Two tests for multivariate normality based on the characteristic function*

Miguel A. Arcones

Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902.

E-mail:arcones@math.binghamton.edu

April 10, 2007

Abstract

We present two tests for multivariate normality. The presented tests are based on the Lévy characterization of the normal distribution and on the BHEP tests. The tests are affine invariant and consistent. We obtain the asymptotic limit null distribution of the test statistics using some results about generalized one–sample U–statistics which are of independent interest.

1. Introduction. A common assumption in many statistical procedures is normality of the observations. Since departure from the model can affect statistical procedures, testing for normality should be done before using several statistical methods. Many authors have presented different normality tests. Reviews of normality tests are Henze (2002), Thode (2002) and Mecklin and Mundfrom (2004). A classical normality test is the one in (SW) Shapiro and Wilk (1965, 1968). Epps and Pulley (1983) introduced a test of normality based on the weighted integral of the squared modulus of the difference the empirical (ch.f.) characteristic function and the normal ch.f. which is competitive with the Shapiro–Wilk test. Baringhaus and Henze (1988) extended the Epps–Pulley test to the multivariate setup and showed that this test is consistent against any alternative. Baringhaus and Henze (1988) also obtained the asymptotic distribution of this test under the null hypotheses. The test in Baringhaus and Henze (1988) is known in the literature as the BHEP test.

In this paper, we present two tests for multivariate normality, which are related to the BHEP test. Let $X_1, \dots, X_n$ be a sequence of i.i.d.r.v.'s with values in $\mathbb{R}^d$ and c.d.f. F. Let X

**Key words and phrases:* Tests of multivariate normality, characteristic function, characterization of normal distribution.

be a copy of X_1. We consider the testing problem

$$H_0 : F \text{ has a normal distribution, versus } H_1 : F \text{ does not,} \tag{1..1}$$

based on a random sample $X_1, \dots, X_n$ of size n from F.

We will use the usual multivariate notation. For example, given $u = (u_1, \dots, u_d)' \in \mathbb{R}^d$ and $v = (v_1, \dots, v_d)' \in \mathbb{R}^d$, $u'v = \sum_{j=1}^d u_j v_j$ and $|u| = (\sum_{j=1}^n u_j^2)^{1/2}$. Given a $d \times d$ matrix A, $\|A\| = \sup_{v_1, v_2 \in \mathbb{R}^d, |v_1|, |v_2| = 1} v_1' A v_2$. Given $\theta \in \mathbb{R}^d$ and $\epsilon > 0$, $B_d(\theta, \epsilon) = \{x \in \mathbb{R}^d : |x - \theta| < \epsilon\}$. I_d denotes the identity matrix in $\mathbb{R}^d$. $M(d \times d)$ is the set of $d \times d$ matrices. c will denote a constant which may vary from occurrence to occurrence.

We say that a $\mathbb{R}^d$–valued r.v. X has a nondegenerate d–variate distribution if $\Pr\{a'X = b\} < 1$, for each $a \in \mathbb{R}^d$ and each $b \in \mathbb{R}$. If $|X|$ has a second finite moment, then we may define the mean $\mu = E[X_1]$ and the covariance matrix $S = E[(X - \mu)(X - \mu)']$. We have that a nondegenerate d–variate r.v. X with $E[|X|^2] < \infty$ has a normal distribution if and only if $S^{-1/2}(X - \mu)$ has a $\mathbb{R}^d$-valued standard normal distribution. This is equivalent to $E[\exp(it'S^{-1/2}(X_1 - \mu))] = \psi(t)$, for each $t \in \mathbb{R}^d$, where $\psi(t) = \exp(-2^{-1}|t|^2)$. The BHEP test is significative for nonnormality for large values of the statistic

$$T_n := n \int_{\mathbb{R}^d} |\hat{\psi}_n(t) - \psi(t)|^2 \phi_\delta(t)\, dt, \tag{1..2}$$

where $\hat{\psi}_n(t) = n^{-1} \sum_{j=1}^n \exp(it'Y_{n,j})$, $Y_{n,j} = \hat{S}_n^{-1/2}(X_j - \bar{X}_n)$, $\bar{X}_n = n^{-1} \sum_{j=1}^n X_n$, $\hat{S}_n = n^{-1} \sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)'$ and $\phi_\delta(t) = (2\pi\delta^2)^{-d/2} \exp(-2^{-1}\delta^{-2}|t|^2)$. This test is well defined because, by Theorem 2.3 in Eaton and Perlman (1973), $\hat{S}_n$ is nonsingular a.s. Observe that in (1.2) the modulus of the complex number $\hat{\psi}_n(t) - \psi(t)$ is taken. From now, we will call the normality test based on the statistic in (1.2), the BHEP(δ) test. ϕ_δ is the pdf of a $\mathbb{R}^d$–valued normal r.v. with zero means and covariance matrix δI_d. By tuning the parameter δ it is possible to increase the power of the test (see Henze and Zirkler, 1990). Henze and Zirkler (1990) proved that for some distributions the BHEP test is more powerful when $\delta = 1/2$.

We present two normality tests based on the previous test. By the Lévy characterization of the normal distribution (see e.g. Theorem 20.2.A in Loève, 1977), given $m \ge 1$, a c.d.f. F has a normal distribution if and only if $m^{-1/2} \sum_{j=1}^m (X_j - E[X_j])$ has a normal distribution, where $X_1, \dots, X_m$ are i.i.d.r.v.'s with c.d.f. F. Hence, a nondegenerate d–variate c.d.f. F with $E_F[|X_1|^2] < \infty$ has a normal distribution if and only if for some $m \ge 1$,

$$D_m(F) := \int_{\mathbb{R}^d} \left| E_F\left[\exp\left(im^{-1/2} t' \left(\sum_{j=1}^m S_F^{-1/2}(X_j - \mu_F)\right)\right)\right] - \psi(t)\right|^2 \phi_\delta(t)\, dt = 0, \tag{1..3}$$

where $\mu_F = E_F[X_1]$ and $S_F = E_F[(X_1 - \mu)(X_1 - \mu)']$ and E_F is the expectation with respect to the probability measure for which the i.i.d.r.v.'s $X_1, \dots, X_m$ have c.d.f. F. Given location and scale estimators $\hat{\mu}_n = \hat{\mu}_n(X_1, \dots, X_n)$ and $\hat{S}_n = \hat{S}_n(X_1, \dots, X_n)$, we define

$$\hat{D}_{n,m} := \int_{\mathbb{R}^d} \left| \hat{\psi}_{n,m}(t) - \psi(t) \right|^2 \phi_\delta(t)\, dt, \tag{1..4}$$

where

$$\hat{\psi}_{n,m}(t) := \frac{(n-m)!}{n!} \sum_{(j_1,\ldots,j_m)\in I_m^n} \exp\left(i m^{-1/2} t' \hat{S}_n^{-1/2} \left(\sum_{p=1}^{m} (X_{j_p} - \hat{\mu}_n) \right) \right) \tag{1..5}$$

and $I_m^n = \{(j_1, \ldots, j_m) \in \mathbb{N}^m : 1 \le j_p \le n, j_p \ne j_q \text{ if } p \ne q\}$. We will see that if $\hat{\mu}_n$ and $\hat{S}_n$ are strong consistent estimators of μ and S respectively, then $\hat{D}_{n,m}$ converges to $D_m(F)$ a.s. (see Theorem 2.2). Possible location and scatter estimators are

$$\hat{\mu}_n := \bar{X}_n := n^{-1} \sum_{j=1}^{n} X_j, \tag{1..6}$$

$$\hat{S}_n := \frac{1}{n} \sum_{j=1}^{n} (X_j - \bar{X}_n)(X_j - \bar{X}_n)', \text{ if } \sum_{j=1}^{n} (X_j - \bar{X}_n)(X_j - \bar{X}_n)' \text{ is positive definite} \tag{1..7}$$

and

$$\hat{S}_n := I_d, \text{ if } \sum_{j=1}^{n} (X_j - \bar{X}_n)(X_j - \bar{X}_n)' \text{ is not positive definite.} \tag{1..8}$$

Note that by Theorem 2.3 in Eaton and Perlman (1973), $\hat{S}_n = n^{-1} \sum_{j=1}^{n} (X_j - \bar{X}_n)(X_j - \bar{X}_n)'$ a.s. But, many authors have introduced several robust estimators for the location and scale multivariate parameters (see e.g. Maronna, 1976; Davies, 1987; Lopuhaä and Rousseeuw, 1991; and Kent and Tyler, 1991). Given $1 > \alpha > 0$, let

$$a_{n,m,\alpha} = \inf\{\lambda \ge 0 : \mathbb{P}_{0,I_d}\{\hat{D}_{n,m} \le \lambda\} \ge 1 - \alpha\}, \tag{1..9}$$

where $\mathbb{P}_{0,I_d}$ is the probability measure when the data has a standard normal d dimensional distribution. Then,

$$\text{the test rejects the hypothesis if } \hat{D}_{n,m} > a_{n,m,\alpha}. \tag{1..10}$$

If $\hat{D}_{n,m}$ is affine invariant, then the distribution of $\hat{D}_{n,m}$ is the same for all normal distributions. Hence, the probability of type I error of the test is less than or equal to α. Conditions on the location and scale estimators to get that $\hat{D}_{n,m}$ is affine invariant are given in Theorem 2.1. The test statistic $\hat{D}_{n,m}$ can be found easily using the expression given by Lemma 2.1. The test is consistent for any alternative (Theorem 2.2). We will obtain the limit distribution of $n\hat{D}_{n,m}$ under the null hypothesis (Theorem 2.3).

A variation in the previous characterization of the normal distribution is as follows. Given $m \ge 2$ and a d–variate c.d.f. F, F has a normal distribution if and only if $m^{-1/2} \sum_{j=1}^{m} (X_j - E[X_j])$ and $X_1 - E[X_1]$ have the same distribution, where $X_1, \ldots, X_m$ are i.i.d.r.v.'s with c.d.f. F. Hence, a nondegenerate d–variate c.d.f. F with $E_F[|X_1|^2] < \infty$ has a normal distribution if and only if for some $m \ge 2$,

$$\begin{aligned} E_m(F) := \int_{\mathbb{R}^d} &|E_F[\exp(i m^{-1/2} t' S_F^{-1/2} (\sum_{j=1}^{m} (X_j - \mu_F)))] \\ &- E_F[\exp(i t' S_F^{-1/2} (X_1 - \mu_F))]|^2 \phi_\delta(t)\, dt = 0. \end{aligned} \tag{1..11}$$

An estimator of the previous quantity is

$$\hat{E}_{n,m} := \int_{\mathbb{R}^d} \left| \hat{\psi}_{n,m}(t) - \hat{\psi}_{n,1}(t) \right|^2 \phi_\delta(t)\,dt. \tag{1..12}$$

where $\hat{\psi}_{n,m}(t)$ is as in (1.5). Given $1 > \alpha > 0$, let

$$b_{n,m,\alpha} = \inf\{\lambda \geq 0 : \mathbb{P}_{0,I_d}\{\hat{E}_{n,m} \leq \lambda\} \geq 1 - \alpha\}. \tag{1..13}$$

Then,

$$\text{the test rejects the hypothesis if } \hat{E}_{n,m} > b_{n,m,\alpha}. \tag{1..14}$$

As before, if $\hat{E}_{n,m}$ is affine invariant, then the type I error of the test is less than or equal to α.

In Section 2, we present several properties of the tests in (1.10) and (1.14). The presented tests are affine invariant and omnibus. The test statistics can be found easily using the expressions given by lemmas 2.1 and 2.2. We obtain the asymptotic limit distribution of the test statistics under the null hypothesis. In Section 3, we study one sample generalized U–statistics. We obtain a theorem on the convergence of U–statistics with estimated parameters, which generalizes Theorem 2.16 in de Wet and Randles (1987). Section 4 contains the outcome of some simulations. Given the large amount of alternatives to the normal distribution, it is impossible to find a test, which outperform every possible test overall. The results of our simulations show that when the distribution is symmetric and unbounded, the power of the presented tests is higher than that of of the SW and BHEP tests. This indicates that knowing that the data is coming from a symmetric unbounded density our test is preferred. This is the case where testing normality is particularly needed. If there exists some indication that the data is not coming from a symmetric unbounded density, the hypothesis of normality should be rejected without testing. The proofs of the theorems are in Section 5.

Arcones and Wang (2005) presented tests for normality based on the empirical distribution function and the used characterization of the normal distribution. These tests are only of interest in the one dimensional situation. In the multivariate case, the tests in Arcones and Wang (2005) are not affine invariant. Simulations show that the presented tests have a higher power than the tests in Arcones and Wang (2005).

2. Properties of the tests.

First we give an alternative expression for $\hat{D}_{n,m}$.

Lemma 2..1. *$\hat{D}_{n,m}$ is equal to*

$$\frac{((n-m)!)^2}{(n!)^2} \sum_{(j_1,\ldots,j_m)\in I_m^n} \sum_{(k_1,\ldots,k_m)\in I_m^n} \exp\left(-2^{-1}\delta^2 m^{-1} \left| \sum_{p=1}^m Y_{n,j_p} - \sum_{q=1}^m Y_{n,k_q} \right|^2 \right) \tag{2..1}$$

$$-2(1+\delta^2)^{-d/2} \frac{(n-m)!}{n!} \sum_{(j_1,\ldots,j_m)\in I_m^n} \exp\left(-2^{-1}\delta^2(\delta^2+1)^{-1} m^{-1} \left| \sum_{p=1}^m Y_{n,j_p} \right|^2 \right) + (1+2\delta^2)^{-d/2},$$

where $Y_{n,j} = \hat{S}_n^{-1/2}(X_j - \hat{\mu}_n)$.

Next theorem gives conditions on the location and scale multivariate parameters to guarantee that $\hat{D}_{n,m}$ is affine invariant.

Theorem 2..1. *Let* $\mu : (\mathbb{R}^d)^n \to \mathbb{R}^d$ *and let* $S : (\mathbb{R}^d)^n \to M_{SPD}(d \times d)$, *where* $M_{SPD}(d \times d)$ *is the collection of the* $d \times d$ *symmetric positive definite matrices. Suppose that for each* $a, x_1 \dots, x_n \in \mathbb{R}^d$ *and each* $B \in M_{NS}(d \times d)$,

$$\mu(a + Bx_1, \dots, a + Bx_n) = a + B\mu(x_1, \dots, x_n) \tag{2..2}$$

and

$$S(a + Bx_1, \dots, a + Bx_n) = BS(x_1, \dots, x_n)B', \tag{2..3}$$

where $M_{NS}(d \times d)$ *is the collection of* $d \times d$ *nonsingular matrices. Let* $\hat{D}_{n,m} : (\mathbb{R}^d)^n \to \mathbb{R}$ *be defined by*

$$\hat{D}_{n,m}(x_1, \dots, x_n) = \int_{\mathbb{R}^d} \left| \tfrac{(n-m)!}{n!} \textstyle\sum_{(j_1,\dots,j_m) \in I_m^n} \exp(im^{-1/2}t'(\sum_{p=1}^m y_{j_p})) - \psi(t) \right|^2 \phi_\delta(t)\, dt,$$

where $y_j = (S(x_1, \dots, x_n))^{-1/2}(x_j - \mu(x_1, \dots, x_n))$.

Then, for each $a, x_1 \dots, x_n \in \mathbb{R}^d$ *and each* $B \in M_{NS}(d \times d)$,

$$\hat{D}_{n,m}(a + Bx_1, \dots, a + Bx_n) = \hat{D}_{n,m}(x_1, \dots, x_n).$$

It follows from the previous theorem that if the estimators $\hat{\mu}_n$ and $\hat{S}_n$ satisfy (2.2) and (2.3), then the statistic $\hat{D}_{n,m}$ is affine invariant. In particular, the distribution of $\hat{D}_{n,m}$ is the same for all nondegenerate normal distributions of the same dimension. It is easy to see that the estimators defined in (1.6)–(1.8) satisfy (2.2) and (2.3).

The following theorem gives the consistency of the first test statistic.

Theorem 2..2. *Let* $\{X_j\}_{j=1}^\infty$ *be a sequence of* $\mathbb{R}^d$*–valued i.i.d.r.v.'s with c.d.f.* F. *Suppose that* $E[|X_1|^2] < \infty$, $\hat{\mu}_n \xrightarrow{a.s.} \mu$ *and* $\hat{S}_n \xrightarrow{a.s.} S$. *Then,* $\hat{D}_{n,m} \xrightarrow{a.s.} D_m(F)$, *as* $n \to \infty$.

It follows from Theorem 2.2 that $a_{n,m,\alpha} \to 0$, as $n \to \infty$. Noting that $D_m(F) = 0$ if and only if F has a normal distribution, Theorem 2.2 implies that if F does not have a normal distribution, then for each $1 > \alpha > 0$ $\mathbb{P}_F\{\hat{D}_{n,m} > a_{n,m,\alpha}\} \to 1$, as $n \to \infty$.

If $E[|X_1|^2 < \infty$, then the estimators in (1.6)–(1.8) satisfy the conditions in the previous theorem. By the strong law of the large numbers $\bar{X}_n \xrightarrow{a.s.} \mu$. Since

$$\hat{S}_n = n^{-1} \sum_{j=1}^n (X_j - \mu)(X_j - \mu)' - (\bar{X}_n - \mu)(\bar{X}_n - \mu)' \ a.s.$$

$\hat{S}_n \xrightarrow{a.s.} S$.

The limit distribution of $n\hat{D}_{n,m}$ is the limit of a degenerate V–statistic. Let $\{X_j\}$ be a sequence of i.i.d.r.v.'s with values in measurable space $(S, \mathcal{S})$. Given a kernel $h : S \times S \to \mathbb{R}$, the U–statistic with kernel h is $(n(n-1))^{-1} \sum_{1 \le j \ne k \le n} h(X_j, X_k)$. The V–statistic with kernel h is $n^{-2} \sum_{j,k=1}^{n} h(X_j, X_k)$. A function $h : S \times S \to \mathbb{R}$ is called symmetric if for each $x, y \in S$, $h(x, y) = h(y, x)$. A function $h : S \times S \to \mathbb{R}$ is called degenerate if $E[h(x, X_1)] = E[h(X_1, x)] = E[h(X_1, X_2)]$ a.s. By Theorem 3.2.2.1 in Lee (1990, page 79), if h is a symmetric degenerate kernel with $E[(h(X_1, X_2))^2] < \infty$ and $E[h(X_1, X_2)] = 0$, then

$$n^{-1} \sum_{1 \le j \ne k \le n} h(X_j, X_k) \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty} \delta_k (g_k^2 - 1),$$

where $\{g_k\}_{k=1}^{\infty}$ is a sequence of i.i.d.r.v.'s with a standard one dimensional normal distribution and $\{\delta_k\}_{k=1}^{\infty}$ denotes the eigenvalues of the integral operator $A : L_2(S, \mathcal{S}, P) \to L_2(S, \mathcal{S}, P)$ defined by $Aq(x) = E[h(x, X_1)q(X_1)]$, where P is the law of X_1 and $L_2(S, \mathcal{S}, P)$ is the collection of measurable functions $\alpha : (S, \mathcal{S}) \to \mathbb{R}$ such that $E[(\alpha(X_1))^2] < \infty$. It follows from the previous result and the law of the large numbers that if h is a symmetric degenerate kernel with $E[(h(X_1, X_2))^2] < \infty$, $E[h(X_1, X_2)] = 0$ and $E[|h(X_1, X_1)|] < \infty$, then

$$n^{-1} \sum_{j,k=1}^{n} h(X_j, X_k) \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty} \delta_k (g_k^2 - 1) + E[h(X_1, X_1)], \tag{2..4}$$

where $\{\delta_k\}$ and $\{g_k\}$ are as before.

Next, we present the limit distribution of $n\hat{D}_{n,m}$ under the null hypothesis.

Theorem 2..3. *Let $\{X_j\}$ be a sequence of i.i.d. $\mathbb{R}^d$–valued r.v.'s with a normal distribution with mean $\mu = E[X]$ and nonsingular covariance matrix $S = E[(X - \mu)(X - \mu)']$. Let $\hat{\mu}_n = \hat{\mu}_n(X_1, \ldots, X_n)$ and let $\hat{S}_n = \hat{S}_n(X_1, \ldots, X_n)$ be estimators of μ and S respectively satisfying (2.2) and (2.3). Suppose that there are functions $\alpha : \mathbb{R}^d \to \mathbb{R}^d$ and $\beta : \mathbb{R}^d \to M_{SPD}(d \times d)$ such that $E[\alpha(S^{-1/2}(X_1 - \mu))] = 0$, $E[|\alpha(S^{-1/2}(X_1 - \mu))|^2] < \infty$, $E[\beta(S^{-1/2}(X_1 - \mu))] = 0$, $E[\|\beta(S^{-1/2}(X_1 - \mu))\|^2] < \infty$,*

$$n^{1/2}(\hat{\mu}_n - \mu) - n^{-1/2} \sum_{j=1}^{n} S^{1/2} \alpha(S^{-1/2}(X_j - \mu)) \xrightarrow{\mathrm{Pr}} 0, \tag{2..5}$$

$$n^{1/2}(\hat{S}_n - S) - n^{-1/2} \sum_{j=1}^{n} S^{1/2} \beta(S^{-1/2}(X_j - \mu)) S^{1/2} \xrightarrow{\mathrm{Pr}} 0. \tag{2..6}$$

Then,

$$n\hat{D}_{n,m} - n^{-1} \textstyle\sum_{j,k=1}^{n} h_{D,m}(S^{-1/2}(X_j - \mu), S^{-1/2}(X_k - \mu)) \xrightarrow{\mathrm{Pr}} 0,$$

where

$$h_{D,m}(x, y) = \int_{\mathbb{R}^d} g_{D,m}(x, t) g_{D,m}(y, t) \phi_\delta(t)\, dt,$$

and

$$g_{D,m}(x,t) = m\exp(-2^{-1}m^{-1}(m-1)|t|^2)\left(\cos(m^{-1/2}t'x)+\sin(m^{-1/2}t'x)\right) \\ +\exp(-2^{-1}|t|^2)(2^{-1}t'\beta(x)t - m^{1/2}t'\alpha(x) - m).$$

Consequently,

$$n\hat{D}_{n,m} \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty}\delta_{D,k}(g_k^2-1)+E[h_{D,m}(S^{-1}(X-\mu),S^{-1}(X-\mu))],$$

where $\{g_k\}$ *is a sequence of i.i.d.r.v.'s with a standard one dimensional normal distribution and* $\{\delta_{D,k}\}$ *denotes the eigenvalues of the operator* $A_D : L_2(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\mathcal{L}(S^{-1}(X_1-\mu))) \to L_2(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\mathcal{L}(S^{-1}(X_1-\mu)))$ *defined by* $A_Dq(x) = E[h_{D,m}(x,S^{-1}(X_1-\mu))q(S^{-1}(X_1-\mu))]$.

The estimators in (1.6)–(1.8) satisfy (2.5)–(2.6) with $\alpha(x)=x$ and $\beta(x)=xx'-I_d$.

If $m=1$, Theorem 2.3 coincides with the theorem in Section 3 in Baringhaus and Henze (1988).

For the second test, we have results similar to the ones for the first test.

Lemma 2..2. $\hat{E}_{n,m}$ *is equal to*

$$\frac{((n-m)!)^2}{(n!)^2}\sum_{(j_1,\ldots,j_m)\in I_m^n}\sum_{(k_1,\ldots,k_m)\in I_m^n}\exp(-2^{-1}\delta^2m^{-1}|\sum_{p=1}^m Y_{n,j_p}-\sum_{q=1}^m Y_{n,k_q}|^2) \\ -2\frac{(n-m)!}{n!n}\sum_{(j_1,\ldots,j_m)\in I_m^n}\sum_{k=1}^n \exp(-2^{-1}\delta^2m^{-1}|\sum_{q=1}^m Y_{n,j_q}-Y_{n,k}|^2) \\ +n^{-2}\sum_{j=1}^n\sum_{k=1}^n \exp(-2^{-1}\delta^2m^{-1}|Y_{n,j}-Y_{n,k}|^2).$$

The next theorem gives conditions on the location and scale multivariate parameters to guarantee that $\hat{E}_{n,m}$ is affine invariant.

Theorem 2..4. *Suppose that* $\mu : (\mathbb{R}^d)^n \to \mathbb{R}^d$ *and* $S : (\mathbb{R}^d)^n \to M_{SPD}(d\times d)$ *satisfy the conditions in Theorem 2.1. Let* $\hat{E}_{n,m} : (\mathbb{R}^d)^n \to \mathbb{R}$ *be defined by*

$$\hat{E}_{n,m}(x_1,\ldots,x_n) = \int_{\mathbb{R}^d}\Big|\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n}\exp(im^{-1/2}t'(\textstyle\sum_{p=1}^m y_{j_p})) \\ -n^{-1}\sum_{k=1}^n\exp(it'y_j))\Big|^2\phi_\delta(t)\,dt,$$

where $y_j = (S(x_1,\ldots,x_n))^{-1/2}(x_j-\mu(x_1,\ldots,x_n))$.

Then, for each $a, x_1\ldots,x_n\in\mathbb{R}^d$*, and each* $B\in M_{NS}(d\times d)$,

$$\hat{E}_{n,m}(a+Bx_1,\ldots,a+Bx_n) = \hat{E}_{n,m}(x_1,\ldots,x_n).$$

The second test is also consistent:

Theorem 2..5. *Let* $\{X_j\}$ *be a sequence of i.i.d.* $\mathbb{R}^d$*–valued r.v.'s with c.d.f.* F*. Suppose that* $E[|X_1|^2]<\infty$, $\hat{\mu}_n \xrightarrow{a.s.} \mu$ *and* $\hat{S}_n \xrightarrow{a.s.} S$*. Then,* $\hat{E}_{n,m} \xrightarrow{a.s.} E_m(F)$*, as* $n\to\infty$.

Theorem 2.5 implies that if F does not have a normal distribution, then, for each $1>\alpha>0$, $\mathbb{P}_F\{\hat{E}_{n,m} > b_{n,m,\alpha}\}\to 1$, as $n\to\infty$.

The limit null distribution of the second test is given by the following theorem:

Theorem 2..6. *Under the conditions in Theorem 2.3,*

$$n\hat{E}_{n,m} - n^{-1}\sum_{j,k=1}^{n} h_{E,m}(S^{-1/2}(X_j-\mu), S^{-1/2}(X_k-\mu)) \xrightarrow{\text{Pr}} 0.$$

where

$$h_{E,m}(x,y) = \int_{\mathbb{R}^d} g_{E,m}(x,t)g_{E,m}(y,t)\phi_\delta(t)\,dt,$$

and

$$\begin{aligned} g_{E,m}(x,t) = {} & m\exp(-2^{-1}m^{-1}(m-1)|t|^2)\left(\cos(m^{-1/2}t'x)+\sin(m^{-1/2}t'x)\right) \\ & -(\cos(t'x)+\sin(t'x))+\exp(-2^{-1}|t|^2)(1-m-(m^{1/2}-1)t'\alpha(x)). \end{aligned}$$

Consequently,

$$n\hat{E}_{n,m} \xrightarrow{\text{d}} \sum_{k=1}^{\infty}\delta_{E,k}(g_k^2-1) + E[h_{E,m}(S^{-1}(X-\mu), S^{-1}(X-\mu))],$$

where $\{g_k\}$ *is a sequence of i.i.d.r.v.'s with a standard one dimensional normal distribution and* $\{\delta_{E,k}\}$ *denotes the eigenvalues of the operator* $A_E : L_2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathcal{L}(S^{-1}(X_1-\mu))) \to L_2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathcal{L}(S^{-1}(X_1-\mu)))$ *defined by* $A_E q(x) = E[h_{E,m}(x, S^{-1}(X-\mu))q(S^{-1}(X-\mu))]$.

Note that the limit distribution of $n\hat{E}_{n,m}$ does not depend on the limit distribution of $n^{1/2}(\hat{S}_n - S)$.

3. Limit theorems for generalized one sample U–statistics. By lemmas 2.1 and 2.2, the statistics $\hat{D}_{n,m}$ and $\hat{E}_{n,m}$ are generalized one sample U–statistics with an estimated parameter. Given r.v.'s $X_1,\dots,X_n$ with values in a measurable space $(S,\mathcal{S})$ and a measurable function $h:(S^m,\mathcal{S}^m)\to\mathbb{R}$, the U–statistic with kernel h is defined by

$$U_{n,m}(h) := \frac{(n-m)!}{n!}\sum_{(j_1,\dots,j_m)\in I_m^n} h(X_{j_1},\dots,X_{j_m}).$$

General references on U–statistics are Lee (1990) and de la Peña and Giné (1999). We say that a kernel h is symmetric if for each $x_1,\dots,x_m$ and each $\sigma\in\text{Perm}(\{1,\dots,m\})$

$$h(x_1,\dots,x_m) = h(x_{\sigma(1)},\dots,x_{\sigma(m)}),$$

where $\text{Perm}(A)$ is the set of permutations on the set A. Given a function h on S^m, let

$$S_m h(x_1,\dots,x_m) = (m!)^{-1}\sum_{\sigma\in\text{Perm}(\{1,\dots,m\})} h(x_{\sigma(1)},\dots,x_{\sigma(m)}).$$

Then, $S_m h$ is a symmetric function and $U_{n,m}(h) = U_{n,m}(S_m h)$.

The main property of U–statistics which we will use is the Hoeffding decomposition, which we describe next. Suppose that $X_1, \ldots, X_n$ are i.i.d.r.v.'s with law P. We define

$$\pi_{k,m}h(x_1, \ldots, x_k) = (\delta_{x_1} - P) \cdots (\delta_{x_m} - P)P^{m-k}h$$

where $Q_1 \cdots Q_m h = \int \cdots \int h(x_1, \ldots, x_m)\, dQ_1(x_1) \cdots dQ_m(x_m)$. Then, it is known that

$$U_{n,m}(h) = \sum_{k=0}^{m} \binom{m}{k} U_{n,k}(\pi_{k,m}h) \tag{3..1}$$

(see Section 3.5 in de la Peña and Giné, 1999). Since the decomposition in (3.1) all the terms are orthogonal, if h is symmetric, then

$$\mathrm{Var}(U_{n,m}(h)) = \sum_{k=1}^{m} \frac{\binom{m}{k}^2}{\binom{n}{k}} E[((\pi_{k,m}h)(X_1, \ldots, X_k))^2]. \tag{3..2}$$

The symmetric function $h(x_1, \ldots, x_m)$ is said to be P–degenerate of order r, $0 \le r \le m-1$, if $\delta_{x_1} \cdots \delta_{x_r} \cdot P^{m-r}h \equiv P^m h$ a.s. but $\delta_{x_1} \cdots \delta_{x_{r+1}} \cdot P^{m-r-1}h$ is not a constant a.s. It is easy to see that h is P–degenerate of order r if and only if $\pi^P_{1,m}h = \cdots = \pi^P_{r,m}h \equiv 0$ and $\pi^P_{r+1,m}h \not\equiv 0$. If follows from (3.2) that if h is a symmetric function which P–degenerate of order r, then

$$\mathrm{Var}(U_{n,m}(h)) \le cn^{-r-1}E[(h(X_1, \ldots, X_m))^2]. \tag{3..3}$$

We will need to work with a certain generalization of U–statistics. Given a measurable function $h : S^{m_1 \times m_2} \to \mathbb{R}$, the generalized one–sample U–statistic with kernel h is

$$\begin{aligned}(GU)_{n,m_1,m_2}(h) := \frac{(n-m_1)!}{n!}\frac{(n-m_2)!}{n!} &\sum_{(i_1,\ldots,i_{m_1}) \in I^n_{m_1}} \sum_{(j_1,\ldots,j_{m_2}) \in I^n_{m_2}} \\ &\times h(X_{i_1}, \ldots, X_{i_{m_1}}, X_{j_1}, \ldots, X_{j_{m_2}}).\end{aligned} \tag{3..4}$$

We will assume that for each $x_1, \ldots, x_{m_1}, y_1, \ldots, y_{m_2} \in S$, each $\sigma_1 \in \mathrm{Perm}(\{1, \ldots, m_1\})$ and each $\sigma_2 \in \mathrm{Perm}(\{1, \ldots, m_2\})$,

$$h(x_1, \ldots, x_{m_1}, y_1, \ldots, y_{m_2}) = h(x_{\sigma_1(1)}, \ldots, x_{\sigma_1(m_1)}, y_{\sigma_2(1)}, \ldots, y_{\sigma_2(m_2)}). \tag{3..5}$$

Given $\max(m_1, m_2) \le k \le m_1 + m_2$, define the function g_k by

$$g_k(x_1, \ldots, x_k) = h(x_1, \ldots, x_{m_1}, x_1, \ldots, x_{m_1+m_2-k}, x_{m_1+1}, x_{m_1+2}, \ldots, x_k).$$

Note that in the previous expression the first $m_1 + m_2 - k$ indices in each group are repeated. We have that

$$\begin{aligned} &(GU)_{n,m_1,m_2}(h) \\ = &\textstyle\sum_{k=\max(m_1,m_2)}^{m_1+m_2} \frac{(n-m_1)!}{n!}\frac{(n-m_2)!}{n!} \binom{m_1}{m_1+m_2-k}\binom{m_2}{m_1+m_2-k}(m_1+m_2-k)! \\ &\times \textstyle\sum_{(j_1,\ldots,j_k) \in I^n_k} g_k(X_{j_1}, \ldots, X_{j_k}) \\ = &\textstyle\sum_{k=\max(m_1,m_2)}^{m_1+m_2} \frac{(n-m_1)!(n-m_2)!m_1!m_2!}{n!(n-k)!(m_1+m_2-k)!(k-m_1)!(k-m_2)!} U_{n,k}(S_k g_k). \end{aligned} \tag{3..6}$$

The previous inequality holds because there are $\binom{m_1}{m_1+m_2-k}$ ways to choose the m_1+m_2-k repeated indices in the first group, there are $\binom{m_2}{m_1+m_2-k}$ ways to choose the m_1+m_2-k repeated indices in the second group, and there are $(m_1+m_2-k)!$ ways to pair the m_1+m_2-k repeated indices of each group.

From the decomposition in (3.6), it is possible to obtain the LLN and the CLT for generalized U–statistics. For each $\max(m_1,m_2) \le k \le m_1+m_2$,

$$n^{m_1+m_2-k}\frac{(n-m_1)!(n-m_2)!(m_1)!(m_2)!}{n!(n-k)!(m_1+m_2-k)!(k-m_1)!(k-m_2)!} \to \frac{m_1!m_2!}{(m_1+m_2-k)!(k-m_1)!(k-m_2)!},$$

as $n \to \infty$. So, under minor conditions, $(GU)_{n,m_1,m_2}(h)$ is asymptotically equivalent to $U_{n,m_1+m_2}(g_{m_1+m_2})$.

In this section, we present several limit theorems over a sequence $\{X_j\}$ of i.i.d.r.v.'s with values in a measurable space $(S,\mathcal{S})$ and law P. The following theorems give sufficient conditions for the strong law of the large numbers and the central limit theorem for one sample generalized U–statistics, respectively.

Theorem 3..1. *Let $h : S^{m_1+m_2} \to \mathbb{R}$ be a measurable function satisfying (3.5). Suppose that for each* $\max(m_1,m_2) \le k \le m_1+m_2$,

$$E[|h(X_1,\dots,X_{m_1},X_1,\dots,X_{m_1+m_2-k},X_{m_1+1},X_{m_1+2},\dots,X_k)|^{k/(m_1+m_2)}] < \infty,$$

then

$$(GU)_{n,m_1,m_2}(h) \overset{a.s.}{\to} E[h(X_1,\dots,X_{m_1+m_2})].$$

Theorem 3..2. *Let $h : S^{m_1+m_2} \to \mathbb{R}$ be a measurable function satisfying (3.5). Suppose that*

$$E[(h(X_1,\dots,X_{m_1+m_2}))^2] < \infty$$

and for each $\max(m_1,m_2) \le k \le m_1+m_2-1$,

$$E[|h(X_1,\dots,X_{m_1},X_1,\dots,X_{m_1+m_2-k},X_{m_1+1},X_{m_1+2},\dots,X_k)|^{k/(m_1+m_2-2^{-1})}] < \infty,$$

then

$$n^{1/2}((GU)_{n,m_1,m_2}(h) - E[h(X_1,\dots,X_{m_1+m_2})]) \overset{\mathrm{d}}{\to} N(0,\mathrm{Var}(m_1h_1(X_1)+m_2h_2(X_1))),$$

where $h_1(x) = E[h(x,X_1,\dots,X_{m_1+m_2-1})]$ *and*
$h_2(x) = E[h(X_1,\dots,X_{m_1},x,X_{m_1+1},\dots,X_{m_1+m_2-1})]$.

The test statistics of the presented tests are one sample generalized U–statistics with an estimated parameter. To obtain the limit distribution of the test statistics under the null hypothesis we will use the following theorem.

Theorem 3..3. *Let $\{X_j\}$ be a sequence of i.i.d.r.v.'s with values in the measurable space $(S,\mathcal{S})$. Let $(T,\mathcal{T},\nu)$ be a measure space. Let $\Theta\subset\mathbb{R}^s$. Let $\theta_0\in\Theta$. For each $l=1,2$, let $g_l: S^{m_l}\times T\times\Theta\to\mathbb{R}$ be a measurable function, where m_l is a positive integer. Let $\hat\theta_n=\hat\theta_n(X_1,\dots,X_n)$ be an estimator of θ_0. Suppose that there exists $\epsilon_0>0$ such that:*

(i) For each $l=1,2$, each $t\in T$, each $\theta\in\Theta$, and each $\sigma_l\in\mathrm{Perm}(\{1,\dots,m_l\})$,

$$g_l(x_1,\dots,x_{m_l},t,\theta)=g_l(x_{\sigma_l(l)},\dots,x_{\sigma_l(m_l)},t,\theta).$$

(ii) For each $l=1,2$, each $t\in T$ and each $\theta\in\Theta$ with $|\theta-\theta_0|\le\epsilon_0$,

$$E[|g_l(X_1,X_2,\dots,X_{m_l},t,\theta)|]<\infty.$$

(iii) For each $l=1,2$ and each $t\in T$, $G_l(t,\theta_0)=0$, where $G_l(t,\theta):=E[g_l(X_1,\dots,X_{m_l},t,\theta)]$.

(iv) For each $l=1,2$,

$$E[\int_T (g_l(X_1,\dots,X_{m_l},t,\theta_0))^4\,d\nu(t)]<\infty$$

and

$$E[\int_T (g_l(X_1,X_1,X_2,\dots,X_{m_l-1},t,\theta_0))^4\,d\nu(t)]<\infty.$$

(v) For each $l=1,2$ and each $m_l\le k\le 2m_l-2$,

$$E[|\int_T g_l(X_1,\dots,X_{m_l},t,\theta_0)g_l(X_1,\dots,X_{2m_l-k},X_{m_l+1},\dots,X_k,t,\theta_0)\,d\nu(t)|]<\infty.$$

(vi) For each $l=1,2$ and each $t\in T$, there exists $\dot G_j(t,\theta_0)\in\mathbb{R}^s$ such that

$$\lim_{\theta\to\theta_0}|\theta-\theta_0|^{-2}\int_T |G_l(t,\theta)-G_l(t,\theta_0)-(\dot G_l(t,\theta_0))'(\theta-\theta_0)|^2\,d\nu(t)=0$$

and $\int_T|\dot G_l(t,\theta_0)|^4\,d\nu(t)<\infty$.

(vii) There exists a function $\alpha: S\to\mathbb{R}^s$ such that $E[\alpha(X)]=0$, $E[|\alpha(X)|^2]<\infty$, and

$$n^{1/2}(\hat\theta_n-\theta_0)-n^{-1/2}\sum_{j=1}^n\alpha(X_j)\xrightarrow{\Pr}0.$$

(viii) For each $l=1,2$ and each $0<\delta\le\epsilon_0$,

$$\int_T E[\sup_{|\theta-\theta_0|\le\epsilon_0,|\theta+\gamma-\theta_0|\le\epsilon_0,|\gamma|\le\delta}|g_l(X_1,\dots,X_{m_l},t,\theta+\gamma)-g_l(X_1,\dots,X_{m_l},t,\theta)|^2]\,d\nu(t)\le A\delta^2,$$

where A is a finite constant.

(ix) For each $|\theta|<\epsilon_0$,

$$E\left[\int_T\left(g(X_1,\dots,X_m,t,\theta_0+n^{-1/2}\theta)-g(X_1,\dots,X_m,t,\theta_0)\right.\right.$$
$$\left.\left.-G(t,\theta_0+n^{-1/2}\theta)\right)^2\,d\nu(t)\right]\to 0.$$

(x) For each $|\theta| < \epsilon_0$,

$$E\left[\int_T \left(g(X_1,\dots,X_m,t,\theta_0+n^{-1/2}\theta) - g(X_1,\dots,X_m,t,\theta_0) \right.\right.$$
$$\left.\left. -G(t,\theta_0+n^{-1/2}\theta)\right)^4 \, d\nu(t)\right] \to 0.$$

(xi) For each $m \le k \le 2m-2$ and each $|\theta| \le \epsilon_0$,

$$n^{1+k-2m}E[| \int_T (g(X_1,\dots,X_m,t,\theta_0+n^{-1/2}\theta) - g(X_1,\dots,X_m,t,\theta_0) - G(t,\theta_0+n^{-1/2}\theta))$$
$$\times(g(X_1,\dots,X_{2m-k},X_{m+1},\dots,X_k,t,\theta_0+n^{-1/2}\theta)$$
$$-g(X_1,\dots,X_{2m-k},X_{m+1},\dots,X_k,t,\theta_0) - G(t,\theta_0+n^{-1/2}\theta))\,d\nu(t)|] \to 0.$$

Then,

$$\begin{aligned} & n\tfrac{(n-m_1)!}{n!}\tfrac{(n-m_2)!}{n!}\textstyle\sum_{(i_1,\dots,i_{m_1})\in I^n_{m_1}}\sum_{(j_1,\dots,j_{m_2})\in I^n_{m_2}}\int_T g_1(X_{i_1},\dots,X_{i_{m_1}},t,\hat\theta_n) \\ & \qquad \times g_2(X_{i_1},\dots,X_{i_{m_1}},t,\hat\theta_n)\,d\nu(t) \\ - & \qquad n^{-1}\textstyle\sum_{j,k=1}^n \int_T (m_1h_1(X_j,t,\theta_0)+(\dot G_1(t,\theta_0))'\alpha(X_j)) \\ & \qquad \times(m_2h_2(X_k,t,\theta_0)+(\dot G_2(t,\theta_0))'\alpha(X_k))\,d\nu(t) \\ & \qquad \overset{\Pr}{\to} 0, \end{aligned} \tag{3..7}$$

where $h_l(x,t,\theta) = E[g_l(x,X_2,\dots,X_{m_l},t,\theta)]$.

Besides,

$$\{n^{-1}\sum_{j,k=1}^n \int_T (m_1h_1(X_j,t,\theta_0)+(\dot G(t,\theta_0))'\alpha(X_j))(m_1h_2(X_k,t,\theta_0)+(\dot G(t,\theta_0))'\alpha(X_k))\,d\nu(t)\}$$

converges in distribution to $\sum_{k=1}^\infty \delta_k(g_k^2-1)+E[\bar h(X_1,X_2)]$, where $\{g_k\}$ is a sequence of i.i.d.r.v.'s with a standard one dimensional normal distribution, $\{\delta_k\}$ denotes the eigenvalues of the operator A defined by $Aq(x) = E[\bar h(x,X)q(X)]$, and

$$\bar h(x,y) = \int_T 2^{-1}\Big((m_1h_1(x,t,\theta_0)+(\dot G_1(t,\theta_0))'\alpha(x))(m_2h_2(y,t,\theta_0)+(\dot G_2(t,\theta_0))'\alpha(y))$$
$$+(m_1h_1(y,t,\theta_0)+(\dot G_1(t,\theta_0))'\alpha(y))(m_2h_2(x,t,\theta_0)+(\dot G_2(t,\theta_0))'\alpha(x))\,d\nu(t)\Big).$$

If the case $m_1 = m_2 = 1$, the previous theorem relaxes Theorem 2.16 in de Wet and Randles (1987).

4. Simulations. We have made simulations to compare the power of the presented tests with the SW test and the BHEP test for some alternative hypothesis. We only consider the presented tests in the case $m = 2$ and $\delta = 1$.

First, we consider one dimensional distributions. The following table shows the values of $na_{n,2,\alpha}$ and $nb_{n,2,\alpha}$ in (1.9) and (1.13), for some values of n. We use $\bar X$ as a location estimator and $\hat{s}_n = \left(n^{-1}\sum_{j=1}^n (X_j - \bar{X}_n)^2\right)^{1/2}$ as the scale parameter. The corresponding values of $a_{n,2,\alpha}$ and $b_{n,m,\alpha}$ when $\hat{s}_n^{\text{unb}} = \left((n-1)^{-1}\sum_{j=1}^n (X_j - \bar{X}_n)^2\right)^{1/2}$ is used instead of the biased sample variance are called $a_{n,2,\alpha}^{\text{unb}}$ and $b_{n,2,\alpha}^{\text{unb}}$. The table was obtained making 10000 simulations from a standard normal distribution. The test statistics were found using the expressions given by lemmas 2.1 and 2.2.

TABLE 1

$na_{n,2,\alpha}(1)$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.04824971	0.07791181	0.14309058
$n = 8$	0.06160079	0.09887000	0.20911663
$n = 10$	0.06804454	0.10876316	0.21063206
$n = 15$	0.07967803	0.12509874	0.24958859
$na_{n,2,\alpha}^{\text{unb}}(1)$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.05228323	0.06914429	0.10640492
$n = 8$	0.06070897	0.09069875	0.17879543
$n = 10$	0.07013934	0.10392443	0.21065929
$n = 15$	0.07892513	0.11329111	0.22407420
$nb_{n,2,\alpha}(1)$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.1574244	0.1902732	0.2590276
$n = 8$	0.1893065	0.2167707	0.2686646
$n = 10$	0.2226303	0.2497585	0.3008117
$n = 15$	0.3132090	0.3306479	0.3764694
$nb_{n,2,\alpha}^{\text{unb}}(1)$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.1228755	0.1489783	0.1972663
$n = 8$	0.1630069	0.1850474	0.2261630
$n = 10$	0.2005697	0.2208427	0.2642223
$n = 15$	0.2953086	0.3138741	0.3575843

The following table shows the power when $\alpha = 0.05$ of the following tests of normality: Shapiro and Wilk (1965), Epps and Pulley, for $\delta = 1$ and $\delta = 1/2$, and the tests in (1.10) and (1.14) for several alternatives:

TABLE 2

	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n = 6$	0.2159	0.2077	0.2078	0.1956	0.1908	0.1866	0.1838
$n = 8$	0.3311	0.3234	0.2929	0.2532	0.2578	0.2865	0.2868
$n = 10$	0.4522	0.3996	0.4052	0.3488	0.3460	0.3840	0.3832
$n = 15$	0.6754	0.6255	0.6106	0.5230	0.5625	0.6095	0.5992
	Alternative: exponential distribution						
	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n = 6$	0.0932	0.1112	0.1183	0.1243	0.1242	0.1112	0.1104
$n = 8$	0.1239	0.1459	0.1495	0.1467	0.1527	0.1484	0.1547
$n = 10$	0.1599	0.1610	0.1827	0.1771	0.1847	0.1688	0.1818
$n = 15$	0.2151	0.2173	0.2265	0.2461	0.2593	0.2098	0.2035
	Alternative: double exponential distribution						
	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n = 6$	0.3651	0.3788	0.3919	0.3919	0.3905	0.3796	0.3752
$n = 8$	0.4888	0.5003	0.5045	0.4856	0.4893	0.5060	0.5142
$n = 10$	0.5910	0.6065	0.5880	0.5911	0.5878	0.5738	0.5918
$n = 15$	0.7558	0.7650	0.7415	0.7445	0.7756	0.6705	0.6487
	Alternative: Cauchy distribution						
	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n = 6$	0.1088	0.1315	0.1333	0.1299	0.1346	0.1288	0.1247
$n = 8$	0.1460	0.1727	0.1725	0.1723	0.1794	0.1705	0.1754
$n = 10$	0.1926	0.1925	0.2199	0.2139	0.2246	0.1959	0.2098
$n = 15$	0.2669	0.2681	0.2898	0.2912	0.3122	0.2508	0.2428
	Alternative: Student's t–distribution with three degrees of freedom						
	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n = 6$	0.0553	0.0671	0.0711	0.0688	0.0668	0.0584	0.0631
$n = 8$	0.0682	0.0774	0.0718	0.0767	0.0779	0.0749	0.0716
$n = 10$	0.0743	0.0719	0.0909	0.0770	0.0852	0.0739	0.0978
$n = 15$	0.0836	0.0826	0.0987	0.0958	0.1096	0.0942	0.0946
	Alternative: Student's t–distribution with ten degrees of freedom						

	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n=6$	0.0834	0.0749	0.0689	0.0698	0.0691	0.0624	0.0604
$n=8$	0.1045	0.0954	0.0723	0.0601	0.0622	0.0754	0.0704
$n=10$	0.1346	0.1033	0.0918	0.0719	0.0685	0.0871	0.0872
$n=15$	0.2145	0.1793	0.1620	0.0840	0.1042	0.1572	0.1486
	Alternative: Beta(2, 1) distribution						

	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n=6$	0.0594	0.0668	0.0763	0.0720	0.0743	0.0704	0.0655
$n=8$	0.0678	0.0852	0.0828	0.0739	0.0868	0.0862	0.0787
$n=10$	0.0872	0.0829	0.0991	0.0943	0.0993	0.0876	0.0872
$n=15$	0.1000	0.0964	0.1112	0.1101	0.1405	0.1093	0.0998
	Alternative: Logistic(1) distribution						

	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n=6$	0.0597	0.0460	0.0418	0.0396	0.0404	0.0356	0.0382
$n=8$	0.0676	0.0501	0.0286	0.0208	0.0264	0.0364	0.0325
$n=10$	0.0828	0.0515	0.0262	0.0189	0.0171	0.0361	0.0358
$n=15$	0.1324	0.0802	0.0192	0.0117	0.0125	0.0384	0.0317
	Alternative: uniform distribution						

	SW	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$D_{n,2}^{\text{unb}}(1)$	$E_{n,2}(1)$	$E_{n,2}^{\text{unb}}(1)$
$n=6$	0.3243	0.3045	0.3141	0.2918	0.2924	0.2805	0.2732
$n=8$	0.4683	0.4681	0.4391	0.3570	0.3911	0.4286	0.4293
$n=10$	0.6167	0.5741	0.5713	0.5119	0.5087	0.5523	0.5542
$n=15$	0.8255	0.7931	0.7926	0.7208	0.7420	0.7940	0.7944
	Alternative: lognormal distribution						

The previous table seems to indicate that the presented tests are competitive with the known tests. The presented tests are the most powerful for a double exponential distribution. The test based on $D_{n,2}^{\text{unb}}$ is the test with the highest power for the t distribution with three degrees of freedom and the logistic distribution.

Next, we consider bivariate distributions. The following table shows the values of $na_{n,2,\alpha}$ and $nb_{n,2,\alpha}$ in (1.9) and (1.13), when $d=2$.

	TABLE 3		
$na_{n,2,\alpha}(1)$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.0988706	0.1172691	0.1518514
$n = 8$	0.1157862	0.1519532	0.2304657
$n = 10$	0.1278453	0.1706654	0.2727697
$n = 15$	0.1264138	0.1662850	0.2903303
$nb_{n,2,\alpha}(1)$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.2722242	0.296694	0.3603040
$n = 8$	0.3251454	0.3523247	0.4066532
$n = 10$	0.3850381	0.4099640	0.4633864
$n = 15$	0.5402228	0.5574175	0.6148627

	TABLE 4			
	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.1921	0.1871	0.1780	0.1564
$n = 8$	0.3379	0.3215	0.2896	0.2836
$n = 10$	0.4575	0.4405	0.3778	0.4021
$n = 15$	0.7338	0.7088	0.6372	0.7161
	Alternative: exponential distribution^2			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.0906	0.0940	0.1001	0.1030
$n = 8$	0.1391	0.1538	0.1621	0.1532
$n = 10$	0.1679	0.1861	0.1921	0.1952
$n = 15$	0.2411	0.2722	0.3263	0.2911
	Alternative: double exponential distribution^2			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.3638	0.3931	0.4456	0.3923
$n = 8$	0.6107	0.6178	0.6269	0.6307
$n = 10$	0.7342	0.7238	0.7384	0.7500
$n = 15$	0.9030	0.8878	0.9001	0.8932
	Alternative: Cauchy distribution^2			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.1071	0.1069	0.1240	0.1107
$n = 8$	0.1726	0.1901	0.1905	0.1758
$n = 10$	0.2149	0.2427	0.2486	0.2450
$n = 15$	0.3348	0.3641	0.4052	0.3604
	Alternative: Student's t–distribution^2 with three degrees of freedom			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.0418	0.0571	0.0607	0.0588
$n = 8$	0.0720	0.0733	0.0800	0.0713
$n = 10$	0.0714	0.0866	0.0795	0.0829
$n = 15$	0.0901	0.0988	0.1136	0.1272
	Alternative: Student's t–distribution^2 with ten degrees of freedom			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.0686	0.0603	0.0595	0.0553
$n = 8$	0.0847	0.0782	0.0571	0.0631
$n = 10$	0.1049	0.0788	0.0576	0.0788
$n = 15$	0.1802	0.1049	0.0730	0.1533
	Alternative: Beta$(2, 1)$ distribution^2			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.0575	0.0621	0.0641	0.0553
$n = 8$	0.0758	0.0850	0.0851	0.0791
$n = 10$	0.0861	0.0975	0.0904	0.0916
$n = 15$	0.0984	0.1242	0.1394	0.1420
	Alternative: Logistic(1) distribution^2			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n = 6$	0.0524	0.0403	0.0410	0.0366
$n = 8$	0.0506	0.0332	0.0277	0.0291
$n = 10$	0.0499	0.0231	0.0180	0.0267
$n = 15$	0.0818	0.0128	0.0093	0.0301
	Alternative: uniform distribution^2			

	BHEP(1)	BHEP(1/2)	$D_{n,2}(1)$	$E_{n,2}(1)$
$n=6$	0.2925	0.2942	0.3048	0.2651
$n=8$	0.5368	0.5146	0.4747	0.4734
$n=10$	0.6918	0.6565	0.6068	0.6466
$n=15$	0.9006	0.8465	0.8582	0.8958
	Alternative: lognormal distribution^2			

As before, the previous table shows the presented tests are competitive. The considered bivariate distributions are the ones determined by two i.i.d.r.v.'s. For the double exponential, Cauchy, t distribution and logistic distributions, the test based on $D_{n,2}(1)$ has the highest power.

5. Proofs.

PROOF OF LEMMA 2.1. We have that

$$\begin{aligned}
&\hat{D}_{n,m}\\
=\ & \tfrac{((n-m)!)^2}{(n!)^2}\sum_{(i_1,\dots,i_m)\in I_m^n}\sum_{(j_1,\dots,j_m)\in I_m^n}\int_{\mathbb{R}^d}\exp(im^{-1/2}t'(\textstyle\sum_{k=1}^m Y_{i_k}-\sum_{l=1}^m Y_{j_l}))\phi_\delta(t)\,dt\\
&-2\tfrac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}\int_{\mathbb{R}^d}\exp(im^{-1/2}t'(\textstyle\sum_{j=1}^m Y_{i_j}))\\
&\times(2\pi\delta^2)^{-d/2}\exp(-2^{-1}|t|^2-2^{-1}\delta^{-2}|t|^2)\,dt\\
&+\int_{\mathbb{R}^d}(2\pi\delta^2)^{-d/2}\exp(-|t|^2-2^{-1}\delta^{-2}|t|^2)\,dt.
\end{aligned}$$

Using that for each nonsingular $d\times d$ matrix S,

$$\int_{\mathbb{R}^d}(2\pi)^{-d/2}(\det(S))^{-1/2}\exp(it'\mu-2^{-1}t'S^{-1}t)\,dt=\exp(-2^{-1}\mu'S\mu), \tag{5..1}$$

we get that

$$\begin{aligned}
&\hat{D}_{n,m}\\
=\ & \tfrac{((n-m)!)^2}{(n!)^2}\sum_{(i_1,\dots,i_m)\in I_m^n}\sum_{(j_1,\dots,j_m)\in I_m^n}\exp(-2^{-1}m^{-1}\delta^2|\textstyle\sum_{k=1}^m Y_{i_k}-\sum_{l=1}^m Y_{j_l}|^2)\\
&-2(1+\delta^2)^{-d/2}\tfrac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}\exp(-2^{-1}m^{-1}\delta^2(\delta^2+1)^{-1}|\textstyle\sum_{k=1}^m Y_{i_k}|^2)\\
&+(1+2\delta^2)^{-d/2}.
\end{aligned}$$

□

PROOF OF THEOREM 2.1. Since $\hat{D}_{n,m}(x_1,\dots,x_n)$ depends on $y_j'y_k$, $1\le j,k\le n$, it suffices to show to that for each $1\le j,k\le n$,

$$y_j'y_k=(x_j-\mu(x_1,\dots,x_n))'(S(x_1,\dots,x_n))^{-1}(x_k-\mu(x_1,\dots,x_n))$$

is affine invariant. Given $a\in\mathbb{R}^d$ and $B\in M_{NS}(d\times d)$, we have that for each $1\le j,k\le n$,

$$\begin{aligned}
&(a+Bx_j-\mu(a+Bx_1,\dots,a+Bx_n))'(S(a+Bx_1,\dots,a+Bx_n))^{-1}\\
&\times(x_k-\mu(a+Bx_1,\dots,a+Bx_n))\\
=\ &(x_j-\mu(x_1,\dots,x_n))'B'(BS(x_1,\dots,x_n)B')^{-1}B(x_k-\mu(x_1,\dots,x_n))\\
=\ &(x_j-\mu(x_1,\dots,x_n))'(S(x_1,\dots,x_n))^{-1}(x_k-\mu(x_1,\dots,x_n)).
\end{aligned}$$

$\square$

Proof of Theorem 2.2. It suffices to prove that

$$\begin{aligned} &\tfrac{((n-m)!)^2}{(n!)^2} \textstyle\sum_{(j_1,\dots,j_m)\in I_m^n} \sum_{(k_1,\dots,k_m)\in I_m^n} \\ &\times \exp(-2^{-1}m^{-1}\delta^2(\textstyle\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})'\hat{S}_n^{-1}(\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})) \\ \overset{a.s.}{\to}\ & E\left[\exp\left(-2^{-1}m^{-1}\delta^2(\textstyle\sum_{p=1}^m X_p - \sum_{q=1}^m X_{m+q})'S^{-1}(\sum_{p=1}^m X_p - \sum_{q=1}^m X_{m+q})\right)\right] \end{aligned} \tag{5..2}$$

and

$$\begin{aligned} &\tfrac{(n-m)!}{n!} \textstyle\sum_{(i_1,\dots,i_m)\in I_m^n}\ \exp(-2^{-1}m^{-1}\delta^2(\delta^2+1)^{-1} \\ &\times(\textstyle\sum_{p=1}^m (X_{j_p}-\hat{\mu}_n)'\hat{S}_n^{-1}(\sum_{p=1}^m (X_{j_p}-\hat{\mu}_n))) \\ \overset{a.s.}{\to}\ & E\left[\exp\left(-2^{-1}m^{-1}\delta^2(\delta^2+1)^{-1}(\textstyle\sum_{p=1}^m (X_p-\mu))'S^{-1}(\sum_{p=1}^m (X_p-\mu))\right)\right]. \end{aligned} \tag{5..3}$$

Using that for $a,b\geq 0$ $|e^{-a}-e^{-b}|\leq |a-b|$, we have that

$$\begin{aligned} &\Big|\tfrac{((n-m)!)^2}{(n!)^2} \textstyle\sum_{(j_1,\dots,j_m)\in I_m^n} \sum_{(k_1,\dots,k_m)\in I_m^n} \\ &\times \exp(-2^{-1}m^{-1}\delta^2(\textstyle\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})'\hat{S}_n^{-1}(\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})) \\ &-\tfrac{((n-m)!)^2}{(n!)^2} \textstyle\sum_{(j_1,\dots,j_m)\in I_m^n} \sum_{(k_1,\dots,k_m)\in I_m^n} \\ &\times \exp(-2^{-1}m^{-1}\delta^2(\textstyle\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})'S^{-1}(\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q}))\Big| \\ \leq\ & \tfrac{((n-m)!)^2}{(n!)^2} \textstyle\sum_{(j_1,\dots,j_m)\in I_m^n} \sum_{(k_1,\dots,k_m)\in I_m^n} 2^{-1}m^{-1}\delta^2 \\ &\times\Big|(\textstyle\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})'\hat{S}_n^{-1}(\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q}) \\ &-(\textstyle\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})'S^{-1}(\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})\Big| \\ \leq\ & \tfrac{((n-m)!)^2}{(n!)^2} \textstyle\sum_{(j_1,\dots,j_m)\in I_m^n} \sum_{(k_1,\dots,k_m)\in I_m^n} 2^{-1}m^{-1}\delta^2|\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q}|^2\|\hat{S}_n^{-1}-S^{-1}\| \\ \overset{\text{a.s.}}{\to}\ & 0, \end{aligned} \tag{5..4}$$

because $\|\hat{S}_n^{-1}-S^{-1}\| \overset{\text{a.s.}}{\to} 0$ and

$$\frac{((n-m)!)^2}{(n!)^2}\sum_{(j_1,\dots,j_m)\in I_m^n}\ \sum_{(k_1,\dots,k_m)\in I_m^n} |\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q}|^2 \overset{\text{a.s.}}{\to} E[|\sum_{p=1}^m X_p - \sum_{q=1}^m X_{m+q}|^2]$$

by Theorem 3.1. Again, by Theorem 3.1,

$$\begin{aligned} &\tfrac{((n-m)!)^2}{(n!)^2} \textstyle\sum_{(j_1,\dots,j_m)\in I_m^n} \sum_{(k_1,\dots,k_m)\in I_m^n} \\ &\times \exp(-2^{-1}m^{-1}\delta^2(\textstyle\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})'S^{-1}(\sum_{p=1}^m X_{j_p} - \sum_{q=1}^m X_{k_q})) \\ \overset{a.s.}{\to}\ & E\left[\exp\left(-2^{-1}m^{-1}\delta^2(\textstyle\sum_{p=1}^m X_p - \sum_{q=1}^m X_{m+q})'S^{-1}(\sum_{p=1}^m X_p - \sum_{q=1}^m X_{m+q})\right)\right]. \end{aligned} \tag{5..5}$$

The limit in (5.2) follows from (5.4) and (5.5). The proof of (5.3) is similar and it is omitted. □

PROOF OF THEOREM 2.3. Without loss of generality, we may assume that $E[X]=0$ and $E[(X-E[X])(X-E[X])']=I_d$. We have that

$$\begin{aligned}
&\hat{D}_{n,m}\\
=\ & \int_{\mathbb{R}^d} \tfrac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m)\in I_m^n} \Big(\exp(im^{-1/2}t'\hat{S}_n^{-1/2}(\textstyle\sum_{p=1}^m (X_{i_p}-\hat{\mu}_n))) - e^{-2^{-1}|t|^2}\Big)\\
&\times \tfrac{(n-m)!}{n!} \sum_{(j_1,\dots,j_m)\in I_m^n} \Big(\exp(-im^{-1/2}t'\hat{S}_n^{-1/2}(\textstyle\sum_{q=1}^m (X_{j_q}-\hat{\mu}_n))) - e^{-2^{-1}|t|^2}\Big)\phi_\delta(t)\,dt\\
=\ & \tfrac{((n-m)!)^2}{(n!)^2} \sum_{(i_1,\dots,i_m)\in I_m^n} \sum_{(j_1,\dots,j_m)\in I_m^n}\\
&\times K(m^{-1/2}\hat{S}_n^{-1/2}\textstyle\sum_{p=1}^m (X_{i_p}-\hat{\mu}_n), m^{-1/2}\hat{S}_n^{-1/2}\sum_{q=1}^m (X_{j_q}-\hat{\mu}_n)),
\end{aligned}$$

where

$$\begin{aligned}
&K(x,y) := \int_{\mathbb{R}^d} (\exp(it'x)-\exp(-2^{-1}|t|^2))(\exp(-it'y)-\exp(-2^{-1}|t|^2))\phi_\delta(t)\,dt\\
=\ & \int_{\mathbb{R}^d} (\cos(t'x)+\sin(t'x)-\exp(-2^{-1}|t|^2))(\cos(t'y)+\sin(t'y)-\exp(-2^{-1}|t|^2))\phi_\delta(t)\,dt.
\end{aligned}$$

So,

$$\begin{aligned}
&\hat{D}_{n,m}\\
=\ & \tfrac{((n-m)!)^2}{(n!)^2} \sum_{(i_1,\dots,i_m)\in I_m^n} \sum_{(j_1,\dots,j_m)\in I_m^n} \int_{\mathbb{R}^d} g(X_{i_1},\dots,X_{i_m},t,\hat{\mu}_n,\hat{S}_n)\\
&\times g(X_{j_1},\dots,X_{j_m},t,\hat{\mu}_n,\hat{S}_n)\phi_\delta(t)\,dt,
\end{aligned}$$

where

$$\begin{aligned}
&g(x_1,\dots,x_m,t,\mu,S) := \cos(m^{-1/2}t'S^{-1/2}\textstyle\sum_{j=1}^m (x_j-\mu))\\
&+\sin(m^{-1/2}t'S^{-1/2}\textstyle\sum_{j=1}^m (x_j-\mu)) - \exp(-2^{-1}|t|^2).
\end{aligned}$$

We apply Theorem 3.3 with $g_1=g_2=g$, $d\nu(t)=\varphi_\delta(t)\,dt$, $\hat{\theta}_n=(\hat{\mu}_n,\hat{S}_n)'$ and $\theta_0=(0,I_d)$. It is obvious that conditions (i), (ii) and (vii) in Theorem 3.3 hold. By (5.1),

$$E[\exp(im^{-1/2}t'S^{-1/2}\sum_{j=1}^m (X_j-\mu))] = \exp(-im^{1/2}t'S^{-1/2}\mu - 2^{-1}t'S^{-1}t).$$

So,

$$\begin{aligned}
&G(t,\mu,S) := E[g(X_1,\dots,X_m,t,\mu,S)]\\
=\ & \mathrm{Real}(E[\exp(im^{-1/2}t'S^{-1/2}\textstyle\sum_{j=1}^m (X_j-\mu))]) + \mathrm{Imag}(E[\exp(im^{-1/2}t'S^{-1/2}\sum_{j=1}^m (X_j-\mu))])\\
&-e^{-2^{-1}|t|^2}\\
=\ & e^{-2^{-1}t'S^{-1}t}(\cos(m^{1/2}t'S^{-1/2}\mu)-\sin(m^{1/2}t'S^{-1/2}\mu)) - e^{-2^{-1}|t|^2}.
\end{aligned}$$

Hence, hypothesis (iii) in Theorem 3.3 follows.

Since g is a bounded function, hypotheses (iv), (v) and (xi) in Theorem 3.3 hold.

We claim that $\dot{G}(t,0,I_d)(\mu,S-I_d)' = e^{-2^{-1}|t|^2}(2^{-1}t'(S-I_d)t - m^{1/2}t'\mu)$. We have that

$$\begin{array}{rl} & G(t,\mu,S) - e^{-2^{-1}|t|^2}(2^{-1}t'(S-I_d)t - m^{1/2}t'\mu) \\ = & (e^{-2^{-1}t'S^{-1}t} - e^{-2^{-1}|t|^2}(1-2^{-1}t'(S^{-1}-I_d)t))(\cos(m^{1/2}t'S^{-1/2}\mu) - \sin(m^{1/2}t'S^{-1/2}\mu)) \\ & +e^{-2^{-1}|t|^2}(\cos(m^{1/2}t'S^{-1/2}\mu) - \sin(m^{1/2}t'S^{-1/2}\mu) - 1 + m^{1/2}t'\mu) \\ & +e^{-2^{-1}|t|^2}2^{-1}t'(S^{-1}-I_d)t(-\cos(m^{1/2}t'S^{-1/2}\mu) + \sin(m^{1/2}t'S^{-1/2}\mu) + 1) \\ & +e^{-2^{-1}|t|^2}(-2^{-1}t'(S^{-1}-I_d)t - 2^{-1}t'(S-I_d)t) \\ =: & I + II + III + IV. \end{array}$$

Using that for each $x \in \mathbb{R}$, $|e^x - 1 - x| \le x^2 e^{|x|}$,

$$\begin{array}{rcl} |I| & \le & 2e^{-2^{-1}|t|^2}|e^{-2^{-1}t'(S^{-1}-I_d)t} - 1 + 2^{-1}t'(S^{-1}-I_d)t| \\ & \le & 2^{-1}|t'(S^{-1}-I_d)t|^2 \exp(-2^{-1}|t|^2 + 2^{-1}|t'(S^{-1}-I_d)t|) \\ & \le & 2^{-1}|t|^4\|S^{-1}\|^2\|S-I_d\|^2 \exp(-2^{-1}|t|^2 + 2^{-1}|t|^2\|S^{-1}-I_d\|). \end{array}$$

We have that

$$\begin{array}{rcl} |II| & \le & e^{-2^{-1}|t|^2}\big(|\cos(m^{1/2}t'S^{-1/2}\mu) - 1| + |\sin(m^{1/2}t'S^{-1/2}\mu) - m^{1/2}t'S^{-1/2}\mu| \\ & & +|m^{1/2}t'S^{-1/2}\mu - m^{1/2}t'\mu|\big) \\ & \le & e^{-2^{-1}|t|^2}\big((a+b)(m^{1/2}t'S^{-1/2}\mu)^2 + m^{1/2}|t||\mu|\|S^{-1/2}-1\|\big) \\ & \le & e^{-2^{-1}|t|^2}\big((a+b)m|t|^2\|S^{-1/2}\|^2|\mu|^2 + m^{1/2}|t||\mu|\|S^{-1/2}-1\|\big), \end{array}$$

where $a := \sup_{x\neq 0} x^{-2}|\cos(x)-1|$ and $b := \sup_{x\neq 0} x^{-2}|\sin(x)-x|$. We also have that

$$|III| \le e^{-2^{-1}|t|^2}2^{-1}|t|^2\|S^{-1}\|\|S-I_d\|(am|t|^2\|S^{-1/2}\|^2|\mu|^2 + m^{1/2}|t|\|S^{-1/2}\||\mu|)$$

and

$$|IV| = e^{-2^{-1}|t|^2}2^{-1}|t'(S-I_d)(S^{-1}-I_d)t| \le e^{-2^{-1}|t|^2}2^{-1}|t|^2\|S-I_d\|^2\|S^{-1}\|.$$

It follows from the previous estimations that

$$\lim_{(\mu,S)\to(0,I_d)} (|\mu|^2 + \|S-I_d\|^2)^{-1}\int_{\mathbb{R}^d} |G(t,\mu,S) - e^{-2^{-1}|t|^2}(2^{-1}t'(S-I_d)t - m^{1/2}t'\mu)|^2\phi_\delta(t)\,dt = 0,$$

i.e. hypothesis (vi) in Theorem 3.3 holds.

Since sin and cos are Lipschitz functions hypotheses (viii)–(x) in Theorem 3.3 hold.

Finally, we have that

$$\begin{array}{rl} & h(x,t,\theta_0) = E[g(x,X_2,\ldots,X_m,t,0,I_d)] \\ = & E[\cos(m^{-1/2}t'(x+\sum_{j=2}^m X_j)) + \sin(m^{-1/2}t'(x+\sum_{j=2}^m X_j)) - \exp(-2^{-1}|t|^2)] \\ = & E[\mathrm{Real}(\exp(im^{-1/2}t'(x+\sum_{j=2}^m X_j))] + E[\mathrm{Imag}(\exp(im^{-1/2}t'(x+\sum_{j=2}^m X_j))] - e^{-2^{-1}|t|^2} \\ = & \mathrm{Real}(\exp(im^{-1/2}t'x - 2^{-1}m^{-1}(m-1)t't)) \\ & +\mathrm{Imag}(\exp(im^{-1/2}t'x - 2^{-1}m^{-1}(m-1)t't)) - e^{-2^{-1}|t|^2} \\ = & \exp(-2^{-1}m^{-1}(m-1)|t|^2)\big(\cos(m^{-1/2}t'x) + \sin(m^{-1/2}t'x)\big) - e^{-2^{-1}|t|^2}. \end{array}$$

So,
$$\begin{aligned}
& mh(x,t,\theta_0)+\dot{G}(t,0,I_d)(\alpha(x),\beta(x))' \\
= \; & m\exp(-2^{-1}m^{-1}(m-1)|t|^2)\left(\cos(m^{-1/2}t'x)+\sin(m^{-1/2}t'x)\right) \\
& +\exp(-2^{-1}|t|^2)(2^{-1}t'\beta(x)t-m^{1/2}t'\alpha(x)-m).
\end{aligned}$$
□

The proofs of Lemma 2.2 and theorems 2.4–2.6 are similar to proofs done before and they are omitted.

Proof of Theorem 3.1. By Proposition 5.2.8 in de la Peña and Giné (1999), if $E[|h(X_1,\dots,X_m)|^{m/s}]<\infty$, for some $s>m$, then $n^{m-s}U_{n,m}(h)\stackrel{a.s}{\to}0$. Hence, for each $\max(m_1,m_2)\le k\le m_1+m_2-1$,
$$n^{k-m_1-m_2}U_{n,k}(g_k)\stackrel{a.s}{\to}0.$$
By the strong law of the large numbers for U–statistics
$$U_{n,m_1+m_2}(g_{m_1+m_2})\stackrel{a.s}{\to}E[h(X_1,\dots,X_{m_1+m_2})].$$
(see e.g. Theorem 4.1.4 in de la Peña and Giné, 1999). The claim follows from the previous limits and (3.6). □

Proof of Theorem 3.2. We proceed as in the proof of Theorem 3.1. By Proposition 5.2.8 in de la Peña and Giné (1999), for each $\max(m_1,m_2)\le k\le m_1+m_2-1$,
$$n^{2^{-1}+k-m_1-m_2}U_{n,k}(g_k)\stackrel{a.s}{\to}0.$$
By the central limit theorem for U–statistics
$$n^{1/2}(U_{n,m_1+m_2}(g_{m_1+m_2})-E[U_{n,m_1+m_2}(g_{m_1+m_2})])\stackrel{\mathrm{d}}{\to}N(0,\mathrm{Var}((m_1+m_2)\pi_{1,m_1+m_2}(S_{m_1+m_2}h))),$$
To finish the proof notice that
$$(m_1+m_2)\pi_{1,m_1+m_2}(S_{m_1+m_2}h)(x)=m_1(h_1(x)-E[h_1(X_1)])+m_2(h_2(x)-E[h_2(X_1)]).$$
□

To prove Theorem 3.3, we will need some previous lemmas.

Lemma 5..1. *Let $h_n:S^{m_1+m_2}\to\mathbb{R}$ be a sequence of measurable functions satisfying (3.5). Suppose that:*

(i) For each $n\ge 1$,
$$E[h_n(x,X_2,\dots,X_{m_1+m_2})]=E[h_n(X_1,\dots,X_{m_1},x,X_{m_1+2},\dots,X_{m_1+m_2})]=0\quad P-a.s.$$
(ii) $\lim_{n\to\infty}E[(h_n(X_1,\dots,X_{m_1+m_2}))^2]=0.$

(iii)
$$\lim_{n\to\infty}E[h_n(X_1,X_1,X_2,X_3,\dots,\dots,X_{m_1+m_2-1})]=0,$$

$$\lim_{n\to\infty} E[h_n(X_1,\ldots,X_{m_1},X_1,X_{m_1+1},\ldots,X_{m_1+m_2-1})]=0$$

and

$$\lim_{n\to\infty} E[h_n(X_1,\ldots,X_{m_1},X_{m_1+1},X_{m_1+1},X_{m_1+2},X_{m_1+3},\ldots,X_{m_1+m_2-1})]=0.$$

(iv)

$$\lim_{n\to\infty} n^{-1}E[(h_n(X_1,X_1,X_2,X_3,\ldots,\ldots,X_{m_1+m_2-1}))^2]=0,$$
$$\lim_{n\to\infty} n^{-1}E[(h_n(X_1,\ldots,X_{m_1},X_1,X_{m_1+1},\ldots,X_{m_1+m_2-1}))^2]=0$$

and

$$\lim_{n\to\infty} n^{-1}E[(h_n(X_1,\ldots,X_{m_1},X_{m_1+1},X_{m_1+1},X_{m_1+2},X_{m_1+3},\ldots,X_{m_1+m_2-1}))^2]=0.$$

(v) For each $\max(m_1,m_2)\le k\le m_1+m_2-2$,

$$n^{1+k-m_1-m_2}E[|h_n(X_1,\ldots,X_{m_1},X_1,\ldots,X_{m_1+m_2-k},X_{m_1+1},X_{m_1+2},\ldots,X_k)|]\to 0.$$

Then,

$$n(GU)_{n,m_1,m_2}(h_n)\overset{\Pr}{\to}0.$$

Proof. Let

$$g_{n,k}(x_1,\ldots,x_k)=h_n(x_1,\ldots,x_{m_1},x_1,\ldots,x_{m_1+m_2-k},x_{m_1+1},x_{m_1+2},\ldots,x_k).$$

By the decomposition in (3.6), it suffices to prove that for each $\max(m_1,m_2)\le k\le m_1+m_2$, $n^{1+k-m_1-m_2}U_{n,k}(S_kg_{n,k})\overset{\Pr}{\to}0$. By (i), $U_{n,m_1+m_2}(S_{m_1+m_2}g_{n,m_1+m_2})$ is a U–statistic degenerated of order 1. Hence, by (3.3) and hypothesis (ii),

$$\begin{aligned}&\operatorname{Var}(nU_{n,m_1+m_2}(g_{n,m_1+m_2}))\le cE[(S_{m_1+m_2}g_{n,m_1+m_2}(X_1,\ldots,X_{m_1+m_2}))^2]\\ =\ &cE[(h_n(X_1,\ldots,X_{m_1+m_2}))^2]\to 0\end{aligned}$$

and $nU_{n,m_1+m_2}(g_{n,m_1+m_2})\overset{\Pr}{\to}0$. By (iii) and (iv),

$$E[U_{n,m_1+m_2-1}(g_{n,m_1+m_2-1})]\to 0$$

and

$$\begin{aligned}&\operatorname{Var}(U_{n,m_1+m_2-1}(g_{m_1+m_2-1}))\le cn^{-1}E[((S_{m_1+m_2-1}g_{n,m_1+m_2-1})(X_1,\ldots,X_{2m-1}))^2]\\ \le\ &cn^{-1}(E[(h_n(X_1,X_1,X_2,X_3,\ldots,\ldots,X_{m_1+m_2-1}))^2]\\ &+E[(h_n(X_1,\ldots,X_{m_1},X_1,X_{m_1+1},\ldots,X_{m_1+m_2-1}))^2]\\ &+E[(h_n(X_1,\ldots,X_{m_1},X_{m_1+1},X_{m_1+1},X_{m_1+2},X_{m_1+3},\ldots,X_{m_1+m_2-1}))^2])\to 0.\end{aligned}$$

Hence, $U_{n,m_1+m_2-1}(g_{n,m_1+m_2-1})\overset{\Pr}{\to}0$. By (v), for each $m_2\le k\le m_1+m_2-2$,

$$E[|n^{1+k-m_1-m_2}U_{n,k}(g_{n,k})|]\le n^{1+k-m_1-m_2}E[|g_{n,k}(X_1,\ldots,X_k)|]\to 0$$

and $n^{1+k-m_1-m_2}U_{n,k}(g_k)\overset{\Pr}{\to}0$. $\square$

Lemma 5..2. *Let $h_n : S^{m_1+m_2} \to \mathbb{R}$ be a sequence of measurable functions satisfying (3.5). Suppose that:*

(i) For each $n \geq 1$,

$$\begin{aligned} &E[h_n(x, y, X_3, \dots, X_{m_1+m_2})] = E[h_n(x, X_2, \dots, X_{m_1}, y, X_{m_1+2}, \dots, X_{m_1+m_2})] \\ &= E[h(X_1, \dots, X_{m_1}, x, y, X_{m_1+3}, \dots, X_{m_1+m_2})] = 0 \ (P \times P) - a.s. \end{aligned}$$

(ii) $\lim_{n\to\infty} n^{-1} E[(h_n(X_1, \dots, X_{m_1+m_2}))^2] = 0.$

(iii)

$$\lim_{n\to\infty} n^{-1} E[(h_n(X_1, X_1, X_2, X_3, \dots, \dots, X_{m_1+m_2-1}))^2] = 0,$$

$$\lim_{n\to\infty} n^{-1} E[(h_n(X_1, \dots, X_{m_1}, X_1, X_{m_1+1}, \dots, X_{m_1+m_2-1}))^2] = 0$$

and

$$\lim_{n\to\infty} n^{-1} E[(h_n(X_1, \dots, X_{m_1}, X_{m_1+1}, X_{m_1+1}, X_{m_1+2}, X_{m_1+3}, \dots, X_{m_1+m_2-1}))^2] = 0.$$

(iv) For each $max(m_1, m_2) \leq k \leq m_1 + m_2 - 2$,

$$n^{1+k-m_1-m_2} E[|h_n(X_1, \dots, X_{m_1}, X_1, \dots, X_{m_1+m_2-k}, X_{m_1+1}, X_{m_1+2}, \dots, X_k)|] \to 0.$$

Then,

$$n(GU)_{n,m_1,m_2}(h_n) \xrightarrow{\Pr} 0.$$

Proof. The proof of this lemma follows similarly to that of Lemma 5.1. The difference is that in the present situation, $U_{n,m_1+m_2}(g_{n,m_1+m_2})$ is a U–statistic degenerate of order 2. So,

$$\begin{aligned} &\mathrm{Var}(nU_{n,m_1+m_2}(g_{n,m_1+m_2})) \leq cn^{-1} E[(S_{m_1+m_2} g_{n,m_1+m_2}(X_1, \dots, X_{m_1+m_2}))^2] \\ = \ &cn^{-1} E[(h_n(X_1, \dots, X_{m_1+m_2}))^2] \to 0. \end{aligned}$$

The limit $n^{1+k-m_1-m_2} U_{n,k}(S_k g_{n,k}) \xrightarrow{\Pr} 0$, for each $m_2 \leq k \leq m_1 + m_2 - 1$ follows exactly as in Lemma 5.1. $\square$

Lemma 5..3. *Let $\{X_j\}$ be a sequence of i.i.d.r.v.'s with values in the measurable space $(S, \mathcal{S})$. Let $(T, \mathcal{T}, \nu)$ be a measure space. Let $\Theta \subset \mathbb{R}^s$. Let $g : S^m \times T \times \Theta \to \mathbb{R}$ be a measurable function. Let $\epsilon_0 > 0$. Suppose that the following conditions are satisfied:*

(i) For each $t \in T$, each $|\theta| \leq \epsilon_0$, each $x_1, \dots, x_{m_1} \in S$ and each $\sigma \in \mathrm{Perm}(\{1, \dots, m\})$,

$$g(x_1, \dots, x_m, t, \theta) = g(x_{\sigma(1)}, \dots, x_{\sigma(m)}, t, \theta).$$

(ii) For each $t \in T$ and each $|\theta| \leq \epsilon_0$,

$$\int_T E[|g(X_1, \dots, X_m, t, \theta)|] \, d\nu(t) < \infty$$

and

$$E[g(X_1,\ldots,X_m,t,\theta)]=0.$$

(iii) For each $x_1,\ldots,x_m,t$,

$$g(x_1,\ldots,x_m,t,0)=0.$$

(iv) For each $0<\delta<\epsilon_0$,

$$\int_T E[\sup_{|\theta|\le\epsilon_0,|\gamma|\le\delta,|\theta+\gamma|\le\epsilon_0}|g(X_1,\ldots,X_m,t,\theta+\gamma)-g(X_1,\ldots,X_m,t,\theta)|^2]\,d\nu(t)\le A\delta^2.$$

where A *is a finite constant.*

(v) For each $|\theta|\le\epsilon_0$,

$$E\left[\left(\int_T g(X_1,\ldots,X_m,t,n^{-1/2}\theta)g(X_{m+1},\ldots,X_{2m},t,n^{-1/2}\theta)\,d\nu(t)\right)^2\right]\to 0.$$

(vi) For each $|\theta|\le\epsilon_0$,

$$E\left[\int_T g(X_1,X_1,\ldots,X_{m-1},t,n^{-1/2}\theta)g(X_m,X_{m+1},\ldots,X_{2m-1},t,n^{-1/2}\theta)\,d\nu(t)\right]\to 0$$

and

$$E\left[\int_T g(X_1,\ldots,X_m,t,n^{-1/2}\theta)g(X_1,X_{m+1},\ldots,X_{2m-1},t,n^{-1/2}\theta)\,d\nu(t)\right]\to 0.$$

(vii) For each $|\theta|\le\epsilon_0$,

$$n^{-1}E\left[\left(\int_T (g(X_1,X_1,X_2,X_3,\ldots,X_{m-1},t,n^{-1/2}\theta)g(X_m,\ldots,X_{2m-1},t,n^{-1/2}\theta)\,d\nu(t)\right)^2\right]\to 0$$

and

$$n^{-1}E\left[\left(\int_T (g(X_1,\ldots,X_m,t,n^{-1/2}\theta)g(X_1,X_{m+1},X_{m+2},\ldots,X_{2m-1},t,n^{-1/2}\theta)\,d\nu(t)\right)^2\right]\to 0.$$

(viii) For each $m\le k\le 2m-2$ *and each* $|\theta|\le\epsilon_0$,

$$n^{1+k-2m}E[|\int_T g(X_1,\ldots,X_m,t,n^{-1/2}\theta)g(X_1,\ldots,X_{2m-k},X_{m+1},\ldots,X_k,t,n^{-1/2}\theta)\,d\nu(t)|]\to 0.$$

Then,

$$\sup_{|\theta|\le\epsilon_0}\|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n}g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta))\|_{L_2(\nu)}\stackrel{\Pr}{\to}0.$$

Proof. Given $\tau, \eta > 0$, we prove that

$$\limsup_{n\to\infty} \Pr\{\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta)\|_{L_2(\nu)} \ge \tau\} \le \eta.$$

Take $\delta > 0$ such that $2^2 A\delta^2\tau^{-2} < \eta$. By compactness, there exists a function $\pi : \overline{B_s(0,\epsilon_0)} \to \overline{B_s(0,\epsilon_0)}$ with finite range such that for each $\theta \in \overline{B_s(0,\epsilon_0)}$, $|\theta - \pi(\theta)| \le \delta$. Then,

$$\begin{array}{rl} & \sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta)\|_{L_2(\nu)} \\ \le & \sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} (g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta) \\ & -g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta)))\|_{L_2(\nu)} \\ & +\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta))\|_{L_2(\nu)}. \end{array}$$

Hence,

$$\begin{array}{rl} & \Pr\{\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta)\|_{L_2(\nu)} \ge \tau\} \\ \le & \Pr\{\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} (g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta) \\ & -g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta)))\|_{L_2(\nu)} \ge 2^{-1}\tau\} \\ & +\Pr\{\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta))\|_{L_2(\nu)} \ge 2^{-1}\tau\} \\ \le & 4\tau^{-2}E[\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} (g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta) \\ & -g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta)))\|_{L_2(\nu)}^2] \\ & +\sup_{|\theta|\le\epsilon_0} \Pr\{\|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta))\|_{L_2(\nu)} \ge 2^{-1}\tau\}. \end{array}$$

By (iv),

$$\begin{array}{rl} & E[\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} (g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta) \\ & -g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta)))\|_{L_2(\nu)}^2] \\ = & E[\sup_{|\theta|\le\epsilon_0} n\frac{((n-m)!)^2}{(n!)^2}\sum_{(i_1,\ldots,i_m)\in I_m^n}\sum_{(j_1,\ldots,j_m)\in I_m^n} \\ & \times\int_T (g(X_{i_1},\ldots,X_{i_m},t,n^{-1/2}\theta) - g(X_{i_1},\ldots,X_{i_m},t,n^{-1/2}\pi(\theta))) \\ & \times(g(X_{j_1},\ldots,X_{j_m},t,\theta) - g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta)))\,d\nu(t)] \\ \le & E[n\frac{((n-m)!)^2}{(n!)^2}\sum_{(i_1,\ldots,i_m)\in I_m^n}\sum_{(j_1,\ldots,j_m)\in I_m^n} \\ & \times\int_T \sup_{|\theta|\le\epsilon_0} |g(X_{i_1},\ldots,X_{i_m},t,\theta) - g(X_{i_1},\ldots,X_{i_m},t,\pi(\theta))| \\ & \times\sup_{|\theta|\le\epsilon_0} |g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta) - g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta))|\,d\nu(t)] \\ = & n\int_T E[\sup_{|\theta|\le\epsilon_0} |g(X_1,\ldots,X_m,t,n^{-1/2}\theta) - g(X_1,\ldots,X_m,t,n^{-1/2}\pi(\theta))|^2]\,d\nu(t) \\ \le & A\delta^2 \end{array}$$

and

$$\begin{array}{rl} & 4\tau^{-2}E[\sup_{|\theta|\le\epsilon_0} \|n^{1/2}\frac{(n-m)!}{n!}\sum_{(j_1,\ldots,j_m)\in I_m^n} (g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta) \\ & -g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\pi(\theta)))\|_{L_2(\nu)}^2] \\ \le & 4\tau^{-2}A\delta^2 \le \eta. \end{array}$$

To finish the proof it suffices to prove that for each $|\theta| \le \epsilon_0$,

$$\begin{aligned}
& \|n^{1/2}\tfrac{(n-m)!}{n!}\textstyle\sum_{(j_1,\ldots,j_m)\in I_m^n} g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta)\|^2_{L_2(\nu)} \\
= \;& n\tfrac{((n-m)!)^2}{(n!)^2}\textstyle\sum_{(i_1,\ldots,i_m)\in I_m^n}\sum_{(j_1,\ldots,j_m)\in I_m^n}\int_T g(X_{i_1},\ldots,X_{i_m},t,n^{-1/2}\theta) \\
& \times g(X_{j_1},\ldots,X_{j_m},t,n^{-1/2}\theta)\,d\mu(t) \\
& \stackrel{\Pr}{\to} 0.
\end{aligned}$$

We claim that this follows from Lemma 5.1 applied to
$h_n(x_1,\ldots,x_m,x_{m+1},\ldots,x_{2m}) = \int_T g(x_1,\ldots,x_m,t,n^{-1/2}\theta)g(x_{m+1},\ldots,x_{2m},t,n^{-1/2}\theta)\,d\nu(t)$. Hypothesis (i) in Lemma 5.1 follows from (ii). Hypothesis (ii)–(v) in Lemma 5.1 follow from (v)–(viii), respectively. □

Proof of Theorem 3.3. Using that for each $x, y \in \mathbb{R}$, $4xy = (x+y)^2 - (x-y)^2$, to prove (3.7), it suffices to prove that

$$\begin{aligned}
& n\int_T \Bigg(\frac{(n-m_1)!}{n!}\sum_{(i_1,\ldots,i_{m_1})\in I_{m_1}^n} g_1(X_{i_1},\ldots,X_{i_{m_1}},t,\hat\theta_n) \\
& +\frac{(n-m_2)!}{n!}\sum_{(j_1,\ldots,j_{m_2})\in I_{m_2}^n} g_2(X_{j_1},\ldots,X_{j_{m_2}},t,\hat\theta_n)\Bigg)^2\,d\nu(t) \\
& -n\int_T\Bigg(n^{-1}\sum_{j=1}^n (m_1h_1(X_j,t,\theta_0)+(\dot G_1(t,\theta_0))'\alpha(X_j) \\
& +m_2h_2(X_j,\theta_0)+(\dot G_2(t,\theta_0))'\alpha(X_j))\Bigg)^2\,d\nu(t) \\
& \stackrel{\Pr}{\to} 0
\end{aligned} \tag{5..6}$$

and

$$\begin{aligned}
& n\int_T \Bigg(\frac{(n-m_1)!}{n!}\sum_{(i_1,\ldots,i_{m_1})\in I_{m_1}^n} g_1(X_{i_1},\ldots,X_{i_{m_1}},t,\hat\theta_n) \\
& -\frac{(n-m_2)!}{n!}\sum_{(j_1,\ldots,j_{m_2})\in I_{m_2}^n} g_2(X_{j_1},\ldots,X_{j_{m_2}},t,\hat\theta_n)\Bigg)^2\,d\nu(t) \\
& -n\int_T\Bigg(n^{-1}\sum_{j=1}^n (m_1h_1(X_j,t,\theta_0)+(\dot G_1(t,\theta_0))'\alpha(X_j) \\
& -m_2h_2(X_j,t,\theta_0)-(\dot G_2(t,\theta_0))'\alpha(X_j))\Bigg)^2\,d\nu(t) \\
& \stackrel{\Pr}{\to} 0.
\end{aligned} \tag{5..7}$$

By the central limit theorem for degenerate V–statistics in (2.4),

$$\begin{aligned}
& n\int_T \Big(n^{-1}\textstyle\sum_{j=1}^n (m_1 h_1(X_j,t,\theta_0) + (\dot G_1(t,\theta_0))'\alpha(X_j) \\
& \quad + m_2 h_2(X_j,\theta_0) + (\dot G_2(t,\theta_0))'\alpha(X_j))\Big)^2 \, d\nu(t) \\
= \; & n^{-1}\textstyle\sum_{j,k=1}^n \int_T (m_1 h_1(X_j,t,\theta_0) + (\dot G_1(t,\theta_0))'\alpha(X_j) + m_2 h_2(X_j,\theta_0) + (\dot G_2(t,\theta_0))'\alpha(X_j)) \\
& \times (m_1 h_1(X_k,t,\theta_0) + (\dot G_1(t,\theta_0))'\alpha(X_k) + m_2 h_2(X_k,\theta_0) + (\dot G_2(t,\theta_0))'\alpha(X_k))\, d\nu(t) \\
= \; & O_P(1).
\end{aligned}$$

Hence, to prove (5.6), it suffices to prove that

$$\begin{aligned}
& n^{1/2}\|\tfrac{(n-m_1)!}{n!}\textstyle\sum_{(i_1,\ldots,i_{m_1})\in I^n_{m_1}} g_1(X_{i_1},\ldots,X_{i_{m_1}},t,\hat\theta_n) \\
& + \tfrac{(n-m_2)!}{n!}\textstyle\sum_{(j_1,\ldots,j_{m_2})\in I^n_{m_2}} g_2(X_{j_1},\ldots,X_{j_{m_2}},t,\hat\theta_n)\|_{L_2(\nu)} \\
& - n^{1/2}\|n^{-1}\textstyle\sum_{j=1}^n (m_1 h_1(X_j,t,\theta_0) + (\dot G_1(t,\theta_0))'\alpha(X_j) \\
& + m_2 h_2(X_j,t,\theta_0) + (\dot G_2(t,\theta_0))'\alpha(X_j))\|_{L_2(\nu)} \\
& \xrightarrow{\Pr} 0.
\end{aligned} \tag{5..8}$$

We have that

$$\begin{aligned}
& |n^{1/2}\|\tfrac{(n-m_1)!}{n!}\textstyle\sum_{(i_1,\ldots,i_{m_1})\in I^n_{m_1}} g_1(X_{i_1},\ldots,X_{i_{m_1}},t,\hat\theta_n) \\
& + \tfrac{(n-m_2)!}{n!}\textstyle\sum_{(j_1,\ldots,j_{m_2})\in I^n_{m_2}} g_2(X_{j_1},\ldots,X_{j_{m_2}},t,\hat\theta_n)\|_{L_2(\nu)} \\
& - n^{-1/2}\|n^{-1}\textstyle\sum_{j=1}^n (m_1 h_1(X_j,t,\theta_0) + (\dot G_1(t,\theta_0))'\alpha(X_j) \\
& + m_2 h_2(X_j,t,\theta_0) + (\dot G_2(t,\theta_0))'\alpha(X_j))\|_{L_2(\nu)}| \\
\le \; & n^{1/2}\|\tfrac{(n-m_1)!}{n!}\textstyle\sum_{(i_1,\ldots,i_{m_1})\in I^n_{m_1}} (g_1(X_{j_1},\ldots,X_{j_{m_1}},t,\hat\theta_n) \\
& - g_1(X_{j_1},\ldots,X_{j_{m_1}},t,\theta_0) - G_1(t,\hat\theta_n))\|_{L_2(\nu)} \\
& + n^{1/2}\|\tfrac{(n-m_1)!}{n!}\textstyle\sum_{(i_1,\ldots,i_{m_1})\in I^n_{m_1}} g_1(X_{i_1},\ldots,X_{i_{m_1}},t,\theta_0) - n^{-1}\sum_{k=1}^n m_1 h_1(X_k,t,\theta_0)\|_{L_2(\nu)} \\
& + n^{1/2}\|G_1(t,\hat\theta_n) - (\dot G_1(t,\theta_0))'(\hat\theta_n-\theta_0)\|_{L_2(\nu)} \\
& + \|n^{1/2}(\dot G_1(t,\theta_0))'(\hat\theta_n-\theta_0) - n^{-1/2}\textstyle\sum_{j=1}^n (\dot G_1(t,\theta_0))'\alpha(X_j)\|_{L_2(\nu)} \\
& + n^{1/2}\|\tfrac{(n-m_2)!}{n!}\textstyle\sum_{(i_1,\ldots,i_{m_2})\in I^n_{m_2}} (g_2(X_{i_1},\ldots,X_{i_{m_2}},t,\hat\theta_n) \\
& - g_2(X_{i_1},\ldots,X_{i_{m_2}},t,\theta_0) - G_2(t,\hat\theta_n))\|_{L_2(\nu)} \\
& + n^{1/2}\|\tfrac{(n-m_2)!}{n!}\textstyle\sum_{(i_1,\ldots,i_{m_2})\in I^n_{m_2}} g_2(X_{i_1},\ldots,X_{i_{m_2}},t,\theta_0) - n^{-1}\sum_{k=1}^n m_2 h_2(X_k,t,\theta_0)\|_{L_2(\nu)} \\
& + n^{1/2}\|G_2(t,\hat\theta_n) - (\dot G_2(t,\theta_0))'(\hat\theta_n-\theta_0)\|_{L_2(\nu)} \\
& + \|n^{1/2}(\dot G_2(t,\theta_0))'(\hat\theta_n-\theta_0) - n^{-1/2}\textstyle\sum_{j=1}^n (\dot G_2(t,\theta_0))'\alpha(X_j)\|_{L_2(\nu)} \\
& \xrightarrow{\Pr} 0.
\end{aligned} \tag{5..9}$$

Since a similar estimation holds for (5.7), to finish the proof, we need to show that for each

$l = 1, 2,$

$$n^{1/2} \| \frac{(n-m_l)!}{n!} \sum_{(i_1,\dots,i_{m_l}) \in I^n_{m_l}} (g_l(X_{j_1}, \dots, X_{j_{m_l}}, t, \hat{\theta}_n) \\ - g_l(X_{j_1}, \dots, X_{j_{m_l}}, t, \theta_0) - G_l(t, \hat{\theta}_n)) \|_{L_2(\nu)} \stackrel{\Pr}{\to} 0. \tag{5..10}$$

$$n^{1/2} \| \frac{(n-m_l)!}{n!} \sum_{(i_1,\dots,i_{m_l}) \in I^n_{m_l}} g_l(X_{i_1}, \dots, X_{i_{m_l}}, t, \theta_0) - n^{-1} \sum_{k=1}^n m_l h_l(X_k, \theta_0) \|_{L_2(\nu)} \stackrel{\Pr}{\to} 0. \tag{5..11}$$

$$n^{1/2} \| G_l(t, \hat{\theta}_n) - (\dot{G}_l(t, \theta_0))' (\hat{\theta}_n - \theta_0) \|_{L_2(\nu)} \stackrel{\Pr}{\to} 0. \tag{5..12}$$

$$\| n^{1/2} (\dot{G}_l(t, \theta_0))' (\hat{\theta}_n - \theta_0) - n^{-1/2} \sum_{j=1}^n (\dot{G}_l(t, \theta_0))' \alpha(X_j) \|_{L_2(\nu)} \stackrel{\Pr}{\to} 0. \tag{5..13}$$

To prove (5.10), it suffices to prove that for each $l = 1, 2,$

$$n \frac{((n-m_l)!)^2}{(n!)^2} \sum_{(i_1,\dots,i_{m_l}) \in I^n_{m_l}} \sum_{(j_1,\dots,j_{m_l}) \in I^n_{m_l}} \tag{5..14}$$
$$\times \int_T (g_l(X_{i_1}, \dots, X_{i_{m_l}}, t, \hat{\theta}_n) - g_l(X_{i_1}, \dots, X_{i_{m_l}}, t, \theta_0) - G_l(t, \hat{\theta}_n))$$
$$\times (g_l(X_{j_1}, \dots, X_{j_{m_l}}, t, \hat{\theta}_n) - g_l(X_{j_1}, \dots, X_{j_{m_l}}, t, \theta_0) - G_l(t, \hat{\theta}_n)) \, d\nu(t) \stackrel{\Pr}{\to} 0.$$

We will prove that for each $0 < M < \infty$,

$$\sup_{|\theta| \le M} | n \frac{((n-m_l)!)^2}{(n!)^2} \sum_{(i_1,\dots,i_{m_l}) \in I^n_{m_l}} \sum_{(j_1,\dots,j_{m_l}) \in I^n_{m_l}} \tag{5..15}$$
$$\times \int_T (g_l(X_{i_1}, \dots, X_{i_{m_l}}, t, \theta_0 + n^{-1/2}\theta) - g_l(X_{i_1}, \dots, X_{i_{m_l}}, t, \theta_0) - G_l(t, \theta_0 + n^{-1/2}\theta))$$
$$\times (g_l(X_{j_1}, \dots, X_{j_{m_l}}, t, \theta_0 + n^{-1/2}\theta) - g_l(X_{j_1}, \dots, X_{j_{m_l}}, t, \theta_0) - G_l(t, \theta_0 + n^{-1/2}\theta)) \, d\nu(t) |$$
$$\stackrel{\Pr}{\to} 0.$$

(5.14) follows plugging $n^{1/2}(\hat{\theta}_n - \theta_0)$ as θ in (5.15). We claim that (5.15) follows from Lemma 5.3 with

$$g(x_1, \dots, x_{m_l}, t, \theta) = g_l(x_1, \dots, x_{m_l}, t, \theta_0 + \theta) - g_l(x_1, \dots, x_{m_l}, t, \theta_0) - G_l(t, \theta_0 + \theta).$$

Hypothesis (i)–(iv) in Lemma 5.3 are obviously satisfied. Hypothesis (v) and (vii) in Lemma 5.3 follow from (x). Hypothesis (vi) in Lemma 5.3 follows from (ix). Hypothesis (viii) in Lemma 5.3 follows from (xi).

Display (5.11) follows from Lemma 5.2 with

$$h_n(x_1, \dots, x_{2m_l}) = \int_T (g_l(x_1, \dots, x_{m_l}, t, \theta_0) - \sum_{p=1}^{m_l} h_l(x_p, t, \theta_0)) \\ \times (g_l(x_{m_l+1}, \dots, x_{2m_l}, t, \theta_0) - \sum_{p=1}^{m_l} h(x_{m_l+p}, t, \theta_0)) \, d\nu(t),$$

using (iv) and (v). (5.12) follows from (vi) and (vii). (5.13) follows from (vi) and (vii). □

References

[1] M. A. Arcones and Y. Wang, Some new tests for normality based on U–processes. *Statist. Probab. Lett.*, **76** (2006), pp. 69–82.

[2] L. Baringhaus, and N. Henze, A consistent test for multivariate normality based on the empirical characteristic function. *Metrika*, **35** (1988), pp. 339–348.

[3] P. L. Davies, Asymptotic behavior of S–estimates of multivariate dispersion matrices. *Ann. Statist.*, **14** (1987), pp. 1269–1292.

[4] V. H. de la Peña, and E. Giné, *Decoupling. From Dependence to Independence.* Springer–Verlag, New York, 1999.

[5] T. de Wet, and R. H. Randles, On the effect of substituting parameter estimators in limiting χ^2 U and V statistics. *Ann. Statist.*, **15** (1987), pp. 398–412.

[6] M. L. Eaton and M. D. Perlman, The non–singularity of generalized sample covariance matrices. *Ann. Statist.*, **1** (1973), pp. 710–717.

[7] T. W. Epps and L. B. Pulley, A test for normality based on the empirical characteristic function. *Biometrika*, **70** (1983), pp. 723–726.

[8] N. Henze, Invariant tests for multivariate normality: a critical review. *Statist. Pap.*, **43** (2002), pp. 467–506.

[9] N. Henze and B. Zirkler, A class of invariant consistent tests for multivariate normality. *Comm. Statist. Theory Methods*, **19** (1990), pp. 3595–3617.

[10] J. Kent, and D. Tyler, Redescending M–estimates of multivariate location and scatter. *Ann. Statist.*, **19** (1991), pp. 2102–2119.

[11] A. J. Lee, *U–statistics, Theory and Practice.* Marcel Dekker, New York, 1990.

[12] M. Loève, *Probability Theory I*, 4–th Edition. Springer, New York, 1977.

[13] H. P. Lopuhaä, and P. J. Rousseeuw, Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Ann. Statist.*, **19** (1991), pp. 229–248.

[14] R. Maronna, Robust M–estimators of multivariate scatter. *Ann. Statist.*, **4** (1976), 51–67.

[15] C. J. Mecklin and D. J. Mundfrom, An appraisal and bibliography of tests for multivariate normality. *Internat. Statist. Rev.*, **72** (2004), pp. 123-138.

[16] S. S. Shapiro and M. B. Wilk, An analysis of variance test for normality: complete samples. *Biometrika*, **52** (1965), pp. 591–611.

[17] S. S. Shapiro and M. B. Wilk, Approximations for the null distribution of the W statistic. *Technometrics*, **10** (1968), pp. 861–866.

[18] H. C. Thode, Jr., *Testing for Normality.* Marcel Dekker, New York, 2002.