Far East Journal of Theoretical Statistics
Volume 28, Number 1, 2009, Pages 57-106
Published Online: June 4, 2009
This paper is available online at http://www.pphmj.com

SOME NEW NORMALITY TESTS FOR THE ERROR OF A LINEAR REGRESSION MODEL

MIGUEL A. ARCONES and YISHI WANG

Department of Mathematical Sciences
Binghamton University
Binghamton, NY 13902, U. S. A.
e-mail: arcones@math.binghamton.edu

Department of Mathematics and Computer Science
Western Carolina University
Cullowhee, NC 28723, U. S. A.

Abstract

We present two new normality tests for the error of a linear regression model. The tests are obtained by applying the normality tests in Arcones and Wang [3] to the residuals obtained using the least squares estimators. We show that the considered tests are omnibus. We also obtain the limit distribution of the considered tests under the null hypothesis. Simulations show that the power of the presented tests is competitive with common normality tests.

1. Introduction

In this paper, we apply the normality tests in Arcones and Wang [3] to the linear regression model. We consider the linear regression model: $Y_{n,j} = \beta' x_{n,j} + \varepsilon_j$, $1 \le j \le n$, where $\{\varepsilon_j\}_{j=1}^n$ 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

2000 Mathematics Subject Classification: Primary 62J05; Secondary 62J20, 62F03.

Keywords and phrases: test of normality, linear regression model, residuals, U-processes.

Received July 9, 2007; Revised November 7, 2007

an unknown parameter. $x_{n,j}$ is called the *regressor* or *predictor variable*. Depending on which characteristic of the linear regression we are interested in, there are different ways to select the regressor design $\{x_{n,j} : 1 \le j \le n\}$ (see Sections 1.8 and 1.9 in Draper and Smith [10]). That is why we allow the regressor design to depend on n. $Y_{n,j}$ is called the *response variable*. ε_j is an error variable with mean zero. Let F be the c.d.f. of the sequence $\{\varepsilon_j\}_{j=1}^{\infty}$. We study the testing problem

$$H_0 : F \text{ has a normal distribution with mean zero,}$$
$$\text{versus } H_1 : F \text{ does not.} \tag{1.1}$$

Let $\mathbf{X}$ be the $n \times p$ matrix having by j-th row $x'_{n,j}$. If $\mathbf{X'X}$ is invertible, then the (LS) least squares estimator of β is uniquely determined and it is $\hat{\beta}_n := (\mathbf{X'X})^{-1}\mathbf{X'Y}$, where $\mathbf{Y} := (Y_{n,1}, \ldots, Y_{n,n})'$ (see, e.g., Section 3.2 in Rao and Toutenburg [22]). The least squares estimator has several optimality properties (see, e.g., Section 3.3 in Rao and Toutenburg [22]). It is easy to see that $\mathbf{X'X} = \sum_{j=1}^{n} x_{n,j}x'_{n,j}$. In order that the LS estimator of β is well uniquely defined, we assume that the design of regressors satisfies the following condition:

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

To estimate the distribution F of the errors we use the (OLS) ordinary least squares residuals $\hat{\varepsilon}_{n,j} := Y_{n,j} - \hat{\beta}'_n x_{n,j}, 1 \le j \le n$. Observe that

$$\hat{\beta}_n - \beta = \sum_{j=1}^{n} A_n^{-1} x_{n,j} Y_{n,j} - \beta = \sum_{j=1}^{n} \varepsilon_j A_n^{-1} x_{n,j} \tag{1.2}$$

and

$$\hat{\varepsilon}_{n,j} = Y_{n,j} - \hat{\beta}'_n x_{n,j} = \varepsilon_j - (\hat{\beta}_n - \beta)' x_{n,j}$$

$$= \varepsilon_j - \sum_{k=1}^{n} \varepsilon_k x'_{n,k} A_n^{-1} x_{n,j} = \varepsilon_j - \sum_{k=1}^{n} \varepsilon_k x'_{n,j} A_n^{-1} x_{n,k}, \tag{1.3}$$

where we have used that A_n is a symmetric matrix. Hence

$$\sum_{j=1}^{n} \hat{\varepsilon}_{n,j} x_{n,j} = \sum_{j=1}^{n} \varepsilon_j x_{n,j} - \sum_{j=1}^{n} \sum_{k=1}^{n} \varepsilon_k x_{n,j} x'_{n,j} A_n^{-1} x_{n,k}$$

$$= \sum_{j=1}^{n} \varepsilon_j x_{n,j} - \sum_{k=1}^{n} \varepsilon_k x_{n,k} = 0. \tag{1.4}$$

This implies that the residual vector $(\hat{\varepsilon}_{n,1}, \ldots, \hat{\varepsilon}_{n,n})' \in \mathbb{R}^n$ lies in a subspace of dimension less than or equal to $n - p$ and $\{\hat{\varepsilon}_{n,j}\}_{j=1}^{n}$ are not independent r.v.'s. Most of the normality tests for the errors of regression model are obtained by applying classical normality tests to the residuals. An alternative to use the residuals is to use the Theil's (BLUS) best linear unbiased scalar residuals. The Theil's BLUS residuals are independent. Huang and Bolch [14] noticed that the power of the Shapiro-Wilk test is higher using the OLS residuals than the Theil's BLUS residuals. Thus, we base our tests on the OLS residuals $\hat{\varepsilon}_{n,j}, 1 \le j \le n$. Notice that if the term $\sum_{k=1}^{n} \varepsilon_k x'_{n,k} A_n^{-1} x_{n,j}$ in (1.3) is asymptotically negligible, the distribution of a statistic based on the residuals is asymptotically equivalent to the distribution of the statistic based on the unknown errors.

By the Lévy characterization of the normal distribution, a c.d.f. F with finite second moment has a normal distribution with mean zero and variance $\sigma^2 > 0$, if and only if for some $m \ge 1$,

$$D_m(F) := \sup_{t \in \mathbb{R}} \left| P_F \left\{ \sigma^{-1} m^{-1/2} \sum_{j=1}^{m} \varepsilon_j \le t \right\} - \Phi(t) \right| = 0, \tag{1.5}$$

where Φ is the c.d.f. of a standard normal distribution and P_F is the probability for which the i.i.d. r.v.'s $\varepsilon_1, \ldots, \varepsilon_m$ have distribution F.

Let

$$D_{n,m} := \sup_{t\in\mathbb{R}} \left| \frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m)\in I_m^n} I\left(\hat{\sigma}_n^{-1} m^{-1/2} \sum_{j=1}^{m} \hat{\varepsilon}_{n,i_j} \le t \right) - \Phi(t) \right|, \quad (1.6)$$

where

$$I_m^n = \{(i_1, \dots, i_m) \in \mathbb{N}^m : 1 \le i_j \le n, \, i_j \ne i_k \text{ if } j \ne k\}$$

and

$$\hat{\sigma}_n^2 := \frac{1}{n-p} \sum\nolimits_{j=1}^{n} \hat{\varepsilon}_{n,j}^2.$$

It is easy to see that the distribution of $D_{n,m}$ is invariant by changes of scale on the error variable. Given a design of regressors $\{x_{n,j}\}_{j=1}^n$, the distribution of $D_{n,m}$ is the same for all normal distributions with mean zero.

Given $1 > \alpha > 0$, let

$$b_{n,m,\alpha} = \inf\{\lambda \ge 0 : P_\Phi\{D_{n,m} < \lambda\} \ge 1 - \alpha\},$$

where P_Φ is the probability measure for which the errors $\varepsilon_1, \dots, \varepsilon_n$ are i.i.d. r.v.'s with a standard normal distribution. Notice that $b_{n,m,\alpha}$ depends on the regressors design $\{x_{n,j}\}_{j=1}^n$. The proposed test rejects the null hypothesis if $D_{n,m} \ge b_{n,m,\alpha}$. Hence, the probability of type I error of the test is less than or equal to α.

We also have that F has a normal distribution with mean zero if and only if for some $m \ge 2$,

$$\widetilde{D}_m(F) := \sup_{t\in\mathbb{R}} \left| P_F\left\{ m^{-1/2} \sum_{j=1}^{m} X_j \le t \right\} - P_F\{X_1 \le t\} \right| = 0. \quad (1.7)$$

An estimator of the previous quantity is

$$\widetilde{D}_{n,m} := \sup_{t\in\mathbb{R}}\left|\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(m^{-1/2}\sum_{j=1}^{m}\hat{\varepsilon}_{n,i_j} \le t\right) - \frac{1}{n}\sum_{j=1}^{n} I(\hat{\varepsilon}_{n,j} \le t)\right|. \quad (1.8)$$

Given $1 > \alpha > 0$, let

$$c_{n,m,\alpha} = \inf\{\lambda \ge 0 : P_{\Phi}\{\widetilde{D}_{n,m} < \lambda\} \ge 1-\alpha\}.$$

Then, the test rejects the null hypothesis if $\widetilde{D}_{n,m} \ge c_{n,m,\alpha}$.

The two tests above are constructed by doing a minor variation of the tests in Arcones and Wang [3]. These normality tests are based on distribution functions as several normality tests do, like the normality test in Lilliefors [16]. There exists a large literature applying empirical processes to goodness-of-fit tests (see Stephens [25], del Barrio et al. [9] and del Barrio [8]).

The asymptotic distribution of the previous test statistics is similar to that of the empirical distribution based on the residuals. Pierce and Kopecky [20] obtained the asymptotic distribution of the empirical distribution function based on the OLS residuals, when $x_i' = (1, x_{i,2}, \ldots, x_{i,p})$, for each $1 \le i \le n$. Pierce and Kopecky [20] obtained that the estimation of regression parameters has no additional effect on the limiting distribution of the normality test based on empirical process. Loynes [17] considered the asymptotic properties of the empirical distribution function obtained from the residuals from a generalized regression model. White and Macdonald [27] argued that for several normality tests (such as the D'Agostino's test) the convergence of the test using the OLS residuals is quite fast. However, Weisberg [26] noticed that the effect of the regressors design is significative for the Shapiro and Wilk test. Jure□ková et al. [15] and Sen et al. [23] considered the asymptotics of normality tests for the residuals of a linear regression model using the approach in Shapiro and Wilk.

In view of the remarks above, the cutpoints of each of our tests are based on a distribution of a test statistic which depends on the regression design. In Section 2, we prove that the two presented tests are omnibus. The power of these tests tends to one for any alternative hypotheses as the sample size goes to infinity. We also obtain the limit distribution of the test statistics under the null hypothesis. As expected, under certain conditions, this limit distribution does not depend on the regressors design. In Section 3, we present the outcome of several simulations. The cutpoints and the power of the tests for several regression design matrices and for sample sizes 8, 16, 20 and 24 are presented. Our results show that the test based on $D_{n,m}$ is competitive with usual common normality tests. However, the test $\widetilde{D}_{n,m}$ seems to behave badly. Section 4 contains several results on the asymptotics of U-processes, which are of independent interest. We study the asymptotic normality of U-processes based on independent (not necessarily identically distributed) r.v.'s and kernels varying from occurrence to occurrence. Our results generalize to the case of U-processes the work in Pollard [21] and Arcones [1]. The proofs of the theorems are given in Section 5.

We will use the usual multivariate notation. For example, given $u = (u_1, \dots, u_d)' \in \mathbb{R}^d$, $|u| = \left(\sum_{j=1}^n u_j^2\right)^{1/2} \cdot$ I_p denotes the $p \times p$ identity matrix. c will denote a constant which may vary from occurrence to occurrence.

We also will denote expectation and sample means using the functional notation common in empirical processes. Suppose that $(S, \mathcal{S}, \mu)$ is a measure space and f is an integrable function, then we denote $\mu(f)$ by $\int_S f(x) d\mu(x)$. Given $x \in S$, δ_x denotes the Dirac measure on $(S, \mathcal{S})$. In particular, given a random sample $X_1, \dots, X_n$, P_n denotes the empirical measure and $P_n f$ denotes $n^{-1} \sum_{j=1}^n f(X_j)$.

Given measurable spaces $(S_1, \mathcal{S}_1), \dots, (S_m, \mathcal{S}_m)$, $\left(\prod_{j=1}^m S_j, \prod_{j=1}^m \mathcal{S}_j\right)$ denotes the product space endowed of the product σ-field. If μ_j is a measure on $(S_j, \mathcal{S}_j)$, for each $1 \le j \le m$, then $\mu_1 \otimes \cdots \otimes \mu_m$ denotes the product measure on $\left(\prod_{j=1}^m S_j, \prod_{j=1}^m \mathcal{S}_j\right)$.

2. Asymptotic Results for the Tests

In this section, we present several results regarding the asymptotics of the considered tests. We obtain results similar to the ones in Arcones and Wang [3]. We assume the following condition:

(B) $\sup_{1 \le j \le n} | A_n^{-1/2} x_{n,j} | \to 0,$ as $n \to \infty.$

Next theorem shows that the first test is omnibus.

Theorem 2.1. *Assume that regressors design satisfies Conditions* (A) *and* (B). *Suppose that the errors* $\{\varepsilon_j\}_{j=1}^{\infty}$ *form a sequence of i.i.d. r.v.'s with mean zero, variance* $\sigma^2 > 0$ *and continuous c.d.f. F. Then, for each* $m \ge 1,$

$$D_{n,m} \xrightarrow{P} \sup_{t \in \mathbb{R}} \left| P_F \left\{ \sigma^{-1} m^{-1/2} \sum_{j=1}^{m} \varepsilon_j \le t \right\} - \Phi(t) \right|, \quad as\ n \to \infty.$$

Theorem 2.1 implies that $b_{n,m,\alpha} \to 0,$ as $n \to \infty.$ Theorem 2.1 also implies that if the c.d.f. F of the sequence of errors does not have a normal distribution, then for each $1 > \alpha > 0,$ $P_F\{D_{n,m} \ge b_{n,m,\alpha}\} \to 1,$ as $n \to \infty.$

Condition (B) is a sort of necessary condition in Theorem 2.1. Suppose that $x'_{n,1} A_n^{-1} x_{n,1} \to c > 0$ and $\sup_{2 \le j \le n} | A_n^{-1/2} x_{n,j} | \to 0,$ then the proof of Theorem 2.1 gives that the residuals are approximately $\varepsilon_j - c\varepsilon_1,$ and

$$D_{n,m} \xrightarrow{P} \sup_{t \in \mathbb{R}} \left| P \left\{ \sigma^{-1} m^{-1/2} \sum_{j=1}^{m} (\varepsilon_j - c\varepsilon_1) \le t \right\} - \Phi(t) \right|, \quad \text{as } n \to \infty.$$

Next, we consider the asymptotic null distribution of the first test. We assume the following condition:

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

Suppose that for each $1 \le i \le n$, $x'_{n,i} = (1, x_{n,i,2}, \ldots, x_{n,i,p})$. Then

$$n^{-1} \sum_{j,k=1}^{n} x'_{n,j} A_n^{-1} x_{n,k} = 1.$$

Observe that

$$n^{-1} \sum_{j,k=1}^{n} x'_{n,j} A_n^{-1} x_{n,k}$$

is the (1, 1)-element of the matrix

$$n^{-1} \sum_{j,k=1}^{n} x_{n,j} x'_{n,j} A_n^{-1} x_{n,k} x'_{n,k} = n^{-1} A_n,$$

which is 1. Hence, if $x'_i = (1, x_{i,2}, \ldots, x_{i,p})$ for each $1 \le i \le n$, then condition (C) holds.

Theorem 2.2. *Assume that the regressors design satisfies* (A), (B) *and* (C). *Suppose that* $\{\varepsilon_j\}_{j=1}^{\infty}$ *is a sequence of i.i.d. r.v.'s from a normal distribution with mean zero and variance* $\sigma^2 > 0$. *Then, for each* $m \ge 2$,

$$n^{1/2} D_{n,m} - \sup_{t \in \mathbb{R}} \left| n^{-1/2} \sum_{j=1}^{n} \left(g(\sigma^{-1}\varepsilon_j, t) - E[g(\sigma^{-1}\varepsilon_j, t)]\right) \right| \xrightarrow{P} 0, \qquad (2.1)$$

where

$$g(\varepsilon, t) = m\Phi((m-1)^{-1/2}(m^{1/2}t - \varepsilon)) + (m^{1/2}\varepsilon + 2^{-1}(\varepsilon^2 - 1)t)\phi(t),$$

and ϕ *is the pdf of a standard normal distribution.*

Consequently

$$n^{1/2} D_{n,m} \xrightarrow{d} \sup_{t \in \mathbb{R}} | U(t) |,$$

where $\{U(t) : t \in \mathbb{R}\}$ *is a Gaussian process with mean zero and covariance given by*

$$E[U(s)U(t)] = \mathrm{Cov}(g(Z_1, s), g(Z_1, t)), \quad s, t \in \mathbb{R},$$

where Z_1 *is a standard normal r.v.*

Without assuming conditions (B) and (C), we may have that $n^{1/2}D_{n,m}$ converges in distribution. But, the limit could be different from the one in the previous theorem. The proof of Theorem 2.2 gives that if (A) and (B) hold, but (C) does not, then

$$n^{1/2}D_{n,m} - \sup_{t\in\mathbb{R}}\Bigg| n^{-1/2}\sum_{j=1}^{n}(\tilde{g}(\sigma^{-1}\varepsilon_j, t) - E[\tilde{g}(\sigma^{-1}\varepsilon_j, t)])$$

$$+ m^{1/2}n^{-1/2}\sum_{j,k=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}\varepsilon_k\phi(t)\Bigg| \xrightarrow{P} 0, \tag{2.2}$$

where

$$\tilde{g}(\varepsilon, t) = m\Phi((m-1)^{-1/2}(m^{1/2}t - \varepsilon)) + 2^{-1}(\varepsilon^2 - 1)t\phi(t).$$

We have that

$$m^{1/2}n^{-1/2}\sum\nolimits_{j,k=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}\varepsilon_k$$

and

$$m^{1/2}n^{-1/2}\sum\nolimits_{k=1}^{n}\varepsilon_k$$

are asymptotically equivalent if only if (C) holds. Notice that by (5.8)

$$m^{1/2}n^{-1/2}\sum_{j,k=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}\varepsilon_k - m^{1/2}n^{-1/2}\sum_{k=1}^{n}\varepsilon_k$$

has a normal distribution with mean zero and variance

$$m\sigma^2 - m\sigma^2 n^{-1}\sum_{k=1}^{n}\sum_{j=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}.$$

For the second proposed test, we have asymptotics similar to those of the first test:

Theorem 2.3. *Assume that the regressors design satisfies* (A) *and* (B). *Suppose that the errors* $\{\varepsilon_j\}_{j=1}^{\infty}$ *form a sequence of i.i.d. r.v.'s with mean zero, variance* $\sigma^2 > 0$ *and continuous c.d.f. F. Then, for each* $m \geq 1$,

$$\widetilde{D}_{n,m} \xrightarrow{P} \sup_{t \in \mathbb{R}} \left| P_F\left\{ m^{-1/2} \sum_{j=1}^{m} \varepsilon_j \leq t \right\} - P_F\{\varepsilon_1 \leq t\} \right|,$$

as $n \to \infty$.

As before Theorem 2.3 implies that $c_{n,m,\alpha} \to 0$, as $n \to \infty$. It also implies that if the distribution of the sequence of errors is not normal, then for each $1 > \alpha > 0$, $P_F\{D_{n,m} \geq c_{n,m,\alpha}\} \to 1$, as $n \to \infty$.

Theorem 2.4. *Assume that the regressors design satisfies* (A), (B) *and* (C). *Suppose that* $\{\varepsilon_j\}_{j=1}^{\infty}$ *is a sequence of i.i.d. r.v.'s from a normal distribution with mean zero and variance* $\sigma^2 > 0$. *Then, for each* $m \geq 1$,

$$n^{1/2}\widetilde{D}_{n,m} - \sup_{t \in \mathbb{R}} \left| n^{-1/2} \sum_{j=1}^{n} (h(\sigma^{-1}\varepsilon_j, t) - E[h(\sigma^{-1}\varepsilon_j, t)]) \right| \xrightarrow{P} 0,$$

where

$$h(\varepsilon, t) = m\Phi((m-1)^{-1/2}(m^{1/2}t - \varepsilon)) - I(\varepsilon \leq t) + (m^{1/2} - 1)\varepsilon\phi(t).$$

Consequently, $n^{1/2}\widetilde{D}_{n,m} \xrightarrow{d} \sup_{s \in \mathbb{R}} |V(s)|$, *where* $\{V(s) : s \in \mathbb{R}\}$ *is a mean zero Gaussian process with covariance given by*

$$E[V(s)V(t)] = \mathrm{Cov}(h(Z_1, s), h(Z_1, t)), \quad s, t \in \mathbb{R}.$$

3. Simulations

In this section, we present simulations of the presented normality tests. Besides the presented tests, we consider the test in (L) Lilliefors [16], in (SW) Shapiro and Wilk [24], in (BJ) Bera and Jarque [5, 6] and in (BHEP) Baringhaus and Henze [4] and Epps and Pulley [12]. The SW and the BJ normality tests are the most often used in the literature in

Statistics and Econometrics, respectively. We include the L test, because the presented tests appear as a modification of this test. The BHEP test is a common test.

We use different regressors designs to see how the power varies with the cutpoints. We only consider regressors design for the simple linear regression model. We assume that $x_{n,j} = (1, x_{n,j,2})'$, where $|x_{n,j,2}| \leq 1$. Our regressors design follows the discussion on Sections 1.8 and 1.9 in Draper and Smith [10].

We only consider the case $m = 2$. By the results in Arcones and Wang [3], there is no gain in using higher order m's.

First, we use $x_{n,j} = (1, (n-1)^{-1}(2j-1-n))'$, for $1 \leq j \leq n$. This regressors design is used when we would like to check whether there exists a linear relation between the variables.

The following tables show the values of $na_{n,2,\alpha}$ and $nb_{n,2,\alpha}$ for some values of n. The tables were obtained by doing 10000 simulations from a standard normal distribution.

$nb_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 8$	0.1431317074	0.1621240463	0.2086544447
$n = 12$	0.1065845579	0.1237184782	0.1640259523
$n = 16$	0.08653575281	0.09999083749	0.13250019197
$n = 20$	0.07497237684	0.08700437101	0.11447884146
$n = 24$	0.06597126745	0.07531215271	0.09620995402

$nc_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 8$	0.2678571429	0.2678571429	0.2857142857
$n = 12$	0.2045454545	0.2272727273	0.2575757576
$n = 16$	0.1791666667	0.1958333333	0.2250000000
$n = 20$	0.1578947368	0.1710526316	0.2000000000
$n = 24$	0.1449275362	0.1576086957	0.1811594203

The following table shows the power when $\alpha = 0.05$ of the mentioned tests of normality.

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0892	0.0932	0.1080	0.1046	0.1266	0.0399
$n = 12$	0.1140	0.1404	0.1727	0.1496	0.1680	0.0293
$n = 16$	0.1478	0.1108	0.2414	0.1846	0.2265	0.0302
$n = 20$	0.1653	0.2209	0.2648	0.2262	0.2510	0.0306
$n = 24$	0.2192	0.2864	0.3282	0.2701	0.3136	0.0258

Alternative: double exponential distribution

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.2601	0.2974	0.3295	0.3206	0.3619	0.0953
$n = 12$	0.4534	0.5688	0.6086	0.5704	0.6032	0.1263
$n = 16$	0.6563	0.5920	0.7525	0.7286	0.7544	0.2007
$n = 20$	0.7507	0.8316	0.8377	0.8287	0.8447	0.2750
$n = 24$	0.8465	0.8947	0.8956	0.8842	0.9103	0.3111

Alternative: Cauchy distribution

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0973	0.1064	0.1228	0.1156	0.1385	0.0460
$n = 12$	0.1322	0.1806	0.2141	0.1917	0.1988	0.0269
$n = 16$	0.1857	0.1616	0.3043	0.2522	0.2777	0.0371
$n = 20$	0.2113	0.3023	0.3621	0.2936	0.3239	0.0383
$n = 24$	0.2665	0.3731	0.4259	0.3450	0.3886	0.0374

Alternative: Student's t-distribution with three degrees of freedom

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0635	0.0599	0.0621	0.0646	0.0780	0.0427
$n = 12$	0.0612	0.0738	0.0829	0.0727	0.0823	0.0357
$n = 16$	0.0676	0.0612	0.1086	0.0852	0.0976	0.0439
$n = 20$	0.0673	0.0887	0.1096	0.0815	0.1008	0.0384
$n = 24$	0.0814	0.1123	0.1339	0.0937	0.1140	0.0368

Alternative: Student's t-distribution with ten degrees of freedom

The following tables show the values of $nb_{n,2,\alpha}$ and $nc_{n,2,\alpha}$ for the regressors design $x_{n,j} = (1, (-1)^j)'$, for $1 \leq j \leq n$. This regressors design is used when we would like to estimate β with the best accuracy possible (assuming that a linear relation between the variables holds).

$nb_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 8$	0.1386297968	0.1568467964	0.1936227425
$n = 12$	0.1050275466	0.1220525521	0.1589839848
$n = 16$	0.08652461426	0.10099724885	0.13081039783
$n = 20$	0.07504055901	0.08644431446	0.11339672468
$n = 24$	0.06596753173	0.07637168722	0.09794838675

$nc_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 8$	0.2678571429	0.2857142857	0.2857142857
$n = 12$	0.2045454545	0.2272727273	0.2500000000
$n = 16$	0.1791666667	0.1958333333	0.2250000000
$n = 20$	0.1578947368	0.1710526316	0.200000000
$n = 24$	0.1449275362	0.1576086957	0.1811594203

The following table shows the power when $\alpha = 0.05$ of the mentioned tests of normality for the second regressors design.

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0752	0.0788	0.1143	0.0871	0.1221	0.0385
$n = 12$	0.1162	0.1383	0.1670	0.1497	0.1642	0.0257
$n = 16$	0.1454	0.0983	0.2383	0.1840	0.2166	0.0301
$n = 20$	0.1765	0.2233	0.2653	0.2332	0.2554	0.0312
$n = 24$	0.2017	0.2596	0.3204	0.2570	0.2948	0.0264
	Alternative: double exponential distribution					

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.2017	0.3426	0.4287	0.3576	0.4323	0.0417
$n = 12$	0.4914	0.5822	0.6116	0.5837	0.6212	0.1207
$n = 16$	0.6547	0.5780	0.7578	0.7395	0.7484	0.2498
$n = 20$	0.7629	0.8209	0.8356	0.8264	0.8481	0.3417
$n = 24$	0.8467	0.8934	0.8939	0.8902	0.8988	0.4038
	Alternative: Cauchy distribution					

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0822	0.0854	0.1356	0.1068	0.1416	0.0412
$n = 12$	0.1373	0.1850	0.2265	0.1867	0.2107	0.0297
$n = 16$	0.1761	0.1543	0.3075	0.2534	0.2674	0.0350
$n = 20$	0.2169	0.3006	0.3470	0.2999	0.3192	0.0390
$n = 24$	0.2525	0.3581	0.4211	0.3475	0.3772	0.1080
	Alternative: Student's t-distribution with three degrees of freedom					

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0543	0.0597	0.0664	0.0611	0.0864	0.0503
$n = 12$	0.0602	0.0714	0.0823	0.0731	0.0845	0.0386
$n = 16$	0.0698	0.0570	0.1108	0.0791	0.0841	0.0392
$n = 20$	0.0710	0.0900	0.1023	0.0900	0.0991	0.0386
$n = 24$	0.0741	0.1009	0.1220	0.0906	0.1064	0.0387
	Alternative: Student's t-distribution with ten degrees of freedom					

The following tables show the values of $nb_{n,2,\alpha}$ and $nc_{n,2,\alpha}$ for the regressors design $x_{n,j} = (1, ((4/3)n - 1)^{-1}(2j - 1 - n(4/3)))'$, for $1 \leq j \leq (3/4)n$; $x_{n,j} = (1, (-1)^j)'$, for $(3/4)n + 1 \leq j \leq n$.

$nb_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 8$	0.1391902200	0.1581479131	0.2032428016
$n = 12$	0.1067211036	0.1242696687	0.1626261962
$n = 16$	0.08794126414	0.10146586222	0.13323187698
$n = 20$	0.07506775110	0.08789944333	0.11266438443
$n = 24$	0.06637517220	0.07669306853	0.10063409877

$nc_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 8$	0.2678571429	0.2678571429	0.2857142857
$n = 12$	0.2121212121	0.2272727273	0.2575757576
$n = 16$	0.1791666667	0.1958333333	0.2208333333
$n = 20$	0.1578947368	0.1710526316	0.1973684211
$n = 24$	0.1449275362	0.1576086957	0.1811594203

The following table shows the power when $\alpha = 0.05$ of the mentioned tests of normality for the third regressors design.

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0766	0.0796	0.1081	0.0981	0.1281	0.0378
$n = 12$	0.1125	0.1358	0.1779	0.1436	0.1613	0.0302
$n = 16$	0.1453	0.1067	0.2363	0.1868	0.2190	0.0364
$n = 20$	0.1695	0.2412	0.2740	0.2200	0.2535	0.0287
$n = 24$	0.2131	0.2680	0.3237	0.2609	0.2973	0.0277
	Alternative: double exponential distribution					

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.2341	0.2473	0.3678	0.2579	0.3239	0.0816
$n = 12$	0.4809	0.5670	0.6118	0.5670	0.6009	0.1109
$n = 16$	0.6418	0.5920	0.7479	0.7227	0.7478	0.1833
$n = 20$	0.7619	0.8278	0.8329	0.8279	0.8397	0.2524
$n = 24$	0.8416	0.8920	0.8982	0.8901	0.9035	0.3194
	Alternative: Cauchy distribution					

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0836	0.0901	0.1182	0.1029	0.1347	0.0433
$n = 12$	0.1389	0.1839	0.2253	0.1832	0.2017	0.0316
$n = 16$	0.1768	0.1498	0.3022	0.2450	0.2764	0.0415
$n = 20$	0.2154	0.3027	0.3590	0.2973	0.3172	0.0408
$n = 24$	0.2636	0.3580	0.4229	0.3513	0.3772	0.0404
	Alternative: Student's t-distribution with three degrees of freedom					

	L	SW	BJ	BHEP	$D_{n,2}$	$\widetilde{D}_{n,2}$
$n = 8$	0.0518	0.0563	0.0666	0.0584	0.0933	0.0457
$n = 12$	0.0634	0.0698	0.0875	0.0661	0.0821	0.0382
$n = 16$	0.0657	0.1498	0.1007	0.0809	0.0859	0.0397
$n = 20$	0.0672	0.0910	0.1148	0.0924	0.0954	0.0389
$n = 24$	0.0700	0.0996	0.1333	0.0938	0.0994	0.0359
	Alternative: Student's t-distribution with ten degrees of freedom					

Previous simulations show that the test based on $D_{n,2}$ is competitive with the other tests. However, the test based on $\widetilde{D}_{n,2}$ is a bad test. It seems that the BJ test is the best test overall. The test based on $D_{n,2}$ is the second best. The ranking of tests does not change with the regressors design. However, the power does. The power of the tests is slightly smaller for the second regressor design than for the first one. The third regressor design is a combination of the first and the second ones. Not surprisingly, the power of the tests for the third regressor design is between that of the first two regressor designs.

4. Several Results on Limit Theorems for *U*-processes

In order to obtain the asymptotic null distribution of the test statistics $D_{n,m}$ and $\widetilde{D}_{n,m}$ we present several results on the central limit theorem for U-processes over a sequence of independent (not necessarily identically distributed) r.v.'s and over kernels varying from occurrence to occurrence. General references on U-processes are Arcones and Giné [2] and de la Peña and Giné [7].

Given r.v.'s $X_1, \ldots, X_n$ 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

$$U_{n,m}(h) := \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} h(X_{i_1}, \ldots, X_{i_m}).$$

We extend this definition by allowing the function h to vary from occurrence to occurrence. We present the following theorem:

Theorem 4.1. *Let* $\{X_{n,1}, \ldots, X_{n,n} : n \geq 1\}$ *be a triangular array of row-wise independent r.v.'s with values in a measurable space* $(S, \mathcal{S})$. *Let* $h_{n, i_1, \ldots, i_m} : (S^m, \mathcal{S}^m) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *be a measurable function for each* $(i_1, \ldots, i_m) \in I_m^n$. *Suppose that the following conditions are satisfied*:

(i) *There exists a sequence* $\{\delta_n\}_{n=1}^{\infty}$ *of positive numbers converging to zero such that for each* $n \geq 1$ *and each* $(i_1, \ldots, i_m) \in I_m^n$,

$$| h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}) | \leq n\delta_n \quad a.s.$$

(ii) $\mathrm{Var}(U_n) \to \sigma^2 < \infty$, *where*

$$U_n := \sum_{(i_1, \ldots, i_m) \in I_m^n} (h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}) - E[h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m})]).$$

Then, $U_n \xrightarrow{d} N(0, \sigma^2)$ *as* $n \to \infty$.

Next, we present a limit theorem for U-processes over a triangular array of row-wise independent r.v.'s satisfying a VC-like condition.

Let (T, d) be a metric space. Given $K \subset T$, the *packing number* of K is defined as

$$D(u, K) := \max\{m \geq 1 : \text{there are } t_1, \ldots, t_m \in K,$$
$$\text{such that } d(t_i, t_j) > u, \text{ for } t_i \neq t_j\}, \quad u > 0.$$

The *covering number* of K is defined as

$$N(u, K) := \min\{m \geq 1 : \text{there are } t_1, \ldots, t_m \in T,$$
$$\text{such that } K \subset \textstyle\bigcup_{\mathrm{j}=1}^{\mathrm{m}} \overline{B}(t_j, u)\}, \quad u > 0,$$

where $\overline{B}(t_j, u) = \{t \in T : d(t, t_j) \leq u\}$. It is easy to see that for each $u > 0$,

$$N(u, K) \leq D(u, K) \leq N(2^{-1}u, K). \tag{4.1}$$

We will use the previous definitions, when $T = \mathbb{R}^n$ and d is the Euclidean distance. By Theorem II.3.1 in Marcus and Pisier [18] (see also Pollard [21, Theorem 3.5]),

$$E\left[\sup_{v\in K}\left|\sum_{j=1}^{n} R_j(v^{(j)} - v_0^{(j)})\right|\right] \leq 9\int_0^D (\log D(u, K))^{1/2} du, \tag{4.2}$$

for any $K \subset \mathbb{R}^n$ and any $v_0 \in K$, where $\{R_j\}_{j=1}^n$ is a sequence of Rademacher r.v.'s, $v' = (v^{(1)}, \ldots, v^{(n)})$ and $D := \sup_{v\in K}|v - v_0|$.

We consider triangular arrays of functions satisfying the following condition:

Definition 4.1. Let $\{X_{n,1}, \ldots, X_{n,n} : n \geq 1\}$ be a triangular array of row-wise independent r.v.'s with values in a measurable space $(S, \mathcal{S})$. Suppose that for each $n \geq m$, each $(i_1, \ldots, i_m) \in I_m^n$ and each $t \in T$, we have a measurable function $h_{n,i_1,\ldots,i_m}(\cdot, t) : (S^m, \mathcal{S}^m) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where T is an index set. We say that the *triangular array* of U-processes

$$\{h_{n,i_1,\ldots,i_m}(X_{n,i_1}, \ldots, X_{n,i_m}, t) : 1 \leq n, (i_1, \ldots, i_m) \in I_m^n, t \in T\}$$

is manageable with respect to the envelope r.v.'s

$$\{H_{n,i_1,\ldots,i_m}(X_{n,i_1}, \ldots, X_{n,i_m}) : 1 \leq n, (i_1, \ldots, i_m) \in I_m^n\},$$

where $H_{n,i_1,\ldots,i_m} : (S^m, \mathcal{S}^m) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a measurable function if:

(i)

$$\sup_{t\in T}|h_{n,i_1,\ldots,i_m}(X_{n,i_1}, \ldots, X_{n,i_m}, t)| \leq H_{n,i_1,\ldots,i_m}(X_{n,i_1}, \ldots, X_{n,i_m}) \text{ a.s.}$$

(ii) The function $M(u)$, defined on $(0, 2^{1/2}]$ by

$M(u)$

$$:= \sup_{n \geq 1, x_1, \ldots, x_n \in S} D\left(u\left(\sum_{(t_1, \ldots, i_m) \in I_m^n} (H_{n, i_1, \ldots, i_m}(x_{i_1}, \ldots, x_{i_m}))^2\right)^{1/2}, \mathcal{G}_n(x_1, \ldots, x_n)\right),$$

satisfies that

$$\int_0^{2^{1/2}} (\log M(u))^{1/2} du < \infty,$$

where

$$\mathcal{G}_n(x_1, \ldots, x_n) := \{(h_{n, i_1, \ldots, i_m}(x_{i_1}, \ldots, x_{i_m}, t))_{(i_1, \ldots, i_m) \in I_m^n}$$

$$\in \mathbb{R}^{n!/(n-m)!} : t \in T\}.$$

Notice that in condition (ii) in Definition 4.1 $D(t, \mathcal{G}_n(x_1, \ldots, x_n))$, $t > 0$, denotes the packing number of $\mathcal{G}_n(x_1, \ldots, x_n)$ when the Euclidean distance of $\mathbb{R}^{n!/(n-m)!}$ is used.

The last definition is an extension to the U-processes case of Definition 7.9 in Pollard [21]. Definition 7.9 in Pollard [21] generalizes to the triangular array case of the concept of VC subgraph classes, which has been studied by several authors (see for example Dudley [11]).

Next, we present an analogous of (4.2) for our situation. Notice that (4.2) uses the entropy with respect to the L_2 of corresponding Rademacher process. In our case, by the Cauchy-Schwartz inequality, for each $s, t \in T$ and each $x_1, \ldots, x_n \in S$

$$E\left[\left| n^{1/2} \sum_{(i_1, \ldots, i_m) \in I_m^n} R_{i_1}(h_{n, i_1, \ldots, i_m}(x_{i_1}, \ldots, x_{i_m}, s) - h_{n, i_1, \ldots, i_m}(x_{i_1}, \ldots, x_{i_m}, t))\right|^2\right]$$

$$= n \sum_{i_1 = 1}^{n} \left(\sum_{(i_2, \ldots, i_m) : (i_1, \ldots, i_m) \in I_m^n} (h_{n, i_1, \ldots, i_m}(x_{i_1}, \ldots, x_{i_m}, s) - h_{n, i_1, \ldots, i_m}(x_{i_1}, \ldots, x_{i_m}, t))\right)^2$$

$$\leq n \sum_{i_1=1}^{n} \sum_{(i_2, \dots, i_m):(i_1, \dots, i_m) \in I_m^n} (h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, s)$$

$$- h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t))^2 \frac{n!}{n(n-m)!}$$

$$= \frac{n!}{(n-m)!} \sum_{(i_1, \dots, i_m) \in I_m^n} (h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, s) - h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t))^2. \quad (4.3)$$

Hence, for a sequence of Rademacher r.v.'s $\{R_j\}_{j=1}^{\infty}$, for each $x_1, \dots, x_n \in S$ and each $t_0 \in T$,

$$n^{1/2} E\left[\sup_{t \in T} \left| \sum_{(i_1, \dots, i_m) \in I_m^n} R_{i_1}(h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t) - h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t_0)) \right|\right]$$

$$\leq 9 \int_0^{(n!/(n-m)!)^{1/2} D_n} (\log D(((n-m)!/n!)^{1/2} u, \mathcal{G}_n(x_1, \dots, x_n))^{1/2} du$$

$$\leq 9 \left(\frac{n!}{(n-m)!} \right)^{1/2} \int_0^{D_n} (\log D(u, \mathcal{G}_n(x_1, \dots, x_n)))^{1/2} du, \quad (4.4)$$

where

$$D_n^2 = \sup_{t \in T} \sum_{(i_1, \dots, i_m) \in I_m^n} (h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t) - h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t_0))^2.$$

Theorem 4.2. *Under the notation on Definition* 4.1, *suppose that for each* $t \in T$, *there exists a measurable function* $f_{n, i_1, \dots, i_m}(\cdot, t) : (S^m, \mathcal{S}^m) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *such that*:

(i) *For each* $x_1, \dots, x_n \in S$ *and each* $(i_1, \dots, i_m) \in I_m^n$,

$$h_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t) = I(f_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t) \geq 0).$$

(ii) *For each* $x_1, \dots, x_n \in S$,

$$\{(f_{n, i_1, \dots, i_m}(x_{i_1}, \dots, x_{i_m}, t))_{(i_1, \dots, i_m) \in I_m^n} \in \mathbb{R}^{n!/(n-m)!} : t \in T\}$$

lies in a subspace of dimension d of $\mathbb{R}^{n!/(n-m)!}$.

Then, the triangular array of U-processes

$$\{h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}, t) : 1 \leq n, (i_1, \ldots, i_m) \in I_m^n, t \in T\}$$

is manageable with respect to the envelope r.v.'s

$$\{H_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}) : 1 \leq n, (i_1, \ldots, i_m) \in I_m^n\},$$

where $H_{n, i_1, \ldots, i_m} \equiv 1.$

The previous theorem follows from Lemma 4.4 and Corollary 4.10 in Pollard [21].

Theorem 4.3. *With the above notation, suppose that*:

(i) *The conditions in Definition* 4.1 *hold.*

(ii) *There exists a finite constant a such that for each* $(i_1, \ldots, i_m) \in I_m^n$,

$$H_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}) \leq a \quad a.s.$$

(iii) *For each* $\delta > 0$, *there are a positive number* n_0 *and* $\pi : T \to T$ *such that* $\# \pi(T) < \infty$, $\pi(\pi(t)) = \pi(t)$, *for each* $t \in T$, *and for each* $n \geq n_0$ *and each* $t \in T$,

$$\frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} E[(h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}, t) - h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}, \pi(t)))^2] \leq \delta^2.$$

(iv) *For each permutation* σ *of* $\{1, \ldots, m\}$, *each* $(i_1, \ldots, i_m) \in I_m^n$, *each* $x_1, \ldots, x_m \in S$, *and each* $t \in T$,

$$h_{n, i_1, \ldots, i_m}(x_1, \ldots, x_m, t) = h_{n, i_{\sigma(1)}, \ldots, i_{\sigma(m)}}(x_{\sigma(1)}, \ldots, x_{\sigma(m)}, t).$$

(v) *For each* $s, t \in T$,

$$\lim_{n \to \infty} n \operatorname{Cov}(U_n(s), U_n(t))$$

exists, where

$$U_n(t) := \frac{(n-m)!}{n!} \sum_{(i_1, \dots, i_m) \in I_m^n} h_{n, i_1, \dots, i_m}(X_{n, i_1}, \dots, X_{n, i_m}).$$

Then

$$\{n^{1/2}(U_n(t) - E[U_n(t)]) : t \in T\} \xrightarrow{w} \{Z(t) : t \in T\},$$

where $\{Z(t) : t \in T\}$ *is a Gaussian process with mean zero and covariance,*

$$\mathrm{Cov}(Z(s), Z(t)) = \lim_{n \to \infty} n\mathrm{Cov}(U_n(s), U_n(t)), \quad s, t \in T.$$

Theorem 4.4. *Let* $\{X_j\}_{j=1}^{\infty}$ *be a sequence of i.i.d. r.v.'s. Let* p *be a positive integer. Let* $h : \mathbb{R}^m \to \mathbb{R}$ *be a measurable function. Let* $f_{n, k, i_1, \dots, i_m} : \mathbb{R}^m \to \mathbb{R}$ *be a measurable function for each* $(i_1, \dots, i_m) \in I_m^n$ *and each integer* $1 \le k \le d$. *Let*

$$\begin{aligned} &U_n(b, t) \\ &= \frac{(n-m)!}{n!} \sum\nolimits_{(i_1, \dots, i_m) \in I_m^n} \Bigg(I\Bigg(h(X_{i_1}, \dots, X_{i_m}) \\ &\quad + \sum\nolimits_{k=1}^{p} f_{n, k, i_1, \dots, i_m}(X_{i_1}, \dots, X_{i_m}) b_k \le t \Bigg) - P\Bigg\{ h(X_{i_1}, \dots, X_{i_m}) \\ &\quad + \sum\nolimits_{k=1}^{p} f_{n, k, i_1, \dots, i_m}(X_{i_1}, \dots, X_{i_m}) b_k \le t \Bigg\} \Bigg), \end{aligned}$$

where $b = (b_1, \dots, b_p)' \in \mathbb{R}^p$, $t \in \mathbb{R}$. *Suppose that*:

(i) *The function* $P\{h(X_1, \dots, X_m) \le t\}$, $t \in \mathbb{R}$, *is continuous.*

(ii) *There exists a sequence of positive numbers* $\{\delta_n\}_{n=1}^{\infty}$ *converging to zero such that for each* $n \ge 1$ *and each* $(i_1, \dots, i_m) \in I_m^n$

$$| f_{n, k, i_1, \dots, i_m}(X_1, \dots, X_m) | \le \delta_n, \quad a.s.$$

(iii) *For each permutation* σ *of* $\{1, \ldots, m\}$ *and each* $x_1, \ldots, x_m \in \mathbb{R}$,

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

(iv) *For each permutation* σ *of* $\{1, \ldots, m\}$, *each* $(i_1, \ldots, i_m) \in I_m^n$, *each* $x_1, \ldots, x_m \in S$, *each* $1 \leq k \leq d$, *and each* $t \in T$,

$$f_{n, k, i_1, \ldots, i_m}(x_1, \ldots, x_m, t) = f_{n, k, i_{\sigma(1)}, \ldots, i_{\sigma(m)}}(x_{\sigma(1)}, \ldots, x_{\sigma(m)}, t).$$

Then, for each $0 < M < \infty$,

$$\{n^{1/2}(U_n(b, t) - E[U_n(b, t)]) : b \in \mathbb{R}^d, |\, b\,| \leq M, t \in \mathbb{R}\}$$

converges weakly to a mean zero Gaussian process $\{Z(b, t) : |\, b\,| \leq M, t \in \mathbb{R}\}$ *with covariance given by*

$$\begin{aligned}\mathrm{Cov}(Z(b, s), Z(d, t)) &= \lim_{n \to \infty} n \mathrm{Cov}(U_n(b, s), U_n(b, t)) \\ &= \mathrm{Cov}(\phi(X_1, s), \phi(X_1, t)),\end{aligned}$$

for each $s, t \in \mathbb{R}$, $b, d \in \mathbb{R}^p$, $|\, b\,|, |\, d\,| \leq M$, *where*

$$\phi(x, t) = \sum_{j=1}^{m} P\{h(X_1, \ldots, X_{j-1}, x, X_{j+1}, \ldots, X_m) \leq t\}.$$

5. Proofs

Proof of Theorem 2.1. By (1.3)

$$\begin{aligned}&\frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} I\left(\hat{\sigma}_n^{-1} m^{-1/2} \sum_{j=1}^{m} \hat{\varepsilon}_{i_j} \leq t\right) \\ &= \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} I\left(\sum_{j=1}^{m} \varepsilon_{i_j} \leq \sum_{j=1}^{m} (\hat{\beta}_n - \beta)' A_n^{1/2} A_n^{-1/2} x_{n, i_j} + t \hat{\sigma}_n m^{1/2}\right).\end{aligned} \tag{5.1}$$

First, we prove that for each $0 < M < \infty$,

$$\sup_{|b|\le M}\sup_{t\in\mathbb{R}}\left|\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(\sum_{j=1}^m \varepsilon_{i_j} \le \sum_{j=1}^m b' A_n^{-1/2} x_{n,i_j} + t\right) - H(t)\right| \overset{a.s.}{\to} 0, \tag{5.2}$$

where

$$H(t) = P\left\{\sum_{j=1}^m \varepsilon_j \le t\right\}, \quad t \in \mathbb{R}.$$

Since the c.d.f. of ε_1 is continuous, so is H. Hence, H is uniformly continuous and for each $\varepsilon > 0$ there exists $\delta > 0$ such that if $|s-t| \le \delta$, then $|H(s) - H(t)| < \varepsilon$. By condition (B), there exists a positive integer n_0 such that if $n \ge n_0$, then $\max_{1\le j\le n} mM|A_n^{-1/2}x_{n,j}| \le \delta$. If $n \ge n_0$, then we have

$$\sup_{|b|\le M}\sup_{t\in\mathbb{R}}\left(\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(\sum_{j=1}^m \varepsilon_{i_j} \le \sum_{j=1}^m b' A_n^{-1/2} x_{n,i_j} + t\right) - H(t)\right)$$

$$\le \sup_{t\in\mathbb{R}}\left(\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(\sum_{j=1}^m \varepsilon_{i_j} \le \delta + t\right) - H(t)\right)$$

$$\le \sup_{t\in\mathbb{R}}|H_n(t) - H(t)| + \sup_{t\in\mathbb{R}}|H(\delta+t) - H(t)| \le \sup_{t\in\mathbb{R}}|H_n(t) - H(t)| + \varepsilon$$

and

$$\inf_{|b|\le M}\inf_{t\in\mathbb{R}}\left(\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(\sum_{j=1}^m \varepsilon_{i_j} \le \sum_{j=1}^m b' A_n^{-1/2} x_{n,i_j} + t\right) - H(t)\right)$$

$$\ge \inf_{t\in\mathbb{R}}\left(\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(\sum_{j=1}^m \varepsilon_{i_j} \le -\delta + t\right) - H(t)\right)$$

$$\ge -\sup_{t\in\mathbb{R}}|H_n(t) - H(t)| - \sup_{t\in\mathbb{R}}|H(\delta+t) - H(t)| \ge -\sup_{t\in\mathbb{R}}|H_n(t) - H(t)| - \varepsilon,$$

where

$$H_n(t) := \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} I\left(\sum_{j=1}^{m} \varepsilon_{i_j} \leq t\right), \quad t \in \mathbb{R}.$$

Hence

$$\sup_{|b| \leq M} \sup_{t \in \mathbb{R}} \left| \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} I\left(\sum_{j=1}^{m} \varepsilon_{i_j} \leq \sum_{j=1}^{m} b' A_n^{-1/2} x_{n,j} + t\right) - H(t) \right|$$

$$\leq \sup_{t \in \mathbb{R}} | H_n(t) - H(t) | + \varepsilon.$$

By the strong law of the large numbers for U-processes,

$$\sup_{t \in \mathbb{R}} | H_n(t) - H(t) | \overset{\text{a.s.}}{\to} 0.$$

Hence, with probability one

$$\limsup_{n \to \infty} \sup_{|b| \leq M} \sup_{t \in \mathbb{R}} \left| \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} I\left(\sum_{j=1}^{m} \varepsilon_{i_j} \leq \sum_{j=1}^{m} b' A_n^{-1/2} x_{n,j} + t\right) - H(t) \right|$$

$\leq \varepsilon$ a.s.

Since $\varepsilon > 0$ is arbitrary, (5.2) follows.

By (1.2)

$$A_n^{1/2}(\hat{\beta}_n - \beta) = \sum_{j=1}^{n} \varepsilon_j A_n^{-1/2} x_{n,j},$$

$$E[A_n^{1/2}(\hat{\beta}_n - \beta)] = \sum_{j=1}^{n} E[\varepsilon_1] A_n^{-1/2} x_{n,j} = 0,$$

$$\text{Var}(A_n^{1/2}(\hat{\beta}_n - \beta)) = \sigma^2 I_p.$$

Hence

$$A_n^{1/2}(\hat{\beta}_n - \beta) = O_p(1). \tag{5.3}$$

Plugging $A_n^{1/2}(\hat{\beta}_n - \beta)$ into b and $m^{1/2}\hat{\sigma}_n t$ into t in (5.2) and using (5.1) and (5.3), we get that

$$\sup_{t\in\mathbb{R}}\left|\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n} I\left(m^{-1/2}\hat{\sigma}_n^{-1}\sum_{j=1}^{m}\hat{\varepsilon}_{i_j}\le t\right)-H(m^{1/2}\hat{\sigma}_n t)\right|\xrightarrow{P}0. \quad (5.4)$$

By (1.3)

$$\hat{\sigma}_n^2 = \frac{1}{n-p}\sum\nolimits_{j=1}^{n}(\varepsilon_j-(\hat{\beta}_n-\beta)' x_{n,j})^2$$

$$= \frac{1}{n-p}\sum\nolimits_{j=1}^{n}\varepsilon_j^2-\frac{2}{n-p}\sum\nolimits_{j=1}^{n}\varepsilon_j(\hat{\beta}_n-\beta)' x_{n,j}$$

$$+\frac{1}{n-p}\sum\nolimits_{j=1}^{n}((\beta-\hat{\beta}_n)' x_{n,j})^2.$$

By the law of the large numbers

$$\frac{1}{n-p}\sum_{j=1}^{n}\varepsilon_j^2\xrightarrow{\text{a.s.}}\operatorname{Var}(\varepsilon_1).$$

By (1.2)

$$\frac{1}{n-p}\sum_{j=1}^{n}\varepsilon_j(\hat{\beta}_n-\beta)' x_{n,j} = \frac{1}{n-p}(\hat{\beta}_n-\beta)' A_n(\hat{\beta}_n-\beta).$$

By the definition of A_n

$$\frac{1}{n-p}\sum_{j=1}^{n}((\beta-\hat{\beta}_n)' x_{n,j})^2 = \frac{1}{n-p}\sum_{j=1}^{n}(\beta-\hat{\beta}_n)' x_{n,j}x_{n,j}'(\beta-\hat{\beta}_n)$$

$$= \frac{1}{n-p}(\hat{\beta}_n-\beta)' A_n(\hat{\beta}_n-\beta).$$

Hence, by (5.3)

$$-\frac{2}{n-p}\sum_{j=1}^{n}\varepsilon_j(\hat{\beta}_n-\beta)' x_{n,j}+\frac{1}{n-p}\sum_{j=1}^{n}((\beta-\hat{\beta}_n)' x_{n,j})^2$$

$$= -\frac{1}{n-p}(\hat{\beta}_n-\beta)' A_n(\hat{\beta}_n-\beta) = o_P(1).$$

Therefore

$$\hat{\sigma}_n^2 \xrightarrow{P} E[\varepsilon_1^2]. \tag{5.5}$$

Using (5.5) and the uniformly continuity of H

$$\sup_{t \in \mathbb{R}} | H(m^{1/2}\hat{\sigma}_n t) - H(m^{1/2}\sigma t) | \xrightarrow{P} 0.$$

Therefore, the claim follows.

Lemma 5.1. *Let* $\{\varepsilon_j\}_{j=1}^{\infty}$ *be a sequence of i.i.d. r.v.'s from a normal distribution with mean zero and variance* $\sigma^2 > 0$. *Assume that conditions* (A) *and* (B) *hold. Let*

$U_n(b, T)$

$$:= n^{1/2} \frac{(n-m)!}{n!} \sum\nolimits_{(i_1, \ldots i_m) \in I_m^n} \left(I\left(\sum\nolimits_{j=1}^{m} \varepsilon_{i_j} \leq \sum\nolimits_{j=1}^{m} b' A_n^{-1/2} x_{n, i_j} + tm^{1/2} \right) \right.$$

$$\left. - \Phi\left(m^{-1/2} \sigma^{-1} \sum\nolimits_{j=1}^{m} b' A_n^{-1/2} x_{n, i_j} + \sigma^{-1} t \right) \right),$$

where $b \in \mathbb{R}^p$ *and* $t \in \mathbb{R}$. *Then, for each* $0 < M < \infty$,

$$\{U_n(b, t) : b \in \mathbb{R}^p, | b | \leq M, t \in \mathbb{R}\}$$

converges weakly to a mean zero Gaussian process $\{W(b, t) : b \in \mathbb{R}^p, | b | \leq M, t \in \mathbb{R}\}$ *with covariance given by*

$$E[W(b_1, s) W(b_2, t)] = Cov(m\Phi((m-1)^{-1/2} \sigma^{-1} (sm^{1/2} - \varepsilon_1)),$$

$$m\Phi((m-1)^{-1/2} \sigma^{-1} (tm^{1/2} - \varepsilon_1))),$$

for each $| b_1 |, | b_2 | \leq M$, $s, t \in \mathbb{R}$.

Proof. The lemma follows directly from Theorem 4.4.

Lemma 5.2. *Under the notation and conditions in the previous lemma, for each* $0 < M < \infty$,

$$\sup_{\substack{|b| \leq M \\ |s| \leq M}} \sup_{t \in \mathbb{R}} | U_n(b, t + n^{-1/2}st) - U_n(0, t) | \xrightarrow{P} 0.$$

Proof. By Theorem 3.7.2 in Dudley [11] for each $\tau > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to \infty} P\{ \sup_{d((b_1, s), (b_1, t)) \leq \delta} | U_n(b_1, s) - U_n(b_2, t) | \geq \tau \} = 0,$$

where

$$d^2((b_1, s), (b_2, t))$$

$$:= \operatorname{Var}(m\Phi((m-1)^{-1/2}\sigma^{-1}(sm^{1/2} - \varepsilon_1)) - m\Phi((m-1)^{-1/2}\sigma^{-1}(tm^{1/2} - \varepsilon_1)))$$

$$\leq E_{\varepsilon_1}[m\Phi((m-1)^{-1/2}\sigma^{-1}(sm^{1/2} - \varepsilon_1)) - m\Phi((m-1)^{-1/2}\sigma^{-1}(tm^{1/2} - \varepsilon_1))^2]$$

$$\leq m^2 E_{\varepsilon_1}\left[\left(E_{\varepsilon_1, \dots, \varepsilon_n}\left[I\left(\sum_{j=1}^m \varepsilon_j \leq sm^{1/2}\right) - I\left(\sum_{j=1}^m \varepsilon_j \leq tm^{1/2}\right)\right]\right)^2\right]$$

$$\leq m^2 E\left[\left(I\left(\sum_{j=1}^m \varepsilon_j \leq sm^{1/2}\right) - I\left(\sum_{j=1}^m \varepsilon_j \leq tm^{1/2}\right)\right)^2\right]$$

$$= m^2 | \Phi(\sigma^{-1}s) - \Phi(\sigma^{-1}t) |.$$

Hence

$$\sup_{\substack{|b| \leq M \\ |s| \leq M}} \sup_{t \in \mathbb{R}} d^2((b, t + n^{-1/2}st), (0, t))$$

$$\leq \sup_{|s| \leq M} \sup_{t \in \mathbb{R}} | \Phi(\sigma^{-1}(t + n^{-1/2}st)) - \Phi(\sigma^{-1}t) | \to 0,$$

as $n \to \infty$. Hence, the claim follows.

Lemma 5.3. *There exists a universal constant c such that for each* $|s| \leq 2^{-2}$ and each $|h| \leq 1$,

$$\sup_{t \in \mathbb{R}} | \Phi(t + st + h) - \Phi(t) - (st + h)\phi(t) | \leq c(s^2 + h^2). \tag{5.6}$$

Proof. By Taylor theorem, for each $s, t \in \mathbb{R}$,

$$|\Phi(s) - \Phi(t) - (s-t)\phi(t)| = \left| \int_t^s (s-u)\phi'(u)du \right|$$

$$= \left| \int_0^{s-t} (s-t-u)\phi'(u+t)du \right|$$

$$\leq 2^{-1}(s-t)^2 \sup\{|\phi'(t+u)| : |u| \leq |s-t|\}.$$

Hence, for each $t \in \mathbb{R}$, $|s| \leq 2^{-2}$ and $|h| \leq 1$,

$$|\Phi(t+st+h) - \Phi(t) - (st+h)\phi(t)|$$

$$\leq 2^{-1}(st+h)^2 \sup\{|\phi'(t+u)| : |u| \leq |st+h|\}.$$

If $|t| \leq 2$, then

$$2^{-1}(st+h)^2 \sup\{|\phi'(t+u)| : |u| \leq |st+h|\}$$

$$\leq (s^2t^2 + h^2)\sup_{x\in\mathbb{R}}|\phi'(x)| \leq (4s^2 + h^2)\sup_{x\geq 0}|\phi'(x)|.$$

If $|st| \leq |h|$, then

$$2^{-1}(st+h)^2 \sup\{|\phi'(t+u)| : |u| \leq |st+h|\}$$

$$\leq (s^2t^2 + h^2)\sup_{x\in\mathbb{R}}|\phi'(x)| \leq 2h^2 \sup_{x\geq 0}|\phi'(x)|.$$

If $|t| > 2$ and $|st| > |h|$, then

$$2^{-1}(st+h)^2 \leq s^2t^2 + h^2 \leq 2s^2t^2,$$

$$|st+h| \leq |st| + |h| \leq 2|st| \leq 2^{-1}|t|$$

and

$$|t+u| \geq |t| - |u| \geq |t| - |st+h| \geq 2^{-1}|t| \geq 1.$$

ϕ' is an odd function and it is negative and increasing in $(1, \infty)$. Hence

$$\sup\{|\phi'(t+u)| : |u| \leq |st+h|\} \leq |\phi'(2^{-1}|t|)|.$$

Therefore, if $|t| > 2$ and $|st| > |h|$, then

$$2^{-1}(st+h)^2 \sup\{|\phi'(t+u)| : |u| \le |st+h|\} \le 2s^2t^2|\phi'(2^{-1}|t|)|$$

$$\le 8s^2 \sup_{x\ge 0} x^2|\phi'(x)|.$$

The claim follows from the previous estimations.

Lemma 5.4. *For each* $0 < M < \infty$,

$$\sup_{\substack{|b|\le M\\ |s|\le M}} \sup_{t\in\mathbb{R}} \Bigg| n^{1/2} \frac{(n-m)!}{n!} \sum\nolimits_{(i_1,\dots,i_m)\in I_m^n}$$

$$\times \left(\Phi(m^{-1/2}\sigma^{-1}\sum\nolimits_{j=1}^m bA_n^{-1/2}x_{n,i_j} + n^{-1/2}\sigma^{-1}st + \sigma^{-1}t) - \Phi(\sigma^{-1}t)\right)$$

$$-\left(n^{-1/2}\sum\nolimits_{j=1}^n m^{1/2}\sigma^{-1}b'A_n^{-1/2}x_{n,j} + \sigma^{-1}st\right)\phi(\sigma^{-1}t)\Bigg| \to 0, \quad \textit{as } n\to\infty.$$

Proof. By Lemma 5.3, for n large enough,

$$\sup_{\substack{|b|\le M\\ |s|\le M}} \sup_{t\in\mathbb{R}} \Bigg| n^{1/2} \frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m)\in I_m^n}$$

$$\times\left(\Phi\left(m^{-1/2}\sigma^{-1}\sum_{j=1}^m b'A_n^{-1/2}x_{n,i_j} + n^{-1/2}\sigma^{-1}st + \sigma^{-1}t\right) - \Phi(\sigma^{-1}t)\right)$$

$$-\left(n^{1/2}\frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}\sum_{j=1}^m m^{-1/2}\sigma^{-1}b'A_n^{-1/2}x_{n,i_j} + \sigma^{-1}st\right)\phi(\sigma^{-1}t)\Bigg|$$

$$\le c\sup_{\substack{|b|\le M\\ |s|\le M}} n^{1/2}\frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}\left(\left|m^{-1/2}\sigma^{-1}\sum_{j=1}^m b'A_n^{-1/2}x_{n,i_j}\right|^2 + n^{-1}\sigma^{-2}s^2\right)$$

$$\le cn^{-1/2} + c\sup_{|b|\le M} n^{1/2}\frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}\sum_{j=1}^m (b'A_n^{-1/2}x_{n,i_j})^2$$

$$= cn^{-1/2} + c\sup_{|b|\le M} n^{-1/2}\sum_{j=1}^n b'A_n^{-1/2}x_{n,j}x_{n,j}'A_n^{-1/2}b = cn^{-1/2}.$$

We also have that

$$n^{1/2}\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n}\sum_{j=1}^{m} m^{-1/2}\sigma^{-1}b'A_n^{-1/2}x_{n,i_j}$$

$$= n^{-1/2}\sum_{j=1}^{n} m^{1/2}\sigma^{-1}b'A_n^{-1/2}x_{n,j}.$$

Proof of Theorem 2.2. By Lemmas 5.2 and 5.4, for each $0 < M < \infty$,

$$\sup_{\substack{|b|\le M\\ |s|\le M}}\sup_{t\in\mathbb{R}}\Bigg| n^{1/2}\frac{(n-m)!}{n!}\sum\nolimits_{(i_i,\ldots,i_m)\in I_m^n}$$

$$\times\left(I\left(\sum\nolimits_{j=1}^{m}\varepsilon_{i_j}\le\sum\nolimits_{j=1}^{m}bA_n^{-1/2}x_{n,i_j}+tm^{1/2}+n^{-1/2}stm^{1/2}\right)\right.$$

$$\left.-I\left(\sum\nolimits_{j=1}^{m}\varepsilon_{i_j}\le tm^{1/2}\right)\right)$$

$$-\left(n^{-1/2}\sum\nolimits_{j=1}^{n}m^{1/2}\sigma^{-1}b'A_n^{-1/2}x_{n,j}+\sigma^{-1}st\right)\phi(\sigma^{-1}t)\Bigg|\xrightarrow{P}0.$$

Plugging $A_n^{1/2}(\hat{\beta}_n-\beta)$ into b, $\sigma^{-1}n^{1/2}(\hat{\sigma}_n-\sigma)$ into s and σt into t in the previous expression, we get

$$\sup_{t\in\mathbb{R}}\Bigg| n^{1/2}\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n}\left(I\left(\hat{\sigma}_n^{-1}m^{-1/2}\sum_{j=1}^{m}\hat{\varepsilon}_{i_j}\le t\right)-I\left(\sigma^{-1}m^{-1/2}\sum_{j=1}^{m}\varepsilon_{i_j}\le t\right)\right)$$

$$-\left(n^{-1/2}\sum_{j=1}^{n}m^{1/2}\sigma^{-1}(\hat{\beta}_n-\beta)'x_{n,j}+n^{1/2}(\hat{\sigma}_n-\sigma)t\right)\phi(t)\Bigg|\xrightarrow{P}0. \qquad (5.7)$$

By (1.2)

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

We have

$$\mathrm{Var}\left(n^{-1/2}\sum_{j,k=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}\varepsilon_k - n^{-1/2}\sum_{k=1}^{n}\varepsilon_k\right)$$

$$= \sigma^2 n^{-1}\sum_{k=1}^{n}\left(\sum_{j=1}^{n} x'_{n,k}A_n^{-1}x_{n,j} - 1\right)^2$$

$$= \sigma^2 n^{-1}\sum_{k=1}^{n}\left(\sum_{j=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}\right)^2$$

$$- 2\sigma^2 n^{-1}\sum_{k=1}^{n}\sum_{j=1}^{n} x'_{n,k}A_n^{-1}x_{n,j} + \sigma^2$$

$$= \sigma^2 n^{-1}\sum_{k=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n} x'_{n,j}A_n^{-1}x_{n,k}x'_{n,k}A_n^{-1}x_{n,l}$$

$$- 2\sigma^2 n^{-1}\sum_{k=1}^{n}\sum_{j=1}^{n} x'_{n,k}A_n^{-1}x_{n,j} + \sigma^2$$

$$= \sigma^2 - \sigma^2 n^{-1}\sum_{k=1}^{n}\sum_{j=1}^{n} x'_{n,k}A_n^{-1}x_{n,j} \to 0, \tag{5.8}$$

using condition (C). Hence

$$n^{-1/2}\sum_{j,k=1}^{n} x'_{n,k}A_n^{-1}x_{n,j}\varepsilon_k - n^{-1/2}\sum_{k=1}^{n}\varepsilon_k = o_P(1). \tag{5.9}$$

By a computation in the proof of Theorem 2.1,

$$\frac{n-p}{n}\hat{\sigma}_n^2 - \frac{1}{n}\sum_{j=1}^{n}\varepsilon_j^2 = -\frac{1}{n}(\hat{\beta}_n - \beta)' A_n(\hat{\beta}_n - \beta) = o_P(1).$$

We have

$$\sigma^{-1}n^{1/2}(\hat{\sigma}_n - \sigma) = \sigma^{-1}n^{1/2}(\hat{\sigma}_n - \sigma)^{-1}(\hat{\sigma}_n^2 - \sigma^2)$$

$$= \sigma^{-1}n^{1/2}(\hat{\sigma}_n + \sigma)^{-1}(n^{-1}(n-p)\hat{\sigma}_n^2 - \sigma^2 + n^{-1}p\hat{\sigma}_n^2).$$

Hence

$$\sigma^{-1}n^{1/2}(\hat{\sigma}_n - \sigma) - 2^{-1}\sigma^{-2}n^{-1/2}\sum_{j=1}^{n}(\varepsilon_j^2 - \sigma^2)$$

$$= \sigma^{-1}n^{1/2}(\hat{\sigma}_n - \sigma)^{-1}\left(n^{-1}(n-p)\hat{\sigma}_n^2 - n^{-1}\sum_{j=1}^{n}\varepsilon_j^2\right)$$

$$+ (\sigma^{-1}(\hat{\sigma}_n - \sigma)^{-1} - 2^{-1}\sigma^{-2})n^{-1/2}$$

$$\times \sum_{j=1}^{n} (\varepsilon_j^2 - \sigma^2) + \sigma^{-1} n^{1/2} (\hat{\sigma}_n + \sigma)^{-1} n^{-1} p \hat{\sigma}_n^2 \xrightarrow{P} 0. \qquad (5.10)$$

By Corollary 4.2 in Arcones and Giné [2], a U-process is asymptotically equivalent to the first term in its Hoeffding decomposition:

$$n^{1/2} \sup_{t \in \mathbb{R}} \left| \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} \left(I\left(m^{-1/2} \sum_{j=1}^{m} \sigma^{-1} \varepsilon_{i_j} \le t \right) - \Phi(t) \right) \right.$$

$$\left. - mn^{-1} \sum_{j=1}^{n} (\Phi((m-1)^{-1/2}(m^{1/2} t - \sigma^{-1}\varepsilon_j)) - \Phi(t)) \right| \xrightarrow{P} 0. \qquad (5.11)$$

Observe that

$$E_{\varepsilon_2, \ldots, \varepsilon_m}\left[I\left(m^{-1/2} \sum_{j=1}^{m} \sigma^{-1}\varepsilon_j \le t \right) \right]$$

$$= E_{\varepsilon_2, \ldots, \varepsilon_m}\left[I\left(\sum_{j=2}^{m} \sigma^{-1}\varepsilon_j \le m^{1/2} t - \sigma^{-1}\varepsilon_1 \right) \right]$$

$$= \Phi((m-1)^{-1/2}(m^{1/2} t - \sigma^{-1}\varepsilon_1)).$$

(The previous argument is exactly (4.4) in Arcones and Wang [3]. We have included it for the reader's convenience.)

From (5.8)-(5.11), we get

$$n^{1/2} \sup_{t \in \mathbb{R}} \left| \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} \left(I\left(\hat{\sigma}_n^{-1} m^{-1/2} \sum_{j=1}^{m} \hat{\varepsilon}_{i_j} \le t \right) - \Phi(t) \right) \right.$$

$$- n^{-1} \sum_{j=1}^{n} m(\Phi((m-1)^{-1/2}(m^{1/2} t - \sigma^{-1}\varepsilon_j)) - \Phi(t))$$

$$\left. - \left(n^{-1} m^{1/2} \sigma^{-1} \sum_{j=1}^{n} \varepsilon_j + 2^{-1}\sigma^{-2} n^{-1} \sum_{j=1}^{n} (\varepsilon_j^2 - \sigma^2) t \right) \phi(t) \right| \xrightarrow{P} 0, \qquad (5.12)$$

which implies Theorem 2.2.

The proof of Theorem 2.3 is similar to that of Theorem 2.1 and it is omitted.

Proof of Theorem 2.4. Equation (5.7) with $m = 1$ gives that

$$\sup_{t\in\mathbb{R}}\left| n^{-1/2}\sum_{j=1}^{m}(I(\hat{\sigma}_n^{-1}\hat{\varepsilon}_j \le t) - I(\sigma^{-1}\varepsilon_j \le t))\right.$$

$$\left. - \left(n^{-1/2}\sum_{j=1}^{n}\sigma^{-1}(\hat{\beta}_n - \beta)' x_{n,j} + n^{1/2}(\hat{\sigma}_n - \sigma)t\right)\phi(t)\right| \xrightarrow{P} 0.$$

Using the previous limit, (5.9) and (5.10), we get

$$\sup_{t\in\mathbb{R}}\left| n^{-1/2}\sum\nolimits_{j=1}^{n}(I(\hat{\sigma}_n^{-1}\hat{\varepsilon}_j \le t) - I(\sigma^{-1}\varepsilon_j \le t))\right.$$

$$\left. - \left(n^{-1/2}\sum\nolimits_{k=1}^{n}\sigma^{-1}\varepsilon_k + 2^{-1}n^{-1/2}\sum\nolimits_{j=1}^{n}\sigma^{-2}(\varepsilon_j^2 - \sigma^2)t\right)\phi(t)\right| \xrightarrow{P} 0. \quad (5.13)$$

(5.12) and (5.13) imply the claim.

We will need the following theorem for triangular arrays. Recall that a triangular array $\{X_{n,j} : 1 \le j \le k_n,\ n \ge 1\}$ is infinitesimal if for each $\varepsilon > 0$,

$$P\{\max_{1\le j\le k_n} |X_{n,j}| \ge \varepsilon\} \to 0, \quad \text{as } n \to \infty.$$

Theorem 5.1 (Gnedenko and Kolmogorov, see, e.g., Theorem 4.3 in Petrov [19]). *Let $\{X_{n,j} : 1 \le j \le k_n,\ n \ge 1\}$ be a triangular array of row-wise independent r.v.'s. Let $\{a_n\}$ be a sequence of real numbers. Let $\tau > 0$. Then, $S_n - a_n \xrightarrow{d} N(\mu, \sigma^2)$ and the triangular array is infinitesimal if and only if*

(i) *For each* $\varepsilon > 0, \lim_{n\to\infty}\sum_{j=1}^{k_n} P\{|X_{n,j}| \ge \varepsilon\} = 0.$

(ii) $\lim_{n\to\infty}\sum_{j=1}^{k_n}\mathrm{Var}(X_{n,j}I(|X_{n,j}| < \tau)) = \sigma^2.$

(iii) $\lim_{n\to\infty}\sum_{j=1}^{k_n} E[X_{n,j}I(|X_{n,j}|<\tau)]-a_n=\mu.$

Proof of Theorem 4.1. Let $P_{n,i}$ be the distribution of $X_{n,i}$. Then

$$\begin{aligned}
U_n &:= \frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}((\delta_{X_{n,i_1}}\otimes\cdots\otimes\delta_{X_{n,i_m}})h_{n,i_1,\dots,i_m} \\
&\qquad - (P_{n,i_1}\otimes\cdots\otimes P_{n,i_m})h_{n,i_1,\dots,i_m}) \\
&= \frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}(((\delta_{X_{n,i_1}}-P_{n,i_1}+P_{n,i_1})\otimes\cdots\otimes \\
&\qquad (\delta_{X_{n,i_m}}-P_{n,i_m}+P_{n,i_m}))h_{n,i_1,\dots,i_m} \\
&\qquad - (P_{n,i_1}\otimes\cdots\otimes P_{n,i_m})h_{n,i_1,\dots,i_m}) \\
&= \frac{(n-m)!}{n!}\sum_{k=1}^{m}\sum_{(i_1,\dots,i_k)\in I_k^n} f^{(k)}_{n,i_1,\dots,i_k}(X_{n,i_1},\dots,X_{n,i_k}),
\end{aligned} \tag{5.14}$$

where

$$\begin{aligned}
&f^{(k)}_{n,i_1,\dots,i_k}(x_1,\dots,x_k) \\
&= \sum_{1\le p_1<\cdots<p_k\le m}\ \sum_{(j_1,\dots,i_m)\in I_k^n: j_{p_1}=i_1,\dots,j_{p_k}=i_k}(\mu_1^{\mathbf{p},\mathbf{i},\mathbf{j}}\otimes\cdots\otimes\mu_m^{\mathbf{p},\mathbf{i},\mathbf{j}})h_{n,j_1,\dots,j_m},
\end{aligned} \tag{5.15}$$

where

$$\mu_{p_1}^{\mathbf{p},\mathbf{i},\mathbf{j}} = \delta_{x_1}-P_{n,i_1},\ \dots,\ \mu_{p_k}^{\mathbf{p},\mathbf{i},\mathbf{j}} = \delta_{x_k}-P_{n,i_k}$$

and $\mu_q^{\mathbf{p},\mathbf{i},\mathbf{j}} = P_{n,j_q}$, for $q\notin\{p_1,\dots,p_k\}$. Notice that (5.15) is obtained by collecting all summands with exactly k summands of the form $(\delta_{X_{n,i}}-P_{n,i})$. The first summand in (5.15) takes care of the choice of the $\binom{m}{k}$ places, where these products are. The second summand in (5.15) takes care of the choice of the places where the products of the form $P_{n,i}$ are. In (5.14), U_n is decomposed into orthogonal summands. If $k_1\neq k_2$,

$(i_1, \ldots, i_{k_1}) \in I_{k_1}^n$ and $(j_1, \ldots, j_{k_2}) \in I_{k_2}^n$, then

$$E[f_{n, i_1, \ldots, i_{k_1}}^{(k_1)}(X_{n, i_1}, \ldots, X_{n, i_{k_1}}) f_{n, i_1, \ldots, i_{k_2}}^{(k_2)}(X_{n, j_1}, \ldots, X_{n, j_{k_2}})] = 0.$$

We also have that if $\{i_1, \ldots, i_k\} \neq \{j_1, \ldots, j_k\}$, then

$$E[f_{n, i_1, \ldots, i_k}^{(k)}(X_{n, i_1}, \ldots, X_{n, i_k}) f_{n, j_1, \ldots, j_k}^{(k)}(X_{n, j_1}, \ldots, X_{n, j_k})] = 0.$$

Hence, only when $(i_1, \ldots, i_k)$ is a permutation of $(j_1, \ldots, j_k)$

$$E[f_{n, i_1, \ldots, i_k}^{(k)}(X_{n, i_1}, \ldots, X_{n, i_k}) f_{n, j_1, \ldots, j_k}^{(k)}(X_{n, j_1}, \ldots, X_{n, j_k})]$$

may be different from zero.

Since there are $\binom{m}{k}$ ways to choose the p's in (5.15) and $\frac{(n-k)!}{(n-m)!}$ to choose the free j's in (5.15), we have

$$|f_{n, j_1, \ldots, j_k}^{(k)}(X_{n, i_1}, \ldots, X_{n, i_k})| \leq \binom{m}{k} \frac{(n-k)!}{(n-m)!} 2^k \delta_n n \quad \text{a.s.} \tag{5.16}$$

By (5.15) and (5.16), for each $k \geq 2$,

$$\operatorname{Var}\left(\frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_k) \in I_k^n} f_{n, i_1, \ldots, i_k}^{(k)}(X_{n, i_1}, \ldots, X_{n, i_k}) \right)$$

$$\leq \left(\frac{(n-m)!}{n!} \right)^2 \binom{n}{k} \left(k! \binom{m}{k} \frac{(n-k)!}{(n-m)!} 2^k \delta_n n \right)^2 \leq c \delta_n^2 n^{2-k} \to 0.$$

By Theorem 5.1

$$\frac{(n-m)!}{n!} \sum_{i=1}^n f_{n, i}^{(1)}(X_{n, i}) \xrightarrow{d} N(0, \sigma^2).$$

Notice that for each $1 \leq i \leq n$,

$$\frac{(n-m)!}{n!} |f_{n, i}^{(k)}(X_{n, i})| \leq \frac{(n-m)!}{n!} \binom{m}{1} \frac{(n-1)!}{(n-m)!} 2^{m-1} \delta_n n \leq m 2^{m-1} \delta_n \text{ a.s.}$$

In the proof of the next theorem, we will use the following lemma:

Lemma 5.5. *Let* $\{X_{n,1}, ..., X_{n,n} : n \geq 1\}$ *be a triangular array of row-wise independent r.v.'s with values in a measurable space* $(S, \mathcal{S})$. *Let* $h_{n,i_1,...,i_m} : (S^m, \mathcal{S}^m) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *be a measurable function for each* $(i_1, ..., i_m) \in I_m^n$. *Suppose that for each* $(i_1, ..., i_m) \in I_m^n$

$$E[(h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m}))^2] < \infty.$$

Then

$$\mathrm{Var}\left(\sum_{(i_1,...,i_m)\in I_m^n}(h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m}) - E[h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m})])\right)$$

$$\leq \left(\frac{n!}{(n-m)!}\frac{n!}{(n-m)!} - \frac{n!}{(n-2m)!}\right)\frac{(n-m)!}{n!}$$

$$\sum_{(i_1,...,i_m)\in I_m^n} E[(h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m}))^2].$$

Proof. Without loss of generality, we may that

$$E[h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m})] = 0,$$

for each $(i_1, ..., i_m) \in I_m^n$. We have

$$\mathrm{Var}\left(\sum_{(i_1,...,i_m)\in I_m^n} h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m})\right)$$

$$= \sum_{\substack{(i_1,...,i_m),(j_1,...,j_m)\in I_m^n \\ \{i_1,...,i_m\}\cap\{j_1,...,j_m\}\neq\varnothing}} E[h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m})h_{n,j_1,...,j_m}(X_{n,j_1}, ..., X_{n,j_m})]$$

$$\leq \sum_{\substack{(i_1,...,i_m),(j_1,...,j_m)\in I_m^n \\ \{i_1,...,i_m\}\cap\{j_1,...,j_m\}\neq\varnothing}} (E[(h_{n,i_1,...,i_m}(X_{n,i_1}, ..., X_{n,i_m}))^2])^{1/2}$$

$$(E[(h_{n,j_1,...,j_m}(X_{n,j_1}, ..., X_{n,j_m}))^2])^{1/2}$$

$$\leq \left(\sum_{\substack{(i_1, \ldots, i_m), (j_1, \ldots, j_m) \in I_m^n \\ \{i_1, \ldots, i_m\} \cap \{j_1, \ldots, j_m\} \neq \varnothing}} E[(h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}))^2] \right)^{1/2}$$

$$\times \left(\sum_{\substack{(i_1, \ldots, i_m), (j_1, \ldots, j_m) \in I_m^n \\ \{i_1, \ldots, i_m\} \cap \{j_1, \ldots, j_m\} \neq \varnothing}} E[(h_{n, j_1, \ldots, j_m}(X_{n, j_1}, \ldots, X_{n, j_m}))^2] \right)^{1/2}$$

$$= \sum_{\substack{(i_1, \ldots, i_m), (j_1, \ldots, j_m) \in I_m^n \\ \{i_1, \ldots, i_m\} \cap \{j_1, \ldots, j_m\} \neq \varnothing}} E[(h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}))^2]$$

$$= \left(\frac{n!}{(n-m)!} \frac{n!}{(n-m)!} - \frac{n!}{(n-2m)!} \right) \frac{(n-m)!}{n!}$$

$$\sum_{(i_1, \ldots, i_m) \in I_m^n} E[(h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}))^2].$$

Notice that there are $\frac{n!}{(n-m)!} \frac{n!}{(n-m)!}$ summands in $(i_1, \ldots, i_m)$, $(j_1, \ldots, j_m) \in I_m^n$ and there are $\frac{n!}{(n-2m)!}$ possible ways to get $\{i_1, \ldots, i_m\} \cap \{j_1, \ldots, j_m\} = \varnothing$. Hence, the total number of summands in $\sum_{\substack{(i_1, \ldots, i_m), (j_1, \ldots, j_m) \in I_m^n \\ \{i_1, \ldots, i_m\} \cap \{j_1, \ldots, j_m\} \neq \varnothing}}$ is $\left(\frac{n!}{(n-m)!} \frac{n!}{(n-m)!} - \frac{n!}{(n-2m)!} \right)$. By symmetry each of $E[(h_{n, i_1, \ldots, i_m}(X_{n, i_1}, \ldots, X_{n, i_m}))^2]$ appears the same number of times. Hence, the last inequality in the last display follows.

Proof of Theorem 4.3. By Theorem 4.1 the finite dimensional distributions of $\{n^{1/2}(U_n(t) - E[U_n(t)]) : t \in T\}$ converge to those of $\{Z(t) : t \in T\}$.

Hence, it suffices to show tightness of the process

$$\{n^{1/2}(U_n(t) - E[U_n(t)]) : t \in T\},$$

i.e., it suffices to show that for each $\tau > 0$, there exists a finite partition π of T such that

$$\limsup_{n\to\infty} P\{\sup_{t\in T} n^{1/2} | U_n(t) - U_n(\pi(t)) - E[U_n(t) - U_n(\pi(t))] | \ge \tau\} \le c\tau, \quad (5.17)$$

where c is some universal constant.

Take $\delta > 0$ small enough such that $2^{-3/2} a^{-1}\delta < 1$,

$$288a \int_0^{2^{-3/2a^{-1}\delta}} (\log M(u))^{1/2} du \le \tau^2 \quad (5.18)$$

and

$$32\tau^{-2} b_m c_m^3 \delta^2 \le 1, \quad (5.19)$$

where

$$b_m := \sup_{n\ge m} n \left(\frac{(n-m)!}{n!}\right)^2 \left(\frac{n!}{(n-m)!}\frac{n!}{(n-m)!} - \frac{n!}{(n-2m)!}\right).$$

Notice $\frac{n!}{(n-m)!}\frac{n!}{(n-m)!} - \frac{n!}{(n-2m)!}$ is a polynomial on n of degree $2m - 1$. Hence, $b_m < \infty$. Take a finite partition π of T satisfying (iii) for this δ.

Let $(X_{n,1}^{(k)}, \ldots, X_{n,n}^{(k)})$, $1 \le k \le m$, be m independent copies of $(X_{n,1}, \ldots, X_{n,n})$. By Theorem 3.4.1 in de la Peña and Giné [7],

$$P\{\sup_{t\in T} n^{1/2} | U_n(t) - U_n(\pi(t)) - E[U_n(t) - U_n(\pi(t))] | \ge \tau\}$$

$$\le c_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t)) - E[U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t))] | \ge c_m^{-1}\tau\}, \quad (5.20)$$

where

$$U_n^{(\mathrm{dec})}(t) := \frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_m^n} h_{n,i_1,\ldots,i_m}(X_{n,i_1}^{(1)}, \ldots, X_{n,i_m}^{(m)}, t).$$

By Lemma 5.5 and (5.19), for each $t \in T$,

$$c_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t)) - E[U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t))] | \geq 2^{-1} c_m^{-1} \tau\}$$

$$\leq 4c_m^3 \tau^{-2} n \left(\frac{(n-m)!}{n!}\right)^3 \left(\frac{n!}{(n-m)!} \frac{n!}{(n-m)!} - \frac{n!}{(n-2m)!}\right)$$

$$\times \sum_{(i_1,\dots,i_m)\in I_m^n} E[(h_{n,i_1,\dots,i_m}(X_{n,i_1}^{(1)}, \dots, X_{n,i_m}^{(m)}, t) - h_{n,i_1,\dots,i_m}(X_{n,i_1}^{(1)}, \dots, X_{n,i_m}^{(m)}, \pi(t)))^2]$$

$$\leq 4c_m^3 \tau^{-2} b_m \frac{(n-m)!}{n!} \times \sum_{(i_1,\dots,i_m)\in I_m^n} E[(h_{n,i_1,\dots,i_m}(X_{n,i_1}^{(1)}, \dots, X_{n,i_m}^{(m)}, t)$$

$$- h_{n,i_1,\dots,i_m}(X_{n,i_1}^{(1)}, \dots, X_{n,i_m}^{(m)}, \pi(t)))^2]$$

$$\leq 16\tau^{-2} b_m c_m^3 \delta^2 \leq 2^{-1}. \tag{5.21}$$

Hence, by Lemma 2.1 in Giné and Zinn [13],

$$c_m P\{n^{1/2} \sup_{t\in T} | U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t)) - E[U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t))] | \geq c_m^{-1} \tau\}$$

$$\leq 2c_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t)) - (\widetilde{U}_n^{(\mathrm{dec})}(t) - \widetilde{U}_n^{(\mathrm{dec})}(\pi(t))) | \geq 2^{-1} c_m^{-1} \tau\}, \tag{5.22}$$

where

$$\widetilde{U}_n^{(\mathrm{dec})}(t) := \frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m)\in I_m^n} h_{n,i_1,\dots,i_m}(Y_{n,i_1}^{(1)}, \dots, Y_{n,i_m}^{(m)}, t)$$

and $(Y_{n,1}^{(k)}, \dots, Y_{n,n}^{(k)})$, $k = 1, \dots, m$, are independent copies of $(X_{n,1}^{(k)}, \dots, X_{n,n}^{(k)})$, $k = 1, \dots, m$.

By Theorem 3.4.1 in de la Peña and Giné [7],

$$2c_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{dec})}(t) - U_n^{(\mathrm{dec})}(\pi(t)) - (\widetilde{U}_n^{(\mathrm{dec})}(t) - \widetilde{U}_n^{(\mathrm{dec})}(\pi(t))) | \geq 2^{-1} c_m^{-1} \tau\}$$

$$= 2c_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{dec, double, sym})}(t) - U_n^{(\mathrm{dec, double, sym})}(\pi(t)) | \ge 2^{-1} c_m^{-1} \tau\}$$

$$\le 4c_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{dec, sym})}(t) - U_n^{(\mathrm{dec, sym})}(\pi(t)) | \ge 2^{-2} c_m^{-1} \tau\}$$

$$\le 4c_m c'_m P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{sym})}(t) - U_n^{(\mathrm{sym})}(\pi(t)) | \ge 2^{-2} c_m^{-1} (c'_m)^{-1} \tau\}, \tag{5.23}$$

where

$$U_n^{(\mathrm{dec, double, sym})}(t) := \frac{(n-m)!}{n!} \sum_{(i_1, \dots, i_m) \in I_m^n} R_{i_1} (h_{n, i_1, \dots, i_m} (X_{n, i_1}^{(1)}, \dots, X_{n, i_m}^{(m)}, t)$$

$$- h_{n, i_1, \dots, i_m} (Y_{n, i_1}^{(1)}, \dots, Y_{n, i_m}^{(m)}, t)),$$

$$U_n^{(\mathrm{dec, sym})}(t) := \frac{(n-m)!}{n!} \sum_{(i_1, \dots, i_m) \in I_m^n} R_{i_1} h_{n, i_1, \dots, i_m} (X_{n, i_1}^{(1)}, \dots, X_{n, i_m}^{(m)}, t),$$

$$U_n^{(\mathrm{sym})}(t) := \frac{(n-m)!}{n!} \sum_{(i_1, \dots, i_m) \in I_m^n} R_{i_1} h_{n, i_1, \dots, i_m} (X_{n, i_1}, \dots, X_{n, i_m}, t)$$

and $\{R_j\}_{j=1}^n$ is a sequence of Rademacher's i.i.d. r.v.'s independent of $(X_{n,1}, \dots, X_{n,n})$, $(X_{n,1}^{(k)}, \dots, X_{n,n}^{(k)})$, $1 \le k \le m$, and $(Y_{n,1}^{(k)}, \dots, Y_{n,n}^{(k)})$, $1 \le k \le m$. Notice that hypothesis (iv) is used in (5.23) in order to apply the reverse decoupling inequality.

We have that

$$P\{\sup_{t\in T} n^{1/2} | U_n^{(\mathrm{sym})}(t) - U_n^{(\mathrm{sym})}(\pi(t)) | \ge 2^{-2} c_m^{-1} (c'_m)^{-1} \tau\}$$

$$= P\left\{ \begin{array}{l} \sup_{t\in T} n^{1/2} | U_n^{(\mathrm{sym})}(t) - U_n^{(\mathrm{sym})}(\pi(t)) | \ge 2^{-2} c_m^{-1} (c'_m)^{-1} \tau, \\ \\ \sup_{t\in T} \frac{(n-m)!}{n!} \sum_{(i_1, \dots, i_m) \in I_m^n} (g_{n, i_1, \dots, i_m, \pi} (X_{n, i_1}, \dots, X_{n, i_m}, t))^2 \le 2\delta^2 \end{array} \right\}$$

$$+ P\left\{\sup_{t\in T}\frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I_m^n}(g_{n,i_1,\dots,i_m,\pi}(X_{n,i_1},\dots,X_{n,i_m},t))^2 > 2\delta^2\right\}$$

$$=: \mathrm{I} + \mathrm{II}, \tag{5.24}$$

where

$$g_{n,i_1,\dots,i_m,\pi}(x_1,\dots,x_m,t) = h_{n,i_1,\dots,i_m}(x_1,\dots,x_m,t) - h_{n,i_1,\dots,i_m}(x_1,\dots,x_m,\pi(t)).$$

By (4.4)

$$I \le 36c_m(c'_m)\tau^{-1}\left(\frac{(n-m)!}{n!}\right)^{1/2} E\left[I_A\int_0^{D'_n}(\log D(u,\mathcal{G}'_n))^{1/2}du\right], \tag{5.25}$$

where

$$A := \left\{\sup_{t\in T}\frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m,\pi)\in I_m^n}(g_{n,i_1,\dots,i_m}(X_{n,i_1},\dots,X_{n,i_m},t))^2 \le 2\delta^2\right\},$$

$$\mathcal{G}'_n := \{(g_{n,i_1,\dots,i_m,\pi}(X_{i_1},\dots,X_{i_m},t))_{(i_1,\dots,i_m)\in I_m^n} \in \mathbb{R}^{n!/(n-m)!} : t\in T\}$$

and

$$D'^2_n := \sup_{t\in T}\sum_{(i_1,\dots,i_m)\in I_m^n}(g_{n,i_1,\dots,i_m,\pi}(X_{n,i_1},\dots,X_{n,i_m},t))^2.$$

In A, $D'^2_n \le 2\delta^2\frac{n!}{(n-m)!}$. By (4.1) and the remark on page 22 in Pollard [21],

$$(\log D(u,\mathcal{G}'_n))^{1/2} \le (\log N(2^{-1}u,\mathcal{G}'_n))^{1/2}$$

$$\le 2^{1/2}(\log N(2^{-2}u,\mathcal{G}_n))^{1/2} \le 2^{1/2}(\log D(2^{-2}u,\mathcal{G}_n))^{1/2}$$

$$\le 2^{1/2}\left(\log M\left(2^{-2}\left(\sum_{(i_1,\dots,i_m)\in I_m^n}(H_{n,i_1,\dots,i_m}(X_{n,i_1},\dots,X_{n,i_m},t))^2\right)^{-1/2}u\right)\right)^{1/2}$$

$$\le 2^{1/2}\left(\log M\left(2^{-2}\left(\frac{(n-m)!}{n!}\right)^{1/2}a^{-1}u\right)\right)^{1/2},$$

where $\mathcal{G}_n$ and M are as in Definition 4.1. By (5.18) and (5.25),

$$I \le 2^{1/2}(36)c_m(c'_m)\tau^{-1}\left(\frac{(n-m)!}{n!}\right)^{1/2}$$

$$\times \int_0^{2^{1/2}\delta\left(\frac{n!}{(n-m)!}\right)^{1/2}} \left(\log M\left(2^{-2}\left(\frac{(n-m)!}{n!}\right)^{1/2} a^{-1}u\right)\right)^{1/2} du$$

$$\le 2^{1/2}(144)ac_m(c'_m)\tau^{-1}\int_0^{2^{-3/2}\delta a^{-1}} (\log M(u))^{1/2} du \le c_m(c'_m)\tau. \qquad (5.26)$$

We have

$$\mathrm{II} \le \delta^{-2}E\Bigg[\sup_{t\in T}\Bigg|\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n}(g_{n,i_1,\ldots,i_m,\pi}(X_{n,i_1},\ldots,X_{n,i_m},t))^2$$

$$- E[(g_{n,i_1,\ldots,i_m,\pi}(X_{n,i_1},\ldots,X_{n,i_m},t))^2]\Bigg|\Bigg]$$

$$\le 2\delta^{-2}E\left[\sup_{t\in T}\left|\frac{(n-m)!}{n!}\sum_{(i_1,\ldots,i_m)\in I_m^n}R_{i_1}(g_{n,i_1,\ldots,i_m,\pi}(X_{n,i_1},\ldots,X_{n,i_m},t))^2\right|\right].$$

By (4.3)

$$E\Bigg[\Bigg|n^{1/2}\sum_{(i_1,\ldots,i_m)\in I_m^n}R_{i_1}((g_{n,i_1,\ldots,i_m,\pi}(x_{n,i_1},\ldots,x_{n,i_m},s))^2$$

$$-(g_{n,i_1,\ldots,i_m,\pi}(x_{n,i_1},\ldots,x_{n,i_m},t))^2)\Bigg|^2\Bigg]$$

$$\le \frac{n!}{(n-m)!}\sum_{(i_1,\ldots,i_m)\in I_m^n}((g_{n,i_1,\ldots,i_m,\pi}(x_{n,i_1},\ldots,x_{n,i_m},s))^2$$

$$-(g_{n,i_1,\ldots,i_m,\pi}(x_{n,i_1},\ldots,X_{n,i_m},t))^2)^2$$

$$\leq 16a^2 \frac{n!}{(n-m)!} \sum_{(i_1,\ldots,i_m)\in I_m^n} (g_{n,i_1,\ldots,i_m,\pi}(x_{n,i_1}, \ldots, x_{n,i_m}, s)$$

$$- g_{n,i_1,\ldots,i_m,\pi}(x_{n,i_1}, \ldots, x_{n,i_m}, t))^2.$$

Hence, by (4.2)

$$E\left[\sup_{t\in T}\left|\frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_m^n} R_{i_1}(g_{n,i_1,\ldots,i_m,\pi}(X_{n,i_1}, \ldots, X_{n,i_m}, t))^2\right|\right]$$

$$\leq 36an^{-1/2}\left(\frac{(n-m)!}{n!}\right)^{1/2} E\left[\int_0^{D_n'} (\log D(u, \mathcal{G}_n'))^{1/2} du\right]$$

$$\leq (72)2^{1/2}a^2n^{-1/2}\int_0^{2^{-3/2}} (\log M(u))^{1/2} du \to 0.$$

Proof of Theorem 4.4. We apply Theorem 4.3, with

$$T := \{(b, t) \in \mathbb{R}^p \times \mathbb{R} : |b| \leq M, t \in R\}$$

and

$$h_{n,i_1,\ldots,i_m}(x_1, \ldots, x_m, (b, t))$$

$$= I\left(h(x_1, \ldots, x_m) + \sum\nolimits_{k=1}^p f_{n,k,i_1,\ldots,i_m}(x_1, \ldots, x_m)b_k \leq t\right),$$

for each $(i_1, \ldots, i_m) \in I_m^n$. Hypothesis (i) in Theorem 4.3 follows from Theorem 4.2 with

$$f_{n,i_1,\ldots,i_m}(x_1, \ldots, x_m, (b, t)) = t - h(x_1, \ldots, x_m) - \sum_{k=1}^p f_{n,k,i_1,\ldots,i_m}(x_1, \ldots, x_m)b_k.$$

Notice that for each $x_1, \ldots, x_n \in \mathbb{R}$,

$$\{(f_{n,i_1,\ldots,i_m}(x_{i_1}, \ldots, x_{i_m}))_{(i_1,\ldots,i_m)\in I_m^n} \in \mathbb{R}^{n!/(n-m)!} : (b, t) \in T\}$$

$$= \left\{\left(t - h(x_{i_1}, \ldots, x_{i_m}) - \sum_{k=1}^p f_{n,k,i_1,\ldots,i_m}(x_{i_1}, \ldots, x_{i_m})b_k\right)_{(i_1,\ldots,i_m)\in I_m^n}\right.$$

$$\left.\in \mathbb{R}^{n!/(n-m)!} : (b,\, t) \in T\right\}$$

lies in a subspace of dimension $p + 2$ of $\mathbb{R}^{n!/(n-m)!}$.

As to hypothesis (iii). Let $\delta > 0$. Let $\Gamma(t) := P\{h(X_1, \ldots, X_m) \le t\}$, $t \in \mathbb{R}$. Since Γ is a continuous nondecreasing function such that $\lim_{t\to-\infty}\Gamma(t) = 0$ and $\lim_{t\to\infty}\Gamma(t) = 1$, there are $t_1 < t_2 < \cdots < t_m$ such that $\Gamma(t_j) - \Gamma(t_{j-1}) \le 2^{-1}\delta$, for each $1 \le j \le m+1$, where $t_0 = -\infty$ and $t_{m+1} = \infty$. There exists also a $\tau > 0$ such that for each $1 \le j \le m$, if $|t - t_j| \le \tau$, then $|\Gamma(t) - \Gamma(t_j)| \le 2^{-2}\delta$. Let $\pi : T \to T$ be defined by $\pi(b, t) = (0, t_j)$, if $t \in [t_j, t_{j+1})$, for some $1 \le j \le m$, let $\pi(b, t) = (0, t_1)$, if $t \in (-\infty, t_1)$. Suppose that $p\delta_n M \le \tau$. If $t \in (-\infty, t_1)$, then

$$-I(h(X_{i_1}, \ldots, X_{i_m}) \le t_1)$$

$$\le I\left(h(X_{i_1},\ldots,X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1},\ldots,X_{i_m})b_k \le t\right) - I(h(X_{i_1},\ldots,X_{i_m}) \le t_1)$$

$$\le I(h(X_{i_1},\ldots,X_{i_m}) \le t_1 + p\delta_n M) - I(h(X_{i_1},\ldots,X_{i_m}) \le t_1),$$

$$\left| I\left(h(X_{i_1},\ldots,X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1},\ldots,X_{i_m})b_k \le t\right) - I(h(X_{i_1},\ldots,X_{i_m}) \le t_1)\right|$$

$$\le I(h(X_{i_1}, \ldots, X_{i_m}) \le t_1 + p\delta_n M)$$

and

$$E\left[\left| I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m})b_k \le t\right)\right.\right.$$

$$\left.\left. - I(h(X_{i_1}, \ldots, X_{i_m}) \le t_1)\right|^2\right]$$

$$\le \Gamma(t_1 + p\delta_n M) \le 2\delta.$$

If $t \in [t_j, t_{j+1})$, for some $1 \le j \le m-1$,

$$I(h(X_{i_1}, \ldots, X_{i_m}) \le t_j - p\delta_n M) - I(h(X_{i_1}, \ldots, X_{i_m}) \le t_j)$$

$$\le I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m})b_k \le t\right)$$

$$- I(h(X_{i_1}, \ldots, X_{i_m}) \le t_j)$$

$$\le I(h(X_{i_1}, \ldots, X_{i_m}) \le t_{j+1} + p\delta_n M) - I(h(X_{i_1}, \ldots, X_{i_m}) \le t_j),$$

$$\left| I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m})b_k \le t\right)\right.$$

$$\left. - I(h(X_{i_1}, \ldots, X_{i_m}) \le t_j)\right|$$

$$\le I(t_j - p\delta_n M < h(X_{i_1}, \ldots, X_{i_m}) \le t_{j+1} + p\delta_n M)$$

and

$$E\left[\left| I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m})b_k \le t\right)\right.\right.$$

$$\left.\left. - I(h(X_{i_1}, \ldots, X_{i_m}) \le t_j)\right|^2\right]$$

$$\le E[(I(t_j - p\delta_n M < h(X_{i_1}, \ldots, X_{i_m}) \le t_{j+1} + p\delta_n M))^2]$$

$$= \Gamma(t_{j+1} + p\delta_n M) - \Gamma(t_j - p\delta_n M) \le \delta.$$

If $t \in [t_m, \infty)$, then

$$I(h(X_{i_1}, \ldots, X_{i_m}) \le t_m - p\delta_n M) - I(h(X_{i_1}, \ldots, X_{i_m}) \le t_m)$$

$$\le I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m})b_k \le t\right)$$

$$- I(h(X_{i_1}, \ldots, X_{i_m}) \leq t_j)$$

$$\leq I(t_m < h(X_{i_1}, \ldots, X_{i_m})),$$

$$\left| I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m}) b_k \leq t \right) - I(h(X_{i_1}, \ldots, X_{i_m}) \leq t_j) \right|$$

$$\leq I(t_m - p\delta_n M < h(X_{i_1}, \ldots, X_{i_m}))$$

and

$$E\left[\left| I\left(h(X_{i_1}, \ldots, X_{i_m}) + \sum_{k=1}^{p} f_{n,k,i_1,\ldots,i_m}(X_{i_1}, \ldots, X_{i_m}) b_k \leq t \right) \right.\right.$$

$$\left.\left. - I(h(X_{i_1}, \ldots, X_{i_m}) \leq t_m) \right|^2 \right]$$

$$\leq E[(I(t_m - p\delta_n M < h(X_{i_1}, \ldots, X_{i_m})))^2] = 1 - \Gamma(t_m - p\delta_n M) \leq \delta.$$

Hence, if $p\delta_n M \leq \tau$, then

$$\sup_{t \in T} \frac{(n-m)!}{n!} \sum_{(i_1, \ldots, i_m) \in I_m^n} E[(h_{n,i_1,\ldots,i_m}(X_{n,i_1}, \ldots, X_{n,i_m}(b, t))$$

$$- h_{n,i_1,\ldots,i_m}(X_{n,i_1}, \ldots, X_{n,i_m}, \pi(b, t)))^2] \leq \delta.$$

Acknowledgement

The authors thank the referee for his/her comments and suggestions on the earlier version of the manuscript.

References

[1] M. A. Arcones, Weak convergence for the row sums of a triangular array of empirical processes indexed by a manageable triangular array of functions, Electronic J. Probab. 4(7) (1999), 1-17.

[2] M. A. Arcones and E. Giné, Limit theorems for U-processes, Ann. Probab. 21 (1993), 1494-1542.

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

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

[5] A. K. Bera and C. M. Jarque, Efficient tests for normality homoscedasticity and serial independence of regression residuals, Econo. Lett. 6 (1980), 255-259.

[6] A. K. Bera and C. M. Jarque, A test for normality of observations and regression residuals, Internat. Statist. Rev. 55 (1987), 163-172.

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

[8] E. del Barrio, Empirical and quantile processes in the asymptotic theory of goodness-of-fit tests, Lectures on Empirical Processes, Theory and Statistical Applications, European Mathematical Society Publishing House, Zürich, Switzerland, 2007, pp. 1-92.

[9] E. del Barrio, J. A. Cuesta-Albertos and C. Matrán, Contributions of empirical and quantile processes to the asymptotic theory of goodness-of-fit tests, With discussion, Test 9 (2000), 1-96.

[10] N. R. Draper and H. Smith, Applied Regression Analysis, Second Edition, Wiley, New York, 1981.

[11] R. M. Dudley, Uniform Central Limit Theorems, Cambridge University Press, Cambridge, UK, 1999.

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

[13] E. Giné and J. Zinn, Lectures on the central limit theorem for empirical processes, Lecture Notes in Math. 1221, Springer, New York, 1986, pp. 50-112.

[14] C. J. Huang and B. W. Bolch, On the testing of regression disturbances for normality, J. Amer. Statist. Assoc. 69 (1974), 330-335.

[15] J. Jure□ková, J. Picek and P. K. Sen, Goodness-of-fit test with nuisance regression and scale, Metrika 58 (2003), 235-258.

[16] H. Lilliefors, On the Kolmogorov-Smirnov test for normality with mean and variance unknown, J. Amer. Statist. Assoc. 62 (1967), 399-402.

[17] R. M. Loynes, The empirical distribution function of residuals from generalized regression, Ann. Statist. 2 (1980), 285-298.

[18] M. B. Marcus and G. Pisier, Random Fourier series with applications to harmonic analysis, Ann. Math. Studies, Vol. 101, Princeton University Press, Princeton, New Jersey, 1981.

[19] V. V. Petrov, Sums of Independent Random Variables, Springer-Verlag, New York, 1975.

[20] D. A. Pierce and K. J. Kopecky, Testing goodness of fit for the distribution of errors in regression models, Biometrika 66 (1979), 1-5.

[21] D. Pollard, Empirical Processes: Theory and Applications, NSF-CBMS Regional Conference Series in Probab. and Statist., Vol. 2, Institute of Mathematical Statistics, Hayward, California, 1990.

[22] C. R. Rao and H. Toutenburg, Linear Models, Least Squares and Alternatives, Springer, New York, 1995.

[23] P. K. Sen, J. Jurečková and J. Picek, Goodness-of-fit test of Shapiro-Wilk type with nuisance regression and scale, Austrian J. Statist. 32(1-2) (2003), 163-177.

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

[25] M. A. Stephens, Tests based on EDF statistics, Goodness of Fit Techniques, R. B. D'Agostino and M. A. Stephens, eds., Marcel-Dekker, New York, 1986, pp. 97-193.

[26] S. Weisberg, Comment on some large-sample tests for nonnormality in the linear regression model, by White and MacDonald, J. Amer. Statist. Assoc. 75(369) (1980), 28-31.

[27] H. White and G. M. MacDonald, Some large-sample tests for nonnormality in the linear regression model, J. Amer. Statist. Assoc. 75 (1980), 16-20.

■