INVARIANT TESTS FOR SYMMETRY ABOUT AN UNSPECIFIED POINT BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION [1] [2]

N. Henze[3], B. Klar, and S. G. Meintanis

Institut für Mathematische Stochastik, Universität Karlsruhe,
Englerstr. 2, 76128 Karlsruhe, Germany
and
Department of Engineering Sciences, University of Patras,
261 10 Patras, Greece

Abstract. This paper considers a flexible class of omnibus affine invariant tests for the hypothesis that a multivariate distribution is symmetric about an unspecified point. The test statistics are weighted integrals involving the imaginary part of the empirical characteristic function of suitably standardized given data, and they have an alternative representation in terms of an L^2-distance of nonparametric kernel density estimators. Moreover, there is a connection with two measures of multivariate skewness. The tests are performed via a permutational procedure that conditions on the data.

Keywords. Test for symmetry, affine invariance, Mardia's measure of multivariate skewness, skewness in the sense of Móri, Rohatgi and Székely, empirical characteristic function, permutational limit theorem.

---

[1] AMS 1991 subject classification numbers: 62G10, 62G20

[2] Proposed running head: INVARIANT TESTS FOR SYMMETRY

[3] Corresponding author

1 Introduction and summary

Suppose $X_1, \ldots, X_n, \ldots$ is a sequence of independent copies of a random d-dimensional column vector X. Writing "$\sim$" for equality in distribution, we consider the problem of testing the hypothesis

$$(1.1) \qquad H_0 : X - \mu \sim \mu - X \text{ for some } \mu \in \mathbb{R}^d$$

of symmetry about some unknown center, against general alternatives. This problem has been a topic of intensive research in the univariate case $d = 1$ (see e.g., Aki (1981), Alemayehu, Giné, and Pena (1993), Antille, Kersting, and Zucchini (1982), Bhattacharya, Gastwirth, and Wright (1982), Boos (1982), Csörgő and Heathcote (1982), Csörgő and Heathcote (1987), Doksum, Fenstad, and Aaberge (1977), Hollander (1988), Koutrouvelis (1985), Koziol (1985), and Schuster and Barker (1987)). In the multivariate case $d > 1$, the testing problem (1.1) is known as testing for *reflected* symmetry, in order to distinguish it from the more special problem of testing for *spherical* symmetry (see, e.g. Baringhaus (1991), Gupta and Kabe (1993), Kariya and Eaton (1977), King (1980), Koltchinskii and Li (1998), Zhu, Fang, and Zhang (1995) and Zhu, Fang, and Li (1997)).

Our approach of tackling (1.1) is similar to that of Heathcote, Rachev, and Cheng (1995) and Neuhaus and Zhu (1998). Unlike these papers, however, we stress the hitherto neglected aspect of *affine invariance*. To put this issue into perspective, notice that the testing problem under discussion is invariant not only with respect to translations, but more generally with respect to transformations of the kind $x \mapsto Ax + b$, $x \in \mathbb{R}^d$, where A is a nonsingular $(d \times d)$-matrix and $b \in \mathbb{R}^d$. Consequently, a decision in favor or against H_0 should be the same for $X_1, \ldots, X_n$ and $AX_1 + b, \ldots, AX_n + b$. This goal is achieved if the test statistic T_n, say, has the property

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

for each nonsingular $(d \times d)$-matrix A and any $b \in \mathbb{R}^d$. To this end, define the standardized data

$$(1.2) \qquad Y_j = S_n^{-1/2}(X_j - \bar{X}_n), \quad j = 1, \ldots, n,$$

where $\bar{X}_n = n^{-1} \sum_{j=1}^{n} X_j$ denotes the sample mean, $S_n^{-1/2}$ is the symmetric square root of the inverse of the sample covariance matrix

$$S_n \;=\; \frac{1}{n}\sum_{j=1}^{n}(X_j-\bar{X}_n)(X_j-\bar{X}_n)',$$

and the prime stands for transpose. We assume that S_n is nonsingular with probability one. This condition holds, e.g., if $n > d$ and the distribution of X puts mass zero on each $(d-1)$-dimensional hyperplane (see e.g. Eaton and Perlman (1973)).

Notice that the distribution of X is symmetric about μ if, and only if, the imaginary part of the characteristic function (c.f.) of $X-\mu$ vanishes, i.e., if

$$E[\sin(t'(X-\mu))] \;=\; 0 \quad \text{for each} \quad t\in\mathbb{R}^d.$$

This fact was the starting point of many papers on testing for symmetry (Csörgő and Heathcote (1982), Csörgő and Heathcote (1987), Feuerverger and Mureika (1977), Ghosh and Ruymgaart (1992), Heathcote et al. (1995), Koutrouvelis (1985), Neuhaus and Zhu (1998)).

In the spirit of the class of BHEP tests for multivariate normality (Henze and Wagner, 1997), our test statistic is

$$\text{(1.3)} \qquad T_{n,a} \;=\; \int_{\mathbf{R}^d}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sin(t'Y_j)\right)^2 \exp(-a\|t\|^2)\;dt,$$

where a is some positive constant. In view of

$$\int_{\mathbf{R}^d}\cos(t'c)\exp(-a\|t\|^2)\;dt \;=\; \left(\frac{\pi}{a}\right)^{d/2}\exp\left(-\frac{\|c\|^2}{4a}\right),$$

the trigonometric identity $\sin u \sin v \;=\; (\cos(u-v)-\cos(u+v))\,/2$ yields

$$\text{(1.4)} \qquad T_{n,a} \;=\; \frac{\pi^{d/2}}{a^{d/2}2n}\sum_{j,k=1}^{n}\left[\exp\left(-\frac{1}{4a}\|Y_j-Y_k\|^2\right)-\exp\left(-\frac{1}{4a}\|Y_j+Y_k\|^2\right)\right],$$

which shows that a computer routine implementing $T_{n,a}$ is readily available. Since

$$\|Y_j-Y_k\|^2=(X_j-X_k)'S_n^{-1}(X_j-X_k),$$

$$\|Y_j+Y_k\|^2=(X_j-\bar{X}_n+X_k-\bar{X}_n)'S_n^{-1}(X_j-\bar{X}_n+X_k-\bar{X}_n),$$

the statistic $T_{n,a}$ is affine invariant. Moreover, not even the square root $S_n^{-1/2}$ of S_n^{-1} is needed.

The introduction of the parameter a in the definition of $T_{n,a}$ allows for some flexibility regarding the power of a test for symmetry that rejects H_0 for large values of $T_{n,a}$. In Section 2, it will be seen that $T_{n,a}$ has an alternative representation in terms of an L^2-distance between two nonparametric kernel density estimators. Moreover, $T_{n,a}$ is related to a linear combination of two measures of multivariate skewness as $a \to \infty$. Section 3 gives theoretical results on the limit behavior of $T_{n,a}$ under H_0 and under contiguous alternatives to symmetry as $n \to \infty$. Since the limit distribution of $T_{n,a}$ under H_0 depends on the unknown underlying distribution, some extra randomization is necessary in order to obtain an asymptotically distribution-free procedure. To this end, a permutational limit theorem for $T_{n,a}$ is given in Section 4. In Section 5, we prove the consistency of the test against general alternatives. The paper concludes with the results of a Monte Carlo study.

2 Discussion of the weight function $\exp(-a\|t\|^2)$

This section sheds some light on the role of the weight function $\exp(-\|t\|^2)$ figuring in (1.3). Our first result shows that $T_{n,a}$ has an alternative representation in terms of an L^2-distance between two nonparametric kernel density estimators.

Proposition 2.1 *Let*

$$\hat{f}_n(x) \;=\; \frac{1}{n}\sum_{j=1}^{n} \frac{1}{(2\pi a)^{d/2}} \exp\left(-\frac{\|x - Y_j\|^2}{2a}\right). \tag{2.1}$$

Then

$$T_{n,a} \;=\; \frac{n(2\pi)^d}{4}\int_{\mathbf{R}^d} \left(\hat{f}_n(x) - \hat{f}_n(-x)\right)^2 dx.$$

PROOF. Let $L^2(\mathbb{R}^d)$ denote the Hilbert space of measurable complex-valued functions on $\mathbb{R}^d$ that are square integrable with respect to Lebesgue measure. The Fourier transform

$$\tilde{u}(x) = \int_{\mathbf{R}^d} \exp(ix't)u(t)\,dt$$

of $u \in L^2(\mathbb{R}^d)$ belongs to $L^2(\mathbb{R}^d)$ and, by Plancherel's theorem,

$$\int_{\mathbf{R}^d} |\tilde{u}(x)|^2\,dx \;=\; (2\pi)^d \int_{\mathbf{R}^d} |u(t)|^2\,dt. \tag{2.2}$$

Now, the key observation is the equation

$$\left(\frac{1}{n}\sum_{j=1}^{n}\sin(t'Y_j)\right)^2 \exp\left(-a\|t\|^2\right) = \frac{1}{4}\left|\frac{1}{n}\sum_{j=1}^{n}\exp\left(it'Y_j - \frac{a}{2}\|t\|^2\right) - \right.$$
$$\left. -\frac{1}{n}\sum_{j=1}^{n}\exp\left(-it'Y_j - \frac{a}{2}\|t\|^2\right)\right|^2.$$

Write $\mathcal{P}_n$ for the empirical distribution of $Y_1, \ldots, Y_n$, and let $\mathcal{Q}_n$ be the empirical distribution of $-Y_1, \ldots, -Y_n$. The function $n^{-1}\sum_{j=1}^{n}\exp\left(it'Y_j - a\|t\|^2/2\right)$ is the Fourier transform of the convolution $\mathcal{P}_n * \mathcal{N}(0, aI_d)$, and $n^{-1}\sum_{j=1}^{n}\exp\left(-it'Y_j - a\|t\|^2/2\right)$ is the Fourier transform of the convolution $\mathcal{Q}_n * \mathcal{N}(0, aI_d)$. Since $\mathcal{P}_n * \mathcal{N}(0, aI_d)$ and $\mathcal{Q}_n * \mathcal{N}(0, aI_d)$ have densities $\hat{f}_n(x)$ and $\hat{f}_n(-x)$, respectively, the assertion follows from (2.2). ■

Notice that $\hat{f}_n(x)$ figuring in (2.1) 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 $Y_1, \ldots, Y_n$, and that $\hat{f}_n(-x)$ is the same density estimator, applied to the data after reflection at the origin. Thus, the role of a is that of a smoothing parameter. However, whereas density estimators let the bandwidth depend on the sample size, we keep a fixed in what follows in order to achieve positive asymptotic power against alternatives that approach the null hypothesis at the rate $n^{-1/2}$ (see Section 3). A similar observation was made in connection with the class of BHEP tests for multivariate normality (see Henze and Wagner (1997) and Gürtler (2000)).

We close this section by revealing a peculiar connection between $T_{n,a}$ and two measures of multivariate skewness. The first measure, introduced by Mardia (1970), is

$$b_{1,d} = \frac{1}{n^2}\sum_{j,k=1}^{n}(Y_j'Y_k)^3.$$

The second measure, which was proposed by Móri, Rohatgi, and Székely (1993) and studied further by Henze (1997b), is

$$\tilde{b}_{1,d} = \frac{1}{n^2}\sum_{j,k=1}^{n} Y_j'Y_k\|Y_j\|^2\|Y_k\|^2.$$

Notice that both $b_{1,d}$ and $\tilde{b}_{1,d}$ reduce to squared (Pearson) sample skewness in the univariate case. The following result follows by straightforward algebra.

Proposition 2.2 *We have*

$$\lim_{a\to\infty} \frac{96}{n\pi^{d/2}}\ a^{3+d/2}\ T_{n,a} \ = \ 2b_{1,d} \ + \ 3\tilde{b}_{1,d}.$$

Thus, for large values of 'the bandwidth' a, $T_{n,a}$ is approximately a weighted sum of $b_{1,d}$ and $\tilde{b}_{1,d}$. Interestingly, apart from a factor, the same weights appear in the context of testing for multivariate normality, when forming the 'limit' of the BHEP class of test statistics (Henze, 1997a). For further examples of 'limit' statistics connected with weighted L^2-type test statistics based on empirical transforms, see Baringhaus, Gürtler, and Henze (2000).

3 Asymptotic distribution theory

We first study the limit distribution of $T_{n,a}$ under H_0. To this end, the distribution of X is supposed to be symmetric about some value. In view of affine invariance, we assume $E[X] = 0$ and $E[XX'] = I_d$, the identity matrix of order d. We make the further assumption $E\|X\|^4 < \infty$.

To prove the convergence in distribution of $T_{n,a}$ under H_0, a convenient setting is the separable Hilbert space $\mathcal{L}^2$ of measurable real-valued functions on $\mathbb{R}^d$ that are square integrable with respect to the measure $\exp(-a\|t\|^2)dt$. The norm in $\mathcal{L}^2$ will be denoted by

$$\|h\|_{\mathcal{L}^2} \ = \ \left(\int_{\mathbf{R}^d} h(t)^2 \exp(-a\|t\|^2)dt\right)^{1/2}.$$

The notation $\xrightarrow{\mathcal{D}}$ means weak convergence of random elements of $\mathcal{L}^2$ and random variables, and $O_P(1)$ stands for a sequence of random variables that is bounded in probability. Likewise, $o_P(1)$ is a sequence of random variables that converges to 0 in probability.

Theorem 3.1 *Let $E[X] = 0$, $E[XX'] = I_d$, $E\|X\|^4 < \infty$, and suppose the distribution of X is symmetric, i.e. $\Phi(t) = E[\cos(t'X)], t \in \mathbb{R}^d$, where $\Phi(\cdot)$ is the characteristic function of X. Furthermore, let*

$$\mathcal{W}_n(t) \ = \ \frac{1}{\sqrt{n}}\sum_{j=1}^{n} \sin(t'Y_j), \quad t \in \mathbb{R}^d,$$

where $Y_1, \ldots, Y_n$ are given in (1.2). Then there exists a centered Gaussian process $\mathcal{W}(\cdot)$ on $\mathcal{L}^2$ having covariance kernel

$$(3.1) \quad K(s,t) = E[\sin(s'X)\sin(t'X)] - \Phi(t)E[t'X\sin(s'X)] - \Phi(s)E[s'X\sin(t'X)] + s't\,\Phi(s)\Phi(t)$$

such that

$$(3.2) \quad \mathcal{W}_n(\cdot) \xrightarrow{\mathcal{D}} \mathcal{W}(\cdot)$$

and

$$(3.3) \quad T_{n,a} \xrightarrow{\mathcal{D}} \int_{\mathbf{R}^d} \mathcal{W}^2(t)\exp(-a\|t\|^2)dt.$$

PROOF. Since the reasoning is similar to that given in Henze and Wagner (1997), it will only be sketched. Notice that $Y_j = X_j + \Delta_j$, where $\Delta_j = (S_n^{-1/2} - I_d)X_j - S_n^{-1/2}\bar{X}_n$. Define the auxiliary processes

$$\tilde{\mathcal{W}}_n(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\sin(t'X_j) + t'\Delta_j\cos(t'X_j)\right),$$

$$(3.4) \quad \mathcal{W}_n^*(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\sin(t'X_j) - t'\Phi(t)X_j\right), \quad t \in \mathbb{R}^d.$$

We will prove

$$(3.5) \quad \|\mathcal{W}_n(\cdot) - \tilde{\mathcal{W}}_n(\cdot)\|_{\mathcal{L}^2} = o_P(1),$$

$$(3.6) \quad \|\tilde{\mathcal{W}}_n(\cdot) - \mathcal{W}_n^*(\cdot)\|_{\mathcal{L}^2} = o_P(1),$$

and

$$(3.7) \quad \mathcal{W}_n^*(\cdot) \xrightarrow{\mathcal{D}} \mathcal{W}(\cdot),$$

whence (3.2) and (3.3) follow.

To prove (3.5), note that $\sin(t'Y_j) = \sin(t'X_j) + t'\Delta_j\cos(t'X_j) + \varepsilon_{n,j}(t)$, where $|\varepsilon_{n,j}(t)| \le \|t\|^2\|\Delta_j\|^2$. Since $n^{-1/2}\sum_{j=1}^{n}\|\Delta_j\|^2 = o_P(1)$ (cf. p.9 of Henze and Wagner

(1997)), we have $|\mathcal{W}_n(t) - \tilde{\mathcal{W}}_n(t)| \;\leq\; \|t\|^2 o_P(1)$ and thus (3.5). To show (3.6), start with

$$\tilde{\mathcal{W}}_n(t) \;-\; \mathcal{W}_n^*(t) \;=\; A_n(t) - B_n(t) \;-\; C_n(t), \tag{3.8}$$

where

$$A_n(t) \;=\; t'\sqrt{n}(S_n^{-1/2} - I_d)\;\frac{1}{n}\sum_{j=1}^n X_j\cos(t'X_j),$$

$$B_n(t) \;=\; t'\sqrt{n}(S_n^{-1/2} - I_d)\bar{X}_n\;\frac{1}{n}\sum_{j=1}^n \cos(t'X_j),$$

$$C_n(t) \;=\; t'\sqrt{n}\bar{X}_n\left(\frac{1}{n}\sum_{k=1}^n \cos(t'X_k) \;-\; \Phi(t)\right).$$

Notice that $|A_n(t)| \leq O_P(1)\|t\|\|n^{-1}\sum_{j=1}^n X_j\cos(t'X_j)\|$ and thus, apart from a factor that is bounded in probability, an upper bound for $\|A_n\|^2_{\mathcal{L}^2}$ is the V-statistic

$$V_{n,1} \;=\; \frac{1}{n^2}\sum_{j,k=1}^n\int_{\mathbf{R}^d}\|t\|^2 X_j'X_k\cos(t'X_j)\cos(t'X_k)\exp(-a\|t\|^2)dt.$$

Since, by the strong law of large numbers for V-statistics, $V_{n,1}$ tends to zero almost surely (note that $E[X\cos(t'X)] = 0$), we have $\|A_n\|_{\mathcal{L}^2} = o_P(1)$. Furthermore, $|B_n(t)| \leq o_P(1)\|t\|\;|n^{-1}\sum_{j=1}^n\cos(t'X_j)|$ and thus $\|B_n\|^2_{\mathcal{L}^2} \;\leq\; o_P(1)V_{n,2}$, where

$$V_{n,2} \;=\; \frac{1}{n^2}\sum_{j,k=1}^n\int_{\mathbf{R}^d}\|t\|^2\cos(t'X_j)\cos(t'X_k)\exp(-a\|t\|^2)dt.$$

Since $V_{n,2} \longrightarrow \int\Phi(t)^2\|t\|^2\exp(-a\|t\|^2)dt$ almost surely, it follows that $\|B_n\|_{\mathcal{L}^2} = o_P(1)$. Finally, $|C_n(t)| \leq O_P(1)\;\|t\|\;|n^{-1}\sum_{j=1}^n(\cos(t'X_j) - \Phi(t))\;|$ and thus $\|C_n\|_{\mathcal{L}^2} \;\leq\; O_P(1)$ $V_{n,3}$, where

$$V_{n,3} \;=\; \frac{1}{n^2}\sum_{j,k=1}^n\int_{\mathbf{R}^d}\|t\|^2\left(\cos(t'X_j) - \Phi(t)\right)\left(\cos(t'X_k) - \Phi(t)\right)\exp(-a\|t\|^2)dt.$$

Since $V_{n,3} \longrightarrow 0$ almost surely (notice that $E[\cos(t'X)] = \Phi(t)$), we have $\|C_n\|_{\mathcal{L}^2} = o_P(1)$. Using (3.8) and the triangle inequality for $\|\cdot\|_{\mathcal{L}^2}$, (3.6) follows.

By a standard central limit theorem for i.i.d. random random elements in Hilbert spaces, $\mathcal{W}_n^*(\cdot)$ converges to some centered Gaussian process on $\mathcal{L}^2$. Since $\mathcal{W}_n^*(\cdot)$ has the covariance kernel given in (3.1), assertion (3.7) follows, and the proof of Theorem 3.1 is completed. ■

We now consider the behavior of $T_{n,a}$ under contiguous alternatives to symmetry.

Theorem 3.2 *Suppose $X_{n1}, \ldots, X_{nn}$, $n \geq d+1$, is a triangular array of rowwise independent and identically distributed random variables having Lebesgue 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 which is symmetric around 0, i.e., we have $f_0(x) = f_0(-x)$, $x \in \mathbb{R}^d$, and h is a bounded function such that $\int h(x) f_0(x) dx = 0$. Then

$$\mathcal{W}_n(\cdot) \;\xrightarrow{\mathcal{D}}\; \mathcal{W}(\cdot) + c(\cdot),$$

where $\mathcal{W}_n(\cdot)$ and the Gaussian process $\mathcal{W}(\cdot)$ are defined in the statement of Theorem 3.1. The shift function $c(\cdot)$ is given by

$$c(t) \;=\; \int_{\mathbf{R}^d} \left[\sin(t'x) - t'\Phi(t)x\right] h(x) f_0(x) dx,$$

where $\Phi(t) = \int_{\mathbf{R}^d} \cos(t'y) f_0(y)\, dy$. Moreover,

$$T_{n,a} \;\xrightarrow{\mathcal{D}}\; \int_{\mathbf{R}^d} (\mathcal{W}(t) + c(t))^2 \exp(-a\|t\|^2) dt. \tag{3.9}$$

Proof. Mutatis mutandis, the reasoning closely follows the proof of Theorem 3.2 of Henze and Wagner (1997) and will thus not be given. Denoting by $Q^{(n)}$ and $P^{(n)}$ the joint distribution of $X_{n1}, \ldots, X_{nn}$ under f_n and under f_0, respectively, the shift function originates as the limit covariance, as $n \to \infty$, of $\mathcal{W}_n^*(t)$and $\log dQ^{(n)}/dP^{(n)}$, where $\mathcal{W}_n^*$ is defined in (3.4). ■

4 A permutational limit theorem for $T_{n,a}$

Since both the finite-sample and the asymptotic null distribution of $T_{n,a}$ depend on the underlying unknown distribution of X, a test that rejects H_0 for large values of $T_{n,a}$ cannot be performed without some sort of additional randomization. We propose to use the following permutation procedure.

Independently of the sequence $X_1, X_2, \ldots$, let $U_1, U_2, \ldots$ be a sequence of i.i.d. random variables such that $P(U_j = 1) = P(U_j = -1) = 1/2$. All random variables are assumed to be defined on a common probability space $(\Omega, \mathcal{A}, P)$. For a fixed $\omega \in \Omega$, the permutation procedure conditions on the realizations $y_j = Y_j(\omega)$ $(j = 1, \ldots, n)$ of the scaled vectors $Y_1, \ldots, Y_n$, which were defined in (1.2). The basic idea is that, under H_0, $Y_1, \ldots, Y_n$ should have a distribution that is approximately symmetric around 0. Consequently, the point pattern $U_1 y_1, \ldots, U_n y_n$, which arises from randomly reflecting a point around 0 with probability 1/2 or otherwise keeping it unchanged, independently of the other points, should also 'look symmetrically distributed' around 0. The permutation statistic we propose is

$$T_{n,a}^P \;=\; \int_{\mathbf{R}^d} \left(\mathcal{W}_n^P(t)\right)^2 \exp(-a\|t\|^2)dt, \tag{4.1}$$

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

$$\mathcal{W}_n^P(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n U_j \left\{ \sin(t'y_j) - \left(\frac{1}{n} \sum_{k=1}^n \cos(t'y_k) \right) t'y_j \right\}. \tag{4.2}$$

At first sight, it seems strange to consider $\mathcal{W}_n^P(\cdot)$ and not the 'obvious' process

$$\mathcal{V}_n^P(t) \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^n \sin(t'U_j y_j) \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^n U_j \sin(t'y_j)$$

and the corresponding permutation statistic $\int_{\mathbf{R}^d} \left(\mathcal{V}_n^P(t)\right)^2 \exp(-a\|t\|^2)dt$ (cf. (1.3)). The simple reason is that, unlike $\mathcal{W}_n^P(\cdot)$, the almost sure (i.e., for almost all sequences $X_1(\omega), X_2(\omega), \ldots$) limit process of $\mathcal{V}_n^P(\cdot)$ under H_0 has a covariance kernel that is different from the kernel (3.1).

We first give a representation of $T_{n,a}^P$ that is suitable for computational purposes.

Proposition 4.1 *Let* $Z_j = U_j y_j$ $(j = 1, \ldots, n)$, *and* $\bar{Z}_n = n^{-1} \sum_{j=1}^n Z_j$. *Then*

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

PROOF. Since $\cos(t'y_k) = \cos(t'Z_k)$, we have

$$\mathcal{W}_n^P(t) \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left\{ \sin(t'Z_j) - \left(\frac{1}{n} \sum_{k=1}^{n} \cos(t'Z_k) \right) t'Z_j \right\}$$

and thus

$$\begin{aligned} T_{n,a}^P \;=\;& \frac{1}{n} \sum_{i,j=1}^{n} \int_{\mathbf{R}^d} \sin(t'Z_i)\sin(t'Z_j)\exp(-a\|t\|^2)dt \\ &- \frac{2}{n^2} \sum_{i,j,k=1}^{n} \int_{\mathbf{R}^d} \sin(t'Z_i)\cos(t'Z_k)t'Z_j \exp(-a\|t\|^2)dt \\ &+ \frac{1}{n^3} \sum_{i,j,k,l=1}^{n} \int_{\mathbf{R}^d} \cos(t'Z_k)\cos(t'Z_l)t'Z_i t'Z_j \exp(-a\|t\|^2)dt \\ =\;& \frac{1}{n} \sum_{i,j=1}^{n} I_1(i,j) - \frac{2}{n^2} \sum_{i,j,k=1}^{n} I_2(i,j,k) + \frac{1}{n^3} \sum_{i,j,k,l=1}^{n} I_3(i,j,k,l) \end{aligned} \tag{4.3}$$

(say). Use the identities $\sin u \sin v = (\cos(u-v) - \cos(u+v))/2$, $\sin u \cos v = (\sin(u-v) + \sin(u+v)))/2$, $\cos u \cos v = (\cos(u-v) + \cos(u+v))/2$ and the formulae

$$\begin{aligned} \int_{\mathbf{R}^d} \cos(t'c)\exp(-a\|t\|^2)\; dt &= \left(\frac{\pi}{a}\right)^{d/2} \exp\left(-\frac{\|c\|^2}{4a}\right), \\ \int_{\mathbf{R}^d} \sin(t'c)t'b\exp(-a\|t\|^2)\; dt &= \frac{\pi^{d/2}}{2a^{d/2+1}} c'b \exp\left(-\frac{\|c\|^2}{4a}\right), \\ \int_{\mathbf{R}^d} \cos(t'c)t'bt'\gamma\exp(-a\|t\|^2)\; dt &= \frac{\pi^{d/2}}{4a^{d/2+2}} (2ab'\gamma - c'b\,c'\gamma)\exp\left(-\frac{\|c\|^2}{4a}\right) \end{aligned}$$

to obtain

$$\begin{aligned} I_1(i,j) &= \frac{\pi^{d/2}}{2a^{d/2}} \left[g_1(Z_i, Z_j) - g_2(Z_i, Z_j) \right], \\ I_2(i,j,k) &= \frac{\pi^{d/2}}{4a^{d/2+1}} \left[(Z_i - Z_k)'Z_j\, g_1(Z_i, Z_k) + (Z_i + Z_k)'Z_j\, g_2(Z_i, Z_k) \right], \\ I_3(i,j,k,l) &= \frac{\pi^{d/2}}{8a^{d/2+2}} \left[(2aZ_i'Z_j - (Z_k - Z_l)'Z_i(Z_k - Z_l)'Z_j)\, g_1(Z_k, Z_l) \right. \\ &\qquad \left. + (2aZ_i'Z_j - (Z_k + Z_l)'Z_i(Z_k + Z_l)'Z_j)\, g_2(Z_k, Z_l) \right], \end{aligned}$$

where $g_1(u,v) = \exp(-\|u-v\|^2/(4a))$ and $g_2(u,v) = \exp(-\|u+v\|^2)/(4a))$. Plugging these expressions into (4.3), the result is obtained after straightforward algebra.

■

For the special case $U_1 = U_2 = \ldots = U_n = 1$, we have $Z_j = y_j$ and $\bar{Z}_n = 0$. Consequently, $T^P_{n,a}$ takes the value

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

which is $T_{n,a}(y_1, \ldots, y_n)$ (as it should be!).

To prove the (conditional) convergence in distribution of the permutation process $\mathcal{W}^P_n$ to the Gaussian process $\mathcal{W}$ figuring in Theorem 3.1, we use the following Hilbert space Central Limit Theorem of Kundu et al. (Kundu, Majumdar, and Mukherjee (2000), Theorem 1.1). Therein, $\mathcal{H}$ denotes a real separable infinite-dimensional Hilbert space.

Lemma 4.2 *Let $\{e_k : k \geq 0\}$ be an orthonormal basis of $\mathcal{H}$. For each $n \geq 1$, let $W_{n1}, W_{n2}, \ldots, W_{nn}$ be a finite sequence of independent $\mathcal{H}$-valued random elements with zero means and finite second moments, and put $W_n = \sum_{j=1}^n W_{nj}$. Let C_n be the covariance operator of W_n. Assume that the following conditions hold:*

a) $\lim_{n\to\infty} \langle C_n e_k, e_l \rangle = a_{kl}$ *(say) exists for all* $k \geq 0$ *and* $l \geq 0$.

b) $\lim_{n\to\infty} \sum_{k=0}^{\infty} \langle C_n e_k, e_k \rangle = \sum_{k=0}^{\infty} a_{kk} < \infty$.

c) $\lim_{n\to\infty} L_n(\varepsilon, e_k) = 0$ *for every* $\varepsilon > 0$ *and every* $k \geq 0$, *where, for* $b \in \mathcal{H}$,
$L_n(\varepsilon, b) = \sum_{j=1}^n E\left(\langle W_{nj}, b\rangle^2 \mathbf{1}\{|\langle W_{nj}, b\rangle| > \varepsilon\}\right)$.

Then $W_n \Rightarrow \mathcal{N}(0, C)$ in $\mathcal{H}$, where the covariance operator C is characterized by $\langle Ch, e_l \rangle = \sum_{j=0}^{\infty} \langle h, e_j \rangle a_{jl}$, for every $l \geq 0$.

The main result of this section is as follows.

Theorem 4.3 *For almost all sample sequences $X_1(\omega), X_2(\omega), \ldots$, we have*

$$\mathcal{W}^P_n(\cdot) \xrightarrow{\mathcal{D}} \mathcal{W}(\cdot)$$

and

$$T^P_{n,a} \xrightarrow{\mathcal{D}} \int_{\mathbf{R}^d} \mathcal{W}^2(t) \exp(-a\|t\|^2) dt,$$

as $n \to \infty$, where $\mathcal{W}$ is the Gaussian process figuring in the statement of Theorem 3.1.

Proof. Let D^* be the set of all $\omega \in \Omega$ for which $\bar{X}_n(\omega) \to 0$, $n^{-1}\sum_{j=1}^n \|X_j(\omega)\|^r \to E\|X\|^r$ for $r = 1, 2$, $S_n(\omega) \to I_d$, and $n^{-1/2}\max_{1\le j\le n}\|X_j(\omega)\| \to 0$ as $n \to \infty$. By the law of large numbers and Theorem 5.2 of Barndorff-Nielsen (1963), D^* has measure one. For $s, t \in \mathbb{R}^d$, put

$$\begin{aligned}
D_s &= \left\{\omega \in \Omega : \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \cos(s'X_j^\omega) = E[\cos(s'X)]\right\}, \\
D_{t,s}^{(1)} &= \left\{\omega \in \Omega : \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \sin(s'X_j^\omega)\sin(t'X_j^\omega) = E[\sin(s'X)\sin(t'X)]\right\}, \\
D_{t,s}^{(2)} &= \left\{\omega \in \Omega : \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n s'X_j^\omega \sin(t'X_j^\omega) = E[s'X\sin(t'X)]\right\}.
\end{aligned}$$

where, for short, $X_j^\omega = X_j(\omega)$. Furthermore, let $D = D^* \cap \{\cap_{s\in T} D_s\} \cap \{\cap_{t,s\in T}(D_{t,s}^{(1)} \cap D_{t,s}^{(2)})\}$, where T is a countable dense set of $\mathbb{R}^d$. Being an intersection of countably many sets of

measure one, D has measure one as well. Then $D = D^* \cap \{\cap_{s\in\mathbf{R}^d} D_s\} \cap \{\cap_{t,s\in\mathbf{R}^d}(D_{t,s}^{(1)} \cap D_{t,s}^{(2)})\}$ by the Lipschitz continuity of the sine and cosine function.

In what follows, fix $\omega \in D$, and put

$$c_n^\omega(t) = \frac{1}{n}\sum_{k=1}^n \cos(t'Y_k^\omega),$$

where $Y_k^\omega = Y_k(\omega)$. By some algebra, it follows that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \left(\sin(t'Y_j^\omega) - c_n^\omega(t)t'Y_j^\omega\right) = 0,$$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \left(\sin(t'Y_j^\omega) - c_n^\omega(t)t'Y_j^\omega\right)\left(\sin(s'Y_j^\omega) - c_n^\omega(s)s'Y_j^\omega\right) = K(t,s).$$

For simplicity of notation, we will omit the superscript ω in the sequel. The proof will be completed by verifying conditions a) - c) of Lemma 4.2 for $W_{n1}, \ldots, W_{nn}$, where $W_{nj}(t) = U_j a_j(t)/\sqrt{n}$ and $a_j(t) = \sin(t'Y_j) - c_n(t)t'Y_j$. To this end, let C_n be the covariance operator of $W_n = \sum_{j=1}^n W_{nj}$ $(= \mathcal{W}_n^P)$, and put

$$K_n(s,t) = E\left[W_n(s)W_n(t)\right] = \frac{1}{n}\sum_{j=1}^n a_j(s)a_j(t).$$

As complete orthonormal set $\{e_k\}$ in $\mathcal{L}_2$, one can choose products of univariate Hermite polynomials (see, e.g., Rayner and Best (1989), p. 100). Since, for $\omega \in D$ and sufficiently large n,

$$\begin{aligned}
|K_n(s,t)| &\leq 1 + \frac{1}{n}\sum_{j=1}^{n}\left(|t'Y_j| + |s'Y_j| + |t'Y_j s'Y_j|\right) \\
&\leq 1 + (\|s\| + \|t\|)\frac{1}{n}\sum_{j=1}^{n}\|Y_j\| + \|t\|\|s\|\frac{1}{n}\sum_{j=1}^{n}\|Y_j\|^2 \\
&\leq 1 + (\|s\| + \|t\|)2E\|X\| + \|s\|\|t\|2E\|X\|^2,
\end{aligned}$$

and since $\lim_{n\to\infty} K_n(s,t) = K(s,t)$ for $\omega \in D$, dominated convergence yields

$$\begin{aligned}
\lim_{n\to\infty}\langle C_n e_k, e_l\rangle &= \lim_{n\to\infty}\int_0^\infty\int_0^\infty K_n(s,t)e_k(s)e_l(t)P_a(ds)P_a(dt) \\
&= \int_0^\infty\int_0^\infty K(s,t)e_k(s)e_l(t)P_a(ds)P_a(dt) \\
&= \langle Ce_k, e_l\rangle,
\end{aligned}$$

where $P_a(dt)$ is shorthand for $\exp(-a\|t\|^2)dt$, and C is the covariance operator of $\mathcal{W}$. Setting $a_{kl} = \langle Ce_k, e_l\rangle$, this proves condition a) of Lemma 4.2.

To verify condition b) of Lemma 4.2, use monotone convergence, Parseval's equality and dominated convergence to show

$$\begin{aligned}
\lim_{n\to\infty}\sum_{k=0}^{\infty}\langle C_n e_k, e_k\rangle &= \lim_{n\to\infty}\sum_{k=0}^{\infty}E\langle e_k, W_n\rangle^2 \\
&= \lim_{n\to\infty}E\|W_n\|^2_{\mathcal{L}^2} \\
&= \int_0^\infty \lim_{n\to\infty}K_n(t,t)P_a(dt) \\
&= \int_0^\infty K(t,t)P_a(dt) \\
&= E\|\mathcal{W}\|^2_{\mathcal{L}^2} \\
&= \sum_{k=0}^{\infty}a_{kk} < \infty.
\end{aligned}$$

To prove condition c) of Lemma 4.2, notice that

$$\begin{aligned}
|\langle W_{nj}, e_k\rangle| &= n^{-\frac{1}{2}}\left|\int U_j a_j(t)e_k(t)P_a(dt)\right| \\
&\leq n^{-\frac{1}{2}}\int |U_j a_j(t)e_k(t)|P_a(dt)
\end{aligned}$$

$$
\begin{aligned}
&\le n^{-\frac{1}{2}} \left(\int |U_j a_j(t)|^2 P_a(dt) \right)^{1/2} \|e_k\|_{\mathcal{L}^2} \\
&\le n^{-\frac{1}{2}} \left(\int (2 + 2\|t\|^2 \|Y_j\|^2) P_a(dt) \right)^{1/2} \\
&\le n^{-\frac{1}{2}} \left(\kappa_1 + \kappa_2 \max_{1\le j\le n} \|Y_j\| \right)
\end{aligned}
$$

for some positive constants κ_1 and κ_2. By the definition of the set D, the last expression converges to zero, whence

$$
E\left(\langle W_{nj}, e_k\rangle^2 \, \mathbf{1}\{|\langle W_{nj}, e_k\rangle| > \varepsilon\} \right) \;=\; 0
$$

for sufficiently large n, and thus $\lim_{n\to\infty} L_n(\varepsilon, e_k) = 0$. By Lemma 4.2, $W_n \Rightarrow \mathcal{N}(0, C)$ in $\mathcal{L}^2$. Since the above reasoning holds for every $\omega \in D$, the assertion follows. ■

5 Consistency

In this section, we prove the consistency of a test of symmetry that rejects H_0 for large values of $T_{n,a}$ against general alternatives. To this end, let X have an arbitrary distribution satisfying $E\|X\|^2 < \infty$. Moreover, we assume that the distribution of X puts mass zero on each $(d-1)$-dimensional hyperplane to ensure the almost sure invertibility of the sample covariance matrix S_n if $n > d$. In view of affine invariance, assume further that $E[X] = 0$ and $E[XX'] = I_d$.

Theorem 5.1 *Suppose the distribution of X is not symmetric (around 0). Then*

$$
\lim_{n\to\infty} P\left(T_{n,a} > c^P_{n,a}(\alpha) \right) \;=\; 1 \;,
$$

where $c^P_{n,a}(\alpha)$ denotes the $(1-\alpha)$-quantile of the distribution of the permutation statistic $T^P_{n,a}$.

Proof. We first prove

$$
\liminf_{n\to\infty} \frac{T_{n,a}}{n} \;\ge\; \int_{\mathbf{R}^d} \left(E(\sin(t'X)) \right)^2 \exp(-a\|t\|^2)\, dt \tag{5.1}
$$

almost surely. Since the right-hand side of (5.1) is strictly positive if the distribution of X is not symmetric (around 0), we have $\liminf_{n\to\infty} T_{n,a} = \infty$ almost surely, which entails consistency under such an alternative provided that the critical value, which is computed from the distribution of the permutation statistic $T^P_{n,a}$, is bounded in probability almost surely. To prove (5.1), notice that, by (1.3) and Fatou's lemma,

$$\liminf_{n\to\infty} \frac{T_{n,a}}{n} \;\geq\; \int_{\mathbf{R}^d} \liminf_{n\to\infty} \left(\frac{1}{n} \sum_{j=1}^{n} \sin(t'Y_j) \right)^2 \exp(-a\|t\|^2)\, dt$$

almost surely. Since, by the definition of Y_j, we have

$$|\sin(t'Y_j) - \sin(t'X_j)| \;\leq\; \|t\| \cdot \|(S_n^{-1/2} - I_d)X_j - S_n^{-1/2}\bar{X}_n\|,$$

use the strong law of large numbers to show $\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} \sin(t'Y_j) = E[\sin(t'X)]$ almost surely, whence (5.1) follows.

It remains to prove that the distribution of the permutation statistic $T^P_{n,a}$ is bounded in probability almost surely as $n \to \infty$. Of course, this implies almost sure boundedness of the critical value $c^P_{n,a}(\alpha)$, which is a quantile from that distribution. By Markov's inequality (notice that $T^P_{n,a}$ is nonnegative), it suffices to prove

$$\limsup_{n\to\infty} E_P[T^P_{n,a}] \;<\; \infty \tag{5.2}$$

almost surely, where $E_P(\cdot)$ denotes expectation with respect to the binary random variables $U_1, \ldots, U_n$ (cf. Section 4).

To prove (5.2), start with the representation of $T^P_{n,a}$ given in Proposition 4.1.Putting

$$Z^-_{j,k} = \exp\left(-\frac{1}{4a}\|Z_j - Z_k\|^2\right), \quad Z^+_{j,k} = \exp\left(-\frac{1}{4a}\|Z_j + Z_k\|^2\right)$$

and $C_n = \pi^{d/2}/(2a^{d/2}n)$, we have

$$E_P(T^P_{n,a}) \;=\; C_n \sum_{\nu=1}^{6} A_\nu(n),$$

where

$$A_1(n) \;=\; \sum_{j,k=1}^{n} E_P[Z^-_{j,k} - Z^+_{j,k}],$$

$$A_2(n) \;=\; \frac{1}{2a} \sum_{j,k=1}^{n} E_P\left[\|\bar{Z}_n\|^2 (Z^-_{j,k} + Z^+_{j,k})\right],$$

$$A_3(n) \;=\; -\frac{1}{a} \sum_{j,k=1}^{n} E_P\left[(Z_j - Z_k)'\bar{Z}_n Z^-_{j,k}\right],$$

$$
\begin{aligned}
A_4(n) &= -\frac{1}{a}\sum_{j,k=1}^{n} E_P\left[(Z_j+Z_k)'\bar{Z}_n Z_{j,k}^{+}\right], \\
A_5(n) &= -\frac{1}{4a^2}\sum_{j,k=1}^{n} E_P\left[\left\{(Z_j-Z_k)'\bar{Z}_n\right\}^2 Z_{j,k}^{-}\right], \\
A_6(n) &= -\frac{1}{4a^2}\sum_{j,k=1}^{n} E_P\left[\left\{(Z_j+Z_k)'\bar{Z}_n\right\}^2 Z_{j,k}^{+}\right].
\end{aligned}
$$

If $j \neq k$, then

$$
\begin{aligned}
E_P[Z_{j,k}^{-}-Z_{j,k}^{+}] &= \frac{1}{4}\exp\left(-\frac{1}{4a}\|y_j-y_k\|^2\right)+\frac{1}{4}\exp\left(-\frac{1}{4a}\|y_j+y_k\|^2\right) \\
&\quad -\left(\frac{1}{4}\exp\left(-\frac{1}{4a}\|y_j+y_k\|^2\right)+\frac{1}{4}\exp\left(-\frac{1}{4a}\|y_j-y_k\|^2\right)\right) \\
&= 0
\end{aligned}
$$

and thus

$$
\begin{aligned}
\Big|A_1(n)\Big| &= \left|\sum_{j=1}^{n}\left(1-E_P\left[\exp\left(-\frac{1}{a}\|Z_j\|^2\right)\right]\right)\right| \\
&= \left|\sum_{j=1}^{n}\left(1-\exp\left(-\frac{1}{a}\|y_j\|^2\right)\right)\right| \\
&\leq n.
\end{aligned}
$$

Writing $tr(\cdot)$ for trace, notice that

$$
\begin{aligned}
\sum_{j=1}^{n}\|y_j\|^2 &= \sum_{j=1}^{n}(x_j-\bar{x}_n)'S_n^{-1}(x_j-\bar{x}_n) \\
&= \sum_{j=1}^{n} tr\left(S_n^{-1}(x_j-\bar{x}_n)(x_j-\bar{x}_n)'\right) = tr\left(S_n^{-1}\, n\, S_n\right) = tr(nI_d) \\
&= nd,
\end{aligned}
$$

whence

$$
\begin{aligned}
E_P\|\bar{Z}_n\|^2 &= \frac{1}{n^2}\sum_{j,k=1}^{n} E_P\left[Z_j'Z_k\right] = \frac{1}{n^2}\sum_{j,k=1}^{n} y_j'y_k E\left[U_jU_k\right] \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\|y_j\|^2 \;=\; \frac{d}{n}.
\end{aligned}
$$

Since $0 \leq Z_{j,k}^{-} \leq 1$ and $0 \leq Z_{j,k}^{+} \leq 1$, it follows that $|A_2(n)| \leq nd/a$.

To tackle $A_3(n)$, notice that $E_P[(Z_j - Z_k)' Z_\nu Z_{j,k}^-] = 0$ if $\nu \notin \{j,k\}$. We therefore have

$$\begin{aligned} A_3(n) &= -\frac{1}{an} \sum_{j,k=1}^{n} E_P\left[(Z_j - Z_k)'(Z_j + Z_k) Z_{j,k}^-\right] \\ &= -\frac{1}{an} \sum_{j,k=1}^{n} E_P\left[(\|Z_j\|^2 - \|Z_k\|^2) Z_{j,k}^-\right] \end{aligned}$$

and thus

$$\begin{aligned} |A_3(n)| &\le \frac{1}{an} \sum_{j,k=1}^{n} E_P\left[\|Z_j\|^2 + \|Z_k\|^2\right] \\ &= \frac{1}{an} \sum_{j,k=1}^{n} \left(\|y_j\|^2 + \|y_k\|^2\right) \;=\; \frac{2nd}{a}. \end{aligned}$$

Likewise,

$$\begin{aligned} |A_4(n)| &\le \frac{1}{an} \sum_{j,k=1}^{n} E_P\left[\|Z_j + Z_k\|^2\right] \\ &\le \frac{2}{an} \sum_{j,k=1}^{n} E_P\left[\|Z_j\|^2 + \|Z_k\|^2\right] \;=\; \frac{4nd}{a}. \end{aligned}$$

An upper bound for $|A_5(n)|$ is

$$\begin{aligned} |A_5(n)| &\le \frac{1}{4a^2} \sum_{j,k=1}^{n} E_P\left[\|Z_j - Z_k\|^2 \|\bar{Z}_n\|^2\right] \\ &\le \frac{1}{2a^2} \sum_{j,k=1}^{n} E_P\left[\left(\|Z_j\|^2 + \|Z_k\|^2\right) \|\bar{Z}_n\|^2\right] \\ &= \frac{1}{2a^2} \sum_{j,k=1}^{n} (\|y_j\|^2 + \|y_k\|^2) E_P \|\bar{Z}_n\|^2 \;=\; \frac{d^2 n}{a^2}. \end{aligned}$$

In the same way, $|A_6(n)| \le (d^2 n)/a^2$. Summarizing, it follows that

$$E_P\left(T_{n,a}^P\right) \;\le\; \frac{\pi^{d/2}}{2a^{d/2}} \left(1 + \frac{7d}{a} + \frac{2d^2}{a^2}\right),$$

proving (5.2) and thus the consistency of the test for symmetry based on $T_{n,a}$ against alternatives satisfying the assumptions stated at the beginning of this section. ■

6 Simulation results

To assess the actual level of the test for symmetry based on $T_{n,a}$, a simulation study was performed for sample sizes $n = 20$, $n = 40$ and $n = 60$, dimensions $d = 2$, $d = 4$, $d = 6$, and the following symmetric distributions:

- the d-variate Standard Normal distribution $\mathcal{N}(0, I_d)$
- a mixture of $\mathcal{N}(0, I_d)$ and a d-variate normal distribution with mean zero, unit variances and equal correlation ρ between components, with mixing probabilities 0.75 and 0.25, respectively. This distribution is denoted by $\mathcal{NM}_1$, for $\rho = 0.25$ and $\mathcal{NM}_2$, for $\rho = 0.50$.
- the multivariate uniform distribution in the hypercube $[-1, 1]^d$, denoted by $\mathcal{U}$,
- the multivariate Student's distribution with ν degrees of freedom, denoted by t_ν.

For each fixed combination of n, d and the underlying distribution as given above, the following procedure was replicated 5 000 times:

1. generate a random sample $x_1, \ldots, x_n$
2. compute the scaled residuals $y_1, \ldots, y_n$ as defined in (1.2)
3. generate 500 independent pseudo-random vectors $(U_1, \ldots, U_n)$, where $U_1, \ldots, U_n$ are i.i.d. and $P(U_1 = 1) = P(U_1 = -1) = 1/2$.
4. calculate the corresponding 500 realizations $T^P_{n,a}(j),\ 1 \le j \le 500$ (say) of the permutation statistic $T^P_{n,a}$ (cf. Proposition 4.1).
5. reject H_0 if $T_{n,a}$, computed on $x_1, \ldots, x_n$, exceeds the empirical 95%-quantile of $T^P_{n,a}(j),\ 1 \le j \le 500$.

Table 1 shows the percentages of rejection of H_0. Notice that the observed level is fairly close to the nominal level 5% if $d = 2$ even for samples of size $n = 20$, but is far below the nominal level for the case $d = 6$ and $a = 1.0$. However, our simulation results indicate that the actual level of significance seems to approach its nominal value

			$a = 1.0$	$a = 2.0$	$a = 3.0$	$a = 4.0$
$\mathcal{N}(0, I_d)$	$n = 20$	$d = 2$	3.9	4.2	4.4	4.3
		$d = 4$	2.1	3.7	4.6	5.1
		$d = 6$	0.3	1.6	3.4	4.5
	$n = 40$	$d = 2$	4.5	4.9	5.2	5.2
		$d = 4$	3.6	4.6	5.2	5.4
		$d = 6$	1.7	3.8	4.8	5.5
	$n = 60$	$d = 2$	4.6	4.8	4.8	4.9
		$d = 4$	4.2	5.1	5.5	5.5
		$d = 6$	2.5	4.0	4.7	5.2
$\mathcal{NM}_1$	$n = 20$	$d = 2$	3.7	4.0	4.1	4.2
		$d = 4$	2.0	3.8	4.5	4.9
		$d = 6$	0.2	1.7	3.4	4.6
	$n = 40$	$d = 2$	4.7	5.0	5.2	5.1
		$d = 4$	4.0	5.0	5.6	5.8
		$d = 6$	1.6	3.8	4.8	5.2
	$n = 60$	$d = 2$	4.5	4.7	4.9	4.9
		$d = 4$	4.1	5.2	5.5	5.8
		$d = 6$	2.7	4.3	4.9	5.2
$\mathcal{NM}_2$	$n = 20$	$d = 2$	4.0	4.4	4.6	4.5
		$d = 4$	2.2	4.2	5.0	5.3
		$d = 6$	0.3	2.0	3.8	5.3
	$n = 40$	$d = 2$	4.8	5.2	5.5	5.4
		$d = 4$	3.9	5.1	5.9	6.2
		$d = 6$	1.8	4.0	5.2	5.8
	$n = 60$	$d = 2$	4.5	4.9	5.3	5.5
		$d = 4$	4.3	4.9	5.3	5.5
		$d = 6$	2.9	4.6	5.3	5.6
$\mathcal{U}$	$n = 20$	$d = 2$	3.3	3.3	3.4	3.4
		$d = 4$	1.3	2.1	2.6	2.7
		$d = 6$	0.1	0.9	1.8	2.4
	$n = 40$	$d = 2$	4.0	3.9	3.9	3.8
		$d = 4$	3.0	3.8	4.1	4.2
		$d = 6$	0.8	2.1	2.6	3.1
	$n = 60$	$d = 2$	4.6	4.8	4.7	4.6
		$d = 4$	3.4	4.4	4.6	4.6
		$d = 6$	1.5	2.7	3.1	3.3
t_{18}	$n = 20$	$d = 2$	4.6	5.3	5.3	5.5
		$d = 4$	2.9	5.2	6.3	6.9
		$d = 6$	0.5	2.8	5.3	6.9
	$n = 40$	$d = 2$	5.6	6.1	6.2	6.2
		$d = 4$	4.5	5.9	6.5	6.9
		$d = 6$	2.5	4.7	5.8	6.7
	$n = 60$	$d = 2$	4.6	5.0	5.1	5.3
		$d = 4$	4.6	5.6	6.2	6.6
		$d = 6$	3.4	5.3	6.4	6.7

Table 1: Estimated level for the permutation test (nominal level: 5%)

5% with increasing sample size, particularly for $a > 1$.

To assess the power of the test based on $T_{n,a}$, we simulated data from the following alternative distributions:

- A multivariate distribution with iid centered χ^2_1 marginals, denoted by χ^2_1,
- a convolution of the distributions $\mathcal{N}(0, I_d)$ and χ^2_1, denoted by $\mathcal{N} + \chi^2_1$,
- a multivariate lognormal distribution, as described in Johnson, Balakrishnan, and Kotz (2000), page 27, denoted by $\mathcal{LN}$. The simulated case corresponds to vectors with uncorrelated components each following (conditionally on the remaining components) a univariate lognormal distribution.
- a multivariate Gamma distribution, as described in Johnson et al. (2000), chapter 48, Sec. 3.1, denoted by $\Gamma(\alpha)$. The simulated cases correspond to vectors with 'practically' uncorrelated components each following a univariate gamma distribution with shape parameter α.

Tables 2-3 show the percentages of rejection of H_0, rounded to the nearest integer. An asterisk denotes power 100%.

Notice that power increases with the sample size. Moreover, the test becomes progressively more powerful as we depart from 'nearly' symmetric distributions (for example, the $\Gamma(10)$) and approach alternative distributions which are more skewed (for example, the $\Gamma(1)$). Hence, based on level and power results, we may suggest that a test corresponding to a larger value of a (perhaps $a = 3$ or 4) would be both powerful and accurate in estimating the nominal level, although we do not claim that this statement would be necessarily true under different sampling situations.

As additional alternative distributions, we considered non-symmetric bivariate normal mixtures. Let $\mathcal{N}(\mu, \rho)$ denote a bivariate normal distribution with mean (μ, μ), unit variances, and correlation ρ. We used the following mixtures:

- $\mathcal{NM}_3 : 0.5\mathcal{N}(0, 0) + 0.5\mathcal{N}(1, 0.5)$
- $\mathcal{NM}_4 : 0.5\mathcal{N}(0, 0) + 0.5\mathcal{N}(1, 0.9)$

			$a = 1.0$	$a = 2.0$	$a = 3.0$	$a = 4.0$
χ_1^2	$n = 20$	$d = 2$	95	96	96	96
		$d = 4$	94	97	98	98
		$d = 6$	76	92	96	97
	$n = 40$	$d = 2$	*	*	*	*
		$d = 4$	*	*	*	*
		$d = 6$	*	*	*	*
	$n = 60$	$d = 2$	*	*	*	*
		$d = 4$	*	*	*	*
		$d = 6$	*	*	*	*
$\mathcal{N}+\chi_1^2$	$n = 20$	$d = 2$	25	29	31	32
		$d = 4$	18	28	33	35
		$d = 6$	5	17	25	31
	$n = 40$	$d = 2$	55	59	60	61
		$d = 4$	57	66	69	71
		$d = 6$	47	63	69	71
	$n = 60$	$d = 2$	78	81	82	83
		$d = 4$	85	89	90	91
		$d = 6$	82	89	91	92
$\mathcal{LN}$	$n = 20$	$d = 2$	90	92	93	94
		$d = 4$	89	94	96	97
		$d = 6$	68	89	94	96
	$n = 40$	$d = 2$	*	*	*	*
		$d = 4$	*	*	*	*
		$d = 6$	*	*	*	*
	$n = 60$	$d = 2$	*	*	*	*
		$d = 4$	*	*	*	*
		$d = 6$	*	*	*	*

Table 2: Estimated power for the permutation test

- $\mathcal{NM}_5 : 0.5\mathcal{N}(0, -0.5) + 0.5\mathcal{N}(1, 0.5)$
- $\mathcal{NM}_6 : 0.5\mathcal{N}(0, -0.5) + 0.5\mathcal{N}(1, 0.9)$
- $\mathcal{NM}_7 : 0.5\mathcal{N}(0, -0.9) + 0.5\mathcal{N}(1, 0.5)$
- $\mathcal{NM}_8 : 0.5\mathcal{N}(0, -0.9) + 0.5\mathcal{N}(1, 0.9)$

Notice that the generalizations of the above covariance matrices with negative correlation ρ to higher dimensions are not positive definite. Hence, this part of the simulation is restricted to dimension 2. Tables 4 shows the percentages of rejection of H_0. In general, power seems to increase with increasing differences of correlations. In contrast to Tables 2-3, however, power does not always increase with increasing value

			$a = 1.0$	$a = 2.0$	$a = 3.0$	$a = 4.0$
$\Gamma(1)$	$n = 20$	$d = 2$	76	79	79	80
		$d = 4$	68	79	83	84
		$d = 6$	34	64	74	79
	$n = 40$	$d = 2$	99	$*$	$*$	$*$
		$d = 4$	$*$	$*$	$*$	$*$
		$d = 6$	$*$	$*$	$*$	$*$
	$n = 60$	$d = 2$	99	$*$	$*$	$*$
		$d = 4$	$*$	$*$	$*$	$*$
		$d = 6$	$*$	$*$	$*$	$*$
$\Gamma(2)$	$n = 20$	$d = 2$	46	49	50	49
		$d = 4$	34	45	50	52
		$d = 6$	9	30	40	46
	$n = 40$	$d = 2$	90	91	92	92
		$d = 4$	93	95	96	96
		$d = 6$	86	93	95	95
	$n = 60$	$d = 2$	99	99	99	99
		$d = 4$	$*$	$*$	$*$	$*$
		$d = 6$	$*$	$*$	$*$	$*$
$\Gamma(3)$	$n = 20$	$d = 2$	33	35	35	35
		$d = 4$	21	30	34	37
		$d = 6$	4	17	25	30
	$n = 40$	$d = 2$	76	78	78	78
		$d = 4$	77	82	84	85
		$d = 6$	64	78	82	84
	$n = 60$	$d = 2$	94	95	96	96
		$d = 4$	96	98	98	98
		$d = 6$	95	98	98	98
$\Gamma(5)$	$n = 20$	$d = 2$	19	21	22	22
		$d = 4$	11	18	21	22
		$d = 6$	1	9	15	18
	$n = 40$	$d = 2$	51	54	55	55
		$d = 4$	47	55	58	59
		$d = 6$	33	50	55	57
	$n = 60$	$d = 2$	76	79	80	80
		$d = 4$	79	84	85	86
		$d = 6$	71	82	84	85
$\Gamma(10)$	$n = 20$	$d = 2$	12	12	12	12
		$d = 4$	5	9	11	12
		$d = 6$	1	5	8	10
	$n = 40$	$d = 2$	26	29	29	29
		$d = 4$	22	28	30	31
		$d = 6$	13	23	27	29
	$n = 60$	$d = 2$	43	47	48	48
		$d = 4$	43	49	52	53
		$d = 6$	32	45	49	51

Table 3: Estimated power for the permutation test

		$a = 1.0$	$a = 2.0$	$a = 3.0$	$a = 4.0$	$a = 10.0$
$\mathcal{NM}_3$	$n = 20$	4.5	5.2	5.5	4.9	4.7
	$n = 40$	5.9	6.1	6.9	7.0	6.4
	$n = 80$	7.5	8.7	7.6	8.6	9.4
$\mathcal{NM}_4$	$n = 20$	8.2	10.3	11.5	11.4	11.6
	$n = 40$	12.0	13.6	15.0	14.8	15.4
	$n = 80$	21.3	22.0	22.7	21.4	23.4
$\mathcal{NM}_5$	$n = 20$	8.1	8.0	9.0	8.2	7.9
	$n = 40$	16.3	17.3	16.3	15.7	16.3
	$n = 80$	36.4	34.1	33.6	33.8	32.0
$\mathcal{NM}_6$	$n = 20$	13.4	16.5	16.7	17.0	17.5
	$n = 40$	22.7	24.5	25.5	25.6	27.7
	$n = 80$	47.4	44.9	45.0	44.9	43.6
$\mathcal{NM}_7$	$n = 20$	26.0	26.6	27.1	27.7	25.8
	$n = 40$	58.9	55.8	56.8	55.6	53.2
	$n = 80$	93.3	89.4	87.8	86.0	82.4
$\mathcal{NM}_8$	$n = 20$	28.7	32.3	33.7	34.7	34.1
	$n = 40$	55.6	55.7	55.4	54.4	53.4
	$n = 80$	89.3	84.7	79.8	78.3	75.0

Table 4: Estimated power for the permutation test, normal mixtures, $d = 2$

of the parameter a. Again, $a = 3$ or $a = 4$ seems to be a good choice, but more work regarding the choice of a 'good' value of a is needed.

Acknowledgement: Research on this topic started while the third author was visiting the University of Karlsruhe. Simos Meintanis would like to thank the Institute of Mathematical Stochastics for its hospitality and strong support.

References

Aki, S. (1981). Asymptotic distribution of a Cramér- von Mises type statistic for testing symmetry when the center is estimated. *Ann. Inst. Statist. Math.*, *33*, 1–14.

Alemayehu, D., Giné, E., and Pena, V. de la. (1993). Bootstrap goodness-of-fit tests based on the empirical charcteristic function. *Comput. Sci. & Statist. Proc. 25th Symp. Interface (M.E. Tarter. M.D. Lock, Eds.)*, 1–9.

Antille, A., Kersting, G., and Zucchini, W. (1982). Testing symmetry. *J. Amer. Statist. Ass.*, *77*, 639–646.

Baringhaus, L. (1991). Testing for spherical symmetry of a multivariate distribution. *Ann. Stat.*, *19*(2), 899–917.

Baringhaus, L., Gürtler, N., and Henze, N. (2000). Weighted integral test statistics and components of smooth tests of fit. *Austral. New Zeal. J. Statist.*, *42*(2), 179–192.

Barndorff-Nielsen, O. (1963). On the limit behaviour of extreme order statistics. *Ann. Math. Stat.*, *34*, 992–1002.

Bhattacharya, P., Gastwirth, J., and Wright, A. (1982). Two modified Wilcoxon tests for symmetry about an unknown location parameter. *Biometrika*, *69*, 377–382.

Boos, D. (1982). A test for symmetry associated with the Hodges-Lehmann estimator. *J. Amer. Statist. Ass.*, *77*, 647–651.

Csörgő, S., and Heathcote, C. (1982). Some results concerning symmetric distributions. *Bull. Austral. Math. Soc.*, *25*, 327–335.

Csörgő, S., and Heathcote, C. (1987). Testing for symmetry. *Biometrika*, *74*, 177–184.

Doksum, K., Fenstad, G., and Aaberge, R. (1977). Plots and tests for symmetry. *Biometrika*, *64*(3), 473–487.

Eaton, M., and Perlman, M. (1973). The non–singularity of generalized sample covariance matrices. *Ann. Stat.*, *1*(4), 710–717.

Feuerverger, A., and Mureika, R. (1977). The empirical characteristic function and its applications. *Ann. Stat.*, *5*(1), 88–97.

Ghosh, S., and Ruymgaart, F. (1992). Applications of empirical characteristic functions in some multivariate problems. *Canad. J. Statist.*, *20*, 429–440.

Gupta, A., and Kabe, D. (1993). Multivariate robust tests for spherical symmetry with applications to multivariate least squares regression. *J. Appl. Stat. Sci.*, *1*, 159–168.

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

Heathcote, C., Rachev, S., and Cheng, B. (1995). Testing multivariate symmetry. *J. Multiv. Anal.*, *54*, 91–112.

Henze, N. (1997a). Extreme smoothing and testing for multivariate normality. *Statist.* & *Prob. Lett.*, *35*, 203–213.

Henze, N. (1997b). Limit laws for multivariate skewness in the sense of Móri, Rohatgi and Székely. *Statist.* & *Prob. Lett.*, *33*, 299–307.

Henze, N., and Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *J. Multiv. Anal.*, *62*(1), 1–23.

Hollander, M. (1988). Testing for symmetry. *In. Encyclopedia of statistical sciences, Vol. 9 (S. Kotz and N.L. Johnson, eds.), Wiley, N.Y.*, *211–216*.

Johnson, N., Balakrishnan, N., and Kotz, S. (2000). *Continuous multivariate distributions* (Vol. 1). New York: Wiley.

Kariya, T., and Eaton, M. (1977). Robust tests for spherical symmetry. *Ann. Stat.*, *5*(1), 206–215.

King, M. L. (1980). Robust tests for sperical symmetry and their applications to least squares regression. *Ann. Stat.*, *8*, 1265–1271.

Koltchinskii, V., and Li, L. (1998). Testing of spherical symmetry of a multivariate distribution. *J. Multiv. Anal.*, *65*, 228–244.

Koutrouvelis, I. (1985). Distribution–free procedures for location and symmetry inference problems based on the empirical characteristic function. *Scand. J. Statist.*, *12*, 257–269.

Koziol, J. (1985). A note on testing symmetry with estimated parameters. *Statist.* & *Prob. Lett.*, *3*, 227–230.

Kundu, S., Majumdar, S., and Mukherjee, K. (2000). Central Limit Theorems revisited. *Statist.* & *Prob. Lett.*, *47*, 265–275.

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

Móri, T., Rohatgi, V., and Székely, G. (1993). On multivariate skewness and kurtosis. *Theory Probab. Appl.*, *38*, 547–551.

Neuhaus, G., and Zhu, L.-X. (1998). Permutation tests for reflected symmetry. *J. Multiv. Anal.*, *67*, 129–153.

Rayner, J., and Best, D. (1989). *Smooth tests of goodness of fit.* New York: Oxford University Press.

Schuster, E., and Barker, R. (1987). Using the bootstrap in testing symmetry versus asymmetry. *Commun. Statist. – Simula.*, *16*, 69–84.

Zhu, L., Fang, K., and Li, R. (1997). A new approach for testing symmetry of a high-dimensional distribution. *Bull. Hong Kong Math. Soc.*, *1*, 35–46.

Zhu, L., Fang, K., and Zhang, J. (1995). A projection NT-type test for spherical symmetry of a multivariate distribution. *In. New trends in probability and statistics (Tiit, E.-M., ed.), Vol.3: Multivariate statistics and matrices in statistics. Proc. 5th Conf., Tartu - Puehajaerve, Estonia 1994. Utrecht: VSP*, 109-122.