A normality test for the errors of the linear model *

Miguel A. Arcones

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

E-mail:arcones@math.binghamton.edu

January 27, 2008

Abstract

We present a normality test for the errors of the linear regression model based on the normality test in Epps and Pulley (1983). We prove that the presented test is consistent against any alternative. We show that the test statistic under the null hypothesis is asympotically equivalent to a degenerate V–statistic. We also consider the asymptotic distribution of the test under contiguous alternatives.

1 Introduction

Many of the statistical methods in linear regression assume that the residuals have a normal distribution. In this paper, we study a normality test for the residuals of a linear regression model. Many authors have studied normality tests for (independent identically distributed) i.i.d. observations. Reviews of normality tests are Henze (2002) and Mecklin and Mundfrom (2004). Classical normality tests are the ones by Shapiro and Wilk (1965, 1968) and Epps and Pulley (1983). Several authors have considered tests of normality for the linear regression model. Jurečková, Picek, and Sen (2003) and Sen, Jurečková, and Picek (2003) considered a test of normality for the residuals of a linear regression model using the approach in Shapiro and Wilk. Here, we apply the normality test in Epps and Pulley (1983) to the linear regression model. The asymptotic distribution of the test statistic of the Epps and Pulley test under (independent identically distributed random variables) i.i.d.r.v.'s was obtained by Baringhaus and Henze (1988). The test in Epps and Pulley (1983) is known in the literature as the BHEP test.

* *Key words and phrases:* normality test, linear regression.

Suppose that $X_1, \ldots, X_n$ are i.i.d.r.v.'s from a (cumulative distribution function) c.d.f. F. Assume that X_1 has a second finite moment, then we may define $\mu = E[X_1]$ and $\sigma^2 = E[(X_1 - \mu)^2]$. We have that X_1 has a nondegenerate normal distribution if and only if $\sigma^{-1}(X_1 - \mu)$ has a standard normal distribution. This is equivalent to $E[\exp(it\sigma^{-1}(X_1 - \mu))] = \exp(-2^{-1}t^2)$, for each $t \in \mathbb{R}$. The BHEP test is significative for nonnormality for large values of the statistic

$$T_n := n \int_{\mathbb{R}} |n^{-1} \sum_{j=1}^{n} \exp(it\hat{\sigma}_n^{-1}(X_j - \bar{X}_n)) - \exp(-2^{-1}t^2)|^2 \phi_\delta(t)\, dt, \tag{1.1}$$

where $\bar{X}_n = n^{-1} \sum_{j=1}^n X_j$, $\hat{\sigma}_n^2 = n^{-1} \sum_{j=1}^n (X_j - \bar{X}_n)^2$ and $\phi_\delta(t) = (2\pi\delta^2)^{-1/2} \exp(-2^{-1}\delta^{-2}t^2)$. Observe that in (1.1) the modulus of the complex number $n^{-1} \sum_{j=1}^n \exp(it\hat{\sigma}_n^{-1}(X_j - \bar{X}_n)) - \exp(-2^{-1}t^2)$ is taken. From now, we will call the normality test based on the statistic in (1.1), the BHEP(δ) test. ϕ_δ is the (probability density function) p.d.f. of a normal r.v. with mean zero and variance $\delta^2 > 0$. By tuning the parameter δ 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 will use the usual multivariate notation. For example, given $u = (u_1, \ldots, u_d)' \in \mathbb{R}^d$ and $v = (v_1, \ldots, v_d)' \in \mathbb{R}^d$, $u'v = \sum_{j=1}^d u_j v_j$ and $|u| = (\sum_{j=1}^n u_j^2)^{1/2}$. Given a $d \times d$ matrix A, $\|A\| = \sup_{v_1, v_2 \in \mathbb{R}^d, |v_1|, |v_2| = 1} v_1' A v_2$ and $\operatorname{trace}(A) = \sum_{j=1}^d a_{j,j}$, where $A = (a_{j,k})_{1 \le j,k \le d}$. Given $\theta \in \mathbb{R}^d$ and $\epsilon > 0$, $B(\theta, \epsilon) = \{x \in \mathbb{R}^d : |x - \theta| < \epsilon\}$. I_d will denote the identity matrix in $\mathbb{R}^d$. c will denote an universal constant which may change from occurrence to occurrence.

In this paper, we present a test of normality for the errors of a linear regression model. We consider the linear regression model: $Y_j = x_{n,j}'\beta + \epsilon_j$, $j \ge 1$, where $\{\epsilon_j\}_{j=1}^\infty$ is a sequence of i.i.d.r.v.'s with mean zero; $x_{n,j}$, $1 \le j \le n$ are p dimensional vectors and $\beta \in \mathbb{R}^p$ is a parameter to be estimated. $x_{n,j}$ is called the regressor or predictor variable. Y_j is called the response variable. ϵ_j is an error variable. Let F be the c.d.f. of the sequence $\{\epsilon_j\}$. We assume that F has mean zero. We study the testing problem

$$H_0 : F \text{ has a normal distribution, versus } H_1 : F \text{ does not.} \tag{1.2}$$

Assuming the condition:

(A) $A_n := \sum_{j=1}^n x_{n,j} x_{n,j}'$ is a nonsingular $p \times p$ matrix,

the least squares estimator of β is $\hat{\beta}_n = A_n^{-1}(\sum_{j=1}^n Y_j x_{n,j})$. To estimate the distribution F of the errors we use the residuals $\hat{\epsilon}_j = Y_j - \hat{\beta}_n' x_{n,j}$, $1 \le j \le n$. Let

$$\begin{aligned} \hat{D}_{n,\delta} &:= \int_{\mathbb{R}} |n^{-1} \sum_{j=1}^n \exp(it\hat{\sigma}_n^{-1}\hat{\epsilon}_j) - \exp(-2^{-1}t^2)|^2 \phi_\delta(t)\, dt \\ &= \|n^{-1} \sum_{j=1}^n \exp(it\hat{\sigma}_n^{-1}\hat{\epsilon}_j) - \exp(-2^{-1}t^2)\|_{L_2(\phi_\delta)}^2, \end{aligned} \tag{1.3}$$

where $\hat{\sigma}_n^2 := (n-p)^{-1}\sum_{j=1}^n \hat{\epsilon}_j^2$. As it is well known $\hat{\sigma}_n^2$ is an unbiased estimator of $\sigma_F^2 := \mathrm{Var}_F(\epsilon)$. We abbreviate $\hat{D}_{n,\delta}$ to $\hat{D}_n$. Note that $\hat{D}_n$ estimates

$$D_F := \int_{\mathbb{R}} |E_F[\exp(it\sigma_F^{-1}\epsilon)] - \exp(-2^{-1}t^2)|^2 \phi_\delta(t)\,dt.$$

Given $1 > \alpha > 0$, let

$$a_{n,\delta,\alpha} = \inf\{\lambda \ge 0 : \mathbb{P}_\Phi\{\hat{D}_n \le \lambda\} \ge \alpha\},$$

where $\mathbb{P}_\Phi$ is the probability measure when the errors have a standard normal distribution. We abbreviate $a_{n,\delta,\alpha}$ to $a_{n,\alpha}$. Since

$$\hat{\beta}_n = A_n^{-1}\sum_{j=1}^n Y_j x_{n,j} = A_n^{-1}\sum_{j=1}^n x_{n,j}(x'_{n,j}\beta + \epsilon_j) = \beta + A_n^{-1}\sum_{j=1}^n \epsilon_j x_{n,j} \tag{1.4}$$

and

$$\hat{\epsilon}_j = Y_j - \hat{\beta}'_n x_{n,j} = \epsilon_j - (\hat{\beta}_n - \beta)' x_{n,j} = \epsilon_j - (\sum_{k=1}^n \epsilon_k x_{n,k})' A_n^{-1} x_{n,j} = \epsilon_j - \sum_{k=1}^n x'_{n,k} A_n^{-1} x_{n,j}\epsilon_k, \tag{1.5}$$

the distribution of $\hat{D}_n$ does not depend on σ. Observe that since the distribution of the residuals depends on the design of regressors, the distribution of the test statistic depends on the design of regressors. Given a design of regressors $x_{n,1}, \dots, x_{n,n}$, the value of $a_{n,\alpha}$ can be estimated via simulations. Therefore,

$$\text{the test rejects the null hypothesis if } \hat{D}_n > a_{n,\alpha} \tag{1.6}$$

can be implemented.

In Section 2, we present the asymptotics of the test statistic $\hat{D}_n$. We show that under certain conditions, the test is consistent against any alternative hypothesis. We obtain the limit distribution of $n\hat{D}_n$ under the null hypothesis and under contiguous alternatives. We show that $n\hat{D}_n$ under the null hypothesis is asympotically equivalent to a degenerate V–statistic. In particular, we get that $na_{n,\alpha}$ converges to a limit. This implies that for n large enough, the test can be done using $\lim_{n\to\infty} na_{n,\alpha}$. In Section 3, we present the outcome of some simulations comparing the presented test with the test in Shapiro and Wilk (1965). The result of the simulations show that the presented test is competitive. The proofs of the theorems in Section 2 are in Section 4.

2 Asymptotics of the test statistic

First, we give a useful representation of test statistic.

Lemma 2.1. *We have that*

$$\begin{aligned}
&\hat{D}_n && (2.1)\\
&= n^{-2}\sum_{j,k=1}^{n}\exp(-2^{-1}\delta^2\hat{\sigma}_n^{-2}(\hat{\epsilon}_j-\hat{\epsilon}_k)^2)\\
&\quad -2(1+\delta^2)^{-1/2}n^{-1}\sum_{j=1}^{n}\exp(-2^{-1}\delta^2(\delta^2+1)^{-1}\hat{\sigma}_n^{-2}\hat{\epsilon}_j^2)+(1+2\delta^2)^{-1/2}.
\end{aligned}$$

From the previous representation it is easy to compute the test statistic.

The following theorem gives the consistency of the test.

Theorem 2.1. *Consider the linear regression model with i.i.d. random errors $\{\epsilon_j\}$. Suppose that $E_F[\epsilon_1]=0$ and $E_F[\epsilon_1^4]<\infty$, where F is the c.d.f. of ϵ_1. Assume that the regressors $x_{n,j}$, $1\le j\le n$, satisfy condition (A). Then,*

$$\hat{D}_n \xrightarrow{\mathbb{P}} D_F := \int_{-\infty}^{\infty}|E_F[e^{it\sigma_F^{-1}\epsilon}]-e^{-2^{-1}t^2}|^2\phi_\delta(t)\,dt,$$

where $\sigma_F^2=\mathrm{Var}_F(\epsilon)$.

We have that $D_F=0$ if and only if F is a normal distribution. Hence, for each $0<\alpha<1$, $a_{n,\alpha}\to 0$, as $n\to\infty$. If F is not a normal c.d.f., then $\mathbb{P}_F\{\hat{D}_n>a_{n,\alpha}\}\to 1$, as $n\to\infty$.

By Lemma 2.1, $\hat{D}_n$ is a V–statistic with an estimated parameter. We will see that $n\hat{D}_n$ is a asymptotically equivalent to a regular V–statistic. So, in order to present the limit distribution of $n\hat{D}_n$ under the null hypothesis, we need to present some results about U–statistics. General references on U–statistics are Lee (1990) and de la Peña and Giné (1999). Given a sequence of i.i.d.r.v.'s $\{X_j\}$ with values in a measurable space $(S,\mathcal{S})$ and a measurable function $h: S\times S\to\mathbb{R}$, the U-statistic (of order 2) with kernel h is

$$U_n(h) := \frac{(n-2)!}{n!}\sum_{1\le j\ne k\le n}h(X_j,X_k).$$

The V–statistic with kernel h is

$$V_n(h) := n^{-2}\sum_{j,k=1}^{n}h(X_j,X_k).$$

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 P–degenerate if $E[h(x,X_1)]=E[h(X_1,x)]=E[h(X_1,X_2)]$ a.s. $[P]$, where $P:=\mathcal{L}(X_1)$ is the law of X_1. By Theorem 3.2.2.1 in Lee (1990, page 79), if h is a symmetric degenerate kernel with $E[h(X_1,X_2)]=0$ and $E[(h(X_1,X_2))^2]<\infty$, then

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

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

$$n^{-1}\sum_{j,k=1}^{n} h(X_j,X_k) \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty} \lambda_k(g_k^2-1) + E[h(X_1,X_1)], \tag{2.2}$$

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

In order to get the asymptotic distribution of the test, we assume the condition:

(B) $n^{-1}\sum_{j,k=1}^{n} x'_{n,j}A_n^{-1}x_{n,k} \to 1$ as $n \to \infty$.

Theorem 2.2. *Consider the linear regression model with i.i.d. random errors $\{\epsilon_j\}$, which have a normal distribution with mean zero and variance $\sigma^2 > 0$. Assume that the regressors $x_{n,j}$, $1 \le j \le n$, satisfy conditions (A) and (B).*

Then,

$$\begin{aligned}&n\hat{D}_n\\ &-\|n^{-1/2}\textstyle\sum_{j=1}^{n}(\cos(\sigma^{-1}t\epsilon_j)+\sin(\sigma^{-1}t\epsilon_j)-e^{-2^{-1}t^2}\left(1+\sigma^{-1}t\epsilon_j-2^{-1}\sigma^{-2}t^2(\epsilon_j^2-\sigma^2)\right))\|^2_{L_2(\phi_\delta)}\\ &\xrightarrow{\Pr} 0.\end{aligned}$$

Besides,

$$\begin{aligned}&\|n^{-1/2}\textstyle\sum_{j=1}^{n}(\cos(\sigma^{-1}t\epsilon_j)+\sin(\sigma^{-1}t\epsilon_j)-e^{-2^{-1}t^2}\left(1+\sigma^{-1}t\epsilon_j-2^{-1}\sigma^{-2}t^2(\epsilon_j^2-\sigma^2)\right))\|^2_{L_2(\phi_\delta)}\\ =\;&\textstyle\sum_{j,k=1}^{n} h(\sigma^{-1}\epsilon_j,\sigma^{-1}\epsilon_k)\\ \xrightarrow{\mathrm{d}}\;&\textstyle\sum_{k=1}^{\infty}\lambda_k(g_k^2-1)+E_\Phi[h(\epsilon,\epsilon)],\end{aligned}$$

where $\{g_k\}$ is a sequence of i.i.d.r.v.'s with a standard normal distribution and $\{\lambda_k\}$ denotes the eigenvalues of the operator $Q_h : L_2(S,\mathcal{S},\Phi) \to L_2(S,\mathcal{S},\Phi)$ defined by $Q_h q(x) = E_\Phi[h(x,\epsilon)q(\epsilon)]$, where Φ is the law of a standard normal r.v. and

$$\begin{aligned}&h(a,b) \qquad (2.3)\\ =\;&\exp(-2^{-1}\delta^2(a-b)^2)\\ &-(\delta^2+1)^{-1/2}\exp(-2^{-1}(\delta^2+1)^{-1}\delta^2a^2)\\ &\times(1+(\delta^2+1)^{-1}\delta^2b-2^{-1}(\delta^2+1)^{-2}\delta^2(b^2-1)(1-(\delta^2+1)^{-2}\delta^2a^2))\\ &-(\delta^2+1)^{-1/2}\exp(-2^{-1}(\delta^2+1)^{-1}\delta^2b^2)\\ &\times(1+(\delta^2+1)^{-1}\delta^2a-2^{-1}(\delta^2+1)^{-2}\delta^2(a^2-1)(1-(\delta^2+1)^{-2}\delta^2b^2))\\ &+(2\delta^2+1)^{-1/2}(1+(2\delta^2+1)^{-1}\delta^2(1-2^{-1}(a-b)^2)+2^{-2}3(2\delta^2+1)^{-2}\delta^4(a^2-1)(b^2-1)),\end{aligned}$$

It is of interest to know the distribution of $n\hat{D}_n$ in the case of sequence of contiguous alternatives. Given a sequence of measurable space spaces $\{(S_n, \mathcal{S}_n)\}$ and two sequences of probabilities measures $\{P_n\}$ and $\{Q_n\}$ such that P_n and Q_n are probability measures on $(S_n, \mathcal{S}_n)$, it is said that $\{Q_n\}$ is contiguous to $\{P_n\}$ if $P_n(A_n) \to 0$ implies $Q_n(A_n) \to 0$ for each sequence of sets $\{A_n\}$, $A_n \in \mathcal{S}_n$ (see e.g. Section VI.1 in Hájek and Šidák, 1967). Suppose that the observations $(x_{n,j}, Y_{n,j})$, $1 \le j \le n$, satisfy $Y_{n,j} = \beta' x_{n,j} + \epsilon_{n,j}$, $j \ge 1$, where $\epsilon_{n,1}, \dots, \epsilon_{n,n}$ are i.i.d.r.v.'s with pdf

$$f_n(x) = \phi_\sigma(x)(1 + n^{-1/2}\alpha_n(x)), \tag{2.4}$$

where $\sigma > 0$. As before, define $\hat{\beta}_n = A_n^{-1}(\sum_{j=1}^n Y_{n,j}x_{n,j})$, $\hat{\epsilon}_{n,j} = Y_{n,j} - \hat{\beta}_n' x_{n,j}$, $1 \le j \le n$, $\hat{\sigma}_n^2 := (n-p)^{-1}\sum_{j=1}^n \hat{\epsilon}_{n,j}^2$ and $\hat{D}_{n,\delta}$ is as in (1.3). Assume that the following condition is satisfied

(C) There exists a function $\alpha : \mathbb{R} \to \mathbb{R}$ such that

$$\int_{\mathbb{R}} (\alpha_n(x) - \alpha(x))^2 \phi_\sigma(x)\, dx \to 0$$

and

$$\int_{\mathbb{R}} (\alpha(x))^2 \phi_\sigma(x)\, dx < \infty.$$

Let $\{\epsilon_j\}$ be sequence of i.i.d.r.v.'s with a normal distribution with mean zero and variance σ^2. Let $P_n = \mathcal{L}(\epsilon_1, \dots, \epsilon_n)$ be the law of $(\epsilon_1, \dots, \epsilon_n)$ in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ–field of $\mathbb{R}^n$. Let $Q_n = \mathcal{L}(\epsilon_{n,1}, \dots, \epsilon_{n,n})$ be the law of $(\epsilon_{n,1}, \dots, \epsilon_{n,n})$ in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. We claim that under Condition (C), the sequence of probability measures $\{Q_n\}$ is contiguous to the sequence $\{P_n\}$. Let p_n be the p.d.f. of P_n with respect to the Lebesgue measure in $\mathbb{R}^n$. Let q_n be the p.d.f. of Q_n with respect to the Lebesgue measure in $\mathbb{R}^n$. Under Condition (C), $n^{-1/2}\sum_{j=1}^n \alpha_n(\epsilon_j) \xrightarrow{\mathrm{d}} N(0, \tau^2)$, where $\tau^2 = \int_{\mathbb{R}} \alpha^2(x)\phi_\sigma(x)\, dx$. By problem VI.5 in page 238 of Hájek and Šidák (1967) this implies that

$$\log \frac{q_n(\epsilon_1, \dots, \epsilon_n)}{p_n(\epsilon_1, \dots, \epsilon_n)} \xrightarrow{\mathrm{d}} N(2^{-1}\tau^2, \tau^2).$$

Hence, by the LeCam's first lemma (see e.g. Section VI.1.2 in Hájek and Šidák, 1967) $\{Q_n\}$ is contiguous to $\{P_n\}$.

The convergence of U–statistics was considered by Gregory (1977). Let $X_1, \dots, X_n$ be i.i.d.r.v.'s with law P and let h be a symmetric P–degenerate kernel with $E[(h(X_1, X_2))^2] < \infty$. Suppose that $X_{n,1}, \dots, X_{n,n}$ are i.i.d.r.v.'s with law P_n, where P_n is absolutely continuous with respect to P with Radon–Nikodym derivative $\frac{dP_n}{dP} = 1 + n^{-1/2}\alpha_n$, where $\alpha_n \in L_2(S, \mathcal{S}, P)$ and $\alpha_n \to \alpha$ in $L_2(S, \mathcal{S}, P)$. By Theorem 2.1 in Gregory (1977) if h is a symmetric degenerate kernel with $E[h(X_1, X_2)] = 0$ and $E[(h(X_1, X_2))^2] < \infty$, then

$$n^{-1} \sum_{1 \le j \ne k \le n} h(X_{n,j}, X_{n,k}) \xrightarrow{d} \sum_{j=1}^\infty \lambda_j((g_j + a_j)^2 - 1),$$

where $\{g_j\}$ is a sequence of i.i.d.r.v.'s with a standard normal distribution, $a_j = E[\psi_j(X_1)\alpha(X_1)]$, and $\{(\lambda_j, \psi_j)\}$ is the finite or infinite collection of pairs of eigenvalues–orthonormal eigenvectors of the operator Q_h. Assuming that $E[|h(X_1, X_1)|] < \infty$, by the strong of the large numbers,

$$n^{-1}\sum_{j=1}^{n}(h(X_j, X_j) - E[h(X_j, X_j)]) \to 0 \ \ a.s.$$

Hence, by contiguity

$$n^{-1}\sum_{j=1}^{n}(h(X_{n,j}, X_{n,j}) - E[h(X_j, X_j)]) \stackrel{\Pr}{\to} 0.$$

Hence, if h is a symmetric degenerate kernel with $E[h(X_1, X_2)] = 0$, $E[(h(X_1, X_2))^2] < \infty$ and $E[|h(X_1, X_1)|] < \infty$, then

$$n^{-1}\sum_{j,k=1}^{n} h(X_{n,j}, X_{n,k}) \stackrel{d}{\to} \sum_{j=1}^{\infty}\lambda_j((g_j + a_j)^2 - 1) + E[h(X_1, X_1)]. \tag{2.5}$$

Theorem 2.3. *Under the notation above, suppose that:*

(i) The regressors $x_{n,j}$, $1 \le j \le n$, satisfy conditions (A) and (B).

(ii) The errors $\epsilon_{n,1}, \ldots, \epsilon_{n,n}$ are i.i.d.r.v.s with p.d.f. given by (2.4).

(iii) Condition (C) is satisfied.

Then,

$$\begin{array}{c} n\hat{D}_n \\ -\|n^{-1/2}\sum_{j=1}^{n}\left(\cos(\sigma^{-1}t\epsilon_{n,j}) + \sin(\sigma^{-1}t\epsilon_{n,j}) - e^{-2^{-1}t^2}\left(1 + \sigma^{-1}t\epsilon_j - 2^{-1}\sigma^{-2}t^2(\epsilon_j^2 - \sigma^2)\right)\right)\|_{L_2(\phi_\delta)}^2 \\ \stackrel{\Pr}{\to} 0. \end{array} \tag{2.6}$$

Besides,

$$\begin{array}{rl} & \|n^{-1/2}\sum_{j=1}^{n}(\cos(\sigma^{-1}t\epsilon_{n,j}) + \sin(\sigma^{-1}t\epsilon_{n,j}) - e^{-2^{-1}t^2}\left(1 + \sigma^{-1}t\epsilon_j - 2^{-1}\sigma^{-2}t^2(\epsilon_j^2 - \sigma^2)\right))\|_{L_2(\phi_\delta)}^2 \\ = & n^{-1}\sum_{j,k=1}^{n} h(\sigma^{-1}\epsilon_{n,j}, \sigma^{-1}\epsilon_{n,k}) \\ \stackrel{d}{\to} & \sum_{j=1}^{\infty}\lambda_j((g_j + a_j)^2 - 1) + E_\Phi[h(\epsilon, \epsilon)], \end{array}$$

where h is as in (2.3), $\{g_j\}$ is a sequence of i.i.d.r.v.'s with a standard normal distribution, $a_j = E_\Phi[\psi_j(\epsilon)\alpha(\epsilon)]$, and $\{(\lambda_j, \psi_j)\}$ is the finite or infinite collection of pairs of eigenvalues-orthonormal eigenvectors of the operator Q_h.

3 Simulations.

We have made simulations to estimate the power of the (Shapiro–Wilk) SW test and the presented test for several distributions. We have used $x_{n,j} = (1, j)'$, for $1 \le j \le n$, and $\sigma = 1$.

Besides the test introduced in the introduction, we also consider the test using the biased estimator of the variance of the error. Let

$$\hat{D}_{n,\delta,bsd} := \|n^{-1}\sum_{j=1}^{n}\exp(it\hat{\sigma}_{n,bsd}^{-1}\hat{\epsilon}_j) - \exp(-2^{-1}t^2)\|^2_{L_2(\phi_\delta)}, \tag{3.1}$$

where $\hat{\sigma}^2_{n,bsd} := n^{-1}\sum_{j=1}^{n}\hat{\epsilon}_j^2$. The test rejects H_0, for $\hat{D}_{n,bsd,\delta} \geq b_{n,\delta,\alpha}$, where

$$b_{n,\delta,\alpha} = \inf\{\lambda \geq 0 : \mathbb{P}_\Phi\{\hat{D}_{n,\delta,bsd} \leq \lambda\} \geq \alpha\}.$$

It is easy to see that theorems 2.1-2.3 hold for $\hat{D}_{n,\delta,bsd}$.

The following tables show the values of $na_{n,,\delta\alpha}$ and $nb_{n,\delta,\alpha}$ for some values of n. The tables were obtained by doing 10000 simulations from a standard normal distribution. The test statistic $\hat{D}_n$ was found using the expression given by Lemma 2.1.

$na_{n,1,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.1723317	0.2263769	0.3625338
$n = 8$	0.2065039	0.2837184	0.4601209
$n = 10$	0.2220333	0.2971192	0.4734248
$n = 15$	0.2555162	0.3432884	0.5263078
$n = 20$	0.2599958	0.3440845	0.5340384
$n = 25$	0.2714118	0.3549643	0.5689468

$na_{n,1/2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.01901845	0.02468509	0.03896278
$n = 8$	0.02047842	0.02686428	0.04253534
$n = 10$	0.02256076	0.03003705	0.04934944
$n = 15$	0.02546400	0.03517977	0.06237715
$n = 20$	0.02679697	0.03730412	0.06435028
$n = 25$	0.02841144	0.03974311	0.06939096

$nb_{n,1,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.2512835	0.3084409	0.4880363
$n = 8$	0.2659785	0.3372054	0.4981566
$n = 10$	0.2709936	0.3503190	0.5136096
$n = 15$	0.2758713	0.3571201	0.5456935
$n = 20$	0.2849200	0.3653052	0.5585711
$n = 25$	0.2895949	0.3711702	0.5712650

$nb_{n,1/2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 6$	0.01831819	0.02512786	0.04152145
$n = 8$	0.02142009	0.03030869	0.05221142
$n = 10$	0.02346332	0.03272659	0.05518262
$n = 15$	0.02610572	0.03617368	0.06352592
$n = 20$	0.02789834	0.03964410	0.06633810
$n = 25$	0.02827096	0.04006235	0.07027089

Since $n\hat{D}_{n,\delta,unbsd}$ and $n\hat{D}_{n,\delta,bsd}$ converge in distribution to the same limit, we have that $na_{n,\alpha,\delta}$ and $nb_{n,\alpha,\delta}$ are close for n large.

The following table shows the power when $\alpha = 0.05$ of the tests of normality in Shapiro and Wilk (1965) and in Epps and Pulley (1983) (using $\delta = 1$ and $\delta = 1/2$ and the unbiased and the biased estimators of the variance).

	SW	$\mathrm{BHEP}_{\mathrm{unbsd}}(1)$	$\mathrm{BHEP}_{\mathrm{unbsd}}(1/2)$	$\mathrm{BHEP}_{\mathrm{bsd}}(1)$	$\mathrm{BHEP}_{\mathrm{bsd}}(1/2)$
$n = 6$	0.1098	0.0989	0.0929	0.1039	0.1050
$n = 8$	0.1807	0.1756	0.1782	0.1915	0.1859
$n = 10$	0.2719	0.2761	0.2621	0.2777	0.2877
$n = 15$	0.5124	0.4768	0.4715	0.5105	0.5162
$n = 20$	0.7022	0.6753	0.6490	0.6823	0.6661
$n = 25$	0.8370	0.8056	0.7721	0.8018	0.8070
	Alternative: exponential distribution with mean one				

	SW	$\mathrm{BHEP}_{\mathrm{unbsd}}(1)$	$\mathrm{BHEP}_{\mathrm{unbsd}}(1/2)$	$\mathrm{BHEP}_{\mathrm{bsd}}(1)$	$\mathrm{BHEP}_{\mathrm{bsd}}(1/2)$
$n = 6$	0.0755	0.0789	0.0652	0.0737	0.0740
$n = 8$	0.0915	0.1013	0.1117	0.0977	0.1059
$n = 10$	0.1057	0.1433	0.1423	0.1169	0.1404
$n = 15$	0.1719	0.1910	0.2148	0.1752	0.2015
$n = 20$	0.2280	0.2625	0.2626	0.2263	0.2294
$n = 25$	0.2748	0.3119	0.2931	0.2719	0.2829
	Alternative: double exponential distribution				

	SW	$\mathrm{BHEP}_{\mathrm{unbsd}}(1)$	$\mathrm{BHEP}_{\mathrm{unbsd}}(1/2)$	$\mathrm{BHEP}_{\mathrm{bsd}}(1)$	$\mathrm{BHEP}_{\mathrm{bsd}}(1/2)$
$n = 6$	0.1471	0.1739	0.1459	0.1523	0.1585
$n = 8$	0.3058	0.3288	0.3462	0.3199	0.3344
$n = 10$	0.4106	0.4898	0.4990	0.4436	0.4872
$n = 15$	0.6922	0.7143	0.7035	0.7045	0.6892
$n = 20$	0.8181	0.8425	0.8353	0.8272	0.8008
$n = 25$	0.9006	0.9139	0.8836	0.8978	0.8765
	Alternative: Cauchy distribution				

	SW	$\text{BHEP}_{\text{unbsd}}(1)$	$\text{BHEP}_{\text{unbsd}}(1/2)$	$\text{BHEP}_{\text{bsd}}(1)$	$\text{BHEP}_{\text{bsd}}(1/2)$
$n=6$	0.0596	0.0583	0.0505	0.0589	0.0586
$n=8$	0.0646	0.0649	0.0706	0.0681	0.0703
$n=10$	0.0629	0.0828	0.0836	0.0677	0.0830
$n=15$	0.0882	0.0900	0.1115	0.0877	0.1060
$n=20$	0.1039	0.1180	0.1281	0.1026	0.1172
$n=25$	0.1242	0.1300	0.1491	0.1114	0.1414
	Alternative: logistic(1) distribution				

	SW	$\text{BHEP}_{\text{unbsd}}(1)$	$\text{BHEP}_{\text{unbsd}}(1/2)$	$\text{BHEP}_{\text{bsd}}(1)$	$\text{BHEP}_{\text{bsd}}(1/2)$
$n=6$	0.0636	0.0506	0.04547	0.0579	0.0541
$n=8$	0.0703	0.0605	0.0578	0.0689	0.0587
$n=10$	0.0844	0.0774	0.0632	0.0878	0.0715
$n=15$	0.1413	0.0970	0.0825	0.1462	0.1069
$n=20$	0.2247	0.1646	0.1141	0.2051	0.1415
$n=25$	0.3171	0.2304	0.1360	0.2837	0.1871
	Alternative: Beta$(2,1)$ distribution				

	SW	$\text{BHEP}_{\text{unbsd}}(1)$	$\text{BHEP}_{\text{unbsd}}(1/2)$	$\text{BHEP}_{\text{bsd}}(1)$	$\text{BHEP}_{\text{bsd}}(1/2)$
$n=6$	0.0565	0.0421	0.0339	0.0491	0.0444
$n=8$	0.0570	0.0285	0.0305	0.0417	0.0343
$n=10$	0.0556	0.0322	0.0232	0.0468	0.0283
$n=15$	0.0840	0.0280	0.0133	0.0670	0.0223
$n=20$	0.1318	0.0477	0.0106	0.0948	0.0162
$n=25$	0.2182	0.0603	0.0105	0.1374	0.0200
	Alternative: uniform $(0,1)$ distribution				

The previous table seems to indicate that the presented tests are competitive with the known tests. The $\text{BHEP}_{\text{unbsd}}(1)$ test is the most powerful for the double exponential, Cauchy and logistic distributions.

4 Proofs

We will use the following lemma:

Lemma 4.1. *Under condition (A),*
(i) For each $1 \leq j \leq n$, $x'_{n,j} A_n^{-1} x_{n,j} \geq 0$.
(ii) $\sum_{j=1}^{n} x'_{n,j} A_n^{-1} x_{n,j} = p$.

Proof. Under condition (A), A_n is a symmetric positive definite matrix. So, $A_n^{-1/2}$ exists and $x'_{n,j} A_n^{-1} x_{n,j} = |A_n^{-1/2} x_{n,j}|^2 \geq 0$, i.e. (i) holds.

Consider the linear regression model with i.i.d. errors with mean zero and variance one. It is well known (see e.g. Section 7.2 in Neter et al., 1996) that the linear regression model can be written as

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon,$$

where $\mathbf{Y} = (Y_1, \ldots, Y_n)'$, $\epsilon = (\epsilon_1, \ldots, \epsilon_n)'$, $\beta = (\beta_1, \ldots, \beta_p)'$,

$$\mathbf{X} = \begin{pmatrix} x_{n,1,1} & x_{n,1,2} & \cdots & x_{n,1,p} \\ x_{n,2,1} & x_{n,2,2} & \cdots & x_{n,2,p} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ x_{n,n,1} & x_{n,n,2} & \cdots & x_{n,n,p} \end{pmatrix}.$$

and $x_{n,j} = (x_{n,j,1}, \ldots, x_{n,j,p})'$. With this matrix notation,

$$A_n = \mathbf{X}'\mathbf{X},$$

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

and

$$\hat{\epsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta} = \epsilon - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon = (I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\epsilon.$$

Assuming that for each $1 \le j \le n$, $E[\epsilon_j] = 0$ and $E[\epsilon_j^2] = 0$, we have that

$$E[\hat{\epsilon}\hat{\epsilon}'] = (I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') = I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

So,

$$\begin{aligned} & \textstyle\sum_{j=1}^n \mathrm{Var}(\hat{\epsilon}_j) = \mathrm{trace}(I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \\ = \;& n - \mathrm{trace}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}) = n - \mathrm{trace}(I_p) = n - p. \end{aligned} \tag{4.1}$$

By (1.5)

$$\begin{aligned} & \mathrm{Var}(\hat{\epsilon}_j) = (1 - x_{n,j}'A_n^{-1}x_{n,j})^2 + \textstyle\sum_{k:1\le k\le n, k\ne j}(x_{n,k}'A_n^{-1}x_{n,j})^2 \\ = \;& 1 - 2x_{n,j}'A_n^{-1}x_{n,j} + \textstyle\sum_{k=1}^n (x_{n,k}'A_n^{-1}x_{n,j})^2 \\ = \;& 1 - 2x_{n,j}'A_n^{-1}x_{n,j} + \textstyle\sum_{k=1}^n x_{n,j}'A_n^{-1}x_{n,k}x_{n,k}'A_n^{-1}x_{n,j} = 1 - x_{n,j}'A_n^{-1}x_{n,j} \end{aligned}$$

and

$$\sum_{j=1}^n \mathrm{Var}(\hat{\epsilon}_j) = n - \sum_{j=1}^n x_{n,j}'A_n^{-1}x_{n,j}. \tag{4.2}$$

Hence, (ii) follows from (4.1) and (4.2). $\square$

Proof of Lemma 2.1. Using that

$$\int_{\mathbb{R}} \exp(it)\phi_\delta(t)\,dt = \exp(-2^{-1}\delta^2),$$

we have that

$$\begin{aligned}
& \hat{D}_n \\
= \;& n^{-2}\textstyle\sum_{j,k=1}^n \int_{\mathbb{R}} (\exp(it\hat{\sigma}_n^{-1}(\hat{\epsilon}_j - \hat{\epsilon}_k)) - 2\exp(it\hat{\sigma}_n^{-1}\hat{\epsilon}_j)e^{-2^{-1}t^2} + e^{-t^2})\phi_\delta(t)\,dt \\
= \;& n^{-2}\textstyle\sum_{j,k=1}^n \int_{\mathbb{R}} \exp(it\hat{\sigma}_n^{-1}(\hat{\epsilon}_j - \hat{\epsilon}_k))\phi_\delta(t)\,dt \\
& -2(1+\delta^2)^{-1/2}n^{-1}\textstyle\sum_{j=1}^n \int_{\mathbb{R}} \exp(it\hat{\sigma}_n^{-1}\hat{\epsilon}_j)\phi_{\delta(1+\delta^2)^{-1/2}}(t)\,dt \\
& + \textstyle\int_{\mathbb{R}} (2\delta^2+1)^{-1/2}\phi_{\delta(2\delta^2+1)^{-1/2}}(t)\,dt \\
= \;& n^{-2}\textstyle\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\hat{\sigma}_n^{-2}(\hat{\epsilon}_j - \hat{\epsilon}_k)^2) \\
& -2(1+\delta^2)^{-1/2}n^{-1}\textstyle\sum_{j=1}^n \exp(-2^{-1}\delta^2(\delta^2+1)^{-1}\hat{\sigma}_n^{-2}\hat{\epsilon}_j^2) + (1+2\delta^2)^{-1/2}.
\end{aligned}$$

□

Proof of Theorem 2.1. We have that

$$\begin{aligned}
D_F \;&=\; \textstyle\int_{-\infty}^{\infty} |E_F[e^{it\sigma_F^{-1}\epsilon}] - e^{-2^{-1}t^2}|^2\phi_\delta(t)\,dt \\
&=\; \textstyle\int_{-\infty}^{\infty} E_F[\exp(it\sigma_F^{-1}(\epsilon_1 - \epsilon_2)) - 2\exp(it\sigma_F^{-1}\epsilon_1 - 2^{-1}t^2) + \exp(2^{-1}t^2)]\phi_\delta(t)\,dt \\
&=\; E_F[\exp(-2^{-1}\delta^2\sigma_F^{-2}(\epsilon_1 - \epsilon_2)^2)] \\
&\quad -2(1+\delta^2)^{-1/2}E_F[\exp(-2^{-1}\delta^2(\delta^2+1)^{-1}\sigma_F^{-2}\epsilon_1^2)] + (1+2\delta^2)^{-1/2}.
\end{aligned}$$

Using (2.1) and the previous equality, it suffices to prove that

$$n^{-2}\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\hat{\sigma}_n^{-2}(\hat{\epsilon}_j - \hat{\epsilon}_k)^2) \xrightarrow{\Pr} E_F[\exp(-2^{-1}\delta^2\sigma_F^{-2}(\epsilon_1 - \epsilon_2)^2)] \tag{4.3}$$

and

$$n^{-1}\sum_{j=1}^n \exp(-2^{-1}\delta^2(\delta^2+1)^{-1}\hat{\sigma}_n^{-2}\hat{\epsilon}_j^2) \xrightarrow{\Pr} E_F[\exp(-2^{-1}\delta^2(\delta^2+1)^{-1}\sigma_F^{-2}\epsilon_1^2)]. \tag{4.4}$$

First, we prove that

$$n^{-1}\sum_{j,k=1}^n x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k \xrightarrow{\Pr} 0. \tag{4.5}$$

Equation (4.5) follows from the following (using Lemma 4.1):

$$E_F[n^{-1}\sum_{j=1}^n x'_{n,j}A_n^{-1}x_{n,j}\epsilon_j^2] = \sigma_F^2 n^{-1}\sum_{j=1}^n x'_{n,j}A_n^{-1}x_{n,j} = \sigma_F^2 n^{-1}p \to 0,$$

$$E_F[n^{-1}\sum_{1\le j<k\le n} x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k] = 0,$$

and

$$\begin{array}{rl} & \mathrm{Var}_F(n^{-1}\sum_{1\le j<k\le n} x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k) = (\mathrm{Var}_F(\epsilon_1^2))^2 n^{-2}\sum_{1\le j<k\le n}(x'_{n,j}A_n^{-1}x_{n,k})^2 \\ \le & 2^{-1}(\mathrm{Var}_F(\epsilon_1^2))^2 n^{-2}\sum_{j,k=1}^n x'_{n,j}A_n^{-1}x_{n,k}x'_{n,k}A_n^{-1}x_{n,j} \\ = & 2^{-1}(\mathrm{Var}_F(\epsilon_1^2))^2 n^{-2}\sum_{j=1}^n x'_{n,j}A_n^{-1}x_{n,j} = 2^{-1}(\mathrm{Var}_F(\epsilon_1^2))^2 n^{-2}p \to 0. \end{array}$$

Next, we prove that $\hat{\sigma}_n^2 \xrightarrow{\Pr} \sigma_F^2$. By (1.4),

$$n^{-1}\sum_{j=1}^n(\hat{\beta}_n-\beta)'x_{n,j}\epsilon_j = n^{-1}\sum_{j=1}^n(\sum_{k=1}^n x_{n,k}\epsilon_k)'A_n^{-1}x_{n,j}\epsilon_j = n^{-1}\sum_{j,k=1}^n x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k$$

and

$$\begin{array}{rl} & n^{-1}\sum_{j=1}^n((\hat{\beta}_n-\beta)'x_{n,j})^2 = n^{-1}\sum_{j=1}^n(\hat{\beta}_n-\beta)'x_{n,j}x'_{n,j}(\hat{\beta}_n-\beta) \\ = & n^{-1}(\hat{\beta}_n-\beta)'A_n(\hat{\beta}_n-\beta) = n^{-1}(\sum_{j=1}^n x_{n,j}\epsilon_j)A_n^{-1}(\sum_{k=1}^n x_{n,k}\epsilon_k) \\ = & n^{-1}\sum_{j,k=1}^n x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k. \end{array} \tag{4.6}$$

Using the previous expressions,

$$\begin{array}{rl} & n^{-1}(n-p)\hat{\sigma}_n^2 \\ = & n^{-1}\sum_{j=1}^n(\epsilon_j-(\hat{\beta}_n-\beta)'x_{n,j})^2 \\ = & n^{-1}\sum_{j=1}^n\epsilon_j^2 - 2n^{-1}\sum_{j=1}^n(\hat{\beta}_n-\beta)'x_{n,j}\epsilon_j + n^{-1}\sum_{j=1}^n((\hat{\beta}_n-\beta)'x_{n,j})^2 \\ = & n^{-1}\sum_{j=1}^n\epsilon_j^2 - n^{-1}\sum_{j,k=1}^n x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k. \end{array} \tag{4.7}$$

Hence, by (4.5) and (4.7),

$$\hat{\sigma}_n^2 \xrightarrow{\Pr} \sigma_F^2. \tag{4.8}$$

Next, we prove that

$$n^{-2}\sum_{j,k=1}^n |(\hat{\epsilon}_j-\hat{\epsilon}_k)^2-(\epsilon_j-\epsilon_k)^2| \xrightarrow{\Pr} 0. \tag{4.9}$$

Using (1.5), we have that

$$\begin{array}{rl} & n^{-2}\sum_{j,k=1}^n |(\hat{\epsilon}_j-\hat{\epsilon}_k)^2-(\epsilon_j-\epsilon_k)^2| \\ = & n^{-2}\sum_{j,k=1}^n |-2(\epsilon_j-\epsilon_k)(\hat{\beta}_n-\beta)'(x_{n,j}-x_{n,k})+((\hat{\beta}_n-\beta)'(x_{n,j}-x_{n,k}))^2| \\ \le & 2(n^{-2}\sum_{j,k=1}^n(\epsilon_j-\epsilon_k)^2)^{1/2}(n^{-2}\sum_{j,k=1}^n((\hat{\beta}_n-\beta)'(x_{n,j}-x_{n,k}))^2)^{1/2} \\ & +n^{-2}\sum_{j,k=1}^n((\hat{\beta}_n-\beta)'(x_{n,j}-x_{n,k}))^2. \end{array}$$

By the law of the large numbers for U–statistics,

$$n^{-2}\sum_{j,k=1}^n(\epsilon_j-\epsilon_k)^2 = n^{-2}\sum_{1\le j\ne k\le n}(\epsilon_j-\epsilon_k)^2 \xrightarrow{\Pr} E[(\epsilon_1-\epsilon_2)^2]. \tag{4.10}$$

By (4.6) and (4.5),

$$\begin{aligned} & n^{-2}\sum_{j,k=1}^n((\hat\beta_n-\beta)'(x_{n,j}-x_{n,k}))^2 \le 4n^{-2}\sum_{j,k=1}^n((\hat\beta_n-\beta)'x_{n,j})^2 \\ = \;& 4n^{-1}\sum_{j,k=1}^n x'_{n,j}A_n^{-1}x_{n,k}\epsilon_j\epsilon_k \overset{\Pr}{\to} 0. \end{aligned}$$

Hence, (4.9) follows.

Using that for $a, b \ge 0$, $|e^{-a}-e^{-b}| \le |a-b|$ and (4.8)–(4.10),

$$\begin{aligned} & |n^{-2}\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\hat\sigma_n^{-2}(\hat\epsilon_j-\hat\epsilon_k)^2) - n^{-2}\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\sigma_F^{-2}(\hat\epsilon_j-\hat\epsilon_k)^2)| \\ \le \;& 2^{-1}\delta^2|\hat\sigma_n^{-2}-\sigma_F^{-2}|n^{-2}\sum_{j,k=1}^n(\hat\epsilon_j-\hat\epsilon_k)^2 \\ \le \;& 2^{-1}\delta^2|\hat\sigma_n^{-2}-\sigma_F^{-2}|n^{-2}\sum_{j,k=1}^n|(\hat\epsilon_j-\hat\epsilon_k)^2-(\epsilon_j-\epsilon_k)^2| \\ & +2^{-1}\delta^2|\hat\sigma_n^{-2}-\sigma_F^{-2}|n^{-2}\sum_{j,k=1}^n(\epsilon_j-\epsilon_k)^2 \overset{\Pr}{\to} 0. \end{aligned} \tag{4.11}$$

By a similar argument,

$$\begin{aligned} & |n^{-2}\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\sigma_F^{-2}(\hat\epsilon_j-\hat\epsilon_k)^2) - n^{-2}\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\sigma_F^{-2}(\epsilon_j-\epsilon_k)^2)| \\ \le \;& 2^{-1}\delta^2\sigma_F^{-2}|n^{-2}\sum_{j,k=1}^n|(\hat\epsilon_j-\hat\epsilon_k)^2-(\epsilon_j-\epsilon_k)^2| \overset{\Pr}{\to} 0. \end{aligned} \tag{4.12}$$

By the law of the large numbers for V–statistics,

$$n^{-2}\sum_{j,k=1}^n \exp(-2^{-1}\delta^2\sigma_F^{-2}(\epsilon_j-\epsilon_k)^2) \overset{\Pr}{\to} E_F[\exp(-2^{-1}\delta^2\sigma_F^{-2}(\epsilon_1-\epsilon_2)^2)]. \tag{4.13}$$

Hence, (4.3) follows from (4.11)–(4.13). The proof of (4.4) is similar and it is omitted. □

Theorem 2.3 with $\alpha_n \equiv 0$ implies Theorem 2.2. Hence, we only prove Theorem 2.3.

Lemma 4.2. *Under the conditions in Theorem 2.3,*

$$A_n^{1/2}(\hat\beta_n-\beta) = O_P(1), \tag{4.14}$$

$$n^{-1/2}\sum_{j=1}^n(\epsilon_{n,j}^2-\sigma^2) = O_P(1) \tag{4.15}$$

and

$$n^{1/2}(\hat\sigma_n^2-\sigma^2) - n^{-1/2}\sum_{j=1}^n(\epsilon_{n,j}^2-\sigma^2) \overset{\Pr}{\to} 0. \tag{4.16}$$

Proof. It is easy to see that

$$n^{1/2}E[\epsilon_{n,1}] \to \int_{\mathbb{R}} x\alpha(x)\phi_\sigma(x)\,dx, \tag{4.17}$$

$$n^{1/2}\left(E[\epsilon_{n,1}^2]-\sigma^2\right) \to \int_{\mathbb{R}} x^2\alpha(x)\phi_\sigma(x)\,dx. \tag{4.18}$$

and

$$n^{1/2}\left(E[\epsilon_{n,1}^4]-3\sigma^4\right)\to\int_{\mathbb{R}} x^4\alpha(x)\phi_\sigma(x)\,dx. \tag{4.19}$$

Using (1.4), we get that

$$|A_n^{1/2}(\hat\beta_n-\beta)|^2=(\hat\beta_n-\beta)'A_n(\hat\beta_n-\beta)=\sum_{j,k=1}^n x_{n,j}'A_n^{-1}x_{n,k}\epsilon_{n,j}\epsilon_{n,k}. \tag{4.20}$$

Using Lemma 4.1 and (4.17)–(4.19), we get that

$$E[\sum_{j=1}^n x_{n,j}'A_n^{-1}x_{n,j}\epsilon_{n,j}^2]=E[\epsilon_{n,1}^2]\sum_{j=1}^n x_{n,j}'A_n^{-1}x_{n,j}=pE[\epsilon_{n,1}^2]=0(1),$$

$$\begin{array}{rl} & E[\sum_{1\le j<k\le n} x_{n,j}'A_n^{-1}x_{n,k}\epsilon_{n,j}\epsilon_{n,k}]\le 2^{-1}(E[\epsilon_{n,1}])^2\sum_{j,k=1}^n x_{n,j}'A_n^{-1}x_{n,k}\\ \le & 2^{-2}(E[\epsilon_{n,1}])^2\sum_{j,k=1}^n(x_{n,j}'A_n^{-1}x_{n,j}+x_{n,k}'A_n^{-1}x_{n,k})\\ = & 2^{-1}pn(E[\epsilon_{n,1}])^2=0(1),\end{array}$$

and

$$\begin{array}{rl} & \mathrm{Var}(\sum_{1\le j<k\le n} x_{n,j}'A_n^{-1}x_{n,k}\epsilon_{n,j}\epsilon_{n,k})=n^{-1}\sum_{1\le j<k\le n}\mathrm{Var}(x_{n,j}'A_n^{-1}x_{n,k}\epsilon_{n,j}\epsilon_{n,k})\\ = & \sum_{1\le j<k\le n}(x_{n,j}'A_n^{-1}x_{n,k})^2\mathrm{Var}(\epsilon_{n,j})\mathrm{Var}(\epsilon_{n,k})\\ \le & 2^{-1}\sum_{j,k=1}^n(x_{n,j}'A_n^{-1}x_{n,k})^2(\mathrm{Var}(\epsilon_{n,1}^2))^2\\ = & 2^{-1}(\mathrm{Var}(\epsilon_{n,1}^2))^2\sum_{j,k=1}^n x_{n,j}'A_n^{-1}x_{n,k}x_{n,k}'A_n^{-1}x_{n,j}\\ = & 2^{-1}(\mathrm{Var}(\epsilon_{n,1}^2))^2\sum_{j=1}^n x_{n,j}'A_n^{-1}x_{n,j}=2^{-1}(\mathrm{Var}(\epsilon_{n,1}^2))^2p=0(1).\end{array}$$

Hence, (4.14) follows.

By (4.18) and (4.19),

$$E[n^{-1/2}\sum_{j=1}^n(\epsilon_{n,j}^2-\sigma^2)]=O(1)$$

and

$$\mathrm{Var}(n^{-1/2}\sum_{j=1}^n(\epsilon_{n,j}^2-\sigma^2))=O_P(1).$$

So, (4.15) holds.

By (4.7),

$$\hat\sigma_n^2=(n-p)^{-1}\sum_{j=1}^n\epsilon_{n,j}^2-(n-p)^{-1}\sum_{j,k=1}^n x_{n,j}'A_n^{-1}x_{n,k}\epsilon_{n,j}\epsilon_{n,k}.$$

Hence, to prove (4.16), it suffices to prove that

$$n^{-3/2}\sum_{j=1}^n\epsilon_{n,j}^2\xrightarrow{\Pr}0 \tag{4.21}$$

and

$$n^{-1/2}\sum_{j,k=1}^{n} x'_{n,j}A_n^{-1}x_{n,k}\epsilon_{n,j}\epsilon_{n,k} \xrightarrow{\Pr} 0. \tag{4.22}$$

By (4.18),

$$E[n^{-3/2}\sum_{j=1}^{n}\epsilon_{n,j}^2] \to 0$$

which implies (4.21). (4.22) follows from (4.14) and (4.20).

□

Lemma 4.3. *Under the conditions in Theorem 2.3, for each* $0 < M < \infty$,

$$\sup_{|h|\le M, |v|\le M} \|U_n(t,h,v) - U_n(t,0,0)\|_{L_2(\phi_\delta)} \xrightarrow{\Pr} 0, \tag{4.23}$$

where

$$\begin{aligned} & U_n(t,h,v) \\ := \;& n^{-1/2}\textstyle\sum_{j=1}^{n}(g(\epsilon_{n,j},t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2}) - G_n(t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2})), \end{aligned}$$

is defined for $n > M^2\sigma^{-4}$, $t \in \mathbb{R}$, $|h| \le M$, $|v| \le M$*;*

$$g(x,t,\mu,s) := \cos(s^{-1}t(x-\mu)) + \sin(s^{-1}t(x-\mu)) - e^{-2^{-1}t^2}, t,\mu \in \mathbb{R}, s > 0,$$

and $G_n(t,\mu,s) := E[g(\epsilon_{n,1},t,\mu,s)]$, $t,\mu \in \mathbb{R}$, $s > 0$.

Proof. We say that a function $f : A \to B$ has finite range if the cardinality of $f(A)$ is finite. Given $\eta > 0$, take $\pi_1 : \{h \in \mathbb{R}^p : |h| \le M\} \to \{h \in \mathbb{R}^p : |h| \le M\}$ with finite range such that $|\pi_1(h) - h| \le \eta$ for each $|h| \le M$; and take $\pi_2 : [-M, M] \to [-M, M]$ with finite range such that $|\pi_2(v) - v| \le \eta$ for each $|v| \le M$. Given $\tau > 0$,

$$\begin{aligned} & \Pr\{\sup_{|h|,|v|\le M} \|U_n(t,h,v) - U_n(t,0,0)\|_{L_2(\phi_\delta)} \ge \tau\} \\ \le\;& \Pr\{\sup_{|h|,|v|\le M} \|U_n(t,h,v) - U_n(t,\pi_1(h),\pi_2(v))\|_{L_2(\phi_\delta)} \ge 2^{-1}\tau\} \\ & + \Pr\{\sup_{|h|,|v|\le M} \|U_n(t,\pi_1(h),\pi_2(v)) - U_n(t,0,0)\|_{L_2(\phi_\delta)} \ge 2^{-1}\tau\}. \end{aligned} \tag{4.24}$$

Since sin and cos are Lipschitz functions with Lipschitz constant one, for each n with $n^{-1/2}M \le 2^{-1}\sigma^2$,

$$\begin{aligned} & |g(\epsilon_{n,j},t,h'A_n^{-1}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2}) - g(\epsilon_{n,j},t,(\pi_1(h))'A_n^{-1}x_{n,j},(\sigma^2+n^{-1/2}\pi_2(v))^{1/2}| \\ \le\;& 2|(\sigma^2+n^{-1/2}v)^{-1/2}t(\epsilon_{n,j} - h'A_n^{-1/2}x_{n,j}) \\ & -(\sigma^2+n^{-1/2}\pi_2(v))^{-1/2}t(\epsilon_{n,j} - (\pi_1(h))'A_n^{-1/2}x_{n,j})| \\ \le\;& 2|(\sigma^2+n^{-1/2}v)^{-1/2}t(\epsilon_{n,j} - h'A_n^{-1/2}x_{n,j}) - (\sigma^2+n^{-1/2}\pi_2(v))^{-1/2}t(\epsilon_{n,j} - h'A_n^{-1/2}x_{n,j})| \\ & +2|(\sigma^2+n^{-1/2}\pi_2(v))^{-1/2}t(\epsilon_{n,j} - h'A_n^{-1/2}x_{n,j}) \\ & -(\sigma^2+n^{-1/2}\pi_2(v))^{-1/2}t(\epsilon_{n,j} - (\pi_1(h))'A_n^{-1/2}x_{n,j})| \\ \le\;& 4\sigma^{-3}n^{-1/2}\eta|t|(|\epsilon_{n,j}| + |h'A_n^{-1/2}x_{n,j}|) + 4\sigma^{-1}\eta|t||A_n^{-1/2}x_{n,j}|. \end{aligned}$$

Hence, for each n large enough,

$$\begin{aligned} & \sup_{|h|,|v|\le M} |U_n(t,h,v) - U_n(t,\pi_1(h),\pi_2(v))| \\ \le \; & n^{-1/2}\sum_{j=1}^n \left(4\sigma^{-3}n^{-1/2}\eta|t||\epsilon_{n,j}| + 4\sigma^{-3}n^{-1/2}\eta|t|M|A_n^{-1/2}x_{n,j}| + 4\sigma^{-1}\eta|t||A_n^{-1/2}x_{n,j}|\right), \end{aligned}$$

and

$$\begin{aligned} & \Pr\{\sup_{|h|,|v|\le M} \|U_n(t,h,v) - U_n(t,\pi_1(h),\pi_2(v))\|_{L_2(\phi_\delta)} \ge 2^{-1}\tau\} \qquad (4.25) \\ \le \; & 2\tau^{-1}E[\sup_{|h|,|v|\le M} \|U_n(t,h,v) - U_n(t,\pi_1(h),\pi_2(v))\|_{L_2(\phi_\delta)}] \\ \le \; & 8\tau^{-1}\sigma^{-1}\eta\|t\|_{L_2(\phi_\delta)}\left(\sigma^{-2}E[|\epsilon_{n,1}|] + \sigma^{-2}Mn^{-1}\sum_{j=1}^n |A_n^{-1/2}x_{n,j}| + n^{-1/2}\sum_{j=1}^n |A_n^{-1/2}x_{n,j}|\right) \\ \le \; & 8\tau^{-1}\sigma^{-1}\eta\|t\|_{L_2(\phi_\delta)}\left(\sigma^{-2}E[|\epsilon_{n,1}|] + \sigma^{-2}Mn^{-1/2}p^{1/2} + p^{1/2}\right) \end{aligned}$$

where we have used that by Lemma 4.1,

$$n^{-1/2}\sum_{j=1}^n |A_n^{-1/2}x_{n,j}| \le \left(\sum_{j=1}^n |A_n^{-1/2}x_{n,j}|^2\right)^{1/2} = \left(\sum_{j=1}^n x'_{n,j}A_n^{-1}x_{n,j}\right)^{1/2} = p^{1/2}.$$

We claim that for each h, v,

$$\|U_n(t,h,v) - U_n(t,0,0)\|_{L_2(\phi_\delta)} \xrightarrow{\Pr} 0. \qquad (4.26)$$

We have that

$$\|U_n(t,h,v) - U_n(t,0,0)\|_{L_2(\phi_\delta)}^2 = n^{-1}\sum_{j,k=1}^n \int_{\mathbb{R}} Y_{n,j}(\epsilon_{n,j},t,h,v)Y_{n,k}(\epsilon_{n,k},t,h,v)\phi_\delta(t)\,dt,$$

where

$$\begin{aligned} & Y_{n,j}(x,t,h,v) := g(x,t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2}) - g(x,t,0,\sigma) \\ & -G_n(t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2}) + G_n(t,0,\sigma). \end{aligned}$$

Since sin and cos are Liptchitz functions, for each $t, \mu \in \mathbb{R}$ and each $s > 0$,

$$\begin{aligned} & |g(\epsilon,t,\mu,s) - g(\epsilon,t,0,\sigma)| \le 2|s^{-1}t(\epsilon-\mu) - \sigma^{-1}t\epsilon)| \qquad (4.27) \\ \le \; & 2|s^{-1}t(\epsilon-\mu) - s^{-1}t\epsilon| + 2|s^{-1}t\epsilon - \sigma^{-1}t\epsilon| \le 2s^{-1}|t||\mu| + 2s^{-1}\sigma^{-1}(s+\sigma)^{-1}|t||\epsilon||s^2-\sigma^2|. \end{aligned}$$

Using (4.27), we have that if $n^{-1/2}|v| \le 2^{-1}\sigma^2$, then

$$\begin{aligned} & E[n^{-1}\sum_{j=1}^n \int_{\mathbb{R}}(Y_{n,j}(\epsilon_{n,j},t,h,v))^2\phi_\delta(t)\,dt] \\ \le \; & E[n^{-1}\sum_{j=1}^n \int_{\mathbb{R}}(g(\epsilon_{n,j},t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2}) - g(\epsilon_{n,j},t,0,\sigma))^2\phi_\delta(t)\,dt] \\ \le \; & E[n^{-1}\sum_{j=1}^n \int_{\mathbb{R}}\Big(8(\sigma^2+n^{-1/2}v)^{-1}t^2(h'A_n^{-1/2}x_{n,j})^2 \\ & +8(\sigma^2+n^{-1/2}v)^{-1}\sigma^{-2}((\sigma^2+n^{-1/2}v)^{-1/2}+\sigma)^{-2}t^2\epsilon_{n,j}^2n^{-1}v^2\Big)\ \phi_\delta(t)\,dt \\ \le \; & E[n^{-1}\sum_{j=1}^n \int_{\mathbb{R}}\left(16\sigma^{-2}t^2(h'A_n^{-1/2}x_{n,j})^2 + 6\sigma^{-6}t^2\epsilon_{n,j}^2n^{-1}v^2\right)\ \phi_\delta(t)\,dt \\ = \; & n^{-1}\sigma^{-2}(16|h|^2 + 6\sigma^{-4}v^2E[\epsilon_{n,1}^2])\int_{\mathbb{R}} t^2\,\phi_\delta(t)\,dt, \end{aligned}$$

$$E[n^{-1}\sum_{1\le j<k\le n}\int_{\mathbb{R}} Y_{n,j}(\epsilon_{n,j},t,h,v)Y_{n,k}(\epsilon_{n,k},t,h,v)\phi_\delta(t)\,dt]=0$$

and

$$\begin{aligned}
& \operatorname{Var}\left(n^{-1}\sum_{1\le j<k\le n}\int_{\mathbb{R}} Y_{n,j}(\epsilon_{n,j},t,h,v)Y_{n,k}(\epsilon_{n,k},t,h,v)\phi_\delta(t)\,dt\right)\\
= & \; n^{-2}\sum_{1\le j<k\le n} E[\left(\int_{\mathbb{R}} Y_{n,j}(\epsilon_{n,j},t,h,v)Y_{n,k}(\epsilon_{n,k},t,h,v)\varphi_\delta(t)\,dt\right)^2]\\
\le & \; n^{-2}\sum_{1\le j<k\le n}\int_{\mathbb{R}} E[(Y_{n,j}(\epsilon_{n,j},t,h,v))^2]E[(Y_{n,k}(\epsilon_{n,k},t,h,v))^2]\phi_\delta(t)\,dt]\\
\le & \; 2^{-1}n^{-2}\sum_{j,k=1}^n\int_{\mathbb{R}} E[(Y_{n,j}(\epsilon_{n,j},t,h,v))^2]E[(Y_{n,k}(\epsilon_{n,k},t,h,v))^2]\phi_\delta(t)\,dt\\
\le & \; 2^{-1}n^{-2}\sum_{j,k=1}^n\int_{\mathbb{R}} E[(g(\epsilon_{n,j},t,h'A_n^{-1/2}x_{n,j},\sigma^2+n^{-1/2}v)-g(\epsilon_{n,j},t,0,\sigma^2))^2]\\
& \times E[(g(\epsilon_{n,k},t,h'A_n^{-1/2}x_{n,k},(\sigma^2+n^{-1/2}v)^{1/2})-g(\epsilon_{n,k},t,0,\sigma))^2]\phi_\delta(t)\,dt\\
\le & \; 2^{-1}n^{-2}\sum_{j,k=1}^n\int_{\mathbb{R}} E\left[16\sigma^{-2}t^2(h'A_n^{-1/2}x_{n,j})^2+6\sigma^{-6}t^2\epsilon_{n,j}^2n^{-1}v^2\right]\\
& \times E\left[16\sigma^{-2}t^2(h'A_n^{-1/2}x_{n,k})^2+6\sigma^{-6}t^2\epsilon_{n,k}^2n^{-1}v^2\right]\phi_\delta(t)\,dt\\
\le & \; cn^{-2}\left(\sum_{j=1}^n(h'A_n^{-1/2}x_{n,j})^2\right)^2+cn^{-2}\sum_{j=1}^n(h'A_n^{-1/2}x_{n,j})^2+cn^{-2}=cn^{-2},
\end{aligned}$$

where we have used that

$$\sum_{j=1}^n(h'A_n^{-1/2}x_{n,j})^2=\sum_{j=1}^n h'A_n^{-1/2}x_{n,j}x_{n,j}'A_n^{-1/2}h=|h|^2.$$

Hence, (4.26) follows. By (4.26),

$$\Pr\{\sup_{|h|,|v|\le M}\|U_n(t,\pi_1(h),\pi_2(v))-U_n(t,0,0)\|_{L_2(\phi_\delta)}\ge 2^{-1}\tau\}\to 0. \tag{4.28}$$

By (4.24), (4.25) and (4.28), for each $\eta>0$,

$$\begin{aligned}
& \limsup_{n\to\infty}\Pr\{\sup_{|h|,|v|\le M}\|U_n(t,h,v)-U_n(t,0,0)\|_{L_2(\phi_\delta)}\ge\tau\}\\
\le & \; 8\tau^{-1}\sigma^{-1}\eta\|t\|_{L_2(\phi_\delta)}(\sigma^{-2}E[|\epsilon_1|]+p^{1/2}).
\end{aligned}$$

Since $\eta>0$ is arbitrary the claim follows. □

Proof of Theorem 2.3. It is easy to see that

$$\begin{aligned}
& \|n^{-1/2}\sum_{j=1}^n(g(\epsilon_{n,j},t,0,\sigma)-e^{-2^{-1}t^2}\left(\sigma^{-1}t\epsilon_{n,j}-2^{-1}\sigma^{-2}t^2(\epsilon_{n,j}^2-\sigma^2)\right))\|_{L_2(\phi_\delta)}^2\\
= & \; n^{-1}\sum_{j,k=1}^n h(\sigma^{-1}\epsilon_{n,j},\sigma^{-1}\epsilon_{n,k}),
\end{aligned}$$

where

$$\begin{aligned}
h(a,b)=&\int_{\mathbb{R}}(\cos(ta)+\sin(ta)-\exp(-2^{-1}t^2)\left(1+ta-2^{-1}t^2(a^2-1)\right))\\
&\times(\cos(tb)+\sin(tb)-\exp(-2^{-1}t^2)\left(1+tb-2^{-1}t^2(b^2-1)\right))\phi_\delta(t)\,dt.
\end{aligned} \tag{4.29}$$

First, we prove that the functions h in (2.3) and in (4.29) agree. Using that ϕ_δ is an even function,

$$
\begin{aligned}
& \int_{\mathbb{R}}(\cos(ta)+\sin(ta)-\exp(-2^{-1}t^2)\,(1+ta-2^{-1}t^2(a^2-1))) \\
& \times(\cos(tb)+\sin(tb)-\exp(-2^{-1}t^2)\,(1+tb-2^{-1}t^2(b^2-1)))\,\phi_\delta(t)\,dt \\
= & \int_{\mathbb{R}}(2^{-1}(1-i)\exp(ita)+2^{-1}(1+i)\exp(-ita)-\exp(-2^{-1}t^2)\,(1+ta-2^{-1}t^2(a^2-1))) \\
& \times(2^{-1}(1-i)\exp(itb)+2^{-1}(1+i)\exp(-itb)-\exp(-2^{-1}t^2)\,(1+tb-2^{-1}t^2(b^2-1)))\,\phi_\delta(t)\,dt \\
= & \int_{\mathbb{R}}\exp(it(b-a))\phi_\delta(t)\,dt-\int_{\mathbb{R}}\exp(ita)(1+itb-2^{-1}t^2(b^2-1))\exp(-2^{-1}t^2)\phi_\delta(t)\,dt \\
& -\int_{\mathbb{R}}\exp(itb)(1+ita-2^{-1}t^2(a^2-1))\exp(-2^{-1}t^2)\phi_\delta(t)\,dt \\
& +\int_{\mathbb{R}}(1+ta-2^{-1}t^2(a^2-1))(1+tb-2^{-1}t^2(b^2-1))\exp(-t^2)\phi_\delta(t)\,dt \\
= & \int_{\mathbb{R}}\exp(it(b-a))\phi_\delta(t)\,dt \\
& -(\delta^2+1)^{-1/2}\int_{\mathbb{R}}\exp(ita)(1+tb-2^{-1}t^2(b^2-1))\phi_{\delta(\delta^2+1)^{-1/2}}(t)\,dt \\
& -(\delta^2+1)^{-1/2}\int_{\mathbb{R}}\exp(itb)(1+ta-2^{-1}t^2(a^2-1))\phi_{\delta(\delta^2+1)^{-1/2}}(t)\,dt \\
& +(2\delta^2+1)^{-1/2}\int_{\mathbb{R}}(1+ta-2^{-1}t^2(a^2-1))(1+tb-2^{-1}t^2(b^2-1))\phi_{\delta(2\delta^2+1)^{-1/2}}(t)\,dt \\
= & \exp(-2^{-1}\delta^2(a-b)^2) \\
& -(\delta^2+1)^{-1/2}\exp(-2^{-1}(\delta^2+1)^{-1}\delta^2a^2) \\
& \times(1+(\delta^2+1)^{-1}\delta^2b-2^{-1}(\delta^2+1)^{-2}\delta^2(b^2-1)\,(1-(\delta^2+1)^{-2}\delta^2a^2)) \\
& -(\delta^2+1)^{-1/2}\exp(-2^{-1}(\delta^2+1)^{-1}\delta^2b^2) \\
& \times(1+(\delta^2+1)^{-1}\delta^2a-2^{-1}(\delta^2+1)^{-2}\delta^2(a^2-1)\,(1-(\delta^2+1)^{-2}\delta^2b^2)) \\
& +(2\delta^2+1)^{-1/2}\,(1+(2\delta^2+1)^{-1}\delta^2(1-2^{-1}(a-b)^2)+2^{-2}3(2\delta^2+1)^{-2}\delta^4(a^2-1)(b^2-1))\,,
\end{aligned}
$$

where we have used that

$$
\begin{aligned}
& \int_{\mathbb{R}}\exp(ita)\phi_\delta(t)\,dt=\exp(-2^{-1}a^2\delta^2), \\
& \int_{\mathbb{R}}\exp(ita)it\phi_\delta(t)\,dt=-\exp(-2^{-1}a^2\delta^2)\delta^2a, \\
& \int_{\mathbb{R}}\exp(ita)t^2\phi_\delta(t)\,dt=\exp(-2^{-1}a^2\delta^2)\delta^2(1-a^2\delta^2), \\
& \int_{\mathbb{R}}t^2\phi_\delta(t)\,dt=\delta^2 \quad\text{and}\quad \int_{\mathbb{R}}t^4\phi_\delta(t)\,dt=3\delta^4.
\end{aligned}
$$

Hence, the functions h in (2.3) and in (4.29) agree.

By the Gregory's theorem for V–statistics over contiguous sequences,

$$
n^{-1}\sum_{j,k=1}^{n}h(\sigma^{-1}\epsilon_{n,j},\sigma^{-1}\epsilon_{n,k})
$$

converges in distribution with the limit given by (2.5). Hence, to prove (2.6), it suffices to prove that

$$
n^{1/2}\hat{D}_n^{1/2}-\|n^{-1/2}\sum_{j=1}^{n}(g(\epsilon_{n,j},t,0,\sigma)+e^{-2^{-1}t^2}\left(2^{-1}\sigma^{-2}t^2(\epsilon_{n,j}^2-\sigma^2)+\sigma^{-1}t\epsilon_{n,j})\right)\|_{L_2(\phi_\delta)} \stackrel{\Pr}{\to} 0. \tag{4.30}
$$

It is easy to see that

$$\begin{aligned}
&\hat{D}_n \\
&\quad= n^{-2}\textstyle\sum_{j,k=1}^n \int_{\mathbb{R}} (\cos(\hat{\sigma}_n^{-1}t(\hat{\epsilon}_{n,j}-\hat{\epsilon}_{n,k})) \\
&\qquad -2\cos(\hat{\sigma}_n^{-1}t\hat{\epsilon}_{n,j})\exp(-2^{-1}t^2)+\exp(-t^2))\,\phi_\delta(t)\,dt \\
&\quad= n^{-2}\textstyle\sum_{j,k=1}^n \int_{\mathbb{R}}(\cos(\hat{\sigma}_n^{-1}t\hat{\epsilon}_{n,j})+\sin(\hat{\sigma}_n^{-1}t\hat{\epsilon}_{n,j})-\exp(-2^{-1}t^2)) \\
&\qquad \times(\cos(\hat{\sigma}_n^{-1}t\hat{\epsilon}_{n,j})+\sin(\hat{\sigma}_n^{-1}t\hat{\epsilon}_{n,j})-\exp(2^{-1}t^2))\phi_\delta(t)\,dt \\
&\quad= n^{-2}\textstyle\sum_{j,k=1}^n \int_{\mathbb{R}} g(\epsilon_{n,j},t,\epsilon_{n,j}-\hat{\epsilon}_{n,j},\hat{\sigma}_n)g(\epsilon_{n,k},t,\epsilon_{n,k}-\hat{\epsilon}_{n,k},\hat{\sigma}_n)\phi_\delta(t)\,dt \\
&\quad= \|n^{-1}\textstyle\sum_{j=1}^n g(\epsilon_{n,j},t,(\hat{\beta}_n-\beta)'x_{n,j},\hat{\sigma}_n)\|_{L_2(\phi_\delta)}^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
&|n^{1/2}\hat{D}_n^{1/2}-\|n^{-1/2}\textstyle\sum_{j=1}^n(g(\epsilon_{n,j},t,0,\sigma)+e^{-2^{-1}t^2}\left(2^{-1}\sigma^{-2}t^2(\epsilon_{n,j}^2-\sigma^2)+\sigma^{-1}t\epsilon_{n,j}\right))\|_{L_2(\phi_\delta)}| \\
\leq\ & \|n^{-1/2}\textstyle\sum_{j=1}^n(g(\epsilon_{n,j},t,(\hat{\beta}_n-\beta)'x_{n,j},\hat{\sigma}_n)-g(\epsilon_{n,j},t,0,\sigma) \\
&\quad -G_n(t,(\hat{\beta}_n-\beta)'x_{n,j},\hat{\sigma}_n)+G_n(t,0,\sigma))\|_{L_2(\phi_\delta)} \\
&\quad +\|n^{-1/2}\textstyle\sum_{j=1}^n(G_n(t,(\hat{\beta}_n-\beta)'x_{n,j},\hat{\sigma}_n)-G_n(t,0,\sigma) \\
&\quad -\exp(-2^{-1}t^2)(2^{-1}\sigma^{-2}t^2(\hat{\sigma}_n^2-\sigma^2)-\sigma^{-1}t(\hat{\beta}_n-\beta)'x_{n,j}))\|_{L_2(\phi_\delta)} \\
&\quad +\|e^{-2^{-1}t^2}2^{-1}\sigma^{-2}t^2(n^{1/2}(\hat{\sigma}_n^2-\sigma^2)-n^{-1/2}\textstyle\sum_{j=1}^n(\epsilon_{n,j}^2-\sigma^2))\|_{L_2(\phi_\delta)} \\
&\quad +\|n^{-1/2}\textstyle\sum_{j=1}^n\sigma^{-1}t\exp(-2^{-1}t^2)\left((\hat{\beta}_n-\beta)'x_{n,j}-\epsilon_{n,j}\right)\|_{L_2(\phi_\delta))} \\
=:\ & I+II+III+IV.
\end{aligned}$$

Hence, to prove (4.30), it suffices to show that $I, II, III, IV \xrightarrow{\Pr} 0$.

By Lemma 4.3, for each $0 < M < \infty$,

$$\begin{aligned}
&\sup_{|h|\leq M,|v|\leq M}\|n^{-1/2}\sum_{j=1}^n(g(\epsilon_{n,j},t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2}) \qquad (4.31)\\
&-G_n(t,h'A_n^{-1/2}x_{n,j},(\sigma^2+n^{-1/2}v)^{1/2})-g(\epsilon_{n,j},t,0,\sigma)+G_n(t,0,\sigma))\|_{L_2(\phi_\delta)} \xrightarrow{\Pr} 0.
\end{aligned}$$

Plugging $(A_n^{1/2}(\hat{\beta}_n-\beta), n^{1/2}(\hat{\sigma}_n^2-\sigma^2))$ as (h,v) in (4.31), we get that

$$I \xrightarrow{\Pr} 0. \qquad (4.32)$$

Notice that by Lemma 4.2, $(A_n^{1/2}(\hat{\beta}_n-\beta), n^{1/2}(\hat{\sigma}_n^2-\sigma^2)) = O_P(1)$.

We have that

$$G_n(t,\mu,s) = G(t,\mu,s)+n^{-1/2}\int_{\mathbb{R}}(\cos(s^{-1}t(y-\mu))+\sin(s^{-1}t(y-\mu)))\alpha_n(y)\phi_\sigma(y)\,dy,$$

where $G(t,\mu,s) = E[g(\epsilon_1,t,\mu,s)]$. To prove that $II \xrightarrow{\Pr} 0$, we show that

$$\begin{aligned}
&\|n^{-1/2}\textstyle\sum_{j=1}^n(G(t,(\hat{\beta}_n-\beta)'x_{n,j},\hat{\sigma}_n)-G(t,0,\sigma) \qquad (4.33)\\
&-\exp(-2^{-1}t^2)(2^{-1}\sigma^{-2}t^2(\hat{\sigma}_n^2-\sigma^2)-\sigma^{-1}t(\hat{\beta}_n-\beta)'x_{n,j}))\|_{L_2(\phi_\delta)} \xrightarrow{\Pr} 0
\end{aligned}$$

and

$$\|n^{-1}\sum_{j=1}^n(\int_{\mathbb{R}}(\cos(\hat{\sigma}_n^{-1}t(y-(\hat{\beta}_n-\beta)'x_{n,j}))+\sin(\hat{\sigma}_n^{-1}t(y-(\hat{\beta}_n-\beta)'x_{n,j})) \\ -\cos(\sigma^{-1}ty)-\sin(\sigma^{-1}ty))\alpha_n(y)\phi_\sigma(y)\,dy\|_{L_2(\phi_\delta)}\overset{\mathrm{Pr}}{\to}0. \tag{4.34}$$

Since

$$\begin{aligned} & E[\exp(is^{-1}t(\epsilon_1-\mu)] = \exp(-is^{-1}t\mu-2^{-1}t^2s^{-2}\sigma^2) \\ = \ & \exp(-2^{-1}t^2s^{-2}\sigma^2)(\cos(s^{-1}t\mu)-i\sin(s^{-1}t\mu)), \end{aligned}$$

$$G(t,\mu,s)=\exp(-2^{-1}t^2s^{-2}\sigma^2)(\cos(s^{-1}t\mu)-\sin(s^{-1}t\mu))-\exp(-2^{-1}t^2).$$

Hence,

$$\begin{aligned} & G(t,\mu,s)-G(t,0,\sigma)-\exp(-2^{-1}t^2)(2^{-1}\sigma^{-2}t^2(s^2-\sigma^2)-\sigma^{-1}t\mu) \qquad (4.35) \\ = \ & \exp(-2^{-1}t^2s^{-2}\sigma^2)(\cos(s^{-1}t\mu)-\sin(s^{-1}t\mu)) \\ & -\exp(-2^{-1}t^2)-\exp(-2^{-1}t^2)(2^{-1}\sigma^{-2}t^2(s^2-\sigma^2)-\sigma^{-1}t\mu) \\ = \ & \exp(-2^{-1}t^2)(\exp(-2^{-1}t^2(s^{-2}\sigma^2-1))-1+2^{-1}t^2(s^{-2}\sigma^2-1))(\cos(s^{-1}t\mu)-\sin(s^{-1}t\mu)) \\ & +\exp(-2^{-1}t^2)(1-2^{-1}t^2(s^{-2}\sigma^2-1))(\cos(s^{-1}t\mu)-\sin(s^{-1}t\mu)-1+s^{-1}t\mu) \\ & +\exp(-2^{-1}t^2)2^{-1}t^2(-(s^{-2}\sigma^2-1)-\sigma^{-2}(s^2-\sigma^2)) \\ & +\exp(-2^{-1}t^2)(s^{-1}t\mu-\sigma^{-1}t\mu) \\ & +\exp(-2^{-1}t^2)2^{-1}t^2(s^{-2}\sigma^2-1)s^{-1}t\mu \\ =: \ & V+VI+VII+VIII+IX. \end{aligned}$$

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

$$|V|\le\exp(-2^{-1}t^2+2^{-1}t^2|s^{-2}\sigma^2-1|)2^{-2}t^4s^{-4}(s^2-\sigma^2)^2.$$

We also have that

$$\begin{aligned} |VI| &\le a\exp(-2^{-1}t^2)(1+2^{-1}t^2s^{-2}|s^2-\sigma^2|)\mu^2, \\ |VII| &\le 2\exp(-2^{-1}t^2)2^{-1}t^2|s^2-\sigma^2||s^{-2}-\sigma^{-2}| = \exp(-2^{-1}t^2)t^2s^{-2}\sigma^{-2}(s^2-\sigma^2)^2, \\ |VIII| &\le 2^{-1}\exp(-2^{-1}t^2)|t|^3s^{-3}|s^2-\sigma^2|\mu|, \end{aligned}$$

where $a:=\sup_{x\neq 0}x^{-2}|\cos x-1|+\sup_{x\neq 0}x^{-2}|\sin x-x|$. Hence,

$$\begin{aligned} & |G(t,\mu,s)-G(t,0,\sigma)-\exp(-2^{-1}t^2)(2^{-1}\sigma^{-2}(s^2-\sigma^2)-\sigma^{-1}t\mu)| \qquad (4.36) \\ \le \ & \exp(-2^{-1}t^2+2^{-1}t^2s^{-2}|s^2-\sigma^2|)2^{-2}t^4s^{-4}(s^2-\sigma^2)^2 \\ & +a\exp(-2^{-1}t^2)(1+2^{-1}t^2)s^{-2}|s^2-\sigma^2|\mu^2, \\ & +\exp(-2^{-1}t^2)t^2s^{-2}\sigma^{-2}(s^2-\sigma^2)^2 \\ & +2^{-1}\exp(-2^{-1}t^2)|t|^3s^{-3}|s^2-\sigma^2|\mu| \end{aligned}$$

and

$$\begin{aligned}
& \|n^{-1/2} \textstyle\sum_{j=1}^n (G(t, (\hat\beta_n - \beta)' x_{n,j}, \hat\sigma_n) - G(t, 0, \sigma) \\
& \quad - \exp(-2^{-1}t^2)(2^{-1}\sigma^{-2}t^2(\hat\sigma_n^2 - \sigma^2) - \sigma^{-1}t(\hat\beta_n - \beta)' x_{n,j}))\|_{L_2(\phi_\delta)} \qquad (4.37) \\
\leq \; & 2^{-2}\| \exp(-2^{-1}t^2 + 2^{-1}t^2\hat\sigma_n^{-2}|\hat\sigma_n^2 - \sigma^2|)t^4\|_{L_2(\phi_\delta)} \hat\sigma_n^{-4} n^{1/2}(\hat\sigma_n^2 - \sigma^2)^2 \\
& + a \left(\| \exp(-2^{-1}t^2)\|_{L_2(\phi_\delta)} + 2^{-1}\| \exp(-2^{-1}t^2)t^2\|_{L_2(\phi_\delta)}\hat\sigma_n^{-2}|\hat\sigma_n^2 - \sigma^2|\right) n^{-1/2} \textstyle\sum_{j=1}^n ((\hat\beta_n - \beta)' x_{n,j})^2 \\
& + \| \exp(-2^{-1}t^2)t^2\|_{L_2(\phi_\delta)} \hat\sigma_n^{-2}\sigma^{-2}(\hat\sigma_n^2 - \sigma^2)^2 \\
& + \| \exp(-2^{-1}t^2)|t|^3\|_{L_2(\phi_\delta)} \hat\sigma_n^{-3}|\hat\sigma_n^2 - \sigma^2| n^{-1/2} \textstyle\sum_{j=1}^n |(\hat\beta_n - \beta)' x_{n,j}|,
\end{aligned}$$

which tends in probability to zero, because, by Lemma 4.2, $n^{1/2}(\hat\sigma_n^2 - \sigma^2) = O_P(1)$,

$$\sum_{j=1}^n ((\hat\beta_n - \beta)' x_{n,j})^2 = \sum_{j=1}^n (\hat\beta_n - \beta)' x_{n,j} x_{n,j}' (\hat\beta_n - \beta) = (\hat\beta_n - \beta)' A_n (\hat\beta_n - \beta) = O_P(1)$$

and

$$n^{-1/2} \sum_{j=1}^n |(\hat\beta_n - \beta)' x_{n,j}| \leq \left(\sum_{j=1}^n ((\hat\beta_n - \beta)' x_{n,j})^2 \right)^{1/2} = O_P(1).$$

Therefore, (4.33) follows.

For each $a, h \in \mathbb{R}$, we have that

$$\begin{aligned}
& |\sin(a+h) - \sin(a) - h\cos(a)| \\
\leq \; & |\sin(a)\cos(h) - \sin(a)| + |\cos(a)\sin(h) - h\cos(a)| \leq ch^2
\end{aligned}$$

and

$$\begin{aligned}
& |\cos(a+h) - \cos(a) + h\sin(a)| \\
\leq \; & |\cos(a)\cos(h) - \cos(a)| + |\sin(a)\sin(h) - h\sin(a)| \leq ch^2.
\end{aligned}$$

So,

$$\begin{aligned}
& \|n^{-1} \textstyle\sum_{j=1}^n (\int_{\mathbb{R}} (\cos(\hat\sigma_n^{-1} t(y - (\hat\beta_n - \beta)' x_{n,j})) + \sin(\hat\sigma_n^{-1} t(y - (\hat\beta_n - \beta)' x_{n,j})) \\
& - \cos(\sigma^{-1}ty) - \sin(\sigma^{-1}ty))\alpha_n(y)\phi_\sigma(y)\,dy\|_{L_2(\phi_\delta)} \\
\leq \; & \|n^{-1} \textstyle\sum_{j=1}^n \int_{\mathbb{R}} (\cos(\sigma^{-1}ty) - \sin(\sigma^{-1}ty)) \\
& \times \left(\hat\sigma_n^{-1} t(y - (\hat\beta_n - \beta)' x_{n,j}) - \sigma^{-1}ty\right) \alpha_n(y)\phi_\sigma(y)\,dy\|_{L_2(\phi_\delta)} \\
& + c\|n^{-1} \textstyle\sum_{j=1}^n \int_{\mathbb{R}} \left(\hat\sigma_n^{-1} t(y - (\hat\beta_n - \beta)' x_{n,j}) - \sigma^{-1}ty\right)^2 \alpha_n(y)\phi_\sigma(y)\,dy\|_{L_2(\phi_\delta)} \\
\leq \; & \int_{\mathbb{R}} |y||\alpha_n(y)|\phi_\sigma(y)\,dy \|t\|_{L_2(\phi_\delta)} |\hat\sigma_n^{-1} - \sigma^{-1}| \\
& + \int_{\mathbb{R}} |\alpha_n(y)|\phi_\sigma(y)\,dy \|t\|_{L_2(\phi_\delta)} |\hat\sigma_n^{-1}| |n^{-1} \textstyle\sum_{j=1}^n (\hat\beta_n - \beta)' x_{n,j}| \\
& + c\int_{\mathbb{R}} y^2 \alpha_n(y)\phi_\sigma(y)\,dy \|t^2\|_{L_2(\phi_\delta)} (\hat\sigma_n^{-1} - \sigma^{-1})^2 \\
& + c\int_{\mathbb{R}} |\alpha_n(y)|\phi_\sigma(y)\,dy \|t^2\|_{L_2(\phi_\delta)} \hat\sigma_n^{-2} n^{-1} \textstyle\sum_{j=1}^n ((\hat\beta_n - \beta)' x_{n,j})^2 \\
& \xrightarrow{\Pr} 0,
\end{aligned}$$

because for $k = 0, 1, 2$,

$$\begin{aligned} & \int_{\mathbb{R}} |y|^k |\alpha_n(y)| \phi_\sigma(y)\,dy \\ \leq\ & \int_{\mathbb{R}} |y|^k |\alpha(y)| \phi_\sigma(y)\,dy + \int_{\mathbb{R}} |y|^k |\alpha_n(y) - \alpha(y)| \phi_\sigma(y)\,dy \\ \leq\ & \left(\int_{\mathbb{R}} (\alpha(y))^2 \phi_\sigma(y)\,dy \int_{\mathbb{R}} y^{2k} \phi_\sigma(y)\,dy\right)^{1/2} + \left(\int_{\mathbb{R}} (\alpha_n(y) - \alpha(y))^2 \phi_\sigma(y)\,dy \int_{\mathbb{R}} y^{2k} \phi_\sigma(y)\,dy\right)^{1/2} \\ =\ & 0(1), \end{aligned}$$

$$n^{-1} \sum_{j=1}^n ((\hat{\beta}_n - \beta)' x_{n,j})^2 = n^{-1} (\hat{\beta}_n - \beta)' A_n (\hat{\beta}_n - \beta) \xrightarrow{\Pr} 0$$

and

$$n^{-1} \sum_{j=1}^n (\hat{\beta}_n - \beta)' x_{n,j} \xrightarrow{\Pr} 0. \tag{4.38}$$

By (1.4),

$$n^{-1/2} \sum_{j=1}^n (\hat{\beta}_n - \beta)' x_{n,j} = n^{-1/2} \sum_{j,k=1}^n x'_{n,k} A_n^{-1} x_{n,j} \epsilon_{n,k}.$$

We have that

$$\begin{aligned} & E[n^{-1/2} \textstyle\sum_{j,k=1}^n x'_{n,k} A_n^{-1} x_{n,j} \epsilon_{n,k} - n^{-1/2} \sum_{k=1}^n \epsilon_{n,k}] \\ =\ & \left(n^{-1} \textstyle\sum_{j,k=1}^n x'_{n,k} A_n^{-1} x_{n,j} - 1\right) n^{1/2} E[\epsilon_{n,1}] \to 0. \end{aligned} \tag{4.39}$$

and

$$\begin{aligned} & \mathrm{Var}(n^{-1/2} \textstyle\sum_{j,k=1}^n x'_{n,k} A_n^{-1} x_{n,j} \epsilon_{n,k} - n^{-1/2} \sum_{k=1}^n \epsilon_{n,k}) \\ =\ & (\sigma^2 + o(1)) n^{-1} \textstyle\sum_{k=1}^n \left(\sum_{j=1}^n x'_{n,k} A_n^{-1} x_{n,j} - 1\right)^2 \\ =\ & (\sigma^2 + o(1)) \left(n^{-1} \textstyle\sum_{k=1}^n \left(\sum_{j=1}^n x'_{n,k} A_n^{-1} x_{n,j}\right)^2 - 2\sigma^2 n^{-1} \sum_{k=1}^n \sum_{j=1}^n x'_{n,k} A_n^{-1} x_{n,j} + 1\right) \\ =\ & (\sigma^2 + o(1)) \left(n^{-1} \textstyle\sum_{k=1}^n \sum_{j=1}^n \sum_{l=1}^n x'_{n,j} A_n^{-1} x_{n,k} x'_{n,k} A_n^{-1} x_{n,l} - 2n^{-1} \sum_{k=1}^n \sum_{j=1}^n x'_{n,k} A_n^{-1} x_{n,j} + 1\right) \\ =\ & (\sigma^2 + o(1)) \left(1 - n^{-1} \textstyle\sum_{k=1}^n \sum_{j=1}^n x'_{n,k} A_n^{-1} x_{n,j}\right) \to 0, \end{aligned} \tag{4.40}$$

using condition (B). Hence, (4.38) follows. Therefore, (4.34) follows.

By Lemma 4.2,

$$\begin{aligned} III &= \| e^{-2^{-1} t^2} 2^{-1} \sigma^{-2} t^2 (n^{1/2} (\hat{\sigma}_n^2 - \sigma^2) - n^{-1/2} \textstyle\sum_{j=1}^n (\epsilon_{n,j}^2 - \sigma^2)) \|_{L_2(\phi_\delta)} \\ &\leq |n^{1/2} (\hat{\sigma}_n^2 - \sigma^2) - n^{-1/2} \textstyle\sum_{j=1}^n (\epsilon_{n,j}^2 - \sigma^2)| \| e^{-2^{-1} t^2} 2^{-1} \sigma^{-2} t^2 \|_{L_2(\phi_\delta)} \xrightarrow{\Pr} 0. \end{aligned} \tag{4.41}$$

By (4.39) and (4.40),

$$\begin{aligned} IV &\leq \| n^{-1/2} \textstyle\sum_{j=1}^n \sigma^{-1} t \exp(-2^{-1} t^2) ((\hat{\beta}_n - \beta)' x_{n,j} - \epsilon_{n,j}) \|_{L_2(\phi_\delta)} \\ &\leq |n^{-1/2} \textstyle\sum_{j=1}^n ((\hat{\beta}_n - \beta)' x_{n,j} - \epsilon_{n,j})| \| \sigma^{-1} t \exp(-2^{-1} t^2) \|_{L_2(\phi_\delta)} \xrightarrow{\Pr} 0. \end{aligned} \tag{4.42}$$

Therefore, (4.30) follows from (4.32)–(4.34), (4.41) and (4.42). □

References

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

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

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

[4] Gregory, G. G. (1977). Large sample theory for U–statistics and tests of fit. *Ann. Statist.* **5** 110–123.

[5] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

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

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

[8] Jurečková, J.; Picek, J. and Sen, P. K. (2003). Goodness–of–fit test with nuisance regression and scale. *Metrika* **58** 235–258.

[9] Lee, A. J. (1990). *U–statistics, Theory and Practice.* Marcel Dekker, Inc., New York.

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

[11] Neter, J.; Wasserman, W.; Kutner, M. H.; Nachtsheim, C.; and Nachtsheim, C. J. (1996). *Applied Linear Statistical Models* 4–th Edition. McGraw–Hill, New York.

[12] Sen, P. K.; Jurečková, J.; and Picek, J. (2003). Goodness–of–fit test of Shapiro–Wilk type with nuisance regression and scale. *Austrian J. of Statist.* **32** No 1 & 2 163-177.

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

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