delta$ . We repeat this procedure several times, say  $L$ , and obtain the empirical power rate as  $L^{-1} \sum_{\ell=1}^L \mathbb{1}_{T_\ell > \hat{c}_\delta}$ , where  $T_\ell$  denotes the test statistic corresponding to the  $\ell^{\text{th}}$  sample, for  $\ell \in \{1, \dots, L\}$ .

### 3.2 Composite null hypothesis

Here we are interested in the (fully) composite goodness-of-fit testing problem whereby all distributional parameters are estimated from the observed data. For reasons of explicitness, we state the null hypothesis as

$$\mathcal{H}_0^c : \text{the law of } X \in \mathcal{F}_{\theta, \lambda}, \quad \text{for some } (\theta, \lambda); \tag{3.3}$$

as well as the alternative

$$\mathcal{H}_1^c : \text{the law of } X \notin \mathcal{F}_{\theta, \lambda}, \quad \text{for any } (\theta, \lambda). \tag{3.4}$$

For this case, the re-sampling scheme is as follows. On the basis of  $\mathcal{X}_n$ , compute the estimator  $\hat{\vartheta}_n = (\hat{\theta}_n, \hat{\lambda}_n)$  of  $\vartheta$  and standardize the observations as  $\hat{X}_j = \hat{X}_j(X_j, \hat{\theta}_n)$ ,  $j \in \{1, \dots, n\}$ , where  $\hat{X}_j(X_j, \hat{\theta}_n)$  denotes a standardization applied to  $X_j$  induced by the estimate  $\hat{\theta}_n$ . Next, generate a random sample  $\hat{\mathcal{X}}_{0,m} := (\hat{X}_{0,1}, \dots, \hat{X}_{0,m})$  under the null hypothesis with  $(\theta, \lambda)$  set equal to  $(\theta_0, \hat{\lambda}_n)$ , where  $\hat{\lambda}_n = \hat{\lambda}_n(\hat{\lambda}_n, \hat{\theta}_n)$  is the parameter estimate of  $\lambda$  induced by the standardization. Then the value of the original test statistic is computed according to (2.7) as  $T = T^\Psi(\hat{\mathcal{X}}_n; \hat{\mathcal{X}}_{0,m})$ , where  $\hat{\mathcal{X}}_n := (\hat{X}_1, \dots, \hat{X}_n)$ . In turn, the critical point against which the value of  $T$  will be compared is computed using a *parametric bootstrap* procedure, the steps of which are outlined below:

- Step 1: Generate a random sample  $\mathcal{X}_n^* := (X_1^*, \dots, X_n^*)$  under the null hypothesis with  $(\theta, \lambda)$  set equal to  $(\theta_0, \hat{\lambda}_n)$ .
- Step 2: On the basis of  $\mathcal{X}_n^*$ , compute the estimator  $\hat{\vartheta}_n^* = (\hat{\theta}_n^*, \hat{\lambda}_n^*)$ .
- Step 3: Standardize the components of  $\mathcal{X}_n^*$  as  $\hat{X}_j^* = \hat{X}_j^*(X_j^*, \hat{\theta}_n^*)$ ,  $j \in \{1, \dots, n\}$ .
- Step 4: Generate a random sample  $\hat{\mathcal{X}}_{0,m}^* := (\hat{X}_{0,1}^*, \dots, \hat{X}_{0,m}^*)$  under the null hypothesis with  $(\theta, \lambda)$  set equal to  $(\theta_0, \hat{\lambda}_n^*)$ , where  $\hat{\lambda}_n^* = \hat{\lambda}_n^*(\hat{\theta}_n^*, \hat{\lambda}_n^*)$  is the bootstrap parameter estimate of  $\lambda_n$ .
- Step 5: Compute the value of the bootstrap test statistic by (2.7) as  $T^* = T^\Psi(\hat{\mathcal{X}}_n^*; \hat{\mathcal{X}}_{0,m}^*)$ , where  $\hat{\mathcal{X}}_n^* := (\hat{X}_1^*, \dots, \hat{X}_n^*)$ .

- Step 6: Steps 1–5 are repeated several times, say  $B$ , and thereby we compute the  $(1 - \delta)\%$  quantile  $c_\delta$ , with  $\delta \in (0, 1)$ , of the empirical distribution of  $(T_b^*, b = 1, \dots, B)$  as the size- $\delta$  critical value of the test statistic.
- Step 7: Repeat Steps 1–6 several times, say  $M$ , and thereby obtain pairs of test statistics and corresponding bootstrap critical points  $(T_m, c_{\delta,m})$ ,  $m \in \{1, \dots, M\}$ .
- Step 8: Compute the empirical rejection rate as  $M^{-1} \sum_{m=1}^M \mathbb{1}_{T_m > c_{\delta,m}}$ .

Because the above parametric bootstrap procedure is time-consuming, we adopt the *warp-speed bootstrap* method of Giacomini et al. (2013). Thus, rather than computing a critical value  $c_{\delta,m}$  for each of the  $M$  Monte Carlo samples, we produce a single critical value that is used for all Monte Carlo samples. To do so, we generate only one single bootstrap sample, i.e., with  $B = 1$  on Step 6, for each of the  $M$  Monte Carlo samples and compute the corresponding bootstrap test statistic, say  $T_m^*$ , from this single bootstrap sample. Then the warp-speed critical value, say  $\tilde{c}_\delta$ , is computed from  $(T_m^*, m \in \{1, \dots, M\})$  analogously as in Step 6 above, and the empirical rejection rate is given by  $M^{-1} \sum_{m=1}^M \mathbb{1}_{T_m > \tilde{c}_\delta}$ .

## 4 Families of skewed distributions

In this section, we consider a collection of five families of skewed distributions and exemplify how our method, described in Sects. 2 and 3, may be applied to perform goodness-of-fit tests with them. In the following Sects. 5 and 6, we shall use these five families of distributions, respectively, in simulation studies and applications to real data sets.

### 4.1 Multivariate skew-normal distribution

The multivariate skew-normal (SN) distribution may be conveniently defined by the CF (see Azzalini and Dalla Valle 1996)

$$\varphi_{\vartheta}(t) = 2e^{it^\top \xi - \frac{1}{2}t^\top \Omega t} \Phi\left(i \frac{\alpha^\top \Omega t}{\sqrt{1 + \alpha^\top \Omega \alpha}}\right), \tag{4.1}$$

where  $\Phi(\cdot)$  is the standard normal cumulative distribution function,  $\vartheta := (\xi, \Omega, \alpha)$  is the associated parameters vector, with  $(\xi, \alpha) \in \mathbb{R}^p \times \mathbb{R}^p$  being, respectively, the location and skewness parameters, and where  $\Omega \in \mathbb{R}^{p \times p}$  is a symmetric positive definite matrix. We shall write  $\text{SN}_p(\xi, \Omega, \alpha)$  to denote this distribution, with  $\alpha = \mathbf{0}$  rendering the  $p$ -variate normal distribution with mean  $\xi$  and covariance matrix equal to  $\Omega$ . A consistent estimator of the parameters in  $\vartheta$  is required to apply our test. There exist a variety of methods for estimating them, including maximum-likelihood (MLE) and

moments-based estimation methods (see, e.g., Azzalini and Capitanio 1999; Azzalini et al. 2010; Flecher et al. 2009), as well as packages available for the same purpose (see Azzalini 2022). There also exist a few goodness-of-fit tests in this case, namely, the tests of Balakrishnan et al. (2014), González-Estrada et al. (2022), Jiménez-Gamero and Kim (2015), and Meintanis and Hlávka (2010), which will be discussed further down in the paper (see Sect. 5.4).

It may be shown that if  $X \sim SN_p(\xi, \Omega, \alpha)$ , then there exists a matrix  $H \in \mathbb{R}^{p \times p}$  such that  $H^T(X - \xi) \sim SN_p(\mathbf{0}, I, \alpha^*)$ , with  $\mathbf{0}$  and  $I$  denoting, respectively, the zero vector and identity matrix in the indicated dimension, and  $\alpha^* = (\alpha^*, 0, \dots, 0)^T$ , with  $\alpha^* = (\alpha^T \bar{\Omega} \alpha)^{1/2}$  and  $\Omega = \omega \bar{\Omega} \omega$ , where  $\omega = \text{diag}(\omega_1, \dots, \omega_p)$  is a positive-definite scale matrix (see eq. (5.2) in Azzalini and Capitanio 2014). In the literature (see, e.g., Azzalini and Capitanio 2014; Capitanio 2020), the  $SN_p(\mathbf{0}, I, \alpha^*)$  distribution is also called the canonical form. Moreover, it was proved by Capitanio (2020) that the choice of

$$H = \Omega^{-1/2} Q, \tag{4.2}$$

where  $\Omega^{-1/2}$  is the unique inverse matrix of the positive definite symmetric square root matrix of  $\Omega$ , and  $Q$  is obtained through the spectral decomposition  $Q \Lambda Q^T = \Omega^{-1/2} \Sigma \Omega^{-1/2}$ , with  $\Sigma$  being the covariance matrix of  $X$ , leads to the conclusion that  $H^T(X - \xi)$  follows a canonical skew-normal distribution. In this connection, write  $\hat{\vartheta}_n = (\hat{\xi}_n, \hat{\Omega}_n, \hat{\alpha}_n)$  for an estimator of  $\vartheta$ , and consider the standardized observations  $\hat{X}_j = \hat{H}_n^T(X_j - \hat{\xi}_n)$ ,  $j \in \{1, \dots, n\}$ , where  $\hat{H}_n = \hat{\Omega}_n^{-1/2} \hat{Q}_n$ . Then the test figuring in (2.7) is readily applied by replacing  $X_j$  by  $\hat{X}_j$ , and where the  $X_{0,j}$  are drawn from a SN distribution with parameters  $(\xi, \Omega) = (\mathbf{0}, I)$  and  $\alpha^*$ , the latter being replaced by  $\hat{\alpha}_n^* = (\hat{\alpha}_n^*, 0, \dots, 0)^T$ , where  $\hat{\alpha}_n^* = (\hat{\alpha}_n^T \hat{\Omega}_n \hat{\alpha}_n)^{1/2}$ . For obtaining the estimates of  $\hat{\xi}_n$ ,  $\hat{H}_n$ , and  $\hat{\alpha}_n^*$ , we suggest the use of the `sn` (Azzalini 2022) package within the R (R Core Team 2022) software environment.

### 4.2 Multivariate skew-t distribution

The multivariate skew-t (ST) distribution is related to the multivariate skew-normal distribution through the stochastic equation  $Y = \xi + \sqrt{\eta}X$ , where  $X$  has a multivariate skew-normal distribution,  $X \sim SN_p(\mathbf{0}, \Omega, \alpha)$ , and  $\eta$  has an inverse-Gamma distribution with shape and scale parameters both equal to  $\nu/2$ , i.e.,  $\eta \sim IG(\nu/2, \nu/2)$ . It was shown by Kim and Genton (2011), theorem 7, that the CF of  $Y$  is given by

$$\varphi_{\vartheta}(t) = \exp(it^T \xi) [\psi_{T_p}(\Omega^{1/2}t) + i\tau^+(\rho, \omega t)], \tag{4.3}$$

where

$$\begin{aligned} \rho &= \Omega \alpha / (1 + \alpha^T \Omega \alpha)^{1/2}, \\ \psi_{T_p}(t) &= \frac{\|\sqrt{\nu}t\|^{v/2}}{\Gamma(v/2) 2^{v/2-1}} K_{v/2}(\|\sqrt{\nu}t\|), \quad \text{for } t \in \mathbb{R}^p, \nu > 0, \\ \tau^+(\rho, \omega t) &= \int_0^\infty \exp(-xt^T \Omega t/2) \tau(\sqrt{x} \rho^T \omega t) dH(x), \\ &\text{for } \rho^T \omega t > 0, \end{aligned}$$

with  $\tau^+(\rho, -\omega t) = -\tau^+(\rho, \omega t)$ ,  $\tau(x) = \int_0^x \sqrt{2/\pi} \exp(u^2/2) du$ , for  $x > 0$ , with  $\tau(-x) = -\tau(x)$ ,  $H(x) = \Gamma(v/2, \nu/(2x)) / \Gamma(v/2)$ , for  $x > 0$ , denoting the cumulative distribution function of  $\eta$ , with  $\Gamma(a) = \Gamma(a, 0)$  and where  $\Gamma(a, b) = \int_b^\infty t^{a-1} e^{-t} dt$ , for  $b \geq 0$ , represents the upper incomplete Gamma function, and  $K_\lambda(\cdot)$  is the integral representation of the modified Bessel function of the third kind, defined as  $K_\lambda(w) = \frac{1}{2} \int_0^\infty x^{\lambda-1} \exp\{-\frac{w}{2}(x + \frac{1}{x})\} dx$ , for  $w > 0$  and  $\lambda \in \mathbb{R}$ . Here  $\vartheta = (\xi, \Omega, \alpha, \nu)$  and we write  $Y \sim ST_p(\xi, \Omega, \alpha, \nu)$ .

Since the multivariate skew-t distribution can be expressed as a scale mixture of a skew-normal distribution, proposition 2 in Capitanio (2020) guarantees that, by taking once again the matrix  $H$  as defined in (4.2), any random vector  $Y \sim ST_p(\xi, \Omega, \alpha, \nu)$  can be transformed into the canonical multivariate skew-t distribution  $H^T Y \sim ST_p(\mathbf{0}, I, \alpha_Y^*, \nu)$ . In particular,  $\xi$  and  $\Omega$ , like in the skew-normal case, are nuisance parameters, so that they can be dismissed for hypothesis testing after the standardization is performed. Additionally, for the simulation studies in Sect. 5, it will suffice to implement the tests for different choices of  $p$ ,  $\nu$ , and  $\alpha^*$  (the unique non-null component of  $\alpha_Y^* = (\alpha^*, 0, \dots, 0)^T$ ), substantially reducing the cases that need a proper investigation. Here the `sn` (Azzalini 2022) package in R (R Core Team 2022) can also be used to retrieve the desired multivariate skew-t's parameters estimates as it was also designed for this purpose. The hypothesis test is then carried out similarly to the skew-normal case, as discussed in the preceding subsection.

### 4.3 Multivariate skew-Laplace distribution

The multivariate skew-Laplace (SL) distribution may be conveniently defined by the CF (see Arslan 2010)

$$\varphi_{\vartheta}(t) = \frac{e^{it^T \xi}}{(1 + t^T \Omega t - 2it^T \alpha)^{(p+1)/2}}, \tag{4.4}$$

with  $\vartheta = (\xi, \Omega, \alpha)$ , where  $(\xi, \alpha) \in \mathbb{R}^p \times \mathbb{R}^p$  are, respectively, location and skewness parameters, and  $\Omega$  is a symmetric positive definite matrix. We will use the notation  $SL_p(\xi, \Omega, \alpha)$  for this distribution. Although the multivariate skew-Laplace distribution proposed by Arslan (2010) has very similar properties to the distinct version introduced by

Kotz et al. (2001), Arslan’s alternative has a simpler probability density function, allowing for uncomplicated estimation methods of its parameters in the multivariate setting. Also, Arslan (2010) proposed an efficient EM algorithm that can be used for estimating  $\xi$ ,  $\Omega$ , and  $\alpha$ .

Analogously to the skew-normal and skew-t distributions, provided that  $\alpha \neq \mathbf{0}$ , it may be shown that  $X \sim SL_p(\xi, \Omega, \alpha)$  can be reduced to a canonical form. This novel result brings down the burden of testing for nuisance parameters and reduces the skewness to a singular one-dimensional component.

**Proposition 4.1** (Canonical form - SL distribution) *Let  $X \sim SL_p(\xi, \Omega, \alpha)$  with  $\alpha \neq \mathbf{0}$  and consider the affine non-singular transform*

$$X^* = H^\top(X - \xi), \tag{4.5}$$

with  $H = \Omega^{-1/2}Q$ , where  $\Omega^{-1/2}$  denotes the inverse of the unique positive definite symmetric square root matrix of  $\Omega$ , and  $Q = [v_1 \dots v_p]$  is the orthogonal matrix with  $v_1 = \Omega^{-1/2}\alpha/\|\Omega^{-1/2}\alpha\|$  as its first column vector and the remaining columns  $v_2, \dots, v_p$  belong to the orthogonal complement of  $v_1$ . Then  $X^* \sim SL_p(\mathbf{0}, \mathbf{I}, \alpha_X^*)$  with  $\alpha_X^* = (\alpha^*, 0, \dots, 0)^\top$  and  $\alpha^* = \|\Omega^{-1/2}\alpha\|$ .

**Proof** From proposition 3 in Arslan (2010), if  $A \in \mathbb{R}^{p \times p}$  is any full rank matrix, it follows that  $A(X - \xi) \sim SL_p(\mathbf{0}, A\Omega A^\top, A\alpha)$ . Since  $\alpha = (\alpha_1, \dots, \alpha_p)^\top \neq \mathbf{0}$  and  $\Omega^{-1/2}$  is non-singular, we have  $\Omega^{-1/2}\alpha \neq \mathbf{0}$ . Hence, there is at least one component  $u_k$  in  $u = (u_1, \dots, u_p)^\top = \Omega^{-1/2}\alpha$  that is non-null. Moreover, if  $e_j$  represents the  $j$ th canonical vector in  $\mathbb{R}^p$ , for  $j \in \{1, \dots, p\}$ , then  $\{u\} \cup \{e_j \in \mathbb{R}^p : j \neq k\}$  is a basis of  $\mathbb{R}^p$ . For ease of reading, let us rename the vectors  $u_1 = u, u_{j+1} = e_j$ , for  $1 \leq j < k$ , and  $u_j = e_j$ , for  $k < j \leq p$ . Then we can apply the Gram-Schmidt process to find an orthonormal basis  $\{v_1, \dots, v_p\}$  of  $\mathbb{R}^p$ . For this, take  $\tilde{v}_1 = u_1 = \Omega^{-1/2}\alpha$  and  $\tilde{v}_j = u_j - \sum_{i=1}^{j-1} \frac{\tilde{v}_i^\top u_j}{\tilde{v}_i^\top \tilde{v}_i} \tilde{v}_i$ , for  $2 \leq j \leq p$ . The desired basis is obtained from the normalizations  $v_j = \tilde{v}_j/\|\tilde{v}_j\|$ , for  $j \in \{1, \dots, p\}$ . Therefore, taking  $Q = [v_1 \dots v_p]$  and  $H = \Omega^{-1/2}Q$ , it follows that  $H^\top(X - \xi) \sim SL_p(\mathbf{0}, H^\top\Omega H, H^\top\alpha)$ , with  $H^\top\Omega H = Q^\top\Omega^{-1/2}\Omega\Omega^{-1/2}Q = \mathbf{I}$  and  $H^\top\alpha = Q^\top\Omega^{-1/2}\alpha = (\|\Omega^{-1/2}\alpha\|, 0, \dots, 0)^\top$ .  $\square$

Analogously to the SN and ST cases, the fact that a canonical form is available for the skew-Laplace distribution is useful in the implementation of the goodness-of-fit test due to a reduced number of parameters to be considered. Moreover, the proof of Proposition 4.1 gives at the same time an algorithm to find the canonical form of a multivariate SL distribution. Taking this in consideration, let  $\hat{X}_j = \hat{H}_n^\top(X_j - \hat{\xi}_n)$ ,  $j \in \{1, \dots, n\}$ , be the standardized observations. Then the test figuring in 2.7 is applied by

replacing  $X_j$  by  $\hat{X}_j$ , where  $X_{0,j}$  are drawn from a SL distribution with parameters  $(\xi, \Omega) = (\mathbf{0}, \mathbf{I})$  and  $\alpha$  is set equal to  $(\|\hat{\Omega}_n^{-1/2}\hat{\alpha}_n\|, 0, \dots, 0)^\top$ . To obtain the estimates of  $\hat{\xi}_n, \hat{\Omega}_n$ , and  $\hat{\alpha}_n$ , we suggest the EM algorithm proposed by Arslan (2010). As there is no package available in R (R Core Team 2022) for this purpose, the algorithms involving estimation of the multivariate skew-Laplace distribution in the present work were implemented by a replication of the steps introduced in Section 3 of Arslan (2010).

#### 4.4 Multivariate Tukey g-and-h distribution

The multivariate Tukey g-and-h distribution (GH) was first introduced by Field and Genton (2006) as a generalization of its univariate counterpart presented by Tukey in 1977. It has been gaining popularity due to its flexible marginal distributions, allowing for the fitting of skewed and heavy-tailed data sets from climate and environmental problems (see, e.g., Jeong et al. 2019; Yan and Genton 2019; Yan et al. 2020). Given two parameter vectors  $g = (g_1, \dots, g_p)^\top \in \mathbb{R}^p$  and  $h = (h_1, \dots, h_p)^\top \in \mathbb{R}_+^p$ , the random vector  $Y \in \mathbb{R}^p$  is said to have a standard multivariate Tukey g-and-h distribution if it can be represented as

$$Y = \tau_{g,h}(Z) := (\tau_{g_1,h_1}(Z_1), \dots, \tau_{g_p,h_p}(Z_p))^\top, \tag{4.6}$$

where  $Z = (Z_1, \dots, Z_p)^\top \sim N_p(\mathbf{0}, \mathbf{I})$  has a standard multivariate normal distribution and, for two given  $g \in \mathbb{R}$  and  $h \in \mathbb{R}_+$ , the univariate function  $\tau_{g,h}$  is defined as

$$\tau_{g,h}(z) = \begin{cases} \left(\frac{\exp(gz) - 1}{g}\right) \exp\left(\frac{hz^2}{2}\right), & g \in \mathbb{R} \setminus \{0\}, \\ z \exp\left(\frac{hz^2}{2}\right), & g = 0, \end{cases} \tag{4.7}$$

for any  $z \in \mathbb{R}$ . The general multivariate Tukey g-and-h distribution is then defined as

$$Y = \Omega \tau_{g,h}(Z) + \xi, \tag{4.8}$$

where  $\Omega \in \mathbb{R}^{p \times p}$  and  $\xi \in \mathbb{R}^p$  are, respectively, a positive definite matrix and a location vector. Here  $\vartheta = (\xi, \Omega, g, h)$  and the nuisance parameters are  $\xi$  and  $\Omega$  as they can be standardized out. We shall use the notation  $GH_p(\xi, \Omega, g, h)$ .

In contrast with the SN, ST, and SL distributions, the Tukey g-and-h distribution does not have a known transformation that allows one to represent it in a reducible canonical form. Also, since the inverse of  $\tau_{g,h}(\cdot)$  does not have a closed expression, classical estimation methods, such as the MLE method, rely on numerical approximations. Likewise, the CF, probability density function, and cumulative distribution

function can only be computed numerically. Several different techniques have been proposed for the estimation and fitting of the univariate Tukey  $g$ -and- $h$  distribution (see the review paper by Möstel et al. 2021). However, for the multivariate case, only a few methods are available. For instance, Field and Genton (2006) used multivariate quantiles for data fitting. He and Raghunathan (2012) assumed that  $\Omega$  is a diagonal matrix and, for this reason, they proposed an algorithm that uses quantiles from the univariate Tukey  $g$ -and- $h$ .

In this work, we opted for the MLE method. For this purpose, we need to find the parameters  $\vartheta$  that maximize the log-likelihood function

$$\ell(\vartheta | Y_1, \dots, Y_n) = -n \ln |\Omega| + \sum_{i=1}^n \sum_{j=1}^p \left[ \ln \left\{ \phi(u_i^{(j)}) \right\} - \ln \left\{ \tau'_{g_j, h_j}(u_i^{(j)}) \right\} \right], \quad (4.9)$$

where  $u_i^{(j)} = \tau_{g_j, h_j}^{-1} \left( \{ \Omega^{-1} (Y_i - \xi) \}^\top e_j \right)$ , for  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, p\}$ ,  $e_j$  denotes the  $j$ th canonical vector in  $\mathbb{R}^p$ ,  $\phi(\cdot)$  is the pdf of a standard univariate normal distribution, and  $\tau'_{g_j, h_j}(z)$  denotes the derivative of  $\tau_{g_j, h_j}(z)$  concerning its argument  $z$ . To give an approximated value of the quantities  $u_i^{(j)}$ , we use the `uniroot` function available in the `stats` (R Core Team and contributors worldwide 2022) package in R (R Core Team 2022). For the maximization procedure, we use the `optim` function available in the `MASS` (Ripley 2022) package in R (R Core Team 2022), together with the “Nelder-Mead” method (see Nelder and Mead 1965).

### 4.5 Multivariate $\alpha$ -stable distributions

The  $\alpha$ -stable distributions (AS), similar to the multivariate skew- $t$  and Tukey  $g$ -and- $h$  distributions, are another possible extension of the multivariate normal distribution that comports skewness. However, while the multivariate skew- $t$  family has finite second moments, the  $\alpha$ -stable distributions are regulated by a parameter  $\alpha \in (0, 2]$ , called the tail index, and it has only finite second moments if  $\alpha = 2$ , which reduces itself to the multivariate normal case. Particularly, if  $X$  has a multivariate  $\alpha$ -stable distribution, then  $\mathbb{E}(\|X\|^s) < \infty$ , if  $0 < s < \alpha$ , and  $\mathbb{E}(\|X\|^s) = \infty$ , if  $s \geq \alpha$ . Several parameterizations for the CF have been proposed in the literature. Here, for numerical reasons, we adopt the  $S^0$  parameterization introduced in Abdul-Hamid and Nolan (1998), for which the CF is given by

$$\varphi_\vartheta(t) = \exp \left( - \int_{S_p} \psi_\alpha(t^\top s) \Gamma(ds) + it^\top \xi \right), \quad t \in \mathbb{R}^p, \quad (4.10)$$

where

$$\psi_\alpha(u) = \begin{cases} |u|^\alpha \left( 1 + i \operatorname{sign}(u) \tan \left( \frac{\pi\alpha}{2} \right) (|u|^{1-\alpha} - 1) \right), & \alpha \neq 1, \\ |u| \left( 1 + i \frac{2}{\pi} \operatorname{sign}(u) \ln |u| \right), & \alpha = 1. \end{cases} \quad (4.11)$$

In this case, we denote  $X \sim S_\alpha^0(\Gamma, \xi)$  to indicate that the random vector  $X$  has an  $\alpha$ -stable distribution with finite spectral measure  $\Gamma(\cdot)$ , defined on the unitary sphere  $S_p := \{s \in \mathbb{R}^p : \|s\| = 1\}$ , and shift vector  $\xi \in \mathbb{R}^p$ . We note from (4.10) that the multivariate  $\alpha$ -stable distributions are a semi-parametric family, being completely defined by the triplet  $(\alpha, \Gamma, \xi)$  and belonging to the more general class of infinitely divisible distributions. Furthermore,  $X - \xi \sim S_\alpha^0(\Gamma, \mathbf{0})$ , so that  $\xi$  is a nuisance parameter vector and it coincides with the mean vector when  $1 < \alpha \leq 2$ .

Byczkowski et al. (1993) proposed a discrete approximation for the spectral measure  $\Gamma$ , which is useful for numerical computations and simulations. This approximation is described as follows: consider a finite partition  $P = \{A_1, \dots, A_m\}$  of  $S_p$  and a set of points  $S = \{s_1, \dots, s_m\} \subseteq S_p$ . Then, by setting  $\gamma_i = \Gamma(A_i)$ , the discrete spectral measure associated to  $(P, S, \Gamma)$  is defined as

$$\Gamma^*(\cdot) = \sum_{i=1}^m \Gamma(A_i) \mathbb{1}_{s_i}(\cdot) = \sum_{i=1}^m \gamma_i \mathbb{1}_{s_i}(\cdot), \quad (4.12)$$

for any Borelian in  $S_p$ . Here  $\Gamma^*$  is implicitly defined by concentrating the mass of  $\Gamma(A_i)$  at each point  $s_i$ , the only requirement being that  $\sup_{s \in A_i} |s - s_i|$  is sufficiently small, for each  $i = 1, \dots, m$ . Hence, the discrete spectral measure  $\Gamma^*$  can be used in practice instead of its continuous counterpart  $\Gamma$ . For simulating multivariate  $\alpha$ -stable random vectors, we shall use the result from Modarres and Nolan (1994) which states that, if  $X \sim S_\alpha^0(\Gamma^*, \xi)$ , with  $\Gamma^*(\cdot)$  defined in (4.12), then

$$X = \begin{cases} \sum_{i=1}^m \gamma_i^{1/\alpha} Z_i s_i + \tilde{\xi}, & \alpha \neq 1, \\ \sum_{i=1}^m \gamma_i \left( Z_i + \frac{2}{\pi} \ln(\gamma_i) \right) s_i + \tilde{\xi}, & \alpha = 1, \end{cases} \quad (4.13)$$

where  $\tilde{\xi} = \xi - \tan \left( \frac{\pi\alpha}{2} \right) \sum_{i=1}^m \gamma_i s_i$  and  $Z_1, \dots, Z_m$  are i.i.d. one-dimensional  $\alpha$ -stable random variables with  $Z_i \sim S_\alpha(1, 1, 0)$  (i.e., scale = skewness = 1 and location = 0). As for the estimation part, we use the projection method proposed by Nolan et al. (2001), which relies on the projections of the multivariate samples into a specifically chosen grid of values from the unitary sphere.

As it is well known,  $X \sim S_\alpha^0(\Gamma, \xi)$  is symmetric if and only if  $\Gamma(\cdot)$  is symmetric on  $S_p$ . So far, no tests available in the literature have been designed specifically for the general asymmetric multivariate  $\alpha$ -stable distributions, except for the

one presented by Meintanis et al. (2015) which covers only the symmetric case. Here we propose a test that can be used for both symmetric and asymmetric cases. Although it is not usual in the literature, we use the notation  $X \sim AS_p(\xi, \Gamma, \alpha)$  to indicate that  $X$  has an asymmetric  $\alpha$ -stable distribution. For a more substantial review of these and further technical details concerning multivariate  $\alpha$ -stable distributions, we suggest reading Karling et al. (2023) and Samorodnitsky and Taqqu (2000).

## 5 Simulation studies

In this section, we present the results of simulation studies that were produced using the tests described in Sects. 2 and 3 with the five families of skewed distributions introduced in Sect. 4. For these simulations, we used  $M = L = 1000$  as a standard value in the steps described in Sects. 3.1 and 3.2. Firstly, we start by calculating the empirically estimated sizes of the tests for each family under a  $\delta = 0.05$  designed nominal level. Then we calculate the power of the test in two distinct situations, the simple hypothesis case, and the composite hypothesis case, respectively, within a second and third round of simulations. In the latter, we test the five families of distributions against the family of sinh-arcsinh distributions (see Jones and Pewsey 2009). It is worth pointing out that, as the sample size  $n$  increases, naturally, the tests require more computational time to run. Also, the efficiency of the test is prone to the number of parameters present in each family and the method used for their estimation. Finally, we close this section with a comparison between our test and a few competitors for the skew-normal family that was already available in the literature.

### 5.1 Estimated sizes

The values presented in Table 1 were generated with the test described in Sect. 3 corresponding to the composite null hypothesis case, for  $\delta = 0.05$  and  $M = 1000$  replications. The dimension considered for the samples in the tests is  $p = 2$ . An analogous table for  $p = 3$  can be found in the Supplement. For simplifying the simulations, in each case, we fixed the value of  $m$  to be equal to  $n$ , the sample size, with  $n \in \{100, 250, 500, 750, 1000\}$ . In a later section, and to have a better understanding of the effect of the size  $m$  of the artificial sample, we present simulation results where we fix the sample size  $n$  and let  $m$  vary; see Simulation 12 in Sect. 5.4.

**Simulation 1** We simulated  $n$  observations from an  $SN_p(\mathbf{0}, \mathbf{I}, \alpha)$  distribution, with  $\alpha = (3, 0, \dots, 0)^\top$ . We calculated the empirical sizes of the test and the results are presented in Table 1. We notice from this table that, as the sample size

increases, the empirical sizes of the test stabilize around 0.05, which corresponds to the designed nominal level.

**Simulation 2** Next, we simulated  $n$  observations from an  $ST_p(\mathbf{0}, \mathbf{I}, \alpha, \nu)$  distribution, with  $\alpha = (3, 0, \dots, 0)^\top$  and  $\nu = 5$ . The estimated sizes of the test are presented in Table 1. Comparing this case with the one in Simulation 1, we note that, for small values of  $n$  (100, 250, 500), the estimated sizes are not that close to 0.05 as the ones observed in the SN case, but they start to converge to the designed nominal level as we increase the value of  $n$ , showing consistency. This behavior could be due to the heavier tails presented by the ST family of distributions, requiring a larger sample size to reach the designed nominal level.

**Simulation 3** Here, we simulated  $n$  observations from an  $SL_p(\mathbf{0}, \mathbf{I}, \alpha)$  distribution, with  $\alpha = (3, 0, \dots, 0)^\top$ . The estimated sizes of the test are presented in Table 1. In this case, similar results to the two preceding simulations can be observed; as the sample size increases, the estimated sizes of the test converge to the designed nominal level.

**Simulation 4** Next, we simulated  $n$  observations from a  $GH_p(\mathbf{0}, \mathbf{I}, \mathbf{g}, \mathbf{h})$  distribution, with  $\mathbf{g} = (1, \dots, 1)^\top$  and  $\mathbf{h} = (0.5, \dots, 0.5)^\top$ . The estimated sizes of the test are presented in Table 1. Since all marginal components of  $\mathbf{h}$  are equal to 0.5, we are in a situation when the variance is not finite and heavier tails than the SN, ST, and SL cases are observed. This, in particular, is reflected in the estimated sizes of the test. We observe that for  $n = 1000$ , the rejection rate is equal to 0.071, which is relatively high. This might be explained due to the wide range dispersion of the observed data sets over the tails. Another case, similar to this one, is shown in the next simulation.

**Simulation 5** Finally, we simulated  $n$  observations from an  $AS_2(\mathbf{0}, \Gamma, \alpha)$  distribution, with discrete spectral measure  $\Gamma(\cdot) = (1/3) \sum_{k=1}^3 \mathbb{1}_{s_k}(\cdot)$ , where  $s_k = (\cos(2\pi k/3), \sin(2\pi k/3))^\top$ , for  $k \in \{1, 2, 3\}$ , and stability index  $\alpha = 1.5$ . For the estimation procedure, we used a grid size of  $N = 24$  projections (see Nolan et al. 2001). The estimated sizes of the test are presented in Table 1. Here, like in Simulation 4, the observations originate from a distribution with infinite variance and heavy tails. However, for this case, as we observed an increasing trend in the estimated size of the test when we increased the sample size  $n$  to 1000, we generated two extra rounds of simulations with  $n = 2500$  and  $n = 5000$  to ensure that it was not diverging from the designed nominal level. The rejection rates obtained for these cases were, respectively, 0.066 and 0.070.

**Table 1** Estimated sizes of the tests correspondent to Simulations 1–5

	$n = 100$	$n = 250$	$n = 500$	$n = 750$	$n = 1000$
SN	0.045	0.047	0.049	0.045	0.053
ST	0.063	0.054	0.062	0.052	0.057
SL	0.063	0.045	0.048	0.055	0.046
GH	0.063	0.046	0.057	0.063	0.071
AS	0.051	0.064	0.063	0.076	0.080

### 5.2 Estimated power functions for the simple null hypothesis case

Consider the goodness-of-fit problem with simple null hypotheses  $\mathcal{H}_0^s$  as given in (3.1) and alternative hypotheses  $\mathcal{H}_1^s$  as in (3.2). In the next five simulations (6–10), we calculate the empirical power functions of the test for a few cases of the GH, SL, SN, ST, and AS distributions. For the four first simulations (6–9), we considered the sample sizes of  $n \in \{100, 250, 500, 750, 1000\}$  and  $p \in \{2, 3\}$  for the dimension of the generated observations. The results are summarized and illustrated in Fig. 1. Whereas for Simulation 10, we consider  $n \in \{250, 500, 750, 1000, 2500\}$  and  $p = 2$ , with the results plotted in Fig. 2.

**Simulation 6** We generated  $n$  observations from a  $GH_p(\mathbf{0}, \mathbf{I}, \mathbf{g}, \mathbf{h})$  distribution. For the null hypothesis, we take  $\lambda_0 = (\mathbf{g}_0, \mathbf{h}_0)$ , with  $\mathbf{g}_0 = (2, \dots, 2)^\top$  and  $\mathbf{h}_0 = (1, \dots, 1)^\top$ . The power functions were calculated for  $\lambda = (\mathbf{g}, \mathbf{h})$  with  $\mathbf{h} = (h, \dots, h)^\top$ , for  $h \in \{0.2, 0.4, 0.6, 0.8\}$ , and  $\mathbf{g} = 2\mathbf{h}$ , so that they only depend on the choice of  $h$ . The results are plotted in Fig. 1, items (a) and (b), respectively, for  $p = 2$  and  $p = 3$ . A quick overview of the plotted functions suggests the obvious, as the sample size increases, the power also increases. Moreover, as  $h$  approaches 1, the power functions converge to the size of the test as theoretically expected.

**Simulation 7** Next, we generated  $n$  observations from an  $SL_p(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha})$  distribution. For the null hypothesis, we take  $\lambda_0 = \boldsymbol{\alpha}_0 = (\alpha_0^*, 0, \dots, 0)^\top$  with  $\alpha_0^* = 3$ . Then we calculated the power functions for  $\lambda = \boldsymbol{\alpha} = (\alpha^*, 0, \dots, 0)^\top$  with  $\alpha^* \in \{0, 1, 2, 5, 8, 13\}$  and the results are plotted in Fig. 1, items (c) and (d), respectively, for  $p = 2$  and  $p = 3$ . We notice from the steepness present in the graphs that, for both dimensional cases, the power function is very sensitive to slight changes in  $\alpha^*$ , producing more power as its argument increases or decreases. Additionally, one can notice a small asymmetry on its graphs about  $\alpha^* = 3$ , precisely where the size of the test is located.

**Simulation 8** Here, we generated  $n$  observations from an  $SN_p(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha})$  distribution. For the null hypothesis, we considered  $\lambda_0 = \boldsymbol{\alpha}_0 = (\alpha_0^*, 0, \dots, 0)^\top$  with  $\alpha_0^* = 3$ , and  $\lambda = \boldsymbol{\alpha} = (\alpha^*, 0, \dots, 0)^\top$  with  $\alpha^* \in \{0, 1, 2, 5, 8, 13\}$

for calculating the empirical power functions. The resulting functions are plotted in Fig. 1, items (e) and (f), respectively, for  $p = 2$  and  $p = 3$ . As one can notice from the graphs, the empirically estimated size of the test acts as an inflection point and the power functions are asymmetrically higher when  $\alpha^* < 3$ . This behavior might be due to the asymmetry of the distributions; the closer  $\alpha^*$  is to 0, the closer the distribution becomes to the multivariate normal distribution. The test has not much power when  $\alpha^*$  increases, showing less steepness in that direction.

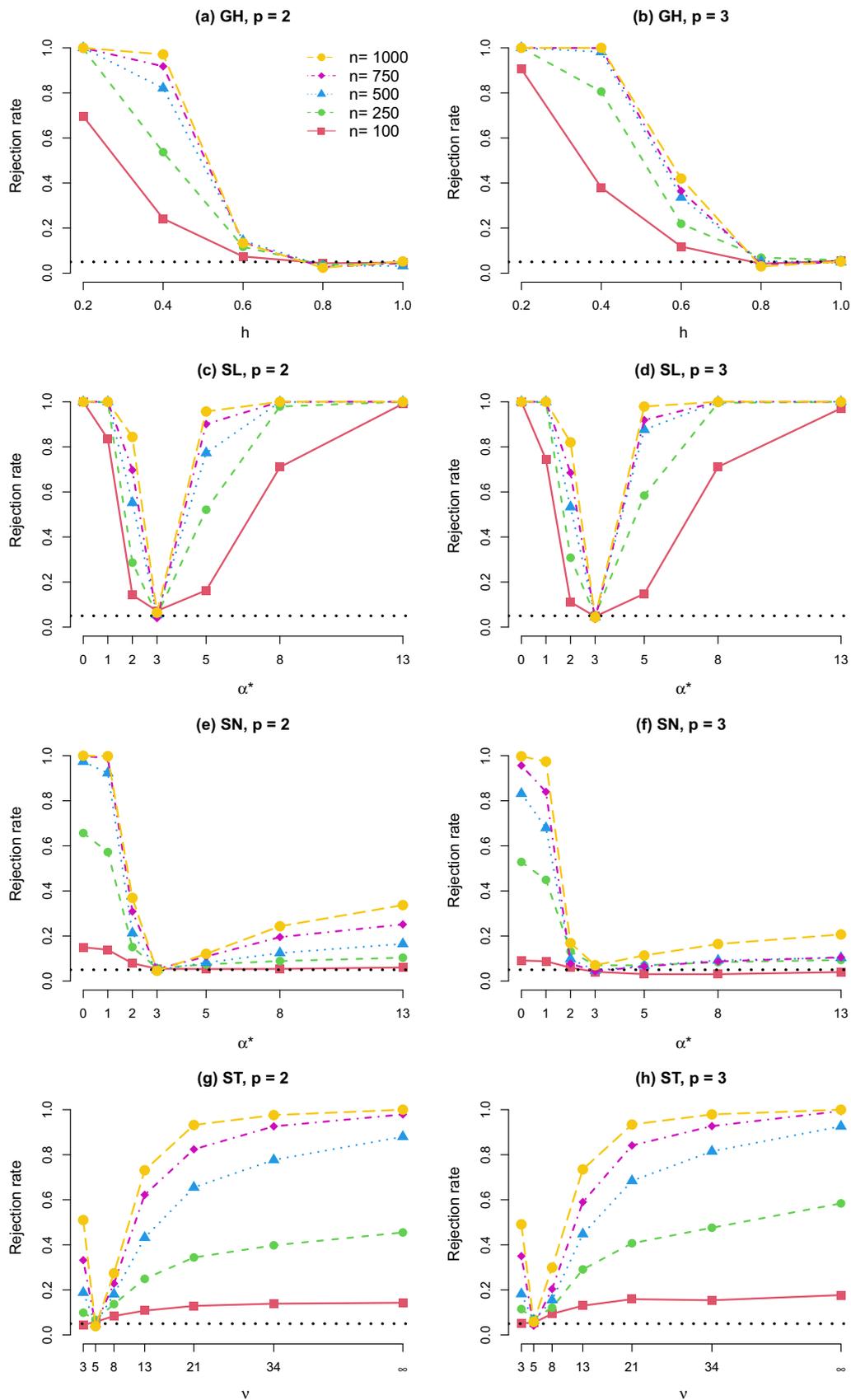
**Simulation 9** Finally, we generated  $n$  observations from an  $ST_p(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha}, \nu)$  distribution with  $\boldsymbol{\alpha} = (3, 0, \dots, 0)^\top$  fixed. We considered  $\nu_0 = 5$  for the null hypothesis and  $\nu \in \{3, 8, 13, 21, 34, \infty\}$  are used for calculating the empirical power functions, plotted in Fig. 1, items (g) and (h), respectively, for  $p = 2$  and  $p = 3$ . By visualizing these figures, one can notice that as  $\nu$  increases the power functions rapidly increase to 1 for large sample sizes and the size of the test is attained at  $\nu = 5$ . Here  $\nu = \infty$  is interpreted as the asymptotic distribution when  $\nu \rightarrow \infty$ .

**Simulation 10** We generated  $n$  observations from an  $AS_2(\mathbf{0}, \Gamma, \boldsymbol{\alpha})$  distribution. For the null hypothesis, we take the spectral measure  $\Gamma_0(\cdot) = (1/3) \sum_{k=1}^3 \mathbb{1}_{s_k}(\cdot)$ , where  $s_k = (\cos(2\pi k/3), \sin(2\pi k/3))^\top$ , for  $k \in \{1, 2, 3\}$ , and  $\alpha_0 = 1.5$ . Then we take two distinct sets for the alternative hypotheses. In the first set, we fix the stability index  $\alpha = \alpha_0$  and calculate the power functions for the alternative spectral measures  $\Gamma_q(\cdot) = (1/q) \sum_{k=1}^q \mathbb{1}_{s_k}(\cdot)$ , with  $s_k = (\cos(2\pi k/q), \sin(2\pi k/q))^\top$ , for  $k \in \{1, \dots, q\}$  and  $q \in \{34, 21, 13, 8, 5\}$ . The resulting estimated power functions are plotted in Fig. 2a. In the second set, we fix the spectral measure as  $\Gamma(\cdot) = (1/3) \sum_{k=1}^3 \mathbb{1}_{s_k}(\cdot)$  and vary the parameter  $\alpha$  instead, for  $\alpha \in \{1.1, 1.3, 1.7, 1.9\}$ . The resulting estimated power functions for this case are plotted in Fig. 2b. In both figures, we also plotted the estimated size of the test with  $\alpha = \alpha_0$  and  $\Gamma(\cdot) \equiv \Gamma_3(\cdot)$ .

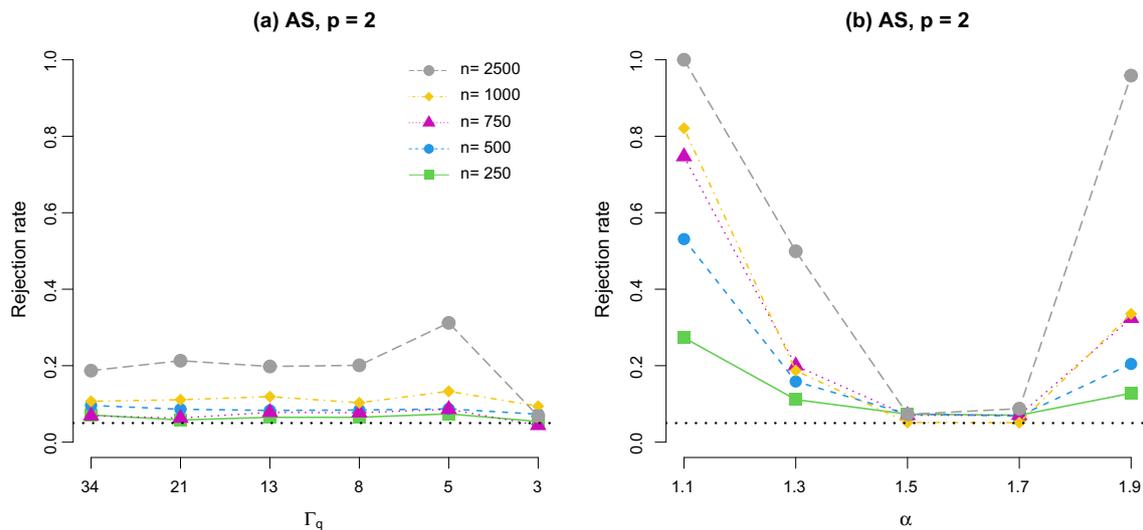
### 5.3 Estimated power functions for the composite null hypothesis case

For the next simulations, consider the goodness-of-fit testing problem with composite null hypotheses  $\mathcal{H}_0^c$  as stated in (3.3) and alternative hypotheses  $\mathcal{H}_1^c$  as in (3.4).

**Simulation 11** We generated  $n \in \{100, 250, 500, 750, 1000\}$  observations from an  $ST_p(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha}, \nu)$  distribution with  $p \in \{2, 3\}$ , a fixed value for  $\boldsymbol{\alpha} = (3, 0, \dots, 0)^\top$ , and  $\nu \in \{1, 2, 3, 5, 8, 13, \infty\}$ . For the null hypothesis, we considered the family of skew-normal distributions. Hence the generated observations belong to the set of alternative hypotheses. Then we calculated the empirical power functions and the results are plotted in Fig. 3. We notice that, as  $\nu \rightarrow \infty$ ,



**Fig. 1** Power functions generated throughout Simulations 6-9: **a, b** Simulation 6  $\rightarrow GH_p(\mathbf{0}, \mathbf{I}, g, h)$ ; **(c),(d)** Simulation 7  $\rightarrow SL_p(\mathbf{0}, \mathbf{I}, \alpha)$ ; **e, f** Simulation 8  $\rightarrow SN_p(\mathbf{0}, \mathbf{I}, \alpha)$ ; **g, h** Simulation 9  $\rightarrow ST_p(\mathbf{0}, \mathbf{I}, \alpha, \nu)$ . The horizontal dotted line corresponds to the 5% significance level



**Fig. 2** Power functions of the test when the null hypothesis is an AS<sub>2</sub>(0, Γ<sub>0</sub>, α<sub>0</sub>) distribution, with α<sub>0</sub> = 1.5 and Γ<sub>0</sub>(·) = (1/3) ∑<sub>k=1</sub><sup>3</sup> 1<sub>s<sub>k</sub></sub>(·), where s<sub>k</sub> = (cos(2πk/3), sin(2πk/3))<sup>⊤</sup>, for k ∈ {1, 2, 3}, corresponding to Simulation 10. On the left-hand side figure, the indexes in the abscissa correspond to the

alternative spectral measures Γ<sub>q</sub>(·) = (1/q) ∑<sub>k=1</sub><sup>q</sup> 1<sub>s<sub>k</sub></sub>(·), where s<sub>k</sub> = (cos(2πk/q), sin(2πk/q))<sup>⊤</sup>, for k ∈ {1, ..., q} and q ∈ {34, 21, 13, 8, 5}. On the right-hand side, we plotted the power of the test when, now, Γ(·) is fixed, but the alternative hypotheses are for α ∈ {1.1, 1.3, 1.7, 1.9}

the ST<sub>p</sub>(0, I, α, ν) distribution converges to the SN<sub>p</sub>(0, I, α) distribution. In particular, this effect is also observed in the power functions, with convergence to the significance level, here set equal to 5%.

**Simulation 12** Now consider the family of multivariate sinh-arcsinh distributions introduced by Jones and Pewsey (2009). Our aim with this simulation is to compute the power functions by considering the sinh-arcsinh as the alternative hypotheses of our test, and as the null hypothesis we shall consider the five families of distribution presented in Sect. 4. To define the multivariate sinh-arcsinh distribution, let us consider the univariate transformation

$$S_{a,b}(z) = \sinh\{b \sinh^{-1}(z) - a\}, \quad a \in \mathbb{R}, b \in \mathbb{R}_+. \quad (5.1)$$

Also, let (e, f) ∈ ℝ<sup>p</sup> × ℝ<sub>+</sub><sup>p</sup>, with e = (e<sub>1</sub>, ..., e<sub>p</sub>)<sup>⊤</sup> and f = (f<sub>1</sub>, ..., f<sub>p</sub>)<sup>⊤</sup>, and Z ~ N<sub>p</sub>(0, I) follow a standard p-variate normal distribution. Then, applying (5.1) component-wise, we say that

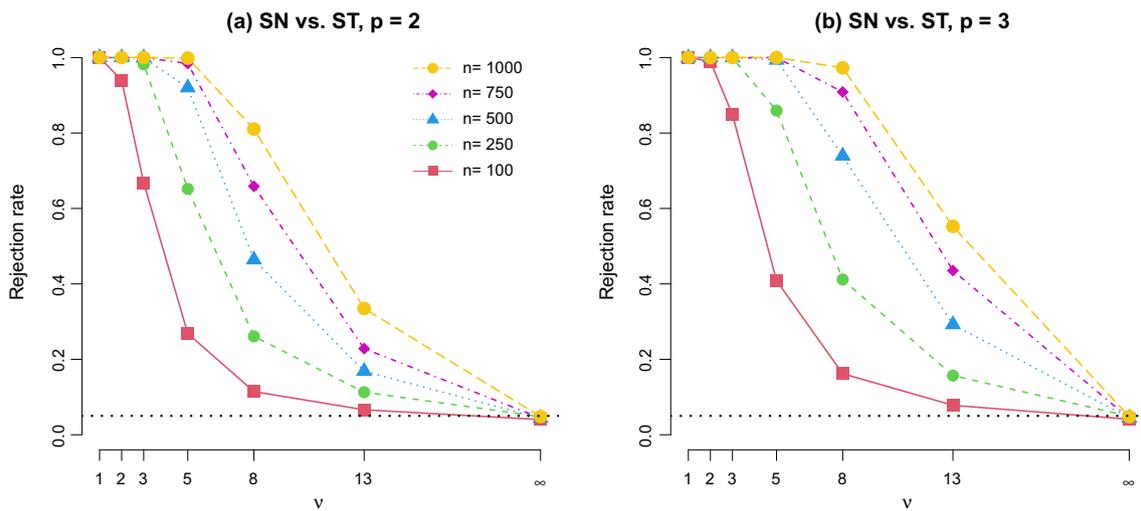
$$Y_{e,f} = S_{-e/f, 1/f}(Z) := \left( S_{-\frac{e_1}{f_1}, \frac{1}{f_1}}(Z_1), \dots, S_{-\frac{e_p}{f_p}, \frac{1}{f_p}}(Z_p) \right)^\top \quad (5.2)$$

has a sinh-arcsinh distribution with parameters (e, f). We simulated n random samples from Y<sub>e,f</sub> when e = (e, ..., e)<sup>⊤</sup>, with e ∈ {0, 0.1, 0.2, 0.3, 0.4, 0.5}, and f = (f, ..., f)<sup>⊤</sup>, with f = (e + 1)<sup>-1</sup>, so that 1/f ∈ {1, 1.1, 1.2, 1.3, 1.4, 1.5} and -e/f ∈ {0, -1/11, -1/6,

-3/13, -2/7, -1/3}. In particular, if e = 0, then Y<sub>e,f</sub> has a standard normal distribution. Moreover, as e increases, its distribution departs rapidly from the standard normal distribution and turns out to be positively skewed on each axis concerning the origin. For the generation of these random samples, we considered two settings. In the first setting, we take n ∈ {50, 100, 250, 500} and m = n in our tests. While in the second set, we fixed the size of the generated samples, with n = 100, and shifted m along the set {100, 250, 500, 1000}, the sample size of the newly generated data needed for the tests, as discussed in Sect. 2. Then we calculated the empirical power functions for each test, with the five different families of distributions considered in Sect. 4 to be the designed null hypothesis. The results found for each of the two different settings are plotted in Figs. 4 and 5. The latter figure shows that, if we have a small sample size data (in this particular simulation, with n = 100), gradually increasing the value of m from 100 to 1000 also slightly increases the power, which is good to know in cases of small sample size; see for instance, the AIS data set from Sect. 6.1.

### 5.4 Comparison over competitor tests

Although the major tests proposed in the literature are restricted to the SN family, we dedicate this subsection to discussing and comparing our test to these alternatives. In a retrospective overview, Meintanis and Hlávka (2010) introduced one of the first goodness-of-fit tests for the family



**Fig. 3** Power of the test when the null hypothesis is assumed to be an  $SN_p(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha})$  distribution, for some  $\boldsymbol{\alpha} = (\alpha^*, 0, \dots, 0)^T$  with  $\alpha^* \in [0, \infty)$ , against the alternative family of  $ST_p(\mathbf{0}, \mathbf{I}, \boldsymbol{\alpha}, \nu)$  dis-

tributions, with  $p \in \{2, 3\}$ ,  $\boldsymbol{\alpha} = (3, 0, \dots, 0)^T$  fixed, and  $\nu \in \{1, 2, 3, 5, 8, 13, \infty\}$ . The horizontal dotted line corresponds to the 5% significance level

of multivariate skew-normal distributions which utilizes the empirical moment-generating function. However, the computational formulas for this test are practically restricted to the two-dimensional case since the case  $p > 2$  requires a solution of a differential equation that is difficult to be numerically evaluated. Later on, Balakrishnan et al. (2014) proposed a test that is based on the skew-normal’s canonical form. Its main advantage relies on the argument that no re-sampling step is needed, saving computational time. However, it is only valid for the skew-normal distribution as the test is based on the fact that the ratios  $X_{i,j}/|X_{i,1}|$ , for  $j \in \{2, \dots, p\}$ , of the observations  $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,p})^T$ , for  $i \in \{1, \dots, n\}$ , are distributed as Cauchy random variables if the  $\mathbf{X}_i$ ’s follow a canonical skew-normal distribution. More recently, González-Estrada et al. (2022) introduced two randomized tests that, similarly to Balakrishnan et al. (2014)’s test, are based on the estimated canonical form of the SN distribution. The first test (W) applies the principle of a generalization of the Shapiro-Wilk test after the sample is transformed into approximately multivariate standard normal observations, whereas the second test (S) relies on a closure property of the sum of univariate independent skew-normal and normal random variables. Jiménez-Gamero and Kim (2015) proposed a pair of re-sampling schemes, one of which is the parametric bootstrap (PB in the paper) which is what we also use in our tests.

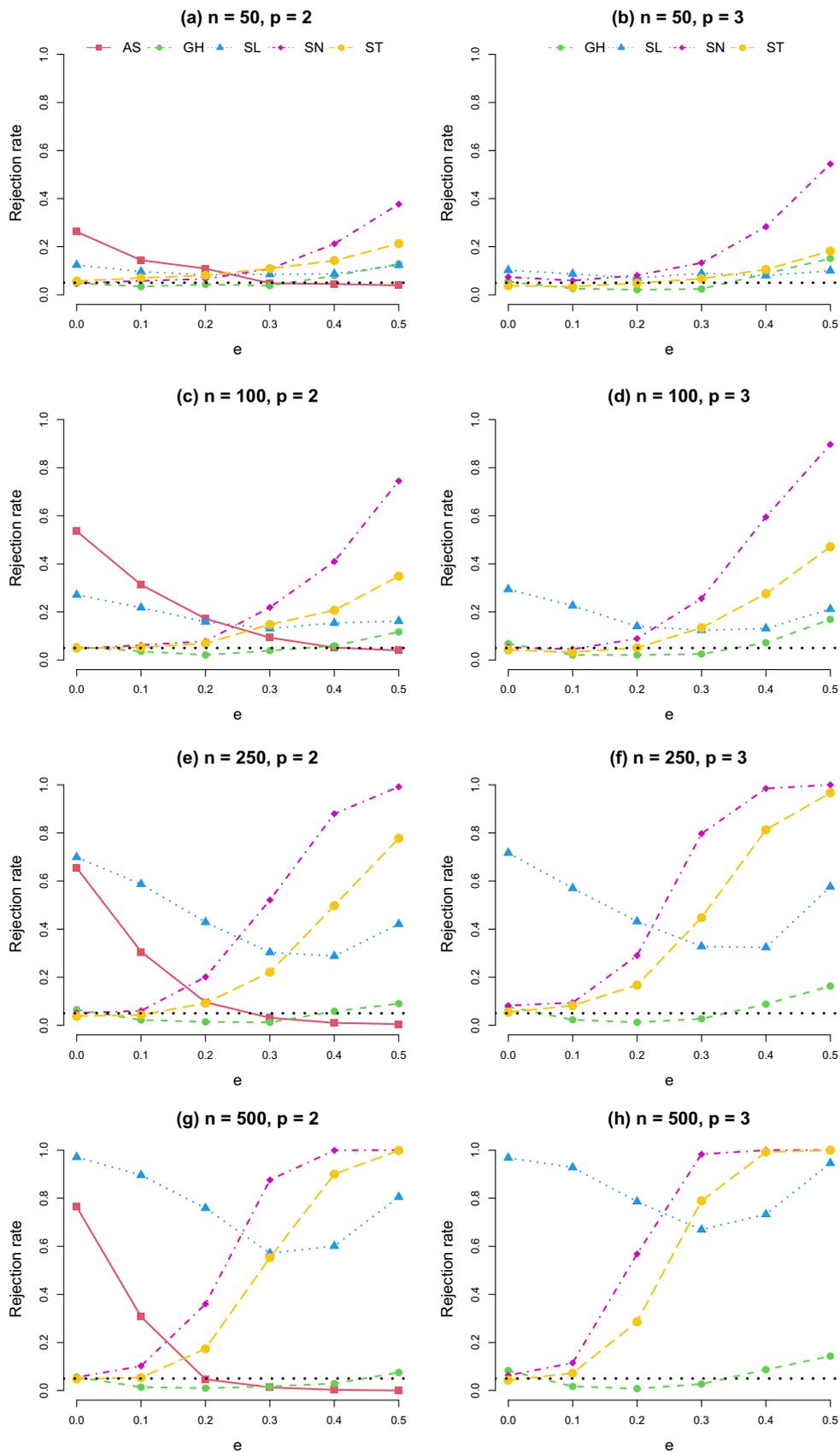
In the next simulation, we replicate one of the original simulation studies from Balakrishnan et al. (2014) and compare the estimated powers of our test with the ones reported by these authors and the ones given in González-Estrada et al. (2022).

**Simulation 13** This study takes into account the 3-variate skew-normal distribution for the composite null hypothesis, and the  $ST_3(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \nu)$  distribution in the alternative set, with  $\boldsymbol{\xi}$ ,  $\boldsymbol{\Omega}$ , and  $\boldsymbol{\alpha}$  defined as follows

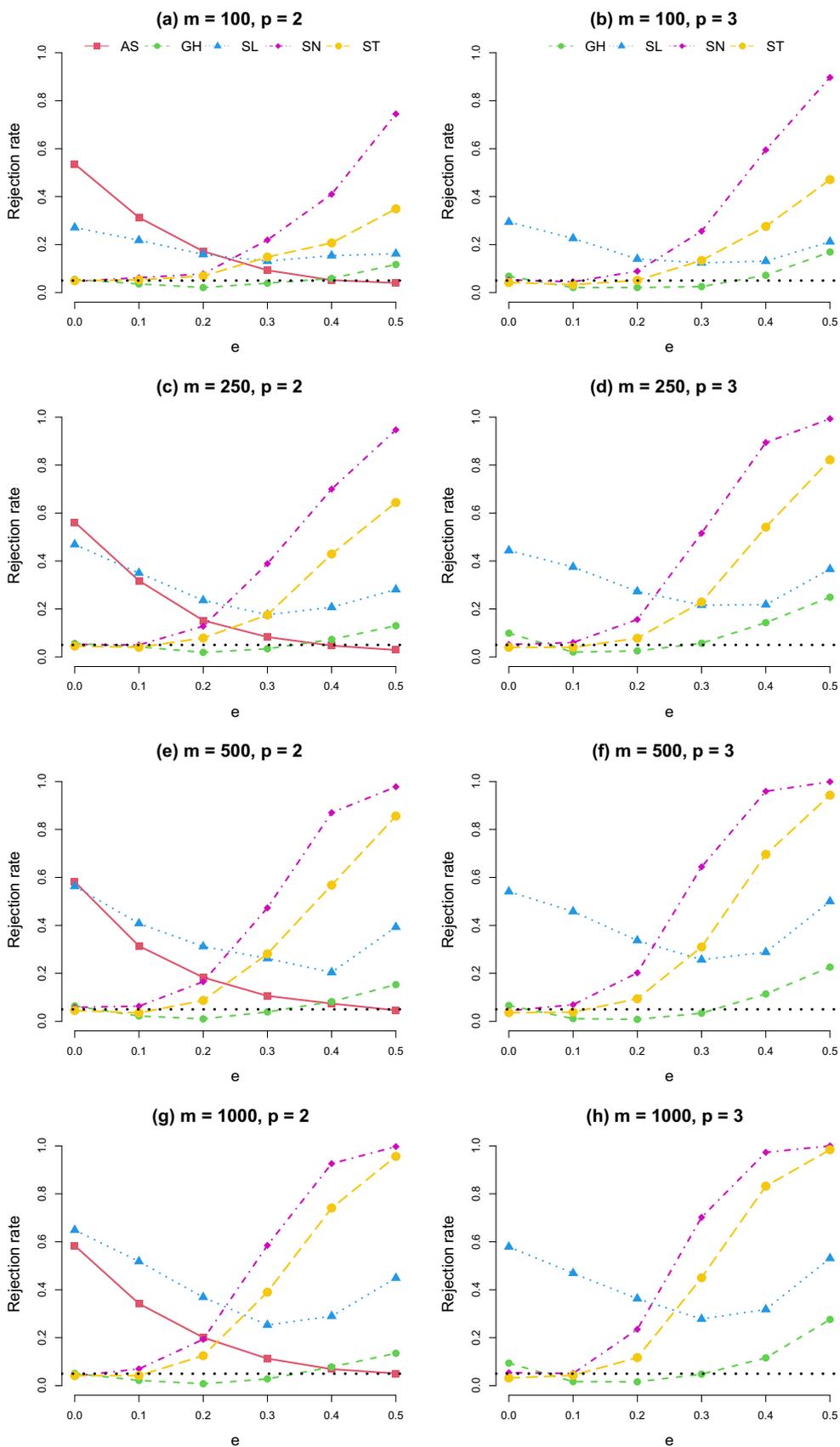
$$\boldsymbol{\xi} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2.5 & 1 \\ 1 & 1 & 5 \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}. \quad (5.3)$$

The degrees of freedom  $\nu$  are taken in the set  $\{1, 2, 3, 5, 10\}$ . We generated ten rounds of replications of our test with different seeds when the sample size is  $n = 100$  and  $m = 1000$ . The rejection rates are presented in Table 2. For comparison reasons, we transcribed the values of the powers reported for this case in Tables 2 and 5 from Balakrishnan et al. (2014) and González-Estrada et al. (2022). As we can see, all ten rounds of tests have shown higher powers than Balakrishnan et al. (2014)’s test. In comparison with González-Estrada et al. (2022)’s test, except for  $\nu = 10$ , our test also has shown higher powers. This shows that, in particular, when the SN distribution is being tested against the ST distribution, our tests are equivalent or even better in terms of powers than the competitor tests presented in the literature.

Before closing this section we wish to point out that while our simulations are extensive we do not wish to overgeneralize our conclusions. Nevertheless due to the wide range of shapes included both under the null hypotheses as well as under alternatives we expect that at least the qualitative aspects of our findings will also hold under other, not drastically different, settings.



**Fig. 4** Power of the tests against the alternative family of sinh-arcsinh distributions correspondent to Simulation 12. Here the sample sizes considered are  $n \in \{50, 100, 250, 500\}$  and  $m = n$



**Fig. 5** Power of the tests against the alternative family of sinh-arcsinh distributions correspondent to Simulation 12 with sample size  $n = 100$  fixed and  $m \in \{100, 250, 500, 1000\}$

**Table 2** Simulated values of power from 1000 replications when the sample size is  $n = 100$  and the data originates from a  $ST_3(\xi, \Omega, \alpha)$  distribution with  $\xi, \Omega,$  and  $\alpha$  defined as in (5.3) and the composite null hypothesis is assumed to be in the skew-normal family

$\nu$ (degrees of freedom)	1	2	3	5	10
Balakrishnan et al. (2014)'s test	0.864	0.475	0.277	0.193	0.165
González-Estrada et al. (2022)'s test	1.000	–	0.979	0.758	0.299
Our test—Round 1	1.000	1.000	0.993	0.838	0.308
Our test—Round 2	1.000	1.000	0.988	0.797	<b>0.247</b>
Our test—Round 3	1.000	1.000	0.992	0.808	<b>0.273</b>
Our test—Round 4	1.000	0.999	0.989	0.801	<b>0.257</b>
Our test—Round 5	1.000	1.000	0.994	0.785	<b>0.281</b>
Our test—Round 6	1.000	1.000	0.990	0.816	0.321
Our test—Round 7	1.000	1.000	0.994	0.791	<b>0.282</b>
Our test—Round 8	1.000	1.000	0.991	0.778	<b>0.265</b>
Our test—Round 9	1.000	1.000	0.995	0.791	<b>0.238</b>
Our test—Round 10	1.000	1.000	0.993	0.806	<b>0.269</b>

Here, values reported in boldface characters indicate the estimated powers that are lower than González-Estrada et al. (2022)'s test

## 6 Data applications

This section considers some examples with real-data samples previously presented in the literature. We apply and discuss the results of the goodness-of-fit tests proposed in the present paper. We considered a 5% confidence level for each test that we performed. The  $p$  values shown in boldface characters throughout Tables 3, 4 and 5 below indicate rejection of the null hypothesis concerning the designated confidence level. Additionally, to obtain more power for each test, we set  $m = \max\{n, 1000\}$  to be in accordance with the results presented in Simulation 12.

### 6.1 AIS data set

The Australian Institute of Sport (AIS) data set is one of the classical examples presented by Azzalini and Capitanio (1999) to illustrate the fitting of a skew-normal distribution. The data consists of biomedical measurements on 100 female and 102 male athletes collected at the Australian Institute of Sport, including body mass index (BMI), body fat percentage (BFP), the sum of skin folds (SSF), and lean body mass (LBM), among others, and it can be retrieved through the `sn` (Azzalini 2022) package in R (R Core Team 2022). These four mentioned indexes were also recently used by Balakrishnan et al. (2014) and González-Estrada et al. (2022) for testing the goodness-of-fit of the skew-normal distribution. Here, in addition to including tests for the skew-normal family, we also include the tests for asymmetric  $\alpha$ -stable (two-dimensional case only), Tukey  $g$ -and- $h$ , skew-Laplace, and skew- $t$  distributions.

We applied our tests on the two-dimensional and four-dimensional data with the athletes segregated by gender,

female and male, and we obtained the estimated  $p$  values shown in Table 3. We observe that the one with the GH distribution was the only test that failed to reject the null hypothesis in the four-dimensional case, for both female and male athletes. As for the SL and ST distribution, the tests suggest the rejection of the null hypotheses only for the data on female athletes. Moreover, the test leads to the conclusion in favor of the SN distribution for the data on female athletes, while for the data on male athletes, the test suggests the rejection of the SN distribution. These two tests are, therefore, in line with the conclusions presented by Balakrishnan et al. (2014) and González-Estrada et al. (2022). For the pairwise two-dimensional case, we observe that most of the tests with the AS distribution suggest rejection of the null hypothesis, with 10 out of 12 pairs of data showing  $p$  values lower than 0.05.

### 6.2 BMI of Australian twin sample biometric data

Nowadays it is clear from a statistical perspective that the BMI's population distribution is not symmetric, usually showing skewness to the right towards a higher ratio of weight to height (see Nuttall 2015). By considering the BMI observations of the AIS data discussed in Sect.6.1, Marchenko and Genton (2010) presented strong evidence that the skewness parameter is different from zero. It has also been pointed out in the literature (see, e.g., Tran et al. 2017; Tsang et al. 2018) that the skew- $t$  distribution is reasonably competitive when describing unimodal BMI data. So, as our second application, we consider the observations of BMI of monozygotic (MZ) twins retrieved from the `twinData` set, available in the `OpenMx` (Boker et al. 2022) package in R (R Core Team 2022). In our analysis, we included individuals

**Table 3** Estimated  $p$  values of the tests for the four-dimensional and pairwise two-dimensional AIS data set

	GH	SL	SN	ST	
<b>Four-dimensional AIS data</b>					
Female athletes	0.715	<b>0.004</b>	0.068	<b>0.022</b>	
Male athletes	0.175	0.503	<b>0.013</b>	0.247	
	AS	GH	SL	SN	ST
<i>Pairwise two-dimensional AIS data</i>					
<b>Female athletes</b>					
BMI & BFP	<b>0.002</b>	0.074	<b>0.003</b>	0.234	0.201
BMI & SSF	<b>0.000</b>	<b>0.000</b>	<b>0.008</b>	<b>0.044</b>	<b>0.010</b>
BMI & LBM	<b>0.000</b>	0.618	0.203	0.102	0.236
BFP & SSF	<b>0.000</b>	0.287	<b>0.001</b>	<b>0.049</b>	<b>0.037</b>
BFP & LBM	<b>0.000</b>	<b>0.000</b>	<b>0.010</b>	0.115	0.087
SSF & LBM	0.271	0.105	<b>0.012</b>	0.162	0.123
<b>Male athletes</b>					
BMI & BFP	<b>0.000</b>	0.258	<b>0.033</b>	<b>0.001</b>	0.648
BMI & SSF	0.257	0.058	0.209	<b>0.003</b>	0.778
BMI & LBM	<b>0.001</b>	0.157	<b>0.038</b>	0.300	0.486
BFP & SSF	<b>0.022</b>	<b>0.018</b>	0.213	<b>0.002</b>	0.204
BFP & LBM	<b>0.000</b>	0.281	0.052	<b>0.032</b>	0.243
SSF & LBM	<b>0.041</b>	<b>0.001</b>	0.118	0.199	0.472

of all ages, separated by gender, with 1171 pairs of females and 532 pairs of males, and we only removed the pairs of twins that showed missing BMI data. There are two reasons why we decided to use this data set rather than the BMI from the AIS data: the first one is because it has more observations and they are more homogeneous (the AIS data includes athletes from mixed categories of sports, which increases the heterogeneity of the data due to distinguished physiological traits); the second reason is that if we singled out the BMI observations from the AIS data it would only be univariate, while the present work is concentrated on multivariate data.

We fitted the two-dimensional vectors of observed BMI to the five distributions introduced in Sect. 4 and applied our goodness-of-fit tests. The estimated  $p$  values are presented in Table 4. For the 5% confidence level, the tests rejected the SN and ST distributions for both female and male MZ twins. Nevertheless, it is interesting to observe that the  $p$  value of the ST test is significantly higher than the one obtained by the SN test, which corroborates the claims found in the literature that the ST distribution is reasonably better. As for the AS and SL distributions, the tests showed ambiguous results for the two genders, rejecting the two distributions in the male case and showing a relatively high  $p$  value for the female case. Lastly, the test with the GH distribution did not show enough evidence to reject the null hypothesis for both data on females and males.

**Table 4** Estimated  $p$  values of the test correspondent to the two-dimensional BMI of Australian monozygotic twin sample data sets

	AS	GH	SL	SN	ST
Female MZ twins	0.458	0.290	0.288	<b>0.000</b>	<b>0.042</b>
Male MZ twins	<b>0.000</b>	0.216	<b>0.000</b>	<b>0.012</b>	<b>0.046</b>

### 6.3 Wind speed data

As a third and final example, we tested the *wind speed* data set presented in Azzalini and Genton (2008), consisting of 278 observations of hourly average wind speed measurements from February 25 to November 30, 2003, recorded at midnight and collected at three meteorological towers: Goodnoe Hills (gh), Kennewick (kw), and Vansycle (vs), located along the Columbia Gorge and the Oregon-Washington border in the US Pacific Northwest. Azzalini and Genton (2008) proposed the fitting of the data by using an i.i.d. skew- $t$  three-dimensional model, claiming that it “brings significant improvements over the normal distribution”. This same data set was also used by Arslan (2010) to illustrate the fitting of the skew-Laplace distribution. The author considered the two-dimensional vectors of wind speed recorded at the towers (gh, kw) and (vs, gh), arguing that the data were satisfactorily fitted to the scatterplots by the skew-Laplace distribution and that it captured the skewness and the apparent heavy tailedness.

We run our goodness-of-fit tests on the tri-dimensional wind speed data set to verify if any of the skewed models introduced in Sect. 4 is inappropriate. To get additional information, we also applied the same tests to the pairwise two-dimensional data sets, now including the  $\alpha$ -stable distribution. The estimated  $p$  values are presented in Table 5. Considering the 5% level of significance, in the tri-dimensional case, only the test for the GH distribution did not show enough evidence for rejecting the null hypothesis, whereas all the other tests, namely, for SL, SN, and ST distributions, presented a  $p$  value lower than 0.05, thus suggesting the rejection of these three distributions. In the pairwise two-dimensional case, most of the tests suggest rejection of the null hypothesis. This conclusion might be because the data from the Goodnoe Hills and Kennewick towers shows signs of bi-modality, as can be seen in their histograms, and perhaps a mixture of distributions is more appropriate to model this data set.

## 7 Conclusion

In this paper, we proposed a goodness-of-fit test for several types of multivariate skewed distributions. On the one hand, the major advantage of the technique addressed in our work

**Table 5** Estimated  $p$  values of the goodness-of-fit tests correspondent to the tri-dimensional and pairwise two-dimensional wind speed data set

GH	SL	SN	ST		
Three-dimensional data					
0.938	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>		
AS	GH	SL	SN	ST	
Pairwise two-dimensional data					
gh & kw	<b>0.006</b>	<b>0.000</b>	<b>0.000</b>	<b>0.002</b>	<b>0.006</b>
gh & vs	<b>0.001</b>	0.067	<b>0.000</b>	<b>0.000</b>	<b>0.004</b>
kw & vs	0.397	<b>0.001</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>

resides in the fact that it is flexible and can be applied to any multivariate parametric family of distributions, provided that a reasonable method of estimation of its parameters is available and that the generation of new replicates is feasible. On the other hand, in terms of computational cost, the implementation is highly demanding since the parametric bootstrap step requires an extra cycle of re-sampling within each Monte Carlo run.

While the need for such nested re-sampling is shared by most goodness-of-fit tests available in the literature, this drawback can be easily circumvented with the use of a parallel algorithm, since the parametric bootstrap does not require any sequential procedures, and with the use of the warp-speed bootstrap method of Giacomini et al. (2013). An important fact to be mentioned is that all tests were run with the help of an Intel Xeon Gold 6230R CPU, of which 100 out of its 104 threads have been intensively used to accelerate even more the completion of the simulations. We demonstrated its effectiveness through five families of multivariate distributions, namely, the multivariate skew-normal, skew-t, asymmetric skew-Laplace, skew  $\alpha$ -stable, and Tukey  $g$ -and- $h$  (for most of which there are no available tests), by utilizing the corresponding canonical forms whenever possible.

As the simulations in Sect. 5.1 show (see Table 1), the estimated sizes of the test are reasonable and consistent for all five families. Similarly, the simulations presented in Sects. 5.2 and 5.3 show that the tests have enough power to detect and reject alternative hypotheses. Compared to the alternative options of tests introduced in the literature, as presented in Sect. 5.4, for the particular case when testing under the composite null hypothesis of an SN distribution, our test has also shown to be better in terms of power when testing against the alternative ST distribution.

The effectiveness of our tests has also been illustrated with real data examples in Sect. 6, showing its applicability and usefulness when applied to biological and natural events observed, respectively, in our daily lives and our environment. In closing, we wish to remind the reader that our test

allows a certain flexibility concerning the actual kernel  $\Psi$  used; see the last paragraph of Sect. 2. In this connection, it would be interesting to investigate the effect that this choice has on the finite-sample properties of our test. More work is needed in this direction.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11222-023-10260-0>.

**Acknowledgements** The authors thank the review team for the comments that improved the manuscript. This research was supported by the King Abdullah University of Science and Technology (KAUST). Simos Meintanis wishes to sincerely thank Marc Genton and the staff of KAUST for the invitation and the hospitality rendered during his visit to KAUST.

**Author contributions** The three authors worked on this version of the manuscript equally.

**Declarations**

**Competing Interests** The authors have no financial or proprietary interests in any material discussed in this article.

**References**

Abdul-Hamid, H., Nolan, J.P.: Multivariate stable densities as functions of one dimensional projections. *J. Multivar. Anal.* **67**(1), 80–89 (1998)

Arslan, O.: An alternative multivariate skew Laplace distribution: properties and estimation. *Stat. Pap.* **51**(4), 865–887 (2010)

Azzalini, A.: sn: the skew-normal and related distributions such as the skew-t and the SUN (2022). R package version 2.1.0. <https://cran.r-project.org/web/packages/sn/>

Azzalini, A., Capitanio, A.: Statistical applications of the multivariate skew normal distribution. *J. Roy. Stat. Soc. B* **61**(3), 579–602 (1999)

Azzalini, A., Capitanio, A.: *The Skew-Normal and Related Families*. Cambridge University Press, New York (2014)

Azzalini, A., Dalla Valle, A.: The multivariate skew-normal distribution. *Biometrika* **83**(4), 715–726 (1996)

Azzalini, A., Genton, M.G.: Robust likelihood methods based on the skew-t and related distributions. *Int. Stat. Rev.* **76**(1), 106–129 (2008)

Azzalini, A., Genton, M.G., Scarpa, B.: Invariance-based estimating equations for skew-symmetric distributions. *Metron* **68**(3), 275–298 (2010)

Balakrishnan, N., Capitanio, A., Scarpa, B.: A test for multivariate skew-normality based on its canonical form. *J. Multivar. Anal.* **128**, 19–32 (2014)

Boker, S.M., Neale, M.C., Maes, H.H., Spiegel, M., Brick, T.R., Estabrook, R., Bates, T.C., Gore, R.J., Hunter, M.D., Pritikin, J.N., Zahery, M., Kirkpatrick, R.M.: *OpenMx: extended structural equation modelling* (2022). R package version 2.20.7. <https://cran.r-project.org/package=OpenMx>

Byczkowski, T., Nolan, J.P., Rajput, B.: Approximation of multidimensional stable densities. *J. Multivar. Anal.* **46**(1), 13–31 (1993)

Capitanio, A.: On the canonical form of scale mixtures of skew-normal distributions. *Statistica (Bologna)* **80**(2), 145–160 (2020)

- Chen, W., Genton, M.G.: Are you all normal? It depends. *Int. Stat. Rev.* **5**, 4 (2022). <https://doi.org/10.1111/insr.12512>
- Chen, F., Jiménez-Gamero, M.D., Meintanis, S., Zhu, L.: A general Monte Carlo method for multivariate goodness-of-fit testing applied to elliptical families. *Comput. Stat. Data Anal.* **175**, 107548 (2022)
- Ebner, B., Henze, N., Strieder, D.: Testing normality in any dimension by Fourier methods in a multivariate Stein equation. *Can. J. Stat.* **50**(3), 992–1033 (2021)
- Fang, K.T., Kotz, S., Ng, K.W.: *Symmetric Multivariate and Related Distributions*. Chapman & Hall/CRC, Boca Raton (1990)
- Field, C., Genton, M.G.: The multivariate g-and-h distribution. *Technometrics* **48**(1), 104–111 (2006)
- Flecher, C., Naveau, P., Allard, D.: Estimating the closed skew-normal distribution parameters using weighted moments. *Stat. Probab. Lett.* **79**(19), 1977–1984 (2009)
- Fragiadakis, K., Meintanis, S.G.: Goodness-of-fit tests for multivariate Laplace distributions. *Math. Comput. Model.* **53**(5–6), 769–779 (2011)
- Giacomini, R., Politis, D.N., White, H.: A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Economet. Theor.* **29**(3), 567–589 (2013)
- González-Estrada, E., Villaseñor, J.A., Acosta-Pech, R.: Shapiro-Wilk test for multivariate skew-normality. *Comput. Stat.* **37**(4), 1985–2001 (2022)
- He, Y., Raghunathan, T.E.: Multiple imputation using multivariate gh transformations. *J. Appl. Stat.* **39**(10), 2177–2198 (2012)
- Henze, N.: Invariant tests for multivariate normality: a critical review. *Stat. Pap.* **43**(4), 467–506 (2002)
- Henze, N., Wagner, T.: A new approach to the BHEP tests for multivariate normality. *J. Multivar. Anal.* **62**(1), 1–23 (1997)
- Henze, N., Jiménez-Gamero, M.D., Meintanis, S.G.: Characterizations of multinormality and corresponding tests of fit, including for GARCH models. *Econom. Theor.* **35**(3), 510–546 (2019)
- Jeong, J., Yan, Y., Castruccio, S., Genton, M.G.: A stochastic generator of global monthly wind energy with Tukey g-and-h autoregressive processes. *Stat. Sin.* **29**(3), 1105–1126 (2019)
- Jiménez-Gamero, M.D., Kim, H.: Fast goodness-of-fit tests based on the characteristic function. *Comput. Stat. Data Anal.* **89**, 172–191 (2015)
- Jones, M.C., Pewsey, A.: Sinh-arcsinh distributions. *Biometrika* **96**(4), 761–780 (2009)
- Karling, M.J., Lopes, S.R.C., de Souza, R.M.: Multivariate  $\alpha$ -stable distributions: VAR(1) processes, measures of dependence and their estimations. *J. Multivar. Anal.* **195**, 105153 (2023)
- Kim, H., Genton, M.G.: Characteristic functions of scale mixtures of multivariate skew-normal distributions. *J. Multivar. Anal.* **102**(7), 1105–1117 (2011)
- Kotz, S., Kozubowski, T.J., Podgórski, K.: *The Laplace Distribution and Generalizations*. Birkhäuser, Boston (2001)
- Kozubowski, T.J., Podgórski, K., Rychlik, I.: Multivariate generalized Laplace distribution and related random fields. *J. Multivar. Anal.* **113**, 59–72 (2013)
- Marchenko, Y.V., Genton, M.G.: A suite of commands for fitting the skew-normal and skew-t models. *Stand. Genom. Sci.* **10**(4), 507–539 (2010)
- Meintanis, S.G., Hlávka, Z.: Goodness-of-fit tests for bivariate and multivariate skew-normal distributions. *Scand. J. Stat.* **37**(4), 701–714 (2010)
- Meintanis, S., Swanepoel, J.: Bootstrap goodness-of-fit tests with estimated parameters based on empirical transforms. *Stat. Probab. Lett.* **77**(10), 1004–1013 (2007)
- Meintanis, S.G., Ngatchou-Wandji, J., Taufer, E.: Goodness-of-fit tests for multivariate stable distributions based on the empirical characteristic function. *J. Multivar. Anal.* **140**, 171–192 (2015)
- Micchelli, C.A., Xu, Y., Zhang, H.: Universal kernels. *J. Mach. Learn. Res.* **7**(95), 2651–2667 (2006)
- Modarres, R., Nolan, J.P.: A method for simulating stable random vectors. *Comput. Stat.* **9**(1), 11–19 (1994)
- Möstel, L., Fischer, M., Pfälzner, F., Pfeuffer, M.: Parameter estimation of Tukey-type distributions: a comparative analysis. *Commun. Stat. Simul. Comput.* **50**(4), 957–992 (2021)
- Nelder, J.A., Mead, R.: A simplex method for function minimization. *Comput. J.* **7**(4), 308–313 (1965)
- Nolan, J.P.: Multivariate elliptically contoured stable distributions: theory and estimation. *Comput. Stat.* **28**(5), 2067–2089 (2013)
- Nolan, J.P., Panorska, A.K., McCulloch, J.H.: Estimation of stable spectral measures. *Math. Comput. Model.* **34**(9–11), 1113–1122 (2001)
- Nuttall, F.Q.: Body mass index: obesity, BMI, and health. *Nutr. Today* **50**(3), 117–128 (2015)
- Pudelko, J.: On a new affine invariant and consistent test for multivariate normality. *Probab. Math. Stat.* **25**(1), 43–54 (2005)
- R Core Team and contributors worldwide. *stats: the R stats package* (2022). R package version 4.3.0. <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html>
- R Core Team. *R: a language and environment for statistical computing* (2022). R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Ripley, B.: *MASS: support functions and datasets for venables and Ripley's MASS* (2022). R package version 7.3-58.1. <https://cran.r-project.org/package=MASS>
- Samorodnitsky, G., Taqqu, M.S.: *Stable Non-Gaussian Random Processes: Stochastic models with infinite variance*. Chapman & Hall/CRC, Boca Raton (2000)
- Székel, G.J., Rizzo, M.L.: Energy statistics: a class of statistics based on distances. *J. Stat. Plan. Inference* **143**(8), 1249–1272 (2013)
- Tran, T., Wiskow, C., Aziz, M.A.: Skewed and flexible skewed distributions: a modern look at the distribution of BMI. *Am. J. Undergrad. Res.* **14**(2), 45–63 (2017)
- Tsang, S., Duncan, G.E., Dinescu, D., Turkheimer, E.: Differential models of twin correlations in skew for body-mass index (BMI). *PLoS ONE* **13**(3), e0194968 (2018)
- Yan, Y., Genton, M.G.: Non-Gaussian autoregressive processes with Tukey g-and-h transformations. *Environmetrics* **30**(2), e2503 (2019)
- Yan, Y., Jeong, J., Genton, M.G.: Multivariate transformed Gaussian processes. *Jpn. J. Stat. Data Sci.* **3**(1), 129–152 (2020)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.