Annals of the Institute of Statistical Mathematics (2020) 72:741–770
https://doi.org/10.1007/s10463-019-00707-5

More good news on the HKM test for multivariate reflected symmetry about an unknown centre

Norbert Henze[1] · Celeste Mayer[1]

Received: 17 July 2018 / Revised: 2 January 2019 / Published online: 8 February 2019

Abstract

We revisit the problem of testing for multivariate reflected symmetry about an unspecified point. Although this testing problem is invariant with respect to full-rank affine transformations, among the few hitherto proposed tests only a class of tests studied in Henze et al. (J Multivar Anal 87:275–297, 2003) that depends on a positive parameter a respects this property. We identify a measure of deviation Δ_a (say) from symmetry associated with the test statistic $T_{n,a}$ (say), and we obtain the limit normal distribution of $T_{n,a}$ as $n \to \infty$ under a fixed alternative to symmetry. Since a consistent estimator of the variance of this limit normal distribution is available, we obtain an asymptotic confidence interval for Δ_a. The test, when applied to a classical data set, strongly rejects the hypothesis of reflected symmetry, although other tests even do not object against the much stronger hypothesis of elliptical symmetry.

Keywords Test for reflected symmetry · Fixed alternatives · Affine invariance · Weighted L^2-statistic · Elliptical symmetry

1 Introduction

Testing for symmetry of a univariate distribution about a specified or unspecified point has been a topic of intensive research, see e.g., Section 3 of Quessy (2016). In the multivariate case, this problem is more complex, since different notions of symmetry are available. Among these are, in increasing order of specialization, *reflected* (*diagonal*) *symmetry*, *elliptical symmetry*, and *spherical symmetry*, see, e.g., Meintanis and Ngatchou-Wandji (2012) or Serfling (2006) for an account on the importance of the assumption of symmetry and a survey on these concepts and corresponding goodness-of-fit tests.

✉ Norbert Henze
henze@kit.edu

1 Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Englerstr. 2, 76131 Karlsruhe, Germany

In this paper, we consider testing for reflected symmetry. To be specific, let $X, X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) d-dimensional random (column) vectors, defined on some common probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and assume $d \geq 1$. Thus, the univariate case is deliberately not excluded in what follows. Writing $\stackrel{\mathcal{D}}{=}$ for equality in distribution, the problem is to test the hypothesis

$$H_0 : X - \mu \stackrel{\mathcal{D}}{=} \mu - X \text{ for some (unknown) } \mu \in \mathbb{R}^d, \qquad (1)$$

of *reflected* (*diagonal*) *symmetry* about an unspecified point, against general alternatives.

The technically less demanding problem of testing for reflected symmetry about a *specified* point has been considered in Aki (1993) and, in the special case $d = 2$, in Dyckerhoff et al. (2015) and Einmahl and Gan (2016). For distributions concentrated on the unit circle, the hypothesis "$X \stackrel{\mathcal{D}}{=} -X$" is called *circular reflective symmetry*, see (Ley and Verdebout 2014) and the references therein. Symmetry of a bivariate distribution about a given line is studied in Madhava Rao and Raghunath (2012).

Notice that if a test of H_0 rejects the hypothesis of reflected symmetry, it is forced to also reject the stronger hypotheses of elliptical or spherical symmetry. Thus, any test of H_0 is in this sense a "necessary test" for elliptical or spherical symmetry, and even for multivariate normality.

There is a further basic issue inherent in the testing problem (1). Suppose $X - \mu \stackrel{\mathcal{D}}{=} \mu - X$, and let A be a regular $(d \times d)$-matrix and $b \in \mathbb{R}^d$. Then,

$$AX + b - (A\mu + b) \stackrel{\mathcal{D}}{=} A\mu + b - (AX + b).$$

This means that the problem of testing for reflected symmetry about an unspecified point is invariant with respect to full-rank affine transformations of X. As a consequence, any genuine test of H_0 based on $X_1, \ldots, X_n$ should respect this property. Hence, if $T_n = T_n(X_1, \ldots, X_n)$ is a test statistic based on $X_1, \ldots, X_n$, we should have affine invariance of T_n, i.e.,

$$T_n(AX_1 + b, \ldots, AX_n + b) = T_n(X_1, \ldots, X_n)$$

for each nonsingular $A \in \mathbb{R}^{d \times d}$, and each $b \in \mathbb{R}^d$. Among the few attempts to tackle problem (1) of testing for reflected symmetry with unknown centre (see Székely and Sen 2002; Heathcote et al. 1995; Henze et al. 2003; Neuhaus and Zhu 1998; Ngatchou-Wandji 2009 and Section 2.1 of Meintanis and Ngatchou-Wandji (2012)), only the test of Henze et al. (2003)— henceforth termed the HKM test—is affine invariant. It is the purpose of this paper to revisit this test, which has the desirable properties of being affine invariant, easy to use, consistent against general alternatives, and able to detect alternatives that approach the hypothesis at the rate $n^{-1/2}$. We sum up these (and more) properties in Sect. 2. In Sect. 3, we consider a fixed alternative distribution to H_0 and identify a measure of deviation Δ_a (say) from symmetry associated with the test statistic of Henze et al. (2003). Moreover, we prove that the test statistic has a limit normal distribution under a fixed alternative distribution to H_0. In Sect. 4, we present a consistent estimator of the variance of this limit distribution, which yields an asymptotic confidence interval for Δ_a. Section 5 presents examples, whereas Sect. 6 applies the test to a data set from a health survey of paint sprayers in a car assembly plant. Section 7 contains some concluding remarks. For the sake of readability, all proofs are deferred to Sect. 8.

2 The HKM test

The test of Henze et al. (2003) shares a similar spirit with the BHEP test for multivariate normality, see (Henze and Wagner 1997). It rejects H_0 for large values of the test statistic

$$T_{n,a} = \int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \sin\left(t^\top Y_{n,j}\right) \right)^2 \exp\left(-a\|t\|^2\right) \mathrm{d}t,$$

where $a > 0$ is some fixed parameter. Here, $\top$ denotes transposition of vectors and matrices, $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$,

$$Y_{n,j} = S_n^{-1/2}(X_j - \overline{X}_n), \quad j = 1, \ldots, n, \tag{2}$$

are the *scaled residuals* of $X_1, \ldots, X_n$, and $\overline{X}_n = n^{-1}\sum_{j=1}^n X_j$, $S_n = n^{-1}\sum_{j=1}^n (X_j - \overline{X}_n)(X_j - \overline{X}_n)^\top$ denote the sample mean and the sample covariance matrix of $X_1, \ldots, X_n$, respectively. The matrix $S_n^{-1/2}$ is the unique symmetric square root of S_n^{-1}. To ensure the almost sure invertibility of S_n, we make the basic tacit assumptions that the distribution of X (henceforth abbreviated by $\mathbb{P}^X$) is absolutely continuous with respect to the Lebesgue measure, and that $n \geq d + 1$, see (Eaton and Perlman 1973). In addition, we will adopt the (minimal) moment assumption $\mathbb{E}\|X\|^2 < \infty$. This assumption guarantees that the covariance matrix Σ (say) of X exists, and that S_n converges almost surely to Σ as $n \to \infty$.

An alternative representation of $T_{n,a}$ is

$$T_{n,a} = \frac{\pi^{d/2}}{2na^{d/2}} \sum_{i,j=1}^{n} \left[\exp\left(-\frac{1}{4a}\|Y_{n,i} - Y_{n,j}\|^2\right) - \exp\left(-\frac{1}{4a}\|Y_{n,i} + Y_{n,j}\|^2\right) \right]$$

(see display (1.4) of Henze et al. (2003)), which is amenable to computational purposes. Notice that $T_{n,a}$ is a function of $Y_{n,i}^\top Y_{n,j} = (X_i - \overline{X}_n)^\top S_n^{-1}(X_j - \overline{X}_n), i, j = 1, \ldots, n$, and is thus affine invariant, see also Section 2 of Henze (2002). Besides, it is not necessary to compute the square root of S_n^{-1}.

A further representation of $T_{n,a}$ is (see Proposition 2.1 of Henze et al. (2003))

$$T_{n,a} = \frac{n(2\pi)^d}{4} \int_{\mathbb{R}^d} \left(\widehat{f}_{n,a}(x) - \widehat{f}_{n,a}(-x)\right)^2 \mathrm{d}x, \tag{3}$$

where

$$\widehat{f}_{n,a}(t) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{(2\pi a)^{d/2}}\exp\left(-\frac{\|t-Y_{n,j}\|^2}{2a}\right), \quad t \in \mathbb{R}^d. \tag{4}$$

Notice that $\widehat{f}_{n,a}(x)$ figuring in (3) is a nonparametric kernel density estimator with Gaussian kernel $(2\pi)^{-d/2}\exp(-\|t\|^2/2)$ and bandwidth $a^{1/2}$, applied to the standardized data $Y_{n,1},\ldots,Y_{n,n}$. Since $\widehat{f}_{n,a}(-x)$ is the same density estimator, applied to the data after reflection at the origin, representation (4) may be regarded as an L^2-distance between two nonparametric kernel density estimators, and the role of a is that of a smoothing parameter. However, in contrast to density estimation where the bandwidth depends on the sample size n, we keep a fixed in order to achieve positive asymptotic power with respect to alternatives that approach H_0 at the rate $n^{-1/2}$. In the spirit of density estimation, it would be tempting to let $a = a_n$ depend on n and have $a_n \to 0$ as $n \to \infty$. In view of results of Gürtler (2000) in connection with the test of Bowman and Foster (1993) for testing for multivariate normality (see also Section 7 of Henze (2002)), we conjecture that such a test would lose positive asymptotic power against "$n^{-1/2}$-close alternatives", and that the limit distribution of T_{n,a_n} under H_0, after a suitable normalization, would be normal.

Some more light on the role of a is cast by the relation

$$\lim_{a\to\infty} a^{d/2+3}\frac{T_{n,a}}{n} = \frac{\pi^{d/2}}{96}\cdot\left(2b_{n,1}+3b_{n,2}\right) \tag{5}$$

(see Proposition 2.2 of Henze et al. (2003)). Here, the limit is elementwise on the underlying probability space, and

$$b_{n,1} = \frac{1}{n^2}\sum_{i,j=1}^{n}\left(Y_{n,i}^{\top}Y_{n,j}\right)^3, \quad b_{n,2} = \frac{1}{n^2}\sum_{i,j=1}^{n}Y_{n,i}^{\top}Y_{n,j}\,\left\|Y_{n,i}\right\|^2\left\|Y_{n,j}\right\|^2$$

denote empirical multivariate skewness in the sense of Mardia (1970) and Móri et al. (1993), respectively. Thus, for large values of a, the test statistic $T_{n,a}$, apart from a scaling factor, is approximately a linear combination of two measures of skewness. In the univariate case $b_{n,1}$ and $b_{n,2}$ coincide, and (5) specializes to give

$$\lim_{a\to\infty} a^{7/2}\frac{T_{n,a}}{n} = \frac{5\sqrt{\pi}}{96}\cdot\left(\frac{1}{n}\sum_{j=1}^{n}\left(\frac{X_j-\overline{X}_n}{s_n}\right)^3\right)^2, \tag{6}$$

where $s_n^2 = n^{-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2$. Hence, apart from a factor, the "limit statistic" $\lim_{a\to\infty} a^{7/2}T_{n,a}$ is just squared sample skewness in the univariate case. A corresponding result for the limit of a measure of asymmetry associated with $T_{n,a}$ will be given in Theorem 3.

Under the more stringent moment assumption $\mathbb{E}\|X\|^4 < \infty$, we have (see Theorem 3.2 of Henze et al. (2003))

$$T_{n,a} \xrightarrow{\mathcal{D}} \int_{\mathbb{R}^d} \mathcal{W}^2(t) \exp\left(-a\|t\|^2\right) \mathrm{d}t,$$

under H_0, where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution, and $\mathcal{W}$ is some centred Gaussian process in the Hilbert space $\mathcal{L}^2 = \mathcal{L}^2(\mathbb{R}^d, \mathcal{B}^d, \exp(-a\,\|t\|^2))$ of (equivalence classes of) measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ that are square integrable with respect to the measure $\exp(-a\,\|t\|^2)\mathrm{d}t$.

Under a triangular array $X_{n,1}, \ldots, X_{n,n}, n \geq d+1$, of row-wise i.i.d. random vectors with density

$$f_n(x) = f_0(x)\left(1 + \frac{h(x)}{\sqrt{n}}\right), \quad x \in \mathbb{R}^d,$$

where f_0 is a density symmetric about 0, and h is a bounded measurable function satisfying $\int_{\mathbb{R}^d} h(x) f_0(x)\mathrm{d}x = 0$, we have

$$T_{n,a} \xrightarrow{\mathcal{D}} \int_{\mathbb{R}^d} (\mathcal{W}(t) + s(t))^2 \exp\left(-a\|t\|^2\right) \mathrm{d}t$$

(see Theorem 3.2 of Henze et al. (2003)), where

$$s(t) = \int_{\mathbb{R}^d} \left[\sin\left(t^\top x\right) - t^\top \psi(t) x\right] h(x) f_0(x)\,\mathrm{d}x, \quad \psi(t) = \int_{\mathbb{R}^d} \cos(t^\top x) f_0(x)\,\mathrm{d}x.$$

Hence, the test has positive asymptotic power against close alternatives that approach the null hypothesis at the rate $n^{-1/2}$, provided that the function $s(\cdot)$ does not vanish.

Since both the finite-sample and the limit null distribution of $T_{n,a}$ depend on the unknown distribution of X, the test is carried out as permutation test. To this end, let $U_1, U_2, \ldots$ be a sequence of i.i.d. random variables, independent of $X_1, X_2, \ldots$, such that $\mathbb{P}(U_j = 1) = \mathbb{P}(U_j = -1) = 1/2$. Conditionally on $Y_{n,j} = y_j$, $j = 1, \ldots, n$, let $Z_j = U_j y_j$, $j = 1, \ldots, n$ and put $\overline{Z}_n = n^{-1} \sum_{j=1}^n Z_j$. Henze et al. (2003) shows that the permutation statistic

$$T_{n,a}^P = \int_{\mathbb{R}^d} \left(\mathcal{W}_n^P(t)\right)^2 \exp(-a\|t\|^2)\,\mathrm{d}t,$$

which is based on the so-called *permutation process*

$$\mathcal{W}_n^P(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n U_j \left\{\sin(t^\top y_j) - \left(\frac{1}{n}\sum_{k=1}^n \cos(t^\top y_k)\right) t^\top y_j\right\},$$

takes the form

$$T_{n,a}^P = \frac{\pi^{d/2}}{2a^{d/2}n} \sum_{i,j=1}^{n} \left[\left(2 + \frac{\|\overline{Z}_n\|^2}{2a} - \left\{ 1 + \frac{(Z_i - Z_j)^\top \overline{Z}_n}{2a} \right\}^2 \right) \exp\left(-\frac{\|Z_i - Z_j\|^2}{4a} \right) \right. \\ \left. + \left(\frac{\|\overline{Z}_n\|^2}{2a} - \left\{ 1 + \frac{(Z_i + Z_j)^\top \overline{Z}_n}{2a} \right\}^2 \right) \exp\left(-\frac{\|Z_i + Z_j\|^2}{4a} \right) \right].$$

Moreover, by Theorem 4.2 of Henze et al. (2003), the limit distribution of $T_{n,a}^P$ under H_0 is the same as that of $T_{n,a}$ for almost all sample sequences $X_1, X_2, \ldots$. Under a fixed alternative distribution satisfying $\mathbb{E}\|X\|^2 < \infty$ (which, in view of affine invariance, is assumed to have zero expectation and unit covariance matrix), we have

$$\lim_{n\to\infty} \mathbb{P}(T_{n,a} > c_{n,a}^P(\alpha)) = 1,$$

where $c_{n,a}^P(\alpha)$ denotes the $(1-\alpha)$-quantile of the distribution of the permutation statistic $T_{n,a}^P$, see Theorem 5.1 of Henze et al. (2003). Since

$$\liminf_{n\to\infty} \frac{T_{n,a}}{n} \geq \int_{\mathbb{R}^d} \left(\mathbb{E}[\sin(t^\top X)] \right)^2 \exp\left(-a\|t\|^2 \right) \mathrm{d}t \tag{7}$$

almost surely (see display (5.1) of Henze et al. (2003)), and since the right-hand side of (7) is strictly positive if the distribution of X is not reflectedly symmetric, we have $\lim_{n\to\infty} T_{n,a} = \infty$ almost surely for any such distribution. In view of the fact that $c_{n,a}^P(\alpha)$ is bounded in probability almost surely, the test based on $T_{n,a}$ is consistent against such an alternative.

We stress that although rejecting H_0 for large values of $T_{n,a}$ yields a globally consistent test for each fixed $a > 0$, this property is "lost in the limit $a \to \infty$", i.e., if one considers the test statistic $2b_{n,1} + 3b_{n,2}$ figuring on the right-hand side of (5).

To carry out the test in practice, one generates M independent pseudo-random vectors $(U_1, \ldots, U_n)$, where $U_1, \ldots, U_n$ are i.i.d. with a uniform distribution on $\{-1, +1\}$, and calculates the corresponding realizations $T_{n,a}^P(j)$, $1 \leq j \leq M$ (say), of the permutation statistic $T_{n,a}^P$. The hypothesis H_0 is rejected at level α, if the value of $T_{n,a}$ exceeds the empirical $(1-\alpha)$-quantile of $T_{n,a}^P(j)$, $1 \leq j \leq M$. In Section 6, we used $M = 100000$ to assess whether the 6-dimensional data set is skew, and the p values given in Table 8 are based on $M = 1000$ pseudo-random vectors.

3 Behaviour under fixed alternatives

In this section, we assume that the distribution of X is *not* symmetric. In view of affine invariance, we further assume without loss of generality that $\mathbb{E}[X] = 0$ and $\mathbb{E}[XX^\top] = \mathrm{I}_d$, where I_d stands for the unit matrix of order d (recall the standing assumption $\mathbb{E}\|X\|^2 < \infty$). In what follows,

$$\mathrm{R}(t) = \mathbb{E}\left[\cos(t^\top X)\right], \quad \mathrm{I}(t) = \mathbb{E}\left[\sin(t^\top X)\right], \quad t \in \mathbb{R}^d, \tag{8}$$

denote the real and the imaginary part of the characteristic function of X, respectively.

The first result shows that the almost sure lower bound of $T_{n,a}/n$ figuring in (7) is the almost sure limit of $T_{n,a}/n$.

Theorem 1 *We have*

$$\frac{T_{n,a}}{n} \xrightarrow{\text{a.s.}} \Delta_a \quad \text{as } n \to \infty,$$

where

$$\Delta_a = \int_{\mathbb{R}^d} I(t)^2 \exp\left(-a\|t\|^2\right) \mathrm{d}t. \tag{9}$$

Interestingly, there is an alternative expression for the measure of distance Δ_a from symmetry figuring in (9).

Theorem 2 *We have*

$$\Delta_a = \frac{1}{4a^d} \int_{\mathbb{R}^d} \left(\mathbb{E}\left[\exp\left(-\frac{\|x - X\|^2}{2a}\right) - \exp\left(-\frac{\|-x - X\|^2}{2a}\right)\right]\right)^2 \mathrm{d}x. \tag{10}$$

The next result complements the finite-n-limit (6) and sheds even more light on the measure Δ_a of asymmetry.

Theorem 3 *Suppose that $d = 1$, and that $\mathbb{E}|X|^3 < \infty$. We then have*

$$\lim_{a\to\infty} a^{7/2}\Delta_a = \frac{5\sqrt{\pi}}{96} \cdot \left(\mathbb{E}[X^3]\right)^2.$$

Notice that, because of $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^2) = 1$, the right-hand side is a factor times squared skewness in the sense of Pearson.

To state a result on the limit distribution of $T_{n,a}$ under fixed alternatives, it will be convenient to introduce the $\mathbb{R}^d$-valued functions

$$\mathrm{C}(t) = \mathbb{E}\left[X\cos\left(t^\top X\right)\right], \quad \mathrm{S}(t) = \mathbb{E}\left[X\sin\left(t^\top X\right)\right] \quad t \in \mathbb{R}^d. \tag{11}$$

Theorem 4 *If $\mathbb{E}\|X\|^4 < \infty$, we have*

$$\sqrt{n}\left(\frac{T_{n,a}}{n} - \Delta_a\right) \xrightarrow{\mathcal{D}} \mathrm{N}\left(0, \sigma_a^2\right),$$

where

$$\sigma_a^2 = 4\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K(s,t)\, I(s) I(t) \exp\left(-a(\|s\|^2 + \|t\|^2)\right) \mathrm{d}s\mathrm{d}t \tag{12}$$

and

$$\begin{aligned}K(s,t) &= \mathbb{E}\left[\sin\left(s^\top X\right)\sin\left(t^\top X\right)\right] - I(s)I(t) - R(t)t^\top S(s) - R(s)s^\top S(t)\\ &\quad + R(s)R(t)s^\top t - \frac{1}{2}t^\top \mathbb{E}\left[\sin\left(s^\top X\right)XX^\top\right]C(t) + \frac{1}{2}I(s)t^\top C(t)\\ &\quad - \frac{1}{2}s^\top \mathbb{E}\left[\sin\left(t^\top X\right)XX^\top\right]C(s) + \frac{1}{2}I(t)s^\top C(s)\\ &\quad + \frac{1}{2}s^\top R(s)\mathbb{E}\left[Xt^\top XX^\top\right]C(t) + \frac{1}{2}t^\top R(t)\mathbb{E}\left[Xs^\top XX^\top\right]C(s)\\ &\quad + \frac{1}{4}\left\{C(s)^\top \mathbb{E}\left[XX^\top st^\top XX^\top\right]C(t) - s^\top C(s)t^\top C(t)\right\}, \quad s,t \in \mathbb{R}^d.\end{aligned}$$

Remark 1 Whereas the moment assumption $\mathbb{E}\|X\|^4 < \infty$ suffices to show asymptotic normality of $T_{n,a}$ for each $a > 0$ under a fixed alternative to H_0, the asymptotics for the "limit statistic" $2b_{n,1} + 3b_{n,2}$ figuring on the right-hand side of (5) require the stronger assumption $\mathbb{E}\|X\|^6 < \infty$, see Theorem 2.1 of Henze (1997b). The same holds for the asymptotics of $b_{n,1}$ and $b_{n,2}$, which are given in Baringhaus and Henze (1992) and Henze (1997a), respectively.

4 Estimation of σ_a^2

Theorem 4 paves the way to an asymptotic confidence interval for Δ_a provided that a consistent estimator $\widehat{\sigma}_{n,a}^2 = \widehat{\sigma}_{n,a}^2(X_1,\ldots,X_n)$ of the variance σ_a^2 figuring in (12) is available. Since Theorem 4 requires $\mathbb{E}(X) = 0$ and $\mathbb{E}(XX^\top) = \mathrm{I}_d$, we base such an estimator on the empirically standardized data defined in (2), where we put $Y_j = Y_{n,j}$ for the sake of brevity in what follows. Moreover, let $w_a(s,t) = \exp\left(-a(\|s\|^2 + \|t\|^2)\right)$. Such an estimator is

$$\widehat{\sigma}_{n,a}^2 = 4\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K_n(s,t)\,\mathrm{I}_n(s)\mathrm{I}_n(t)\,w_a(s,t)\,\mathrm{d}s\mathrm{d}t. \tag{13}$$

Here, $K_n(s,t)$ is the empirical version of $K(s,t)$ figuring in the statement of Theorem 4. This version originates from $K(s,t)$ by replacing the functions $\mathrm{R}(\cdot)$, $\mathrm{I}(\cdot)$, $\mathrm{C}(\cdot)$ and $\mathrm{S}(\cdot)$ defined in (8) and (11) with their respective empirical counterparts

$$\mathrm{R}_n(t) = \frac{1}{n}\sum_{j=1}^n \cos\left(t^\top Y_j\right), \quad \mathrm{I}_n(t) = \frac{1}{n}\sum_{j=1}^n \sin\left(t^\top Y_j\right),$$
$$\mathrm{C}_n(t) = \frac{1}{n}\sum_{j=1}^n Y_j\cos\left(t^\top Y_j\right), \quad \mathrm{S}_n(t) = \frac{1}{n}\sum_{j=1}^n Y_j\sin\left(t^\top Y_j\right), \quad t \in \mathbb{R}^d,$$

and doing the same with each of the five explicitly designated expectations figuring in the definition of $K(s,t)$. Thus, for example, $\mathbb{E}\left[\sin\left(s^\top X\right)\sin\left(t^\top X\right)\right]$ is replaced with $n^{-1}\sum_{j=1}^n \sin(s^\top Y_j)\sin(t^\top Y_j)$, etc. To give an expression of $\widehat{\sigma}^2_{n,a}$ that does not involve any integration and is thus amenable to computational purposes, we put

$$\rho_{1,a}(u,v) := \int_{\mathbb{R}^d} \sin(u^\top t)\sin(v^\top t)\exp(-a\|t\|^2)\,\mathrm{d}t,$$
$$\rho_{2,a}(u,v) := \int_{\mathbb{R}^d} t\cos(u^\top t)\sin(v^\top t)\exp(-a\|t\|^2)\,\mathrm{d}t, \quad u,v\in\mathbb{R}^d.$$

These integrals can be evaluated to give

$$\rho_{1,a}(u,v) = \frac{1}{2}\left(\frac{\pi}{a}\right)^{d/2}\left(\exp\left(-\frac{\|u-v\|^2}{4a}\right) - \exp\left(-\frac{\|u+v\|^2}{4a}\right)\right),$$
$$\rho_{2,a}(u,v) = \frac{1}{4a}\left(\frac{\pi}{a}\right)^{d/2}\left((v-u)\exp\left(-\frac{\|v-u\|^2}{4a}\right) + (v+u)\exp\left(-\frac{\|v+u\|^2}{4a}\right)\right).$$

Notice that the function $\rho_{2,a}$ takes values in $\mathbb{R}^d$. Suppressing the dependence on a, let

$$V_{n,r} := \frac{1}{n^2}\sum_{i,j=1}^n \rho_{r,a}(Y_i,Y_j), \quad \overline{V}_{n,r}(y) := \frac{1}{n}\sum_{\ell=1}^n \rho_{r,a}(y,Y_\ell), \quad y\in\mathbb{R}^d,\ r\in\{1,2\},$$
$$\varSigma_n := \frac{1}{n^2}\sum_{i,k=1}^n \rho_{1,a}(Y_i,Y_k)Y_iY_i^\top, \quad \varGamma_n := \frac{1}{n^2}\sum_{i,\ell=1}^n \rho_{2,a}(Y_i,Y_\ell)Y_i^\top,$$

a computationally feasible expression for $\widehat{\sigma}^2_{n,a}$ is given as follows.

Proposition 1 *We have*

$$\begin{aligned}\widehat{\sigma}^2_{n,a} = {} & \frac{4}{n}\sum_{j=1}^n \overline{V}_{n,1}(Y_j)^2 - 4V_{n,1}^2 - 8\left(\frac{1}{n}\sum_{j=1}^n \overline{V}_{n,1}(Y_j)Y_j\right)^\top V_{n,2} + 4\|V_{n,2}\|^2 \\ & - 4\,tr\left(\varSigma_n\left(\frac{1}{n}\sum_{j=1}^n \overline{V}_{n,2}(Y_j)Y_j^\top\right)\right) + 4V_{n,1}\frac{1}{n}\sum_{j=1}^n Y_j^\top\overline{V}_{n,2}(Y_j) \\ & + 4V_{n,2}^\top\left(\frac{1}{n^2}\sum_{j,k=1}^n Y_jY_j^\top Y_kY_j^\top\overline{V}_{n,2}(Y_k)\right) \\ & + \frac{1}{n}\sum_{j=1}^n\left(Y_j^\top\varGamma_nY_j\right)^2 - \left(\frac{1}{n}\sum_{i=1}^n Y_i^\top\overline{V}_{n,2}(Y_i)\right)^2.\end{aligned}$$

The next result shows that $\widehat{\sigma}^2_{n,a}$ defined in (13) is a consistent estimator of σ^2_a defined in (12).

Theorem 5 *If* $\mathbb{E}\|X\|^4 < \infty$, *we have*

$$\widehat{\sigma}_{n,a}^2 \xrightarrow{\mathbb{P}} \sigma_a^2.$$

The proof is extremely tedious but in principle straightforward. A similar problem was encountered in Gürtler (2000) in the context of estimating the variance of the limit normal distribution of the BHEP test for multivariate normality under a fixed alternative distribution. Details are given in the final Section 8.

From Theorem 5 and Theorem 4, we obtain the following asymptotic confidence interval for Δ_a.

Corollary 1 *For* $\alpha \in (0, 1)$, *let* $z_{1-\alpha/2}$ *be the* $(1-\alpha/2)$-*quantile of the standard normal distribution. Then,*

$$I_{n,a,\alpha} = \left[\frac{T_{n,a}}{n} - \frac{\widehat{\sigma}_{n,a}}{\sqrt{n}} z_{1-\alpha/2}, \frac{T_{n,a}}{n} + \frac{\widehat{\sigma}_{n,a}}{\sqrt{n}} z_{1-\alpha/2}\right] \tag{14}$$

is an asymptotic confidence interval for Δ_a *at level* $1-\alpha$, *i.e., we have* $\lim_{n\to\infty} \mathbb{P}(I_{n,a,\alpha} \ni \Delta_a) = 1 - \alpha$ *if* $\mathbb{E}\|X\|^4 < \infty$.

5 Discussion and Examples

5.1 Δ_a as a measure of asymmetry

Since the distribution of a d-dimensional random vector X is symmetric with respect to the origin if, and only if, the imaginary part of its characteristic function vanishes, $\Delta_a =: \Delta_a(X)$ (say) defined in (9) is zero if, and only if, we have $X \stackrel{\mathcal{D}}{=} -X$. Thus, Δ_a, without any assumption on the distribution of X, may be regarded as a measure of deviation from reflected symmetry with respect to the origin. In what follows, we almost exclusively confine to the case $d = 1$. As an example, consider a random variable with the normal distribution $\mathrm{N}(\mu, \sigma^2)$. Some straightforward algebra shows that, for this distribution, Δ_a takes the value

$$\frac{\sqrt{\pi}}{2\sqrt{a+\sigma^2}}\left(1 - \exp\left(-\frac{\mu^2}{a+\sigma^2}\right)\right).$$

In particular, as may have been anticipated, the degree of asymmetry with respect to 0 increases with increasing $|\mu|$ and, for fixed μ, it decreases as σ^2 increases.

Now suppose, as in the previous sections, that $\mathbb{E}\|X\|^2 < \infty$, and that the covariance matrix Σ of X is not degenerate. Furthermore, put $\mu = \mathbb{E}(X)$. If one computes Δ_a for the standardized random vector $Y := \Sigma^{-1/2}(X - \mu)$, then $\Delta_a(Y) = 0$ if, and only if, the distribution of X is reflectedly symmetric around μ. Since the test statistic $T_{n,a}$ is affine invariant, we will (as before) assume without loss of generality $\mu = 0$ and $\Sigma = \mathrm{I}_d$.

For the sake of illustration, we will by analogy with (Partlett and Prakash 2017) consider several asymmetric univariate distributions. In contrast to Partlett and Prakash (2017), however, our distributions are *standardized*. The distributions under discussion are an exponential distribution (denoted by E), an adjusted lognormal distribution (denoted by LN), and a modified folded normal distribution, denoted by |N|. Moreover, analogously to Partlett and Prakash (2017), we consider the following normal mixture, in order to investigate cases of fairly weak asymmetry. Since there is not much extra work involved, we present this mixture for a general $d \geq 1$. Suppose that

$$X \stackrel{\mathcal{D}}{=} TY_1 + (1-T)Y_2,$$

where T, Y_1, Y_2 are independent, $\mathbb{P}(T = 1) = p = 1 - \mathbb{P}(T = 0)$, $0 \leq p < 1/2$, $Y_1 \stackrel{\mathcal{D}}{=} \mathrm{N}(\mathrm{e}_1, \mathrm{I}_d - \frac{p}{1-p}\mathrm{e}_1\mathrm{e}_1^\top)$ and $Y_2 \stackrel{\mathcal{D}}{=} \mathrm{N}(-p/(1-p)\mathrm{e}_1, \mathrm{I}_d - \frac{p}{1-p}\mathrm{e}_1\mathrm{e}_1^\top)$, where $\mathrm{e}_1 = (1, 0, \ldots, 0)^\top$ is the first canonical unit vector in $\mathbb{R}^d$. In view of $T^2 = T$ and independence, we have

$$\mathbb{E}(X) = \mathbb{E}(T)\mathbb{E}(Y_1) + (1 - \mathbb{E}(T))\mathbb{E}(Y_2) = 0,$$
$$\mathbb{E}(XX^\top) = \mathbb{E}(T)\mathbb{E}(Y_1Y_1^\top) + (1 - \mathbb{E}(T))\mathbb{E}(Y_2Y_2^\top) = \mathrm{I}_d.$$

The addition theorem for the sine function gives

$$\sin(t^\top X) \stackrel{\mathcal{D}}{=} \sin(Tt^\top Y_1)\cos((1-T)t^\top Y_2) + \cos(Tt^\top Y_1)\sin((1-T)t^\top Y_2),$$

and conditioning on T it follows that

$$\mathrm{I}(t) = \mathbb{E}[\sin(t^\top X)] = p\,\mathbb{E}[\sin(t^\top Y_1)] + (1-p)\,\mathbb{E}[\sin(t^\top Y_2)].$$

Writing $t = (t_1, \ldots, t_d)^\top$, we have

$$t^\top Y_1 \stackrel{\mathcal{D}}{=} \mathrm{N}\left(t_1, \|t\|^2 + \left(\frac{1-2p}{1-p} - 1\right)t_1^2\right),$$
$$t^\top Y_2 \stackrel{\mathcal{D}}{=} \mathrm{N}\left(-pt_1/(1-p), \|t\|^2 + \left(\frac{1-2p}{1-p} - 1\right)t_1^2\right).$$

Since the characteristic function of the normal distribution $\mathrm{N}(\mu, \sigma_a^2)$ is $\exp(\mathrm{i}\xi\mu - \sigma_a^2\xi^2/2)$, $\xi \in \mathbb{R}$, it follows that

$$\mathrm{I}(t) = \exp\left(-\frac{\|t\|^2 + \left(\frac{1-2p}{1-p} - 1\right)t_1^2}{2}\right)\left(p\sin t_1 - (1-p)\sin\left(\frac{pt_1}{1-p}\right)\right).$$

Thus,

$$\begin{aligned}\Delta_a &= p^2 \int_{\mathbb{R}^d} \sin^2(t_1) \exp\left(-(1+a)\|t\|^2 - \left(\frac{1-2p}{1-p} - 1\right) t_1^2\right) \mathrm{d}t \\ &+(1-p)^2 \int_{\mathbb{R}^d} \sin^2\left(\frac{pt_1}{1-p}\right) \exp\left(-(1+a)\|t\|^2 - \left(\frac{1-2p}{1-p} - 1\right) t_1^2\right) \mathrm{d}t \\ &-2p(1-p) \int_{\mathbb{R}^d} \sin(t_1)\, \sin\left(\frac{pt_1}{1-p}\right) \exp\left(-(1+a)\|t\|^2 - \left(\frac{1-2p}{1-p} - 1\right) t_1^2\right) \mathrm{d}t.\end{aligned}$$

Since $\int_{-\infty}^{\infty} \exp(-(1+a)\xi^2)\,\mathrm{d}\xi = \sqrt{\pi/(1+a)}$, the computation of Δ_a boils down to the calculation of integrals of the type

$$\begin{aligned}\int_{-\infty}^{\infty} \sin(\alpha\xi) \sin(\beta\xi) \exp(-\gamma\xi^2)\mathrm{d}\xi &= \frac{\sqrt{\pi}}{2\sqrt{\gamma}} \\ &\times \left(\exp\left(-\frac{(\alpha-\beta)^2}{4\gamma}\right) - \exp\left(-\frac{(\alpha+\beta)^2}{4\gamma}\right)\right),\end{aligned}$$

where $\alpha, \beta \in \mathbb{R}$ and $\gamma > 0$. After tedious but straightforward calculations, one obtains

$$\begin{aligned}\Delta_a = \left(\frac{\pi}{a+1}\right)^{(d-1)/2} &\sqrt{\frac{\pi}{\gamma_a}} \left[\frac{p^2}{2}\left(1 - \exp\left(-\frac{1}{\gamma_a}\right)\right)\right. \\ &+\frac{(1-p)^2}{2}\left(1 - \exp\left(-\frac{p^2}{(1-p)^2\gamma_a}\right)\right) \\ &\left.-p(1-p)\left(\exp\left(-\frac{(1-2p)^2}{4(1-p)^2\gamma_a}\right) - \exp\left(-\frac{1}{4(1-p)^2\gamma_a}\right)\right)\right],\end{aligned}$$

where $\gamma_a = a + (1-2p)/(1-p)$. For the following comparison, we choose the values $p = 0.25$ and $p = 0.4$. The resulting normal mixtures are denoted with N1 and N2, respectively. The values of Δ_a for the distributions E, LN and |N| were computed using numerical integration. The results are given in Table 1 for different values of a.

Suppose that random vectors X and Y have distributions P and Q, respectively. Writing $P <_a Q$ if $\Delta_a(X) < \Delta_a(Y)$, i.e., if the distribution of X is less asymmetric than that of Y, when measured by in terms of Δ_a, Table 1 exhibits the ordering

$$N1 <_a N2 <_a |N| <_a E <_a LN \tag{15}$$

of asymmetry, at least if $a \in \{0.01, 0.1, 4\}$. On the other hand, Δ_a(E) is a bit larger than Δ_a(LN) if $a = 1$. Ordering (15) is mostly compatible with an intuitive notion of the strength of asymmetry of a distribution, as seen from Figure 1 and 2. These figures display the densities of the distributions N1, N2, |N| , E, and LN. Furthermore, we plotted corresponding histograms based on 10 000 000 samples in order to enhance the impression of asymmetry.

Table 1 Values of Δ_a, $a \in \{0.01, 0.1, 1, 4\}$, for N1, N2, |N|, LN and E

	$a = 0.01$	$a = 0.1$	$a = 1$	$a = 4$
N1	0.01039	0.00713	0.00061	0.00002
N2	0.05062	0.02889	0.00115	0.00002
|N|	0.19063	0.11333	0.00972	0.00034
E	0.55786	0.29080	0.01961	0.00084
LN	1.14876	0.45110	0.01912	0.00104

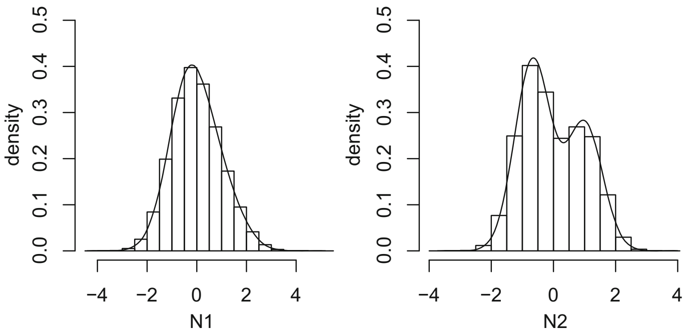

Fig. 1 Densities and histograms: N1 (left) and N2 (right)

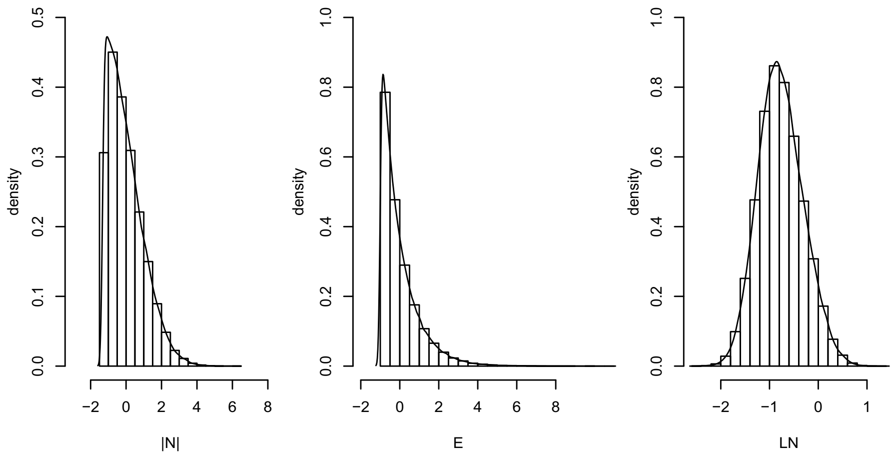

Fig. 2 Densities and histograms: |N| (left), E (middle) and LN (right)

5.2 Confidence intervals for Δ_a

Using the above normal mixture with $p = 0.25$ and $p = 0.4$ in the case $d = 1$, we investigated whether the estimator $T_{n,a}/n$ of Δ_a is useful for practical purposes. Since the normal mixture exhibits fairly weak asymmetry, we studied the performance of $T_{n,a}/n$ also on centred Exp(1) distributed samples, which represent a much stronger degree of asymmetry. To obtain a reasonable conclusion, we computed the underlying

Table 2 Estimated values based on 1000 samples of the distributions N1, N2 and E, $a = 0.01$

	n	$\varnothing T_{n,a}/n$	$\varnothing\widehat{\sigma}^2_{n,a}$	$\widehat{\sigma}^2_{n,a} < 0$	estimated $1-\alpha$	Relative MSE
N1	40	0.1945	1.4770	1	93.7	4.3572
	80	0.1022	0.6431	1	96.6	1.0978
	100	0.0846	0.4946	0	96.3	0.7377
	250	0.0391	0.1986	0	97.9	0.1221
	500	0.0240	0.1211	0	98.0	0.0344
N2	40	0.2355	1.9090	1	97.0	1.0283
	80	0.1465	1.0602	0	98.4	0.3180
	100	0.1300	0.9362	0	98.2	0.2428
	250	0.0798	0.6020	0	97.7	0.0579
	500	0.0641	0.5073	0	95.9	0.0218
E	40	0.7120	6.2629	0	98.4	0.2087
	80	0.6442	5.1665	0	97.2	0.1095
	100	0.6244	4.9267	0	97.7	0.0771
	250	0.5866	4.5024	0	97.2	0.0288
	500	0.5726	4.3538	0	96.8	0.0134

value of σ_a^2 for the latter distribution by means of numerical integration and made use of the results of the previous subsection.

Regarding the choice of the parameter a, note that small values of a entail bigger values for both Δ_a and $T_{n,a}/n$, and likewise for σ_a^2 and $\widehat{\sigma}^2_{n,a}$. To bypass computational inaccuracies and to avoid negative values of $\widehat{\sigma}^2_{n,a}$ that sometimes show up in small samples, we used mainly small values for a, which seems to have no disadvantages at all. Nevertheless, the qualitative behaviour of the estimates is similar if the sample size is big enough. To judge the effect of a, the outcome of the simulation study is displayed in Table 2 for $a = 0.01$ and in Table 3 for $a = 0.1$. For each combination of the sample size n, the parameter a, and the underlying distribution, we performed 1000 simulations and computed the sample mean of $T_{n,a}/n$ (denoted by $\varnothing T_{n,a}/n$) and of the sample variance $\widehat{\sigma}^2_{n,a}$ (denoted by $\varnothing\widehat{\sigma}^2_{n,a}$) as estimates of Δ_a and σ_a^2, respectively. Thereby, using (14) with $\alpha = 0.05$, we calculated an approximation for the $(1-\alpha)$-confidence interval and observed how often the interval contained Δ_a. The average number per 100 samples is displayed in the columns called "estimated $1-\alpha$". Furthermore, we highlighted the total number of negative estimates for σ^2 as "$\widehat{\sigma}^2_{n,a} < 0$". Finally, the relative mean squared error of $T_{n,a}/n$, i.e.,

$$\frac{\frac{1}{1000}\sum_{j=1}^{1000}\left(\frac{T_n}{n} - \Delta_a\right)}{\Delta_a},$$

is denoted as "relative MSE". In Table 2, the true values of Δ_a are $\Delta_a = 0.0104$ for N1, $\Delta_a = 0.0506$ for N2 and $\Delta_a = 0.5577$ for E. Furthermore, the value of σ_a^2 is 3.0409 for the distribution E. Since there is no known formula to represent σ_a^2

Table 3 Estimated values based on 1000 samples of the distributions N1, N2 and E, $a = 0.1$

	n	$\varnothing T_{n,a}/n$	$\varnothing \widehat{\sigma}^2_{n,a}$	$\widehat{\sigma}^2_{n,a} < 0$	Estimated $1-\alpha$	Relative MSE
N1	40	0.0436	0.1452	14	95.6	0.4535
	80	0.0240	0.0800	6	98.0	0.1104
	100	0.0220	0.0733	3	98.0	0.0895
	250	0.0130	0.0422	2	95.3	0.0213
	500	0.0100	0.0327	1	91.5	0.0087
N2	40	0.0614	0.2721	24	95.0	0.1897
	80	0.0488	0.2086	11	92.1	0.0791
	100	0.0475	0.2066	3	91.9	0.0694
	250	0.0344	0.1554	1	92.1	0.0188
	500	0.0321	0.1491	0	93.2	0.0153
E	40	0.3052	0.9429	1	91.9	0.0190
	80	0.3026	0.9135	0	94.0	0.0113
	100	0.2994	0.9028	0	94.7	0.0084
	250	0.2949	0.8679	0	95.5	0.0033
	500	0.2941	0.8541	0	95.2	0.0017

as a composition of sums, the exact computation is time-consuming and is therefore omitted. In Table 3, the true values of Δ_a are $\Delta_a = 0.0071$ for N1, $\Delta_a = 0.0289$ for N2 and $\Delta_a = 0.2908$ for the centred standard exponential distribution E. For the latter distribution, the value of σ_a^2 is $\sigma_a^2 = 0.8875$.

As each table indicates, the desired properties can also be seen in practical applications. Even for small sample sizes, the computed intervals maintain the nominal level, and the estimator $T_{n,a}/n$ quantifies the departure from symmetry for fixed a. Furthermore, the relative mean squared error decreases quickly as the sample size increases.

5.3 An alternative test procedure

We now assess whether the hypothesis of symmetry can be rejected or accepted by looking at the confidence interval $I_{n,a,\alpha}$ for Δ_a, as given in (14). To this end, one might think of testing for symmetry by rejecting the hypothesis H_0 in (1) if $0 \notin I_{n,a,\alpha}$. Our study was carried out using the distributions of Subsection 5.1. Furthermore, we applied the method on the standard normal distribution N(0,1) and on a uniform distribution in [−1 ,1], denoted by U[−1 ,1], to check whether the nominal level was maintained. The same reasoning as in Sect. 5.2 led to the choices $a = 0.01$ and $a = 0.1$. Again, we performed 1000 simulations and used the sample sizes $n = 40$, $n = 60$, $n = 80$, $n = 100$ and $n = 250$. As Tables 4 and 5 indicate, this test procedure shows quite good results. To compare the procedure to the permutation test, as studied in Henze et al. (2003), we expanded the simulations of Henze et al. (2003) by a simulation study for $d = 1$ and the choices $a = 1$ and $a = 4$. The results of Tables 6 and 7 indicate that it depends on the underlying distribution whether the test based on the confidence interval or the permutation test is more favourable.

Table 4 Percentages of rejection of H_0 based on 1000 samples, $a = 0.01$ (confidence interval procedure)

	$n = 40$	$n = 60$	$n = 80$	$n = 100$	$n = 250$
N(0,1)	5.9	5.3	5.5	4.1	3.5
U[−1 ,1]	3	2.9	2.5	3.1	4.1
N1	74.9	94.7	98.3	99.8	100
N2	63.1	68.9	72.8	81	95.4
|N|	22.7	30.2	57.2	71.8	99.8
LN	39.4	58.3	76.7	87.1	99.2
E	45.1	69	86.9	95.4	100

Table 5 Percentages of rejection of H_0 based on 1000 samples, $a = 0.1$ (confidence interval procedure)

	$n = 40$	$n = 60$	$n = 80$	$n = 100$	$n = 250$
N(0,1)	5	4.2	3.3	2.8	1.6
U[−1 ,1]	2.5	1.9	1.7	2	4.1
N1	0.3	16	71.3	90.8	100
N2	19	27.2	33.5	44.1	72.8
|N|	18	39.8	46.8	61.4	99.6
LN	58.6	72	83.2	91.5	100
E	50.9	73	89.3	96	100

Table 6 Percentages of rejection of H_0 based on 1000 samples, $a = 1$ (permutation test)

	$n = 40$	$n = 60$	$n = 80$	$n = 100$	$n = 250$
N(0,1)	5.6	5.6	6.9	6.4	4.1
U[−1 ,1]	4.2	3.9	5.2	6.2	5.4
N1	8.3	13.9	12.2	15.1	37.1
N2	14.4	19.6	25	29.6	64
|N|	74.9	89	97.1	100	100
LN	99.5	100	100	100	100
E	97.1	100	100	100	100

Table 7 Percentages of rejection of H_0 based on 1000 samples, $a = 4$ (permutation test)

	n=40	$n = 60$	$n = 80$	$n = 100$	$n = 250$
N(0,1)	5.5	5.5	6.5	6.6	4.2
U[−1 ,1]	3.5	3.9	5.2	5.9	5.2
N1	8.2	13.2	11.9	15.6	40
N2	12.9	17.4	21.8	27.2	60
|N|	74.7	89.8	97.6	100	100
LN	99.5	100	100	100	100
E	97.2	100	100	100	100

6 A real data example

In Henze et al. (2003), the HKM test was suggested, and several mathematical properties have been obtained. In what follows, we apply the test to a data set that originated from a health survey of paint sprayers in a car assembly plant. This data set, which is given in Royston (1983), contains 103 observations, each consisting of 6 variates, namely:

1. haemoglobin concentration,
2. PCV packed cell volume,
3. white blood cell count,
4. lymphocyte count,
5. neutrophil count,
6. serum lead concentration.

As is a common procedure for haematological data (see, e.g., Royston (1983)), we applied a logarithmic transformation to each of the variates 3.–6., since these exhibit skewed distributions. Royston (1983) first investigated whether the transformed data arise from a normal distribution. Since three observations seem to be outliers, they were removed. By applying a multivariate generalization of the Shapiro–Wilk test for univariate normality, Royston deduced that the 6-dimensional data showed significant departures from normality, although such a conclusion could not be drawn for any of the bivariate marginal distributions.

From an application of a covariance-matrix-based Wald test to the transformed full data set, Schott (2002) arrived at the same result. Since a test for elliptical symmetry, applied to the same data set, gave a p value of 0.11, Schott argued that it is not unreasonable to assume that the sample originates from an elliptical distribution.

Using a Chi-square type statistic for testing for elliptical symmetry, Batsidis et al. (2014) even obtained a p value larger than 0.9 and thus did not find any evidence for rejecting the hypothesis of elliptical symmetry. The latter findings are in stark contrast to the results that originate when applying the HKM test to the full data set. Astonishingly, the test rejected the hypothesis of *central symmetry* with a p value of $7 \cdot 10^{-5}$ using $a = 1$. Taking $a = 0.5$, $a = 2$ and $a = 4$ leads to p values of a similar magnitude. Since central symmetry is a necessary condition for elliptical symmetry, we can also strongly reject the hypothesis of elliptical symmetry of the 6-variate full data set.

To investigate whether the declared outliers are responsible for rejecting symmetry, we removed these values (observations 21, 47, and 52 in the data set given in Royston (1983)) and applied the HKM test. Again taking $a = 0.5$, $a = 1$, $a = 2$, and $a = 4$, we obtained p values of magnitude 10^{-3}. Consequently, also the remaining data exhibit strong asymmetry.

We finally addressed the question whether any bivariate combination of the 6-dimensional logarithmically transformed data (without outliers) is compatible with the hypothesis of reflected symmetry. Looking at the two plots in Fig. 3, both combinations seem to be equally symmetric or rather skew. However, taking $a = 1$ we obtained the p values given in Table 8. Apparently, the desired 5% level of significance is only exceeded for the combinations “haemoglobin concentration—white blood cell

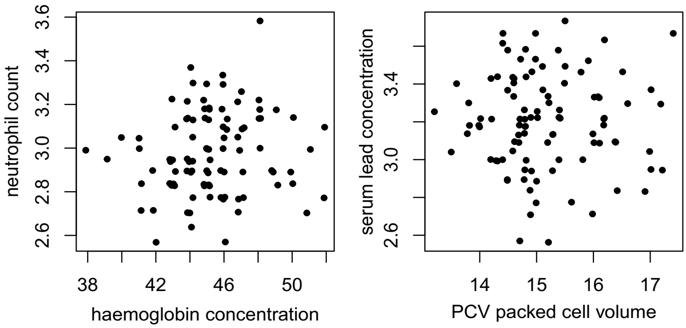

Fig. 3 Scatterplots of haemoglobin concentration— neutrophil count (left) and PCV packed cell volume—serum lead concentration (right)

Table 8 p values of the bivariate HKM test with parameter $a = 1$

	p.c.v.	w.b.c. count	l. count	n. count	s.l. con.
haem. con.	0.158	0.027	0.227	0.028	0.111
p.c.v.		0.252	0.694	0.531	0.699
w.b.c. count			0.164	0.076	0.286
l. count				0.732	0.381
n. count					0.645

count" and "haemoglobin concentration—neutrophil count". Consequently, there is no evidence of departure from symmetry for the right-hand combination in Fig. 3, whereas the left-hand one is certainly skew. We stress that these results only serve for descriptive purposes. Using a false discovery rate controlling procedure, e.g., the Bonferroni correction or the Benjamini–Hochberg procedure, see (Benjamini and Yekutieli 2001), we did not obtain any rejection.

7 Concluding remarks

Like with many other goodness-of-fit tests, also the HKM test for reflective symmetry contains an element of arbitrariness. In this case, it is the parameter a. Although the test is consistent for each fixed $a > 0$, power may sometimes depend heavily on a. From the results of a small simulation study given in Henze et al. (2003) for the case $n \in \{20, 40, 60, 80\}$, $d \in \{2, 4, 6\}$ and $a \in \{1, 2, 3, 4\}$, the choice $a = 1$ and $a = 4$ seems to be promising, and it seems that smaller values of a generally do not increase power. On the other hand, small values of a, such as $a = 0.1$ or $a = 0.01$, lead to reliable confidence intervals for Δ_a. In this respect, much more simulation work is needed. As for a "good choice of a", it may also be tempting to let a depend on $X_1, \ldots, X_n$. However, a basic issue with any goodness-of-fit test is that it has a preference for some finite-dimensional space of alternatives, and there is no test which pays equal attention to an infinite number of orthogonal alternatives, see, e.g., Janssen (2000). Thus, also a "data driven test" will have its limitations.

8 Proofs

8.1 Proof of Theorem 1

Putting

$$\widetilde{T}_{n,a} = \int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \sin\left(t^\top X_j\right) \right)^2 \exp\left(-a\|t\|^2\right) \mathrm{d}t,$$

we have

$$\frac{\widetilde{T}_{n,a}}{n} = \left\| \frac{1}{n} \sum_{j=1}^{n} \sin\left(\bullet^\top X_j\right) \right\|_{\mathcal{L}^2}^2,$$

where $\|\cdot\|_{\mathcal{L}^2}$ denotes the norm in $\mathcal{L}^2$. The strong law of large numbers in Banach spaces yields

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{j=1}^{n} \sin\left(\bullet^\top X_j\right) - \mathbb{E}\left[\sin\left(\bullet^\top X\right)\right] \right\|_{\mathcal{L}^2}^2 = 0$$

$\mathbb{P}$-almost surely, whence

$$\lim_{n\to\infty} \frac{\widetilde{T}_{n,a}}{n} = \Delta_a \tag{16}$$

$\mathbb{P}$-almost surely. Since

$$\left| \frac{1}{n} \sum_{j=1}^{n} \sin\left(t^\top X_j\right) + \frac{1}{n} \sum_{j=1}^{n} \sin\left(t^\top Y_{n,j}\right) \right| \le 2,$$

it follows that

$$\frac{|\widetilde{T}_{n,a} - T_{n,a}|}{n} \le 2 \int_{\mathbb{R}^d} \left| \frac{1}{n} \sum_{j=1}^{n} \sin\left(t^\top X_j\right) - \frac{1}{n} \sum_{j=1}^{n} \sin\left(t^\top Y_{n,j}\right) \right| \exp\left(-a\|t\|^2\right) \mathrm{d}t.$$

Putting

$$\Delta_{n,j} = Y_{n,j} - X_j = S_n^{-1/2}(X_j - \overline{X}_n) - X_j, \tag{17}$$

the inequalities $|\sin a - \sin b| \le |a - b|$ and $|t^\top z| \le \|t\| \cdot \|z\|$ give

$$\frac{|\widetilde{T}_{n,a} - T_{n,a}|}{n} \le 2 \left(\frac{1}{n} \sum_{j=1}^{n} \|\Delta_{n,j}\| \right) \int_{\mathbb{R}^d} \|t\| \, \exp(-a\|t\|^2) \, \mathrm{d}t.$$

Writing $\operatorname{tr}(A)$ for the trace of a square matrix A, we have

$$\frac{1}{n}\sum_{j=1}^{n}\|\Delta_{n,j}\|^2 = \operatorname{tr}\left((S_n^{-1/2}-\mathrm{I}_d)^2\frac{1}{n}\sum_{j=1}^{n}X_jX_j^\top\right) \\ -2\overline{X}_n^\top S_n^{-1/2}(S_n^{-1/2}-\mathrm{I}_d)\overline{X}_n+\overline{X}_n^\top S_n^{-1}\overline{X}_n.$$

Since $S_n^{-1/2}\xrightarrow{\text{a.s.}}\mathrm{I}_d$ and $\overline{X}_n\xrightarrow{\text{a.s.}}0$, it follows that

$$\frac{1}{n}\sum_{j=1}^{n}\|\Delta_{n,j}\|^2\xrightarrow{\text{a.s.}}0. \tag{18}$$

In view of (16) and the Cauchy–Schwarz estimate

$$\frac{1}{n}\sum_{j=1}^{n}\|\Delta_{n,j}\|\le\left(\frac{1}{n}\sum_{j=1}^{n}\|\Delta_{n,j}\|^2\right)^{1/2},$$

we have

$$\frac{1}{n}\sum_{j=1}^{n}\|\Delta_{n,j}\|\xrightarrow{\text{a.s.}}0. \tag{19}$$

Invoking (16), the proof is completed. □

8.2 Proof of Theorem 2.

Denote the right-hand side of (10) by $\widetilde{\Delta}_a$. From (3), we have

$$\frac{T_{n,a}}{n}=\frac{(2\pi)^d}{4}\int_{\mathbb{R}^d}\left(\widehat{f}_{n,a}(x)-\widehat{f}_{n,a}(-x)\right)^2\mathrm{d}x.$$

We show $\lim_{n\to\infty}\mathbb{E}[T_{n,a}/n]=\widetilde{\Delta}_a$ and $\lim_{n\to\infty}\mathbb{V}(T_{n,a}/n)=0$, where $\mathbb{V}$ denotes variance. Since a constant stochastic limit is uniquely determined, the assertion follows. Fubini's theorem gives

$$\mathbb{E}\left[\frac{T_{n,a}}{n}\right]=\frac{(2\pi)^d}{4}\int_{\mathbb{R}^d}\mathbb{E}\left[\left(\widehat{f}_{n,a}(x)-\widehat{f}_{n,a}(-x)\right)^2\right]\mathrm{d}x.$$

Using (4) and expanding the round bracket, we obtain

$$\widehat{f}_{n,a}(x)^2=\frac{1}{(2\pi a)^d}\frac{1}{n^2}\sum_{i,j=1}^{n}\exp\left(-\frac{\|x-Y_{n,i}\|^2}{2a}\right)\exp\left(-\frac{\|x-Y_{n,j}\|^2}{2a}\right).$$

Taking expectations, symmetry arguments, the inequality $\exp(-\xi) \leq 1, \xi \geq 0$, almost sure convergence of $Y_{n,j}$ to X_j for fixed j, dominated convergence and independence yield

$$\lim_{n\to\infty} \mathbb{E}\left[\widehat{f}_{n,a}(x)^2\right] = \frac{1}{(2\pi a)^d}\,\mathbb{E}\left[\exp\left(-\frac{\|x-X\|^2}{2a}\right)\right]^2.$$

The other terms are treated similarly, and thus $\lim_{n\to\infty}\mathbb{E}[T_{n,a}/n] = \widetilde{\Delta}_a$. To prove $\lim_{n\to\infty}\ \mathbb{V}(T_{n,a}/n) = 0$, start with

$$\left(\frac{T_{n,a}}{n}\right)^2 = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\widehat{f}_{n,a}(x) - \widehat{f}_{n,a}(-x)\right)^2\left(\widehat{f}_{n,a}(y) - \widehat{f}_{n,a}(-y)\right)^2 \mathrm{d}x\,\mathrm{d}y$$

and use the techniques indicated above to show that $\lim_{n\to\infty}\mathbb{E}[(T_{n,a}/n)^2] = \widetilde{\Delta}_a^2$. Hence, $\lim_{n\to\infty}\mathbb{V}(T_{n,a}/n) = 0$, and the assertion follows. □

8.3 Proof of Theorem 3.

The proof uses the following Abelian theorem for Laplace transforms (see Widder 1959, p. 182): Suppose $g : [0,\infty) \to \mathbb{R}$ is a measurable function that is integrable over compact intervals. Assume further that $\int_0^\infty g(t)\mathrm{e}^{-at}\mathrm{d}t$ is finite for each $a > 0$. If for some $\gamma \geq 0$ and some real constant A

$$\lim_{s\to 0}\frac{\Gamma(\gamma+1)}{s^\gamma}\int_0^s g(t)\mathrm{d}t = A, \tag{20}$$

then

$$\lim_{a\to\infty} a^\gamma \int_0^\infty g(t)\mathrm{e}^{-at}\mathrm{d}t = A. \tag{21}$$

Notice that (20) holds if

$$\lim_{u\to 0}\frac{\Gamma(\gamma)g(u)}{u^{\gamma-1}} = A. \tag{22}$$

We use the above result in the following way: Starting with (9), the fact that $\mathrm{I}(t) = \mathbb{E}[\sin(tX)]$ satisfies $\mathrm{I}(t) = \mathrm{I}(-t)$ and a change of variable yield

$$\Delta_a = \int_0^\infty g(u)\mathrm{e}^{-au}\,\mathrm{d}u,$$

where

$$g(u) = \frac{\mathrm{I}^2(\sqrt{u})}{\sqrt{u}}, \quad u > 0,$$

and $g(0) := 0$. From Taylor's theorem, we have

$$\sin x = x - \frac{\cos(\vartheta x)}{6} \cdot x^3, \quad x \in \mathbb{R},$$

where $\vartheta = \vartheta(x)$ and $|\vartheta| \le 1$. Hence,

$$\sin(tX) = tX - \frac{\cos(\Theta t X)}{6} \cdot t^3 X^3,$$

where Θ is a random variable that depends on tX and satisfies $|\Theta| \le 1$. Since $\mathbb{E}(X) = 0$, Lebesgue's dominated convergence theorem yields

$$\mathrm{I}(t) = \mathbb{E}[\sin(tX)] = -\frac{t^3 \mathbb{E}(X^3)}{6} + o(t^3)$$

as $t \to 0$ and hence

$$g(u) = \frac{u^{5/2}\left(\mathbb{E}[X^3]\right)^2}{36} + o\left(u^{5/2}\right)$$

as $u \to 0$. It follows that (22) holds with $\gamma = 7/2$ and $A = 5\sqrt{\pi}(\mathbb{E}[X^3])^2/96$, as was to be shown. □

8.4 Proof of Theorem 4.

We use Theorem 1 of Baringhaus et al. (2017), with $\mathrm{I}(t)$ corresponding to $z(t)$ in that paper. Putting

$$W_n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\sin\left(t^\top Y_{n,j}\right) - \mathrm{I}(t)\right), \quad t \in \mathbb{R}^d,$$

we will show that $W_n(\cdot) \xrightarrow{\mathcal{D}} W(\cdot)$ in $\mathcal{L}^2$, where $W(\cdot)$ is a centred Gaussian random element of $\mathcal{L}^2$ having covariance kernel $K(s,t)$ figuring in the statement of Theorem 4. Denoting by $\langle\cdot,\cdot\rangle$ the inner product in $\mathcal{L}^2$ and observing that, with $\mathrm{I}(\cdot)$ defined in (8),

$$\sqrt{n}\left(\frac{T_{n,a}}{n} - \Delta_a\right) = 2\langle W_n, \mathrm{I}\rangle + \frac{1}{\sqrt{n}}\|W_n\|^2_{\mathcal{L}^2},$$

the continuous mapping theorem yields $\langle W_n, \mathrm{I}\rangle \xrightarrow{\mathcal{D}} \langle W, \mathrm{I}\rangle$ as well as $\|W_n\|^2_{\mathcal{L}^2} \xrightarrow{\mathcal{D}} \|W\|^2_{\mathcal{L}^2}$, whence

$$\sqrt{n}\left(\frac{T_{n,a}}{n} - \Delta_a\right) \xrightarrow{\mathcal{D}} 2\langle W, \mathrm{I}\rangle.$$

The distribution of $2\langle W, \mathrm{I}\rangle$ is the required normal distribution $\mathrm{N}(0, \sigma_a^2)$. The proof of $W_n(\cdot) \xrightarrow{\mathcal{D}} W(\cdot)$ will only be sketched since it closely parallels the proof of Theorem 3.1 of Henze et al. (2003). Let

$$\overline{W}_n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\sin\left(t^\top X_j\right) + t^\top \Delta_{n,j} \cos(t^\top X_j) - \mathrm{I}(t) \right),$$

$$W_n^*(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\sin\left(t^\top X_j\right) - \mathrm{I}(t) - t^\top \mathrm{R}(t) X_j - \frac{1}{2} t^\top \left(X_j X_j^\top - \mathrm{I}_d\right) \mathrm{C}(t) \right),$$

where $\Delta_{n,j}$ is given in (17). Since $W_n = (W_n - \overline{W}_n) + (\overline{W}_n - W_n^*) + W_n^*$, the main steps of the proof are to show $\|W_n - \overline{W}_n\|_{\mathcal{L}^2} = o_{\mathbb{P}}(1)$, $\|\overline{W}_n - W_n^*\|_{\mathcal{L}^2} = o_{\mathbb{P}}(1)$ and $W_n^* \xrightarrow{\mathcal{D}} W$ in $\mathcal{L}^2$. The details are omitted. Notice that the convergence $W_n^* \xrightarrow{\mathcal{D}} W$ follows from the Lindeberg–Lévy type central limit theorem in separable Hilbert spaces (see, e.g., Bosq 2000), since the summands comprising W_n^* are i.i.d. centred random elements of $\mathcal{L}^2$. □

8.5 Proof of Proposition 1.

In what follows, ρ_j is shorthand for $\rho_{j,a}$, $j \in \{1, 2\}$. Starting with (13), the proof follows from straightforward but tedious calculations and symmetry arguments using

$$\iint \frac{1}{n} \sum_j \sin(s^\top Y_j) \sin(t^\top Y_j)\, \mathrm{I}_n(s) \mathrm{I}_n(t)\, w_a(s,t)\, \mathrm{d}s\mathrm{d}t$$
$$= \frac{1}{n^3} \sum_{j,k,\ell} \rho_1(Y_j, Y_k) \rho_1(Y_j, Y_\ell) = \frac{1}{n} \sum_j \overline{V}_{n,1}(Y_j)^2,$$

$$\iint \mathrm{I}_n^2(s) \mathrm{I}_n^2(t) w_a(s,t)\, \mathrm{d}s\mathrm{d}t = \left(\frac{1}{n^2} \sum_{j,k} \rho_1(Y_j, Y_k) \right)^2 = V_{n,1}^2,$$

$$\iint \mathrm{R}_n(t) t^\top \mathrm{S}_n(s) \mathrm{I}_n(s) \mathrm{I}_n(t) w_a(s,t)\, \mathrm{d}s\mathrm{d}t$$
$$= \left(\frac{1}{n^2} \sum_{j,k} \rho_1(Y_j, Y_k) Y_j \right)^\top \left(\frac{1}{n^2} \sum_{i,\ell} \rho_2(Y_i, Y_\ell) \right) = \left(\frac{1}{n} \sum_j \overline{V}_{n,1}(Y_j) Y_j \right)^\top V_{n,2},$$

$$\iint \mathrm{R}_n(s) \mathrm{R}_n(t) s^\top t\, \mathrm{I}_n(s) \mathrm{I}_n(t) w_a(s,t)\, \mathrm{d}s\mathrm{d}t = \left\| \frac{1}{n^2} \sum_{i,k} \rho_2(Y_i, Y_k) \right\|^2 = \|V_{n,2}\|^2,$$

$$\iint t^\top \frac{1}{n} \sum_i \sin(s^\top Y_i) Y_i Y_i^\top \mathrm{C}_n(t) \mathrm{I}_n(s) \mathrm{I}_n(t)\, w_a(s,t)\, \mathrm{d}s\mathrm{d}t$$

$$= \frac{1}{n^4} \sum_{i,j,k,\ell} \rho_1(Y_i, Y_k) Y_i^\top Y_j Y_i^\top \rho_2(Y_j, Y_\ell) = \operatorname{tr}\left(\Sigma_n \left(\frac{1}{n} \sum_j \overline{V}_{n,2}(Y_j) Y_j^\top \right)\right),$$

$$\iint \mathrm{I}_n(s) t^\top \mathrm{C}_n(t) \mathrm{I}_n(s) \mathrm{I}_n(t) w_a(s,t)\, \mathrm{d}s\mathrm{d}t$$

$$= \left(\frac{1}{n^2} \sum_{i,k} \rho_1(Y_i, Y_k) \right) \left(\frac{1}{n^2} \sum_{j,\ell} Y_j^\top \rho_2(Y_j, Y_\ell) \right) = V_{n,1} \frac{1}{n} \sum_j Y_j^\top \overline{V}_{n,2}(Y_j),$$

$$\iint s^\top \mathrm{R}_n(s) \frac{1}{n} \sum_j Y_j t^\top Y_j Y_j^\top \mathrm{C}_n(t) \mathrm{I}_n(s) \mathrm{I}_n(t)\, w_a(s,t)\, \mathrm{d}s\mathrm{d}t$$

$$= \left(\frac{1}{n^2} \sum_{i,\ell} \rho_2(Y_i, Y_\ell) \right)^\top \left(\frac{1}{n^3} \sum_{j,k,m} Y_j Y_j^\top Y_k Y_j^\top \rho_2(Y_k, Y_m) \right)$$

$$= V_{n,2}^\top \left(\frac{1}{n^2} \sum_{j,k} Y_j Y_j^\top Y_k Y_j^\top \overline{V}_{n,2}(Y_k) \right),$$

$$\iint \mathrm{C}_n(s)^\top \frac{1}{n} \sum_j Y_j Y_j^\top s t^\top Y_j Y_j^\top \mathrm{C}_n(t)\, \mathrm{I}_n(s) \mathrm{I}_n(t)\, w_a(s,t)\, \mathrm{d}s\mathrm{d}t$$

$$= \frac{1}{n} \sum_j \left(Y_j^\top \Gamma_n Y_j \right)^2,$$

and

$$\iint s^\top \mathrm{C}_n(s) t^\top \mathrm{C}_n(t)\, \mathrm{I}_n(s) \mathrm{I}_n(t)\, w_a(s,t)\, \mathrm{d}s\mathrm{d}t = \left(\frac{1}{n^2} \sum_{i,k} Y_i^\top \rho_2(Y_i, Y_k) \right)^2.$$

Here, summation is from 1 to n for each of the indices, and each integral is over $\mathbb{R}^d$. □

8.6 Proof of Theorem 5.

The first observation is the following: Again suppressing the dependence on a, put

$$\widehat{\sigma}_{n,0}^2 = 4 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_n^0(s,t)\, \mathrm{I}_n^0(s) \mathrm{I}_n^0(t)\, w_a(s,t)\, \mathrm{d}s\mathrm{d}t,$$

where $K_n^0(s,t)$ originates from $K(s,t)$ by replacing the functions $\mathrm{R}(\cdot)$, $\mathrm{I}(\cdot)$, $\mathrm{C}(\cdot)$ and $\mathrm{S}(\cdot)$ with their respective "estimator-free" empirical counterparts

$$\mathrm{R}_n^0(t) = \frac{1}{n} \sum_{j=1}^n \cos\left(t^\top X_j\right), \quad \mathrm{I}_n^0(t) = \frac{1}{n} \sum_{j=1}^n \sin\left(t^\top X_j\right),$$

$$\mathrm{C}_n^0(t) = \frac{1}{n}\sum_{j=1}^{n} X_j \cos\left(t^\top X_j\right), \quad \mathrm{S}_n^0(t) = \frac{1}{n}\sum_{j=1}^{n} X_j \sin\left(t^\top X_j\right), \quad t \in \mathbb{R}^d,$$

and do the same with each of the five explicitly designated expectations figuring in the definition of $K(s,t)$. Hence, $\mathbb{E}\left[\sin\left(s^\top X\right)\sin\left(t^\top X\right)\right]$ is replaced with $n^{-1}\sum_{j=1}^{n}\sin(s^\top X_j)\sin(t^\top X_j)$ etc. It is then straightforward to see that

$$\widehat{\sigma}_{n,0}^2 \xrightarrow{\mathbb{P}} \sigma_a^2. \tag{23}$$

For example, apart from the factor 4, the contribution of the first summand of the representation of $K(s,t)$ to σ_a^2 is

$$J := \iint \mathbb{E}\left[\sin(s^\top X)\sin(t^\top X)\right] \mathrm{I}(s)\mathrm{I}(t)\, w_a(s,t)\,\mathrm{d}s\mathrm{d}t$$

(say). For the empirical version

$$J_n = \iint \frac{1}{n}\sum_i \sin(s^\top X_i)\sin(t^\top X_i)\, \mathrm{I}_n^0(s)\mathrm{I}_n^0(t)\, w_a(s,t)\,\mathrm{d}s\mathrm{d}t$$

(say), Fubini's theorem gives

$$\mathbb{E}(J_n) = \frac{1}{n^3}\sum_{i,j,k} \iint \mathbb{E}\left[\sin(s^\top X_i)\sin(t^\top X_i)\sin(s^\top X_j)\sin(t^\top X_k)\right] w_a(s,t)\,\mathrm{d}s\mathrm{d}t.$$

If all indices are different, then, by symmetry and independence, the expectation beneath the integral sign is $\mathbb{E}[\sin(s^\top X)\sin(t^\top X)]\mathrm{I}(s)\mathrm{I}(t)$. Since the case that at least two of the three indices coincide are asymptotically negligible, we have $\lim_{n\to\infty}\mathbb{E}(J_n) = J$. Likewise, $\lim_{n\to\infty}\mathbb{V}(J_n) = 0$ and thus $J_n \xrightarrow{\mathbb{P}} J$. Since the other terms can be treated similarly, (23) follows.

The much more difficult part of the proof is to show

$$\widehat{\sigma}_{n,a}^2 - \widehat{\sigma}_{n,0}^2 \xrightarrow{\mathbb{P}} 0. \tag{24}$$

In view of the definitions of $\widehat{\sigma}_{n,a}^2$ and $\widehat{\sigma}_{n,0}^2$, this boils down to prove

$$\iint \left(K_n(s,t)\mathrm{I}_n(s)\mathrm{I}_n(s) - K_n^0(s,t)\mathrm{I}_n^0(s)\mathrm{I}_n^0(s)\right) w_a(s,t)\,\mathrm{d}s\mathrm{d}t \xrightarrow{\mathbb{P}} 0.$$

To this end, we have to consider each term of the various summands comprising $K_n(s,t)$ and compare this with the corresponding term in $K_n^0(s,t)$. As an example, we choose the empirical versions of the first summand of $K(s,t)$ that involves moments of X which, apart form the minus sign and the factor 1/2, is $t^\top \mathbb{E}[\sin(s^\top X)XX^\top]$. Putting

$$L_n(s,t) = t^\top \frac{1}{n^4} \sum_{j,k,\ell,m} \sin(s^\top Y_j)Y_jY_j^\top Y_k \cos(t^\top Y_k) \sin(s^\top Y_\ell) \sin(t^\top Y_m),$$
$$L_n^0(s,t) = t^\top \frac{1}{n^4} \sum_{j,k,\ell,m} \sin(s^\top X_j)X_jX_j^\top X_k \cos(t^\top X_k) \sin(s^\top X_\ell) \sin(t^\top X_m),$$

we have to prove

$$\iint \big(L_n(s,t) - L_n^0(s,t)\big) w_a(s,t)\,\mathrm{d}s\mathrm{d}t \xrightarrow{\mathbb{P}} 0.$$

Notice that

$$\iint L_n(s,t)\, w_a(s,t)\,\mathrm{d}s\mathrm{d}t = \frac{1}{n^4} \sum_{j,k,\ell,m} D(Y_j, Y_\ell)\, E(Y_j, Y_k, Y_m),$$

where

$$D(Y_j, Y_\ell) = \int \sin(t^\top Y_j) \sin(s^\top Y_\ell) \exp(-a\|s\|^2)\,\mathrm{d}s,$$
$$E(Y_j, Y_k, Y_m) = \int t^\top Y_j Y_j^\top Y_k \cos(t^\top Y_k) \sin(t^\top Y_m) \exp(-a\|t\|^2)\,\mathrm{d}t.$$

Likewise,

$$\iint L_n^0(s,t)\, w_a(s,t)\,\mathrm{d}s\mathrm{d}t = \frac{1}{n^4} \sum_{j,k,\ell,m} D(X_j, X_\ell)\, E(X_j, X_k, X_m),$$

where

$$D(X_j, X_\ell) = \int \sin(t^\top X_j) \sin(s^\top X_\ell) \exp(-a\|s\|^2)\,\mathrm{d}s,$$
$$E(X_j, X_k, X_m) = \int t^\top X_j X_j^\top X_k \cos(t^\top X_k) \sin(t^\top X_m) \exp(-a\|t\|^2)\,\mathrm{d}t.$$

It follows that

$$\iint \big(L_n(s,t) - L_n^0(s,t)\big) w_a(s,t)\,\mathrm{d}s\mathrm{d}t = K_{n,1} + K_{n,2},$$

where

$$K_{n,1} = \frac{1}{n^4} \sum_{j,k,\ell,m} \Big(E(Y_j, Y_k, Y_m) - E(X_j, X_k, X_m) \Big) D(Y_j, Y_\ell),$$
$$K_{n,2} = \frac{1}{n^4} \sum_{j,k,\ell,m} E(X_j, X_k, X_m) \Big(D(Y_j, Y_\ell) - D(X_j, X_\ell) \Big).$$

We first prove $K_{n,2} \xrightarrow{\mathbb{P}} 0$. Putting

$$c_\nu := \int \|t\|^\nu \exp(-a\|t\|^2)\, \mathrm{d}t, \quad \nu \in \{1, 2\},$$

the fact that $|\cos(t^\top X_k)\sin(t^\top X_m)| \le 1$ and the Cauchy–Schwarz inequality yields

$$|E(X_j, X_k, X_m)| \le c_1 \|X_j\|^2 \|X_k\|.$$

Since $\sin(t^\top Y_j) = \sin(t^\top X_j) + \xi_j t^\top \Delta_j$, where $|\xi_j| \le 1$ (and likewise for $\sin(t^\top Y_\ell)$), the Cauchy–Schwarz inequality gives

$$\begin{aligned} |D(Y_j, Y_\ell) - D(X_j, X_\ell)| &\le \int \Big(\|t\|(\|\Delta_j\| + \|\Delta_\ell\|) + \|t\|^2 \|\Delta_j\| \|\Delta_\ell\| \Big) e^{-a\|t\|^2} \mathrm{d}t \\ &= c_1 \Big(\|\Delta_j\| + \|\Delta_\ell\| \Big) + c_2 \|\Delta_j\| \|\Delta_\ell\|. \end{aligned}$$

We therefore have

$$|K_{n,2}| \le \frac{1}{n^4} \sum_{j,k,\ell,m} c_1 \|X_j\|^2 \|X_k\| \Big(c_1(\|\Delta_j\| + \|\Delta_\ell\|) + c_2 \|\Delta_j\| \|\Delta_\ell\| \Big).$$

Since $n^{-1} \sum_{j=1}^n \|X_j\|^\nu = O_{\mathbb{P}}(1)$ if $\nu \in \{1, 2, 3, 4\}$ (recall the assumption $\mathbb{E}\|X\|^4 < \infty$) and

$$\frac{1}{n} \sum_{j=1}^n \|X_j\|^2 \|\Delta_j\| \le \left(\frac{1}{n} \sum_{j=1}^n \|X_j\|^4 \cdot \frac{1}{n} \sum_{j=1}^n \|\Delta_j\|^2 \right)^{1/2},$$

$K_{n,2} \xrightarrow{\mathbb{P}} 0$ follows from (18) and (19).

As for $K_{n,1}$, first notice that $|D(Y_j, Y_\ell)| \le c_1$ and thus

$$|K_{n,1}| \le \frac{c_1}{n^4} \sum_{j,k,\ell,m} |E(Y_j, Y_k, Y_m) - E(X_j, X_k, X_m)|.$$

Next, we have

$$Y_j Y_j^\top Y_k = X_j X_j^\top X_k + \widetilde{\Delta}_{j,k},$$

where

$$\widetilde{\Delta}_{j,k} = X_j \Delta_j^\top X_k + X_j X_j^\top \Delta_k + X_j \Delta_j^\top \Delta_k + \Delta_j X_j^\top X_k \\ + \Delta_j \Delta_j^\top X_k + \Delta_j X_j^\top \Delta_k + \Delta_j \Delta_j^\top \Delta_k.$$

Therefore,

$$\begin{aligned} &E(Y_j, Y_k, Y_m) - E(X_j, X_k, X_m) \\ &\quad = \int t^\top X_j X_j^\top X_k \Big(\cos(t^\top Y_k) \sin(t^\top Y_m) - \cos(t^\top X_k) \sin(t^\top X_m) \Big) e^{-a\|t\|^2} \mathrm{d}t \\ &\quad + \int t^\top \widetilde{\Delta}_{j,k} \cos(t^\top Y_k) \sin(t^\top Y_m) e^{-a\|t\|^2} \mathrm{d}t. \end{aligned} \tag{25}$$

Since

$$\|\widetilde{\Delta}_{j,k}\| \le \|X_j\|\|\Delta_j\|\|X_k\| + \|X_j\|^2\|\Delta_k\| + \|X_j\|\|\Delta_j\|\Delta_k\| + \|\Delta_j\|\|X_j\|\|X_k\| \\ + \|\Delta_j\|^2\|X_k\| + \|\Delta_j\|\|X_j\|\|\Delta_k\| + \|\Delta_j\|^2\|\Delta_k\|,$$

the inequality $|\cos(t^\top Y_k)\sin(t^\top Y_m)| \le 1$ and the same reasoning as above show that

$$\frac{1}{n^4} \sum_{j,k,\ell,m} \left| \int t^\top \widetilde{\Delta}_{j,k} \cos(t^\top Y_k) \sin(t^\top Y_m) e^{-a\|t\|^2} \mathrm{d}t \right| = o_{\mathbb{P}}(1).$$

Regarding the term figuring in (25), we have

$$|\cos(t^\top Y_k)\sin(t^\top Y_m) - \cos(t^\top X_k)\sin(t^\top X_m)| \\ \le \|t\|(\|\Delta_k\| + \|\Delta_m\|) + \|t\|^2\|\Delta_k\|\|\Delta_m\|,$$

and it follows by the same reasoning as above that

$$\frac{1}{n^4} \sum_{j,k,\ell,m} \left| \int t^\top X_j X_j^\top X_k \Big(\cos(t^\top Y_k) \sin(t^\top Y_m) - \cos(t^\top X_k) \sin(t^\top X_m) \Big) e^{-a\|t\|^2} \mathrm{d}t \right|$$

is asymptotically negligible. Consequently, $K_{n,1} = o_{\mathbb{P}}(1)$. Since all the other summands comprising K_n and K_n^0 can be tackled in the same way, (24) follows. □

Acknowledgements The authors wish to thank two anonymous referees and a member of the Editorial Board for their careful reading of the manuscript and for many helpful suggestions.

References

Aki, S. (1993). On nonparametric tests for symmetry in $\mathbb{R}^m$. *Annals of the Institute of Statistical Mathematics*, *45*, 787–800.

Baringhaus, L., Henze, N. (1992). Limit distributions for Mardia's measure of multivariate skewness. *Annals of Statistics*, *20*, 1889–1902.

Baringhaus, L., Ebner, B., Henze, N. (2017). The limit distribution of weighted L^2-goodness-of-fit statistics under fixed alternatives, with applications. *Annals of the Institute of Statistical Mathematics*, *69*, 969–995.

Batsidis, A., Martin, N., Pardo, L., Zografos, K. (2014). A necessary power divergence-type family of tests for testing elliptical symmetry. *Journal of Statistical Computation and Simulation*, *84*, 57–83.

Benjamini, Y., Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency. *Annals of Statistics*, *29*, 1165–1188.

Bosq, D. (2000). *Linear processes in function spaces*. New York: Springer.

Bowman, A. W., Foster, P. J. (1993). Adaptive smoothing and density-based tests of multivariate normality. *Journal of the American Statistical Association*, *88*, 529–537.

Dyckerhoff, R., Ley, C., Paindaveine, D. (2015). Depth-based run tests for bivariate central symmetry. *Annals of the Institute of Statistical Mathematics*, *67*, 917–941.

Eaton, M. L., Perlman, M. D. (1973). The non-singularity of generalized sample covariance matrices. *Annals of Statistics*, *1*, 710–717.

Einmahl, J. H. J., Gan, Z. (2016). Testing for central symmetry. *Journal of Statistical Planning and Inference*, *169*, 27–33.

Gürtler, N. (2000). Asymptotic results on the BHEP tests for multivariate normality with fixed and variable smoothing parameter (in German). Doctoral dissertation. University of Karlsruhe.

Heathcote, C. R., Rachev, S. T., Cheng, B. (1995). Testing multivariate symmetry. *Journal of Multivariate Analysis*, *54*, 91–112.

Henze, N. (1997a). Limit laws for multivariate skewness in the sense of Móri, Rohatgi and Székely. *Statistics & Probability Letters*, *33*, 299–307.

Henze, N. (1997b). Extreme smoothing and testing for multivariate normality. *Statistics & Probability Letters*, *35*, 203–213.

Henze, N., Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *Journal of Multivariate Analysis*, *62*, 1–23.

Henze, N. (2002). Invariant tests for multivariate normality: a critical review. *Statistical Papers*, *43*, 467–506.

Henze, N., Klar, B., Meintanis, S. (2003). Invariant tests for symmetry about an unspecified point based on the empirical characteristic function. *Journal of Multivariate Analysis*, *87*, 275–297.

Janssen, A. (2000). Global power functions of goodness of fit tests. *Annals of Statistics*, *28*, 239–253.

Ley, C., Verdebout, T. (2014). Simple optimal tests for circular reflective symmetry about a specified median direction. *Statistica Sinica*, *24*, 1319–1339.

Madhava Rao, K. S., Raghunath, M. (2012). A simple nonparametric test for bivariate symmetry about a line. *Journal of Statistical Planning and Inference*, *142*, 430–444.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530.

Meintanis, S., Ngatchou-Wandji, J. (2012). Recent tests for symmetry with multivariate and structured data: a review. Nonparametric statistical methods and related topics, 35–73, World Scientific Publishing, Hackensack, NJ.

Móri, T. F., Rohatgi, V. K., Székely, G. F. (1993). On multivariate skewness and kurtosis. *Theory of Probability and its Applications*, *38*(1), 547–551.

Neuhaus, G., Zhu, L.-X. (1998). Permutation Tests for Reflected Symmetry. *Journal of Multivariate Analysis*, *67*, 129–153.

Ngatchou-Wandji, J. (2009). Testing for symmetry in multivariate distributions. *Statistical Methodology*, *6*, 230–250.

Partlett, C., Prakash, P. (2017). Measuring asymmetry and testing symmetry. *Annals of the Institute of Statistical Mathematics*, *96*(2), 429–460.

Quessy, J.-F. (2016). On consistent nonparametric statistical tests of symmetry hypotheses. Symmetry 8, No.5, Art. 31, 19 pp.

Royston, J. P. (1983). Some techniques for assessing multivariate normality based on the Shapiro-Wilk W. *Applied Statistics*, *32*, 121–133.

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

Serfling, R. (2006). Multivariate symmetry and asymmetry. In S. Kotz, N. Balakrishnan, C. B. Read & B. Vidacovic (Eds.), *Encyclopedia of statistical sciences*. 2nd ed., Vol. 8, pp. 5338–5345. New York: Wiley.

Székely, G. J., Sen, P. K. (2002). Characterization of diagonal symmetry: location unknown, and a test based on allied U-processes. *Journal of Statistical Planning and Inference*, *102*, 349–358.

Widder, D. V. (1959). *The Laplace transform, 5th printing*. Princeton: Princeton University Press.